ForamEcoGEnIE 2.0: incorporating symbiosis and spine traits into a trait-based global planktic foraminiferal model

Rui Ying[1]**, Fanny M. Monteiro**[2]**, Jamie D. Wilson**[1]**, and Daniela N. Schmidt**[1]

[1]School of Earth Sciences, University of Bristol, Bristol, BS8 1RJ, UK
[2]School of Geographical Sciences, University of Bristol, Bristol, BS8 1SS, UK

Correspondence: Rui Ying (rui.ying@bristol.ac.uk)

Abstract. TS1 Planktic foraminifera are major marine calcifiers in the modern ocean, regulating the marine inorganic carbon pump, and generating marine fossil archives of past climate change. The foraminifera contain ecogroups with and without spines and symbiosis, creating functional trait diversity which expands their ecological niches. Here, we incorporate symbiosis and spine traits into the ForamEcoGEnIE CE1 model, focusing on functional traits rather than individual species. We calibrated the modelled new traits using Latin Hypercube Sampling (LHS) and identified the optimal model parameters from an ensemble of 1200 runs compared with global observations from core-top sediment samples, sediment traps, and plankton nets. The model successfully describes the global distribution and seasonal abundance variation of the four major foraminiferal functional groups. The model reproduces the dominance of the symbiont-obligate group in subtropical gyres and of the symbiont-barren types in the productive subpolar oceans. Global annual mean biomass and foraminifer-derived CE2 carbon export rate are correctly predicted compared to data, with biomass ranging from 0.001 to 0.010 mmol C m^{-3} and organic carbon export 0.002–0.031 mmol C m^{-2} d^{-1}. The model captures the seasonal peak time of biomass and organic carbon export but struggles to reproduce the amplitude of both in productive areas. The sparseness and uneven distribution of observations and the model's limitation in upwelling regions likely contribute to this discrepancy. Our model overcomes the lack of major groups in the previous ForamEcoGEnIE 1.0 version and offers the potential to explore foraminiferal ecology dynamics and its impact on biogeochemistry in modern, future, and paleogeographic environments.

1 Introduction

Planktic foraminifera are marine-calcifying zooplankton that have populated the surface ocean since the mid-Jurassic period ($\sim$ 175 Ma). They produce calcite shells (or "tests") preserved in vast amounts of sediments. These sediments provide proxy archives (e.g. ^{13}C, ^{18}O, Mg/Ca) which are commonly used to reconstruct past climate conditions (Tierney et al., 2020), ocean carbonate chemistry (Hönisch et al., 2012), and to study the biotic response to environmental change (Todd et al., 2020). In the modern oceans, foraminifera contribute to 23 %–56 % of the total open-ocean CaCO$_3$ export (Schiebel, 2002) alongside the other major calcifiers, such as coccolithophores (Daniels et al., 2018) and pteropods (Buitenhuis et al., 2019). However, understanding the impacts of environmental change on foraminifera and their role in the carbon cycle is challenged by their low standing stocks in the surface ocean, a (semi)lunar reproductive cycle driving abundances and difficulties in culturing to ground truth physiology (Schiebel and Hemleben, 2017). Modelling planktic foraminifera and their ecology, therefore, has a critical role in increasing and testing our understanding of their biological and ecological influence on the marine inorganic carbon cycle and their role as a paleoproxy carrier.

Significant developments of global foraminifera models have been driven by the increasing number and extent of flux and community structure observations (Siccha and Kucera, 2017; Buitenhuis et al., 2013; Sunagawa et al., 2020). Most existing models are either empirically-based or focus on selected extant species. For example, Waterson et al. (2017) built a Maxent species distribution model based on sediment core data to study the niche variability during the Last

Glacial Maximum (LGM) compared to the Holocene. Žarić et al. (2006) constructed a statistical model that correlated hydrographical factors with sediment trap abundance of 18 dominant species. Correlative models, though, are limited for extensive future projections as they assume a constant environmental niche, neglecting adaptation and acclimation (Buckley et al., 2010). In addition, niche models do not resolve biological interactions which have an important role in shaping species distribution (Anderson, 2017).

Fraile et al. (2008, 2009), Lombard et al. (2011) and Kretschmer et al. (2016, 2018) built and extended ecophysiology-based models (PLAFOM and FORAM-CLIM CE3) to overcome these limitations. They successfully reconstructed planktic foraminifera's geographical distribution, seasonal and vertical population dynamics, and simulated distributions in different climatologies such as the LGM (Fraile et al., 2009) and future high-emission scenarios (Roy et al., 2015). Both models are species-based and therefore cannot be applied in the deeper geological record older than the Miocene (>5 Ma) (Kucera and Schonfeld, 2007) due to a high number of extinct species and cryptic taxa with unknown novel ecologies (Renaud and Schmidt, 2003). Additionally, FORAMCLIM uses experimental growth rates to simulate foraminifera abundance and does not resolve top-down controls on foraminifera biomass. To fill the model gap and to leverage the abundant foraminiferal fossil information, a mechanistic model not limited to species is needed.

Trait-based plankton models are an alternative approach focusing on organismal traits including morphological and physiological properties instead of taxonomic identities. They provide a mechanistic way to mimic the complex ocean ecology characterising the functional groups, their traits, and associated benefits and costs (i.e. trade-offs) (Zakharova et al., 2019; Kiørboe et al., 2018). Models adopting trait-based approaches have successfully reconstructed the biomass distribution of diverse marine community including cyanobacteria (Follows et al., 2007) and diazotrophs (Monteiro et al., 2010). This modelling strategy is well suited to be applied to planktic foraminifera as functional traits such as body size (Schmidt et al., 2004a), size-normalised weight (Todd et al., 2020; Barker and Elderfield, 2002), and symbiosis (Spero and Parker, 1985) are widely measured and studied. The evolution of these functional traits has been described in detail across the Cenozoic (Ezard et al., 2011).

Critical traits of planktic foraminifera include calcification, body size, and the presence and absence of spines and symbiosis. While calcification and body size are universal traits for all foraminifera, the evolution of spines and symbiosis determine the species-level discrepancies (Aze et al., 2011). Based on the presence of symbionts and spines, foraminifera can be divided into four functional groups: (1) symbiont-barren non-spinose; (2) symbiont-barren spinose; (3) symbiont-facultative non-spinose; (4) and symbiont-obligate spinose (Table 1). Roughly 19 out of the 50 modern foraminifera species are symbiotic, bearing eukaryotic

algae such as dinoflagellates, chrysophytes, and haptophytes (Takagi et al., 2019), though this important relationship is not established for all taxa. Photosynthesising symbionts provide extra energy to foraminifera in nutrient-depleted regions (LeKieffre et al., 2018; Ortiz et al., 1995; Uhle et al., 1999). Consequently, symbiotic species dominate tropical to subtropical regions, while non-symbiont species (termed as "symbiont-barren") reach high abundance in temperate and polar oceans (Fig. 3). Some symbiont-bearing taxa cannot live without their symbionts (termed as "symbiont-obligate") (Bé et al., 1982) while others are flexible (termed as "symbiont-facultative").

The presence of calcareous spine influences the foraminifera's feeding behaviour. Non-spinose foraminifera rely on rhizopodia to capture prey. Spinose foraminifera have spines extruding from the test that increases their effective reach range and ability to active prey; this in turn increases the ability to caption more prey types and larger cell sizes like copepods (Anderson et al., 1979). Laboratory observations show that spinose carnivory foraminifera prefer food with a high zooplankton-to-phytoplankton protein ratio (Schiebel and Hemleben, 2017). The effective encounter rate of a spinose taxon can be 3 orders of magnitude higher than non-spinose species (Gaskell et al., 2019). Roughly half of modern species are spinose, but existing models have not taken this trait advantage into consideration.

Recently, Grigoratou et al. (2019) developed the first mechanistic and trait-based 0D model (ForamEcoGEnIE 1) for the symbiont-barren non-spinose foraminiferal group and coupled it to a carbon-centric Grid-ENabled Integrated Earth system model (cGEnIE; Grigoratou et al., 2021a), a 3D Earth System Model of Intermediate Complexity (EMIC), allowing for fast computational time and widely applied to past climates including the Palaeocene–Eocene Thermal Maximum (PETM; Ridgwell and Schmidt, 2010), Last Glacial Maximum (LGM; Rae et al., 2020) and Cretaceous–Paleogene (K–Pg) boundary (Henehan et al., 2019). The computational efficiency and application to a wide range of geological periods mean ForamEcoGEnIE can be used to explore foraminiferal diversity in past climates beyond the limits of other models (Ezard et al., 2011). Here, we extend the model to ForamEcoGEnIE 2.0 by resolving three more critical functional groups of planktic foraminifera by adding the traits of symbiosis and spines (the latter tested in Grigoratou et al., 2021b). This development therefore focuses on solving foraminiferal diversity rather than marine carbonate chemistry. We tuned the model by comparing it with three global observational data compilations (sediment core-tops, plankton nets, and sediment traps) and test its ability to reproduce surface biomass, organic carbon and calcite flux, and geographic distribution in the modern climate.

2 cGEnIE ocean and atmosphere physics

ForamEcoGEnIE is based on cGEnIE (carbon-centric Grid-ENabled Integrated Earth system model). The fast climate and ocean physics of cGEnIE are based on a coarse-resolution 3D frictional geostrophic ocean model coupled to a 2D energy-moisture-balance atmospheric model and a dynamic–thermodynamic sea-ice model (Edwards and Marsh, 2005; Marsh et al., 2011). The ocean has a 36×36 equal-area horizontal grid (uniform in longitude and sine-uniform in latitude) with 16 logarithmically spaced vertical levels as defined in Cao et al. (2009). The physical model is coupled with a model of ocean biogeochemical cycles (Ridgwell et al., 2007; van de Velde et al., 2021), sea-floor sedimentary processes (Ridgwell and Hargreaves, 2007), and marine ecosystem processes (Ward et al., 2018). The plankton ecosystem is resolved in the surface layer (0–80.8 m). The model presented in this study is configured with a seasonally forced pre-industrial climate state and an atmospheric CO_2 concentration restored to 278 ppm.

3 Size-based plankton ecosystem framework EcoGEnIE

3.1 Biogeochemical tracers

The model has three main state variables: inorganic resources (i_r), living biomass (i_b), and detritus (i_d). Each state variable contains multiple biogeochemical tracers: carbon, phosphorus, and iron. Plankton populations are counted in notation j, and each plankton includes the three tracers above, although autotroph plankton (phytoplankton and symbiotic foraminifera) have an extra tracer of chlorophyll (noted in Chl). Figure 1 shows a schematic of the plankton types including foraminifera and denotes elements in different colours.

3.2 Plankton cell size and quota

In EcoGEnIE, individual body size determines key physiological processes including nutrient uptake, photosynthesis, grazing gain, and predation through allometric scaling (West et al., 1997) because of its role as a master trait among pelagic organisms (Andersen et al., 2016). The modelled size-dependent parameters (except for photosynthesis) follow a generic power law: $P = aV^b$, where P is the size-based parameter, V the spherical biovolume, and a and b the allometric intercept and exponent, respectively.

A fundamental size-based concept of EcoGEnIE is the plankton cell quota for various elements. The carbon quota content (Q_C) follows the same power law as per Eq. (1). Variable stoichiometry (Q_{i_b}, Eq. 2) is achieved by the ratio of assimilated nutrients biomass (B_{i_b}, where i_b stands for P, Fe, or chlorophyll) to carbon biomass (B_C) (Droop, 1968; Flynn, 2008). This stoichiometry limits nutrient uptake rate

($Q_{i_b}^{stat}$, Eq. 3) as per Geider et al. (1998), with a higher value close to its maximum ($Q_{i_b}^{max}$) lowering the nutrient uptake or chlorophyll synthesis rate. The nutrient quota range ($Q_{i_b}^{min}$, $Q_{i_b}^{max}$) is proportional to the carbon quota (Q_C):

$$Q_C = aV^b), \tag{1}$$

$$Q_{i_b} = \frac{B_{i_b}}{B_C}, \quad i_b = \text{P, Fe, Chl}, \tag{2}$$

$$Q_{i_b}^{stat} = \left(\frac{Q_{i_b}^{max} - Q_{i_b}}{Q_{i_b}^{max} - Q_{i_b}^{min}} \right)^{0.1}. \tag{3}$$

3.3 Plankton biomass dynamics

The biomass of any plankton group (j) and element (i_b), B_{j,i_b}, varies due to a combination of potential physiological processes that are determined by the type of organism: nutrient uptake, grazing gains, grazing losses, mortality, and respiration loss (Eq. 5). Foraminifera-related specific processes will be introduced in following sections. We refer readers to Ward et al. (2018) for the detailed description of EcoGEnIE that expands on the description below. TS3

$$\frac{\partial B_{j,i_b}}{\partial t} = \underbrace{\mu_{j,i_b} \cdot B_{j,C}}_{\text{nutrient uptake}} + \underbrace{B_{j,C} \cdot \lambda_{i_b} \sum_{j_{prey}=1}^{J} G_{j,j_{prey},i_b}}_{\text{grazing gains}}$$

$$- \underbrace{B_{j_{pred},C} \cdot \sum_{j_{prey}=1}^{J} G_{j_{pred},j,i_b}}_{\text{grazing losses}} - \underbrace{m_j \cdot B_{j,i_b}}_{\text{mortality loss}} - \underbrace{r_{j,C} \cdot B_{j,i_b}}_{\text{respiration loss}}. \tag{4}$$

3.4 Inorganic nutrient dynamics

The inorganic resource state variable (R_{i_r}) varies with nutrient uptake (V_{j,i_r}) and dissolved inorganic carbon (DIC) with the living organisms' respiration ($r_{j,C}$):

$$\frac{\partial R_{i_r}}{\partial t} = \begin{cases} \sum_{j=1}^{J} -\mu_{j,i_r} \cdot B_{j,C}, & i_r = \text{Fe, P} \\ \sum_{j=1}^{J} -\mu_{j,i_r} \cdot B_{j,C} + \sum_{j=1}^{J} r_{j,C}, & i_r = \text{C}. \text{ TS4} \end{cases} \tag{5}$$

Additional sources and sinks of nutrients such as remineralisation of organic matter and air–sea gas exchange are computed in the biogeochemical module BIOGEM (Ridgwell et al., 2007).

3.5 Particulate organic matter dynamics

Particulate organic matter (POM) flux (F) is a combination of predators' messy feeding (the first term) and the mortality

loss (the second term) from all plankton groups (Eq. 6):

$$F = \sum_{j_{\text{pred}}=1}^{J} \sum_{j_{\text{prey}}=1}^{J} (1 - \beta_{j_{\text{pred}}, i_{\text{d}}})(1 - \lambda_{j_{\text{pred}}, i_{\text{b}}}) G_{j_{\text{pred}}, j_{\text{prey}}} B_{j_{\text{pred}}, \text{C}}$$
$$+ \sum_{j=1}^{J} (1 - \beta_j) m_j B_{j, i_{\text{d}}}, \tag{6}$$

where $\beta_{j_{\text{pred}}}$ is the fraction of dissolved organic matter (DOM) subject to diffusion and advection by ocean circu-
5 lation. The remaining fraction $(1 - \beta_{j_{\text{pred}}})$ is the particulate organic matter (POM) subject to redistribution through the water column by sinking. The parameter β is a sigmoid function depending on maximum and minimum DOM fraction $(\beta_{\text{max}}, \beta_{\text{min}})$ of predators' equivalent sphere diameter
10 (ESD) CE4 and the size β_{s} at which DOM / POM ratio equals 1 (Ward and Follows, 2016). Smaller cell sizes are associated with greater proportion as DOM.

$$\beta = \beta_{\text{max}} - \frac{\beta_{\text{max}} - \beta_{\text{min}}}{1 + \beta_{\text{s}}/\text{ESD}}. \tag{7}$$

Messy feeding behaviour is modelled as the unassimilated
15 fraction $(1 - \lambda_{j_{\text{pred}}})$ of prey which is limited by the size-independent maximum efficiency coefficient (λ_{m}) and the nutrient limitation (Fe or P):

$$\lambda = \lambda_{\text{m}} \cdot \min\left[Q_{\text{P}}^{\text{stat}}, Q_{\text{Fe}}^{\text{stat}}\right]. \tag{8}$$

4 ForamEcoGEnIE 1 brief description

20 ForamEcoGEnIE 1 accounted for the feeding behaviour and calcification of foraminifera (Grigoratou et al., 2019, 2021a). It implemented a predator–prey interaction ($G_{j_{\text{pred}}, j_{\text{prey}}}$, Eq. 9) using a Holling type II model (Holling, 1965), where the overall grazing rate depends on the total available prey
25 ($F_{j_{\text{pred}}}$), the maximum grazing rate of predators ($G_{\text{pred}}^{\text{m}}$), and the half-saturation concentration of available food ($k_{j_{\text{prey}}}$); it is regulated by temperature limitation (γ_{T}), a prey-switching term (Φ), and a prey-refuge protection $(1 - e^{\Lambda F_{j_{\text{pred}}}})$. The cal­cification trait was included by reducing foraminifera palata-
30 bility (P_{p} which influences $F_{j_{\text{pred}}}$ (Eq. 10) and mortality rate (m_j, Eq. 5) to account for higher protection against predators and infections at the expense of a lower $G_{\text{foram}}^{\text{m}}$ (Eq. 9). We also introduce the ForamEcoGEnIE 2 parameters (spine ef­fect τ and a mixotrophy limitation λ_{h} TS5) here which is set
35 to 1, i.e. not functioning in ForamEcoGEnIE 1: TS6

$$G_{j_{\text{pred}}, j_{\text{prey}}} = \underbrace{\gamma_{\text{T}} \cdot \lambda_h}_{\text{limitations}} \cdot \underbrace{\frac{G_{\text{pred}}^{\text{m}} F_{j_{\text{pred}}}}{\tau k_{j_{\text{prey}}} + F_{j_{\text{pred}}}}}_{\text{overall grazing rate}} \cdot \underbrace{\Phi_{j_{\text{pred}}, j_{\text{prey}}}}_{\text{Switching}}$$
$$\cdot \underbrace{(1 - e^{\Lambda F_{j_{\text{pred}}}})}_{\text{prey refuge}}, \tag{9}$$

$$F_{j_{\text{pred}}} = P_{\text{p}} \cdot B_{j_{\text{prey}}}$$
$$\cdot \exp\left[-\left(\ln\left(\frac{\mu_{j_{\text{pred}}, j_{\text{prey}}}}{\mu_{\text{opt}}}\right)\right)^2 / (2\sigma_{j_{\text{pred}}}^2)\right]. \tag{10}$$

Predators select their prey (Eq. 10) based on the predator–prey size ratio $\mu_{j_{\text{pred}}, j_{\text{prey}}}$ relative to the optimal value μ_{opt}, the predators' food range $\sigma_{j_{\text{pred}}}^2$, and the calcification protec­
40 tion P_{p}. Foraminifera in both ForamEcoGEnIE 1 and 2 are set as herbivores.

The grazing process like other metabolic processes in EcoGEnIE is temperature-dependent, following the universal metabolic theory (Brown et al., 2004). The body temperature
45 of ectothermic plankton is determined by the ambient sea­water environment (T). Temperature regulation γ_{T} acts on metabolic processes including respiration, nutrient uptake, and predation. It is modelled through an Arrhenius-like func­tion (Eq. 12), where the parameter A determines temperature
50 sensitivity and reference temperature (T_{ref}) is the temperature allowing $\gamma_{\text{T}} = 1$:

$$\gamma_{\text{T}} = e^{A(T - T_{\text{ref}})}. \tag{11}$$

The prey-switching term ($\Phi_{j_{\text{pred}}, j_{\text{prey}}}$) simulates the feeding habitat of zooplankton (Eq. 9). The exponential s defines the
55 active level of zooplankton predators which capture abundant prey with higher priority when s increases. Foraminifera in both ForamEcoGEnIE 1 and 2 are assumed to be ambush passive predators with $s = 1$:

$$\Phi_{j_{\text{pred}}, j_{\text{prey}}} = \frac{\left(F_{j_{\text{pred}}}\right)^s}{\sum_{j_{\text{prey}}=1}^{J} \left(F_{j_{\text{pred}}}\right)^s}. \tag{12}$$
60

A refuge term $(1 - e^{\Lambda F_{j_{\text{pred}}}})$ in Eq. (9) is added to decrease the grazing rate when prey availability lowers. The coefficient Λ determines the strength of such protection.

5 ForamEcoGEnIE 2: improved calcification and more functional groups
65

In ForamEcoGEnIE 2, we add symbiosis and spine traits for foraminifera to result in four functional groups (Table 1, Fig. 1). We also implement a new calcification energetic cost by using a respiration term rather than a reduced maximum growth rate in ForamEcoGEnIE 1.
70

5.1 Calcification trait trade-offs

5.1.1 Benefit: mortality protection

The mortality loss term for zooplankton scales with a basal rate constant m_{b} (Eq. 5). As per Grigoratou et al. (2019, 2021a), this is downscaled for foraminifera by a protection
75 term P_{m}, where a lower value of m_j indicates a higher pro­tection from the foraminiferal test against viral and bacterial

Table 1. The four modelled functional groups of planktic foraminifera and their species representative in ForamEcoGEnIE 2.0.

Spine trait	Symbiosis trait	Species example	Species number*	Model implementation
Spinose	Symbiont-obligate	*Globigerinoides ruber*	17	This study
Spinose	Symbiont-barren	*Globigerina bulloides*	2 TS7	This study
Non-spinose	Symbiont-facultative	*Neogloboquadrina dutertrei*	5	This study
Non-spinose	Symbiont-barren	*Neogloboquadrina pachyderma*	23	Extended from ForamEcoGEnIE 1

* Count from Schiebel and Hemleben (2017).
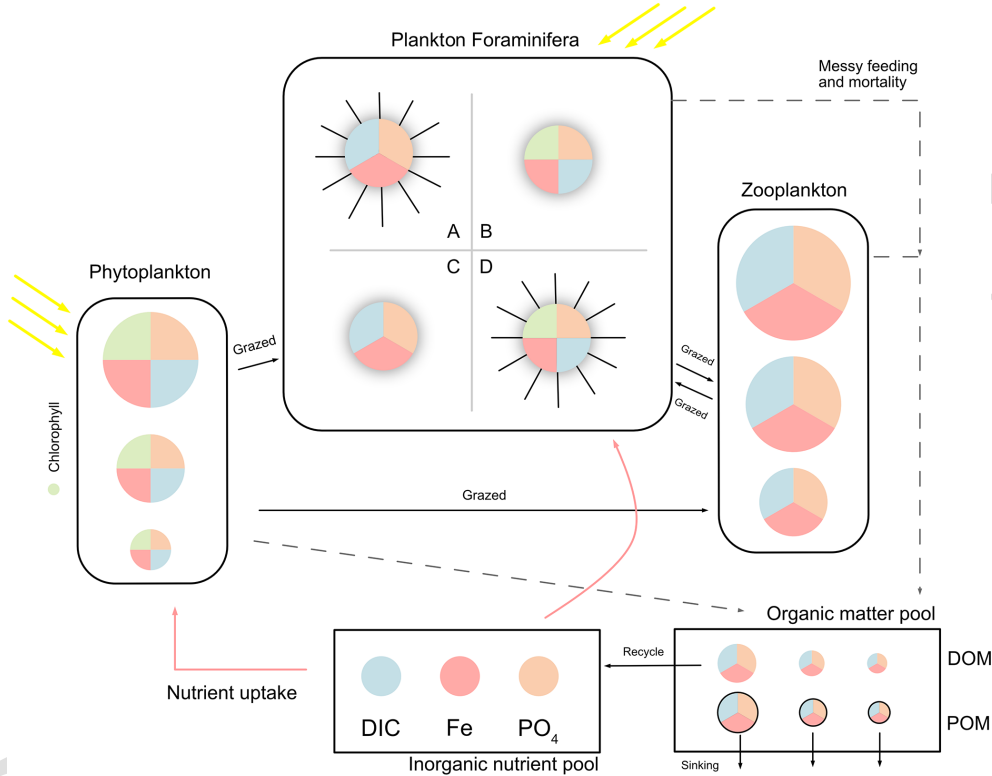
Figure 1. Schematic representation of the ForamEcoGEnIE 2.0 model structure. The model includes the biogeochemical cycles of C, Fe, and P (shown in different colours), various plankton size classes and four main groups of planktic foraminifera: A – symbiont-barren spinose group; B – Symbiont-facultative non-spinose group; C – symbiont-barren non-spinose group; D – symbiont-obligate spinose group. DIC stands for dissolved inorganic carbon and PO_4 for phosphate. The model represents nutrient uptake (red arrows), dissolved and particulate organic matter production (DOM and POM) caused by messy feeding and mortality (dashed arrows), and zooplankton grazing (black arrows).

infections:

$$m_j = P_m \cdot m_b. \tag{13}$$

5.1.2 Benefit: protection from predators (palatability)

As per ForamEcoGEnIE 1.0, calcification protects from grazing and is defined by P_p which reduces the biomass loss from predation (Eq. 10).

5.1.3 Cost: higher metabolic cost

We modified the metabolic cost of calcification defined in Grigoratou et al. (2019, 2021a) by replacing the original reduced maximum growth rate (or specifically maximum grazing rate) with a temperature-dependent respiration loss term. We made this change because (1) extra respiration is a more biologically realistic cost with (2) this temperature-dependent term reconciling the model with the low-latitude biomass observation. The respiration r_j present in Eq. (5) scales with carbon biomass and is multiplied by constant r_b and temperature limitation (Eq. 11). We assumed that the lost carbon from respiration is instantly recycled back to dissolved inorganic carbon (DIC) pool:

$$r_j = r_b \cdot \gamma_T. \tag{14}$$

5.2 Spine trait trade-offs

Spines are an important part of foraminiferal taxonomy. Spines, like the overall test, are made of calcite. A range of biological functions are assumed to be linked to symbiosis and feeding behaviour (Schiebel and Hemleben, 2017).

5.2.1 Benefit: enhanced grazing

Studies show that spinose foraminifera are more efficient in capturing and digesting prey due to the spine and rhizopodia networks (Anderson and Bé, 1976). Spines widen the prey availability of immotile foraminifera and facilitate the capture of larger preys. Non-spinose species cannot hold active prey and only accept smaller particles of copepods in the laboratory observations (Anderson et al., 1979; Hemleben et al., 1989). Grigoratou et al. (2021b) modelled such benefit by reducing the half-saturation constant (conventionally noted as k TS8 in a Michaelis–Menten model). Here, we adopt this approach by reducing $k_{j_{prey}}$ to a scaling parameter CE5 τ ($0 < \tau < 1$; Eq. 10).

5.2.2 Other trade-offs as calcification: higher metabolic cost and reduced palatability

We assume that the metabolic cost and protection from the spines are characterised the same way as for calcification (Eqs. 13–14). Spinose foraminifera have a higher cost for calcification due to the slightly higher amounts of carbonate and a stronger protection than non-spinose taxa (Table 2). We did not reduce the mortality term as this was not supported by direct evidence.

5.3 Symbiosis trait trade-offs

Symbiosis is a novel trait in the model, commonly seen in marine organisms including foraminifera. Many planktic foraminifera harbour algae (e.g. dinoflagellate, diatom) within their cells (Takagi et al., 2019). We represent these symbiotic species in the model as a single organism which combines heterotrophy and autotrophy, equivalent to a calcifying mixotroph. We use the trait-based representation of mixotrophy of Ward and Follows (2016), where any plankton can "naturally" predate and photosynthesise. While mixotrophs have this ability in the model, this is turned off for the rest of plankton (i.e. V_m is 0 for zooplankton and G_m is 0 for phytoplankton).

5.3.1 Benefit: enabled autotrophy for planktic foraminifera

The symbiont has a cell size that is defined via a symbiont/foraminifera size ratio ψ (Eq. 16) to characterise the symbiont's affinity in taking up nutrients and light. Photomicrograph observations showed that foraminifera symbionts are about 1 : 20 smaller in size than the host cell (Takagi et al., 2019):

$$V_s = \psi^3 V_h. \tag{15}$$

The generic nutrient uptake of symbionts follows a Michaelis–Menten function limited by mixotrophy (λ_s), quota ($Q_{i_r}^{stat}$), and temperature (γ_T), where the variable (**R**) represents nutrient resources. The half-saturation constant is replaced by nutrient affinity, a more mechanistic parameter for nutrient uptake α_{i_r} TS9. Nutrient affinity is often referred to "clearance rate" and regarded as a proxy of competitive strength (Fiksen et al., 2013). According to Edwards et al. (2012) review on phytoplankton trait trade-offs, nutrient affinity is negatively related to cell size because of lower surface to volume ratio, while the maximum uptake rate (V_m) is positively related:

$$\mu_{i_r} = \lambda_s \cdot Q_{i_r}^{stat} \cdot \gamma_T \cdot \frac{V_{i_r}^m \alpha_{i_r} \mathbf{R}}{V_{i_r}^m + \alpha_{i_r} \mathbf{R}}. \tag{16}$$

The symbionts' photosynthesis growth is modelled following a size-dependent unimodal equation (Geider et al., 1998; Moore et al., 2001). This equation has higher explanatory power for eukaryotic phytoplankton cells than a power law (Bec et al., 2008). The maximum photosynthesis rate P_C^m is determined by dimensionless parameter P_a, P_b, P_c and the biovolume of symbiont V_s, and the mixotrophy cost λ_s:

$$P_C^m = \frac{\lambda_s \left(P_a + \log_{10} V_s \right)}{P_b + P_c \log_{10} V_s + \log_{10} V_s^2}. \tag{17}$$

The practical photosynthesis rate is further constrained by nutrient availability (the smallest between γ_{Fe} and γ_P), temperature (γ_T), and light intensity (γ_I):

$$P_C = P_C^m \cdot \min\left[\gamma_P, \gamma_{Fe} \right] \cdot \gamma_T \cdot \gamma_I. \tag{18}$$

Nutrient limitation γ_{i_r} (i_r is either P or Fe, see the definition in Eq. 2) is determined by the minimal value of the phosphorus or iron limitation term, which follows the quota relationship in Droop (1968):

$$\gamma_{i_r} = \frac{1 - Q_{i_r}^{min}/Q_{i_r}}{1 - Q_{i_r}^{min}/Q_{i_r}^{max}}, \qquad i_r = \text{Fe, P}. \tag{19}$$

Light limitation follows the model of Moore et al. (2001), where I represents light intensity, α is initial slope of $P - I$ TS10 curve limited by Fe content (γ_{Fe}), and Q_{Chl} is chlorophyll quota.

$$\gamma_I = 1 - \exp\left(\frac{-\alpha \cdot \gamma_{Fe} \cdot Q_{Chl} \cdot I}{P_C^m \cdot \gamma_T \cdot \min\left[\gamma_P, \gamma_{Fe} \right]} \right) \tag{20}$$

5.3.2 Cost: downgrading autotroph and heterotroph efficiency

The cost of mixotrophy is that both autotrophic and heterotrophic processes (i.e. photosynthesis and grazing rates)

are scaled down (by multiplying factor λ_s and λ_h for symbionts and hosts, respectively: $0 < \lambda_s, \lambda_h < 1$, Eqs. 9 and 16) compared to the pure autotroph/CE6 heterotroph specialist (Castellani et al., 2013; Våge et al., 2013; Ward and Follows, 2016). We distinguish between symbiont-obligate and symbiont-facultative foraminifera using different λ_s/λ_h parameter values to reflect their different dependency on symbionts (Table 2).

5.4 Approximating foraminiferal calcite export using fixed PIC/POC ratio

Planktic foraminifera produce organic carbon in the subsurface water column (Salter et al., 2014) and sequester inorganic carbon into the deep oceans via their dead tests (Schiebel, 2002). The organic carbon flux derived from foraminifera is treated the same way as in EcoGEnIE as discussed in Sect. 3.4. The calcite export, specific to foraminifera, is approximated by multiplying the foraminiferal bulk organic carbon export with a globally uniform particle inorganic carbon (PIC) to organic carbon (POC) molar ratio of 0.36 based on the empirical data by Schiebel and Movellan (2012).

6 Model parameterisation

6.1 Plankton community size structure

We resolve eight size classes of phytoplankton, seven size classes of zooplankton, and one size class for each of the four foraminiferal groups. Phytoplankton and zooplankton size classes include 0.6, 1.9, 6.0, 19.0, 60.0, 190.0, 600.0, and 1900.0 µm, with zooplankton missing the smallest class due to minimum prey size. While the size structure of these plankton is fixed, we tested the foraminiferal ESD ranging from pre-adult (60 µm) to adult (600 µm) using the ensemble described below. Each test contains one randomly assigned foraminiferal size and this is same for each foraminiferal group. However, we found that the size (190 µm) from a previous study (Grigoratou et al., 2021 TS11) still achieved the best score (Table 2).

6.2 Experiments with sampled parameters

We run an ensemble of 1200 model experiments, each testing a different combination of parameter values (Table 2), to explore all possible trait values and select the best trait combinations to match available foraminiferal observations (Sect. 3.2). The parameter sets are generated using a Latin Hypercube Sampling (LHS) algorithm that samples values of 12 foraminiferal parameters from uniform parameter distributions (Table 2; Sarrazin et al., 2016). However, several rules are set in the sampling: (1) the spinose ones always own higher palatability and mortality protection (and corresponding respiration cost) than the non-spinose ones;

(2) the symbiont size for both symbiont-facultative and symbiont-obligate groups are set to the same value; (3) all the foraminifera have the same size. Each simulation is run for 250 years continuing from a 10 000-year spin-up simulation as the ecosystem structure typically reaches equilibrium after ~ 50 years. The other ecosystem parameters are the same as Ward et al. (2018) (Table S3).

6.3 Observations for comparison

We used multisource data compilations (sampled from sediment core-tops, sediment traps, and plankton nets) to calibrate the model (see below). To simplify the model–data comparison, we assume that the biomass/export changes between the pre-industrial age (the model) and the present climate represented in plankton net and sediment trap samples are negligible. This is because (1) planktic foraminifera have stably low biomass; and (2) sediment trap and plankton net data were collected over a wide time range (1970s–2010s) with changing climatologies.

6.3.1 Relative abundance

We used a sediment core-top census data compilation (Siccha and Kucera, 2017) to represent a long-term mixed Late Holocene baseline (pre-industrial) to validate the spatial abundance patterns of each modelled foraminiferal group. We calculated the modelled relative abundance of each group based on its carbon export production.

To determine the observed relative abundance, we compiled species into functional groups using species traits defined by Schiebel and Hemleben (2017) and Takagi et al. (2019) (Table S4). We regridded the observations into the model grid (averaging each data grid point onto the cGEnIE grid). We ignored species with less than 3 % local abundance (a few specimen) to avoid uncertainties caused by transport via ocean currents (van Sebille et al., 2015) and taxonomic uncertainties of rare taxa. This threshold is determined by the standard error of Fisher's diversity index (Fisher et al., 1943). We used the Mielke measure (details in Sect. 6.4) to quantify the model–data fit. We omitted the Arctic Ocean and the Mediterranean Sea in the model–data comparison because the model resolution in these regions is too low to represent adequate ocean physics.

6.3.2 Annual average biomass and export

To validate the modelled biomass and organic carbon export, we compiled two global datasets: (1) plankton net data from the first 100 m (if sampled, otherwise the nearest depth that is no more than 120 m) for biomass and (2) sediment trap data for carbon export. We converted the units of plankton net ("number m^{-3}") and sediment trap data ("number m^{-2} d^{-1}") into "mmol C m^{-3}" and "mmol C m^{-2} d^{-1}" using the empirical factor of 0.845 µg per specimen TS13 from Schiebel and Movellan (2012). We converted modelled carbon export

Table 2. List of the foraminifera-relevant model parameters tested in the global sensitivity analysis (GSA) and identified optimal parameter values for each foraminiferal group. CE7

Related trait(s)	Parameter	Description	Tested range[a]	Unit	Optimal values (bn[b])	Optimal values (bs[b])	Optimal values (sn[b])	Optimal values (ss[b])
Foraminiferal size	ESD	Equivalent sphere diameter	[60, 600]	μm	**190[c]**	190	190	190
Calcification/spine	p_m	Protection from mortality	[0–1]		**0.6**	0.6	0.6	0.6
	p_p	Protection from grazing	[0–1]		**0.8**	0.7	0.8	0.7
	r	Respiration rate	[0–0.02]	mmol C d^{-1}	**0.04**	0.06	0.04	0.06
Spine	τ	Coefficient of grazing half saturation	[0–1]		/	**0.9**	/	0.9
Symbiosis	ψ	Symbiont to foraminiferal size ratio	[0–0.05]		/	/	**0.0015**	0.0015
	λ_s	Symbiont autotroph efficiency	[0–1]		/	/	**0.2**	**0.8**
	λ_h	Foraminiferal heterotroph efficiency	[0–1]		/	/	**0.8**	**0.55**

[a] TS12 All scaling parameters are sampled from values of 0 to 1; respiration terms are as per Ward et al. (2018); the symbiont cell size ratio upper bound is calculated from Takagi et al. (2019). For any other plankton group where these traits are not relevant, scaling parameters are set to 1 and cost parameters are set to 0. [b] bn – symbiont-barren non-spinose; bs – symbiont-barren spinose; sn – symbiont-facultative non-spinose; ss – symbiont-obligate spinose. [c] The bold parameters are also shown in other groups with same trait(s).

production (mmol C m^{-3} d^{-1}) into "mmol C m^{-2} d^{-1}" multiplying it by the surface-layer depth (80.8 m) to compare with sediment-trap-generated export observations. The full list of plankton net and sediment trap data sources is in Tables S1 and S2.

Both datasets are classified by species and were regridded into the model resolution following the methods of the core-top data. We calculated the annual average at each grid point to remove seasonality and interannual variability. However, the plankton nets are mostly sampled within 1 month (Fig. S1) and represent a day's snapshot, such that the annual mean biomass is likely overestimated as the nets would be typically sampled during higher production times. In contrast, sediment traps are deployed over 6 months or more (Fig. S1), thereby capturing seasonal variation. Sediment traps were deployed at different depths, typically over 1000 m and thereby deeper than our surface layer. We assume that sediment trap depth has negligible impact on foraminiferal export because foraminiferal tests sink relatively fast due to large size (Takahashi and Be, 1984; Caromel et al., 2014).

6.3.3 Seasonality

To complement the annual comparison, we analysed the modelled seasonal pattern by finding each group's first month with peaking production. We also provided a comparison with plankton net and sediment trap data for most sampled locations in the Supplementary Material. We did not attempt to calculate an M-score TS14 for seasonal model–data comparisons because (1) the temporal coverage of observations is too low at most locations, and (2) the number of available locations is insufficient, creating large spatial bias towards specific oversampled locations.

6.4 Model performance metrics

We used the Mielke measure, or M-score (Watterson, 1996; Watterson et al., 2014), to quantify the model–data fit in (1) relative abundance and (2) annual average biomass/carbon export (Eq. 21). This metric is a non-dimensional transformed mean square error combining mean and variance information CE8 (Gregoire et al., 2011; Hemer and Trenham, 2016). The score spans from −1 (low performance) to 1 (high performance) with 0 representing no predictive skill, and negative values representing negative correlation:

$$M = \frac{2}{\pi} \arcsin \left[\frac{\sum_{i=1}^{n} (M_i - O_i)^2 / n}{\sigma_m^2 + \sigma_o^2 + (\mu_m - \mu_o)^2} \right]. \tag{21}$$

The numerator corresponds to the mean square error, with M_i and O_i denoting the model and observational value in the ith grid point, respectively, and n the total number of grid points. The variance and mean are respectively denoted as σ^2 and μ, with superscripts m and o representing model and observed fields, respectively.

6.5 Global sensitivity analysis

We conduct a global sensitivity analysis (GSA) to explore the model robustness of our 1200 experiments using the PAWN method (Pianosi and Wagener, 2015). This method measures the sensitivity of model outputs (focusing on the M-score here) to different values of 12 input parameters (shown in Table 2). A total M-score is calculated by summing scores of each foraminiferal group in biomass, POC export, and relative abundance (i.e. the total score ranges from −12 to 12). To further measure the uncertainty and robustness of the GSA results, we also apply a bootstrapping method with 1000 resamples. This approach enables us to understand the

confidence intervals of the sensitivity indices without running more experiments (Wagener and Pianosi, 2019). We bootstrapped our data using the *rsample* package (Silge et al., 2021) in the R software environment v4.1 (R Core Team, 2021).

7 Model results

7.1 Model ensemble results

The 1200-member ensemble shows the ability to reproduce the observed POC export and relative abundance in terms of spatial pattern and values (both with the highest total M-score > 1.0) but struggles with capturing the observed biomass (total M-score < 0.5) (Fig. 2). The M-score heatmap (Fig. 2) shows that the model runs cluster into four groups when compared to the three observational datasets. Cluster C, covering most parameter combinations, has an overall low performance in predicting foraminiferal metrics. Cluster D shows an inverse abundance distribution compared to the observation. Cluster B only predicts POC export. Cluster A achieves the highest (i.e. the best) relative-abundance M-score with good predictions for biomass and POC export. Cluster A is also the only cluster with low foraminiferal export, suggesting that low export is associated with parameter values required to have a high total M-score. The sensitivity analysis confirms this, as model performance is sensitive to those parameters controlling the source/sink of foraminiferal export: symbiont size (ψ), autotrophy efficiency (λ_s), and palatability reduction (P_p) (Fig. 3). Models with low export production and higher M-scores tend to have smaller foraminiferal size and symbiont-to-host size ratio (for symbiotic groups) that facilitates the survival of foraminifera in the low-nutrient regions like subtropical gyres. These runs in cluster A also tend to have less than 20 % decreased palatability caused by the shell and a high respiration cost, driving low biomass and export (Fig. S2). In contrast, the runs with negative scoring (Cluster D) have larger foraminiferal size and higher protection against grazing (Fig. S3). These results suggest that foraminiferal body size and the calcification trait have a crucial role in foraminiferal distributions to achieve a match to the observed data. Questions addressing the size trait in more detail, like life history and geographic size distribution (Schmidt et al., 2004b), cannot be answered with this model ensemble as all foraminiferal groups are assigned the same narrow cell size per run, even though they vary between runs.

We selected the model with the parameter set (Table 2) that leads to the highest total M-score (Table 3), hereafter termed the optimal model. This optimal model also has the highest M-score for the relative abundance (group mean $= 0.3$) for each group (Fig. S5) and POC export (group mean $= 0.16$; Fig. S4). More details are discussed in the next sections.
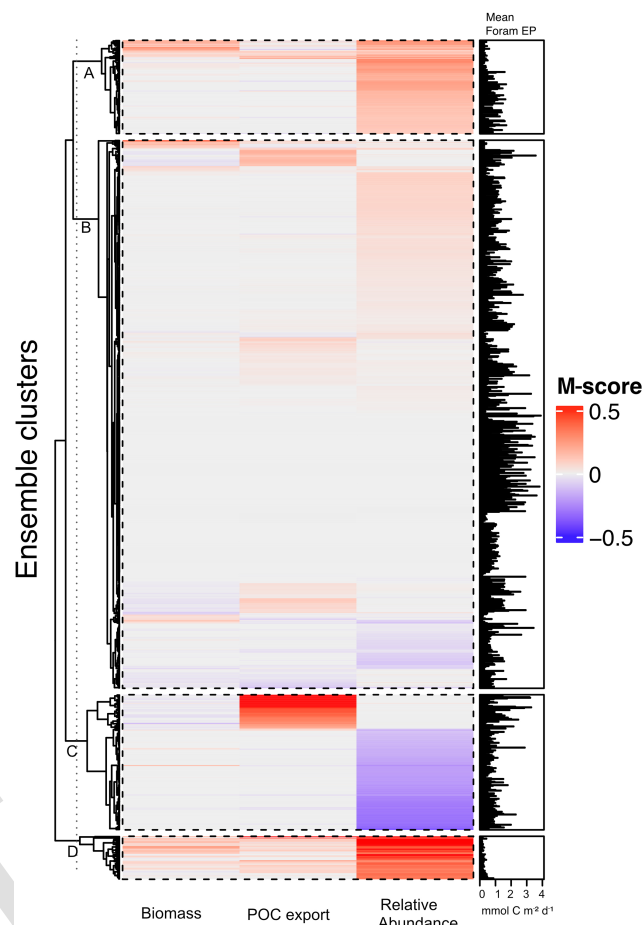
Figure 2. Foraminiferal M-score heatmap of the model ensemble for foraminiferal biomass (plankton net data), POC export (sediment trap data), relative abundance (sediment core-top data). Each of the first three columns shows the M-score sum of the four foraminiferal groups, and the fourth column shows the sum of the left three columns. The right panel shows the global annual mean export production of all foraminiferal groups. The ensemble cluster was derived from a complete linkage clustering algorithm (Jarman, 2020). Higher M-scores have a better performance against observations, whilst negative values stand for negative correlation.

7.2 Relative abundance distribution of foraminiferal groups

Our optimal model run compares generally well with the core-top data showing the relative spatial distribution of the four foraminiferal functional groups (Fig. 4; Table 3). The model agrees with the presence/absence of most groups in the sediment trap and plankton net studies (Figs. 5 and 6). The model suggests that the symbiont-obligate spinose group is the most abundant with a global abundance of 60.7 % (Fig. 4g), dominating the tropical open oceans. In contrast, the symbiont-barren non-spinose (Fig. 4a) and spinose groups (Fig. 4c) dominate in the mid-to-high latitudes, contributing 25.5 % and 9.4 % of the global foraminiferal abun-

Table 3. M-score values across foraminiferal groups for the optimal parameter set. The total foraminiferal M-score is the sum of the M-scores of the four functional groups.

Groups	M-scores				
	Symbiont-barren non-spinose	Symbiont-barren spinose	Symbiont-facultative non-spinose	Symbiont-obligate spinose	Total foraminifera
Biomass	0.19	0.08	−0.05	−0.05	0.17
POC export	0.11	0.07	0.43	0.02	0.63
Relative abundance	0.51	0.35	0.02	0.32	1.20
Row sum	0.81	0.50	0.40	0.29	2.00
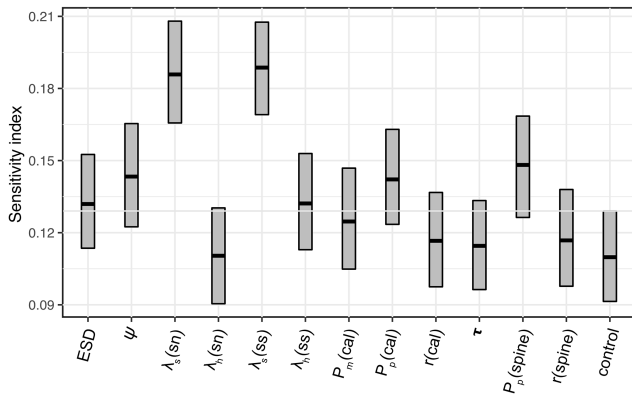
Figure 3. Model parameter sensitivity for overall model performance (summed M-scores). Bar boundaries indicate the 95 % TS15 confidence interval with the thick line showing the mean value. The grey line indicates the non-influential upper limit of the index value as control group. sn – symbiont-facultative non-spinose; ss – symbiont-bearing spinose. cal – the abbreviation of calcification; τ is the spine effect on grazing rate.

dance, respectively (note that the symbiont-barren spinose type contains a small number of taxa with a relatively high contribution to the abundance). The model underestimates the symbiont-facultative group (Fig. 4e) with visible model–data disparities in the eastern equatorial pacific where the sediment data show high abundance. This discrepancy may be due to the resistance to dissolution of some species (e.g. *N. dutertrei*) in high productivity settings as suggested in a previous model study (Lombard et al., 2011). Importantly though, it is not very clear what triggers the presence or absence of symbionts, why this relationship changes and often the taxa are less well studied. The fact that the summed abundances of these two symbiotic groups agree with the observations points to the ability of the symbiont-facultative group to exploit the same benefits as the obligate symbiont-obligate one CE9. It also highlights our need to better understand how often symbiosis is used by the former group and what triggers the switch to the loss of symbionts.

Overall, the modelled root mean square error (RMSE) of relative abundance varies between 12 % and 42 % (Table S5).

This result is comparable to the previous species-based models, like FORAMCLIM (5 %–23 %, Lombard et al., 2011) and PLAFOM (22 %–25 %, Fraile et al., 2009), which rely on well-established foraminiferal species observations. Our results affirm that symbionts and spines and their assumed trade-offs are sufficient to explain significant parts of the relative abundance's geographic distribution. The distribution of non-symbiotic foraminifera in the model follows the biogeography of the prey abundance with high numbers in high-nutrient areas (Fig. S6). In contrast, symbiotic foraminifera grow in low-nutrient areas because they have small-sized symbionts with high nutrient affinities. The model underestimates symbiont-barren spinose foraminifera (mainly *G. bulloides*) in the Arabian Sea and South China Sea (Fig. 4b, d TS16), probably because the model does not include their carnivorous feeding strategy.

7.3 Annual average biomass of foraminiferal groups

The model reproduces low biomass in planktic foraminifera in agreement with the plankton net data (Fig. 5). The global annual mean biomass ranges from 0.001 to 0.010 mmol C m^{-3}, equivalent to 0.08–0.8 mmol C m^{-2}, with the largest contribution from the symbiont-barren non-spinose group (Fig. 7). Integrating across all groups, the model estimates a global foraminiferal biomass of 6.83 Tg C (Fig. 7). Our annual mean biomass estimate is within the MAREDAT project result (0.24–0.94 mmol C m^{-2}) (Schiebel and Movellan, 2012).

The optimal model M-score is low for biomass when compared to the plankton net tow (<0.2; Table 3), possibly because of the low data coverage and the previously mentioned intrinsic seasonal bias in the data compared to annual averages. Our ensemble resulted in higher M-scores for biomass, but at the cost of a lower M-score for relative abundance and export.

7.4 Annual average POC and calcite export of foraminiferal groups

The model reproduces consistent distributions and magnitude of POC export compared to sediment trap data for all
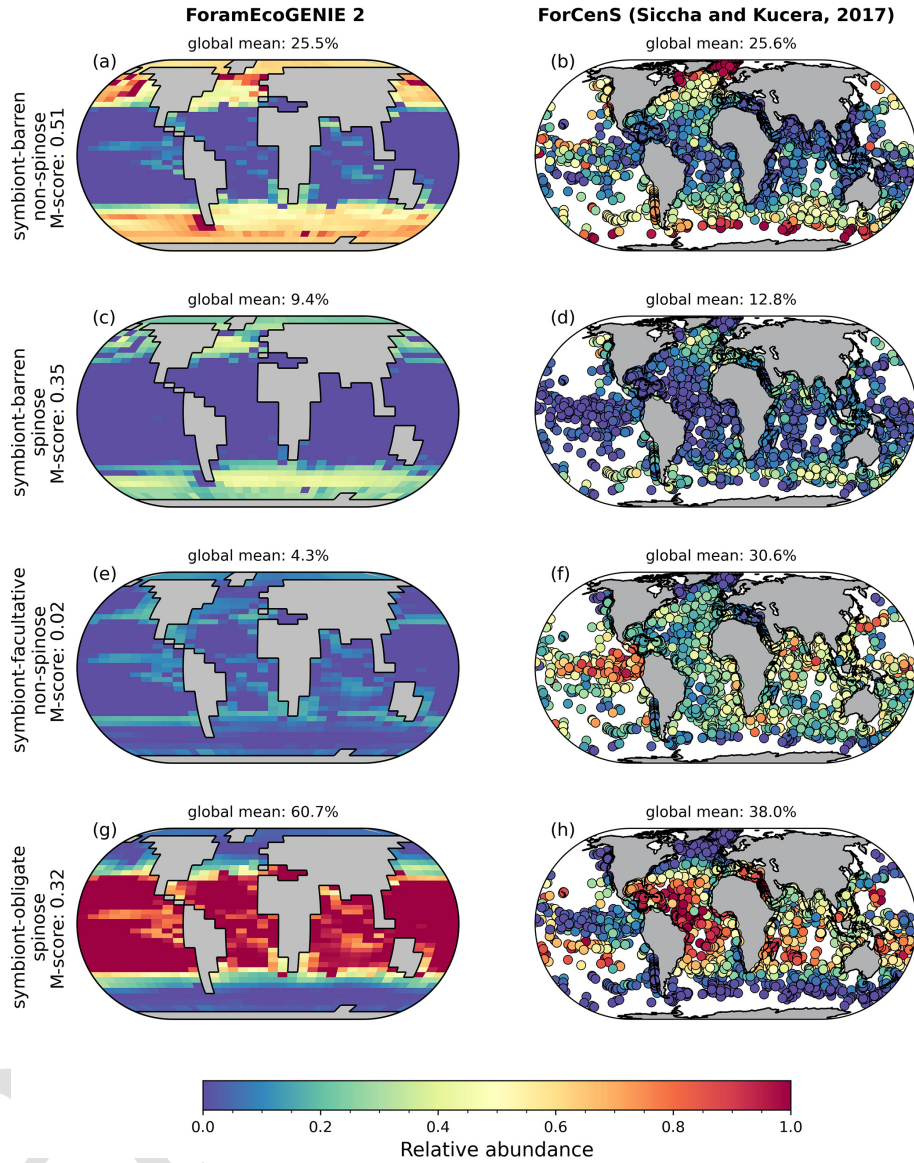
Figure 4. Relative abundance of the four modelled (**a, c, e, g**) planktic foraminiferal function groups, compared to the ForCenS sediment core-top dataset (**b**, **d**, **f**, **h**; Siccha and Kucera, 2017). Subplot titles show the M-scores derived relative to observations and the global mean of relative abundance. TS17

four groups (Figs. 6 and 7). The model estimates a POC export of 0.002–0.031 mmol C m^{-2} d^{-1}, which agrees well with 0.001–0.026 mmol C m^{-2} d^{-1} for the sediment trap data, despite a medium total M-score for the model POC export (0.63) likely caused by the limited geographic coverage akin to the biomass comparison.

Globally, the model suggests a foraminifera-derived organic carbon export of 0.1 Gt C yr^{-1}, dominated by the symbiont-barren non-spinose group (55 %), followed by the symbiont-barren spinose, symbiont-facultative, and symbiont-obligate groups (30 %, 3 %, and 11 %, respectively). Integrating across the ecogroups and using the empirically averaged PIC:POC ratio of 0.36 (Schiebel and Movel-lan, 2012), our model estimates a total calcite flux of pelagic foraminifera of 0.033 Gt PIC yr^{-1} (Fig. 8). This model estimate is at least 5 times smaller than Schiebel (2002)'s estimate of 0.16–0.39 Gt PIC yr^{-1} (for the top 100 m TS19). There are a number of possible reasons for this: (1) a field site selection bias to avoid regions which have very low abundance, (2) our calibration of modelled surface export with deep sediment traps data characterised by typically lower export (as deployed at about 2 km depth), and (3) the temporal variability of observation which is not well captured in the model.
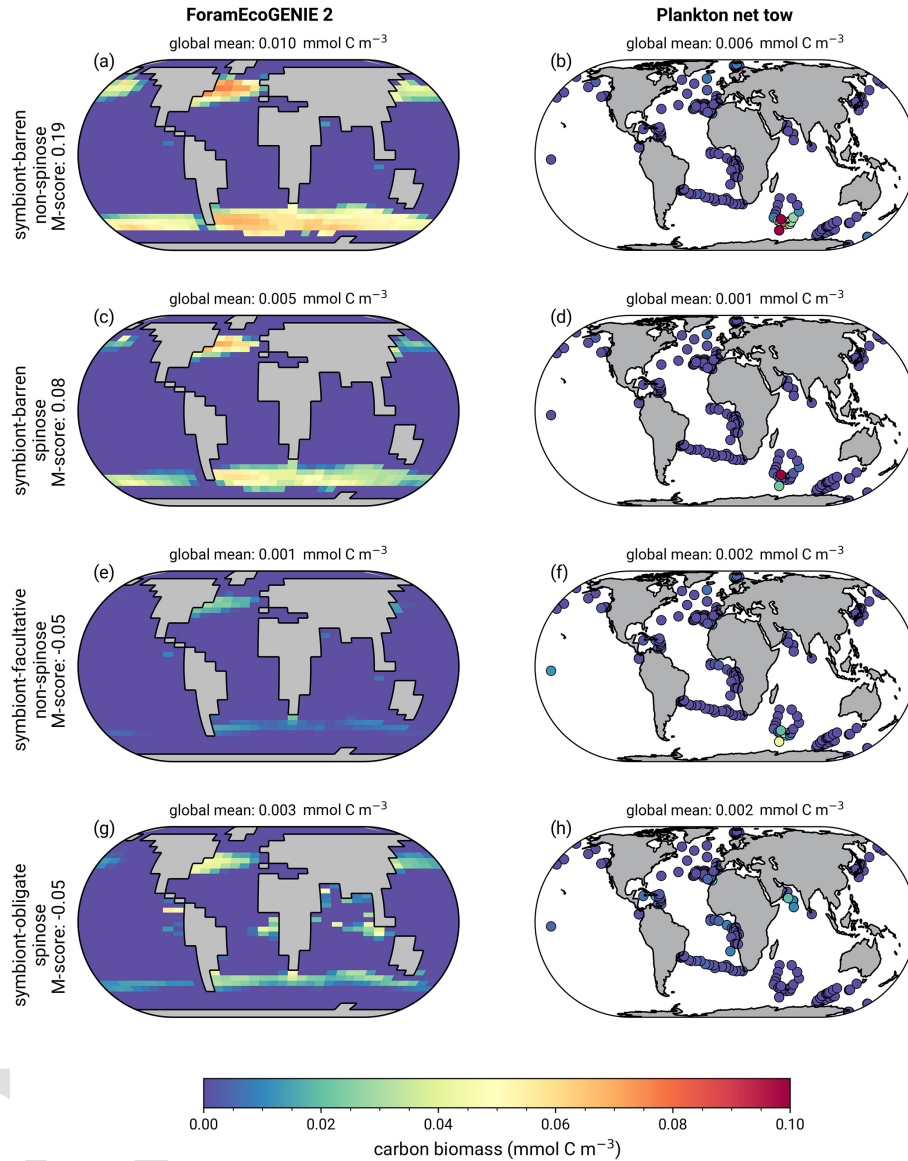
Figure 5. ForamEcoGEnIE 2.0 annual average foraminiferal biomass (mmol C m^{-3}) compared with plankton net data (**b, d, f, h**) for the four main functional groups of planktic foraminifera. TS18

7.5 Seasonal variations of foraminiferal biomass and POC export

Our model shows the different seasonal patterns for each foraminiferal group (Fig. 9), generally consistent with sediment trap study (Jonkers and Kučera, 2015). Jonkers and Kučera (2015) divide the foraminiferal assemblages into a warm group (representing the symbiont-bearing group), cool and temperate group (representing the two symbiont-barren groups), and deep-dwelling group according to their seasonal cycle patterns. The cool/temperate group blooms in spring or summer (Fig. 9a), while the warm group in tropical oceans shows weak and random seasonality (Fig. 9d). The model also captures the earlier-when-warmer signature in the cool/temperate group, i.e. the peaking time is strongly coupled to temperature gradient from high to low latitude (Fig. 9a).

The model generally underestimates seasonal amplitudes of export production (Fig. S7). Plankton net data cannot be compared seasonally due to the very short nature of data collection, despite the general agreement (Fig. S8). The low model export production is not unique to our model and also evident in PLAFOM 2.0 (Kretschmer et al., 2018). Intra-annual variabilities in abundance are driven by the seasonal environmental changes which determine how optimal foraminifera are in the ecological niche. While temperature is often assumed as the primary driver for foraminiferal ecology (Schmidt et al., 2004b; Be and Hamlin, 1967), many
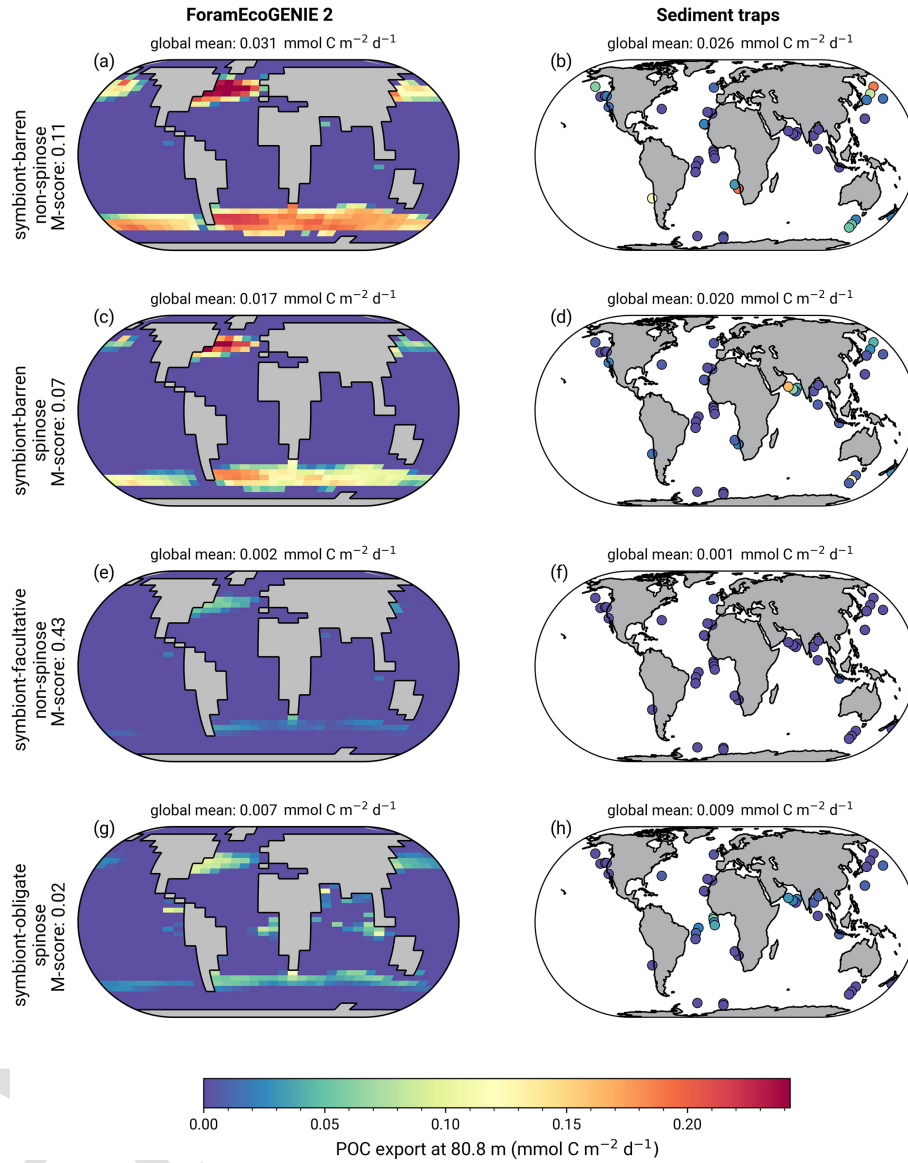
Figure 6. ForamEcoGEnIE 2.0 foraminiferal annual average POC export (mmol C m^{-2} d^{-1}) below the euphotic zone (80.8 m) in comparison to sediment trap samples **(b, d, f, h)**. TS20

other parameters such as primary productivity are correlated with temperature and hence difficult to separate their effects (Jonkers and Kučera, 2015). We suggest that additional functional trait data collections assessing temporal variability, increased geographic coverage, information on deeper-dwelling species, and life history cycle CE10 will contribute to resolve this gap in the future.

8   Comparison to prior model iterations

By comparing ForamEcoGEnIE 2.0 with EcoGEnIE (Ward et al., 2018) and ForamEcoGEnIE 1.0 (Grigoratou et al., 2021a), we find that adding foraminiferal groups increases the total plankton mean body size in the tropical and subtropical regions by roughly 20 % due to the larger size of foraminifera (modelled as 190 μm, Fig. 10c). At the same time, the new iteration decreases the plankton mean size in subpolar areas (<10 %) due to additional grazing pressure by foraminifera on the plankton. In contrast, the total plankton biomass stays almost the same between the model versions because of the low standing stocks of foraminifera. Net primary productivity (NPP) and POC export also do not change, apart from a small drop in the subpolar regions due to enhanced foraminiferal grazing. Therefore, ForamEcoGEnIE 2.0 performs as well as the previous version in terms of total plankton size, biomass, carbon export, and NPP, while capturing foraminiferal diversity and biogeography.
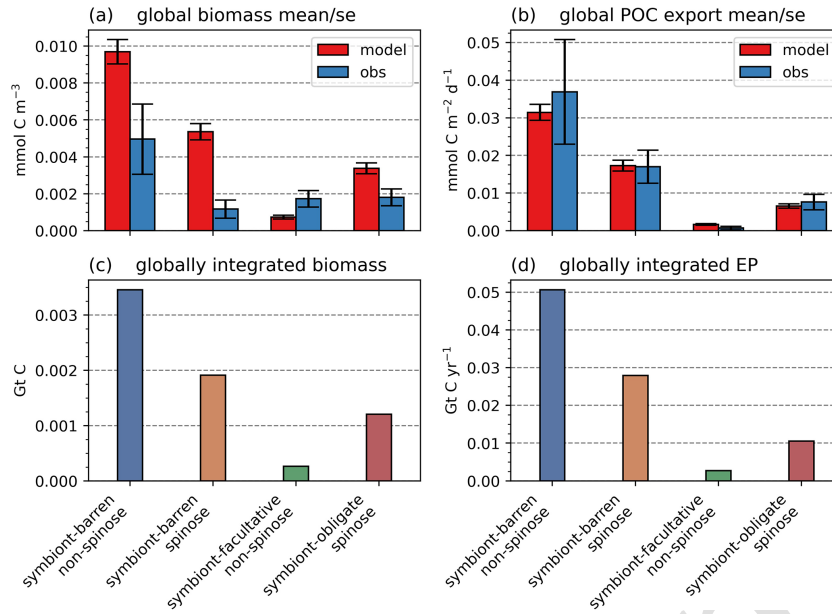
Figure 7. Global ForamEcoGEnIE 2.0 annual average biomass and POC export produced by the four foraminiferal groups: **(a)** modelled (red) and observational (blue) biomass (mmol C m^{-3}); **(b)** POC export below the euphotic zone (mmol C m^{-3} d^{-1}). Bar height and error bar represents the spatial mean value and standard error, respectively. Panels **(c)** and **(d)** show the globally integrated model estimates for **(c)** carbon biomass (Gt C) and **(d)** export production (EP, Gt C yr^{-1}).
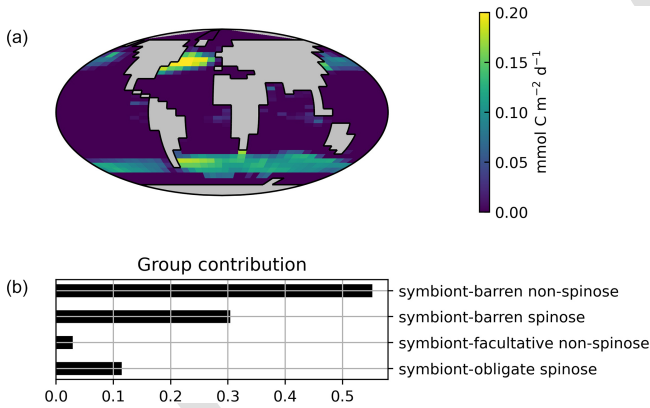
Figure 8. Global ForamEcoGEnIE 2.0 estimates for **(a)** surface foraminiferal calcite flux (at 80.8 m; mmol C m^{-2} yr^{-1}) and **(b)** groups contribution.

Figure 9. The peak month of modelled biomass annual time series of each foraminiferal group in our best ForamEcoGEnIE 2.0 run. Note that the months in Southern Hemisphere indicate the opposite seasonality of the Northern Hemisphere.

While ForamEcoGEnIE 2.0 developments focused on improving diversity in plankton ecology, it also lays the foundation for future studies on the ocean carbon cycle under different climates, past or future. For example, the inclusion of spinose foraminifera is important for particle sinking as they produce and export more calcite than their non-spinose counterpart (Takahashi and Be, 1984). It also opens the door for studies of past climates by expanding the foraminiferal global niche, which may influence the ocean carbon cycles by changing the locations of calcite export and distribution of surface alkalinity. So far, no Earth system model has included

foraminiferal groups acting on the carbon cycle, which would be an important avenue for future research.

9   Model limitation

While making explicit progress in including planktic foraminifera into a modelling framework with a range of traits, ForamEcoGEnIE 2.0 is limited by the non-explicit im-
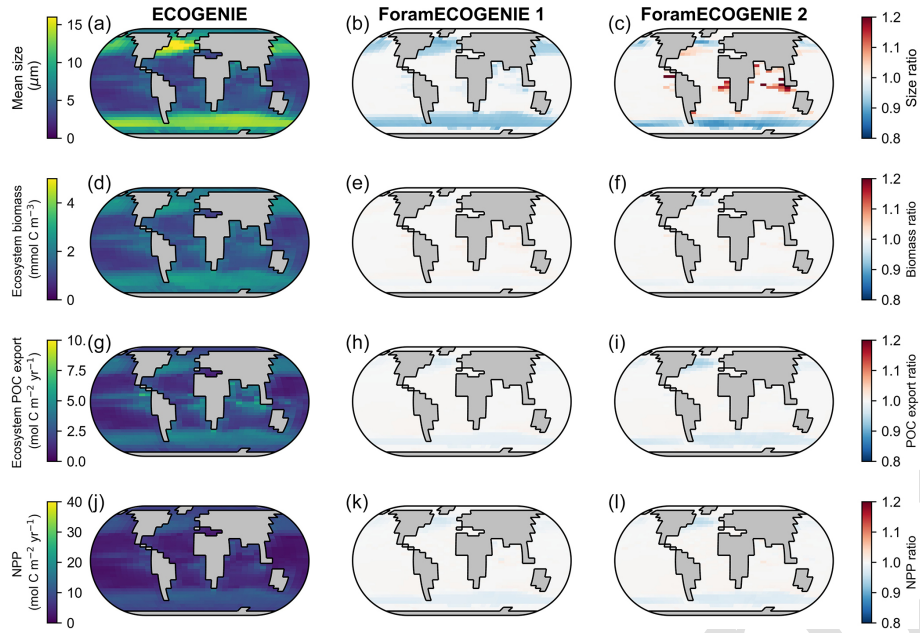
Figure 10. Comparison of the tuned ForamEcoGEnIE 2.0 (third column, with four foraminiferal groups) with EcoGEnIE (first column; Ward et al., 2018) and ForamEcoGEnIE 1.0 (second column, with non-spinose non-symbiont foraminifera only; Grigoratou et al., 2021a) for **(a–c)** total plankton mean size, **(d–f)** total plankton biomass, **(g–i)** total POC export, and **(j–l)** net primary production (NPP). The first column displays absolute values while the latter two show the ratio relative to the first column.

plementations of spines and symbiosis. Currently, our model represents symbiosis based on mixotrophy implementation. According to the definition of mixotrophy types in Mitra et al. (2016), our modelling approach falls within consti-
5 tutive (inherent or innate) mixotrophy rather than the non-constitutive mixotroph grouping. Such indirect photosymbi-otic relationships in the model might miss any differential climate sensitivities of symbiont and host. Furthermore, the current parameterisation of calcification, spines, and symbio-
10 sis will not respond directly to environmental changes, such as bleaching at high temperatures (Edgar et al., 2013) or reduced weight under high CO_2 (Barker and Elderfield, 2002). However, relying on parameterisation is common in EMICs (Claussen et al., 2002), as quantitative experimental studies
15 are lacking now to define the trade-offs and benefits. Furthermore, this lack of understanding of trade-off and their change during development currently makes it impossible to model the life cycle, though further development would be the inclusion of size classes other than 190 μm.
20 Some potentially important trait interactions and physiological variation are not included in the model. For example, the model assumes that the spine and symbiosis are independent. However, observations suggest that foraminifera symbionts are placed along spines during daytime (LeKieffre et
25 al., 2018), increasing the efficiency of the symbiont's photosynthesis due to a higher surface area relative to non-spinose species by avoiding shading.

10   Ecosystem model implementation and complexity

Current coupled Earth system and ecosystem models mostly rely on nutrient–plankton–zooplankton–detritus (NPZD) 30 (Keller et al., 2012; Watanabe et al., 2011) or plankton functional type (PFT) (Moore et al., 2001; Aumont et al., 2015) approaches. The NPZD models focus on biogeochemical fluxes and ignore diversity of phytoplankton and zooplankton. In contrast, PFT models explicitly represent plankton 35 functional types (e.g. diatoms, coccolithophores) and size classes (e.g. picoplankton, nanoplankton, microplankton), improving performance in reconstructing observed patterns like Chl *a* (Quéré et al., 2005) or peak production in oligotrophic areas (Tréguer et al., 2018). Additional traits beyond 40 size, like symbiosis (Suggett et al., 2017) or body extension (Ohman, 2019), play an important role in determining plankton feeding, metabolism, and export efficiency but are often missing in the current generation of coupled models. Trait-based models, such as Darwin (Follows et al., 2007) and 45 EcoGEnIE (Ward et al., 2018), resolve higher plankton diversity by linking key traits with trade-offs (e.g. the allometric relationships for size), allowing a more continuous representation based on physiology (Follows and Dutkiewicz, 2011). This approach enables the inclusion of non-culturable 50 species or species with limited laboratory data. Uniquely, this modelling approach also allows us to characterise extinct taxa and past geological records with different physiologies and niches.

There is still a debate on whether higher ecosystem complexity is needed (Anderson, 2005; Quéré et al., 2005) as more parameters introduce more freedom and longer run time. However, recent studies highlight the importance of biodiversity in the marine biological pumps (Tréguer et al., 2018). The presence of functional groups like diazotroph can alter the response of primary productivity to global warming (Bopp et al., 2022). Therefore, compared to the simple food web structure in current models, ecosystem implementation is very likely going to improve the future prediction of biological carbon pump and carbon cycle (Wilson et al., 2022) building on novel additions in models of ecosystem complexities such as more functional types, variable stoichiometry, and nutrient co-limitations (Séférian et al., 2020).

11 Summary

In this study, we extended the trait-based planktic foraminiferal model, ForamEcoGEnIE, to include symbiosis and spine traits and thereby resolve all main foraminiferal functional groups. Using Latin Hypercube Sampling, we generated 1200 parameter samples and compared these with three global observational sources: sediment surface core-top, plankton nets, and sediment traps. We assessed the model performance describing biogeographic distributions, and quantifying carbon biomass and foraminifera-derived carbon export. Our global sensitivity analysis shows that the symbiosis and the palatability reduction due to the spinose test strongly influences model performance. Our best set of model parameters successfully reproduces the modern biogeographical distribution of the four foraminiferal ecogroups and produces a global annual mean biomass (0.001 to 0.010 mmol C m^{-3}) and foraminifera-derived organic carbon export (0.002–0.031 mmol C m^{-2} d^{-1}) close to observations. The two symbiont-barren groups account for 85 % of standing stocks and foraminifera-derived carbon export. The model accurately reproduces the peak time of seasonal time-series observations of foraminiferal biomass and organic carbon flux but performs poorer in seasonal amplitudes, particularly in upwelling regions. These results provide confidence in the model's ability to explore foraminiferal ecology and diversity in the geological record, for example of the last glacial maximum, as well as helping to solve riddles about their ecology in the past. The trait-based framework of the cGEnIE ecosystem provides the potential to extend the model by presenting more traits such as life history and differential calcification rates across groups.

Code and data availability. The source codes and data are archived in https://doi.org/10.5281/zenodo.6808760 (Ridgwell, 2022 TS21). Experiment configuration and observational dataset can be found in genie-userconfig/MS/yingetal.GMD.2022 TS22. A general manual for cGEnIE model can be found in https://github.com/derpycode/muffindoc TS23.

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/gmd-16-1-2023-supplement.

Author contributions. RY, FMM, JDW, and DNS designed the study. RY, FMM, and JDW developed the model code. RY performed the experiments, data collection, and visualisation. All authors interpreted the data and wrote and edited the original draft.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Acknowledgements. We thank two anonymous reviewers for their helpful and constructive comments. We acknowledge funding from NERC (NE/P019439/1 for DNS and Jamie D. Wilson TS24), NE/N011708/1 (Fanny M. Monteiro TS25), TMS Angelina Messina Grant (Rui Ying TS26), China Scholarship Council (Rui Ying TS27, No. 202006380070), and AXA Research Fund Postdoctoral Fellowship (Jamie D. Wilson TS28). We would like to thank Maria Grigoratou for her insights on the model development and Ruby Barrett for the proofreading.

Financial support. This research has been supported by the Natural Environment Research Council (grant nos. NE/P019439/1 and NE/N011708/1), the China Scholarship Council (grant no. 202006380070), and the AXA Research Fund (AXA Research Fund Postdoctoral Fellowship). TS29

Review statement. This paper was edited by Paul Halloran and reviewed by two anonymous referees.

References

TS30

Adloff, M., Ridgwell, A., Monteiro, F. M., Parkinson, I. J., Dickson, A. J., Pogge von Strandmann, P. A. E., Fantle, M. S., and Greene, S. E.: Inclusion of a suite of weathering tracers in the cGENIE Earth system model – muffin release v.0.9.23, Geosci. Model Dev., 14, 4187–4223, https://doi.org/10.5194/gmd-14-4187-2021, 2021. TS31

Andersen, K. H., Berge, T., Gonçalves, R. J., Hartvig, M., Heuschele, J., Hylander, S., Jacobsen, N. S., Lindemann, C., Martens, E. A., Neuheimer, A. B., Olsson, K., Palacz, A., Prowe, A. E. F., Sainmont, J., Traving, S. J., Visser, A. W., Wadhwa, N., and Kiørboe, T.: Characteristic Sizes of Life in the Oceans, from Bacteria to Whales, Annu. Rev. Mar. Sci., 8, 217–241, https://doi.org/10/f3pdzr, 2016.

Anderson, O. R. and Bé, A. W. H.: The ultrastructure of a planktonic foraminifer, Globigerinoides sacculifer (Brady), and its symbiotic dinoflagellates, J. Foramin. Res., 6, 1–21 TS32, https://doi.org/10.2113/gsjfr.6.1.1, 1976.

Anderson, O. R., Spindler, M., Bé, A. W. H., and Hemleben, Ch.: Trophic activity of planktonic foraminifera, J. Mar. Biol. Ass., 59, 791–799, https://doi.org/10.1017/S002531540004577X, 1979.

Anderson, R. P.: When and how should biotic interactions be considered in models of species niches and distributions?, J. Biogeogr., 44, 8–17, https://doi.org/10.1111/jbi.12825, 2017.

Anderson, T. R.: Plankton functional type modelling: running before we can walk?, J. Plankton Res., 27, 1073–1081, https://doi.org/10.1093/plankt/fbi076, 2005.

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

Aze, T., Ezard, T. H. G., Purvis, A., Coxall, H. K., Stewart, D. R. M., Wade, B. S., and Pearson, P. N.: A phylogeny of Cenozoic macroperforate planktonic foraminifera from fossil data, Biol. Rev., 86, 900–927, https://doi.org/10/dvjwx6, 2011.

Barker, S. and Elderfield, H.: Foraminiferal Calcification Response to Glacial-Interglacial Changes in Atmospheric CO_2, Science, 297, 833–836, https://doi.org/10.1126/science.1072815, 2002.

Be, A. W. H. and Hamlin, W. H.: Ecology of Recent Planktonic Foraminifera: Part 3: Distribution in the North Atlantic during the Summer of 1962, Micropaleontology, 13, 87 TS33, https://doi.org/10/b99v5t, 1967.

Bé, A. W. H., Spero, H. J., and Anderson, O. R.: Effects of symbiont elimination and reinfection on the life processes of the planktonic foraminifer Globigerinoides sacculifer, Mar. Biol., 70, 73–86, https://doi.org/10.1007/BF00397298, 1982.

Bec, B., Collos, Y., Vaquer, A., Mouillot, D., and Souchu, P.: Growth rate peaks at intermediate cell size in marine photosynthetic picoeukaryotes, Limnol. Oceanogr., 53, 863–867, https://doi.org/10.4319/lo.2008.53.2.0863, 2008.

Bopp, L., Aumont, O., Kwiatkowski, L., Clerc, C., Dupont, L., Ethé, C., Gorgues, T., Séférian, R., and Tagliabue, A.: Diazotrophy as a key driver of the response of marine net primary productivity to climate change, Biogeosciences, 19, 4267–4285, https://doi.org/10.5194/bg-19-4267-2022, 2022.

Brovkin, V., Brücher, T., Kleinen, T., Zaehle, S., Joos, F., Roth, R., Spahni, R., Schmitt, J., Fischer, H., Leuenberger, M., Stone, E. J., Ridgwell, A., Chappellaz, J., Kehrwald, N., Barbante, C., Blunier, T., and Dahl Jensen, D.: Comparative carbon cycle dynamics of the present and last interglacial, Quaternary Sci. Rev., 137, 15–32, https://doi.org/10.1016/j.quascirev.2016.01.028, 2016. TS34

Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., and West, G. B.: Toward a Metabolic Theory of Ecology, Ecology, TS35, 1771–1789, https://doi.org/10.1890/03-9000@10.1002/(ISSN)1939-9170.MacArthurAward TS36, 2004.

Buckley, L. B., Urban, M. C., Angilletta, M. J., Crozier, L. G., Rissler, L. J., and Sears, M. W.: Can mechanism inform species' distribution models?, Ecol. Lett., 13, 1041–1054, https://doi.org/10.1111/j.1461-0248.2010.01479.x, 2010.

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.

Buitenhuis, E. T., Quéré, C. L., Bednaršek, N., and Schiebel, R.: Large Contribution of Pteropods to Shallow $CaCO_3$ Export, Global Biogeochem. Cy., 33, 458–468, https://doi.org/10/gjpnzt, 2019.

Cao, L., Eby, M., Ridgwell, A., Caldeira, K., Archer, D., Ishida, A., Joos, F., Matsumoto, K., Mikolajewicz, U., Mouchet, A., Orr, J. C., Plattner, G.-K., Schlitzer, R., Tokos, K., Totterdell, I., Tschumi, T., Yamanaka, Y., and Yool, A.: The role of ocean transport in the uptake of anthropogenic CO_2, Biogeosciences, 6, 375–390, https://doi.org/10.5194/bg-6-375-2009, 2009. TS37

Caromel, A. G. M., Schmidt, D. N., Phillips, J. C., and Rayfield, E. J.: Hydrodynamic constraints on the evolution and ecology of planktic foraminifera, Mar. Micropaleontol., 106, 69–78, https://doi.org/10.1016/j.marmicro.2014.01.002, 2014.

Castellani, M., Våge, S., Strand, E., Thingstad, T. F., and Giske, J.: The Scaled Subspaces Method: A new trait-based approach to model communities of populations with largely inhomogeneous density, Ecol. Model., 251, 173–186, https://doi.org/10.1016/j.ecolmodel.2012.12.006, 2013.

Claussen, M., Mysak, L., Weaver, A., Crucifix, M., Fichefet, T., Loutre, M.-F., Weber, S., Alcamo, J., Alexeev, V., Berger, A., Calov, R., Ganopolski, A., Goosse, H., Lohmann, G., Lunkeit, F., Mokhov, I., Petoukhov, V., Stone, P., and Wang, Z.: Earth system models of intermediate complexity: closing the gap in the spectrum of climate system models, Clim. Dynam., 18, 579–586, https://doi.org/10.1007/s00382-001-0200-1, 2002.

Daniels, C. J., Poulton, A. J., Balch, W. M., Marañón, E., Adey, T., Bowler, B. C., Cermeño, P., Charalampopoulou, A., Crawford, D. W., Drapeau, D., Feng, Y., Fernández, A., Fernández, E., Fragoso, G. M., González, N., Graziano, L. M., Heslop, R., Holligan, P. M., Hopkins, J., Huete-Ortega, M., Hutchins, D. A., Lam, P. J., Lipsen, M. S., López-Sandoval, D. C., Loucaides, S., Marchetti, A., Mayers, K. M. J., Rees, A. P., Sobrino, C., Tynan, E., and Tyrrell, T.: A global compilation of coccolithophore calcification rates, Earth Syst. Sci. Data, 10, 1859–1876, https://doi.org/10.5194/essd-10-1859-2018, 2018.

Droop, M. R.: Vitamin B_{12} and Marine Ecology. IV. The Kinetics of Uptake, Growth and Inhibition in *Monochrysis Lutheri*, J. Mar. Biol. Ass., 48, 689–733, https://doi.org/10.1017/S0025315400019238, 1968.

Edgar, K. M., Bohaty, S. M., Gibbs, S. J., Sexton, P. F., Norris, R. D., and Wilson, P. A.: Symbiont "bleaching" in planktic foraminifera during the Middle Eocene Climatic Optimum, Geology, 41, 15–18, https://doi.org/10/f4jwbp, 2013.

Edwards, K. F., Thomas, M. K., Klausmeier, C. A., and Litchman, E.: Allometric scaling and taxonomic variation in nutrient utilization traits and maximum growth rate of phytoplankton, Limnol. Oceanogr., 57, 554–566, https://doi.org/10.4319/lo.2012.57.2.0554, 2012.

Edwards, N. R. and Marsh, R.: Uncertainties due to transport-parameter sensitivity in an efficient 3-D ocean-climate model, Clim. Dynam., 24, 415–433, https://doi.org/10/fcvq9k, 2005.

Ezard, T. H. G., Aze, T., Pearson, P. N., and Purvis, A.: Interplay Between Changing Climate and Species' Ecology Drives Macroevolutionary Dynamics, Science, 332, 349–351, https://doi.org/10/bd77gm, 2011.

Fiksen, Ø., Follows, M. J., and Aksnes, D. L.: Trait-based models of nutrient uptake in microbes extend the Michaelis-Menten framework, Limnol. Oceanogr., 58, 193–202, https://doi.org/10.4319/lo.2013.58.1.0193, 2013.

Fisher, R. A., Corbet, A. S., and Williams, C. B.: The Relation Between the Number of Species and the Number of Individuals in a Random Sample of an Animal Population, J. Anim. Ecol., 12, 42–58, https://doi.org/10.2307/1411, 1943.

Flynn, K. J.: The importance of the form of the quota curve and control of non-limiting nutrient transport in phytoplankton models, J. Plankton Res., 30, 423–438, https://doi.org/10.1093/plankt/fbn007, 2008.

Follows, M. J. and Dutkiewicz, S.: Modeling Diverse Communities of Marine Microbes, Annu. Rev. Mar. Sci., 3, 427–451, https://doi.org/10/b3w27x, 2011.

Follows, M. J., Dutkiewicz, S., Grant, S., and Chisholm, S. W.: Emergent Biogeography of Microbial Communities in a Model Ocean, Science, 315, 1843–1846, https://doi.org/10/bf6j95, 2007.

Fraile, I., Schulz, M., Mulitza, S., and Kucera, M.: Predicting the global distribution of planktonic foraminifera using a dynamic ecosystem model, Biogeosciences, 5, 891–911, https://doi.org/10.5194/bg-5-891-2008, 2008. TS38

Fraile, I., Schulz, M., Mulitza, S., Merkel, U., Prange, M., and Paul, A.: Modeling the seasonal distribution of planktonic foraminifera during the Last Glacial Maximum, Paleoceanography, 24, TS39, https://doi.org/10/dh2z9t, 2009.

Gaskell, D. E., Ohman, M. D., and Hull, P. M.: Zooglider-Based Measurements of Planktonic Foraminifera in the California Current System, J. Foramin. Res., 49, 390–404, https://doi.org/10.2113/gsjfr.49.4.390, 2019.

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

Gregoire, L. J., Valdes, P. J., Payne, A. J., and Kahana, R.: Optimal tuning of a GCM using modern and glacial constraints, Clim. Dynam., 37, 705–719, https://doi.org/10.1007/s00382-010-0934-8, 2011.

Grigoratou, M., Monteiro, F. M., Schmidt, D. N., Wilson, J. D., Ward, B. A., and Ridgwell, A.: A trait-based modelling approach to planktonic foraminifera ecology, Biogeosciences, 16, 1469–1492, https://doi.org/10.5194/bg-16-1469-2019, 2019. TS40

Grigoratou, M., Monteiro, F. M., Wilson, J. D., Ridgwell, A., and Schmidt, D. N.: Exploring the impact of climate change on the global distribution of non-spinose planktonic foraminifera using a trait-based ecosystem model, Glob. Change Biol., gcb.15964 TS41, https://doi.org/10.1111/gcb.15964, 2021a.

Grigoratou, M., Monteiro, F. M., Ridgwell, A., and Schmidt, D. N.: Investigating the benefits and costs of spines and diet on planktonic foraminifera distribution with a trait-based ecosystem model, Mar. Micropaleontol., TS42, 102004, https://doi.org/10/gkbn65, 2021b.

Hemer, M. A. and Trenham, C. E.: Evaluation of a CMIP5 derived dynamical global wind wave climate model ensemble, Ocean Model., 103, 190–203, https://doi.org/10.1016/j.ocemod.2015.10.009, 2016.

Hemleben, C., Spindler, M., and Erson, O. R.: Modern planktonic foraminifera, 363 pp., 1989. TS43

Henehan, M. J., Ridgwell, A., Thomas, E., Zhang, S., Alegret, L., Schmidt, D. N., Rae, J. W. B., Witts, J. D., Landman, N. H., Greene, S. E., Huber, B. T., Super, J. R., Planavsky, N. J., and Hull, P. M.: Rapid ocean acidification and protracted Earth system recovery followed the end-Cretaceous Chicxulub impact, P. Natl. Acad. Sci. USA, 116, 22500–22504, https://doi.org/10/ggbnrm, 2019.

Holling, C. S.: The Functional Response of Predators to Prey Density and its Role in Mimicry and Population Regulation, Mem. Entomol. Soc. Can., 97, 5–60, https://doi.org/10/fhjtms, 1965.

Hönisch, B., Ridgwell, A., Schmidt, D. N., Thomas, E., Gibbs, S. J., Sluijs, A., Zeebe, R., Kump, L., Martindale, R. C., Greene, S. E., Kiessling, W., Ries, J., Zachos, J. C., Royer, D. L., Barker, S., Marchitto, T. M., Moyer, R., Pelejero, C., Ziveri, P., Foster, G. L., and Williams, B.: The Geological Record of Ocean Acidification, Science, 335, 1058–1063, https://doi.org/10/gdj3zf, 2012.

Jarman, A. M.: Hierarchical cluster analysis: Comparison of single linkage, complete linkage, average linkage and centroid linkage method, Georgia Southern University, 2020. TS44

Jonkers, L. and Kučera, M.: Global analysis of seasonality in the shell flux of extant planktonic Foraminifera, Biogeosciences, 12, 2207–2226, https://doi.org/10.5194/bg-12-2207-2015, 2015.

Keller, D. P., Oschlies, A., and Eby, M.: A new marine ecosystem model for the University of Victoria Earth System Climate Model, Geosci. Model Dev., 5, 1195–1220, https://doi.org/10.5194/gmd-5-1195-2012, 2012.

Kiørboe, T., Visser, A., and Andersen, K. H.: A trait-based approach to ocean ecology, ICES J. Mar. Sci., 75, 1849–1863, https://doi.org/10.1093/icesjms/fsy090, 2018.

Kretschmer, K., Kucera, M., and Schulz, M.: Modeling the distribution and seasonality of Neogloboquadrina pachyderma in the North Atlantic Ocean during Heinrich Stadial 1, Paleoceanography, 31, 986–1010, https://doi.org/10.1002/2015PA002819, 2016.

Kretschmer, K., Jonkers, L., Kucera, M., and Schulz, M.: Modeling seasonal and vertical habitats of planktonic foraminifera on a global scale, Biogeosciences, 15, 4405–4429, https://doi.org/10.5194/bg-15-4405-2018, 2018. TS45

Kucera, M. and Schonfeld, J.: The origin of modern oceanic foraminiferal faunas and Neogene climate change, in: Deep-Time Perspectives on Climate Change: Marrying the Signal from Computer Models and Biological Proxies, edited by: Williams, M., Haywood, A. M., Gregory, F. J., and Schmidt, D. N., The Geological Society of London on behalf of The Micropalaeontological Society, 409–425, https://doi.org/10.1144/TMS002.18, 2007.

LeKieffre, C., Spero, H. J., Russell, A. D., Fehrenbacher, J. S., Geslin, E., and Meibom, A.: Assimilation, translocation, and utilization of carbon between photosynthetic symbiotic dinoflagellates and their planktic foraminifera host, Mar. Biol., 165, 104 TS46, https://doi.org/10/gj35br, 2018.

Lombard, F., Labeyrie, L., Michel, E., Bopp, L., Cortijo, E., Retailleau, S., Howa, H., and Jorissen, F.: Modelling planktic foraminifer growth and distribution using an ecophysiological multi-species approach, Biogeosciences, 8, 853–873, https://doi.org/10.5194/bg-8-853-2011, 2011.

Marsh, R., Müller, S. A., Yool, A., and Edwards, N. R.: Incorporation of the C-GOLDSTEIN efficient climate model into the GENIE framework: "eb_go_gs" configurations of GENIE,

Geosci. Model Dev., 4, 957–992, https://doi.org/10.5194/gmd-4-957-2011, 2011.

Meyer, K. M., Kump, L. R., and Ridgwell, A.: Biogeochemical controls on photic-zone euxinia during the end-Permian mass extinction, Geol. TS47, 36, 747 TS48, https://doi.org/10.1130/G24618A.1, 2008. TS49

Mitra, A., Flynn, K. J., Tillmann, U., Raven, J. A., Caron, D., Stoecker, D. K., Not, F., Hansen, P. J., Hallegraeff, G., Sanders, R., Wilken, S., McManus, G., Johnson, M., Pitta, P., Våge, S., Berge, T., Calbet, A., Thingstad, F., Jeong, H. J., Burkholder, J., Glibert, P. M., Granéli, E., and Lundgren, V.: Defining Planktonic Protist Functional Groups on Mechanisms for Energy and Nutrient Acquisition: Incorporation of Diverse Mixotrophic Strategies, Protist, 167, 106–120, https://doi.org/10/f3p5h2, 2016.

Monteiro, F. M., Follows, M. J., and Dutkiewicz, S.: Distribution of diverse nitrogen fixers in the global ocean, Global Biogeochem. Cy., 24, n/a–n/a TS50, https://doi.org/10/ctkc4h, 2010.

Moore, J. K., Doney, S. C., Kleypas, J. A., Glover, D. M., and Fung, I. Y.: An intermediate complexity marine ecosystem model for the global domain, Deep-Sea Res. Pt. II, 49, 403–462, https://doi.org/10/bp99zn, 2001.

Ohman, M. D.: A sea of tentacles: optically discernible traits resolved from planktonic organisms in situ, ICES J. Mar. Sci., 76, 1959–1972, https://doi.org/10.1093/icesjms/fsz184, 2019.

Ortiz, J. D., Mix, A. C., and Collier, R. W.: Environmental control of living symbiotic and asymbiotic foraminifera of the California Current, Paleoceanography, 10, 987–1009, https://doi.org/10/ft8jc7, 1995.

Pianosi, F. and Wagener, T.: A simple and efficient method for global sensitivity analysis based on cumulative distribution functions, Environ. Modell. Softw., 67, 1–11, https://doi.org/10/f677qs, 2015.

Quéré, C. L., Harrison, S. P., Prentice, I. C., Buitenhuis, E. T., Aumont, O., Bopp, L., Claustre, H., Cunha, L. C. D., Geider, R., Giraud, X., Klaas, C., Kohfeld, K. E., Legendre, L., Manizza, M., Platt, T., Rivkin, R. B., Sathyendranath, S., Uitz, J., Watson, A. J., and Wolf-Gladrow, D.: Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models, Glob. Change Biol., 11, 2016–2040, https://doi.org/10/cm9nzc, 2005.

Rae, J. W. B., Gray, W. R., Wills, R. C. J., Eisenman, I., Fitzhugh, B., Fotheringham, M., Littley, E. F. M., Rafter, P. A., Rees-Owen, R., Ridgwell, A., Taylor, B., and Burke, A.: Overturning circulation, nutrient limitation, and warming in the Glacial North Pacific, Science Advances, 6, eabd1654, https://doi.org/10/ghrj7m, 2020.

R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2021. TS51

Renaud, S. and Schmidt, D. N.: Habitat tracking as a response of the planktic foraminifer *Globorotalia truncatulinoides* to environmental fluctuations during the last 140 kyr, Mar. Micropaleontol., 49, 97–122, https://doi.org/10/bgp3cz, 2003.

Ridgwell, A. and Hargreaves, J. C.: Regulation of atmospheric CO_2 by deep-sea sediments in an Earth system model: Regulation of CO_2 by Deep-Sea Sediments TS52, Global Biogeochem. Cy., 21, n/a–n/a TS53, https://doi.org/10.1029/2006GB002764, 2007.

Ridgwell, A. and Schmidt, D. N.: Past constraints on the vulnerability of marine calcifiers to massive carbon dioxide release, Nat. Geosci., 3, 196–200, https://doi.org/10.1038/ngeo755, 2010.

Ridgwell, A., Hargreaves, J. C., Edwards, N. R., Annan, J. D., 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. TS54

Ridgwell, A., ruiying-ocean, Reinhard, C., van de Velde, S., Adloff, M., Monteiro, F., Hülse, D., Wilson, J., Ward, B., Vervoort, P., Kirtland, S., Turner, and Li, M. TS55: ruiying-ocean/cgenie.muffin: ForamEcoGENIE (v0.9.26), Zenodo [code and data set], https://doi.org/10.5281/zenodo.7093077 TS56, 2022. TS57

Roy, T., Lombard, F., Bopp, L., and Gehlen, M.: Projected impacts of climate change and ocean acidification on the global biogeography of planktonic Foraminifera, Biogeosciences, 12, 2873–2889, https://doi.org/10.5194/bg-12-2873-2015, 2015. TS58

Salter, I., Schiebel, R., Ziveri, P., Movellan, A., Lampitt, R., and Wolff, G. A.: Carbonate counter pump stimulated by natural iron fertilization in the Polar Frontal Zone, Nat. Geosci., 7, 885–889, https://doi.org/10.1038/ngeo2285, 2014.

Sarrazin, F., Pianosi, F., and Wagener, T.: Global Sensitivity Analysis of environmental models: Convergence and validation, Environ. Modell. Softw., 79, 135–152, https://doi.org/10/f8n8kp, 2016.

Schiebel, R.: Planktic foraminiferal sedimentation and the marine calcite budget, Global Biogeochem. Cy., 16, 3-1–3-21 TS59, https://doi.org/10/bdxfhs, 2002.

Schiebel, R. and Hemleben, C.: Planktic Foraminifers in the Modern Ocean, Springer Berlin Heidelberg, Berlin, Heidelberg, https://doi.org/10.1007/978-3-662-50297-6, 2017.

Schiebel, R. and Movellan, A.: First-order estimate of the planktic foraminifer biomass in the modern ocean, Earth Syst. Sci. Data, 4, 75–89, https://doi.org/10.5194/essd-4-75-2012, 2012.

Schmidt, D. N., Thierstein, H. R., Bollmann, J., and Schiebel, R.: Abiotic Forcing of Plankton Evolution in the Cenozoic, Science, 303, 207–210, https://doi.org/10/b37mvn, 2004a.

Schmidt, D. N., Renaud, S., Bollmann, J., Schiebel, R., and Thierstein, H. R.: Size distribution of Holocene planktic foraminifer assemblages: biogeography, ecology and adaptation, Mar. Micropaleontol., 50, 319–338, https://doi.org/10/b6nqrj, 2004b.

van Sebille, E., Scussolini, P., Durgadoo, J. V., Peeters, F. J. C., Biastoch, A., Weijer, W., Turney, C., Paris, C. B., and Zahn, R.: Ocean currents generate large footprints in marine palaeoclimate proxies, Nat. Commun., 6, 6521 TS60, https://doi.org/10/f67xqv, 2015.

Séférian, R., Berthet, S., Yool, A., Palmiéri, J., Bopp, L., Tagliabue, A., Kwiatkowski, L., Aumont, O., Christian, J., Dunne, J., Gehlen, M., Ilyina, T., John, J. G., Li, H., Long, M. C., Luo, J. Y., Nakano, H., Romanou, A., Schwinger, J., Stock, C., Santana-Falcón, Y., Takano, Y., Tjiputra, J., Tsujino, H., Watanabe, M., Wu, T., Wu, F., and Yamamoto, A.: Tracking Improvement in Simulated Marine Biogeochemistry Between CMIP5 and CMIP6, Curr. Clim. Change Rep., 6, 95–119, https://doi.org/10.1007/s40641-020-00160-0, 2020.

Siccha, M. and Kucera, M.: ForCenS, a curated database of planktonic foraminifera census counts in ma-

rine surface sediment samples, Sci. Data, 4, 170109, https://doi.org/10.1038/sdata.2017.109, 2017.

Silge, J., Chow, F., Kuhn, M., and Wickham, H.: rsample: General resampling infrastructure, 2021. TS61

Spero, H. J. and Parker, S. L.: Photosynthesis in the symbiotic planktonic foraminifer Orbulina universa, and its potential contribution to oceanic primary productivity, J. Foramin. Res., 15, 273–281, https://doi.org/10/c2rt2q, 1985.

Suggett, D. J., Warner, M. E., and Leggat, W.: Symbiotic Dinoflagellate Functional Diversity Mediates Coral Survival under Ecological Crisis, Trends Ecol. Evol., 32, 735–745, https://doi.org/10.1016/j.tree.2017.07.013, 2017.

Sunagawa, S., Acinas, S. G., Bork, P., Bowler, C., Eveillard, D., Gorsky, G., Guidi, L., Iudicone, D., Karsenti, E., Lombard, F., Ogata, H., Pesant, S., Sullivan, M. B., Wincker, P., and de Vargas, C.: Tara Oceans: towards global ocean ecosystems biology, Nat. Rev. Microbiol., 18, 428–445, https://doi.org/10.1038/s41579-020-0364-5, 2020.

Takagi, H., Kimoto, K., Fujiki, T., Saito, H., Schmidt, C., Kucera, M., and Moriya, K.: Characterizing photosymbiosis in modern planktonic foraminifera, Biogeosciences, 16, 3377–3396, https://doi.org/10.5194/bg-16-3377-2019, 2019. TS62

Takahashi, K. and Be, A. W. H.: Planktonic foraminifera: factors controlling sinking speeds, Deep-Sea Res., 31, 1477–1500, https://doi.org/10.1016/0198-0149(84)90083-9, 1984.

Tierney, J. E., Poulsen, C. J., Montañez, I. P., Bhattacharya, T., Feng, R., Ford, H. L., Hönisch, B., Inglis, G. N., Petersen, S. V., Sagoo, N., Tabor, C. R., Thirumalai, K., Zhu, J., Burls, N. J., Foster, G. L., Goddéris, Y., Huber, B. T., Ivany, L. C., Kirtland Turner, S., Lunt, D. J., McElwain, J. C., Mills, B. J. W., Otto-Bliesner, B. L., Ridgwell, A., and Zhang, Y. G.: Past climates inform our future, Science, 370, eaay3701, https://doi.org/10/gh6c3g, 2020.

Todd, C. L., Schmidt, D. N., Robinson, M. M., and Schepper, S. D.: Planktic Foraminiferal Test Size and Weight Response to the Late Pliocene Environment, Paleoceanography and Paleoclimatology, 35, e2019PA003738, https://doi.org/10/ghrd4r, 2020.

Tréguer, P., Bowler, C., Moriceau, B., Dutkiewicz, S., Gehlen, M., Aumont, O., Bittner, L., Dugdale, R., Finkel, Z., Iudicone, D., Jahn, O., Guidi, L., Lasbleiz, M., Leblanc, K., Levy, M., and Pondaven, P.: Influence of diatom diversity on the ocean biological carbon pump, Nat. Geosci., 11, 27–37, https://doi.org/10/gcxznd, 2018.

Uhle, M. E., Macko, S. A., Spero, H. J., Lea, D. W., Ruddiman, W. F., and Engel, M. H.: The fate of nitrogen in the Orbulina universa foraminifera-symbiont system determined by nitrogen isotope analyses of shell-bound organic matter, Limnol. Oceanogr., 44, 1968–1977, https://doi.org/10/ffgtfw, 1999.

Våge, S., Castellani, M., Giske, J., and Thingstad, T. F.: Successful strategies in size structured mixotrophic food webs, Aquat. Ecol., 47, 329–347, https://doi.org/10.1007/s10452-013-9447-y, 2013.

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), Geosci. Model Dev., 14, 2713–2745, https://doi.org/10.5194/gmd-14-2713-2021, 2021.

Ward, B. A. and Follows, M. J.: Marine mixotrophy increases trophic transfer efficiency, mean organism size, and vertical carbon flux, P. Natl. Acad. Sci. USA, 113, 2958–2963, https://doi.org/10/ggnmm5, 2016.

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, Geosci. Model Dev., 11, 4241–4267, https://doi.org/10.5194/gmd-11-4241-2018, 2018. TS63

Watanabe, S., Hajima, T., Sudo, K., Nagashima, T., Takemura, T., Okajima, H., Nozawa, T., Kawase, H., Abe, M., Yokohata, T., Ise, T., Sato, H., Kato, E., Takata, K., Emori, S., and Kawamiya, M.: MIROC-ESM 2010: model description and basic results of CMIP5-20c3m experiments, Geosci. Model Dev., 4, 845–872, https://doi.org/10.5194/gmd-4-845-2011, 2011.

Waterson, A. M., Edgar, K. M., Schmidt, D. N., and Valdes, P. J.: Quantifying the stability of planktic foraminiferal physical niches between the Holocene and Last Glacial Maximum: Niche Stability of Planktic Foraminifera, Paleoceanography, 32, 74–89, https://doi.org/10/f9vbtb, 2017.

Watterson, I. G.: Non-Dimensional Measures of Climate Model Performance, Int. J. Climatol., 16, 379–391, https://doi.org/10.1002/(SICI)1097-0088(199604)16:4<379::AID-JOC18>3.0.CO;2-U, 1996.

Watterson, I. G., Bathols, J., and Heady, C.: What Influences the Skill of Climate Models over the Continents?, B. Am. Meteorol. Soc., 95, 689–700, https://doi.org/10.1175/BAMS-D-12-00136.1, 2014.

West, G. B., Brown, J. H., and Enquist, B. J.: A General Model for the Origin of Allometric Scaling Laws in Biology, Science, 276, 122–126, https://doi.org/10.1126/science.276.5309.122, 1997.

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

Zakharova, L., Meyer, K. M., and Seifan, M.: Trait-based modelling in ecology: A review of two decades of research, Ecol. Model., 407, 108703, https://doi.org/10.1016/j.ecolmodel.2019.05.008, 2019.

Žarić, S., Schulz, M., and Mulitza, S.: Global prediction of planktic foraminiferal fluxes from hydrographic and productivity data, Biogeosciences, 3, 187–207, https://doi.org/10.5194/bg-3-187-2006, 2006.

Remarks from the language copy-editor

Remarks from the typesetter

TS23   Please clarify whether the data set is your own. If yes, please provide a DOI in addition to your GitHub URL since our reference standard includes DOIs rather than URLs. If you have not yet created a DOI for your data set, please issue a Zenodo DOI (https://help.github.com/en/github/creating-cloning-and-archiving-repositories/referencing-and-citing-content). If the data set is not your own, please inform us accordingly. In any case, please ensure that you include a reference list entry corresponding to the data set including creators, title, and date of last access.

identifier, and publication year. Regarding the publisher/repository, please add "[data set]" or "[code]" to the entry (e.g. Zenodo [code]).

TS31   This reference is not mentioned in the text.

TS32   Please check page range.

TS33   Please check page range/article number.

TS34   This reference is not mentioned in the text.

TS35   Please provide volume number.

TS36   Please check DOI.

TS37   Please confirm DOI.

TS38   Please confirm DOI.

TS39   Please provide page range/article number.

TS40   Please confirm DOI.

TS41   Please check page range/article number and provide volume number.

TS42   Please provide volume number.

TS43   Please provide more information.

TS44   Please provide more information.

TS45   Please confirm DOI.

TS46   Please check page range/article number.

TS47   Please check journal name.

TS48   Please check page range/article number.

TS49   This reference is not mentioned in the text.

TS50   Please provide article number.

TS51   Please provide more information.

TS52   Please check title.

TS53   Please provide article number.

TS54   Please confirm DOI.

TS55   Please check author list and provide full names for the nicknames if possible.

TS56   Please check DOI; should the current one be used, or https://doi.org/10.5281/zenodo.6808760?

TS57   Please confirm reference list entry.

TS58   Please confirm DOI.

TS59   Please check page range.

TS60   Please check page range/article number.

TS61   Please provide more information.

TS62   Please confirm DOI.

TS63   Please confirm DOI.