Global Biogeochemical Cycles

RESEARCH ARTICLE
10.1029/2022GB007685

The Impact of Zooplankton Calcifiers on the Marine Carbon Cycle

Nielja S. Knecht[1], Fabio Benedetti[1], Urs Hofmann Elizondo[1], Nina Bednaršek[2,3], Sonia Chaabane[4,5,6], Catharina de Weerd[7], Katja T. C. A. Peijnenburg[7,8], Ralf Schiebel[6], and Meike Vogt[1]

[1]Environmental Physics, Institute of Biogeochemistry and Pollutant Dynamics, ETH Zurich, Zurich, Switzerland, [2]National Institute of Biology, Marine Biological Station, Piran, Slovenia, [3]Cooperative Institute for Marine Resources Studies, Oregon State University, Corvallis, OR, USA, [4]Aix-Marseille Université, CNRS, IRD, INRAE, CEREGE, Aix-en-Provence, France, [5]French Foundation for Research on Biodiversity (FRB-CESAB), Paris, France, [6]Department of Climate Geochemistry, Max-Planck-Institute for Chemistry, Mainz, Germany, [7]Plankton Diversity and Evolution, Naturalis Biodiversity Center, Leiden, The Netherlands, [8]Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands

Key Points:

- Shelled pteropods and planktic foraminifers contribute 0.2%–3.2% and 0.1%–3.8% to global annual total inorganic carbon surface export fluxes
- Temperature is the strongest environmental covariate associated with their global biomass distributions, followed by chlorophyll-a
- Sampling data characteristics and simplified biomass conversions introduce uncertainty and our estimates are likely lower bounds

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

N. S. Knecht,
nknecht@ethz.ch

Citation:

Knecht, N. S., Benedetti, F., Hofmann Elizondo, U., Bednaršek, N., Chaabane, S., de Weerd, C., et al. (2023). The impact of zooplankton calcifiers on the marine carbon cycle. *Global Biogeochemical Cycles*, *37*, e2022GB007685. https://doi.org/10.1029/2022GB007685

Received 27 DEC 2022
Accepted 11 MAY 2023

Abstract Shelled pteropods and planktic foraminifers are calcifying zooplankton that contribute to the biological carbon pump via the sinking of their calcareous shells. However, their importance for regional and global plankton biomass and carbon fluxes is not well understood. Here, we modeled global annual patterns of pteropod and foraminifer total carbon (TC) biomass and total inorganic carbon (TIC) export fluxes over the top 200 m using five species distribution models (SDMs). An extended version of the MARine Ecosystem DATa (MAREDAT) of zooplankton abundance observations was used to estimate the biomass of both plankton groups. We found hotspots of mean annual pteropod biomass in the high Northern latitudes and the global upwelling systems, and in the high latitudes of both hemispheres and the tropics for foraminifers. This largely agrees with previously observed distributions. For both groups, temperature is the strongest environmental correlate, followed by chlorophyll-a. We found mean annual standing stocks of 52 Tg TC (48 to 57 Tg TC) and 0.9 Tg TC (0.6 to 1.1 Tg TC) for pteropods and foraminifers, respectively. This translates to mean annual TIC fluxes of 14 Tg TIC yr^{-1} (9 to 22 Tg TIC yr^{-1}) for pteropod shells and 11 Tg TIC yr^{-1} (3 to 27 Tg TIC yr^{-1}) for foraminifer tests. These results are similar to previous estimates for foraminifers, but approximately a factor of ten lower for pteropods. Pteropods contribute 0.2%–3.2% and foraminifers 0.1%–3.8% to global surface carbonate fluxes. Including global coccolithophore fluxes, this leaves 40%–60% of the global carbonate fluxes unaccounted for. Our figures are likely lower-bound estimates due to sampling data characteristics and uncertainty associated with organism growth rates.

1. Introduction

Marine calcifying plankton play a key role in the ocean's carbon cycle, particularly through the formation, sinking, and dissolution of their $CaCO_3$ shells (Sarmiento & Gruber, 2006b). These processes impact the carbonate system throughout the water column (Sarmiento & Gruber, 2006a; Takahashi & Bé, 1984). At the ocean surface, the formation of the calcium carbonate shells consumes the buffering carbonate ion CO_3^{2-} and hence causes outgassing of CO_2 (Sarmiento & Gruber, 2006a). In contrast, the sinking shells of calcifying organisms constitute a downward flux of inorganic carbon from the surface ocean amounting to 0.7 to 4.7 Pg C yr^{-1} (Ziveri et al., 2023). At depth, a large fraction of these shells dissolves, thereby increasing ocean alkalinity (Sarmiento & Gruber, 2006a). The export of calcium carbonate shells is responsible for approximately 20% of global surface ocean carbon export fluxes, with soft-tissue export making up 70% of the fluxes and solubility effects responsible for 10% (Sarmiento & Gruber, 2006a). However, there are significant uncertainties regarding the spatial and seasonal calcium carbon flux patterns and the relative contribution of the different plankton groups to global calcification rates. The major groups of calcifying plankton are coccolithophores, shelled pteropods and planktic foraminifers (Bednaršek, Mozina, et al., 2012; Lalli & Gilmer, 1989; Schiebel, 2002; Schiebel & Hemleben, 2017; Stepien, 1980). Shelled pteropods from the suborder Thecosomata (in the following referred to as pteropods) build shells of aragonite, a metastable form of calcium carbonate (Lalli & Gilmer, 1989) with adults ranging from 1 to 30 mm in size (Bednaršek, Mozina, et al., 2012; Bednaršek, Tarling, et al., 2012). Aragonite is 50% more soluble than calcite (Mucci, 1983), which makes pteropods more sensitive to ocean acidification, that

is, the long-term reduction of ocean pH due to the dissolution of excess atmospheric CO_2 in the water (Caldeira & Wickett, 2003), than calcite-shelled organisms (Bednaršek et al., 2016; Doney et al., 2009; Fabry et al., 2008; Manno et al., 2016). Pteropods are flux feeders, that is, they secrete a floating mucus web to trap sinking organic particles (Gilmer & Harbison, 1986). They are active swimmers and some species perform diel vertical migration (DVM), feeding at night at the surface and spending the day at depths between 100 m and in some cases up to 1,000 m (Bé & Gilmer, 1977; Bednaršek, Tarling, et al., 2012) to avoid predation.

Foraminifers build calcareous tests that can reach diameters ranging between 100 μm and 1 mm (Frerichs et al., 1972; Schiebel & Hemleben, 2017). They are generally omnivorous and can capture prey actively, but feeding preferences differ between species (Anderson et al., 1979; Caron & Bé, 1984; Rhumbler, 1911; Spindler et al., 1984) with some species also harboring facultative photosymbionts (Hemleben et al., 1989). The global abundances and habitat suitability of pteropods and foraminifers are known to be controlled by a range of environmental parameters, including temperature (Beaugrand et al., 2010; Bednaršek et al., 2022; Helaouët & Beaugrand, 2009; Jonkers et al., 2019), chlorophyll-a as a proxy for food availability (Pinkerton et al., 2020; Thibodeau et al., 2019; Vereshchaka et al., 2022), and parameters related to physical mixing that influence phytoplankton growth through light availability and particle sinking rates (Bednaršek et al., 2022; Boyce et al., 2010; Govoni et al., 2010; Longhurst, 2007; Mackas et al., 2005; Rothschild & Osborn, 1988; Seuront et al., 2001).

The relative importance of the different calcifying plankton groups for global carbonate fluxes remains uncertain. Coccolithophores were long thought to dominate the inorganic carbon export (Anglada-Ortiz et al., 2021; Iglesias-Rodriguez et al., 2002; Rembauville et al., 2016; Rost & Riebesell, 2004; Schiebel, 2002). However, in global observation-based estimates, they only accounted for 26%–52% of global carbonate fluxes, which leaves a significant fraction of the carbonate fluxes unattributed (Buitenhuis, Vogt, et al., 2013; C. J. O'Brien, 2015). This discrepancy shifted the attention towards the contribution of the two calcifying zooplankton groups, pteropods and foraminifers. Recent observational studies estimated pteropods to contribute more than previously thought to global surface carbonate fluxes, making up 20%–42% of the global annual mean carbonate fluxes from the surface to depths below 200 m (Bednaršek, Mozina, et al., 2012). Foraminifer carbon flux estimates vary by a factor of 100 (Buitenhuis et al., 2019; Buitenhuis, Vogt, et al., 2013; Schiebel, 2002; Schiebel & Movellan, 2012). However, recent studies based on newly available observations find significantly lower fluxes. Finally, a recent mechanistic modelling study found pteropods to dominate upper subsurface $CaCO_3$ export, with contributions ranging between 33% and 89% (Buitenhuis et al., 2019). These results further suggest the key role of pteropods and foraminifers for the oceanic inorganic carbon cycle.

To derive the magnitude of carbon export mediated by zooplankton calcifiers, we first need to quantify the global biomass standing stocks and characterize the global distribution patterns of these groups. Earlier descriptions of the global patterns based on global plankton sampling data were made by the MARine Ecosystem DATa (MARE-DAT) project (Buitenhuis, Vogt, et al., 2013). Additionally, large-scale observational data sets have been collected by the Continuous Plankton Recorder (CPR) survey (Richardson et al., 2006). However, the existing observations are usually confined to specific ocean regions and have an overall low data coverage in the central oceanic basins (Bednaršek, Mozina, et al., 2012; de Garidel-Thoron et al., 2022; Schiebel & Movellan, 2012). Furthermore, plankton distributions are generally patchy in space and time (Beckmann et al., 1987; Boltovskoy, 1971; Buitenhuis, Vogt, et al., 2013; Siccha et al., 2012), which causes high variability in the observed abundances. Different sampling techniques and varying sampling depths and mesh sizes introduce additional variation (Wells, 1973). The derivation of continuous global biomass maps and standing stock estimates for zooplankton calcifiers hence requires us to account for these data gaps and biases by employing statistical methods.

Species distribution models (SDMs) empirically derive the relation between the target variable and a range of environmental predictors through response curves and can then extrapolate said target variable to un-sampled regions by projecting these response curves on predictor values (Elith & Leathwick, 2009; Guisan & Zimmermann, 2000; Merow et al., 2014). They have been successfully used in marine macroecology to model plankton species distributions based on occurrence data (presence/absence) (Barton et al., 2016; Bednaršek et al., 2022; Benedetti et al., 2021; Brun et al., 2015, 2016; Righetti et al., 2019) and are increasingly being used to model continuous abundance values (De Broyer et al., 2014; Pinkerton et al., 2010; Waldock et al., 2022). In the present work, we apply an SDM framework to estimate global biomasses for calcifying zooplankton.

To this end, we use newly compiled global data sets of pteropod and foraminifer abundances and species-specific biomass conversion methods to calculate biomass concentrations over the top 200 m. We combine the global

gridded biomass data with an ensemble of SDMs to address the following questions: (a) What are the biogeographic patterns and main environmental covariates of global total carbon (TC) biomass for pteropods and foraminifers (on a monthly, $1 \times 1°$ gridded scale of the upper open ocean)? (b) What is the magnitude and range of uncertainty of the associated annual total inorganic carbon (TIC) fluxes from pteropods and foraminifers? Answering these questions will improve our understanding of the role calcifying zooplankton play in the global marine carbon cycle and provide a new versatile modeling pipeline to map other quantitative, continuous variable types.

2. Methods

We model the biomass patterns and associated carbon fluxes of pteropods and foraminifers at a global scale using SDMs and updated abundance data sets for the two groups. To this end, we use a multi-step modeling pipeline as shown in Figure 1.

2.1. Plankton Data

2.1.1. Data Collection and Pre-Processing

We updated the original MAREDAT pteropod and foraminifer abundance and biomass data sets of Schiebel and Movellan (2012) and Bednaršek, Mozina, et al. (2012) by aggregating abundance concentration data from large scale sampling campaigns, existing data compilation efforts, and unpublished sampling data (Figure 1). The main data sources (Figure S1 in Supporting Information S1) for both plankton groups included the Southern Ocean Continuous Plankton Recorder (SO-CPR; Hosie, 2021), the Australian CPR (Aus-CPR; IMOS, 2022), the North Atlantic and North Pacific CPR (NA-NP CPR; Johns, 2021), and the Coastal and Oceanic Plankton Ecology, Production and Observation Database (COPEPOD; T. D. O'Brien, 2010). For pteropods, we added data from the Tara Oceans expeditions (Brandão et al., 2021), the Atlantic Meridional Transect (AMT24; Burridge et al., 2017) and AMT27 (Peijnenburg, 2021, Personal communication), as well as unpublished sampling data from the North Atlantic (Schiebel, 2021, Personal communication). For foraminifers, we also gathered data from various individual sampling campaigns (Jentzen et al., 2018; Schiebel, 2002; Schiebel et al., 1995, 2001, 2002, 2004; Schiebel & Hemleben, 2000).

We took several pre-processing steps to ensure the quality of the biological observations. To harmonize all classifications across data sets and correct for potential deprecated scientific species names, we matched all taxonomic information against the list of accepted taxon names of the World Register of Marine Species (WoRMS; Horton et al., 2017). Observations lacking complete sampling metadata (date, depth, longitude, latitude, and abundance value) and observations of body parts were removed (21,303 points for pteropods, mainly due to observations of body parts and larvae, and 522 for foraminifers). Additionally, pteropod abundance values from the Ecosystem Monitoring—Ships Of OPportunity surveys (EcoMon-SOOP) in the Gulf of Maine from the COPEPOD data set were corrected by dividing them by a factor of 100 as the units in the original data set had been erroneously reported (Hofmann Elizondo, 2022, Personal communication). We did not standardize the abundance estimates between the various mesh sizes used in the different sampling cruises as there were not yet any published correction factors that we were aware of for these two specific plankton groups.

The final, quality-controlled pteropod abundance data set (Figure S2 in Supporting Information S1) contains 841,239 data points at 309,921 individual locations, collected at a mean sampling depth (±sd) of 38.15 ± 190.89 m over the 1938–2021 period (2001.25 ± 15.23). Abundances range between 0 ind/m^3 and 1,066.67 ind/m^3, with a mean of 4.38 ± 79.86 ind/m^3. The median abundance (0.00 ind/m^3) is low due to the CPR data sets which make up 91.15% of the data, and contain 92.06% absence observations. 50.19% of the data is resolved only to the order-level, whereas 24.03% of the observations are species-resolved and 22.41% resolved to the genus level (see Table S2 in Supporting Information S1). The data set contains observations on 33 species out of 165 currently recognized pteropod species (Peijnenburg et al. (2020); see Table S2 in Supporting Information S1). The largest contributions to total abundance summed over all observations stem from *Limacina helicina* sensu lato (47.7% of the total species-resolved abundance), *Heliconoides inflatus* (26.7%), and *L. retroversa* s.l. (10.0%).

The final, quality-controlled foraminifer abundance data set (Figure S2 in Supporting Information S1) consists of 1,021,283 points at 308,641 unique locations, with a mean sampling depth of 108.06 ± 340.49 m and collected

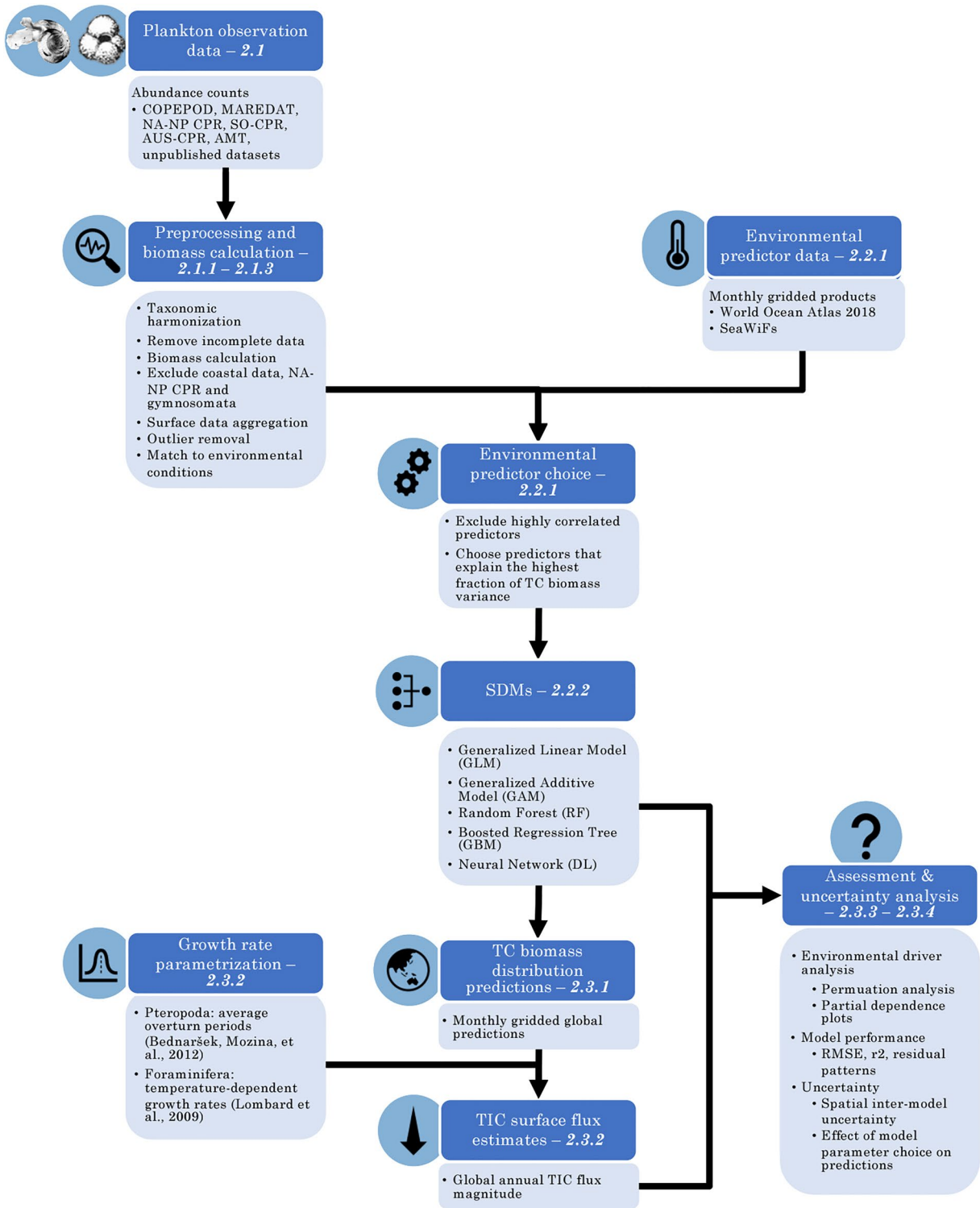

Figure 1. Flow diagram illustrating the pipeline of numerical analyses implemented for the present study. The various steps taken from the raw data to the final total carbon (TC) biomass distributions and total inorganic carbon (TIC) flux estimates using species distribution models (SDMs) are shown. The numbers in italics indicate the subsection of the Methods where the corresponding step is described.

during the 1939–2021 period (mean 2000.36 ± 13.30). Foraminifer abundances range between 0 ind/m^3 and 152,170.00 ind/m^3, with a mean abundance of 3.63 ± 163.08 ind/m^3. There is a high prevalence of CPR data (74.35% of the total data) with 89.72% zero abundance observations, which causes a low median abundance value of 0.00 ind/m^3. 59.79% of the data are species resolved, followed by 33.07% of the observations on a phylum

level (see Table S4 in Supporting Information S1). This data set contains observations on 42 of the 47 extant foraminifer species (Schiebel & Hemleben, 2017). Most of the total abundance is composed of *Globigerina bulloides* (25.6% of the total species-resolved abundance), *Neogloboquadrina incompta* (23.7%), *Turborotalita quinqueloba* (13.3%), and *Globigerinita glutinata* (11.3%).

For model training, we performed additional data quality controls to ensure sensible relations between environmental predictors and biomass values could be derived. The NA-NP CPR data set was flagged and discarded for modeling as it contained discrete medians of abundance bins instead of continuous values (removal of 340,250 points for pteropods and 250,620 points for foraminifers). Additionally, we excluded data from neritic sampling locations associated with a climatological salinity <30 PSU from the analysis to avoid observations influenced by terrestrial freshwater and nutrient inputs (Brun et al., 2015; removal of 18,725 data points for pteropods and 17,207 points for foraminifers). Lastly, observations for pteropods from the clades Gymnosomata and Pseudothecosomata were removed for modeling, as only some of the latter are calcifiers (Lalli & Gilmer, 1989), and there is very little literature on their role in the carbon cycle (removal of 106,929 points). The final data sets used for modeling contain 375,336 points for pteropods and 770,663 points for foraminifers as shown in Figure 2.

2.1.2. Biomass Calculations

To estimate calcifying zooplankton biomass and subsequent carbon fluxes, we converted the abundances to biomass data based on morphology-based conversion factors (cf. Figure 1). To this end, we grouped species of similar morphology into shape groups and derived biomass as a function of average body size (maximum elongation) based on shape-specific conversion equations. Generally, we applied all conversions on the lowest taxonomic level available and only used shape-group or phylum-wise averages where the species identification was not available.

2.1.2.1. Biomass Calculation for Pteropods

To convert pteropod abundance into carbon biomass, we used corrected species-specific biomass conversion equations from Bednaršek, Mozina, et al. (2012) to calculate wet weight (WW) as shown in Table S1 in Supporting Information S1. These equations are based on six different morphological shape groups and relate an individual species' body length in millimeters to its biomass. For observations without morphometric data (99.8%), we used the species-average lengths from Bednaršek, Mozina, et al. (2012). We used pteropod shell length whenever given in Bednaršek, Mozina, et al. (2012), otherwise we used the body length values from the same source. Table S2 in Supporting Information S1 shows the average length value used for each species, their respective shape group, and the number of observations for each species. WW was then transformed to dry weight (DW) as per C. S. Davis and Wiebe (1985):

$$DW = WW \cdot 0.28, \tag{1}$$

and subsequently transformed to total carbon (TC) following Larson (1986):

$$TC = DW \cdot 0.25, \tag{2}$$

Finally, total inorganic carbon (TIC) was computed following Bednaršek, Mozina, et al. (2012):

$$TIC = 0.27 \cdot TC. \tag{3}$$

This TC-TIC conversion factor is based on data for *L. helicina antarctica* and hence probably not representative for all pteropod species and life stages (Hofmann Elizondo & Vogt, 2022). To account for the lack of species-specific TC-TIC conversion factors in literature, we added an uncertainty range of ±20% to the conversion factor, based on the range of TIC values reported in Bednaršek, Tarling, et al. (2012). The effect of this parameter choice is assessed according to the methodology in Section 2.3.4.

2.1.2.2. Biomass Conversion for Foraminifers

A morphological approach was carried out for converting foraminifer abundances to TC concentrations. We were not aware of any published shape class definitions for foraminifers. Thus, we defined eight morphological shape groups based on similar adult test shape and structure as shown in Table S3 in Supporting Information S1.

To derive biovolume-to-biomass conversion equations, we constructed species and group-specific maximum test length to biomass functions from the literature. We collected species-specific test weight measurements per

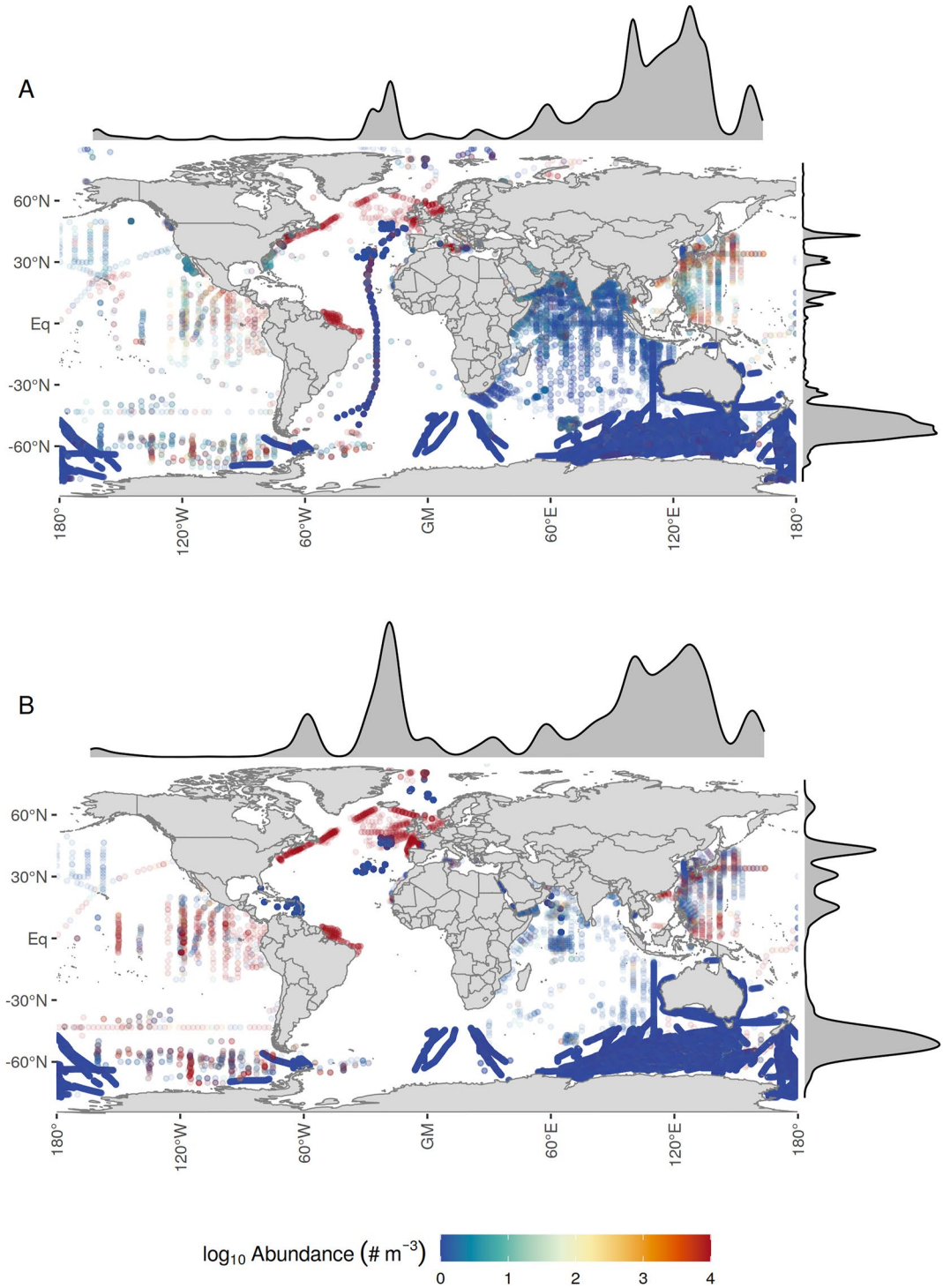

Figure 2. Global distribution of the final quality-controlled observations of pteropod (a) and foraminifer (b) abundance used for modeling. The marginal plots show the density of observations and highlight the dominant role of the Southern Ocean Continuous Plankton Recorder (SO-CPR) survey as well as a spatially confined, highly resolved data set in the North Atlantic. This plot shows the data set used for modeling, that is, the data set after removing the North Atlantic and North Pacific CPR data, coastal observations with surface salinity ≤30, and observations of naked pteropods (Gymnosomata) as described above. For the full collected data set, see Figure S2 in Supporting Information S1.

plankton size class from Schiebel and Hemleben (2000) and Takahashi and Bé (1984). We fitted linear functions to calculate biomass as a function of length per species and per shape group, where the biomass of a shape group is calculated as the average of all species within the group (Figure S3 in Supporting Information S1). To compare the ranges of the conversion factors to published equations for the entire foraminifera phylum, we used the equation provided by Michaels et al. (1995) (Figure S3 in Supporting Information S1). This function computes foraminifer cytoplasm carbon (i.e., total organic carbon, TOC) as a function of test length. The TC biomass is calculated based on the following conversion factors (Schiebel & Movellan, 2012):

$$TIC = 0.36 \cdot TOC, \tag{4}$$

$$TIC = 0.265 \cdot TC. \tag{5}$$

As for pteropods, we added an uncertainty range of ±20% to the TIC-TC conversion factor to account for variability in calcification intensity associated with different species and environmental conditions.

To compute TC from the test weight measurements of Schiebel and Hemleben (2000) and Takahashi and Bé (1984), we used the following molar relationship:

$$m(C) = \frac{m(CaCO_3)}{M(CaCO_3)} \cdot M(C) = \frac{m(CaCO_3)}{100.09\,\mathrm{g\,mol^{-1}}} \cdot 12.01\,\mathrm{g\,mol^{-1}}, \tag{6}$$

where m denotes the mass and M the molar weight.

The biomass conversion factors (BCF) shown in Table S5 in Supporting Information S1 are the coefficients of the linear relation between foraminifer TC biomass and their biovolume. To apply the conversion factors, the BCF values were substituted into the following equation:

$$TC = a * L^3 * BCF, \tag{7}$$

where TC represents the TC biomass of foraminifers in µg, a denotes foraminifer abundance and L the species' length in µm.

We collected average length values for all species from the images of Schiebel and Hemleben (2017). These average length values as well as the number of observations per species can be found in Table S4 in Supporting Information S1.

2.1.3. Surface Ocean Aggregation

To reduce spatio-temporal patchiness and noise in the data, we conducted a surface ocean aggregation (C. J. O'Brien, 2015). To this end, we re-gridded all data onto the 1 × 1° grid of the World Ocean Atlas 2018 (WOA18; Boyer et al., 2018). For each grid cell, we summed all biomass concentrations from the same sampling event, as different species were sometimes counted as separate measurements. Next, we averaged all biomass and abundance values per grid cell and month of the year over the top 200 m. This depth cutoff was deemed reasonable as 99.1% and 99.4% of the summed abundance of pteropods and foraminifers, respectively, stem from the top 200 m (Figure S4 in Supporting Information S1).

To better approximate a normal distribution, TC mass values were log-transformed with a $\log_{10}(TC + 1)$ transformation for further analyses. Lastly, to dampen the effect of plankton patchiness and bloom dynamics, we flagged outliers in the surface aggregated values based on the z-score criterion (Burba & Anderson, 2005). Hence, for modeling, we excluded high biomass observations with a score of $z > 3$, that is, more than three standard deviations away from the sample mean.

2.2. Modeling

2.2.1. Environmental Predictor Selection

To identify the set of predictors used for training the biomass-based SDMs, we collected gridded monthly climatologies of meaningful environmental predictors as shown in Table 1 and Figure 1. Whenever necessary, the fields were averaged and re-gridded to monthly climatologies at a 1 × 1° resolution. Depth-resolved predictors from the WOA18 were averaged over the climatological mixed layer depth (MLD). As many pteropods actively

Table 1
Environmental Predictor Variables Used in the Univariate Predictor Evaluation

Predictor	Source	Reference
Temperature	WOA18	Locarnini et al. (2018)
Chlorophyll-a	SeaWiFs	NASA OB.DAAC (2018a)
Mixed layer depth (MLD)	SODA3.4.2	Carton et al. (2018)
Eddy kinetic energy (EKE)	Copernicus	Copernicus (2021)
Salinity	WOA18	Zweng et al. (2019)
Dissolved oxygen	WOA18	Garcia et al. (2019b)
Nitrate	WOA18	Garcia et al. (2019a)
Phosphate	WOA18	Garcia et al. (2019a)
Depth of the euphotic layer (z_{eu})	SeaWiFS	NASA OB.DAAC (2018c)
Photosynthetically active radiation (PAR)	SeaWiFS	NASA OB.DAAC (2018d)
Particulate backscattering coefficient at 443 nm (BBP_{443})	SeaWiFS	NASA OB.DAAC (2018d)
Diffuse attenuation coefficient for downwelling irradiance at 490 nm (Kd_{490})	SeaWiFS	NASA OB.DAAC (2018b)
Total alkalinity (TA)	OceanSODA-ETHZ	Gregor and Gruber (2021)
Dissolved inorganic carbon (DIC)	OceanSODA-ETHZ	Gregor and Gruber (2021)
Partial pressure of CO2 (pCO_2)	OceanSODA-ETHZ	Gregor and Gruber (2021)
Calcite saturation state (Ω_{Ca})	OceanSODA-ETHZ	Gregor and Gruber (2021)
Aragonite saturation state (Ω_{Ar})	OceanSODA-ETHZ	Gregor and Gruber (2021)

Note. WOA18 refers to the 2018 edition of the World Ocean Atlas (Boyer et al., 2018), SeaWiFS denotes the Sea-viewing Wide Field-of-view Sensor satellite data (NASA OB.DAAC, 2018a) and SODA describes the Simple Ocean Data Assimilation project (Carton et al., 2018).

migrate vertically (on a daily or seasonal basis) and both groups are passively vertically mixed within the water column (Lalli & Gilmer, 1989; Mackas et al., 2005; Myers, 1968; Schiebel & Hemleben, 2017; Wormuth, 1981), the depth-averaged environmental predictors are more representative of the conditions they experience rather than the surface values. However, as sampling devices are often towed vertically or obliquely, the reported water depth interval of each observation is not directly representative of the depth where an organism dwells over the entire day or even through its life span. Hence, we assume that pteropods and foraminifers move within the mixed layer, where the majority of the organic matter is present (Sallée et al., 2021; Soviadan et al., 2022). For all depth-resolved environmental predictors considered, the average over the top 200 m, the values at the surface and the MLD-averaged predictor values are each correlated with a Pearson correlation coefficient of $r > 0.99$, so this simplification is deemed reasonable. For dissolved oxygen concentration, we used the value at 200 m depth to avoid the strong collinearity with the sea surface temperature (SST) values. The distribution of chlorophyll-a concentrations, nutrient variables, MLD, and eddy kinetic energy (EKE) were right-skewed (Figure S5 in Supporting Information S1), therefore we log-transformed those variables so their distribution was closer to a normal one to improve model performance and resilience. Then, we collocated the environmental parameters with the gridded monthly pteropod and foraminifer biomass fields.

To select the most meaningful environmental predictors for the final biomass-based SDMs we used a multi-step approach for each zooplankton group. First, we identified clusters of collinear predictors (Pearson correlation coefficient $|r| > 0.7$ calculated from the values matched up with the monthly biomass climatologies, Brun et al., 2020). Second, we excluded all but one predictor in each cluster, which improves model performance (Brun et al., 2020; Dormann et al., 2013; Figures S6 and S7 in Supporting Information S1). Thus, for each cluster we first chose the most normally distributed predictor as assessed by the Shapiro-Wilk test (Shapiro & Wilk, 1965), and second, we chose predictors whose effect are easier to interpret from an ecological point of view (e.g., chlorophyll-a over Kd_{490}, the remotely sensed light attenuation at a wave length of 490 nm, which is an indirect measure of surface productivity and turbidity). This selection procedure resulted in the following seven candidate predictors for both foraminifers and shelled pteropods: surface chlorophyll-a, MLD, temperature averaged over the MLD, surface EKE, oxygen at 200 m depth, salinity averaged over the MLD, partial pressure of CO_2 (pCO_2), photosynthetically active radiation (PAR), and particulate backscattering coefficient at 443 nm (BBP_{443}).

The exclusion of a predictor variable does not mean that it is not ecologically relevant for the organisms modeled. The calcite and aragonite saturation states are known to influence habitat suitability for foraminifers and pteropods, respectively (Bednaršek et al., 2016, 2022; Lischka et al., 2011; Lischka & Riebesell, 2012; Manno et al., 2016). However, the matched saturation states were determined to be highly correlated with water temperature averaged over the MLD (Pearson $r > 0.99$). As previous studies have shown temperature to be more biologically relevant in influencing large-scale biogeographic distribution patterns (Beaugrand et al., 2013; Bednaršek et al., 2018; Howes et al., 2015; MacKas & Galbraith, 2012), we excluded the saturation states as predictors. Exchanging temperature for the aragonite saturation state in the pteropod models does not have a significant effect on the biomass distribution pattern or the annual TC fluxes (Figure S8 in Supporting Information S1).

To choose the final predictor set, we assessed the variance of the TC biomass explained by each of the seven candidate predictors using univariate regression models (Figure S9 in Supporting Information S1). For this, we calculated both 1° pixel-wise and latitudinal 10°, 5°, and 1° monthly means of the TC biomass and the environmental predictors to identify the large-scale effects of the environmental predictors. To model variations in the TC biomass as a function of each environmental predictor, we trained two Generalized Linear Models (GLMs) with a Gaussian response function (one with only a linear term and the second with both a linear and a quadratic term) and a Generalized Additive Model (GAM) with a cubic regression spline. Then, we assessed the percentage of deviance explained by each predictor (Hosmer Jr et al., 2013; Nelder & Wedderburn, 1972). We retained all predictors that explained ≥5% of variability at any of the spatial aggregation levels. For pteropods, the resulting set of predictors included: MLD-averaged temperature, surface chlorophyll-a, and MLD. For foraminifers, we retained the MLD-averaged temperature, surface chlorophyll-a, and EKE (Figures S9 and S10 in Supporting Information S1 for mean annual maps of the predictors).

To assess the impact of this predictor selection procedure on SDM outputs, we also trained the models for both plankton groups on a Principle Component Analysis (PCA) transformation of the full initial predictor set (Table 1). There was no significant difference between the PCA-based global annual TIC fluxes and those calculated based on our final choice of predictors ($p > 0.05$ for both plankton types as assessed with a Kruskal-Wallis test (Kruskal & Wallis, 1952), see Figure S11 in Supporting Information S1). This shows that the selected predictors are sufficient to represent the full set of predictors.

2.2.2. Multivariate Modeling

We used the identified predictors to train an ensemble of five SDMs of increasing complexity: a GLM, a GAM, a Random Forest (RF), a Gradient Boosting Machine (GBM) and a Neural Network/Deep Learning Model (DL; see Figure 1). GLMs, GAMs and RFs have been widely and successfully used in the modeling of global marine plankton distributions (Benedetti et al., 2021; Brun et al., 2016; Righetti et al., 2019). The more complex models have also been used for modeling plankton distributions, though less frequently (GBMs in Pinkerton et al. (2010, 2020), DL models in Benedetti et al. (2021); C. J. O'Brien et al. (2016)). For an extensive description of the more complex model types, we refer to Boehmke and Greenwell (2019a, 2019b, 2019e) and sources within. All modeling was conducted with the h2o 3.36.0.3 R package (H2O.ai, 2021).

For the GLM, we included both first and second-order dependencies on the predictors and assumed a normal distribution of the target variable with an identity link function (Nelder & Wedderburn, 1972). In the GAM, we fitted smoothing terms for all predictor variables using cubic regression splines, the most common smoothing algorithm (Hastie & Tibshirani, 1990), and a normal distribution with the identity function as link for the target variable. For the RF, GBM, and DL, the hyperparameters were tuned using a grid search (Boehmke & Greenwell, 2019d). Tables S6, S7, and S8 in Supporting Information S1 show the grid of parameters evaluated for each model. The final setup of the RF as determined from the tuning process (Table S6 in Supporting Information S1) included 830 trees for pteropods and 330 for foraminifers. At each tree node, one and two environmental predictors were evaluated (m_{try}) for pteropods and foraminifers, respectively, and the minimum number of rows at each final node (min_{rows}) was set to three and two. The maximum tree size was constrained to 30 for pteropods and 10 for foraminifers. For each bootstrap replicate of the tree, we chose a fraction (r_{sample}) of 0.8 and 0.632 of the total data set. For the GBM, we determined a maximum depth ($max_{depth} = 5$) and minimum number of observations per terminal node ($min_{rows} = 1$) for each individual tree for both plankton groups (see also Table S7 in Supporting Information S1). The learning rate (r_{learn}) was determined to be 0.01 and each individual tree is trained on a fraction of 0.75 and 0.5 of the total data set for pteropods and foraminifers, respectively, using all of the predictor columns ($r_{samplecolumns}$). The DL (see also Table S8 in Supporting Information S1) was determined to

have a Tanh activation function for both plankton groups. The pteropod model has two hidden layers with 20 neurons each and the foraminifer model has two hidden layers of 15 neurons each. To avoid overfitting, L_1 and L_2 regularizations were included (Boehmke & Greenwell, 2019a) with weight factors $\lambda_{L_1} = 0$ and $\lambda_{L_2} = 1 * 10^{-3}$ for pteropods, and $\lambda_{L_1} = 1 * 10^{-3}$ and $\lambda_{L_1} = 1 * 10^{-5}$ for foraminifers.

We assessed the effect of the hyperparameter tuning on the global annual TIC fluxes by comparing the fluxes calculated using the tuned models (for the RF, GBM, and DL) to those based on the untuned models with standard hyperpameter set-up (see Tables S6–S8 in Supporting Information S1). As expected, the tuned models showed a better model performance, but the global annual TIC fluxes did not differ significantly ($p > 0.05$ as assessed with a Kruskal-Wallis test (Kruskal & Wallis, 1952) for each plankton group), which points to a more accurate spatial representation of the patterns in the tuned models. Tuning the models hence does not introduce unfounded model complexity or biases.

To train the SDMs and assess their performance, we split the data set into a training and a testing set (Boehmke & Greenwell, 2019d). For a conservative estimate of model performance, we randomly assigned 75% of the values to the training data set. On the training data set, we performed a 5-fold cross validation, where we (a) split the training data set into five equally sized, randomly chosen, non-overlapping subsets, (b) train the SDMs on four of the subsets, and (c) evaluate the model performance of the trained SDM on the remaining subset based on the average root mean squared error (RMSE). This procedure was repeated until each of the five subsets of the data were used four times for training and once for validation. Finally, we evaluated the trained SDM on the testing set.

2.2.3. Model Performance

We assessed model performance using three metrics (Figure 1). The root mean squared error (RMSE) is an error metric estimating the deviation between predicted and true values. Pearson's coefficient of correlation, R^2 indicates the magnitude of correspondence between trends in the predicted and observed values. Finally, the Nash-Sutcliffe-efficiency (NSE; Nash & Sutcliffe, 1970) compares the model performance to a null model, that is, the mean of all observations. Positive NSE values indicate that the assessed model performs better than the null model. Each performance metric was calculated on both the training and the testing set of the data (cf. Section 2.2.2).

2.3. Model Inference

2.3.1. Global Total Carbon (TC) Biomass Patterns

We used the SDMs to project global monthly TC biomass values as a function of the monthly climatological environmental predictors (see Figure 1). Projections were made for each grid cell and month where environmental data were available. We flagged and excluded all predictions of negative biomass values, because they correspond to unrealistic predictions (0.33% of all predicted values for pteropods and 0.06% for foraminifers). Many complex SDMs suffer from low transferability into novel environmental conditions due to non-linear response curves (Bell & Schlaepfer, 2016; Elith et al., 2010; Qiao et al., 2019). Thus, for each grid cell we evaluated whether the environmental conditions lie within the range of the training data set or are considered non-analog using a Multivariate Environmental Similarity Surfaces (MESS) analysis (Elith et al., 2010). The MESS analysis assesses the similarity between the environmental conditions at any given point and the training data set of each SDM. To avoid including unrealistically high values in the flux calculations and global summaries of calcifying zooplankton biomass, we excluded the biomass values from regions where non-analog environmental conditions were detected by the MESS analysis (3.25% of the values for pteropods and 4.03% for foraminifers). To analyze the spatial biomass patterns, we defined hotspots as unusually high biomass concentrations that lie above the 90th percentile for each plankton group.

2.3.2. Annual Total Inorganic Carbon (TIC) Export Fluxes

We computed TIC fluxes from the projected global TC biomass values and environmental conditions (see Figure 1). To compare our results to those of Buitenhuis et al. (2019), biomass values were calculated for TC, while export fluxes were based only on the inorganic shells, that is, on TIC. Hence, we assumed that the carbon export flux is dominated by the sinking and empty shells.

2.3.2.1. TIC Export Flux Calculation for Pteropods

To compute the annual pteropod TIC flux, we applied a simplified approach based on an average overturn time of 1 year, following the methodology of Bednaršek, Mozina, et al. (2012). Based on grid cell-wise mean annual

biomass concentrations, we computed the global annual mean biomass as the spatially weighted mean of the average concentrations of each grid cell, multiplied by the TIC-TC factor (Bednaršek, Mozina, et al., 2012), the depth of 200 m, and the global open ocean area excluding shelf seas ($362*10^6$ km^2; Bednaršek, Mozina, et al., 2012; Dietrich et al., 1975). To represent the variability and uncertainty in turnover times between various pteropod species and regions, we added an uncertainty factor of ±20% to the flux conversion equation based on the range of values given in the review study by Wang et al. (2017). The effect of this parameter choice is evaluated according to the uncertainty analysis described in Section 2.3.4.

2.3.2.2. TIC Export Flux Calculation for Foraminifers

To calculate foraminifer TIC fluxes, we used the phylum-resolved temperature-dependent growth rates from Lombard et al. (2009). To calculate annual TIC fluxes, we multiplied the daily growth rate at each grid cell and month by the current biomass concentration, the TIC-TC factor (Equation 4 in Section 2.1.2), and the depth of 200 m, and weighted the result by grid-cell area. To represent uncertainty in the growth rate, we calculated the minimum and maximum growth rates by computing all combinations within the parameter uncertainty range. Then, we chose those parameter combinations that would minimize or maximize the integral of the growth rate as a function of temperature from 0° to 30°C, while maintaining ecologically sensible response shapes (see Figure S12 in Supporting Information S1 for an illustration of the growth rate options). The effect of this choice was evaluated according to the methodology described in Section 2.3.4.

2.3.3. Environmental Predictor Analysis

To examine how underlying ecological processes were captured by the SDMs (Figure 1), we assessed the models' dependence on the predictor variables in two ways. First, we assessed the overall effect of each environmental predictor based on a permutation analysis using the Fisher-Yates algorithm (Fisher & Yates, 1953). Second, we characterized the biological relevance of the response curve learned by each SDM using partial dependence plot (PDP) curves. The PDP curves were calculated by computing biomass predictions at 25 evenly spaced points across each predictor's range while keeping all other predictors constant at their mean value (Boehmke & Greenwell, 2019c).

2.3.4. Uncertainty Quantification

We assessed the three main sources of uncertainty underlying our SDMs predictions: SDM choice (Thuiller et al., 2019), TIC-TC factor, and growth rate parametrization (Figure 1). First, we identified potential non-normal relationships based on the patterns of the model residuals. Second, we quantified the effect of different model and parameter choices (see Sections 2.1.2, 2.2.2 and 2.3.2 for details on the uncertainty setup) on the carbon flux predictions using a multivariate Analysis of Variance (mANOVA; Weinfurt, 1995) whose target variable was the monthly TIC flux values at each grid cell. We used the model type, the growth rate definition, the TIC-TC conversion factor, and the interactions between these three factors as input for the mANOVA.

3. Results

3.1. Global Biogeographic Total Carbon (TC) Biomass Patterns

The global mean annual TC biomass (±sd) is 0.701 ± 0.648 mg TC m^{-3} for pteropods, and 13.5 ± 28.7 µg TC m^{-3} for foraminifers, implying that pteropod biomass is a factor of ∼50 larger than foraminifer biomass. The projected global mean biomass patterns are shown in Figures 3a and 3b for pteropods and foraminifers (see also Figures S13 and S14 in Supporting Information S1 for the projected patterns per SDM). For both plankton groups, high biomass concentrations are found in the tropics and at latitudes ≥50°N. Lower biomass concentrations (mean values of 0.31 mg TC m^{-3} and 5 µg TC m^{-3}) are found between 40° and 50°S for pteropods and between 30° and 40° in both hemispheres for foraminifers. Contrary to pteropods, we find high biomass concentrations of up to 880 µg TC m^{-3} for foraminifers in the Southern Ocean south of 50°S.

On a regional scale, the North Atlantic Ocean is associated with biomass hotspots (values above the 90th percentile) for both plankton groups, but particularly for foraminifers. A trail of high foraminifer biomasses with a mean value of 150 µg TC m^{-3} is found across the North Atlantic. Other regions of high biomass are associated with tropical and coastal upwelling systems. Pteropod biomass concentrations are particularly high in the coastal Eastern Boundary Upwelling Systems (EBUS) with an average concentration of 3 mg TC m^{-3}. For foraminifers, regions of high biomass are associated with the equatorial upwelling region.

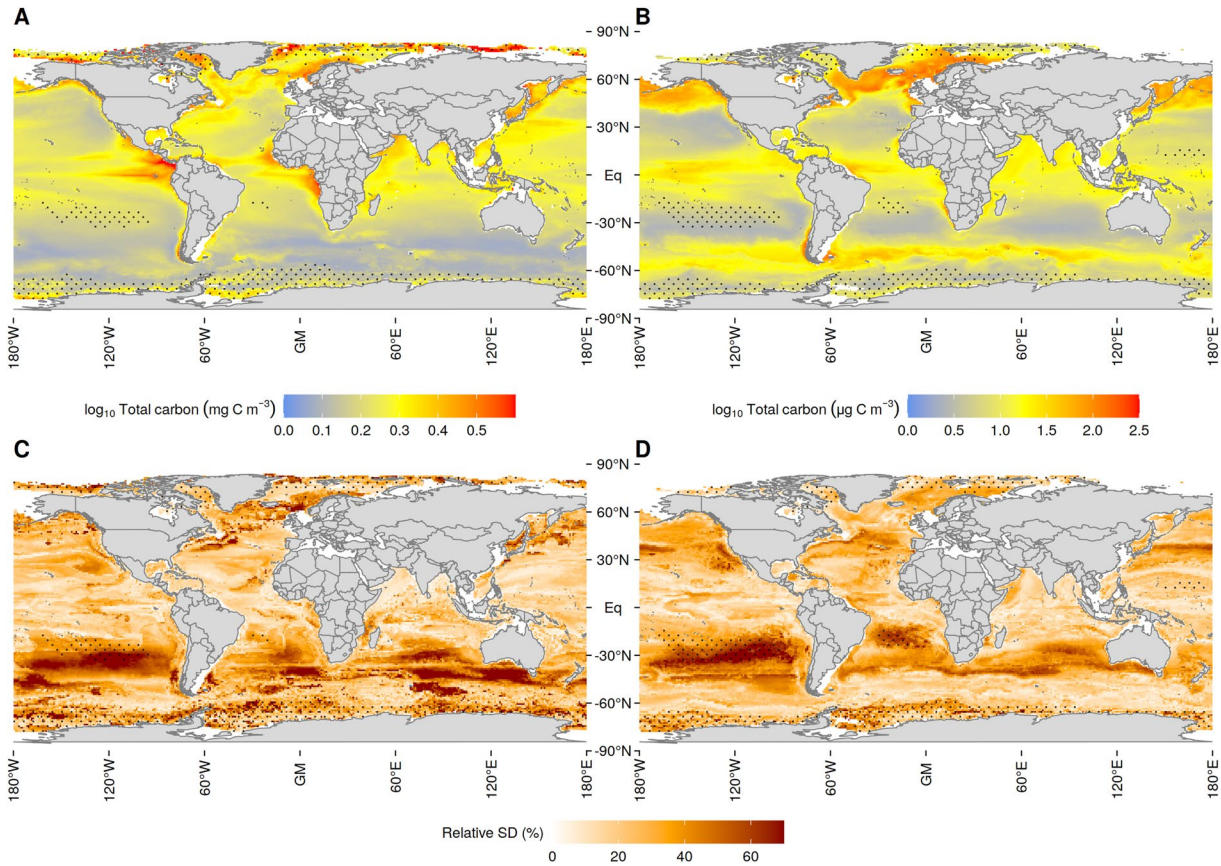

Figure 3. Global mean annual total carbon (TC) biomass concentration for pteropods (left, a) and foraminifers (right, b), averaged over all months and models. Values are shown as $\log_{10}(TC + 1)$, note also the different color scales for pteropods and foraminifers. Stippled regions in plots a–d indicate grid points where the environmental conditions were outside the training data set for more than 6 months of the year as calculated with the Multivariate Environmental Similarity Surfaces (MESS) analysis. The lower panel plots c and d show the mean annual relative standard deviation of the model predictions for pteropods (left) and foraminifers (right), normalized with the mean prediction value at each grid point to facilitate comparability.

On a seasonal scale, biomass hotspots shift toward high latitudes during global summer (Figures S15 and S16 in Supporting Information S1). The seasonal variation in biomass V for each grid point and model is defined as the difference between the maximum and minimum monthly surface ocean biomass concentration at that point. On average, V is stronger in the Northern Hemisphere (NH) than in the Southern Hemisphere (SH) with a difference in variability ($V_{NH} - V_{SH}$) of +0.73 mg TC m^{-3} for pteropods and of +45.59 µg TC m^{-3} for foraminifers ($p < 2*10^{-16}$ for both groups using a t-test; Student, 1908). Foraminifers display a higher seasonal variation than pteropods (+0.28, $p < 2*10^{-16}$ when comparing the maximum seasonal variation at each grid point normalized by the mean global biomass between the plankton groups with a t-test).

3.2. Model Performance

To assess model performance of the five SDMs, we evaluated each model using the root mean squared error (RMSE), the R^2 and the Nash-Sutcliffe-Efficiency (NSE) as shown in Table 2 for both plankton groups. Compared to the GLM and GAM, the more complex model types (RF, GBM, and DL) have a lower RMSE, a higher R^2, and a higher NSE, that is, they generally perform better across all three metrics (Table 2, see also Section 2.2.3 for a description of the metrics). For both pteropods and foraminifers, the RF performs best, followed by the GBM. However, the GBM's R^2 is significantly higher on the training set than on the testing set, which indicates model overfitting. The same pattern is visible for the RMSE (Table 2). In contrast, the RF achieves similar performances on the training an testing set, which indicates a robustly high performance. All model types perform better than using the mean observation value as prediction, which is indicated by the positive NSE values (Table 2). Comparing the R^2 values between the plankton groups shows that the pteropod models generally perform better and can

Table 2
Model Performance for the Pteropod and Foraminifer Models

	Model	R^2_{train}	R^2_{test}	$RMSE_{train}$	$RMSE_{test}$	NSE_{train}	NSE_{test}	Ranking
Pteropods	GLM	0.1113	−0.6427	0.2612	0.3633	0.1113	0.1442	5
	GAM	0.1299	0.1678	0.2585	0.2586	0.1299	0.1678	4
	RF	0.2332	0.2805	0.2408	0.2404	0.5581	0.2805	1
	GBM	0.409	0.2674	0.2114	0.2426	0.3652	0.2674	2
	DL	0.1597	0.1822	0.2521	0.2563	0.1625	0.1822	3
Foraminifers	GLM	0.0503	−0.0279	0.8554	0.8789	0.0503	0.0491	5
	GAM	0.1116	0.0823	0.8274	0.8304	0.1116	0.0823	4
	RF	0.2424	0.2003	0.7586	0.7752	0.4252	0.2003	1
	GBM	0.3999	0.1926	0.6751	0.7789	0.3594	0.1926	2
	DL	0.1718	0.1367	0.7931	0.8054	0.1780	0.1367	3

Note. Each model metric was calculated on both the training set (X_{train}) and the testing set (X_{test}). R^2 ranges from $-\infty$ to +1, with a perfect fit of the model and full variance explained indicated by a value of +1. The root mean squared error (RMSE) is an error measure, hence smaller values show higher accuracy. The Nash-Sutcliffe-efficiency (NSE) indicates improvement of the model predictions over using the observation mean, with perfect model performance indicated by a value of +1 and a value of 0 indicating that the models perform no better than the observation mean. The models are ranked by their performance over the five metrics, where rank 1 refers to the best performing model and rank 5 to the lowest performing one.

explain a higher fraction of the biomass variability (Table 2). For the complex non-parametric models (RF, GBM, DL), R^2 is not an optimal metric (Spiess & Neumeyer, 2010). However, as it is frequently reported in plankton studies as a measure of the fraction of variance explained (Zurell et al., 2020; Pinkerton et al., 2010, 2020), we chose to still include it.

All models tend to underestimate the total biomass on a global scale (−35% for pteropods and −5% for foraminifers of log-transformed biomass), with a stronger underestimation of the top 10th percentile biomass hotspots (on average −78% for pteropods and −53% for foraminifers). However, this underestimation is less pronounced in the more complex models (Figures S17 and S18 in Supporting Information S1). On a basin-scale, highly productive regions are generally underestimated and low productivity areas overestimated with an average overestimation of the lowest 50% of log-transformed biomass by a factor of 8.7 for pteropods and a factor of 2.5 for foraminifers. Hence, biomass concentrations of both plankton groups are underestimated in the North Atlantic Ocean and the tropical Pacific and Atlantic, whereas predictions in the Indian Ocean and the region around Australia are on average too high (Figures S17 and S18 in Supporting Information S1).

3.3. Environmental Covariates

In general, the modeled responses of biomass to the fitted predictors converge across the ensemble members, except near the outer ranges of the predictor values, and for EKE (Figure 4). Temperature shows an overall positive relation to pteropod biomass and a bimodal relation for foraminifer biomass with peaks around 5° to 7°C and above 25°C. Chlorophyll-a is positively related to both pteropod and foraminifer biomass. At high chlorophyll-a concentrations (Chl-a >1 mg m^{-3}), biomass concentrations stagnate for pteropods and decrease slightly for foraminifers. MLD has a negative parabolic relation to pteropod biomass. A deepening of the mixed layer up to a maximum of 30 m depth is associated with a decrease in biomass while a further deepening of the MLD relates to an increase in biomass concentrations. The effect of EKE on foraminifer biomass varies across the models, with a strong positive effect in the simpler GLM and GAM, a near neutral effect in the RF and GBM, and a negative influence in the DL (see Figure 4).

3.4. Global Annual Total Inorganic Carbon (TIC) Export Fluxes

Global mean annual biomass standing stocks are 52.2 Tg TC (ranging from 49.2 to 57.3 Tg TC across SDM types) for pteropods and 0.9 Tg TC (0.6 to 1.1 Tg TC) for foraminifers (Table 3).

The corresponding global annual TIC fluxes were calculated based on growth rate parametrizations (Section 2.3.2) and are on average 14.1 Tg TIC yr^{-1} (13.3 to 15.5 Tg TIC yr^{-1} for the standard TIC-TC conversion rate and growth

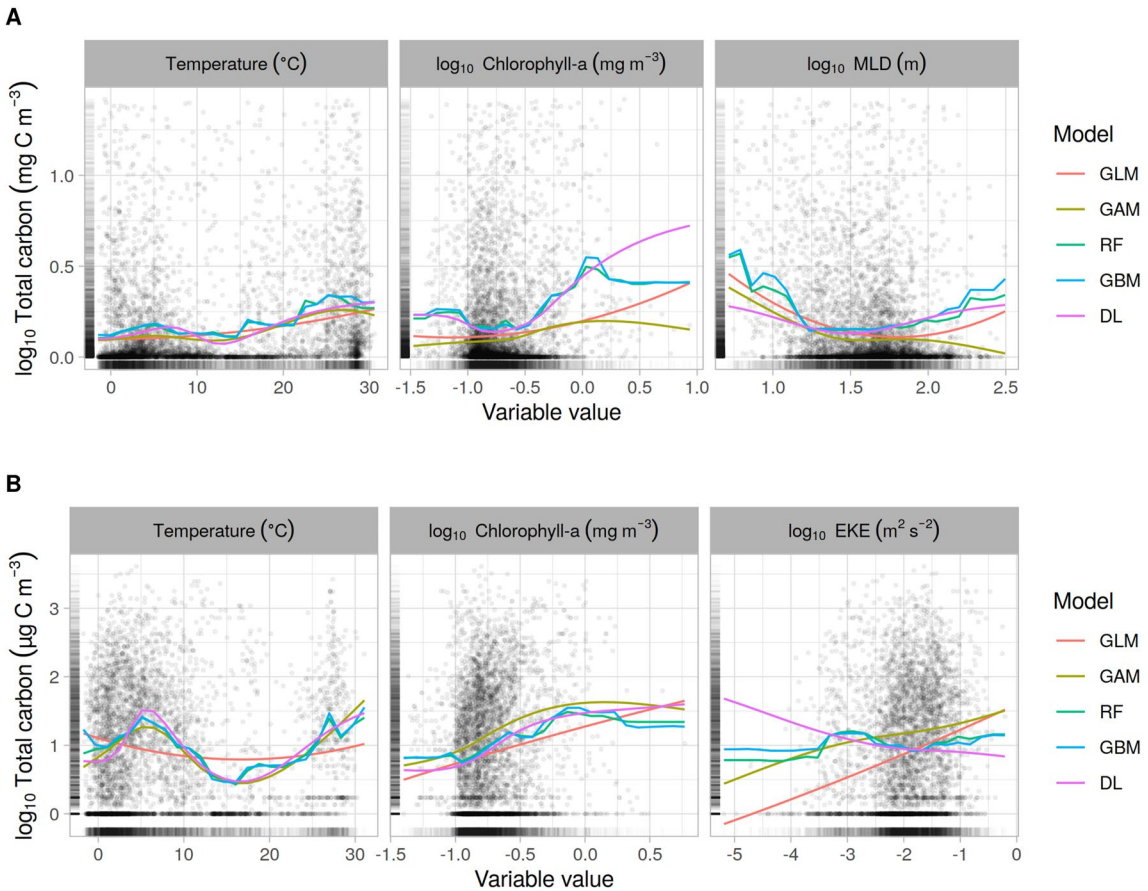

Figure 4. Partial dependence plots (PDP) for the environmental predictors in the pteropod (a) and foraminifer (b) models. The curves indicate the relations learned by the different SDMs and the rug on the *x*- and *y*-axis represents the distribution of the training data. MLD refers to the mixed layer depth, EKE to the eddy kinetic energy. The different model types are the Generalized Linear Model (GLM), Generalized Additive Model (GAM), Random Forest (RF), Boosted Regression Tree (GBM) and Neural Network (DL).

Table 3
Global Mean Annual Surface Pteropod and Foraminifer Total Carbon (TC) Biomass Standing Stocks and Annual Total Inorganic Carbon (TIC) Flux Estimates as Calculated by the Five Species Distribution Model (SDMs)

	Pteropods		Foraminifers	
Model	Standing stock (Tg TC)	Carbon flux (Tg TIC yr^{-1})	Standing stock (Tg TC)	Carbon flux (Tg TIC yr^{-1})
GLM	49.2	13.3 (8.5–19.1)	0.6	8.9 (4.0–16.4)
GAM	49.2	13.3 (8.5–19.1)	0.7	8.5 (3.0–19.0)
RF	57.3	15.5 (9.9–22.3)	1.0	14.2 (5.0–26.8)
GBM	56.9	15.4 (9.8–22.1)	1.1	13.2 (4.8–24.7)
DL	48.3	13.1 (8.4–18.9)	1.1	9.8 (3.7–20.3)
Average	52.2	14.1	0.9	10.9

Note. The uncertainties shown in parentheses for the TIC fluxes correspond to the minimum and maximum flux values derived from the range of TIC-TC conversion factors and growth rate parametrizations used.

rate setup; Table 3) for pteropods. Foraminifer TIC fluxes amount to on average 10.9 Tg TIC yr^{-1} (8.5 to 14.3 Tg TIC yr^{-1}; Table 3). The inter-SDMs range of the TIC fluxes increases by a factor of approximately 4–5 if the modeling uncertainty associated with the TIC-TC factor and the growth rate parametrization are included (Table 3, see also Section 3.5).

3.5. Uncertainty Quantification

To assess the effects of SDM choice, growth rate parametrization and TIC-TC factor parametrization on the TIC flux predictions, we conducted a mANOVA and evaluated spatial patterns of standard deviation between model predictions. The main sources of variability in global mean annual TIC fluxes differ between the plankton groups (Figure S19 in Supporting Information S1). For pteropods, the growth rate and TIC-TC conversion factor choice are the major sources of uncertainty, as each explains 27% of the variability. SDM choice explains 10% of the variability in fluxes for pteropods. In contrast, the TIC flux variability for foraminifers is dominated by the parametrization of the foraminifer growth rate (71%), followed by the model choice (11%), and the TIC-TC factor (<10%).

From a spatial point of view, relative inter-SDMs variability is highest in regions of low productivity and where environmental conditions are outside

the range of the training data set (Figure 3, panels C and D). This encompasses the Southern Hemispheric (SH) oceanic gyres and the low-productivity latitudinal band around 45°S for pteropods and around 30°S for foraminifers. Absolute biomass predictions differ the most in regions of high biomass, that is, mainly the North Atlantic for both groups (Figure 3, panels A and B).

4. Discussion

4.1. Biogeographic Biomass Patterns

The biogeographic distribution patterns found for pteropods and foraminifers largely agree with previous findings (Bednaršek, Mozina, et al., 2012; Buitenhuis et al., 2019; Lalli & Gilmer, 1989; Lombard et al., 2011; Schiebel, 2002). We found high biomass concentrations for both plankton groups in the warm tropical waters, at the high northern latitudes and in the upwelling systems.

The global warm-water belt around the equator has previously been identified as a region of high biomass for pteropods (Bednaršek, Mozina, et al., 2012; Burridge et al., 2017; Lalli & Gilmer, 1989) and foraminifers (Schiebel & Movellan, 2012). High biomass concentrations in the equatorial region for the two plankton groups are representative of total global mesozooplankton distribution patterns (Moriarty et al., 2013; Strömberg et al., 2009), which also show peaks in the tropical ocean.

Earlier studies also found the high latitudes to be regions of high biomass for both plankton groups (Bednaršek, Mozina, et al., 2012; Hunt et al., 2008; Lalli & Gilmer, 1989; Schiebel & Movellan, 2012). Contrary to previous studies (Bednaršek, Mozina, et al., 2012; Hunt et al., 2008; Lalli & Gilmer, 1989), the Southern Ocean was not identified as a region of major pteropod productivity in our study. This is likely due to the influence of the SO-CPR data set, which included a high fraction (95.8%) of absences. Removing all CPR data from our training data set (i.e., SO-CPR and Aus-CPR) significantly increases biomass concentrations for pteropods by a factor of 4–8 in the Southern Ocean but not in other basins (Figures S20 and S21 in Supporting Information S1). However, removing the CPR data also leads to significantly less well constrained PDP curves for low temperatures, which increases the uncertainty of these CPR-depleted SDMs projections. As previous studies were based on much fewer and spatially confined observations in the Southern Ocean (e.g., 141 data points south of 60°S in Bednaršek, Mozina, et al., 2012), it is possible that they constitute local upper bound estimates of pteropod abundance and biomass in the Southern Ocean.

Similar to our findings, upwelling regions have previously been associated with high abundances of pteropods (Burridge et al., 2017; Dadon & Masello, 1999; Koppelmann et al., 2013; McGowan, 1967) and foraminifers (Ivanova et al., 1999; Naidu & Malmgren, 1996; Schiebel et al., 2004). Upwelling systems are characterized by recurrent nutrient inputs that trigger high local primary productivity (Kämpf & Chapman, 2016), and thus optimal conditions for opportunistic foraminifer and pteropod species (Kucera, 2007; Schiebel & Hemleben, 2017). However, the upwelling systems are also associated with the upwelling of low pH waters (Hauri et al., 2013; Joint et al., 2011), and the shoaling of the calcite and aragonite saturation horizon (Frenger et al., 2018; Leinweber & Gruber, 2013), with potentially deleterious effects.

4.2. Environmental Drivers

In agreement with other studies (Beaugrand et al., 2013; Jentzen et al., 2018; Meilland et al., 2016; Pinkerton et al., 2020), temperature was the strongest statistical covariate for the biomass distributions of pteropods and foraminifers in our study. This is not surprising since temperature influences all scales of biological processes, from intra-cellular reaction rates to species interactions (Brown et al., 2004; Chapperon & Seuront, 2011; Kirby & Beaugrand, 2009; Schmidt-Nielsen, 1997). Temperature is also related to the water column stratification, which in turn can affect plankton biomass by influencing nutrient availability (see Section 2.2.1) and primary productivity (Chiswell et al., 2014). The present global dependencies of biomass on temperature can differ from the results of local studies (e.g., a negative dependency of pteropod biomass on temperature as found in Bednaršek et al. (2022) and species-specific responses of foraminifers to a range of environmental parameters as found by Weinkauf et al. (2016)), as our models capture large-scale climatological effects contrary to local models which reflect physiological responses to environmental gradients on a much smaller scale. Such scale-dependent effects of environmental drivers on biogeographic patterns are common in macroecology (Thuiller et al., 2015), but remain to be investigated for marine taxa.

The modelled bimodal structure of the biomass dependency of our SDMs on temperature for foraminifers, and—to a lesser extent—pteropods (peaks around 5°–7°C and above 25°C; Figure 4), likely reflects the existence of distinct assemblages of warm-water species and cold-water species within these groups (Bradshaw, 1959). Some foraminifer species are associated with one end of the temperature spectrum—for instance, *Neogloboquadrina pachyderma* is associated with temperatures below 10°C and *Globigerinoides ruber (white)* with temperatures above 18°C (Kucera, 2007; G. A. Schmidt & Mulitza, 2002), which broadly matches our identified peaks (Antell et al., 2021; Morard et al., 2015; Rillo et al., 2022). Furthermore, the temperature interval around 17°C constitutes a minimum in the foraminifer biomass dependency curve and it can be associated with the subtropical front (D. N. Schmidt et al., 2004). As very dynamic dispersal barriers, fronts are regions of significant environmental variability where foraminifer body sizes were found to be significantly smaller (D. N. Schmidt et al., 2004). This could help explain why the subtropical front was associated with lower foraminifer biomass. However, most foraminifer species display wide thermal tolerances of around 10°C (Schiebel & Hemleben, 2017). Additionally, the sampling density was high in the cold regions of the Southern Ocean and at high temperatures in the tropics, but few data points (18.0% for pteropods and 23.1% for foraminifers) stem from the intermediate temperature range between 10°C and 20°C for both plankton groups. Hence, the bimodal structure might also be skewed due to an uneven sampling distribution (as seen in the uneven density of the *x*-axis rug plot in Figure 4a).

In our SDMs, surface chlorophyll-a concentration emerged as the second-most important environmental covariate for the biomass of both plankton groups, which is also supported by the literature (Meilland et al., 2016; Pinkerton et al., 2020; Schiebel et al., 1995, 2001). Generally, a positive near-linear relationship between chlorophyll-a concentrations and pteropod and foraminifer biomasses is observed, particularly in the well-constrained range of the PDP curve (Figure 4). As a measure of food availability, chlorophyll-a can be directly positively linked to zooplankton abundances and biomass (Pinkerton et al., 2020; Schiebel et al., 2001; Strömberg et al., 2009). However, both pteropods and foraminifers also feed on non-phytoplankton prey and organic particles to varying degrees (Caron & Bé, 1984; Lalli & Gilmer, 1989; Rhumbler, 1911; Spindler et al., 1984). Pteropods typically feed on particles that are one 100–1000th of their own size (Conley et al., 2018), while some foraminifers can digest prey larger than themselves (Schiebel & Hemleben, 2017). This can explain the smaller-scale deviations of the PDP curves from the near-linear trend and a certain decoupling at low chlorophyll-a concentrations (Figure 4) as the zooplankton can feed on alternative organic particles.

The overall importance of the environmental variables driving biomass in models may vary with the spatio-temporal scale at which the analysis is conducted (Corney et al., 2006). Both MLD and EKE were found to be of minor importance as driving variables in our SDMs, which might be due to their predominantly mesoscale effect on mixing and food availability. MLD negatively influences pteropod biomass concentrations over most of the assessed range (Figure 4). As flux-feeders, pteropods rely on a steady downward flux of particles, which can be hindered by a deep and turbulent water column mixing (Tsurumi et al., 2005). From a viewpoint of ecological successions over seasons, the shoaling of the deep winter mixed layer in spring is one of the main factors triggering spring phytoplankton blooms (Chiswell et al., 2014). Following these blooms, zooplankton abundance increases as the zooplankton feed on the phytoplankton (Romagnan et al., 2015). This temporal succession might explain the increase in pteropod biomass for shallow MLD values as an indirect consequence. EKE shows a slight positive impact on foraminifer biomass in the simpler models (Figure 4). At the mesoscale, eddies can sustain increases in foraminifer biomass, as they can drive the mixing of the deep chlorophyll-a maximum into shallower surface layers, that is, into the habitat of foraminifers (Beckmann et al., 1987; Fallet et al., 2011; Kupferman et al., 1986; Schiebel et al., 1995; Steinhardt et al., 2014; Turner, 2015). However, the effect of eddies varies as their direction of rotation determines the dominant vertical direction of water movement (Dufois et al., 2016). The direct large-scale effects of MLD and EKE on biomass patterns are not frequently assessed in the literature (exceptions for MLD are Pinkerton et al. (2020) and Schiebel et al. (2001)). On a local, short-term scale, however, they might have a strong influence on zooplankton biomass that cannot be captured by our global-scale, $1 \times 1°$, monthly model.

Previous work identified carbonate chemistry as an important predictor for net calcification on a local scale (Bednaršek & Ohman, 2015; Bednaršek et al., 2022; Lischka et al., 2011; Manno et al., 2017; Mekkes, Renema, et al., 2021). CO_2-rich waters characterized by low pH, low calcite, and low aragonite saturation states may negatively affect certain calcifying organisms by increasing their dissolution and lowering their calcification rate (Bednaršek, Feely, et al., 2017; Bednaršek et al., 2022; Mekkes, Renema, et al., 2021; Mekkes, Sepúlveda-Rodríguez, et al., 2021). For pteropods, these changes in water chemistry can reduce their metabolic

activity, increase shell dissolution, and decrease their growth and survival (Bednaršek et al., 2016; Bednaršek et al., 2022; Bednaršek, Feely, et al., 2017; Bednaršek, Klinger, et al., 2017; Gardner et al., 2017; Lischka et al., 2011; Lischka & Riebesell, 2012; Maas et al., 2015; Manno et al., 2007). Foraminifers are less sensitive to changes in saturation states as their shells are made of calcite (Orr et al., 2005; Weinkauf et al., 2016), but laboratory and field studies point to reduced calcification rates and metabolic depression for them as pH decreases (C. V. Davis et al., 2017; Iwasaki et al., 2019; Moy et al., 2009; Osborne et al., 2020), even though calcite saturation states are higher than aragonite saturation states. As our statistical models, by design, did not include any indicator of organism physiology or biomineralization and were based on climatological environmental conditions, we could not account for such effects in the way that for example, equation-based numerical models with an explicit representation of organismal physiology, such as individual-based models do (Hofmann Elizondo & Vogt, 2022). Based on field studies, changes in carbonate chemistry have not been shown to cause large-scale abundance decreases of either pteropods or foraminifers (Howes et al., 2015; Ohman et al., 2009; Thibodeau et al., 2019). Nonetheless, in combination with thermal stress, decreased saturation states were shown to negatively affect population sizes on regional scales (Bednaršek et al., 2022). Such combined effects cannot be resolved with our approach as highly correlated predictors must be removed from our models. Furthermore, the fitness reduction of individual organisms leads to delayed responses on the population level, such that large-scale changes detectable by our models may happen only in the near future under climate change (Bednaršek et al., 2022).

Overall, the relative importance and response curves shapes (Figure 4 and Figure S22 in Supporting Information S1) of the various environmental predictors are in line with our current state of knowledge. Some response curves are affected by uneven sampling across environmental and geographic space and scale dependencies, but within the most commonly observed ranges of environmental conditions, the response curves of the five SDMs agree well with each other for both plankton groups.

4.3. Current Global Surface Ocean Biomass and TIC Export Fluxes

Estimates of global plankton standing stocks from observations have only become possible during the past decade (Buitenhuis, Vogt, et al., 2013) due to paucity in the available information about marine ecosystems. Hence, there are still large uncertainties, particularly for organisms such as zooplankton with patchy abundance patterns (Buitenhuis, Hashioka, & Quéré, 2013) and strongly uneven sampling distributions and methodologies (see also Section 4.4 and Figures S23 and S24 in Supporting Information S1 for an assessment of data patchiness). Estimates of standing stocks are highly uncertain, though less so than in marine systems than terrestrial ones (Bar-On et al., 2018; de Garidel-Thoron et al., 2022). In this context, we deem the partly large deviations of our estimates from previous studies as plausible.

On a global mean annual scale, our estimates of total plankton biomass standing stocks are a factor of 10 lower than previous MAREDAT observation-based estimates for pteropods (Bednaršek, Mozina, et al., 2012) and in the same range for foraminifers (Schiebel & Movellan, 2012) as shown in Table 4. For both plankton groups, the previous standing stock estimates were (a) calculated using globally averaged, unweighted biomass concentrations, (b) based on a spatiotemporal subset of our current observational data set, and (c) based only on non-zero abundance observations.

The discrepancy between our results and those of previous studies decreases when we calculate global standing stocks based on these different configurations (see Table 4). Calculating standing stocks based on (a) MAREDAT methodology does not change the standing stock estimates strongly (46 to 57 Tg TC for pteropods and 0.5 to 1.1 Tg TC for foraminifers). Additionally (b) subsetting our prediction fields at the original MAREDAT sampling points increases pteropod standing stock estimates by approximately 50% to 62 to 95 Tg TC, while foraminifer estimates remain near constant at 0.6 to 1.9 Tg TC. Finally, (c) excluding zero abundance observations before modeling increases standing stock estimates to 91 to 140 Tg TC for pteropods and 2 to 3 Tg TC for foraminifers. Combining all three modifications causes an increase of factor 3–4 for both plankton types. Following the same methodology, pteropod biomass estimates are still a factor of 2–4 lower than the MAREDAT estimates, while foraminifer biomass estimates are in the same range. A potential reason for this difference between the two plankton groups might be the variation in patchiness due to the larger body size of pteropods. The original MAREDAT pteropod abundance observations are nearly four times as patchy as those of foraminifers (Buitenhuis, Vogt, et al., 2013), which could have led to a higher bias in the pteropod standing stock estimate. The exclusion of high biomass outliers in our analysis (cf. Section 2.1.3) served to exclude erroneously reported observations to not

Table 4
Comparison Between Modeled Total Carbon (TC) Standing Stocks and Total Inorganic Carbon (TIC) Fluxes for Pteropods and Foraminifers With Previous Studies

	Pteropods		Foraminifers	
Source	Standing stock (Tg TC)	Carbon flux (Tg TIC yr^{-1})	Standing stock (Tg TC)	Carbon flux (Tg TIC yr^{-1})
Estimates based on mechanistic modeling studies				
Buitenhuis et al. (2019)		152–4183[a]		100–141[a]
Gangstø et al. (2008)		300		
Estimates based on observational data				
Bednaršek, Tarling, et al. (2012)	444–505[b,c]	112–150[b,c]		
Schiebel and Movellan (2012)			1–5[b]	3–12[b]
Schiebel (2002)				157–389[b,d]
Our results	49–57	8–22	1–2	3–35
1: MAREDAT methodology	*46–57*	*12–15*	*0.5–1.1*	*5–10*
2: Sampled at MAREDAT points & methodology	*62–95*	*17–26*	*1–2*	*1–4*
3a: W/o zeros	*91–140*	*24–38*	*2–3*	*16–33*
3b*: W/o zeros, MAREDAT points & methodology*	*132–220*	*35–60*	*2–9*	*4–20*
W/o CPR data	*90–155*	*25–42*	*1–3*	*18–33*

Note. All values were converted to represent TC and TIC, respectively. The results of the mechanistic studies from Gangstø et al. (2008) and Buitenhuis et al. (2019) denote the reported $CaCO_3$ production and not the export flux. The export flux calculations include dissolution of the sinking calcium carbonate shells. However, we do not take this into account in the current study. Thus, we compare the production terms before dissolution. The sensitivity analyses are shown in italics. For the comparisons to MAREDAT, the projected biomass maps were sampled at the MAREDAT observation points of the respective plankton group (Bednaršek, Mozina, et al., 2012; Schiebel & Movellan, 2012). To be consistent with the methodology used in Bednaršek, Mozina, et al. (2012) and Schiebel and Movellan (2012), the total standing stocks and fluxes were calculated from global non-weighted mean biomass concentrations and assuming one and nine complete overturn periods for pteropods and foraminifers, respectively.

[a]Based on calcite production, not flux. [b]Based on subset of observations used in this study. [c]Estimates based on non-zero observations only. [d]Flux at 100 m.

skew the analysis. While this may have led to the exclusion of high biomass events, these are considered local and short-term events that cannot be reliably captured and modeled at the scale of our study.

In the context of the marine trophic foodweb, pteropods constitute approximately 6%–8% of total macrozooplankton biomass, whereas foraminifers make up 0.1%–0.6% of microzooplankton biomass as shown in Figure 5 (Buitenhuis, Vogt, et al., 2013). Each plankton size class encompasses a broad range of taxonomic groups, so that the relatively small contributions of pteropods and foraminifers is logical. In contrast to the other PFTs estimates and the earlier MAREDAT estimates for pteropods and foraminifers, our results are based on global climatological biomass estimates instead of spatially discrete observation data. This causes a lower discrepancy between our mean and median estimates as well as a lower total standard deviation (Figure 5), because high biomass extreme events are not as prevalent in our results as in the raw field observations (cf. also Section 4.4 and Figure S23 in Supporting Information S1).

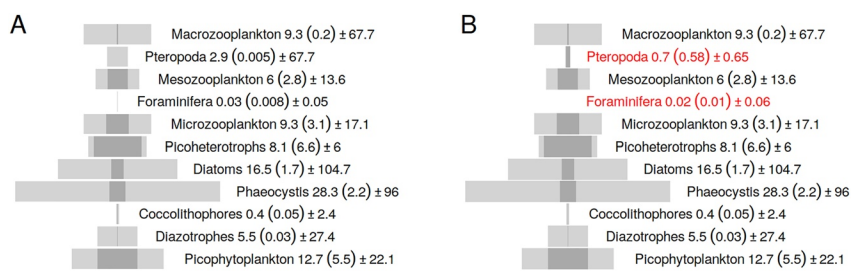

Figure 5. Trophic pyramid of autotrophic and heterotrophic plankton functional types (PFT). The bars show the mean (light gray filling) and median (dark gray filling, value in parentheses) biomass concentrations in µg TC L^{-1} in the surface 200 m. The standard deviation is denoted for each PFT. (a) shows the original results from the MAREDAT project presented in Buitenhuis, Vogt, et al. (2013). (b) shows our updated estimates for pteropods and foraminifers as highlighted in red.

Estimated pteropod TIC fluxes are a factor of 5–100 lower than in previous numerical modeling studies (Table 4). The estimates by Buitenhuis et al. (2019) and Gangstø et al. (2008) are based on mechanistic models which used published laboratory evidence for model calibration and observational data from MAREDAT for model evaluation. However, the parametrization of the growth rate is based on copepod observations instead of pteropods in Buitenhuis, Vogt, et al. (2013) and hence could have introduced a bias. The difference in depth at which TIC-fluxes are reported (100 m in Buitenhuis et al. (2019) and 200 m in our study) likely introduce further uncertainties, however as neither of the results include dissolution effects, these are deemed minor. An additional reason for the discrepancy could be an incomplete representation of the true abundances in our observation data due to sampling biases (cf. Section 4.4). Examples of such biases include net avoidance, diel vertical migration (DVM), and the use of sub-optimal mesh sizes for the target group (e.g., in the CPR), which can lead to underestimated abundances in our observational data (Doubek et al., 2020; Pinkerton et al., 2020; Zamelczyk et al., 2021). Excluding CPR data from our models approximately doubles the estimated TIC fluxes (Table 4), however it also increases uncertainty in the environmental driver dependencies (cf. Section 4.1).

Our foraminifer TIC flux estimates are of the same order of magnitude as the most recent observation-based estimates and mechanistic model-based studies, albeit on the lower end for the latter (Table 4). The earlier observational study by Schiebel (2002) is based on much smaller data sets with a spatial bias toward the highly productive North Atlantic Ocean and found substantially higher TIC fluxes. However, our results align well with the flux estimate calculated by more recent studies, such as Schiebel and Movellan (2012). In our work, we only account for large adults due to mesh size limitations, but including the biomass of earlier life-stages might double foraminifer biomass and flux estimates (Schiebel & Movellan, 2012; see also Section 4.4). This uncertainty could also explain the deviations of our results from the mechanistic model-based estimate by Buitenhuis et al. (2019), which is a factor of 1.5–50 higher than our global annual TIC flux estimate. Similar as for pteropods, excluding CPR data prior to modeling approximately doubles the estimated global annual TIC fluxes (cf. Table 4), which might be indicative of non-optimal representation of foraminifer abundances in this data set.

Pteropods contribute 0.2%–3.2% to total annual global carbonate fluxes and foraminifers contribute 0.1%–3.8%, assuming annual global fluxes amount to 0.7 to 4.7 Pg TIC yr^{-1} (Ziveri et al., 2023). We can assume that the carbon fluxes calculated in our study represent a lower bound estimate due to biases and incompleteness of the observation data set (see Section 4.4). Coccolithophores are estimated to contribute 26%–52% to global carbonate fluxes (C. J. O'Brien, 2015), which leaves 40%–70% of global carbonate fluxes unaccounted for and points to an underestimation of the contribution from the calcifying zooplankton. Additional minor contributors to the marine $CaCO_3$ budget are fishes, atlantid heteropods, pseudothecosomes (particularly the fully shelled *Peracle* species), calcifying ostracods, dinoflagellates, ciliates and the larvae of both benthic molluscs and gymnosomates (Buitenhuis et al., 2019). However, their contribution to global carbonate fluxes is not well constrained, but may range between 3% and 15% per group (Buitenhuis et al., 2019; Schiebel, 2002; Wilson et al., 2009), and hence warrants further investigation.

4.4. Limitations and Uncertainties

Here, we use large global data sets and an exhaustive model ensemble approach to estimate pteropod and foraminifer biomass. We quantify and discuss the uncertainty arising from the model choice and key parametrizations of the growth rate and TIC-TC factor. However, our biomass and carbon flux estimates are affected by the characteristics and errors underlying the observational data and the simplifying assumptions made for the model setup. These include the interaction of spatio-temporal biases in sampling effort with the inherent patchiness of plankton distribution, variations in sampling net mesh sizes, and limited taxonomic resolution for biomass conversions (de Garidel-Thoron et al., 2022).

Patchy sampling across space and time leads to spatio-temporal biases in the training data set (Figure 2 and Figures S4, S23 and S24 in Supporting Information S1). Data coverage is low in the low productivity oligotrophic gyres and during the less productive months (Figure S23 in Supporting Information S1). We find that a large fraction of the inter-model variability is due to environmental conditions outside of or at the outer ranges of the training data (Section 3.5 and Figures 3 and 4). Nonetheless, a large fraction of the global environmental space of our predictor variables is covered by the abundance data sets, which allows us to predict biomass values with higher certainty (Figure S25 in Supporting Information S1). Furthermore, plankton distributions are generally characterized by a high level of seasonal and spatial patchiness (Figures S23 and S24 in Supporting Information S1;

Beckmann et al., 1987; Boltovskoy, 1971; Buitenhuis, Vogt, et al., 2013; Siccha et al., 2012). Patchiness introduces high variance in the observed abundances (Figure S23 in Supporting Information S1), which poses the risk of underestimating the importance of high productivity events occurring on smaller scales. By aggregating the biomass observations over the top 200 m of the ocean and averaging across grid cells, we dampen the variability in our biomass observations and deem this a reasonable approximation for mean annual export fluxes on a global scale. The observation patchiness also causes a mismatch between the gridded monthly climatologies used as environmental predictors and the mesoscale-affected biomass patterns (Benedetti et al., 2021; Righetti et al., 2019). However, previous studies found no significant benefit of using highly temporally resolved data over climatologies (Pinkerton et al., 2020), as the environmental conditions an organism experiences are based on their Lagrangian movement over time (e.g., Hofmann Elizondo & Vogt, 2022). Of course, additional aspects that could not be resolved by our current simplified bottom-up approach are species-specific responses to environmental parameters, as well as top-down controls of the zooplankton biomass by predators. Finally, the use of coarse mesh sizes for sampling relatively small zooplankton can underestimate the true abundances as small and/or mobile individuals are missed (Fabry, 1989; Mack et al., 2012; Miloslavić et al., 2014; Skjoldal et al., 2013; Tseng et al., 2011; Wells, 1973; Zamelczyk et al., 2021). This is particularly relevant for the SO-CPR and Aus-CPR observations, which use a large mesh size of 270 µm (Richardson et al., 2006) and make up 91% and 73% of our training data for pteropods and foraminifers, respectively (Section 2.1.1). These sampling data constraints hence cause our biomass and flux estimates to be lower-end estimates.

Further uncertainties in the standing stock and flux estimates come from the simplified abundance to biomass conversions and the biomass to carbon flux derivation. We assumed species-level or group-level averages for the size-based biomass conversion functions (Section 2.1.2). Yet, in practice these values vary based on ontogenetic stage, subspecies, cryptic species, the presence of symbionts for foraminifers (Schiebel & Hemleben, 2017; Takagi et al., 2019), ambient temperature (Bradshaw, 1959), and food availability (Meilland et al., 2016; Schiebel et al., 2001; Schiebel & Hemleben, 2005), as well as interactive effects of multiple drivers. These factors vary with latitude and we could not account for them explicitly in the present carbon conversions (cf. Section 2.1.2) due to a lack of available parametrizations. Therefore, we likely underestimated the global latitudinal variability in our biomass predictions. Additionally, we could not include the effects of combined temperature stress and ocean acidification on the physiological and population level (Bednaršek et al., 2022). This might have caused an overestimation of calcification rates and populations in specific regions such as upwelling systems that are affected by multiple stressors. To convert biomasses to TIC fluxes, growth rates and the TIC-TC conversion factor were based on spatially constrained data and a limited number of species due to data availability (cf. Section 2.1.2). The choices made for the growth rate function and the TIC-TC factor had a significant impact on flux estimates for pteropods, and for foraminifers to a lesser extent (Section 3.5). Plankton observations identified at a finer taxonomic level and species-specific laboratory-based conversion factors and growth rates would likely increase the accuracy of our calculations. To estimate export fluxes at depth, particle sinking velocities and dissolution rates need to be considered (Schiebel et al., 2007; Takahashi & Bé, 1984). During periods of peak biomass production, high pulses of fast-sinking organisms occur and can drive higher export efficiency (Bach et al., 2019; Schiebel, 2002). However, the relative species abundances observed in our upper ocean foraminifera data (Figure S26 in Supporting Information S1) are in good agreement with those found in sediment trap data in previous studies (e.g., Kretschmer et al., 2018; Lombard et al., 2011). This shows that the foraminifer surface export fluxes and patterns found in our study are representative of export patterns found in the deeper ocean. For pteropods, to our knowledge, no comprehensive global sediment trap data analysis has yet been conducted. To assess comparability between fluxes at the surface and the deep ocean, such an analysis is hence much needed.

5. Conclusion

The aim of this study was to predict global monthly and annual patterns and drivers of shelled pteropod and planktic foraminifer TC biomass distributions, their associated TIC fluxes, and to assess the importance of these groups for the global biogeochemical cycling of carbon and $CaCO_3$.

Globally, pteropods contribute 6%–8% and foraminifers 0.1%–0.6% to total global macrozooplankton and microzooplankton TC standing stocks, respectively. The sinking of their shells and tests constitutes 0.2%–3.2% of the total global annual surface TIC fluxes for pteropods and 0.1%–3.8% for foraminifers. These estimates are associated with significant uncertainty due to sampling data characteristics and simplifications in parametrization and

likely constitute lower-bound estimates. We found biomass hotspots for both plankton groups in the high Northern latitudes, around the equator, and in the upwelling systems. Temperature and chlorophyll-a concentrations were the two most important environmental covariates for modeling the biomass patterns.

Our modeling pipeline, together with novel abundance data from projects such as AtlantECO, can be used to project global biomass patterns for various plankton functional groups. These estimates can be used to validate newly developed mechanistic marine ecosystem models (Clerc et al., 2023; Le Quéré et al., 2016) of increased complexity and higher diversity in zooplankton functional types. Additionally, the models can be employed to assess future changes in plankton biomass by projecting the present models on future environmental fields under climate change scenarios (Benedetti et al., 2021; Tittensor et al., 2021). This is particularly relevant considering the high sensitivity of, for example, pteropods to ocean acidification and warming (Bednaršek et al., 2016; Manno et al., 2016). Thus, we can identify hotspots of future biomass changes (complementary to future changes in diversity as modeled in Benedetti et al., 2021) and potentially link these to risk assessments based on other ocean health indices (Halpern et al., 2012).

Further steps should be taken to close the gaps in our understanding of and quantification of the biomass and biogeography of all taxa that are important for global TIC export fluxes. Research in the field should increase sampling activity in currently under-represented areas of the ocean such as the Southern Hemisphere as well as the oligotrophic gyres to better cover the ecological niches of the modeled taxa. Additionally, the effects of environmental conditions on the TIC-TC conversion factors and the growth rate parametrizations should be investigated in more detail, to better account for latitudinal and regional differences. The modeling pipeline developed here can also be used to model other types of quantitative data, such as sediment trap data (Kucera et al., 2005) or measurements based on novel approaches like underwater imaging techniques (Pesant et al., 2015). Using these data, we could calculate the flux contributions from other calcifying organism groups such as fish and shelled heteropods (Buitenhuis et al., 2019; Wall-Palmer et al., 2016; Wilson et al., 2009) based on a range of independent observations to improve our current uncertainty estimates. Comparing estimates based on upper ocean data with those based on sediment traps could help improve our understanding of export patterns driven by different groups as a function of depth as well as the effects of carbonate dissolution and sinking rates.

Acknowledgments
We thank all contributors involved in the zooplankton field sampling and identification and we acknowledge the efforts made by the community to share such data through publicly available archives. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 862923. This output reflects only the author's view, and the European Union cannot be held responsible for any use that may be made of the information contained therein. We thank David Johns for preparing and sharing the calcifying pteropod counts of the CPR survey. We are grateful to Anthony J. Richardson and Claire Davies for facilitating the access to the Aus-CPR survey. The Aus-CPR data was sourced from Australia's Integrated Marine Observing System (IMOS)—IMOS is enabled by the National Collaborative Research Infrastructure Strategy (NCRIS). We are also very grateful to John Kitchener for facilitating the access to the SO-CPR survey. The Atlantic Meridional Transect is funded by the United Kingdom Natural Environment Research Council through its National Capability Long-term Single Centre Science Program, Climate Linked Atlantic Sector Science (Grant NE/R015953/1). Part of this research was funded by a Vidi Grant (016.161531) from the Dutch Research Council (NWO) awarded to Katja T.C.A. Peijnenburg. Nina Bednaršek acknowledges support from the Slovene Research Agency (ARRS "Biomarkers of subcellular stress in the Northern Adriatic under global environmental change," #J12468), from the NOAA Multi-stressor project, and the NOAA Ocean Acidification and Global Ocean Observing and Monitoring Programs. Open access funding provided by Eidgenossische Technische Hochschule Zurich.

Data Availability Statement

The observational datasets used to train the models as well as the model outputs are publicly available on the data platform PANGAEA (https://doi.org/10.1594/PANGAEA.957258; Knecht et al., 2023). An adapted version of the modelling pipeline, applicable to any species abundance or biomass dataset in the AtlantECO format is available on Zenodo (https://doi.org/10.5281/zenodo.7888452; Knecht, 2023).

References

Anderson, O. R., Spindler, M., Bé, A. W. H., & Hemleben, C. (1979). Trophic activity of planktonic foraminifera. *Journal of the Marine Biological Association of the United Kingdom*, *59*(3), 791–799. https://doi.org/10.1017/S002531540004577X

Anglada-Ortiz, G., Zamelczyk, K., Meilland, J., Ziveri, P., Chierici, M., Fransson, A., & Rasmussen, T. L. (2021). Planktic foraminiferal and pteropod contributions to carbon dynamics in the Arctic Ocean (North Svalbard Margin). *Frontiers in Marine Science*, *8*(June). https://doi.org/10.3389/fmars.2021.661158

Antell, G. S., Fenton, I. S., Valdes, P. J., & Saupe, E. E. (2021). Thermal niches of planktonic foraminifera are static throughout glacial-interglacial climate change. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(18), 1–9. https://doi.org/10.1073/pnas.2017105118

Bach, L. T., Stange, P., Taucher, J., Achterberg, E. P., Algueró-Muñiz, M., Horn, H., et al. (2019). The influence of plankton community structure on sinking velocity and remineralization rate of marine aggregates. *Global Biogeochemical Cycles*, *33*(8), 971–994. https://doi.org/10.1029/2019GB006256

Bar-On, Y. M., Phillips, R., & Milo, R. (2018). The biomass distribution on Earth. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(25), 6506–6511. https://doi.org/10.1073/pnas.1711842115

Barton, A. D., Irwin, A. J., Finkel, Z. V., & Stock, C. A. (2016). Anthropogenic climate change drives shift and shuffle in North Atlantic phytoplankton communities. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(11), 2964–2969. https://doi.org/10.1073/pnas.1519080113

Bé, A. W. H., & Gilmer, R. W. (1977). A zoogeographic and taxonomic review of euthecosomatous Pteropoda. *Oceanic micropaleontology*, *1*(6), 733–808.

Beaugrand, G., Edwards, M., & Legendre, L. (2010). Marine biodiversity, ecosystem functioning, and carbon cycles. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(22), 10120–10124. https://doi.org/10.1073/pnas.0913855107

Beaugrand, G., Mcquatters-Gollop, A., Edwards, M., & Goberville, E. (2013). Long-term responses of North Atlantic calcifying plankton to climate change. *Nature Climate Change*, *3*(3), 263–267. https://doi.org/10.1038/nclimate1753

Beckmann, W., Auras, A., & Hemleben, C. (1987). Cyclonic cold-core eddy in the eastern North Atlantic. III. Zooplankton. *Marine Ecology Progress Series*, *39*, 165–173. https://doi.org/10.3354/meps039165

Bednaršek, N., Carter, B. R., Mccabe, R. M., Feely, R. A., Howard, E., Chavez, F., et al. (2022). Pelagic calcifiers face increased mortality and habitat loss with warming and ocean acidification. *Authorea*. https://doi.org/10.22541/au.164865092.29568156/v1

Bednaršek, N., Feely, R. A., Beck, M. W., Glippa, O., Kanerva, M., & Engström-Öst, J. (2018). El Niño-related thermal stress coupled with upwelling-related ocean acidification negatively impacts cellular to population-level responses in pteropods along the California current system with implications for increased bioenergetic costs. *Frontiers in Marine Science*, *5*(DEC), 1–17. https://doi.org/10.3389/fmars.2018.00486

Bednaršek, N., Feely, R. A., Tolimieri, N., Hermann, A. J., Siedlecki, S. A., Waldbusser, G. G., et al. (2017). Exposure history determines pteropod vulnerability to ocean acidification along the US West Coast article. *Scientific Reports*, *7*(1), 1–12. https://doi.org/10.1038/s41598-017-03934-z

Bednaršek, N., Harvey, C. J., Kaplan, I. C., Feely, R. A., & Možina, J. (2016). Pteropods on the edge: Cumulative effects of ocean acidification, warming, and deoxygenation. *Progress in Oceanography*, *145*, 1–24. https://doi.org/10.1016/j.pocean.2016.04.002

Bednaršek, N., Klinger, T., Harvey, C. J., Weisberg, S., McCabe, R. M., Feely, R. A., et al. (2017). New ocean, new needs: Application of pteropod shell dissolution as a biological indicator for marine resource management. *Ecological Indicators*, *76*, 240–244. https://doi.org/10.1016/j.ecolind.2017.01.025

Bednaršek, N., Mozina, J., Vogt, M., O'Brien, C., & Tarling, G. A. (2012). The global distribution of pteropods and their contribution to carbonate and carbon biomass in the modern ocean. *Earth System Science Data*, *4*(1), 167–186. https://doi.org/10.5194/essd-4-167-2012

Bednaršek, N., & Ohman, M. D. (2015). Changes in pteropod distributions and shell dissolution across a frontal system in the California Current System. *Marine Ecology Progress Series*, *523*, 93–103. https://doi.org/10.3354/meps11199

Bednaršek, N., Tarling, G. A., Fielding, S., & Bakker, D. C. (2012). Population dynamics and biogeochemical significance of Limacina helicina Antarctica in the Scotia Sea (Southern Ocean). *Deep-Sea Research Part II Topical Studies in Oceanography*, *59–60*, 105–116. https://doi.org/10.1016/j.dsr2.2011.08.003

Bell, D. M., & Schlaepfer, D. R. (2016). On the dangers of model complexity without ecological justification in species distribution modeling. *Ecological Modelling*, *330*, 50–59. https://doi.org/10.1016/j.ecolmodel.2016.03.012

Benedetti, F., Vogt, M., Elizondo, U. H., Righetti, D., Zimmermann, N. E., & Gruber, N. (2021). Major restructuring of marine plankton assemblages under global warming. *Nature Communications*, *12*(1), 1–15. https://doi.org/10.1038/s41467-021-25385-x

Boehmke, B., & Greenwell, B. (2019a). Deep learning. In *Hands-on machine learning with r* (1st ed.). Chapman and Hall/CRC. chap. 12. https://doi.org/10.1201/9780367816377

Boehmke, B., & Greenwell, B. (2019b). Gradient boosting. In *Hands-on machine learning with r* (1st ed.). Chapman and Hall/CRC. chap. 11. https://doi.org/10.1201/9780367816377

Boehmke, B., & Greenwell, B. (2019c). Interpretable machine learning. In *Hands-on machine learning with r*. Chapman and Hall/CRC. chap. 16. https://doi.org/10.1201/9780367816377

Boehmke, B., & Greenwell, B. (2019d). Modeling process. In *Hands-on machine learning with r2* (1st ed.). Chapman and Hall/CRC. chap. 2.

Boehmke, B., & Greenwell, B. (2019e). Random forests. In *Hands-on machine learning with r* (1st ed.). Chapman and Hall/CRC. chap. 11. https://doi.org/10.1201/9780367816377

Boltovskoy, E. (1971). Patchiness in the distribution of planktonic foraminifera. In *Proceedings of the 2. planktonic conference* (pp. 107–115). Tecnoscienza.

Boyce, D. G., Lewis, M. R., & Worm, B. (2010). Global phytoplankton decline over the past century. *Nature*, *466*(7306), 591–596. https://doi.org/10.1038/nature09268

Boyer, T. P., García, H. E., Locarnini, R. A., Zweng, M. M., Mishonov, A. V., Reagan, J. R., et al. (2018). World Ocean Atlas 2018. Retrieved from https://www.ncei.noaa.gov/archive/accession/NCEI-WOA18

Bradshaw, J. S. (1959). Ecology of living planktonic foraminifera in the north and equatorial Pacific Ocean. *Cushman Foundation Foraminiferal Research Contribution*, *10*(2), 25–64.

Brandão, M. C., Benedetti, F., Martini, S., Soviadan, Y. D., Irisson, J. O., Romagnan, J. B., et al. (2021). Macroscale patterns of oceanic zooplankton composition and size structure. *Scientific Reports*, *11*(1), 1–19. https://doi.org/10.1038/s41598-021-94615-5

Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., & West, G. B. (2004). Toward a metabolic theory of ecology. *Ecology*, *85*(7), 1771–1789. https://doi.org/10.1890/03-9000

Brun, P., Kiørboe, T., Licandro, P., & Payne, M. R. (2016). The predictive skill of species distribution models for plankton in a changing climate. *Global Change Biology*, *22*(9), 3170–3181. https://doi.org/10.1111/gcb.13274

Brun, P., Thuiller, W., Chauvier, Y., Pellissier, L., Wüest, R. O., Wang, Z., & Zimmermann, N. E. (2020). Model complexity affects species distribution projections under climate change. *Journal of Biogeography*, *47*(1), 130–142. https://doi.org/10.1111/jbi.13734

Brun, P., Vogt, M., Payne, M. R., Gruber, N., O'Brien, C. J., Buitenhuis, E. T., et al. (2015). Ecological niches of open ocean phytoplankton taxa. *Limnology & Oceanography*, *60*(3), 1020–1038. https://doi.org/10.1002/lno.10074

Buitenhuis, E. T., Hashioka, T., & Quéré, C. L. (2013). Combined constraints on global ocean primary production using observations and models. *Global Biogeochemical Cycles*, *27*(3), 847–858. https://doi.org/10.1002/gbc.20074

Buitenhuis, E. T., Le Quéré, C., Bednaršek, N., & Schiebel, R. (2019). Large contribution of pteropods to shallow $CaCO_3$ export. *Global Biogeochemical Cycles*, *33*(3), 458–468. https://doi.org/10.1029/2018GB006110

Buitenhuis, E. T., Vogt, M., Moriarty, R., Bednaršek, N., Doney, S. C., Leblanc, K., et al. (2013). MAREDAT: Towards a world atlas of MARine Ecosystem DATa. *Earth System Science Data*, *5*(2), 227–239. https://doi.org/10.5194/essd-5-227-2013

Burba, G., & Anderson, D. (2005). *A brief practical guide to eddy covariance flux measurements (Tech. Rep.)*. LI-COR Biosciences. https://doi.org/10.1076/ceyr.18.1.62.5393

Burridge, A. K., Goetze, E., Wall-Palmer, D., Le Double, S. L., Huisman, J., & Peijnenburg, K. T. C. A. (2017). Diversity and abundance of pteropods and heteropods along a latitudinal gradient across the Atlantic Ocean. *Progress in Oceanography*, *158*, 213–223. https://doi.org/10.1016/j.pocean.2016.10.001

Caldeira, K., & Wickett, M. E. (2003). Anthropogenic carbon and ocean pH. *Nature*, *425*(6956), 365. https://doi.org/10.1038/425365a

Caron, D. A., & Bé, A. W. H. (1984). Predicted and observed feeding rates of the spinose planktonic foraminifer Globigerinoides sacculifer. *Bulletin of Marine Science*, *35*(1), 1–10.

Carton, J. A., Chepurin, G. A., & Chen, L. (2018). SODA3: A new ocean climate reanalysis. *Journal of Climate*, *31*(17), 6967–6983. https://doi.org/10.1175/jcli-d-18-0149.1

Chapperon, C., & Seuront, L. (2011). Behavioral thermoregulation in a tropical gastropod: Links to climate change scenarios. *Global Change Biology*, *17*(4), 1740–1749. https://doi.org/10.1111/j.1365-2486.2010.02356.x

Chiswell, S. M., Calil, P. H., & Boyd, P. W. (2014). Spring blooms and annual cycles of phytoplankton: A unified perspective. *Journal of Plankton Research*, *37*(3), 500–508. https://doi.org/10.1093/plankt/fbv021

Clerc, C., Bopp, L., Benedetti, F., Vogt, M., & Aumont, O. (2023). Including filter-feeding gelatinous macrozooplankton in a global marine biogeochemical model: Model–data comparison and impact on the ocean carbon cycle. *Biogeosciences*, *20*, 869–895. https://doi.org/10.5194/bg-20-869-2023

Conley, K. R., Lombard, F., & Sutherland, K. R. (2018). Mammoth grazers on the ocean's minuteness: A review of selective feeding using mucous meshes. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1878), 20180056. https://doi.org/10.1098/rspb.2018.0056

Copernicus (2021). Climate data store of the copernicus marine environment. https://doi.org/10.24381/cds.4c328c78

Corney, P. M., Duc, M. G., Smart, S. M., Kirby, K. J., Bunce, R. G., & Marrs, R. H. (2006). Relationships between the species composition of forest field-layer vegetation and environmental drivers, assessed using a National Scale Survey. *Journal of Ecology*, *94*(2), 383–401. https://doi.org/10.1111/j.1365-2745.2006.01094.x

Dadon, J. R., & Masello, J. F. (1999). Mechanisms generating and maintaining the admixture of zooplanktonic molluscs (Euthecosomata: Opistobranchiata: Gastropoda) in the subtropical front of the South Atlantic. *Marine Biology*, *135*(1), 171–179. https://doi.org/10.1007/s002270050614

Davis, C. S., & Wiebe, P. H. (1985). Macrozooplankton biomass in a warm-core Gulf Stream ring: Time series changes in size structure, taxonomic composition, and vertical distribution. *Journal of Geophysical Research*, *90*(C5), 8871–8884. https://doi.org/10.1029/JC090iC05p08871

Davis, C. V., Rivest, E. B., Hill, T. M., Gaylord, B., Russell, A. D., & Sanford, E. (2017). Ocean acidification compromises a planktic calcifier with implications for global carbon cycling. *Scientific Reports*, *7*(1), 1–8. https://doi.org/10.1038/s41598-017-01530-9

de Garidel-Thoron, T., Chaabane, S., Giraud, X., Meilland, J., Jonkers, L., Kucera, M., et al. (2022). The foraminiferal response to climate stressors project: Tracking the community response of planktonic foraminifera to historical climate change. *Frontiers in Marine Science*, *9*(May), 1–6. https://doi.org/10.3389/fmars.2022.827962

De Broyer, C., Koubbi, P., Griffiths, H. J., Raymond, B., d'Udekem d'Acoz, C., Van de Putte, A. P., et al. (Eds.) (2014). *Biogeographic Atlas of the Southern Ocean*. SCAR.

Dietrich, G., Kalle, K., Krauss, W., & Siedler, G. (1975). *Allgemeine Meereskunde* (3rd ed.). Gebrüder Bornträger.

Doney, S. C., Fabry, V. J., Feely, R. A., & Kleypas, J. A. (2009). Ocean acidification: The other CO_2 problem. *Annual Review of Marine Science*, *1*(1), 169–192. https://doi.org/10.1146/annurev.marine.010908.163834

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., et al. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*(1), 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Doubek, J. P., Goldfarb, S. K., & Stockwell, J. D. (2020). Should we be sampling zooplankton at night? *Limnology and Oceanography Letters*, *5*(4), 313–321. https://doi.org/10.1002/lol2.10151

Dufois, F., Hardman-Mountford, N. J., Greenwood, J., Richardson, A. J., Feng, M., & Matear, R. J. (2016). Anticyclonic eddies are more productive than cyclonic eddies in subtropical gyres because of winter mixing. *Science Advances*, *2*(5), 1–7. https://doi.org/10.1126/sciadv.1600282

Elith, J., Kearney, M., & Phillips, S. (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution*, *1*(4), 330–342. https://doi.org/10.1111/j.2041-210x.2010.00036.x

Elith, J., & Leathwick, J. R. (2009). Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology Evolution and Systematics*, *40*(1), 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

Fabry, V. J. (1989). Aragonite production by pteropod molluscs in the subarctic Pacific. *Deep-Sea Research, Part A: Oceanographic Research Papers*, *36*(11), 1735–1751. https://doi.org/10.1016/0198-0149(89)90069-1

Fabry, V. J., Seibel, B. A., Feely, R. A., & Orr, J. C. (2008). Impacts of ocean acidification on marine fauna and ecosystem processes. *ICES Journal of Marine Science*, *65*(3), 414–432. https://doi.org/10.1093/icesjms/fsn048

Fallet, U., Ullgren, J. E., Castañeda, I. S., van Aken, H. M., Schouten, S., Ridderinkhof, H., & Brummer, G.-J. A. (2011). Contrasting variability in foraminiferal and organic paleotemperature proxies in sedimenting particles of the Mozambique Channel (SW Indian Ocean). *Geochimica et Cosmochimica Acta*, *75*(20), 5834–5848. https://doi.org/10.1016/j.gca.2011.08.009

Fisher, R. A., & Yates, F. (1953). *Statistical tables for biological, agricultural and medical research*. Hafner Publishing Company.

Frenger, I., Bianchi, D., Stührenberg, C., Oschlies, A., Dunne, J., Deutsch, C., et al. (2018). Biogeochemical role of subsurface coherent eddies in the ocean: Tracer cannonballs, hypoxic storms, and microbial stewpots? *Global Biogeochemical Cycles*, *32*(2), 226–249. https://doi.org/10.1002/2017GB005743

Frerichs, W. E., Heiman, M. E., Borgman, L. E., & Be, A. W. H. (1972). Latitudal variations in planktonic foraminiferal test porosity; Part 1, Optical studies. *Journal of Foraminiferal Research*, *2*(1), 6–13. https://doi.org/10.2113/gsjfr.2.1.6

Gangstø, R., Gehlen, M., Schneider, B., Bopp, L., Aumont, O., & Joos, F. (2008). Modeling the marine aragonite cycle: Changes under rising carbon dioxide and its role in shallow water $CaCO_3$ dissolution. *Biogeosciences*, *5*(4), 1057–1072. https://doi.org/10.5194/bg-5-1057-2008

Garcia, H. E., Weathers, K. W., Paver, C. R., Smolyar, I., Boyer, T. P., Locarnini, M. M., et al.(2019a). *World Ocean Atlas 2018: Dissolved oxygen, apparent oxygen utilization, and dissolved oxygen saturation. (Tech. Rep.)*. (Vol. 3). NOAA. Retrieved from https://www.ncei.noaa.gov/access/world-ocean-atlas-2018/bin/woa18oxnu.pl?parameter=o

Garcia, H. E., Weathers, K. W., Paver, C. R., Smolyar, I., Boyer, T. P., Locarnini, M. M., et al. (2019). *World ocean Atlas 2018: Dissolved inorganic nutrients (phosphate, nitrate and nitrate + nitrite, silicate) (Tech. Rep.)* (Vol. 4). NOAA. Retrieved from https://data.nodc.noaa.gov/woa/WOA18/DOC/woa18_vol4.pdf

Gardner, J., Manno, C., Bakker, D. C. E., Peck, V. L., & Tarling, G. A. (2017). Southern Ocean pteropods at risk from ocean warming and acidification. *Marine Biology*, *165*(1), 8. https://doi.org/10.1007/s00227-017-3261-3

Gilmer, R. W., & Harbison, G. R. (1986). Morphology and field behavior of pteropod molluscs: Feeding methods in the families Cavoliniidae, Limacinidae and Peraclididae (Gastropoda: Thecosomata). *Marine Biology*, *91*(1), 47–57. https://doi.org/10.1007/bf00397570

Govoni, J. J., Hare, J. A., Davenport, E. D., Chen, M. H., & Marancik, K. E. (2010). Mesoscale, cyclonic eddies as larval fish habitat along the southeast United States shelf: A Lagrangian description of the zooplankton community. *ICES Journal of Marine Science*, *67*(3), 403–411. https://doi.org/10.1093/icesjms/fsp269

Gregor, L., & Gruber, N. (2021). OceanSODA-ETHZ: A global gridded data set of the surface ocean carbonate system for seasonal to decadal studies of ocean acidification. *Earth System Science Data*, *13*(2), 777–808. https://doi.org/10.5194/essd-13-777-2021

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, *135*(2–3), 147–186. https://doi.org/10.1016/S0304-3800(00)00354-9

H2O.ai. (2021). R interface for H_2O. Retrieved from https://github.com/h2oai/h2o-3

Halpern, B. S., Longo, C., Hardy, D., McLeod, K. L., Samhouri, J. F., Katona, S. K., et al. (2012). An index to assess the health and benefits of the global ocean. *Nature*, *488*(7413), 615–620. https://doi.org/10.1038/nature11397

Hastie, T. J., & Tibshirani, R. J. (1990). Generalized additive models. In *Generalized additive models* (1st ed., pp. 136–173). Routledge. https://doi.org/10.1201/9780203753781

Hauri, C., Gruber, N., McDonnell, A. M., & Vogt, M. (2013). The intensity, duration, and severity of low aragonite saturation state events on the California continental shelf. *Geophysical Research Letters*, *40*(13), 3424–3428. https://doi.org/10.1002/grl.50618

Helaouët, P., & Beaugrand, G. (2009). Physiology, ecological niches and species distribution. *Ecosystems*, *12*(8), 1235–1245. https://doi.org/10.1007/s10021-009-9261-5

Hemleben, C., Spindler, M., & Anderson, O. R. (1989). Host and symbiont relationships. In *Modern planktonic foraminifera* (pp. 86–111). Springer.

Hofmann Elizondo, U., & Vogt, M. (2022). Individual-based modeling of shelled pteropods. *Ecological Modelling*, *468*, 109944. https://doi.org/10.1016/j.ecolmodel.2022.109944

Horton, T., Gofas, S., Kroh, A., Poore, G. C., Read, G., Rosenberg, G., et al. (2017). Improving nomenclatural consistency: A decade of experience in the world register of marine species. *European Journal of Taxonomy*, *389*, 1–24. https://doi.org/10.5852/ejt.2017.389

Hosie, G. (2021). *Southern Ocean continuous plankton recorder zooplankton records, ver. 9*. Australian Antarctic Data Centre. https://doi.org/10.26179/ksds-s610

Hosmer, D. W., Jr., Lemeshow, S., & Sturdivant, R. X. (2013). Assessing the fit of the model. In *Applied logistic regression* (Vol. 398, pp. 153–225). John Wiley & Sons.

Howes, E. L., Stemmann, L., Assailly, C., Irisson, J. O., Dima, M., Bijma, J., & Gattuso, J. P. (2015). Pteropod time series from the North Western Mediterranean (1967-2003): Impacts of pH and climate variability. *Marine Ecology Progress Series*, *531*, 193–206. https://doi.org/10.3354/meps11322

Hunt, B. P., Pakhomov, E. A., Hosie, G. W., Siegel, V., Ward, P., & Bernard, K. (2008). Pteropods in Southern Ocean ecosystems. *Progress in Oceanography*, *78*(3), 193–221. https://doi.org/10.1016/j.pocean.2008.06.001

Iglesias-Rodriguez, M. D., Armstrong, R., Feely, R., Hood, R., Kleypas, J., Milliman, J. D., et al. (2002). Progress made in study of ocean's calcium carbonate budget. *Eos*, *83*(34), 2000–2002. https://doi.org/10.1029/2002EO000267

IMOS. (2022). IMOS - AusCPR: Zooplankton abundance. NASA/Global Change Master Directory (GCMD) earth science keywords. Version 6.0.0.0.0. Retrieved from https://catalogue-imos.aodn.org.au/geonetwork/srv/eng/catalog.search#/metadata/c1344e70-480e-0993-e044-00144f7bc0f4

Ivanova, E. M., Conan, S. M.-H., Peeters, F. J. C., & Troelstra, S. R. (1999). Living Neogloboquadrina pachyderma sin and its distribution in the sediments from Oman and Somalia upwelling areas. *Marine Micropaleontology*, *36*(2–3), 91–107. https://doi.org/10.1016/s0377-8398(98)00027-9

Iwasaki, S., Kimoto, K., Sasaki, O., Kano, H., & Uchida, H. (2019). Sensitivity of planktic foraminiferal test bulk density to ocean acidification. *Scientific Reports*, *9*(1), 1–9. https://doi.org/10.1038/s41598-019-46041-x

Jentzen, A., Schönfeld, J., & Schiebel, R. (2018). Assessment of the effect of increasing temperature on the ecology and assemblage structure of modern planktic foraminifers in the Caribbean and surrounding seas. *Journal of Foraminiferal Research*, *48*(3), 251–272. https://doi.org/10.2113/gsjfr.48.3.251

Johns, D. M. B. A. O. T. U. M. (2021). Continuous plankton recorder data for all coccolithophores, foraminifera and thecosomata - All areas. *The Archive for Marine Species and Habitats Data (DASSH)*. https://doi.org/10.17031/1763

Joint, I., Doney, S. C., & Karl, D. M. (2011). Will ocean acidification affect marine microbes. *ISME Journal*, *5*(1), 1–7. https://doi.org/10.1038/ismej.2010.79

Jonkers, L., Hillebrand, H., & Kucera, M. (2019). Global change drives modern plankton communities away from the pre-industrial state. *Nature*, *570*(7761), 372–375. https://doi.org/10.1038/s41586-019-1230-3

Kämpf, J., & Chapman, P. (2016). *Upwelling systems of the world*. Springer.

Kirby, R. R., & Beaugrand, G. (2009). Trophic amplification of climate warming. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1676), 4095–4103. https://doi.org/10.1098/rspb.2009.1320

Knecht, N. (2023). AtlantECO pipeline v1.0.0. https://doi.org/10.5281/zenodo.7888452

Knecht, N., Benedetti, F., Hofmann Elizondo, U., Bednaršek, N., Chaabane, S., de Weerd, C., et al. (2023). Global distributions of calcifying zooplankton abundance and biomass - curated observations (as.CSV) and gridded products (biomass climatologies and model outputs, as NetCDF). https://doi.org/10.1594/PANGAEA.957258

Koppelmann, R., Kullmann, B., Lahajnar, N., Martin, B., & Mohrholz, V. (2013). Onshore - Offshore distribution of Thecosomata (Gastropoda) in the benguela current upwelling region off Namibia: Species diversity and trophic position. *Journal of the Marine Biological Association of the United Kingdom*, *93*(6), 1625–1640. https://doi.org/10.1017/S0025315413000052

Kretschmer, K., Jonkers, L., Kucera, M., & Schulz, M. (2018). Modeling seasonal and vertical habitats of planktonic foraminifera on a global scale. *Biogeosciences*, *15*(14), 4405–4429. https://doi.org/10.5194/bg-15-4405-2018

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, *47*(260), 583–621. https://doi.org/10.1080/01621459.1952.10483441

Kucera, M. (2007). Chapter six planktonic foraminifera as tracers of past oceanic environments. *Developments in Marine Geology*, *1*(07), 213–262. https://doi.org/10.1016/S1572-5480(07)01011-1

Kucera, M., Weinelt, M., Kiefer, T., Pflaumann, U., Hayes, A., Chen, M., et al. (2005). Reconstruction of the glacial Atlantic and Pacific sea-surface temperatures from assemblages of planktonic foraminifera: Multi-technique approach based on geographically constrained calibration datasets. *Quaternary Science Reviews*, *24*(7–9), 951–998. https://doi.org/10.1016/j.quascirev.2004.07.014

Kupferman, S. L., Becker, G. A., Simmons, W. F., Schauer, U., Marietta, M. G., & Nies, H. (1986). An intense cold core eddy in the North-East Atlantic. *Nature*, *319*(6053), 474–477. https://doi.org/10.1038/319474a0

Lalli, C. M., & Gilmer, R. W. (1989). *Pelagic snails: The biology of holoplanktonic gastropod mollusks*. Stanford University Press.

Larson, R. J. (1986). Water content, organic content, and carbon and nitrogen composition of medusae from the northeast Pacific. *Journal of Experimental Marine Biology and Ecology*, *99*(2), 107–120. https://doi.org/10.1016/0022-0981(86)90231-5

Leinweber, A., & Gruber, N. (2013). Variability and trends of ocean acidification in the Southern California Current System: A time series from Santa Monica Bay. *Journal of Geophysical Research: Oceans*, *118*(7), 3622–3633. https://doi.org/10.1002/jgrc.20259

Le Quéré, C., Buitenhuis, E. T., Moriarty, R., Alvain, S., Aumont, O., Bopp, L., et al. (2016). Role of zooplankton dynamics for Southern Ocean phytoplankton biomass and global biogeochemical cycles. *Biogeosciences*, *13*(14), 4111–4133. https://doi.org/10.5194/bg-13-4111-2016

Lischka, S., Büdenbender, J., Boxhammer, T., & Riebesell, U. (2011). Impact of ocean acidification and elevated temperatures on early juveniles of the polar shelled pteropod Limacina helicina: Mortality, shell degradation, and shell growth. *Biogeosciences*, *8*(4), 919–932. https://doi.org/10.5194/bg-8-919-2011

Lischka, S., & Riebesell, U. (2012). Synergistic effects of ocean acidification and warming on overwintering pteropods in the Arctic. *Global Change Biology*, *18*(12), 3517–3528. https://doi.org/10.1111/gcb.12020

Locarnini, M. M., Mishonov, A. V., Baranova, O. K., Boyer, T. P., Zweng, M. M., Garcia, H. E., et al. (2018). World Ocean Atlas 2018. In *Temperature (tech. Rep.)* (Vol. 1). NOAA. Retrieved from https://www.ncei.noaa.gov/access/world-ocean-atlas-2018/bin/woa18.pl?parameter=t

Lombard, F., Labeyrie, L., Michel, E., Bopp, L., Cortijo, E., Retailleau, S., et al. (2011). Modelling planktic foraminifer growth and distribution using an ecophysiological multi-species approach. *Biogeosciences*, *8*(4), 853–873. https://doi.org/10.5194/bg-8-853-2011

Lombard, F., Labeyrie, L., Michel, E., Spero, H. J., & Lea, D. W. (2009). Modelling the temperature dependent growth rates of planktic foraminifera. *Marine Micropaleontology*, *70*(1–2), 1–7. https://doi.org/10.1016/j.marmicro.2008.09.004

Longhurst, A. R. (2007). Physical control of ecological processes. In A. R. Longhurst (Ed.), *Ecological geography of the sea* (2nd ed., pp. 51–70). Academic Press. https://doi.org/10.1016/b978-012455521-1/50005-x

Maas, A. E., Lawson, G. L., & Tarrant, A. M. (2015). Transcriptome-wide analysis of the response of the thecosome pteropod Clio pyramidata to short-term CO_2 exposure. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, *16*, 1–9. https://doi.org/10.1016/j.cbd.2015.06.002

Mack, H. R., Conroy, J. D., Blocksom, K. A., Stein, R. A., & Ludsin, S. A. (2012). A comparative analysis of zooplankton field collection and sample enumeration methods. *Limnology and Oceanography: Methods*, *10*(JANUARY), 41–53. https://doi.org/10.4319/lom.2012.10.41

MacKas, D. L., & Galbraith, M. D. (2012). Pteropod time-series from the NE Pacific. *ICES Journal of Marine Science*, *69*(3), 448–459. https://doi.org/10.1093/icesjms/fsr163

Mackas, D. L., Tsurumi, M., Galbraith, M. D., & Yelland, D. R. (2005). Zooplankton distribution and dynamics in a North Pacific Eddy of coastal origin: II. Mechanisms of eddy colonization by and retention of offshore species. *Deep-Sea Research Part II Topical Studies in Oceanography*, *52*(7–8), 1011–1035. https://doi.org/10.1016/j.dsr2.2005.02.008

Manno, C., Bednaršek, N., Tarling, G. A., Peck, V. L., Comeau, S., Adhikari, D., et al. (2017). Shelled pteropods in peril: Assessing vulnerability in a high CO_2 ocean. *Earth-Science Reviews*, *169*(April), 132–145. https://doi.org/10.1016/j.earscirev.2017.04.005

Manno, C., Peck, V. L., & Tarling, G. A. (2016). Pteropod eggs released at high pCO_2 lack resilience to ocean acidification. *Scientific Reports*, *6*(November 2015), 1–10. https://doi.org/10.1038/srep25752

Manno, C., Sandrini, S., Tositti, L., & Accornero, A. (2007). First stages of degradation of Limacina helicina shells observed above the aragonite chemical lysocline in Terra Nova Bay (Antarctica). *Journal of Marine Systems*, *68*(1–2), 91–102. https://doi.org/10.1016/j.jmarsys.2006.11.002

McGowan, J. A. (1967). Distributional atlas of pelagic molluscs in the California Current region. *CalCOFI Atlas*, *6*.

Meilland, J., Fabri-Ruiz, S., Koubbi, P., Monaco, C. L., Cotte, C., Hosie, G. W., et al. (2016). Planktonic foraminiferal biogeography in the Indian sector of the Southern Ocean: Contribution from CPR data. *Deep-Sea Research Part I Oceanographic Research Papers*, *110*, 75–89. https://doi.org/10.1016/j.dsr.2015.12.014

Mekkes, L., Renema, W., Bednaršek, N., Alin, S. R., Feely, R. A., Huisman, J., et al. (2021). Pteropods make thinner shells in the upwelling region of the California Current Ecosystem. *Scientific Reports*, *11*(1), 1–11. https://doi.org/10.1038/s41598-021-81131-9

Mekkes, L., Sepúlveda-Rodríguez, G., Bielkinaitė, G., Wall-Palmer, D., Brummer, G. J. A., Dämmer, L. K., et al. (2021). Effects of ocean acidification on calcification of the sub-Antarctic pteropod Limacina retroversa. *Frontiers in Marine Science*, *8*(March), 1–12. https://doi.org/10.3389/fmars.2021.581432

Merow, C., Smith, M. J., Edwards, T. C., Guisan, A., Mcmahon, S. M., Normand, S., et al. (2014). What do we gain from simplicity versus complexity in species distribution models? *Ecography*, *37*(12), 1267–1281. https://doi.org/10.1111/ecog.00845

Michaels, A. F., Caron, D. A., Swanberg, N. R., & Howse, F. A. (1995). Primary productivity by symbiont-bearing planktonic sarcodines (Acantharia, Radiolaria, Foraminifera) in surface waters near Bermuda. *Journal of Plankton Research*, *17*(1), 103–129. https://doi.org/10.1093/plankt/17.1.103

Miloslavić, M., Lučić, D., Gangai, B., & Onofri, I. (2014). Mesh size effects on mesozooplankton community structure in a semi-enclosed coastal area and surrounding sea (South Adriatic Sea). *Marine Ecology*, *35*(4), 445–455. https://doi.org/10.1111/maec.12101

Morard, R., Darling, K. F., Mahé, F., Audic, S., Ujiié, Y., Weiner, A. K. M., et al. (2015). PFR2: A curated database of planktonic foraminifera 18S ribosomal DNA as a resource for studies of plankton ecology, biogeography and evolution. *Molecular Ecology Resources*, *15*(6), 1472–1485. https://doi.org/10.1111/1755-0998.12410

Moriarty, R., Buitenhuis, E. T., Le Quéré, C., & Gosselin, M. P. (2013). Distribution of known macrozooplankton abundance and biomass in the global ocean. *Earth System Science Data*, *5*(2), 241–257. https://doi.org/10.5194/essd-5-241-2013

Moy, A. D., Howard, W. R., Bray, S. G., & Trull, T. W. (2009). Reduced calcification in modern Southern Ocean planktonic foraminifera. *Nature Geoscience*, *2*(4), 276–280. https://doi.org/10.1038/ngeo460

Mucci, A. (1983). The solubility of calcite and aragonite in seawater at various salinities, temperatures, and one atmosphere total pressure. *American Journal of Science*, *283*(7), 780–799. https://doi.org/10.2475/ajs.283.7.780

Myers, T. D. (1968). *Horizontal and vertical distribution of thecosomatous pteropods off Cape Hatteras*. Duke University.

Naidu, P. D., & Malmgren, B. A. (1996). A high-resolution record of late quaternary upwelling along the Oman Margin, Arabian Sea based on planktonic foraminifera. *Paleoceanography*, *11*(1), 129–140. https://doi.org/10.1029/95pa03198

NASA OB.DAAC. (2018a). Sea-viewing Wide Field-of-view Sensor (SeaWiFS) chlorophyll data. https://doi.org/10.5067/ORBVIEW-2/SEAWIFS/L3M/CHL/2018

NASA OB.DAAC. (2018b). Sea-viewing Wide Field-of-view Sensor (SeaWiFS) downwelling diffuse attenuation coefficient data. https://doi.org/10.5067/ORBVIEW-2/SEAWIFS/L3M/KD/2018

NASA OB.DAAC. (2018c). Sea-viewing Wide Field-of-view Sensor (SeaWiFS) Euphotic depth data. https://doi.org/10.5067/ORBVIEW-2/SEAWIFS/L3M/ZLEE/2018

NASA OB.DAAC. (2018d). Sea-viewing Wide Field-of-view Sensor (SeaWiFS) Garver-Siegel-Maritorena (GSM) model data. https://doi.org/10.5067/ORBVIEW-2/SEAWIFS/L3M/GSM/2018

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I - A discussion of principles. *Journal of Hydrology*, *10*(3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society: Series A*, *135*(3), 370–384. https://doi.org/10.2307/2344614

O'Brien, C. J. (2015). *Global-scale distributions of marine haptophyte phytoplankton (Unpublished doctoral dissertation)*. ETH Zurich.

O'Brien, C. J., Vogt, M., & Gruber, N. (2016). Global coccolithophore diversity: Drivers and future change. *Progress in Oceanography*, *140*, 27–42. https://doi.org/10.1016/j.pocean.2015.10.003

O'Brien, T. D. (2010). COPEPOD, a global plankton database: A review of the 2010 database contents, processing methods, and access interface. Retrieved from https://repository.library.noaa.gov/view/noaa/5040

Ohman, M. D., Lavaniegos, B. E., & Townsend, A. W. (2009). Multi-decadal variations in calcareous holozooplankton in the California current system: Thecosome pteropods, heteropods, and foraminifera. *Geophysical Research Letters*, *36*(18), 2–6. https://doi.org/10.1029/2009GL039901
Orr, J. C., Fabry, V. J., Aumont, O., Bopp, L., Doney, S. C., Feely, R. A., et al. (2005). Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms. *Nature*, *437*(7059), 681–686. https://doi.org/10.1038/nature04095
Osborne, E. B., Thunell, R. C., Gruber, N., Feely, R. A., & Benitez-Nelson, C. R. (2020). Decadal variability in twentieth-century ocean acidification in the California current ecosystem. *Nature Geoscience*, *13*(1), 43–49. https://doi.org/10.1038/s41561-019-0499-z
Peijnenburg, K. T. C. A., Janssen, A. W., Wall-Palmer, D., Goetze, E., Maas, A. E., Todd, J. A., & Marlétaz, F. (2020). The origin and diversification of pteropods precede past perturbations in the Earth's carbon cycle. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(41), 25609–25617. https://doi.org/10.1073/pnas.1920918117
Pesant, S., Not, F., Picheral, M., Kandels-Lewis, S., Le Bescot, N., Gorsky, G., et al. (2015). Open science resources for the discovery and analysis of Tara Oceans data. *Scientific Data*, *2*(1), 150023. https://doi.org/10.1038/sdata.2015.23
Pinkerton, M. H., Décima, M., Kitchener, J. A., Takahashi, K. T., Robinson, K. V., Stewart, R., & Hosie, G. W. (2020). Zooplankton in the Southern Ocean from the continuous plankton recorder: Distributions and long-term change. *Deep-Sea Research Part I Oceanographic Research Papers*, *162*(May 2019), 103303. https://doi.org/10.1016/j.dsr.2020.103303
Pinkerton, M. H., Smith, A. N., Raymond, B., Hosie, G. W., Sharp, B., Leathwick, J. R., & Bradford-Grieve, J. M. (2010). Spatial and seasonal distribution of adult Oithona similis in the Southern Ocean: Predictions using boosted regression trees. *Deep-Sea Research Part I Oceanographic Research Papers*, *57*(4), 469–485. https://doi.org/10.1016/j.dsr.2009.12.010
Qiao, H., Feng, X., Escobar, L. E., Peterson, A. T., Soberón, J., Zhu, G., & Papeş, M. (2019). An evaluation of transferability of ecological niche models. *Ecography*, *42*(3), 521–534. https://doi.org/10.1111/ecog.03986
Rembauville, M., Meilland, J., Ziveri, P., Schiebel, R., Blain, S., & Salter, I. (2016). Planktic foraminifer and coccolith contribution to carbonate export fluxes over the central Kerguelen Plateau. *Deep-Sea Research Part I Oceanographic Research Papers*, *111*, 91–101. https://doi.org/10.1016/j.dsr.2016.02.017
Rhumbler, L. (1911). *Die foraminiferen (Thalamophoren) der Plankton expedition, Pt. 1, die Allgemeinen Organizationsverhaltnisse der Foraminifera* (p. 331). Lipsius & Tischer.
Richardson, A. J., Walne, A. W., John, A. W., Jonas, T. D., Lindley, J. A., Sims, D. W., et al. (2006). Using continuous plankton recorder data. *Progress in Oceanography*, *68*(1), 27–74. https://doi.org/10.1016/j.pocean.2005.09.011
Righetti, D., Vogt, M., Gruber, N., Psomas, A., & Zimmermann, N. E. (2019). Global pattern of phytoplankton diversity driven by temperature and environmental variability. *Science Advances*, *5*(5), 1–11. https://doi.org/10.1126/sciadv.aau6253
Rillo, M. C., Woolley, S., & Hillebrand, H. (2022). Drivers of global pre-industrial patterns of species turnover in planktonic foraminifera. *Ecography*, *2022*(1), 1–11. https://doi.org/10.1111/ecog.05892
Romagnan, J.-B., Legendre, L., Guidi, L., Jamet, J.-L., Jamet, D., Mousseau, L., et al. (2015). Comprehensive model of annual plankton succession based on the whole-plankton time series approach. *PLoS One*, *10*(3), e0119219. https://doi.org/10.1371/journal.pone.0119219
Rost, B., & Riebesell, U. (2004). Coccolithophores and the biological pump: Responses to environmental changes. In H. R. Thierstein & J. R. Young (Eds.), *Coccolithophores* (pp. 99–125). Springer. https://doi.org/10.1007/978-3-662-06278-4_5
Rothschild, B. J., & Osborn, T. R. (1988). Small-scale turbulence and plankton contact rates. *Journal of Plankton Research*, *10*(3), 465–474. https://doi.org/10.1093/plankt/10.3.465
Sallée, J. B., Pellichero, V., Akhoudas, C., Pauthenet, E., Vignes, L., Schmidtko, S., et al. (2021). Summertime increases in upper-ocean stratification and mixed-layer depth. *Nature*, *591*(7851), 592–598. https://doi.org/10.1038/s41586-021-03303-x
Sarmiento, J., & Gruber, N. (2006a). Calcium carbonate cycle. In *Ocean biogeochemical dynamics* (pp. 359–391). Princeton University Press.
Sarmiento, J., & Gruber, N. (2006b). Carbon cycle. In *Ocean biogeochemical dynamics* (pp. 318–358). Princeton University Press.
Schiebel, R. (2002). Planktic foraminiferal sedimentation and the marine calcite budget. *Global Biogeochemical Cycles*, *16*(4), 3–1. https://doi.org/10.1029/2001gb001459
Schiebel, R., Barker, S., Lendt, R., Thomas, H., & Bollmann, J. (2007). Planktic foraminiferal dissolution in the twilight zone. *Deep Sea Research Part II: Topical Studies in Oceanography*, *54*(5–7), 676–686. https://doi.org/10.1016/j.dsr2.2007.01.009
Schiebel, R., & Hemleben, C. (2000). Interannual variability of planktic foraminiferal populations and test flux in the eastern North Atlantic Ocean (JGOFS). *Deep-Sea Research Part II Topical Studies in Oceanography*, *47*(9–11), 1809–1852. https://doi.org/10.1016/S0967-0645(00)00008-4
Schiebel, R., & Hemleben, C. (2005). Modern planktic foraminifera. *Paläontologische Zeitschrift*, *79*(1), 135–148. https://doi.org/10.1007/bf03021758
Schiebel, R., & Hemleben, C. (2017). *Planktic foraminifers in the modern ocean* (2nd ed.). Springer. https://doi.org/10.1007/978-3-662-50297-6
Schiebel, R., Hiller, B., & Hemleben, C. (1995). Impacts of storms on recent planktic foraminiferal test production and $CaCO_3$ flux in the North Atlantic at 47°N, 20°W (JGOFS). *Marine Micropaleontology*, *26*(1–4), 115–129. https://doi.org/10.1016/0377-8398(95)00035-6
Schiebel, R., & Movellan, A. (2012). First-order estimate of the planktic foraminifer biomass in the modern ocean. *Earth System Science Data*, *4*(1), 75–89. https://doi.org/10.5194/essd-4-75-2012
Schiebel, R., Waniek, J., Bork, M., & Hemleben, C. (2001). Planktic foraminiferal production stimulated by chlorophyll redistribution and entrainment of nutrients. *Deep-Sea Research Part I Oceanographic Research Papers*, *48*(3), 721–740. https://doi.org/10.1016/S0967-0637(00)00065-0
Schiebel, R., Waniek, J., Zeltner, A., & Alves, M. (2002). Impact of the Azores Front on the distribution of planktic foraminifers, shelled gastropods, and coccolithophorids. *Deep-Sea Research Part II Topical Studies in Oceanography*, *49*(19), 4035–4050. https://doi.org/10.1016/S0967-0645(02)00141-8
Schiebel, R., Zeltner, A., Treppke, U. F., Waniek, J. J., Bollmann, J., Rixen, T., & Hemleben, C. (2004). Distribution of diatoms, coccolithophores and planktic foraminifers along a trophic gradient during SW monsoon in the Arabian Sea. *Marine Micropaleontology*, *51*(3–4), 345–371. https://doi.org/10.1016/j.marmicro.2004.02.001
Schmidt, D. N., Renaud, S., Bollmann, J., Schiebel, R., & Thierstein, H. R. (2004). Size distribution of Holocene planktic foraminifer assemblages: Biogeography, ecology and adaptation. *Marine Micropaleontology*, *50*(3–4), 319–338. https://doi.org/10.1016/S0377-8398(03)00098-7
Schmidt, G. A., & Mulitza, S. (2002). Global calibration of ecological models for planktic foraminifera from coretop carbonate oxygen-18. *Marine Micropaleontology*, *44*(3–4), 125–140. https://doi.org/10.1016/S0377-8398(01)00041-X
Schmidt-Nielsen, K. (1997). *Animal physiology: Adaptation and environment*. Cambridge university press.
Seuront, L., Schmitt, F., & Lagadeuc, Y. (2001). Turbulence intermittency, small-scale phytoplankton patchiness and encounter rates in plankton: Where do we go from here? *Deep-Sea Research Part I Oceanographic Research Papers*, *48*(5), 1199–1215. https://doi.org/10.1016/S0967-0637(00)00089-3
Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3–4), 591–611. https://doi.org/10.1093/biomet/52.3-4.591

Siccha, M., Schiebel, R., Schmidt, S., & Howa, H. (2012). Short-term and small-scale variability in planktic foraminifera test flux in the Bay of Biscay. *Deep Sea Research Part I: Oceanographic Research Papers*, *64*, 146–156. https://doi.org/10.1016/j.dsr.2012.02.004

Skjoldal, H. R., Wiebe, P. H., Postel, L., Knutsen, T., Kaartvedt, S., & Sameoto, D. D. (2013). Intercomparison of zooplankton (net) sampling systems: Results from the ICES/GLOBEC sea-going workshop. *Progress in Oceanography*, *108*, 1–42. https://doi.org/10.1016/j.pocean.2012.10.006

Soviadan, Y. D., Benedetti, F., Brandão, M. C., Ayata, S. D., Irisson, J. O., Jamet, J. L., et al. (2022). Patterns of mesozooplankton community composition and vertical fluxes in the global ocean. *Progress in Oceanography*, *200*(September 2021), 102717. https://doi.org/10.1016/j.pocean.2021.102717

Spiess, A.-N., & Neumeyer, N. (2010). An evaluation of R2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: A Monte Carlo approach. *BMC Pharmacology*, *10*(1), 6. https://doi.org/10.1186/1471-2210-10-6

Spindler, M., Hemleben, C., Salomons, J. B., & Smit, L. P. (1984). Feeding behavior of some planktonic foraminifers in laboratory cultures. *Journal of Foraminiferal Research*, *14*(4), 237–249. https://doi.org/10.2113/gsjfr.14.4.237

Steinhardt, J., Cléroux, C., Ullgren, J., de Nooijer, L., Durgadoo, J. V., Brummer, G.-J., & Reichart, G.-J. (2014). Anti-cyclonic eddy imprint on calcite geochemistry of several planktonic foraminiferal species in the Mozambique Channel. *Marine Micropaleontology*, *113*, 20–33. https://doi.org/10.1016/j.marmicro.2014.09.001

Stepien, J. C. (1980). The occurrence of chaetognaths, pteropods and euphausiids in relation to deep flow reversals in the Straits of Florida. Deep Sea Research Part A. *Oceanographic Research Papers*, *27*(12), 987–1011. https://doi.org/10.1016/0198-0149(80)90061-8

Strömberg, K. H., Smyth, T. J., Allen, J. I., Pitois, S., & O'Brien, T. D. (2009). Estimation of global zooplankton biomass from satellite ocean colour. *Journal of Marine Systems*, *78*(1), 18–27. https://doi.org/10.1016/j.jmarsys.2009.02.004

Student (1908). The probable error of a mean. *Biometrika*, *6*(1), 1–25. https://doi.org/10.2307/2331554

Takagi, H., Kimoto, K., Fujiki, T., Saito, H., Schmidt, C., Kucera, M., & Moriya, K. (2019). Characterizing photosymbiosis in modern planktonic foraminifera. *Biogeosciences*, *16*(17), 3377–3396. https://doi.org/10.5194/bg-16-3377-2019

Takahashi, K., & Bé, A. W. (1984). Planktonic foraminifera: Factors controlling sinking speeds. *Deep-Sea Research, Part A: Oceanographic Research Papers*, *31*(12), 1477–1500. https://doi.org/10.1016/0198-0149(84)90083-9

Thibodeau, P. S., Steinberg, D. K., Stammerjohn, S. E., & Hauri, C. (2019). Environmental controls on pteropod biogeography along the Western Antarctic Peninsula. *Limnology & Oceanography*, *64*(S1), S240–S256. https://doi.org/10.1002/lno.11041

Thuiller, W., Guéguen, M., Renaud, J., Karger, D. N., & Zimmermann, N. E. (2019). Uncertainty in ensembles of global biodiversity scenarios. *Nature Communications*, *10*(1), 1–9. https://doi.org/10.1038/s41467-019-09519-w

Thuiller, W., Pollock, L. J., Gueguen, M., & Münkemüller, T. (2015). From species distributions to meta-communities. *Ecology Letters*, *18*(12), 1321–1328. https://doi.org/10.1111/ele.12526

Tittensor, D. P., Novaglio, C., Harrison, C. S., Heneghan, R. F., Barrier, N., Bianchi, D., et al. (2021). Next-generation ensemble projections reveal higher climate risks for marine ecosystems. *Nature Climate Change*, *11*(11), 973–981. https://doi.org/10.1038/s41558-021-01173-9

Tseng, L. C., Dahms, H. U., Hung, J. J., Chen, Q. C., & Hwang, J. S. (2011). Can different mesh sizes affect the results of copepod community studies? *Journal of Experimental Marine Biology and Ecology*, *398*(1–2), 47–55. https://doi.org/10.1016/j.jembe.2010.12.007

Tsurumi, M., Mackas, D. L., Whitney, F. A., DiBacco, C., Galbraith, M. D., & Wong, C. S. (2005). Pteropods, eddies, carbon flux, and climate variability in the Alaska Gyre. *Deep Sea Research Part II: Topical Studies in Oceanography*, *52*(7–8), 1037–1053. https://doi.org/10.1016/j.dsr2.2005.02.005

Turner, J. T. (2015). Zooplankton fecal pellets, marine snow, phytodetritus and the ocean's biological pump. *Progress in Oceanography*, *130*, 205–248. https://doi.org/10.1016/j.pocean.2014.08.005

Vereshchaka, A. L., Lunina, A. A., & Mikaelyan, A. S. (2022). Surface chlorophyll concentration as a mesoplankton biomass assessment tool in the Southern Ocean region. *Global Ecology and Biogeography*, *31*(3), 405–424. https://doi.org/10.1111/geb.13435

Waldock, C., Stuart-Smith, R. D., Albouy, C., Cheung, W. W. L., Edgar, G. J., Mouillot, D., et al. (2022). A quantitative review of abundance-based species distribution models. *Ecography*, *2022*(1). https://doi.org/10.1111/ecog.05694

Wall-Palmer, D., Smart, C. W., Kirby, R., Hart, M. B., Peijnenburg, K. T. C. A., & Janssen, A. W. (2016). A review of the ecology, palaeontology and distribution of Atlantid heteropods (Caenogastropoda: Pterotracheoidea: Atlantidae). *Journal of Molluscan Studies*, *82*(2), 221–234. https://doi.org/10.1093/mollus/eyv063

Wang, K., Hunt, B. P., Liang, C., Pauly, D., & Pakhomov, E. A. (2017). Reassessment of the life cycle of the pteropod *Limacina helicina* from a high resolution interannual time series in the temperate North Pacific. *ICES Journal of Marine Science*, *74*(7), 1906–1920. https://doi.org/10.1093/icesjms/fsx014

Weinfurt, K. P. (1995). Multivariate analysis of variance. In G. L. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 245–276). American Psychological Association.

Weinkauf, M. F., Kunze, J. G., Waniek, J. J., & Kučera, M. (2016). Seasonal variation in shell calcification of planktonic foraminifera in the ne atlantic reveals species-specific response to temperature, productivity, and optimum growth conditions. *PLoS One*, *11*(2), 1–33. https://doi.org/10.1371/journal.pone.0148363

Wells, F. E. (1973). Effects of mesh size on estimation of population densities of tropical euthecosomatous pteropods. *Marine Biology*, *20*(4), 347–350. https://doi.org/10.1007/BF00354276

Wilson, R. W., Millero, F. J., Taylor, J. R., Walsh, P. J., Christensen, V., Jennings, S., & Grosell, M. (2009). Contribution of fish to the marine inorganic carbon cycle. *Science*, *323*(5912), 359–362. https://doi.org/10.1126/science.1157972

Wormuth, J. H. (1981). Vertical distributions and diel migrations of Euthecosomata in the northwest Sargasso Sea. *Deep Sea Research Part A. Oceanographic Research Papers*, *28*(12), 1493–1515. https://doi.org/10.1016/0198-0149(81)90094-7

Zamelczyk, K., Fransson, A., Chierici, M., Jones, E., Meilland, J., Anglada-Ortiz, G., & Lødemel, H. H. (2021). Distribution and abundances of planktic foraminifera and shelled pteropods during the polar night in the sea-ice covered Northern Barents Sea. *Frontiers in Marine Science*, *8*(October), 1–19. https://doi.org/10.3389/fmars.2021.644094

Ziveri, P., Gray, W. R., Anglada-Ortiz, G., Manno, C., Grelaud, M., Incarbona, A., et al. (2023). Pelagic calcium carbonate production and shallow dissolution in the North Pacific Ocean. *Nature Communications*, *14*(1), 805. https://doi.org/10.1038/s41467-023-36177-w

Zurell, D., Franklin, J., König, C., Bouchet, P. J., Dormann, C. F., Elith, J., et al. (2020). A standard protocol for reporting species distribution models. *Ecography*, *43*(9), 1261–1277. https://doi.org/10.1111/ecog.04960

Zweng, M. M., Seidov, D., Boyer, T., Locarnini, M., Garcia, H., Mishonov, A., et al. (2019). *World Ocean Atlas 2018* (Vol. 2). Salinity (Tech. Rep.). NOAA. Retrieved from https://www.ncei.noaa.gov/access/world-ocean-atlas-2018/bin/woa18.pl?parameter=s