Environmental Controls of Size Distribution of Modern Planktonic Foraminifera in the Tropical Indian Ocean

Michael B Adebayo, Clara T Bolton, Ross Marchant, Franck Bassinot, Sandrine Conrod, Thibault de Garidel-thoron

▶ **To cite this version:**

Michael B Adebayo, Clara T Bolton, Ross Marchant, Franck Bassinot, Sandrine Conrod, et al.. Environmental Controls of Size Distribution of Modern Planktonic Foraminifera in the Tropical Indian Ocean. Geochemistry, Geophysics, Geosystems, 2023, 24 (4), 10.1029/2022gc010586 . hal-04087217

HAL Id: hal-04087217
https://hal.science/hal-04087217

Submitted on 3 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Check for updates

Geochemistry, Geophysics, Geosystems®

AGU ADVANCING EARTH AND SPACE SCIENCE

RESEARCH ARTICLE
10.1029/2022GC010586

Environmental Controls of Size Distribution of Modern Planktonic Foraminifera in the Tropical Indian Ocean

Michael B. Adebayo[1], Clara T. Bolton[1], Ross Marchant[2], Franck Bassinot[3], Sandrine Conrod[1], and Thibault de Garidel-Thoron[1]

[1]Centre Européen de Recherche et d'Enseignement des Géosciences de l'Environnement (CEREGE), CNRS, IRD, Collège de France, INRAE, Aix-Marseille Université, Aix-en-Provence, France, [2]School of Electrical Engineering & Robotics, Queensland University of Technology, Brisbane, QLD, Australia, [3]Laboratoire des Sciences du Climat et de l'Environnement (IPSL), CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France

Key Points:

- Optimum size-hypothesis holds true in planktonic foraminifera if one considers the main parameters driving each species' size distribution
- Size variations in planktonic foraminifera are linked to species' niches and diversity does not increase with productivity
- Within-species size is driven by CO_3^{2-} concentration, temperature, and salinity; assemblage size by CO_3^{2-} concentration and temperature

Supporting Information:

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

Correspondence to:

M. B. Adebayo,
adebayo@cerege.fr;
adebayomichael4@gmail.com

Citation:

Adebayo, M. B., Bolton, C. T., Marchant, R., Bassinot, F., Conrod, S., & de Garidel-Thoron, T. (2023). Environmental controls of size distribution of modern planktonic foraminifera in the tropical Indian Ocean. *Geochemistry, Geophysics, Geosystems*, *24*, e2022GC010586. https://doi.org/10.1029/2022GC010586

Received 22 JUN 2022
Accepted 15 JAN 2023

Author Contributions:

Conceptualization: Michael B. Adebayo, Clara T. Bolton, Thibault de Garidel-Thoron
Data curation: Michael B. Adebayo
Formal analysis: Michael B. Adebayo
Funding acquisition: Michael B. Adebayo, Clara T. Bolton, Thibault de Garidel-Thoron
Investigation: Michael B. Adebayo, Clara T. Bolton, Thibault de Garidel-Thoron

© 2023 The Authors.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Abstract Paleoceanographic studies often rely on abundance changes in microfossil species, with little consideration for characteristics such as organism size, which may also be related to environmental changes. Using a tropical Indian Ocean (TIO) core-top data set, we test the Optimum size-hypothesis (OSH), investigating whether relative abundance or environmental variables are better descriptors of planktonic foraminifera species' optimum conditions. We also investigate the environmental drivers of whole-assemblage planktonic foraminiferal test size variation in the TIO. We use an automated imaging and sorting system (MiSo) to identify planktonic foraminiferal species, analyze their morphology, and quantify fragmentation rate using machine learning techniques. Machine model accuracy is confirmed by comparison with human classifiers (97% accuracy). Data for 33 environmental parameters were extracted from modern databases and, through exploratory factor analysis and regression models, we explore relationships between planktonic foraminiferal size and oceanographic parameters in the TIO. Results show that the size frequency distribution of most planktonic foraminifera species is unimodal, with some larger species showing multimodal distributions. Assemblage $size_{95/5}$ (95th percentile size) increases with increasing species diversity, and this is attributed to vertical niche separation induced by thermal stratification. Our test for the OSH reveals that relative abundance is not a good predictor of species' optima and within-species $size_{95/5}$ response to environmental parameters is species-specific, with parameters related to carbonate ion concentration, temperature, and salinity being primary drivers. At the species and assemblage levels, our analyses indicate that carbonate ion concentration and temperature play important roles in determining size trends in TIO planktonic foraminifera.

Plain Language Summary In core-top samples from the tropical Indian Ocean (TIO), we investigate the optimum size-hypothesis, testing whether species' relative abundance or environmental parameter(s) are better descriptors of planktonic foraminifera species' optimum conditions. Further, we investigate the main environmental drivers of size variations in planktonic foraminifera at the assemblage-level, given that temperature has been reported to primarily drive assemblage size trends. We use a state-of-the-art machine (MiSo) to automatically identify planktonic foraminiferal species, analyze their size, and quantify fragmentation using machine learning techniques. When compared to identification carried out by human experts across 21 species, the machine classified the species accurately 97% of the time. The MiSo-generated size data was similar to that by other researchers. The frequency distributions of the species' size spectra show that most species have distributions that form bell-shaped curves. As species diversity increased, so did the assemblage size (95th percentile size); we attribute this observation to the effect of temperature-dependent niche separation. We find that, in the TIO, environmental parameters are better descriptors of optimum conditions in planktonic foraminifera than relative abundance. Our results also reveal that size variation at the species and assemblage levels is mostly driven by ambient carbonate chemistry and temperature.

1. Introduction

Increasing anthropogenic emissions of carbon dioxide (CO_2) and other greenhouse gases are drastically affecting the physico-chemical characteristics of the ocean, resulting in higher temperatures, lower pH, stronger stratification, and reduced biological productivity (Bindoff et al., 2019; Caldeira & Wickett, 2003; Gao et al., 2012; Stips et al., 2016; Zeebe & Wolf-Gladrow, 2001). Over the last century, the surface ocean has warmed by on average 0.9°C while upper-ocean stratification has increased by 4.9% ± 1.5% between 1970 and 2018 (Fox-Kemper

Methodology: Michael B. Adebayo, Clara T. Bolton, Ross Marchant, Franck Bassinot, Sandrine Conrod, Thibault de Garidel-Thoron
Project Administration: Michael B. Adebayo, Clara T. Bolton, Thibault de Garidel-Thoron
Resources: Franck Bassinot, Sandrine Conrod
Software: Michael B. Adebayo, Ross Marchant
Supervision: Clara T. Bolton, Thibault de Garidel-Thoron
Visualization: Michael B. Adebayo
Writing – original draft: Michael B. Adebayo
Writing – review & editing: Michael B. Adebayo, Clara T. Bolton, Ross Marchant, Franck Bassinot, Thibault de Garidel-Thoron

et al., 2021). Under more stratified conditions, upper ocean vertical mixing decreases, leading to reduced nutrient inputs and primary productivity (PP) (Gruber, 2011; He & Mahadevan, 2021). These surface ocean changes impact calcifying marine plankton, including planktonic foraminifera, which are major contributors to the global carbonate pump driven by the production of $CaCO_3$ in surface waters and its subsequent export to deep waters and, ultimately, sediments (Berner & Raiswell, 1983; Fabry et al., 2008; Henehan et al., 2017; Manno et al., 2012; Moy et al., 2009; Orr et al., 2005; Schiebel, 2002). Planktonic foraminifera are unicellular marine protozoans that precipitate calcite tests, accounting for ~25% of the total calcite exported to the deep ocean (Buitenhuis et al., 2019; Schiebel, 2002; Schiebel et al., 2007). Their growth, calcification, and reproduction processes are directly linked to ambient environmental conditions (Schiebel & Hemleben, 2017; Weinkauf et al., 2013, 2022). Fossil planktonic foraminiferal assemblages and the geochemical signatures of tests are routinely used to reconstruct past upper ocean conditions, including temperature and productivity (e.g., Cayre et al., 1999; Kucera, 2007; Kucera et al., 2005; Tachikawa et al., 2008). Another key variable that we can extract from the planktonic foraminiferal fossil record is test size, yet comparatively little is known about how it varies in response to the environment, despite existing studies showing that test size distribution can provide valuable insight into present and past oceanographic conditions (Moller et al., 2013; Schmidt, Renaud, et al., 2004).

The size of an organism is a functional trait that scales with ecological parameters such as population growth rate, competition, and reproduction (Arendt, 2007; Millien et al., 2006), and influences its metabolic rate, its ability to coexist with other organisms, and its response to changes in ambient conditions (Brown et al., 2007; Hart et al., 2016; Violle et al., 2012). Different rules have been proposed to explain the relationship between an organism's size and temperature. First, Bergmann's rule, originally proposed for endotherms in the context of heat conservation, states that larger organisms tend to inhabit colder environments (Bergmann, 1847). This rule has since been recast for groups including marine crustaceans, birds, and some ectotherms (Ashton & Feldman, 2003; Fan et al., 2019; Ollala-Tàrraga et al., 2006). Second, the temperature–size rule states that the body size of ectothermic organisms (including protozoans, copepods, and fish) generally decreases with increasing temperature (Atkinson, 1994). Whilst neither of these rules have been applied specifically to planktonic foraminifera, a positive temperature–size relationship on global biogeographical and evolutionary scales has been documented (Schmidt, Renaud, et al., 2004; Schmidt, Thierstein, et al., 2004); that is, modern planktonic foraminifera are larger in warm tropical waters than in colder high latitude waters, and Paleocene (greenhouse climate) planktonic foraminifera are larger than Pleistocene (icehouse climate) planktonic foraminifera.

Another theory suggests that size variations in planktonic foraminifera follow the optimum size-hypothesis (OSH), which states that populations are largest at their ecological optimum (Hecht, 1976). Testing the OSH by examining the relationship between a species' size and a measure of its optimum conditions is challenging, and multiple factors can contribute to the diversity of results obtained (e.g., Beer et al., 2010; Hecht, 1976; Kahn, 1981; Kennett, 1976; Malmgren & Kennett, 1976, 1977; Rillo et al., 2020; Schmidt, Renaud, et al., 2004; Schmidt, Thierstein, et al., 2004; Weinkauf et al., 2016; Zarkogiannis et al., 2020). These factors include differences in the population unit investigated (single or multiple populations), in the methodology or statistics adopted, in the choice of metric used to define size in a population (e.g., mean, maximum, $size_{95/5}$ (95th percentile size: the size separating the top 5% from the remaining 95% size within a population or assemblage), shell weight, size normalized weight), and lastly how optimum growth conditions are defined (maximum absolute or relative abundance, environmental preferences). In addition, the range of environmental gradients covered within the samples of a study may also influence the outcome with respect to the OSH (Kontakiotis et al., 2021; Weinkauf et al., 2016). Aside from the global study of Schmidt, Renaud, et al. (2004), most studies that found support for the OSH investigated single or very few species, and defined optimum conditions based on species' absolute or relative abundance (Hecht, 1976; Malmgren & Kennett, 1976; Moller et al., 2013). Schmidt, Renaud, et al. (2004) investigated the OSH using the correlation between temperature at maximum $size_{95/5}$ and temperature at maximum relative abundance, under the assumption that convergence between the two temperatures indicated an ecological optimum. While this is plausible, this approach uses only a small proportion of the available data, and because different species have different thermal niches, a 1:1 temperature relationship as observed in their multispecies study will still occur even if there is no relationship between size and abundance in individual species (Rillo et al., 2020). In Weinkauf et al. (2016) and Rillo et al. (2020), correlations between size (mean size and $size_{95/5}$) and environmental parameters such as sea surface temperature (SST), net PP, and chlorophyll a showed no support for the OSH. Therefore, there is a need to evaluate whether optimum conditions in planktonic foraminifera are better inferred from species' relative abundances or from their preferred environmental conditions using robust statistical methods.

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Within the context of planktonic foraminifera size response to abiotic factors, it can be difficult to define which parameter serves as the good measure of optimum conditions for each species, since different species have varying ecological preferences (Schiebel & Hemleben, 2017). Whilst SST has been implicated as a primary driver of size variation in planktonic foraminifera (Schmidt, Renaud, et al., 2004), laboratory experiments have shown that size variations can also be linked to salinity, carbonate saturation, nutrients, and oxygen availability (Bé et al., 1981; Bijma et al., 1990; Caron et al., 1982; Hemleben et al., 1987; Kuroyanagi et al., 2013), making the SST–size relationship difficult to discern at the species level (Rillo et al., 2020). The lack of consensus on which physico-chemical parameter(s) exert the strongest controls on planktonic foraminiferal size may in part result from: (a) differential size responses to ambient conditions at the assemblage level (Schmidt, Renaud, et al., 2004) and the species level (Colombo & Cita, 1980; Hecht, 1974; Moller et al., 2013; Rillo et al., 2020) due to species-specific ecological adaptation and resource requirements; (b) investigation of few parameters ($n < 8$) rather than a broader range of variables that more closely represent species' living conditions (e.g., Rillo et al., 2020); (c) methodological differences arising from the use of different size metrics such as diameter, area, aspect ratio, and roundness (Brombacher et al., 2017); (d) differences in sample material, that is, whether data from modern samples (laboratory culture experiments, plankton tow, or sediment trap) or fossil materials (core-tops and down-core sediment samples) are used; and (e) local or regional effects (Moller et al., 2013). This lack of consensus warrants further study to understand which environmental variable(s) exert major controls on planktonic foraminiferal size distribution at the species and assemblage levels.

In this study, we quantified planktonic foraminiferal species compositions and analyzed their size variations in a suite of 82 core-top samples from the tropical Indian Ocean (TIO). We use state-of-the-art automated morphometric and identification methods to collect a large and robust data set, and our new high-throughput workflow based on machine-learning techniques is validated by comparison with human counts. Using statistical modeling, we test the OSH, investigating whether species' optimum conditions are best inferred from their relative abundances or from the environmental parameters that mostly drive variations in their sizes (33 parameters are tested). We investigated the environmental drivers of size variations in TIO planktonic foraminifera at the species and assemblage levels. We focus on the TIO because a significant amount of data on planktonic foraminifera distribution in this region already exists, and because this region contains highly contrasted oceanic environments in terms of physico-chemical characteristics (oxygenation, stratification, PP, carbonate chemistry, salinity). In the TIO region encompassing our samples, mean annual sea surface temperatures range from 19 to 30°C and average mixed layer (0–200 m) temperatures range from 12 to 22°C (Locarnini et al., 2018), thus the influence of temperature on size is unlikely to override the influence of other environmental factors.

2. Materials and Methods

2.1. Oceanographic Setting

The TIO, with its unique seasonally-reversing monsoons and transient cyclonic storms, is landlocked in the north, and represents an area of intense air–sea interaction (R. R. Rao et al., 1991; Schott et al., 2009). The northern TIO is divided into two parts by the Indian Peninsula: the Arabian Sea and the Bay of Bengal. Due to its proximity to the Persian Gulf and Red Sea, the Arabian Sea has an upper layer of highly saline waters and experiences more evaporation than precipitation (Levitus, 1982; L. V. G. Rao & Ram, 2005). The Bay of Bengal receives high freshwater inputs via river discharge from abundant precipitation during the southwest (summer) monsoon, resulting in a low-salinity upper layer and a stratification dominated by salinity gradients rather than temperature gradients (L. V. G. Rao & Ram, 2005; Wyrtki, 1971). In the Arabian Sea, wind-driven coastal upwelling in the west is strong during the southwest monsoon and vertical mixing is also enhanced during the northeast monsoon, promoting higher productivity (Barber et al., 2001). In the southwestern TIO, the Mozambique Channel is important for water exchanges between the Indian and Atlantic Oceans via the Agulhas Current (Schott et al., 2009; Ternon et al., 2014). During the summer monsoon from April to October, southeasterly winds from Australia induce Ekman pumping in the southeastern TIO, causing upwelling along the coasts of Sumatra and Java (Susanto et al., 2001). Within the TIO (31°N, 31°S, 20–120°E), mean annual SST ranges from 18°C (southern part; July–October) to >30°C (northern part; April–May) (Locarnini et al., 2018; L. V. G. Rao & Ram, 2005). Meanwhile, sea surface salinity ranges from 30 to 36 psu, dissolved oxygen concentration ranges from 189 to 255 μmol kg^{-1}, surface Chl-a concentration ranges from 0.1 to 0.5 mg m^{-3}, and carbonate ion content of surface waters ranges from 200 to 280 μmol kg^{-1}, with an average of >220 μmol/kg (Broecker & Sutherland, 2000; Garcia et al., 2018a; Lovenduski et al., 2015; L. V. G. Rao & Ram, 2005; Wyrtki, 1971; Zweng et al., 2018).

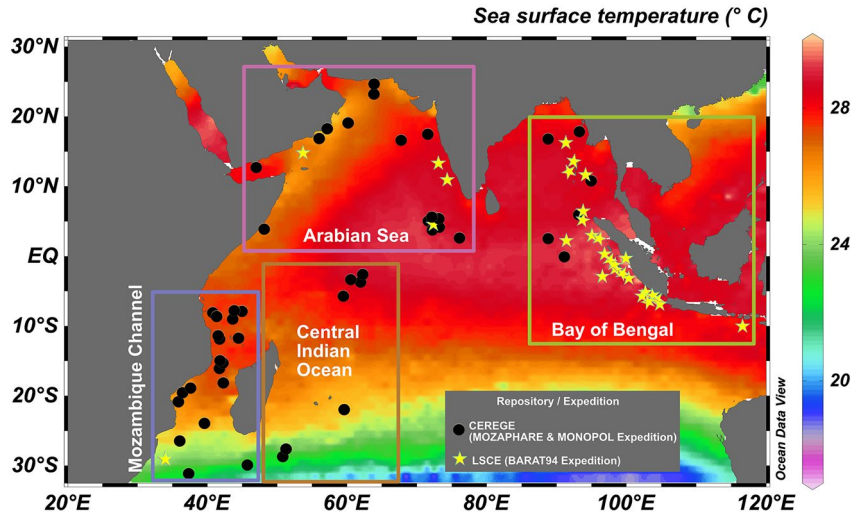

Figure 1. Map showing the location of the core-top samples used in this study as well as the laboratories where they are archived. Underlaid is the zonal sea surface temperature gradient of the tropical Indian Ocean. Shown also are the categories of the regions denoted in this study: Arabian Sea (pink box), Bay of Bengal (yellow box), central Indian Ocean (brown box), and Mozambique Channel (blue box).

2.2. Sample Material

We analyzed 82 core-top samples collected in the TIO (Figure 1), stored in core repositories at the Center Européen de Recherche et d'Enseignement de Géosciences de l'Environnement (CEREGE, France) and at the Laboratoire des Sciences du Climat et de l'Environnement (LSCE, France). Samples were taken during expeditions carried out on the RV *Marion Dufresne* and RV *Baruna Jaya I* research vessels. The sample set covers the Arabian Sea, Bay of Bengal (including offshore Java), Mozambique Channel, and central Indian Ocean (31° 30′S, 36° 45′E to 24° 50′N, 60° 36′E) and core-tops are from water depths between 364 and 4,425 m (mean depth 2,346 m). Surface sediment samples (top 2–3 cm) were dried at 50°C then washed over a 150 μm sieve, and the coarse fractions were oven-dried at 50°C. Using a SIO microsplitter, two splits were prepared per sample; one split served as an archive half while the other was used as the working half. The coarse fraction (>150 μm) for the working half of each sample was used for automated imaging on the Microfossil Sorter (MiSo, patent pending) at CEREGE. This automation takes multiple (generally 12) images of single foraminifera tests at several focal depths (vertical resolution of 11 μm). Processed, Z-stacked foraminifera images have a resolution of 1.1594 pixels per micron.

2.3. Environmental Data

The environmental data for the ocean overlying each site were retrieved from oceanographic databases using mean annual values from different sources as reported in Table 1. Environmental parameters were selected to capture a wide variety of variables including (but not limited to) parameters that have been suggested to influence planktonic foraminifera size and assemblage distribution: temperature, nutrient concentrations, salinity, oxygen concentrations, carbonate system parameters, and thermal gradients (Cayre et al., 1999; Clemens et al., 1996; Kuroyanagi et al., 2013; Moller et al., 2013; Schmidt, Renaud, et al., 2004). Since the average living depth of most planktonic foraminifera is within the upper 200 m of the water column (Lessa et al., 2020), most of the environmental data used in this work come from 0 to 200 m, so as to capture both surface and sub-surface conditions. The environmental data sets that we used are archived in the French Open Access database, SEANOE, and can be accessed at: https://doi.org/10.17882/86211 (Adebayo, Bolton, Marchant, Bassinot, Sandrine, et al., 2022). The spatial variability in surface ocean parameters in the TIO was analyzed using a covariance matrix principal component analysis (PCA) to identify the surface hydrographic settings of our study area. The paleontological statistical software Past (version 4.01, Hammer et al., 2001) was used for this analysis.

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1
List of Environmental Parameters Included in the Analysis for This Study

	Variable name/resolution	Unit	Description	Var_code	Source
1	Sea surface temperature/1°	°C	1955–2017 annual mean temperature at 0 m	Sst	World Ocean Atlas (WOA 18) (Locarnini et al., 2018)
2	Sea temperature/1°	°C	1955–2017 annual mean temperature at 10 m	T10	World Ocean Atlas (WOA 18) (Locarnini et al., 2018)
3	Sea temperature/1°	°C	1955–2017 annual mean at 50 m	T50	World Ocean Atlas (WOA 18) (Locarnini et al., 2018)
4	Sea temperature/1°	°C	1955–2017 annual mean temperature at 100 m	T100	World Ocean Atlas (WOA 18) (Locarnini et al., 2018)
5	Sea temperature/1°	°C	1955–2017 annual mean temperature at 200 m	T200	World Ocean Atlas (WOA 18) (Locarnini et al., 2018)
	Sea temperature/1°	°C	Δ Temperature between 0 and 200 m	ΔT	World Ocean Atlas (WOA 18) (Locarnini et al., 2018)
6	Sea temperature/1°	°C	1955–2017 Summer SST (July–September)	Sumsst	World Ocean Atlas (WOA 18) (Locarnini et al., 2018)
7	Sea Temperature/1°	°C	1955–2017 Winter SST (January–March)	Winsst	World Ocean Atlas (WOA 18) (Locarnini et al., 2018)
8	Sea temperature/1°	°C	Δ Temperature between Winter and Summer Temperatures	Winsum	World Ocean Atlas (WOA 18) (Locarnini et al., 2018)
9	Phosphate/1°	μmol/kg	1955–2017 annual mean phosphate concentration at 0 m	Phos	World Ocean Atlas (WOA 18) (Garcia et al., 2018a)
10	Phosphate/1°	μmol/kg	1955–2017 annual mean phosphate concentration at 200 m	Phos200	World Ocean Atlas (WOA 18) (Garcia et al., 2018a)
11	Phosphate/1°	μmol/kg	Δ Phosphate between 0 and 200 m	ΔPhos	World Ocean Atlas (WOA 18) (Garcia et al., 2018a)
12	Nitrate/1°	(μmol/kg)	1955–2017 annual mean nitrate at 0 surface	Nit	World Ocean Atlas (WOA 18) (Garcia et al., 2018a)
13	Nitrate/1°	μmol/kg	1955–2017 annual mean nitrate at 200 m	Nit200	World Ocean Atlas (WOA 18) (Garcia et al., 2018a)
14	Nitrate/1°	μmol/kg	Δ Nitrate between 0 and 200 m	ΔNit	World Ocean Atlas (WOA 18) (Garcia et al., 2018a)
15	Silicate/1°	μmol/kg	1955–2017 annual mean silicate concentration at 0 m	Sil	World Ocean Atlas (WOA 18) (Garcia et al., 2018a)
16	Salinity/1°		1955–2017 annual mean salinity at 0 m	Sal	World Ocean Atlas (WOA 18) (Zweng et al., 2018)
17	Dissolved oxygen/1°	μmol/kg	1955–2017 annual mean oxygen at 0 m	Ox	World Ocean Atlas (WOA 18) (Garcia et al., 2018b)
18	Dissolved oxygen/1°	μmol/kg	1955–2017 annual mean oxygen at 200 m	Ox200	World Ocean Atlas (WOA 18) (Garcia et al., 2018b)
19	Dissolved oxygen/1°	μmol/kg	Δ oxygen between 0 and 200 m	ΔOx	World Ocean Atlas (WOA 18) (Garcia et al., 2018b)
20	Apparent oxygen saturation/1°	μmol/kg	1955–2017 annual mean apparent oxygen saturation at 0 m	aos	World Ocean Atlas (WOA 18) (Garcia et al., 2018b)
21	Primary productivity/2°	g C/m^2/yr	Annual Mean Primary Productivity	pp	Ocean Productivity (National Center for Atmospheric Research Staff) (Eds)/ https://climatedataguide.ucar.edu/climate-data/ocean-productivity-phytoplankton-size-estimates-satellite

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1
Continued

	Variable name/resolution	Unit	Description	Var_code	Source
22	Carbonate ion concentration/1°	$[CO_3^{2-}]$	Calculated annual mean $[CO_3^{2-}]$ at 0 m	Scarb	Calculated in CO2SYS using temperature, phosphate, depth, silicate, and salinity data from WOA 18, and alkalinity and carbon dioxide data from Global Ocean Data Analysis (GLODAP) as input data (Key et al., 2015; Olsen et al., 2016; Pierrot et al., 2006)
23	Carbonate ion concentration/1°	$[CO_3^{2-}]$	Calculated annual mean $[CO_3^{2-}]$ at core depth	Carbcd	Same as source for Scarb
24	Carbonate ion concentration/1°	$[CO_3^{2-}]$	Calculated annual mean $[CO_3^{2-}]$ at 200 m	Carb200	Same as source for Scarb
25	Carbonate ion concentration/1°	$[CO_3^{2-}]$	Δ $[CO_3^{2-}]$ between 0 and 200 m	Carb0-200	Same as source for Scarb
26	Carbonate ion concentration/1°	$[CO_3^{2-}]$	Calculated annual mean $[CO_3^{2-}]$ at 500 m	Carb500	Same as source for Scarb
27	Carbonate ion concentration/1°	$[CO_3^{2-}]$	Δ $[CO_3^{2-}]$ between 0 and 200 m	ΔCarb	Same as source for Scarb
28	Carbonate ion concentration/1°	$[CO_3^{2-}]$	Δ $[CO_3^{2-}]$ between 0 and 500 m	Carb0-500	Same as source for Scarb
29	Thermocline depth/2°	m	Defined as the depth above which the temperature is superior to the surface temperature (at 10 m depth) minus 0.2°C	Thermdepth	IFREMER/LOS Mixed Layer Depth Climatology website (www.ifremer.fr/cerweb/deboyer/mld)/Mignot et al. (2007)
30	Pycnocline depth/2°	m	Values obtained for each core location	pycdepth	NOAA/(https://data.noaa.gov/dataset/dataset/mixed-layer-depth-isothermal-layer-depth-barrier-layer-depth-and-other-upper-ocean-thermohaline)
31	Mixed layer depth/1°	m	Surface mixed layer depth	Mld_ifremer	IFREMER/LOS Mixed Layer Depth Climatology website (www.ifremer.fr/cerweb/deboyer/mld/surface_mixed_layer_depth)/de Boyer Montégut et al. (2004)
32	Mixed layer depth/1°	m	Surface mixed layer depth	Mld_NOAA	NOAA/(https://data.noaa.gov/dataset/dataset/mixed-layer-depth-isothermal-layer-depth-barrier-layer-depth-and-other-upper-ocean-thermohaline)
33	Core depth/m	m	Values retrieved from cruise reports	Depth	Cruise Reports

2.4. Automated Imaging, Analysis, and Convolutional Neural Network (CNN) Training

Following Marchant et al. (2020), splits of the >150 μm size fraction were transferred to MiSo and images of each particle (primarily/exclusively whole foraminifera or foraminiferal fragments) were acquired (total $N = 311{,}380$, average N per sample = 3,797). With the aid of ParticleTrieur, a software developed at CEREGE (Marchant et al., 2020), 20,734 images were then manually labeled. Six convolutional neural networks ("Base Cyclic 16," "ResNet Cyclic 4," "ResNet Cyclic 8," "ResNet50 Cyclic Gain TL," "ResNet Cyclic TL," and "VGG19") with different topologies were then trained by random selection of 80% of the manually labeled images which served

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

as the training image set, while the remaining 20% served as the validation set that was used to evaluate CNN performance on three metrics:

1. accuracy, expressed as $\left(\frac{P_{\text{true}}+N_{\text{true}}}{P_{\text{true}}+P_{\text{false}}+N_{\text{true}}+N_{\text{false}}}\right)$;
2. recall, expressed as $\left(\frac{P_{\text{true}}}{P_{\text{true}}+N_{\text{false}}}\right)$; and
3. precision expressed as $\left(\frac{P_{\text{true}}}{P_{\text{true}}+P_{\text{false}}}\right)$.

In the three above expressions, P_{true} = true positives, P_{false} = false positives, N_{true} = true negatives, and N_{false} = false negatives. The best performing CNN was then adopted in the automatic identification of the remaining images. To ascertain the reliability of our model, all 311,380 images were validated manually, and comparison was made between the model classification and that collectively performed by two human classifiers (a graduate student and a senior researcher). Morphometric parameters (area, perimeter, convex area, convex perimeter, major axis, minor axis, minimum enclosing circle area, minimum enclosing circle radius, eccentricity, solidity, roundness, circularity, mean, standard deviation, standard deviation invariant, skewness, kurtosis, fifth moment, and sixth moment) as well as the abundance of each species within each assemblage were extracted from ParticleTrieur. For more details on ParticleTrieur workflows, see Marchant et al. (2020). The image data sets generated in this study using MiSo are archived in the SEANOE database and can be accessed at: https://doi.org/10.17882/86411 (Adebayo, Bolton, Marchant, Bassinot, Conrod, et al., 2022).

2.5. Size and Relative Abundance Data

We measured individual foraminiferal size based on the test maximum diameter. The test maximum diameter was calculated from the area using the formula:

$$\text{MD} = \left(\sqrt{\frac{\text{Area}\left(\text{mm}^2\right)}{\pi}}\right) * 2.$$

We chose to measure test size using the maximum diameter because it is less affected by random object orientation (Schmidt, Renaud, et al., 2004) and has been reported to be a consistent size proxy for two-dimensional images (Brombacher et al., 2018). The 95th percentile size of each assemblage, hereafter referred to as $\text{size}_{95/5}$, was selected to investigate the size response of planktonic foraminifera assemblages to environmental variables. We chose $\text{size}_{95/5}$ because, compared to mean and maximum size, it is less sensitive to single outliers (Rillo et al., 2020; Schmidt, Renaud, et al., 2004). To validate our size spectra, we compared our $\text{size}_{95/5}$ data with published measurements reported for the TIO from core-top samples (Rillo et al., 2019, 2020). We also compared the relative abundances of the species recorded in this study with those reported for the closest sites to our core locations in the ForCenS Database, which compiles planktonic foraminifera assemblages from core-top sediments (Siccha & Kucera, 2017).

2.6. Dissolution

To estimate the degree of dissolution in our samples based on fragmentation intensity, we used the index of Suárez-Ibarra et al. (2021):

$$\text{FI} = \frac{\frac{(fb)}{(fb+pf)}}{\frac{(\text{area})}{(\text{perimeter})}}$$

where *fb* represents the total number of broken tests and fragments, *pf* is the number of whole planktonic foraminifera tests, and area and perimeter are the averages for the broken and fragmented specimens in each sample. Values closer to zero indicate well preserved samples while values closer to one indicate that tests are more affected by dissolution. In addition, we explored other dissolution indicators by investigating the relationship between $\text{size}_{95/5}$ versus depth, $\text{size}_{95/5}$ versus carbonate saturation state at core depth ($\Delta[CO_3^{2-}]$), and $\text{size}_{95/5}$ versus fragmentation rate.

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.7. Statistical Analyses

Before analyzing the planktonic foraminiferal size and assemblage data, we excluded rare species (i.e., species with <50 individuals per sample in all samples). We also excluded samples with low whole specimen counts (<300 whole planktonic foraminifera in a sample); this brought the total number of core-top samples analyzed from 82 to 62. We used the Jarque–Bera normality test (Jarque, 2011) in the R package "tseries" (version 0.10–51) to evaluate whether our size data set showed a normal distribution (based on skewness and kurtosis). This enabled us to decide whether parametric or non-parametric correlation and regression analyses would be preferable to explore relationships between variables. Results of the Jarque–Bera normality test confirmed that the distribution of our size data set is non-parametric (X-squared = 478,591, $p < 2.2e{-}16$); therefore, only non-parametric correlations and regression analyses were adopted going forward.

For comparison purposes, we restricted the Rillo et al. (2020) data set to the TIO (30°S–30°N, 54–112°E) and utilized shell diameter data instead of shell area. Using the Spearman's rank-order correlation ("ρ"), we compared our $size_{95/5}$ data with that contained in the Rillo et al. (2019) data set. Likewise, we compared the species' relative abundances recorded in this study with those reported for the closest sites to our core locations in the ForCenS Database (REF). Aside from making comparisons based on size measurements, through robust regression analysis, we also assessed the relationship between the $size_{95/5}$ of nine planktonic foraminifera species (*Pulleniatina obliquiloculata*, *Globorotalia menardii*, *Neogloboquadrina dutertei*, *Globigerinoides ruber albus*, *Trilobatus sacculifer*, *Globigerinella siphonifera*, *Globigerinoides conglobatus*, *Globorotalia truncatulinoides*, and *Globoconella inflata*) and mean annual SST as contained in the Rillo et al. (2019) data set for the TIO. We used the "robustbase" package (version 0.95-0) in R for this and subsequent robust regression analyses.

A size frequency distribution (SFD) was used to show the size ranges, the number of observations within a given size, as well as the number of species represented across the study sites. We determined the modal distribution of each species' SFD using the "LaplacesDemon" package (version 16.1.6) in R which tests for unimodality, bimodality, and multimodality in a given distribution. Through robust regression analyses, we tested the OSH by: (a) examining the relationship between individual species' $size_{95/5}$ and their relative abundance, assuming that relative abundance is a measure of optimum growth conditions, and (b) investigating the relationship between individual species' $size_{95/5}$ and environmental parameters. For these tests, p-values were adjusted using the Holm-Bonferroni correction method (Holm, 1979) to control the family-wise error rate and reduce the chances of obtaining false positive significance levels. In visualizing the species-specific environmental controls on the individual species $size_{95/5}$ (hypothesis test (b) above), because the number of variables tested were many, we binned the variables under umbrella parameters as follows: *carbonate* (includes surface carbonate concentration and carbonate concentration at core depth); *temperature* (includes SST, temperature at 10 m, temperature at 50 m, temperature at 100 m, temperature at 200 m, and difference between winter and summer temperatures); *nutrient* (surface phosphate concentration, surface nitrate concentration, nitrate concentration at 200 m); *oxygen* (includes Δoxygen concentration between 200 and 0 m and oxygen concentration at 200 m); and *salinity* (surface salinity).

We further explored the relationship between $size_{95/5}$, species richness (defined as the number of species in each sample), and species diversity through robust regression analyses. Species diversity was calculated using the classical Shannon–Weiner diversity index (Shannon & Weaver, 1949):

$$H = \sum_{i=1}^{s} Pi(\ln pi)$$

where,
H = Shannon–Weiner diversity index
S = Total number of species encountered
Pi = proportion of ith species in the population
ln = natural logarithm
$\sum$ = sum from species 1 to species S

We assessed the influence of assemblage composition on $size_{95/5}$ in two ways: first, in each sample, we investigated the size spectrum of the 13 most abundant species relative to the sample $size_{95/5}$ using the density function in the R ggplot2 (version 3.3.6) package (Wickham, 2016); second, through robust linear regression analysis, we evaluated the relationship between $size_{95/5}$, the total number of foraminifera, and the percentage contribution of species with sizes greater than the regional $size_{95/5}$ in our data set. The regional $size_{95/5}$ was calculated by binning

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 2
Principal Component Analysis Factor Loading of Surface Hydrographic Parameters in the Tropical Indian Ocean

	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7
SST	−0.158	**0.735**	0.384	−0.455	0.278	0.046	0.021
NIT	0.063	−0.161	−0.114	0.202	**0.941**	0.112	−0.135
PHOS	−0.094	0.127	0.636	**0.748**	−0.059	0.071	0.039
PP	−0.060	−0.015	−0.055	−0.049	−0.141	**0.959**	−0.225
CO_3^{2-}	**0.978**	0.163	0.107	−0.003	−0.036	0.056	−0.037
OXY	−0.034	0.625	**−0.644**	0.435	−0.063	−0.021	−0.021
SAL	0.039	−0.027	−0.073	0.006	0.092	0.237	**0.963**

Note. Keys—SST (sea surface temperature); NIT (surface nitrate concentration), PHOS (surface phosphate concentration); PP (primary productivity); CO_3^{2-} (surface carbonate ion concentration); OXY (surface oxygen concentration); and SAL (salinity). The values in bold indicate the parameter with the highest weighting on each PC axis.

the individual sizes of planktonic foraminifera from all core-tops and computing the 95th percentile size. Species with sizes higher than the regional $size_{95/5}$ are regarded as "large" because we assume that the size of adult-sized species differ and the maximum size they can attain is fixed. Hence, the percentage contribution of species with sizes greater than the regional $size_{95/5}$ represents the ratio of the total number of species where at least one specimen in a sample is larger than the regional size95/5 to the total number of species in the sample.

To investigate the environmental controls on planktonic foraminiferal assemblage $size_{95/5}$ in the TIO, we first characterized planktonic foraminiferal $size_{95/5}$ distribution by computing a factor analysis to identify the discernible ecological patterns in the assemblage-wide planktonic foraminifera $size_{95/5}$ records. The factor analysis was carried out only on the species' $size_{95/5}$ data, with the species name on the horizontal axis and the core IDs on the vertical axis. Each value in the matrix represents the $size_{95/5}$ of each species within the core-top sample they were found. The output of the factor analysis are the factor scores and factor loadings. Thereafter, we regressed the factor scores against environmental data to identify the primary and secondary drivers of size at the assemblage level using a robust regression model. We assessed the reliability of our analytical procedure by carrying out factor analysis on the species' relative abundance data, identifying their latent variables, and comparing our results to a similar work by Cayre et al. (1999), who reconstructed paleoproductivity of the TIO from planktonic foraminifera assemblage distribution. Our goal was to check whether the same parameters found by Cayre et al. (1999) to primarily drive planktonic foraminifera assemblage distribution within the TIO were also identified using our machine learning approach. All analyses were carried out in R version 4.0.3 (R Core Team, 2020) and all packages used are listed in Table S1 of Supporting Information S1.

3. Results

3.1. PCA of Surface Oceanographic Parameters

Results of PCA run on TIO surface hydrographic data show that the first two principal components explain 63% of the total variance (Figure S1 in Supporting Information S1). PC1 alone explains 41% of the variance and is likely associated with seawater carbonate chemistry, indicated by its large component loading on carbonate ion concentration (97%). PC2 explains 22% of the variation and mainly relates to variations in SST as shown by its large component loading on SST (74%; Table 2). We present the full component loadings of the surface parameters analyzed in Table 2.

3.2. Ground-Truthing Our High Throughput Workflow

3.2.1. CNN Performance and Comparison With Human Classifier

Of the six CNNs trained, the "Base Cyclic 16" network gives the best results, reaching an accuracy of 89.9%; a precision of 77.7%; and a recall of 72.6% over five iterations (Figure 2a). "Base Cyclic 16" is robust in identifying classes such as whole planktonic foraminifera, fragments (broken foraminifera), *Uvigerina peregrina* (a benthic foraminiferal species), aggregates (including radiolarians, acritarchs, and other organic materials such as broken non-foraminiferal shells), and doubles (i.e., more than one particle in an image). This is shown by

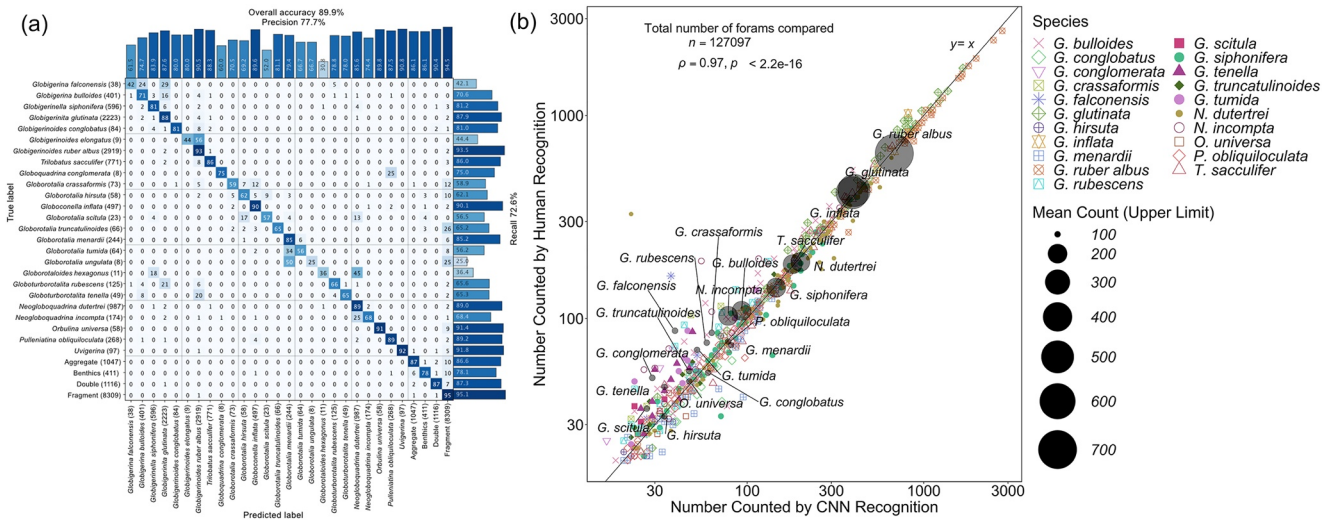

Figure 2. (a) Confusion matrix showing the performance of "Base Cyclic 16" convolutional neural network on the validation data set. We observed that poorly represented samples in the validation set were more frequently misidentified (more false positives) compared to well-represented species. (b) Count comparability between human and machine classifiers for the most abundant species. Although identified, rare species such as *Globigerinella calida* are excluded from this analysis. *Globigerinoides ruber albus* and *Globigerinita glutinata* were the most abundant species in the tropical Indian Ocean data set. Values on both *x*- and *y*-axis are represented as log-scales. The mean count (upper limit) represents the average count per sample within a range binning of 100.

its high accuracy scores ranging between 86% and 95%, and a recall of 78%–95%. Other CNNs trained typically ranged between 81% and 88% in accuracy, 64%–76% in precision, 60%–70% in recall. The results of the performance of the "Base Cyclic 16" network tested over five iterations is presented in Table S2 of Supporting Information S1.

After deploying the "Base Cyclic 16" CNN model in the automatic classification of all images, comparison with the classification by a human classifier shows that the machine successfully labeled most of the species and achieved near full accuracy compared to a human classifier ($\rho = 0.97$, $p < 2.2e{-}16$, Figure 2b). Using this automated classification scheme, we plot the spatial distribution of planktonic foraminifera from our core-top data set (Figures 3a–3c). While *G. ruber albus*, *G. inflata*, and *Globigerinita glutinata* are the most abundant species in the TIO, we observe a dominance of *G. glutinata* in the northern TIO reaching 85%, while the southern TIO is dominated by *G. ruber* up to 69%, as shown in Figure S2 of Supporting Information S1.

3.2.2. Relative Abundance Comparison With ForCenS Database

Comparison of the relative abundance of the 21 most abundant species from our study sites with the corresponding closest sites (pairing was based on a single site) in the *ForCenS Database* (Siccha & Kucera, 2017) shows a significant correlation (Figure 4a). While the data fall near the $y = x$ line and the correlation between both data sets is strong ($\rho = 0.77$, $p < 2.2e{-}16$), some differences in the relative abundances of some species occur. These could result from differences in the total number of individuals counted in this study (2,000 individuals per sample on average) compared to that of the *ForCenS* database (300 individuals per sample on average). Nevertheless, the similarity between the two data sets becomes more evident when we observe the regions of highest relative abundance of some of the species. For example, in both studies, the relative abundance of *G. glutinata* is highest in the Arabian Sea while the highest relative abundance of *G. ruber albus* is seen around the Mozambique Channel. Likewise, *G. inflata* relative abundance is highest around the Southern Mozambique Basin (Figure 3 and Figure S3 in Supporting Information S1).

3.2.3. Morphometry Data Comparison With Rillo et al. (2020) Data Set

There is a strong correlation ($\rho = 0.94$, $p = 5.9e{-}06$) between the $size_{95/5}$ of the TIO planktonic foraminiferal species measured in the present study and those recorded in the Rillo et al. (2020) data set, confirming the reliability of the morphometric data generated using our workflow (Figure 4b). However, the size of many of the species fall below the $y = x$ line, suggesting that the Rillo et al. (2020) data set is biased toward larger sizes. This bias, noted by Rillo et al. (2020), comes from a collector bias who predominantly picked, identified, and curated the largest foraminifera in their samples.

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

10.1029/2022GC010586

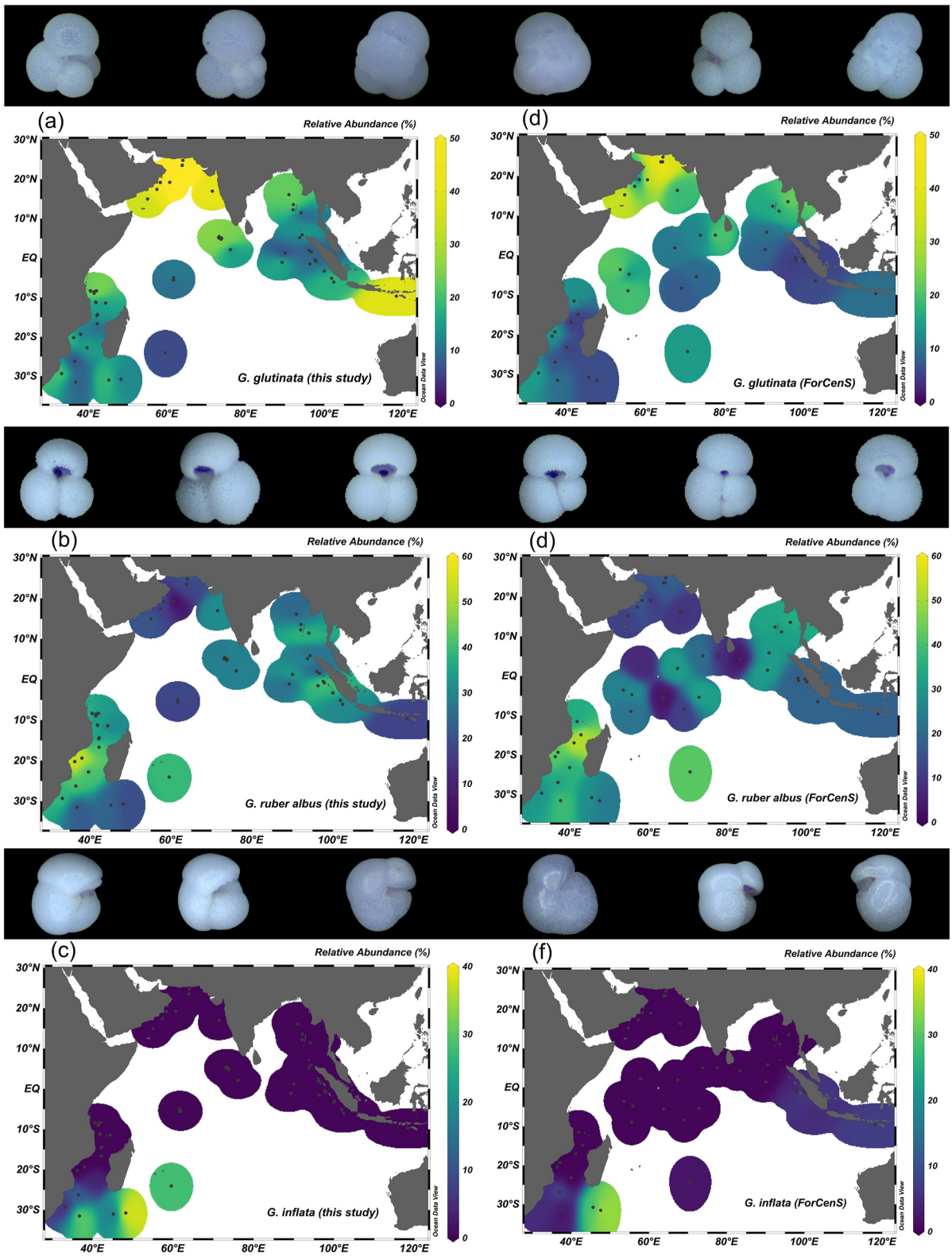

Figure 3. Similarity in the areas of highest relative abundance of *G. glutinata*, *G. ruber albus*, and *G. inflata* in this study versus ForCenS database. Purple color represents areas of low relative abundance while yellow color represents areas of high relative abundance. Gradient colors between purple and yellow colors represent intermediate relative abundance (See Figure S3 in Supporting Information S1 for other species). Representative images of each species are shown above their distribution map.

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

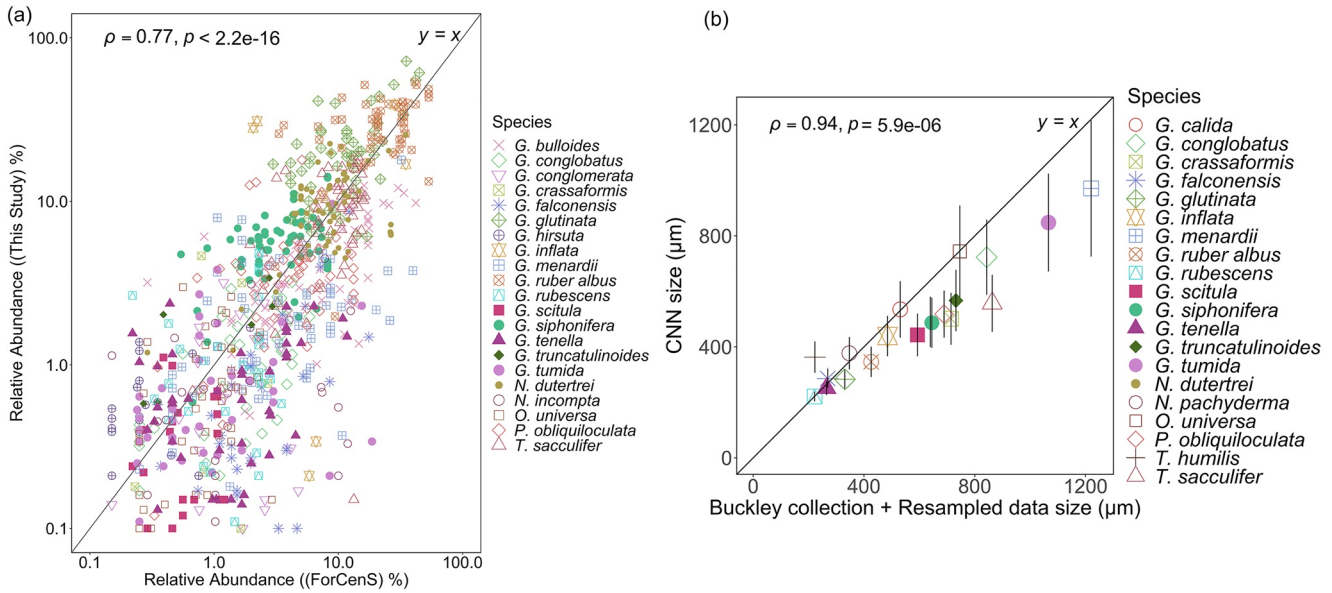

Figure 4. (a) Comparison between the relative abundance of planktonic foraminifera species (this study) versus the ForCenS database. Site selection was based on nearest-neighbor that is, the site closest to our core locations in the ForCenS Database (Siccha & Kucera, 2017). (b) Comparison between planktonic foraminifera $size_{95/5}$ recorded in this study and those from the tropical Indian Ocean samples in the Rillo et al. (2020) data set. Line is identity function. Error bars represent 1 standard deviation in our measurements. The Buckley collection represents samples from the Henry Buckley Collection of Planktonic Foraminifera stored at The Natural History Museum in London, UK while the "resampled data" refers to the data from the bulk sediments from which Rillo et al. (2020) validated the Buckley collection. The resampling was done on the Ocean Bottom Deposits collection which were bulk sediments from which Buckley built his collections.

3.2.4. Assemblage Distribution Factor Analysis Comparison With Cayre et al. (1999)

Factor analysis on species' relative abundance data from our study sites reveal that 10 axes explain 72% of the total variance in the assemblage distribution of the species studied, with the first two axes accounting for 32% of the total variance. Thus, we make interpretations based on the first two axes. High negative factor scores for the first factorial axis are related to species such as *Globigerina falconensis*, *Globigerina bulloides*, *G. glutinata*, and *T. sacculifer*. In the second factorial axis, high positive factor scores are represented by species such as *Globorotalia crassaformis*, *G. truncatulinoides*, and *G. inflata* (see Table S3 in Supporting Information S1); interestingly, these same species are highly weighted in the second axis of the size factor loading. Regression between

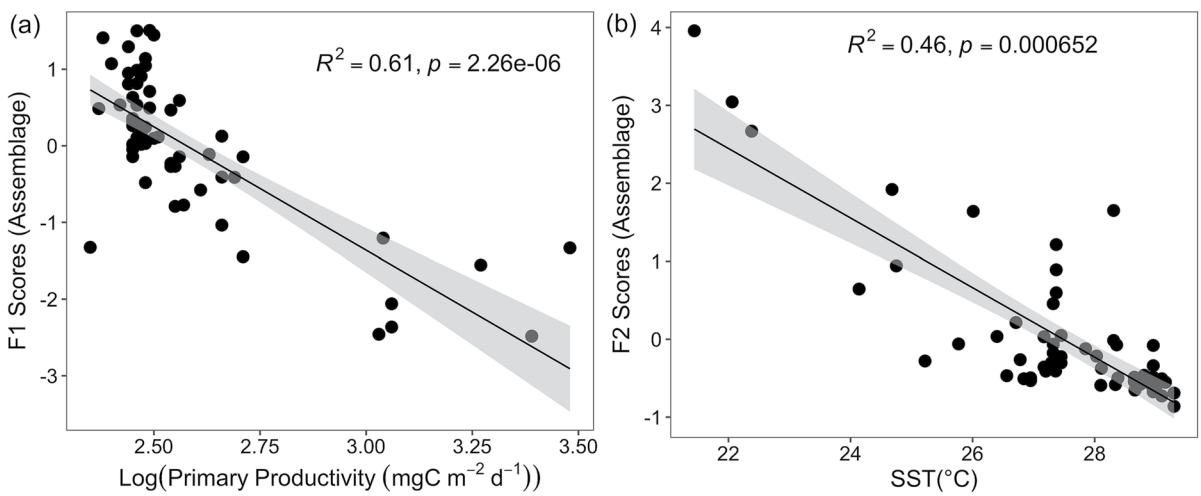

Figure 5. Robust regression analysis between the *F*1-Axis factor scores computed from the factor analysis of 26 planktonic foraminifera species' relative abundance data and environmental parameters showed best fit with log of primary productivity (PP) (a), while the *F*2-Axis factor scores showed best fit with sea surface temperature (b). Note that the PP data was log-transformed because of the non-normal distribution of the data and the larger data range relative to other variables.

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

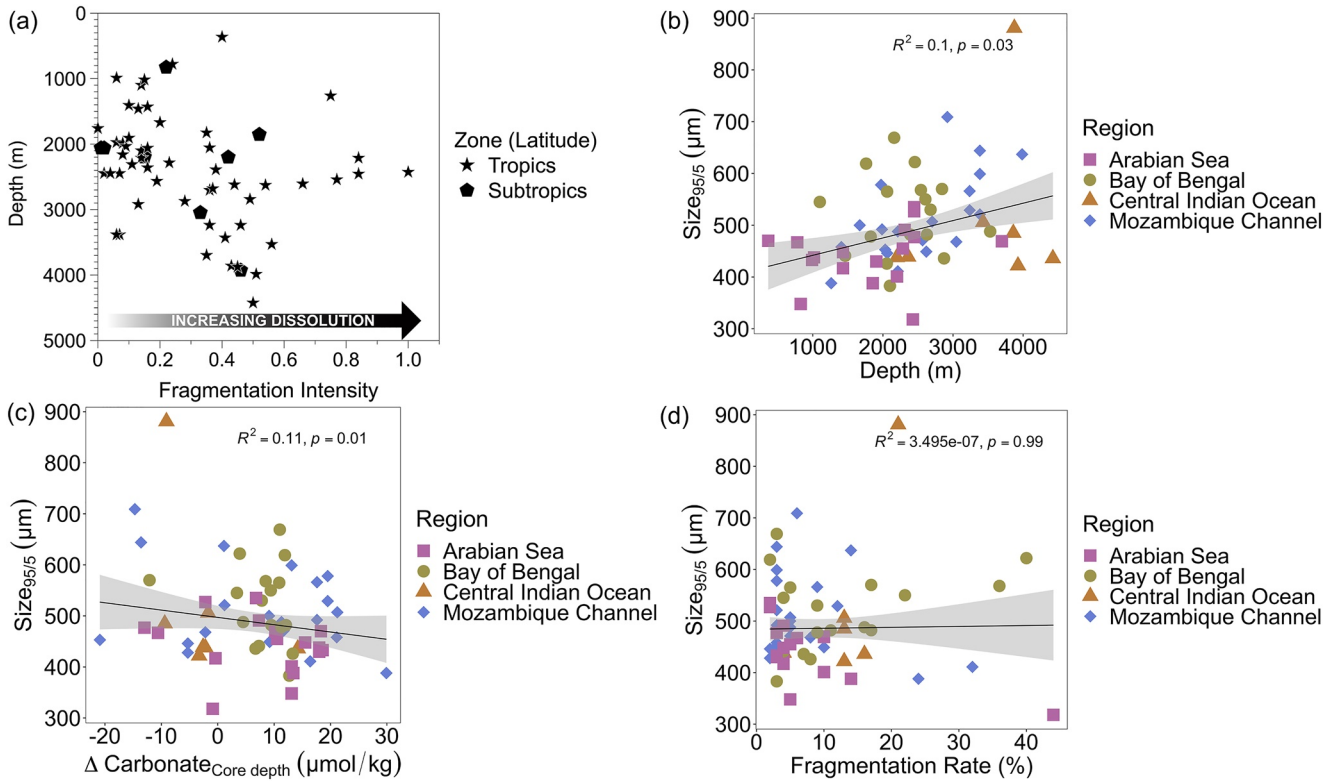

Figure 6. Dissolution proxies: (a) Fragmentation intensity according to Suárez-Ibarra et al. (2021) versus depth; (b) $size_{95/5}$ versus depth; (c) $size_{95/5}$ versus carbonate saturation state at core depth; (d) $size_{95/5}$ versus fragmentation rate. Shaded intervals represent 95% confidence interval.

the $F1$-Axis and environmental parameters show highest R^2 values with (log transformed) PP (mgC m^{-2} d^{-1}) ($R^2 = 0.61$, $p < 0.05$, Figure 5a) while the $F2$-Axis show highest R^2 values with SST ($R^2 = 0.46$, $p < 0.05$, Figure 5b). This result is similar to that of Cayre et al. (1999) whose first factorial axis after PCA on planktonic foraminifera assemblage distribution in core-top samples from the TIO showed best fit with PP. This result gave us confidence in our subsequent analysis on the assemblage size data set.

3.3. Size Distribution of Planktonic Foraminifera in the TIO

3.3.1. Dissolution

The fragmentation intensity dissolution index indicates good preservation in most of the samples (Figure 6a). However, it suggests relatively high levels of dissolution (index >0.6) in six out of 62 samples, representing about 10% of the samples. A weak relationship is observed between $size_{95/5}$ and depth ($R^2 = 0.10$, $p = 0.03$) (Figure 6b). Similarly, changes in carbonate saturation state at core depth ($\Delta[CO_3^{2-}]$) show a weak relationship with $size_{95/5}$ ($R^2 = 0.11$, $p = 0.01$) (Figure 6c). No relationship is found between $size_{95/5}$ and fragmentation rate ($R^2 = 3.50e{-}07$, $p > 0.99$) (Figure 6d). Given that changes in seafloor depth and bottom water carbonate saturation state respectively explain 10% and 11% in the variation of our size metric ($size_{95/5}$), and only 10% of the samples were identified by the fragmentation intensity dissolution index as being significantly affected by dissolution, we conclude that the influence of dissolution on our species' size records is small and that our results in relation to environmental parameters are likely robust. This is supported by the fact that the $size_{95/5}$ values in some deeper sites are smaller than those in shallower waters, whereas preferential dissolution of smaller individuals might be expected at deeper sites with lower carbonate saturation. Thus, we did not exclude any of the sites identified by the fragmentation intensity index from our analysis because their species size spectra and $size_{95/5}$ do not indicate any significant bias toward larger individuals or more robust species, and the total whole foraminifera counts from these samples were quite high.

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

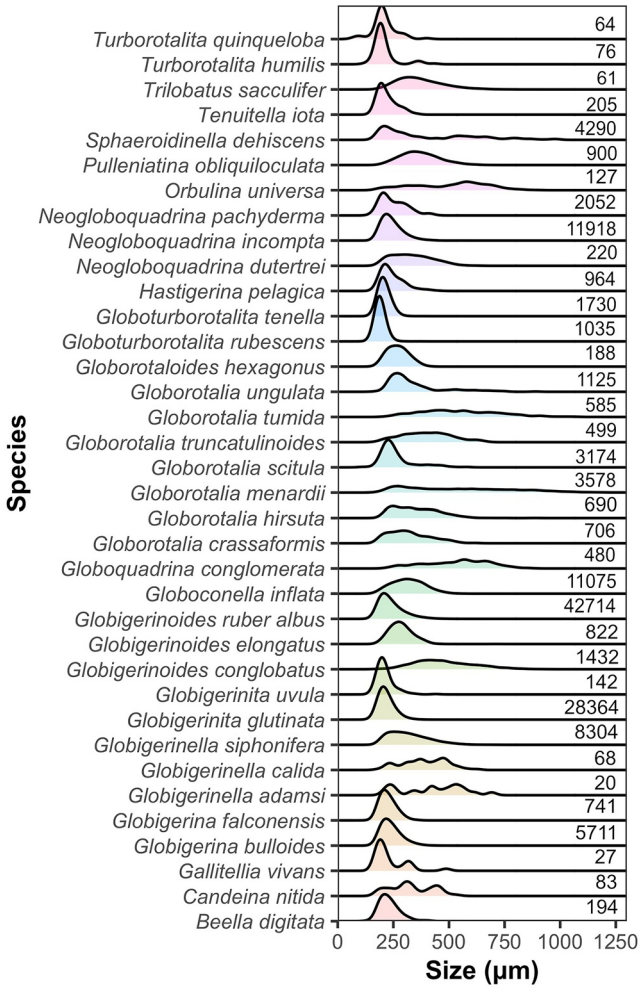

Figure 7. An overview of the size frequency distribution (SFD) of the planktonic foraminifera species recorded in this study. About 56% of the species showed unimodal distribution for adult sized species (>150 μm), while 44% of the species showed multimodal distributions. The numbers to the right of each species' SFD represent the total number of individuals represented in their distribution.

3.3.2. Species Size Frequency Distribution

Of the 36 species recorded, the size-frequency distribution (SFD) of 23 species appear to possess unimodal distributions, while the remaining 13 species show multimodal distributions (Figure 7). Bimodal distribution is observed in species such as *N. pachyderma and Candeina nitida*. Meanwhile, the SFD of the largest species such as *G. menardii*, *Globoquadrina conglomerata,* and *Globorotalia tumida* are widely spread, showing no clear peaks. Statistical confirmation of the modal distribution types observed for each species is presented in Table S4 of Supporting Information S1.

3.4. Test for the OSH

3.4.1. Size Versus Relative Abundance

Out of the 26 species tested, only *G. conglobatus* increased in $size_{95/5}$ with increasing relative abundance ($R^2 = 0.21$, $p = 0.01$) (Figure 8). No significant relationship is observed between the $size_{95/5}$ and relative abundance of most species, including the abundant species such as *G. glutinata*, *G. elongatus*, *N. dutertrei*, and *T. sacculifer*.

3.4.2. Size Versus Abiotic Forcing (Environmental Parameters)

Evaluated based on the parameter that showed the best coefficient of determination (R^2), the response of individual species' $size_{95/5}$ to environmental parameters investigated in this study reveals that 25% of the species show a low to moderate significant response (R^2 values range from 0.10 to 0.35, $p < 0.025$) to temperature-related parameters excluding SST (Figure 9). On the other hand, parameters related to carbonate ion concentration (R^2 values range from 0.10 to 0.29, $p < 0.025$) explain $size_{95/5}$ variation in 35% of the species, especially in deep-dwelling and large-sized species (e.g., *G. menardii*, *G. tumida*, *G. truncatulinoides,* and *P. obliquiloculata*). Factors such as salinity (R^2 values range from 0.22 to 0.51, $p < 0.025$) and oxygen (R^2 values range from 0.18 to 0.23, $p < 0.025$) explain $size_{95/5}$ variation in 20% and 10% of the species respectively. Meanwhile, nutrient availability explains $size_{95/5}$ variation in 10% of species (R^2 values range from 0.19 to 0.31, $p < 0.025$). Only *Neogloboquadrina incompta* $size_{95/5}$ shows no significant relationship with any environmental parameter (see Table S5 in Supporting Information S1 for details of the exact parameters that are found to drive the variation in individual species' $size_{95/5}$).

With a focus on the impact of SST on the size variations of the species examined, nine out of 20 species show weak but significant relationships with SST including *G. bulloides* ($R^2 = 0.07$, $p = 0.01$), *G. siphonifera* ($R^2 = 0.07$, $p = 0.002$), *T. sacculifer* ($R^2 = 0.05$, $p = 0.03$), *G. conglomerata* ($R^2 = 0.13$, $p = 0.01$), *G. crassaformis* ($R^2 = 0.18$, $p = 0.01$), *Globorotalia hirsuta* ($R^2 = 0.19$, $p = 0.001$), *G. inflata* ($R^2 = 0.30$, $p = 2.79\text{e}{-05}$), *G. menardii* ($R^2 = 0.05$, $p = 0.03$), and *N. dutertrei* ($R^2 = 0.04$, $p = 0.05$). However, none of these R^2 values are greater than the parameters that we find to primarily drive size changes in all these species. See Table S6 in Supporting Information S1 for details of the R^2 and p-values describing the relationship between individual planktonic foraminifera species $size_{95/5}$ and mean annual SST.

Our re-analyses of TIO planktonic foraminifera in the Rillo et al. (2020) data set reveal that, out of the nine species investigated, only three show significant relationships between their $size_{95/5}$ and SST. The three species are *T. sacculifer* ($R^2 = 0.22$, $p = 0.03$), *G. truncatulinoides* ($R^2 = 0.58$, $p = 0.02$), and *G. inflata* ($R^2 = 0.69$, $p = 0.03$), although we note that the number of samples per species are significantly lower than in our study (62 vs. 14 for *T. sacculifer*, 27 vs. 9 for *G. truncatulinoides,* and 25 vs. 4 for *G. inflata*) (Figure S4 in Supporting Information S1).

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

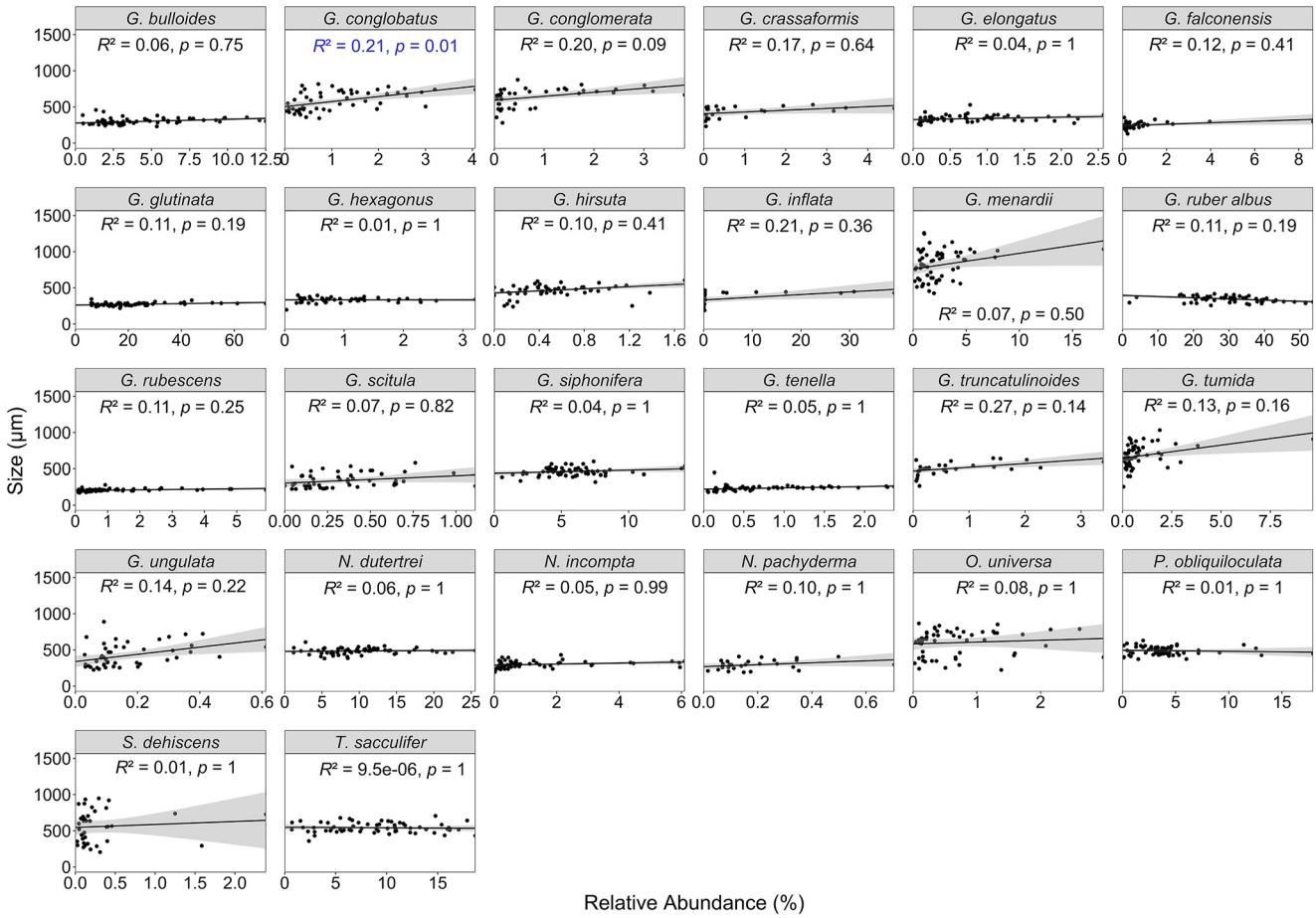

Figure 8. Single-species robust regression analysis showing the relationship between individual species' size$_{95/5}$ and their relative abundances. We assume that their relative abundances indicate optimum growth conditions. *P*-values were corrected using the Holm–Bonferroni correction method.

3.5. Size Versus Species Diversity and Species Richness

Species richness shows no relationship with assemblage size$_{95/5}$ ($R^2 = 0.013$, $p = 0.546$) (Figure 10a). Conversely, the Shannon–Weiner diversity index increases with increasing assemblage size$_{95/5}$ ($R^2 = 0.49$, $p = 6.43e{-}09$) (Figure 10b). Highest Shannon diversity (2.34) is found in a core-top located near the Madingley Rise (central Indian Ocean) at a water depth of 3,875 m while minimum Shannon diversity (1.21) is recorded in a core located in the Arabian Sea at a water depth of 2,427 m.

3.6. Influence of Assemblage Composition on Assemblage Size$_{95/5}$

Generally, the percentage contributions of species with sizes greater than the regional size$_{95/5}$ in our samples are around 20%–40%, with the largest contributions (50%) observed predominantly in core-tops from the central Indian Ocean and Bay of Bengal (including offshore Java) (Figure 11). Above a size$_{95/5}$ of 550 µm, assemblage size$_{95/5}$ seems to be impacted by the increasing percentage contribution of species with sizes greater than the regional size$_{95/5}$. Within the central Indian Ocean, irrespective of the total number of foraminifera, size$_{95/5}$ values do not exceed ~500 µm except in one core-top. Thus, this observation suggests that foraminifera growing in this region are smaller. Finally, we observe that maximum size$_{95/5}$ values are attained at around a count of 2,000 individuals. We suggest that this might be a minimum count limit for foraminiferal size analysis using automated systems. Further, we find no clear regional character in the modal size and size$_{95/5}$ in the studied sites based on assemblage composition. In all cores, the modal size is lower than the size$_{95/5}$, thus indicating that the size distribution in all the cores is not skewed by large sized species and that small sized species contribute significantly to the assemblages (see Figure S5 in Supporting Information S1).

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

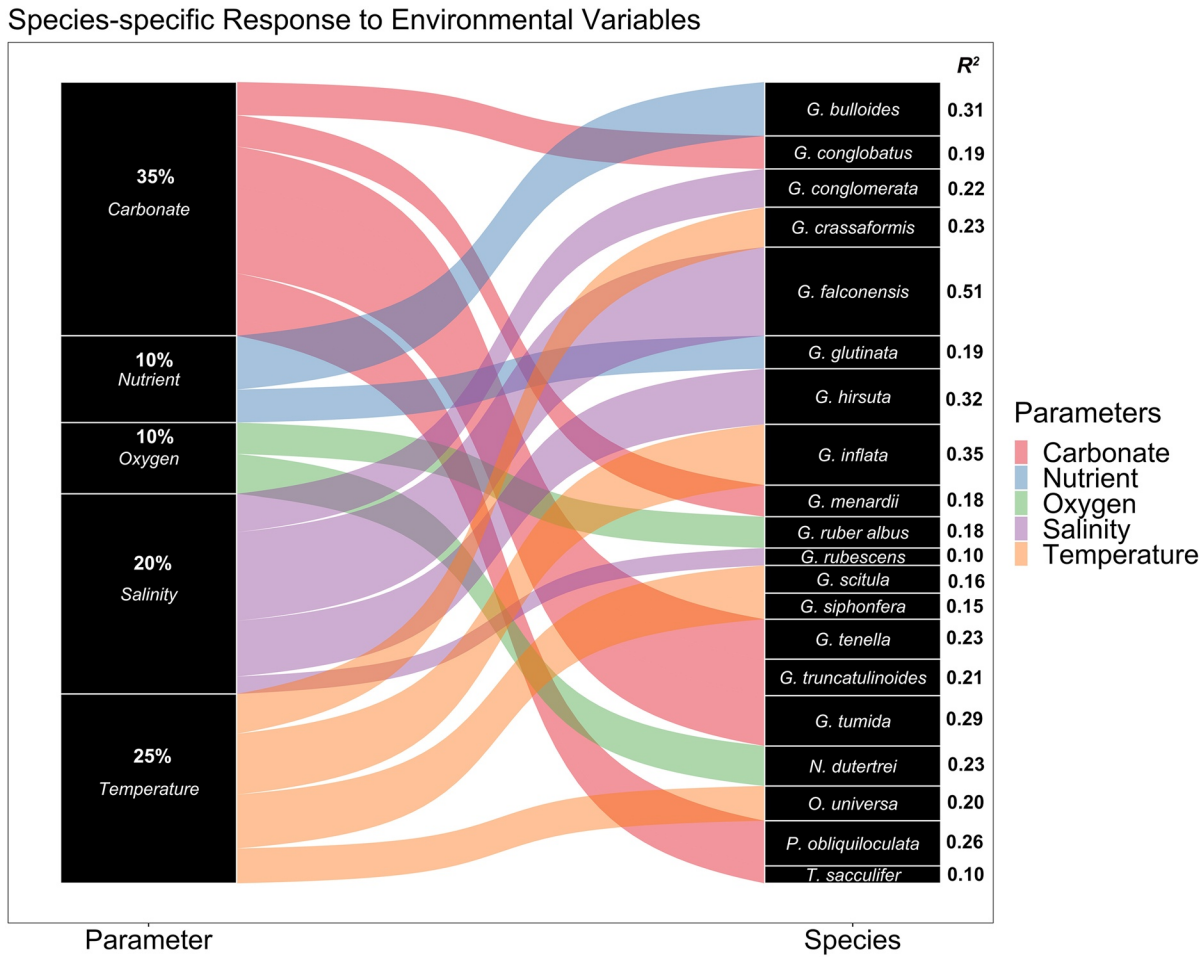

Figure 9. Robust regression analysis between individual planktonic foraminifera species' size$_{95/5}$ and environmental parameters. Values contained in each parameter box represent the percentage proportion explained by that parameter among the five broad parameters tested. Values to the right of each species represent the coefficient of determination between that species' size$_{95/5}$ and the parameter it showed the best fit with. Because the number of variables tested were many, we binned the variables under umbrella parameters as follows: **Carbonate** (includes surface carbonate concentration, carbonate concentration at 200 m, and ΔCarbonate concentration between 0 and 500 m); **Temperature** (includes temperature at 100 m, temperature at 200 m, winter sea surface temperature (SST), and summer SST); **Nutrient** (includes surface phosphate concentration and phosphate concentration at 200 m); **Oxygen** (includes Δoxygen concentration between 200 and 0 m); and **Salinity** (surface). See Table S6 in Supporting Information S1 for details about the exact parameters that were found to drive individual species' size$_{95/5}$ variation.

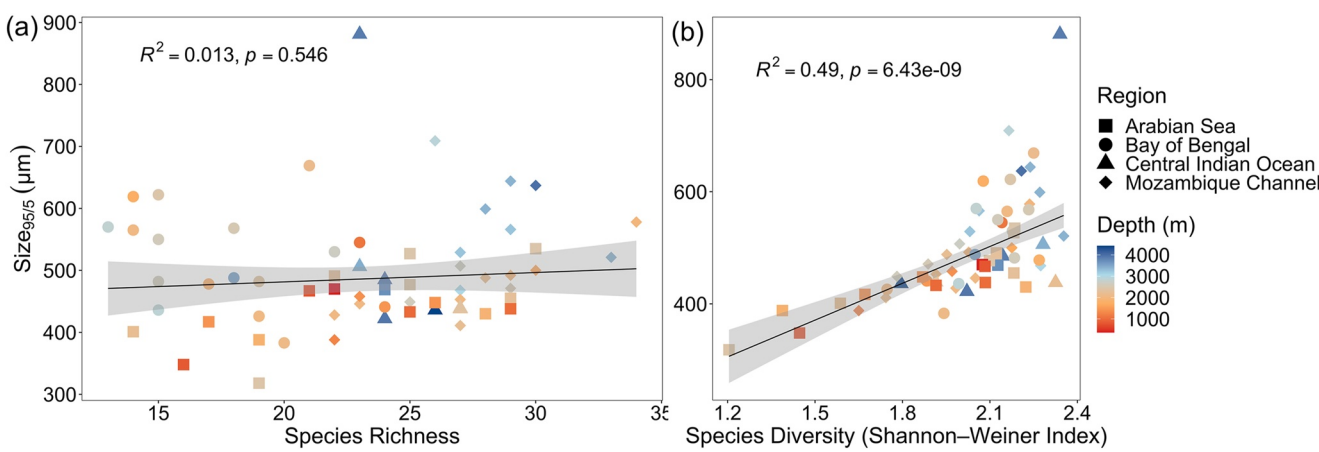

Figure 10. (a) Relationship between size$_{95/5}$ and species richness; (b) relationship between size$_{95/5}$ and species diversity. Shaded intervals represent the 95% confidence interval.

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

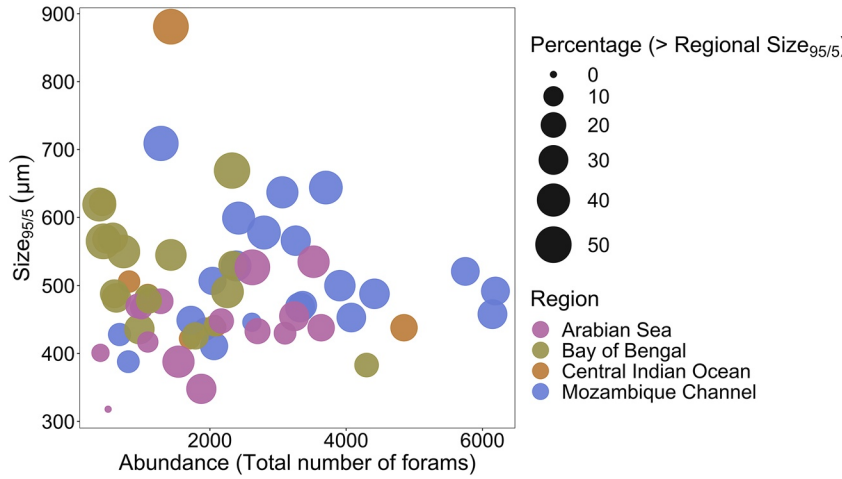

Figure 11. $Size_{95/5}$ versus total numbers of planktonic foraminifera per sample. The size of the circles corresponds to the percentage contribution of species with sizes greater than the regional $size_{95/5}$ in the Indian Ocean. The regional $size_{95/5}$ here is taking as the arbitrary size cut-off that defines the size above which species are regarded as "large" in the tropical Indian Ocean. Increasing percentage representation show the effect of assemblage composition on local $size_{95/5}$ values.

3.7. Assemblage Size Response (Factor Analysis)

The size spectra of 26 planktonic foraminiferal species were used for this analysis. Ten factors explain 70% of the total size variance. The first and second axes explain 45% of the variance while axes three to 10 explain only 25% and were often monospecific, thus, we only used the first and second axes. The first factorial coordinate of the *F*-Matrix derived from the factor analysis of planktonic foraminifera assemblage size attributes high positive factor scores mostly to species that either bear calcite crusts or are large-sized. These include *Globigerinoides conglobatus*, *G. menardii*, *G. tumida*, *Pulleniatina obliquiloculata*, *Neogloboquadrina dutertrei*, *and Orbulina universa*. On the other hand, high negative factor scores are mainly attributed to deep-dwelling species such as *G. inflata*, *G. crassaformis*, and *G. truncatulinoides* on the second factorial coordinate (see Table S7 in Supporting Information S1). Among the environmental parameters explored, the first factorial coordinate shows the best fit with surface carbonate ion concentration (Figure 12a), while the second factorial coordinate shows strongest fit with SST (Figure 12b).

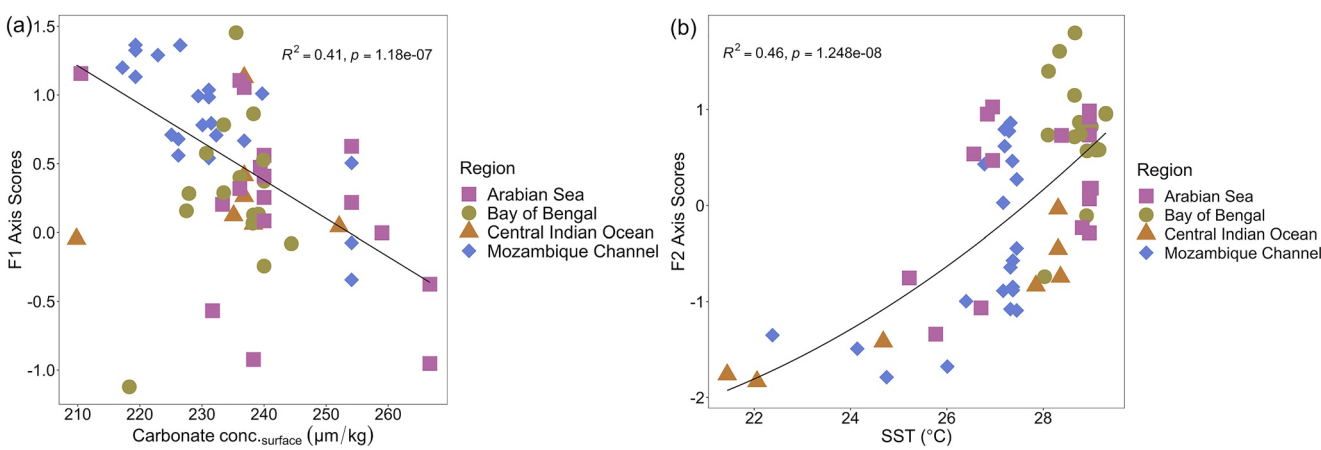

Figure 12. (a) Robust regression analysis between *F*1-Axis factor scores computed from the Factor Analysis of 26 planktonic foraminifera species morphometric data versus surface carbonate concentration: (b) second order polynomial relationship between *F*2-Axis scores versus sea surface temperature.

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

4. Discussion

4.1. CNN Classifies Species Images With High Accuracy

A review of our neural network "Base Cyclic 16" CNN confusion matrix (Figure 2a) reveals that despite its high accuracy, it misclassified some planktonic foraminifera species. The major identified biases by the CNN were inductive (i.e., when a model incorrectly predicts a species' class due to the non-inclusion of the species in the training data set or favoring of highly sampled classes over less-represented classes), numerical, and taxonomic. For instance, *G. elongatus* was often misidentified as *G. ruber albus*, *G. tumida as G. menardii*, *Globorotalia ungulata* as either *G. tumida* or *G. menardii,* and *Globorotaloides hexagonus* as *N. dutertrei*.

In comparison to the other models tested, our "Base Cyclic 16" model was less susceptible to biases resulting from orientation, over-representation of a single species, and pigmentation (a situation where two morphologically similar species are distinguishable by color). Due to misclassification of species with low abundance, the CNN classifier underestimates the population of minor species, but this bias is noticed by the human classifier. Similar to human classifiers, the CNN classifier accuracy improves as the average species count increases. The strong correlation ($\rho = 0.97$, $p < 2.2e{-}16$) between the number of individuals counted by the CNN versus human classifiers across 21 species highlights the reliability of the data produced using our "Base Cyclic 16" neural network (Figure 2b). This is further supported by our results correlating the relative abundance data generated by the model with the data retrieved from the closest stations to the cores used in this study from the modern surface sediment in the ForCenS database (Figure 4a); this also showed significant positive correlation ($\rho = 0.77$, $p < 2.2e{-}16$). Species such as *G. ruber albus*, *G. glutinata*, and *G. inflata*, recorded the highest relative abundance in the ForCenS database at similar locations in our study (Figure 3). Although skewed toward large sizes, the relatively comparable size spectra of the Rillo et al. (2019) and the data set produced by MiSo provides confidence in the use of automated workflows such as MiSo/ParticleTrieur in the identification, classification, and measurement of planktonic foraminiferal species (Figure 4b) (Marchant et al., 2020). Our workflow and CNN model show a better accuracy (89.9%, Figure 3a) than other workflows utilizing supervised learning models in the automatic identification of planktonic foraminifera, for example, the "VGG 16" model test (27,737 training images, 6,903 validation data set) by Hsiang et al. (2019) achieved an accuracy of 87.4% when used to classify modern planktonic foraminifera images. Likewise, the "ResNet50 + VGG 16" model test (1,258 training images, 180 validation data set) by Mitra et al. (2019) achieved an accuracy of ~86% when used to classify images of modern planktonic foraminifera.

4.2. TIO Planktonic Foraminifera Mostly Exhibit Unimodal Size Frequency Distribution at the Species Level

The SFD of a species is a measure of the density of the species' size along a size spectrum within a sample (Peeters et al., 1999). In this study, about half of the species in our data set exhibit a unimodal distribution while others show flat to multimodal distributions (Figure 7, Table S4 in Supporting Information S1). The species *Neogloboquadrina pachyderma* and *Candeina nitida* show bimodality in their SFDs. SFDs in this study compare favorably with those reported in the literature, though existing data sets are sparse. For example, consistent with our findings, Peeters et al. (1999) reported unimodal SFDs for *G. bulloides*, *G. ruber albus*, *T. sacculifer,* and *G. glutinata* in Arabian Sea surface sediments. The modal sizes of these species were found to be very similar to those in our study, with values of 210 µm (*G. bulloides*), 206 µm (*G. ruber*), 299 µm (*T. sacculifer*), and 189 µm (*G. glutinata*) reported for the Arabian Sea (Peeters et al., 1999), compared to 226 µm (*G. bulloides*), 198 µm (*G. ruber*), 290 µm (*T. sacculifer*), and 193 µm (*G. glutinata*) reported in this study.

N. pachyderma and *C. nitida* showed bimodal SFDs, however, we note that these species have the lowest counts (127 and 83 specimens respectively) in the entire data set. Hence, although we have confirmed the bimodality of both species via a distribution test, we think that the observed bimodal SFDs may have also resulted from the low representation of these species in our data set, and larger populations need to be analyzed to confirm this result. Nevertheless, in a study investigating the distribution of planktonic foraminifera in the Arctic Ocean, Carstens et al. (1997) reported a weak bimodal SFD in *N. pachyderma* populations ($n = 6{,}768$). Planktonic foraminiferal SFD can be controlled by multiple processes including surface water physical and chemical properties (Naidu & Malmgren, 1995), reproductive cycle (Bijma et al., 1994), size-dependent weight, morphology (spinose vs. non-spinose) and sinking speed (Bijma et al., 1994; Takahashi & Be, 1984), dissolution (Peeters et al., 1999),

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

cryptic diversity (Huber et al., 1997), winnowing after settling on the sea floor (Peeters et al., 1999), and the tendency for the morphological characteristics of a species to vary depending on how the species reproduces (that is, sexually or asexually) (Meilland et al., 2022). Since no published water column size spectra are available to the best of our knowledge to compare our data set with living communities, we infer that the SFDs of the well-represented species reported here are representative of their regional SFDs.

4.3. Vertical Niche Partitioning May Provide a Link Between Assemblage Size and Diversity

In this study restricted to the TIO, increasing species richness does not influence assemblage $size_{95/5}$ ($R^2 = 0.013$, $p = 0.546$) (Figure 10a). This result is supported by Zarkogiannis et al. (2020), who also observed no relationship between size and species richness in core-top samples from the Eastern Mediterranean Sea. Meanwhile, planktonic foraminiferal $size_{95/5}$ has been reported to positively correlate with species richness on a global scale (Schmidt, Renaud, et al., 2004). This suggests that the size versus species richness relationship is only observable on a global scale and that at the regional to sub-regional scale, the signal of this relationship is lost because of the lower variability in prevailing hydrographic conditions.

Conversely, we found a positive relationship between assemblage $size_{95/5}$ and species diversity ($R^2 = 0.49$, $p = 6.43e{-}09$) (Figure 10b). Two models could explain this relationship. First, the diversity–productivity model suggests that diversity in planktonic foraminifera is a unimodal function of biomass and this in turn affects species' size distribution as communities try to achieve balance between food availability and selective predation (Irigoien et al., 2004). In this case, a positive relationship between species diversity and PP is expected. Secondly, the thermal stratification model suggests that vertical heterogeneity supports niche and depth-parapatric speciation, resulting in the gradual addition of new niches and larger sized species (Whittaker et al., 2001). In this case, the relationship between species diversity and thermal stratification (here we use the temperature difference between 0 and 200 m, ΔT) should be positive. Our results from the TIO lend support to the thermal stratification model rather than the diversity–productivity model because the correlation between species diversity and PP yielded a negative relationship ($\rho = -0.27$, $p = 0.03$), while its relationship with ΔT is positive ($\rho = 0.31$, $p = 0.02$). In our study, the association of higher planktonic foraminiferal diversities and larger $size_{95/5}$ at sites in oligotrophic areas (and lower diversities and smaller $size_{95/5}$ in highly productive upwelling areas), contradicts the diversity–productivity model. This is consistent with the findings of Schmidt, Renaud, et al. (2004) on a global scale. Al-Sabouni et al. (2007) also found a similar negative relationship between species diversity and PP in the Atlantic, whilst Rutherford et al. (1999) reported that changes in upper-ocean thermal structure strongly drive planktonic foraminiferal diversity. Seasonality may also play a role in explaining the size–diversity relationship in the TIO. Although the correlation between species diversity and temperature seasonality (ΔT summer–winter) reveals no relationship ($\rho = 0.08$, $p = 0.56$), PP in the TIO is highly seasonal (Longhurst, 2007), increasing the range of ecological niches over the year that species can adapt to. This warrants further investigation.

4.4. Abundance and Ecological Optima

The OSH postulates that planktonic foraminiferal species attain largest size at their ecological optimum (Hecht, 1976). Thus, if a species' abundance or relative abundance is a biotic measure of stress (Drake & Griffen, 2010), and by extension ecological optimum, then its largest size should be attained where it is most abundant and least stressed. In this study, among the 26 species where the relationship between $size_{95/5}$ and relative abundance was investigated, only *G. conglobatus* ($R^2 = 0.21$, $p = 0.013$) significantly increased in $size_{95/5}$ with increasing abundance (Figure 8). This implies that relative abundance is not a good measure of optimum growth conditions in planktonic foraminifera, or that size does not increase under optimum conditions. This result is consistent with Rillo et al. (2020), who found no relationship between $size_{95/5}$ and relative abundance in eight out of nine planktonic foraminifera species in core-top samples. Similarly, Weinkauf et al. (2019) found no relationship between the test size of large populations of *O. universa* and their abundance. Although previous studies found support for the OSH based on the relationship between species' size and abundance (Hecht, 1976; Kahn, 1981; Malmgren & Kennett, 1976; Moller et al., 2013), our results do not support the use of relative abundance as a determinant of optimum conditions, if one assumes that $size_{95/5}$ indeed does increase under optimum conditions.

The lack of relationship between $size_{95/5}$ and abundance in nearly all species could be related to adaptation strategies, or the multitude of factors influencing the (relative) abundance of a species. Under high stress conditions,

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

a species may adapt by adopting a stabilizing selection strategy that allows only the most tolerant phenotype (in this case, size ranges) to thrive (van Valen, 1965; Weinkauf et al., 2019). As a result, total abundance of the species will be reduced, but its "survival size ranges" will persist and attain high relative abundances within the species even under sub-optimal conditions and at the expense of less-tolerant phenotypes (Hecht, 1976; Rillo et al., 2020). Furthermore, a species' abundance can be influenced by many biotic (competition, predation) and abiotic (temperature, salinity, oxygen, nutrients) factors within its living environment as well as by post-depositional dissolution that could impact the representation of a species in a sample (Hecht, 1976; Schiebel & Hemleben, 2017).

4.5. Environmental Forcing of Planktonic Foraminiferal Size at the Species Level

Based on regression of temperature at maximum size with temperature at maximum relative abundance, Schmidt, Renaud, et al. (2004) proposed that mean annual SST primarily drives planktonic foraminifera size at the species level. Our results from the TIO show that SST did not primarily drive size changes in any of the species studied (see Table S6 in Supporting Information S1). Although the $size_{95/5}$ of some species showed weak but significant relationships with SST, in all cases, the coefficients of determination were markedly lower than the parameters with which they showed the best fit (Table S5 in Supporting Information S1). In the global core-top study of Rillo et al. (2020), the size (log of shell area) of nine planktonic foraminifera species showed little to no response to SST and PP. This led to the conclusion that environmental parameters cannot consistently predict planktonic foraminifera shell size. Whilst results from our study support this general conclusion, they also reveal species-specific environmental controls on the size of some of the most abundant species, including *O. universa*, *G. bulloides*, *G. ruber albus*, *G. inflata*, and *G. glutinata* (Figure 9). In Figure 9, we binned the parameters (e.g., all temperature parameters, all carbonate chemistry parameters) to identify broad environmental controls on species' size distribution. Parameters related to carbonate ion concentration explain size variation in 35% of the species while parameters related to other factors such as temperature (25%), salinity (20%), nutrient concentration (10%), and oxygenation (10%) also explain some of the size variations in individual planktonic foraminiferal species (see Table S5 in Supporting Information S1).

Results from re-analysis of the Rillo et al. (2020) data set show that three of the nine species studied (*T. sacculifer*, $R^2 = 0.22$, $p = 0.03$; *G. truncatulinoides*, $R^2 = 0.58$, $p = 0.02$; and *G. inflata*, $R^2 = 0.69$, $p = 0.03$) showed significant relationships with SST in the TIO (see Figure S4 in Supporting Information S1). In this study, *T. sacculifer* ($R^2 = 0.05$, $p = 0.03$) and *G. inflata* ($R^2 = 0.30$, $p = 0.01$) were also found to show significant relationships with SST. Whether the high coefficients of determinations of the three species reported for the Rillo et al. (2020) data set would be realized when a higher number of samples are investigated remain in doubt because the number of samples per species in our study were significantly higher than theirs (62 vs. 14 for *T. sacculifer*, 27 vs. 9 for *G. truncatulinoides,* and 25 vs. 4 for *G. inflata*). Nevertheless, we argue that the relationships observed between the $size_{95/5}$ of these three species (and the species observed to show similar relationships in our study) and SST indicate a possible secondary effect of SST on their size variations. Thus, two inferences can be drawn from the results of our study combined with those of Rillo et al. (2020). Regionally and globally, size variation in individual species is linked to multiple processes such as calcification (initial test growth), respiration, trophic strategy, and resource availability, and no single parameter can predict intra-specific size variation (Kontakiotis et al., 2021; Kuroyanagi et al., 2013; Rillo et al., 2020). In addition, while SST may affect the rates of the aforementioned processes differently among species (e.g., Lombard et al., 2009; Weinkauf et al., 2016), in tropical environments with predominantly warm SSTs, the net impact of SST on intra-specific size variation does not outweigh the impact of other physiochemical parameters. By inference, SST is not a suitable parameter for testing the OSH at the species level.

In our data set, the identified drivers of size in some species agree with described ecological and environmental preferences (Figure 9). For instance, phosphate concentration (an indicator of the nutritional landscape or "water fertility") was found to be a significant driver of size in *G. bulloides* ($R^2 = 0.31$, $p = 2.4e{-}07$) and *G. glutinata* ($R^2 = 0.19$, $p = 2.94e{-}05$), and these species are known to thrive in nutrient-rich and upwelling environments (Bé & Hutson, 1977; Conan & Brummer, 2000; Ivanova et al., 2003; Kroon & Ganssen, 1988; Schiebel & Hemleben, 2017; Seears et al., 2012). Interestingly, calcification in *G. bulloides* in the North Atlantic has also been reported to be linked to phosphate concentration (Aldridge et al., 2012). Therefore, phosphate concentration may be key to providing optimum conditions for these species. Carbonate ion concentration was found to

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

significantly drive size variation in sub-surface species which bear gametogenic calcite such as *G. conglobatus* ($R^2 = 0.19$, $p = 0.01$), *G. tumida* ($R^2 = 0.29$, $p = 0.001$), and *P. obliquiloculata* ($R^2 = 0.26$, $p = 0.047$). On the other hand, tropical to sub-tropical species such as *G. ruber albus* and *T. sacculifer*, which are surface dwellers with a preference for warm water masses (Bé, 1977), showed no relationship with any temperature-related parameter. Using a small data set from the Caribbean Sea, Stone (1956) reported that test size in *O. universa* showed no relationship with SST while test size in *G. ruber albus* is weakly sensitive to SST compared to *T. sacculifer* which showed a higher sensitivity. In this study, similarly, $size_{95/5}$ variation in *O. universa* showed no sensitivity to SST ($R^2 = 0.06$, $p = 0.06$) and *T. sacculifer* showed more sensitivity to SST ($R^2 = 0.05$, $p = 0.03$), albeit weak. Meanwhile, *G. ruber albus* $size_{95/5}$ is not impacted by SST changes ($R^2 = 0.02$, $p = 0.15$) in this study, and its size related more to oxygenation. Our results suggest that, in the TIO, $size_{95/5}$ variation in *O. universa* is in part driven by sub-surface temperatures (at 100 m water depth; $R^2 = 0.20$, $p = 1.2e{-}05$) while $size_{95/5}$ in *T. sacculifer* is influenced by carbonate concentration ($R^2 = 0.10$, $p = 0.01$).

Culture experiments on two subtropical–transitional planktonic foraminifer species (*O. universa* and *G. bulloides*) show increased shell length under increasing dissolved oxygen conditions (Kuroyanagi et al., 2013). Symbiont-bearing *O. universa* not only increased in length but also in pore diameter and size under varying dissolved oxygen concentrations and low temperature variability. This may explain the positive correlation we found between size and oxygenation in *G. ruber albus* (also a symbiont-bearing species like *O. universa*) and we speculate that it may imply that under low temperature variability, growth in *G. ruber albus* could be linked to increased gas exchange even though some of its nutrients are derived from organic carbon supply due to enhanced photosynthetic activity of its symbionts (Jørgensen et al., 1985; LeKieffre et al., 2018).

Broadly, our results suggest that environmental parameters have limited but significant impacts on species-specific geographical size distribution in TIO planktonic foraminifera. Building upon the conclusions of Rillo et al. (2020), we infer that no single environmental parameter can consistently predict intra-specific size variations in planktonic foraminifera, but that specific environmental parameters might define each species' ecological optimum, supporting the OSH.

4.6. Primary and Secondary Environmental Controls on Assemblage Size Trends

4.6.1. Surface Carbonate Ion Concentration Exerts Primary Controls on Assemblage Size Trends

In a previous global study, SST was reported to primarily drive size changes in planktonic foraminifera assemblages (Schmidt, Renaud, et al., 2004). Our investigation of this hypothesis suggests that SST does not primarily drive assemblage size changes regionally in the TIO. The *F*-matrix loadings from the factor analysis on planktonic foraminifera assemblage size shows characteristic faunal aggregation on the first factor axis. The high weighting and aggregation of tropical and sub-tropical species such as *T. sacculifer*, *P. obliquiloculata*, *G. menardii*, *G. tumida*, *G. elongatus*, *N. dutertrei*, and *O. universa*, along Factor axis 1 of the assemblage size data makes it difficult to decipher a clear temperature-related size distribution (see Table S7 in Supporting Information S1). Likewise, the low weighting on species with preferences for nutrient rich environments (upwelling zones) such as *G. glutinata*, *G. falconensis,* and *G. bulloides* on this axis also rules out the possibility that this axis is related to resource availability. However, notably, large-sized and relatively more resistant species as defined by Berger (1970), have higher loadings on this axis. We note that the Berger (1970) solubility ranking included most species but not all. In addition, some of the species with high positive factor scores such as *P. obliquiloculata*, *G. conglobatus*, *G. menardii*, and *G. tumida* bear gametogenic calcite. Robust regression analysis between this first factor axis and hydrographic parameters revealed a best fit with surface carbonate ion concentration ($R^2 = 0.41$, $p = 1.18e{-}07$) (Figure 12a), consistent with the importance of surface carbonate ion concentration in the TIO revealed by our PCA of surface hydrographic parameters (Figure S1 in Supporting Information S1, Table 2). The weighting of different species on this axis might reflect the different sensitivities of thermocline/sub-thermocline species versus mixed layer species to carbonate ion concentration.

Our results imply that, at the assemblage level, upper water column carbonate ion concentration is significantly inversely correlated with assemblage $size_{95/5}$ in TIO planktonic foraminifera. While this result may suggest a possible influence of carbonate ion in driving growth and indirectly calcification at the assemblage level, the use of assemblage size as a proxy of calcification is not straightforward because it integrates all single-species responses, and their full life histories. More so, carbonate ion has only so far been reported to influence calcification at the

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

species level (Barker and Elderfield, 2002; de Moel et al., 2009; Henehan et al., 2017; Naik et al., 2011; Weinkauf et al., 2013), and it remains to be tested whether the same observation will apply at the assemblage level. Nevertheless, if one posits that the assemblage $size_{95/5}$ signal were linked to community calcification, our result yields a negative relationship compared to the positive relationship observed in foraminifera at the species level (Barker & Elderfield, 2002; Bijma et al., 1999, 2002). We propose two possible explanations for this observation. First, it might be that in assemblages in more supersaturated waters in the TIO, calcification carries less energetic costs to the foraminifera (i.e., reduced energetic costs of elevating pH at the calcification site) such that species (a) produce their full test more quickly and (b) have more energy to spend on other metabolic processes (e.g., reproduction); by reserving energy for such an activity, species complete their life or reproductive cycle more efficiently and quickly, but do so within the lower limits of their survival size ranges. This hypothesis cannot be tested using our data set, as further data including size-normalized shell weight or CT-scans of multiple species would be needed to fully understand the link between size, growth and calcification in each assemblage.

Second, the negative size to carbonate ion relationship observed may be linked to differing carbonate production patterns between upwelling and more oligotrophic environments. Upwelling environments have been reported to be characterized by high carbonate fluxes, lower surface carbonate ion concentration, and a dominance of small-sized foraminiferal species (e.g., *G. glutinata*, *G. bulloides,* and *N. incompta*) (Peeters & Brummer, 2002; Ramaswamy & Nair, 1994; Thunell et al., 2007). Conversely, oligotrophic environments are generally characterized by lower carbonate fluxes, higher surface carbonate ion concentration, and high relative abundance of larger-sized species (e.g., *O. universa*, *T. sacculifer,* and *G. siphonifera*) (Peeters & Brummer, 2002; Ramaswamy & Nair, 1994; Schiebel & Hemleben, 2017). We test this hypothesis through correlation and "correlation cluster" analyses to examine the relationship between the relative abundances of small-sized upwelling species and large-sized oligotrophic species versus surface carbonate ion concentration in the TIO (Figure S6 in Supporting Information S1). In our full data set, we find that higher relative abundances of small-sized upwelling species occur under higher surface carbonate ion concentration and that higher relative abundances of large-sized oligotrophic species occur under lower surface carbonate condition (Figures S6a and S6b in Supporting Information S1). This leads to an overall lower assemblage $size_{95/5}$ under high surface carbonate condition and a higher assemblage $size_{95/5}$ under low surface carbonate ion conditions. However, this result is sensitive to the inclusion of the Arabian Sea data set, and its exclusion leads to opposite results (i.e., high relative abundances of small-sized species under low surface carbonate ion conditions and vice versa) (Figures S6c–S6d and S7 in Supporting Information S1). Overall, in the TIO, our results show that surface carbonate ion concentration exerts primary control on assemblage $size_{95/5}$ variations, and we suggest that this response might be due to energetic optimization and/or changes in the dominance of the main species linked to nutrient regimes, though this pattern remains to be confirmed in other ocean basins.

4.6.2. SST Exerts Secondary Controls on Assemblage Size Trends

A temperature effect on assemblage size might be expected for different reasons. First, previous studies have shown that test size decreases from the tropics toward the poles (de Villiers, 2004; Renaud & Schmidt, 2003). This implies that planktonic foraminifera assemblages are sensitive to temperature changes and as communities shift from the poles toward the equator, assemblage $size_{95/5}$ increases. Support for this is found in Schmidt, Renaud, et al. (2004) who suggested that SST primarily drives assemblage size changes in planktonic foraminifera on a global scale (Schmidt, Renaud, et al., 2004). Second, in fossil sediments and plankton net samples, temperature has been reported to influence shell calcification (shell weight and SNW) in planktonic foraminifera (Aldridge et al., 2012; Gonzalez-Mora et al., 2008). Third, the second axis of the PCA analysis on surface hydrographic parameters show that SST is an important hydrographic parameter in the TIO after surface carbonate ion concentration. We test this assumption by correlating our *F*2 scores with hydrographic parameters.

On the second factor axis of the *F*-matrix loading of the assemblage size factor analysis, unlike the first factor axis, a clear niche partitioning is observed. Factor 2 is characterized by deep-dwellers including *G. inflata*, *G. crassaformis*, and *G. truncatulinoides* (see Table S7 in Supporting Information S1), suggesting a possible relationship with a temperature-related parameter based on the common thermal niche preference of these species. Assemblage size *F*2 scores indeed showed the best relationship with SST ($R^2 = 0.46$, $p = 1.24e{-}08$) among the environmental variables tested (Figure 12b). In agreement with previous studies, our result shows that assemblage size increased with temperature (Schmidt, Renaud, et al., 2004). The influence of SST may be more pronounced on assemblage $size_{95/5}$ and not species' $size_{95/5}$ because the aggregate net effect of SST on assemblage calcification

processes becomes more evident when species composition is heterogenous. Therefore, based on our results, we suggest that ambient carbonate ion concentration primarily influences assemblage $size_{95/5}$, but temperature plays a significant role in modulating the ontogenic calcification process. The influence of both carbonate ion concentration and temperature on assemblage $size_{95/5}$ (and not species' $size_{95/5}$) has not been reported previously, thus our results constitute the first in this regard. Yet, while most species-specific studies on the factors that influence shell calcification have shown that carbonate ion concentration drives shell calcification (Barker & Elderfield, 2002; Bijma et al., 2002; Lombard et al., 2010; Spero et al., 1997), other studies have shown that shell calcification can be attributed to the influence of carbonate ion concentration and temperature (Aldridge et al., 2012; Gonzalez-Mora et al., 2008); our results support the latter. Therefore, we conclude that SST plays a secondary but significant role in driving planktonic foraminifera assemblage $size_{95/5}$ trends in the TIO.

5. Summary and Conclusions

In this study, through automated foraminiferal analysis, we test the OSH to determine whether relative abundance or environmental parameters are better indicators of optimum conditions in TIO planktonic foraminiferal species. We also investigate the dominant environmental controls on TIO planktonic foraminifera assemblage $size_{95/5}$ distribution. Combined, our high-throughput imaging machine as well as our CNN-trained model achieved classification and measurement accuracies similar to human experts. Although most species show unimodal SFD, we report flat to multimodal SFDs in larger species such as *G. menardii*, *G. conglomerata*, and *G. tumida.* We infer that vertical niche separation explains the relationship between planktonic foraminiferal species' size and diversity. Using environmental parameters rather than relative abundances as indicators of optimum conditions in planktonic foraminiferal species, we found support for the OSH. However, we report that the environmental drivers of size in planktonic foraminifera are species-specific, with parameters related to carbonate ion concentration and temperature exerting significant controls on species-level $size_{95/5}$ variations. Therefore, no single parameter can predict intra-specific size variation in planktonic foraminifera, as suggested by Rillo et al. (2020). Further, in contrast to previous reports that planktonic foraminifera assemblage size is primarily driven by temperature, results from this study suggest that, within tropical oceans where surface ocean temperature variability is restricted, carbonate ion concentration exerts the strongest control on planktonic foraminifera size variations, although temperature was found to exert a secondary control on assemblage $size_{95/5}$ variation. In support of previous work (Cayre et al., 1999) and as a validation of our automated method, we found that PP exerts primary control on TIO species assemblage distribution, but temperature was found to be a secondary driver of species assemblage distribution in the TIO. Although it was found to be a secondary driver, temperature influence on planktonic foraminiferal assemblage size variations and assemblage distribution is significant. Because of multicollinearity and complexity in relationships between environmental variables, we conclude that the isolated drivers of $size_{95/5}$ variation identified in this study act as reasonable proxies for a larger suite of drivers. Overall, if our findings apply to other tropical ocean planktonic foraminifera populations and communities, changes in planktonic foraminiferal test size variations are mostly linked to the effects of parameters related to carbonate ion and temperature.

Data Availability Statement

Data sets associated with this paper are available in the SEANOE open-access database at https://doi.org/10.17882/86211 (Adebayo, Bolton, Marchant, Bassinot, Sandrine, et al., 2022) and https://doi.org/10.17882/86411 (Adebayo, Bolton, Marchant, Bassinot, Conrod, et al., 2022), published under a CC-BY-NC license. The scripts for the code used in this work can be accessed at https://github.com/mikolinton. The *ParticleTrieur* software used in this work is available with a tutorial from http://particle-classification.readthedocs.io, and the Python scripts to train the CNNs in this paper are available from http://www.github.com/microfossil/particle-classification and can be installed using the pip Python package.

Acknowledgments
We acknowledge the Flotte Océanographique Française for recovery and access to all Marion-Dufresne samples. The authors thank Jean Charles Mazur for laboratory assistance. This project is funded by an AMU President's Scholarship, the CNRS-INSU LEFE IndSO Grant, ANR Grants FIRST (ANR-15-CE4-0006-01) and iMonsoon (ANR-16-CE01-0004-01).

References

Adebayo, M., Bolton, C., Marchant, R., Bassinot, F., Conrod, S., & De-Garidel Thoron, T. (2022). Tropical Indian ocean annotated planktonic foraminifera image dataset from surface sediments [Dataset]. SEANOE. https://doi.org/10.17882/86411

Adebayo, M., Bolton, C., Marchant, R., Bassinot, F., Sandrine, C., & De-Garidel Thoron, T. (2022). Recent planktonic foraminifera size and environmental data from the tropical Indian Ocean [Dataset]. SEANOE. https://doi.org/10.17882/86211

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Aldridge, D., Beer, C. J., & Purdie, D. A. (2012). Calcification in the planktonic foraminifera *Globigerina bulloides* linked to phosphate concentrations in surface waters of the North Atlantic Ocean. *Biogeosciences*, *9*(5), 1725–1739. https://doi.org/10.5194/bg-9-1725-2012

Al-Sabouni, N., Kucera, M., & Schmidt, D. N. (2007). Vertical niche separation control of diversity and size disparity in planktonic foraminifera. *Marine Micropaleontology*, *63*(1–2), 75–90. https://doi.org/10.1016/j.marmicro.2006.11.002

Arendt, J. (2007). Ecological correlates of body size in relation to cell size and cell number: Patterns in flies, fish, fruits and foliage. *Biological Reviews*, *82*(2), 241–256. https://doi.org/10.1111/j.1469-185x.2007.00013.x

Ashton, K. G., & Feldman, C. R. (2003). Bergmann's rule in nonavian reptiles: Turtles follow it, lizards and snakes reverse it. *Evolution*, *57*(5), 1151–1163. https://doi.org/10.1554/0014-3820(2003)057[1151:brinrt]2.0.co;2

Atkinson, D. (1994). Temperature and organism size: A biological law for ectotherms? *Advances in Ecolological Research*, *25*, 1–58.

Barber, R. T., Marra, J., Bidigare, R. C., Codispoti, L. A., Halpern, D., Johnson, Z., et al. (2001). Primary productivity and its regulation in the Arabian Sea during 1995. *Deep Sea Research Part II: Topical Studies in Oceanography*, *48*(6–7), 1127–1172. https://doi.org/10.1016/s0967-0645(00)00134-x

Barker, S., & Elderfield, H. (2002). Foraminiferal calcification response to glacial–interglacial changes in atmospheric CO_2. *Science*, *297*(5582), 833–836. https://doi.org/10.1126/science.1072815

Bé, A. W. H. (1977). An ecological, zoogeographic and taxonomic review of recent planktonic foraminifera. In A. T. S. Ramsay (Ed.), *Oceanic micropaleontology* (pp. 1–100). Academic Press.

Bé, A. W. H., Caron, D. A., & Anderson, O. R. (1981). Effects of feeding frequency on life processes of the planktonic foraminifer Globigerinoides sacculifer in laboratory culture. *Journal of the Marine Biological Association of the United Kingdom*, *61*(1), 257–277. https://doi.org/10.1017/S002531540004604X

Bé, A. W. H., Hutson, W. H., & Be, A. W. H. (1977). Ecology of planktonic foraminifera and biogeographic patterns of life and fossil assemblages in the Indian Ocean. *Micropaleontology*, *23*(4), 369–414. https://doi.org/10.2307/1485406

Beer, C. J., Schiebel, R., & Wilson, P. A. (2010). Testing planktic foraminiferal shell weight as a surface water [CO_3^{2-}] proxy using plankton net samples. *Geology*, *38*(2), 103–106. https://doi.org/10.1130/G30150.1

Berger, W. H. (1970). Planktonic foraminifera: Selective solution and the lysocline. *Marine Geology*, *8*(2), 111–138. https://doi.org/10.1016/0025-3227(70)90001-0

Bergmann, C. (1847). Über die Verhältnisse der Wärmeökonomie der Thiere zu ihrer Grösse. *Gottinger Studien*, *3*, 595–708.

Berner, R. A., & Raiswell, R. (1983). Burial of organic carbon and pyrite sulfur in sediments over phanerozoic time: A new theory. *Geochimica et Cosmochimica Acta*, *47*(5), 855–862. https://doi.org/10.1016/0016-7037(83)90151-5

Bijma, J., Faber, W. W., & Hemleben, C. (1990). Temperature and salinity limits for growth and survival of some planktonic foraminifers in laboratory cultures. *Journal of Foraminiferal Research*, *20*(2), 95–116. https://doi.org/10.2113/gsjfr.20.2.95

Bijma, J., Hemleben, C., & Wellnitz, K. (1994). Lunar-influenced carbonate flux of the planktic foraminifer *Globigerinoides sacculifer* (Brady) from the central Red Sea. *Deep Sea Research Part I: Oceanographic Research Papers*, *41*(3), 511–530. https://doi.org/10.1016/0967-0637(94)90093-0

Bijma, J., Hönisch, B., & Zeebe, R. E. (2002). Impact of the ocean carbonate chemistry on living foraminiferal shell weight: Comment on "Carbonate ion concentration in glacial-age deep waters of the Caribbean Sea" by W. S. Broecker and E. Clark. *Geochemistry, Geophysics, Geosystems*, *3*(11), 1064–1071. https://doi.org/10.1029/2002GC000388

Bijma, J., Spero, H. J., & Lea, D. W. (1999). Reassessing foraminiferal stable isotope geochemistry: Impact of the oceanic carbonate system. In G. Fisher & G. Wefer (Eds.), *Use of proxies in paleoceanography: Examples of the South Atlantic* (pp. 489–512). Springer-Verlag.

Bindoff, N. L., Cheung, W. W. L., Kairo, J. G., Arístegui, J., Guinder, V. A., Hallberg, R., et al. (2019). Changing ocean, marine ecosystems, and dependent communities. In H. O. Pörtner, D. C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, et al. (Eds.), *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate* (pp. 447–587). Cambridge University Press. https://doi.org/10.1017/9781009157964.007

Broecker, W. S., & Sutherland, S. (2000). Distribution of carbonate ion in the deep ocean: Support for a post-Little Ice Age change in Southern Ocean ventilation? *Geochemistry, Geophysics, Geosystems*, *1*(7), 1023. https://doi.org/10.1029/2000GC000039

Brombacher, A., Elder, L. E., Hull, P. M., Wilson, P. A., & Ezard, T. H. G. (2018). Calibration of test diameter and area as proxies for body size in the planktonic foraminifer *Globoconella Puncticula*ta. *Journal of Foraminiferal Research*, *48*(3), 241–245. https://doi.org/10.2113/gsjfr.48.3.241

Brombacher, A., Wilson, P. A., & Ezard, T. H. G. (2017). Calibration of the repeatability of foraminiferal test size and shape measures with recommendations for future use. *Marine Micropaleontology*, *133*, 21–27. https://doi.org/10.1016/j.marmicro.2017.05.003

Brown, J. H., Allen, A. P., & Gillooly, J. F. (2007). The metabolic theory of ecology and the role of body size in marine and freshwater ecosystems. In A. G. Hildrew, D. G. Raffaelli, & R. Edmonds-Brown (Eds.), *Body size: The structure and function of aquatic ecosystems* (pp. 1–15). Cambridge University Press. https://doi.org/10.1017/CBO9780511611223.002

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

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., & Anderson, O. R. (1982). Effects of variations in light intensity on life processes of the planktonic foraminifer *Globigerinoides sacculifer* in laboratory culture. *Journal of the Marine Biological Association of the United Kingdom*, *62*(2), 435–451. https://doi.org/10.1017/S0025315400057374

Carstens, J., Hebbeln, D., & Wefer, G. (1997). Distribution of planktic foraminifera at the ice margin in the Arctic (Fram Strait). *Marine Micropaleontology*, *29*(3–4), 257–269. https://doi.org/10.1016/s0377-8398(96)00014-x

Cayre, O., Beaufort, L., & Vincent, E. (1999). Paleoproductivity in the Equatorial Indian Ocean for the last 260,000 yr: A transfer function based on planktonic foraminifera. *Quaternary Science Reviews*, *18*(6), 839–857. https://doi.org/10.1016/S0277-3791(98)00036-5

Clemens, S. C., Murray, D. W., & Prell, W. L. (1996). Nonstationary phase of the Plio-Pleistocene Asian monsoon. *Science*, *274*(5289), 943–948. https://doi.org/10.1126/science.274.5289.943

Colombo, M. R., & Cita, M. B. (1980). Changes in size and test porosity of *Orbulina universa* d'Orbigny in the Pleistocene record of Cape Bojador (DSDP Site 397, Eastern North Atlantic). *Marine Micropaleontology*, *5*, 13–29. https://doi.org/10.1016/0377-8398(80)90004-3

Conan, S. M. H., & Brummer, G. J. A. (2000). Fluxes of planktic foraminifera in response to monsoonal upwelling on the Somalia Basin margin. *Deep-Sea Research II*, *47*(9–11), 2207–2227. https://doi.org/10.1016/s0967-0645(00)00022-9

de Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., & Iudicone, D. (2004). Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology [Dataset]. Journal of Geophysical Research, *109*(C12), C12003. https://doi.org/10.1029/2004JC002378

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

de Moel, H., Ganssen, G. M., Peeters, F. J. C., Jung, S. J. A., Kroon, D., Brummer, G. J. A., & Zeebe, R. E. (2009). Planktic foraminiferal shell thinning in the Arabian Sea due to anthropogenic ocean acidification? *Biogeosciences*, *6*(9), 1917–1925. https://doi.org/10.5194/bg-6-1917-200

de Villiers, S. (2004). Optimum growth conditions as opposed to calcite saturation as a control on the calcification rate and shell-weight of marine foraminifera. *Marine Biology*, *144*(1), 45–49. https://doi.org/10.1371/journal.pone.0148363

Drake, J. M., & Griffen, B. D. (2010). Early warning signals of extinction in deteriorating environments. *Nature*, *467*(7314), 456–459. https://doi.org/10.1038/nature09389

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

Fan, L., Cai, T., Xiong, Y., Song, G., & Lei, F. (2019). Bergmann's rule and Allen's rule in two passerine birds in China. *Avian Research*, *10*(1), 34. https://doi.org/10.1186/s40657-019-0172-7

Fox-Kemper, B., Hewitt, H. T., Xiao, C., Aðalgeirsdóttir, G., Drijfhout, S. S., Edwards, T. L., et al. (2021). Ocean, cryosphere and sea level change. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, et al. (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 1211–1362). Cambridge University Press. https://doi.org/10.1017/9781009157896.011

Gao, K., Xu, J., Gao, G., Li, Y., Hutchins, D. A., Huang, B., et al. (2012). Rising CO_2 and increased light exposure synergistically reduce marine primary productivity. *Nature Climate Change*, *2*(7), 519–523. https://doi.org/10.1038/nclimate1507

Garcia, H. E., Weathers, K., Paver, C. R., Smolyar, I., Boyer, T. P., Locarnini, R. A., et al. (2018a). Dissolved Inorganic Nutrients (phosphate, nitrate and nitrate+nitrite, silicate). In A. Mishonov (Ed.), (Technical Ed.), *World Ocean Atlas 2018* (Vol. 4, p. 35). *NOAA Atlas NESDIS 84*.

Garcia, H. E., Weathers, K., Paver, C. R., Smolyar, I., Boyer, T. P., Locarnini, R. A., et al. (2018b). Dissolved oxygen, apparent oxygen utilization, and oxygen saturation. In A. Mishonov (Ed.), (Technical Ed.), *World Ocean Atlas 2018* (Vol. 3, p. 38). *NOAA Atlas NESDIS 83*.

Gonzalez-Mora, B., Sierro, F. J., & Flores, J. A. (2008). Controls of shell calcification in planktonic foraminifers. *Quaternary Science Reviews*, *27*(9–10), 956–961. https://doi.org/10.1016/j.quascirev.2008.01.008

Gruber, N. (2011). Warming up, turning sour, losing breath: Ocean biogeochemistry under global change. *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, *369*(1943), 1980–1996. https://doi.org/10.1098/rsta.2011.0003

Hammer, O., Harper, D. A. T., & Ryan, P. D. (2001). PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica*, *4*(1), 1–9.

Hart, S. P., Schreiber, S. J., & Levine, J. M. (2016). How variation between individuals affects species coexistence. *Ecology Letters*, *19*(8), 825–838. https://doi.org/10.1111/ele.12618

He, J., & Mahadevan, A. (2021). How the source depth of coastal upwelling relates to stratification and wind. *Journal of Geophysical Research: Oceans*, *126*(12), e2021JC017621. https://doi.org/10.1029/2021JC017621

Hecht, A. D. (1974). Intraspecific variation in recent populations of *Globigerinoides ruber* and *Globigerinoides trilobus* and their application to paleoenvironmental analysis. *Journal of Paleontology*, *48*(6), 1217–1234.

Hecht, A. D. (1976). An ecologic model for test size variation in Recent planktonic foraminifera; applications to the fossil record. *Journal of Foraminiferal Research*, *6*(4), 295–311. https://doi.org/10.2113/gsjfr.6.4.295

Hemleben, C., Spindler, M., Breitinger, I., & Ott, R. (1987). Morphological and physiological responses of *Globigerinoides sacculifer* (Brady) under varying laboratory conditions. *Marine Micropaleontology*, *12*, 305–324. https://doi.org/10.1016/0377-8398(87)90025-9

Henehan, M. J., Evans, D., Shankle, M., Burke, J. E., Foster, G. L., Anagnostou, E., et al. (2017). Size-dependent response of foraminiferal calcification to seawater carbonate chemistry. *Biogeosciences*, *14*(13), 3287–3308. https://doi.org/10.5194/bg-14-3287-2017

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*, 65–70.

Hsiang, A. Y., Brombacher, A., Rillo, M. C., Mleneck-Vautravers, M. J., Conn, S., Lordsmith, S., et al. (2019). Endless Forams: >34,000 modern planktonic foraminiferal images for taxonomic training and automated species recognition using convolutional neural networks. *Paleoceanography and Paleoclimatology*, *34*(7), 1157–1177. https://doi.org/10.1029/2019PA003612

Huber, B. T., Bijma, J., & Darling, K. F. (1997). Cryptic speciation in the living planktonicforaminifer *Globigerinella siphonifera* (d'Orbigny). *Paleobiology*, *23*(1), 33–62. https://doi.org/10.1017/s0094837300016638

Irigoien, X., Huisman, J., & Harris, R. P. (2004). Global biodiversity patterns of marine phytoplankton and zooplankton. *Nature*, *429*(6994), 863–867. https://doi.org/10.1038/nature02593

Ivanova, E. M., Schiebel, R., Singh, A. D., Schmiedl, G., Niebler, H. S., & Hemleben, C. (2003). Primary production in the Arabian Sea during the last 135,000 years. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *197*(1–2), 61–82. https://doi.org/10.1016/s0031-0182(03)00386-9

Jarque, C. M. (2011). Jarque-Bera test. In M. Lovric (Ed.), *International Encyclopedia of Statistical Science*. Springer. Retrieved from https://doi-org.lama.univ-amu.fr/10.1007/978-3-642-04898-2_319

Jørgensen, B. B., Erez, J., Revsbech, P., & Cohen, Y. (1985). Symbiotic photosynthesis in a planktonic foraminiferan, *Globigerinoides sacculifer* (Brady), studied with microelectrodes. *Limnology & Oceanography*, *30*(6), 1253–1267. https://doi.org/10.4319/lo.1985.30.6.1253

Kahn, M. I. (1981). Ecological and paleoecological implications of the phenotypic variation in three species of living planktonic foraminifera from the northeastern Pacific Ocean (50°N, 145°W). *Journal of Foraminiferal Research*, *11*(3), 203–211. https://doi.org/10.2113/gsjfr.11.3.203

Kennett, J. P. (1976). Phenotypic variation in some recent and Late Cenozoic planktonic foraminifera. In R. H. Hedley & C. G. Adams (Eds.), *Foraminifera* (Vol. 2, pp. 111–170). Academic Press.

Key, R., Olsen, A., van Heuven, S., Lauvset, S. K., Velo, A., Lin, X., et al. (2015). *Global ocean data analysis project, version 2 (GLODAPv2), ORNL/CDIAC-162, ND-P093*. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy. https://doi.org/10.3334/CDIAC/OTG.NDP093_GLODAPv2

Kontakiotis, G., Efstathiou, E., Zarkogiannis, S. D., Besiou, E., & Antonarakou, A. (2021). Latitudinal differentiation among modern planktonic foraminiferal populations of central Mediterranean: Species–specific distribution patterns and size variability. *Journal of Marine Science and Engineering*, *9*(5), 551. https://doi.org/10.3390/jmse9050551

Kroon, D., & Ganssen, G. (1988). Northern Indian Ocean upwelling cells and the stable isotope composition of living planktic foraminifera. In G. J. A. Brummer & D. Kroon (Eds.), *Planktonic foraminifers as tracers of ocean-climate history* (pp. 219–238). Free University Press.

Kucera, M. (2007). Chapter six planktonic foraminifera as tracers of past oceanic environments. *Developments in Marine Geology*, *1*, 213–262.

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

Kuroyanagi, A., da Rocha, R. E., Bijma, J., Spero, H. J., Russell, A. D., Eggins, S. M., & Kawahata, H. (2013). Effect of dissolved oxygen concentration on planktonic foraminifera through laboratory culture experiments and implications for oceanic anoxic events. *Marine Micropaleontology*, *101*, 28–32. https://doi.org/10.1016/j.marmicro.2013.04.005

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

LeKieffre, C., Spero, H. J., Russell, A. D., Fehrenbacher, J. S., Geslin, E., & Meibom, A. (2018). Assimilation, translocation, and utilization of carbon between photosynthetic symbiotic dinoflagellates and their planktic foraminifera host. *Marine Biology*, *165*(6), 1–15. https://doi.org/10.1007/s00227-018-3362-7

Lessa, D., Morard, R., Jonkers, L., Venancio, I. M., Reuter, R., Baumeister, A., et al. (2020). Distribution of planktonic foraminifera in the Subtropical South Atlantic: Depth hierarchy of controlling factors. *Biogeosciences*, *17*(16), 4313–4342. https://doi.org/10.5194/bg-17-4313-2020

Levitus, S. (1982). *Climatological atlas of the world oceans*. In *NOAA Professional Paper 13* (p. 173). US. government Printing office.

Locarnini, R. A., Mishonov, A. V., Baranova, O. K., Boyer, T. P., Zweng, M. M., García, H. E., et al. (2018). Temperature. In A. Mishonov (Ed.), (Technical Ed.), *World Ocean Atlas 2018* (Vol. 1, p. 52). *NOAA Atlas NESDIS 81*.

Lombard, F., da Rocha, R. E., Bijma, J., & Gattuso, J. P. (2010). Effect of carbonate ion concentration and irradiance on calcification in planktonic foraminifera. *Biogeosciences*, *7*(1), 247–255. https://doi.org/10.5194/bg-7-247-2010

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). *Ecological geography of the sea* (2nd ed., p. 542). Academic Press.

Lovenduski, N. S., Long, L., & Lindsay, K. (2015). Natural variability in the surface ocean carbonate ion concentration. *Biogeosciences*, *12*(21), 6321–6335. https://doi.org/10.5194/bg-12-6321-2015

Malmgren, B. A., & Kennett, J. P. (1976). Size variations in *Globigerina bulloides* d'Orbigny as a Quaternary paleoclimatic index in the Southern Ocean. *Antarctic Journal*, *113*, 177–178.

Malmgren, B. A., & Kennett, J. P. (1977). Biometric differentiation between recent *Globigerina bulloides* and *Globigerina falconensis* in the southern Indian Ocean. *Journal of Foraminiferal Research*, *7*(2), 131–148. https://doi.org/10.2113/gsjfr.7.2.131

Manno, C., Morata, N., & Bellerby, R. (2012). Effect of ocean acidification and temperature increase on the planktonic foraminifer *Neogloboquadrina pachyderma* (sinistral). *Polar Biology*, *35*(9), 1311–1319. https://doi.org/10.1007/s00300-012-1174-7

Marchant, R., Tetard, M., Pratiwi, A., Adebayo, M., & de Garidel-Thoron, T. (2020). Automated analysis of foraminifera fossil records by image classification using a convolutional neural network. *Journal of Micropalaeontology*, *39*(2), 183–202. https://doi.org/10.5194/jm-39-183-2020

Meilland, J., Ezat, M. M., Westgard, A., Manno, C., Morard, R., Siccha, M., & Kucera, M. (2022). Rare but persistent asexual reproduction explains the success of planktonic foraminifera in polar oceans. *Journal of Plankton Research*, *45*(1), 15–32. https://doi.org/10.1093/plankt/fbac069

Mignot, J., de Boyer Montégut, C., Lazar, A., & Cravatte, S. (2007). Control of salinity on the mixed layer depth in the world ocean [Dataset]. Journal of Geophysical Research, *112*(C10), C10010. https://doi.org/10.1029/2006JC003954

Millien, V., Kathleen Lyons, S., Olson, L., Smith, F. A., Wilson, A. B., & Yom-Tov, Y. (2006). Ecotypic variation in the context of global climate change: Revisiting the rules. *Ecology Letters*, *9*(7), 853–869. https://doi.org/10.1111/j.1461-0248.2006.00928.x

Mitra, R., Marchitto, T. M., Ge, Q., Zhong, B., Kanakiya, B., Cook, M. S., et al. (2019). Automated species-level identification of planktic foraminifera using convolutional neural networks, with comparison to human performance. *Marine Micropaleontology*, *147*, 16–24. https://doi.org/10.1016/j.marmicro.2019.01.005

Moller, T., Schulz, H., & Kucera, M. (2013). The effect of sea surface properties on shell morphology and size of the planktonic foraminifer *Neogloboquadrina pachyderma* in the North Atlantic. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *391*, 34–48. https://doi.org/10.1016/j.palaeo.2011.08.014

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

Naidu, P. D., & Malmgren, B. A. (1995). Monsoon upwelling effects on test size of some planktonic foraminiferal species from the Oman Margin, Arabian Sea. *Paleoceanography*, *10*(1), 117–122. https://doi.org/10.1029/94PA02682

Naik, S. S., Godad, S. P., & Naidu, P. D. (2011). Does carbonate ion control planktonic foraminifera shell calcification in upwelling regions? *Current Science*, *101*(10), 1370–1375.

Ollala-Tárraga, M. Á., Rodríguez, M. Á., & Hawkins, B. A. (2006). Broad-scale patterns of body size in squamate reptiles of Europe and North America. *Journal of Biogeography*, *33*(5), 781–793. https://doi.org/10.1111/j.1365-2699.2006.01435.x

Olsen, A., Key, R. M., van Heuven, S., Lauvset, S. K., Velo, A., Lin, X., et al. (2016). The Global Ocean Data Analysis Project version 2 (GLODAPv2) – An internally consistent data product for the world ocean [Dataset]. Earth System Science Data, *8*, 297–323. https://doi.org/10.5194/essd-8-297-2016

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

Peeters, F., Ivanova, E., Conan, S., Brummer, G. J., Ganssen, G., Troelstra, S., & van Hinte, J. (1999). A size analysis of planktic foraminifera from the Arabian Sea. *Marine Micropaleontology*, *36*(1), 31–63. https://doi.org/10.1016/S0377-8398(98)00026-7

Peeters, F. J., & Brummer, G. J. A. (2002). The seasonal and vertical distribution of living planktic foraminifera in the NW Arabian Sea. In P. D. Clift, D. Kroon, C. Gaedicke, & J. Craig (Eds.), *The tectonic and climatic evolution of the Arabian Sea region* (Vol. 195, pp. 463–497). Geological Society, London, Special Publications.

Pierrot, D. E., Lewis, E., & Wallace, D. W. R. (2006). *MS Excel Program Developed for C_O2 System Calculations*. *ORNL/CDIAC-105*a. Carbon Dioxide Information analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy. https://doi.org/10.3334/CDIAC/otg.CO2SYS_XLS_CDIAC105a

Ramaswamy, V., & Nair, R. R. (1994). Fluxes of material in the Arabian Sea and Bay of Bengal—Sediment trap studies. *Proceedings of the Indian Academy of Sciences - Earth & Planetary Sciences*, *103*(2), 189–210. https://doi.org/10.1007/bf02839536

Rao, L. V. G., & Ram, P. S. (2005). *Upper ocean physical processes in the tropical Indian Ocean* (p. 595). Visakhapatnam.

Rao, R. R., Molinari, R. L., & Festa, J. F. (1991). Surface meteorological and near surface oceanographic atlas of the tropical Indian ocean. *NOAA Technical Memorandum ERL AOML-69* (pp. 1–63).

R Core Team. (2020). R: A language and environment for statistical computing, version (4.0.3) [Software]. R Foundation for Statistical Computing. Retrieved from http://R-project.org

Renaud, S., & Schmidt, D. N. (2003). Habitat tracking as a response of the planktic foraminifer *Globorotalia truncatulinoides* to environmental fluctuations during the last 140 kyr. *Marine Micropaleontology*, *49*(1–2), 97–122. https://doi.org/10.1016/S0377-8398(03)00031-8

Rillo, M. C., Miller, C. G., Kucera, M., & Ezard, T. H. G. (2019). Supplementary Material: Intraspecific size variation in planktonic foraminifera cannot be consistently predicted by the environment [CSV]. *Natural History Museum*. https://doi.org/10.5519/0056541

Rillo, M. C., Miller, C. G., Kucera, M., & Ezard, T. H. G. (2020). Intraspecific size variation in planktonic foraminifera cannot be consistently predicted by the environment. *Ecology and Evolution*, *10*(20), 11579–11590. https://doi.org/10.1002/ece3.6792

Rutherford, S., D'Hondt, S., & Prell, W. (1999). Environmental controls on the geographic distribution of zooplankton diversity. *Nature*, *400*(6746), 749–753. https://doi.org/10.1038/23449

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Schiebel, R. (2002). Planktic foraminiferal sedimentation and the marine calcite budget. *Global Biogeochemical Cycles*, *16*(4), 3-1–3-21. 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. (2017). *Planktic foraminifers in the modern ocean*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-50297-6

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, D. N., Thierstein, H. R., Bollmann, J., & Schiebel, R. (2004). Abiotic forcing of plankton evolution in the Cenozoic. *Science*, *303*(5655), 207–210. https://doi.org/10.1126/science.1090592

Schott, F. A., Xie, S. P., & McCreary, J. P., Jr. (2009). Indian Ocean circulation and climate variability. *Reviews of Geophysics*, *47*(1), 1–46. https://doi.org/10.1029/2007rg000245

Seears, H. A., Darling, K. F., & Wade, C. M. (2012). Ecological partitioning and diversity in tropical planktonic foraminifera. *BMC Evolutionary Biology*, *12*(1), 1–16. https://doi.org/10.1186/1471-2148-12-54

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. University of Illinois Press.

Siccha, M., & Kucera, M. (2017). ForCenS, a curated database of planktonic foraminifera census counts in marine surface sediment samples. *Scientific Data*, *4*(1), 170109. https://doi.org/10.1038/sdata.2017.109

Spero, H. J., Bijma, J., Lea, D. W., & Bemis, B. E. (1997). Effect of seawater carbonate concentration on foraminiferal carbon and oxygen isotopes. *Nature*, *390*(6659), 497–500. https://doi.org/10.1038/37333

Stips, A., Macias, D., Coughlan, C., Garcia-Gorriz, E., & Liang, X. S. (2016). On the causal structure between CO_2 and global temperature. *Scientific Reports*, *6*(1), 21691. https://doi.org/10.1038/srep21691

Stone, S. W. (1956). Some ecologic data relating to pelagic foraminifera [Atlantic Ocean and Jamaica]. *Micropaleontology*, 2(4), 361–370. https://doi.org/10.2307/1484352

Suárez-Ibarra, J. Y., Frozza, C. F., Petró, S. M., & Gómez Pivel, M. A. (2021). Fragment or broken? Improving the planktonic foraminifera fragmentation assessment. *Palaois*, *36*(5), 165–172. https://doi.org/10.2110/palo.2020.062

Susanto, R. D., Gordon, A. L., & Zheng, Q. (2001). Upwelling along the coasts of Java and Sumatra and its relation to ENSO. *Geophysical Research Letters*, *28*(8), 1599–1602. https://doi.org/10.1029/2000GL011844

Tachikawa, K., Sépulcre, S., Toyofuku, T., & Bard, E. (2008). Assessing influence of diagenetic carbonate dissolution on planktonic foraminiferal Mg/Ca in the southeastern Arabian Sea over the past 450 ka: Comparison between *Globigerinoides ruber* and *Globigerinoides sacculifer*. *Geochemistry, Geophysics, Geosystems*, *9*(4), Q04037. https://doi.org/10.1029/2007GC001904

Takahashi, K., & Be, A. W. H. (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

Ternon, J. F., Bach, P., Barlow, R., Huggett, J., Jaquemet, S., Marsac, F., et al. (2014). The Mozambique Channel: From physics to upper trophic levels. *Deep Sea Research Part II: Topical Studies in Oceanography*, *100*, 1–9. https://doi.org/10.1016/j.dsr2.2013.10.012

Thunell, R., Benitez-Nelson, C., Varela, R., Astor, Y., & Muller-Karger, F. (2007). Particulate organic carbon fluxes along upwelling-dominated continental margins: Rates and mechanisms. *Global Biogeochemical Cycles*, *21*(1), 1–12. https://doi.org/10.1029/2006gb002793

van Valen, L. (1965). Morphological variation and width of ecological niche. *The American Naturalist*, *99*(908), 377–390. https://doi.org/10.1086/282379

Violle, C., Enquist, B. J., McGill, B. J., Jiang, L., Albert, C. H., Hulshof, C., et al. (2012). The return of the variance: Intraspecific variability in community ecology. *Trends in Ecology & Evolution*, *27*(4), 244–252. https://doi.org/10.1016/j.tree.2011.11.014

Weinkauf, M. F. G., Bonitz, F. G. W., Martini, R., & Kučera, M. (2019). An extinction event in planktonic foraminifera preceded by stabilizing selection. *PLoS One*, *14*(10), e0223490. https://doi.org/10.1371/journal.pone.0223490

Weinkauf, M. F. G., 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), e0148363. https://doi.org/10.1371/journal.pone.0148363

Weinkauf, M. F. G., Moller, T., Koch, M. C., & Kučera, M. (2013). Calcification intensity in planktonic foraminifera reflects ambient conditions irrespective of environmental stress. *Biogeosciences*, *10*(10), 6639–6655. https://doi.org/10.5194/bg-10-6639-2013

Weinkauf, M. F. G., Siccha, M., & Weiner, A. K. M. (2022). Reproduction dynamics of planktonic microbial eukaryotes in the open ocean. *Journal of the Royal Society Interface*, *19*(187), 20210860. https://doi.org/10.1098/rsif.2021.0860

Whittaker, R. J., Willis, K. J., & Field, R. (2001). Scale and species richness: Towards a general, hierarchical theory of species diversity. *Journal of Biogeography*, *28*(4), 453–470. https://doi.org/10.1046/j.1365-2699.2001.00563.x

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis* (2nd ed.). Springer International Publishing; Springer. https://doi.org/10.1007/978-3-319-24277-4

Wyrtki, K. (1971). *Oceanographic Atlas of the International Indian Ocean Expedition* (p. 531). National Science Foundation.

Zarkogiannis, S., Kontakiotis, G., & Antonarakou, A. (2020). Recent planktonic foraminifera population and size response to Eastern Mediterranean hydrography. *Revue de Micropaleontologie*, *69*, 100450. https://doi.org/10.1016/j.revmic.2020.100450

Zeebe, R. E., & Wolf-Gladrow, D. A. (2001). *CO_2 in seawater: Equilibrium, kinetics, isotopes*. Elsevier.

Zweng, M. M., Reagan, J. R., Seidov, D., Boyer, T. P., Locarnini, R. A., García, H. E., et al. (2018). Salinity. In A. Mishonov (Ed.), (Technical Ed.), *World Ocean Atlas 2018* (Vol. 2, p. 50). *NOAA Atlas NESDIS 82*.

References From the Supporting Information

Alboukadel, K. (2020). ggpubr: 'Ggplot2' based publication ready plots. Retrieved from https://rpkgs.datanovia.com/ggpubr/

Brunson, J. C., & Read, Q. D. (2020). ggalluvial: Alluvial Plots in 'ggplot2'. *R package version 0.12.3*. Retrieved from http://corybrunson.github.io/ggalluvial/

Chang, W. (2002). extrafont: Tools for using fonts. Version 0.18. Retrieved from https://github.com/wch/extrafont

Garnier, S., Ross, N., Rudis, B., Sciaini, M., Camargo, A. P., & Scherer, C. (2021). viridis: Colorblind-friendly color maps for R (Version 0.6.1) [Software]. Github. https://github.com/sjmgarnier/viridis/

Hyndman, R., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild, M., et al. (2021). Forecasting functions for time series and linear models (Version 8.15) [Software package]. forecast. https://pkg.robjhyndman.com/forecast/

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Maechler, M., Rousseeuw, P., Croux, C., Todorov, V., Ruckstuhl, A., Salibian-Barrera, M., et al. (2022). robustbase: Basic robust statistics. *R package version 0.95-0*. Retrieved from http://robustbase.r-forge.r-project.org/

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R. B., et al. (2007). Community ecology package (version 2.9) [Software]. Github. https://github.com/vegandevs/vegan

Rudis, B., Kennedy, P., Reiner, P., Wilson, D., Adam, X., Google (Roboto Condensed & Titillium Web Fonts), et al. (2020). hrbrthemes: Additional themes, theme components and utilities for 'ggplot2' (version, 0.8.0). [Software]. Github. https://hrbrmstr.github.io/hrbrthemes/

Sievert, C. (2020). *Interactive web-based data visualization with R, plotly, and shiny*. Chapman and Hall/CRC.

Statisticat, L. L. C. (2021). LaplacesDemon: Complete environment for Bayesian inference. *R package version 16.1.6*. Retrieved from https://web.archive.org/web/20150206004624/http://www.bayesian-inference.com/software

Trapletti, A., & Hornik, K. (2022). tseries: Time series analysis and computational finance. *R package version 0.10-51*. Retrieved from https://CRAN.R-project.org/package=tseries

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). Springer.

Wickham, H., Averick, M., Bryan, J., Chang, W., D'Agostino McGowan, L., François, R., et al. (2019). Welcome to the tidyverse. *Journal of Open-Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., François, R., Henry, L., & Müller, K. (2018). dplyr: A grammar of data manipulation, (version 0.7.6) [Software]. Github. Retrieved from https://CRAN.R-project.org/package=dplyr

Wickham, H., François, R., Henry, L., & Müller, K. (2021). dplyr: A grammar of data manipulation (version 1.0.7) [Software]. Github. Retrieved from https://github.com/tidyverse/dplyr

Wickham, H., Seidel, D., & RStudio. (2020). scales: Scale functions for visualization (version 1.1.1) [Software]. Github. Retrieved from https://github.com/r-lib/scales

Wilke, C. O. (2021). Ridgeline plots in 'ggplot2' (version 0.5.3) [Software]. View on CRAN. Retrieved from https://wilkelab.org/ggridges/

15252027, 2023, 4, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2022GC010586 by Cochrane France, Wiley Online Library on [03/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License