Size normalizing planktonic Foraminifera abundance in the water column

Sonia Chaabane\textsuperscript{1,2,3}, Thibault De Garidel-Thoron\textsuperscript{1}, Xavier Giraud\textsuperscript{1}, Julie Meilland\textsuperscript{4}, Geert-Jan A Brummer\textsuperscript{5}, Lukas Jonkers\textsuperscript{4}, P Graham Mortyn\textsuperscript{6}, Mattia Greco\textsuperscript{7}, Nicolas Casajus\textsuperscript{3}, Michal Kucera\textsuperscript{4}, Olivier Sulpis\textsuperscript{1}, Azumi Kuroyanagi\textsuperscript{8}, Hélène Howa\textsuperscript{9}, Gregory Beaugrand\textsuperscript{10}, and Ralf Schiebel\textsuperscript{2}

\textsuperscript{1}Aix-Marseille Université, CNRS, INRAE, CEREGE
\textsuperscript{2}Department of Climate Geochemistry, Max Planck Institute for Chemistry
\textsuperscript{3}Fondation pour la recherche sur la biodiversité (FRB-CESAB)
\textsuperscript{4}Center for Marine Environmental Sciences, MARUM, University of Bremen
\textsuperscript{5}NIOZ, Department of Ocean Systems, Netherlands Institute for Sea Research
\textsuperscript{6}Universitat Autònoma de Barcelona, ICTA and Dept. of Geography
\textsuperscript{7}Institute of Marine Sciences (ICM), CSIC
\textsuperscript{8}Tohoku University Museum, Tohoku University
\textsuperscript{9}LPG-BIAF, UMR-CNRS 6112, University of Angers
\textsuperscript{10}UMR 8187, LOG, Laboratoire d’Océanologie et de Géosciences, Université Littoral Côte d’Opale, Univ. Lille, CNRS

March 13, 2024
Size normalizing planktonic Foraminifera abundance in the water column

Sonia Chaabane¹,²,³*, Thibault de Garidel-Thoron¹, Xavier Giraud¹, Julie Meillard⁴, Geert-Jan A. Brummer⁵, Lukas Jonkers⁴, P. Graham Mortyn⁶, Mattia Greco⁷, Nicolas Casajus³, Michal Kucera⁴, Olivier Sulpis¹, Azumi Kuroyanagi⁸, Hélène Howa⁹, Gregory Beaugrand¹⁰, Ralf Schiebel²

¹ Aix-Marseille Université, CNRS, IRD, INRAE, CEREGE, Aix-en-Provence, France
² Department of Climate Geochemistry, Max Planck Institute for Chemistry, Mainz, Germany
³ Fondation pour la recherche sur la biodiversité (FRB-CESAB), Montpellier, France
⁴ MARUM, Center for Marine Environmental Sciences, University of Bremen, Bremen, Germany
⁵ NIOZ, Royal Netherlands Institute for Sea Research, Department of Ocean Systems, Texel, The Netherlands
⁶ Universitat Autònoma de Barcelona, ICTA and Dept. of Geography, Spain
⁷ Institute of Marine Sciences (ICM), CSIC, Passeig Marítim de la Barceloneta, 37-49, Barcelona 08003, Spain
⁸ Tohoku University Museum, Tohoku University, Japan
⁹ LPG-BIAF, UMR-CNRS 6112, University of Angers, France
¹⁰ Université Littoral Côte d’Opale, Univ. Lille, CNRS, UMR 8187, LOG, Laboratoire d’Océanologie et de Géosciences, France

* Corresponding author

E-mail: sonia.chaabane@gmail.com
chaabane@cerege.fr
sonia.chaabane@mpic.de
Abstract

Planktonic Foraminifera have been collected from the water column with different plankton sampling devices equipped with nets of various mesh sizes, which impedes direct comparison of observed quantifications. Here, we use data on the community size structure of planktonic Foraminifera to assess the impact of mesh size on the measured abundance (ind/m³) of planktonic Foraminifera. We use data from the FORCIS database (Chaabane et al., 2023) on the global ocean at different sampling depths over the past century. We find a global cumulative increase in abundance with size, which is best described using a Michaelis-Menten function. This function yields multiplication factors by which one size fraction can be normalized to any other size fraction equal to or larger than 100 µm. The resulting size normalization model is calibrated over a range of different depth intervals, and validated with an independent dataset from various depth ranges. The comparison to the Berger, 1969 equivalent catch approach shows a significant increase in the predictive skill of the model. The new size normalization scheme enables comparison of Foraminifera abundance data sampled with plankton nets of different mesh sizes, such as compiled in the FORCIS database. The correction methodology may be effectively employed for various other plankton groups such as diatoms and dinoflagellates.

Keywords

Planktonic Foraminifera, size fraction, size-normalized catch model, vertical distribution, mesh size

1. Introduction

Planktonic Foraminifera are ubiquitous marine unicellular mesoplankton calcifiers, contributing to the global carbon cycle by exporting carbon to the seafloor. They live in the upper hundreds of meters of the ocean (Rebotim et al. 2017) where they grow by sequential addition of calcareous chambers to their shell (called tests), ranging from ~100 to ~1 mm, varying across species. While calcifying, Foraminifera encode ambient environmental conditions in their shells, enabling key proxies for paleoceanographic reconstructions. Upon death or reproduction, the empty shells sink through the water column to the ocean floor where they accumulate and eventually fossilize (Bé and Anderson, 1976; Hemleben et al., 1989). Their
ocean-wide distribution, species diversity, and shell geochemistry (isotope and trace metal composition) facilitate the reconstruction of past ocean and climate conditions using sedimentary records (e.g., Anderson and Prell, 1993; Bé et al., 1977; Bonneau et al., 1980; Caulet et al., 1992; CLIMAP, 1976; Ganssen, 1991; Garidel-Thoron et al., 2005; Grazzini et al., 1995; Kroon and Ganssen, 1989; Kucera et al., 2005; Mortyn and Martinez-Boti, 2007; Steens et al., 1992; Strack et al. 2022).

Planktonic Foraminifera begin their life with a single initial calcified chamber, the proloculus, which ranges 5-50 µm in diameter (Sverdlove and Bé, 1985; Brummer et al., 1987). The calcified shell which houses the cytoplasm grows by adding chambers to the test and reaches an adult size and ecological conditions (Schmidt et al., 2004). Therefore, the size spectrum of the planktonic Foraminifera community at a specific location and time is indicative of the present species, their average maturity levels and fitness with the environmental parameters. When collecting samples with a mesh larger than the smallest specimens, both the number of individuals per volume of seawater and species composition data will be skewed, as smaller specimens will be missed (Berger, 1969; Berger, 1971; Bé and Hutson, 1977; Brummer and Kroon, 1988).

Over the last century, living planktonic Foraminifera were collected from the water column using various devices equipped with nets of different mesh sizes (from 30 µm to 650 µm). Originally, samples were collected with coarse-meshed nets and gradually changed over time to finer nets that caught smaller specimens and species. For instance, King and Demond (1953) used a mesh size of 650 µm while Bé et al. (1985) and Bé et al. (1971) used mesh sizes of 333 µm and 202 µm, respectively. More recent studies have used smaller mesh sizes from 30 µm up to the classical 150 µm sieve size traditionally used in paleoceanographic studies (e.g., Ufkes et al. (1998), 150 µm; Kuroyanagi and Kawahata (2004), Mallo et al. (2017), 125 µm; Schiebel et al., (1995), 100 µm; Sousa et al. (2014), Lessa et al. (2020), Keigwin et al. (2005), 63 µm; Boltovskoy et al. (2000), 30 µm). However, there are currently no available standards for the sieve mesh sizes used in paleoceanographic studies.

The resulting sampling bias in measuring planktonic Foraminifera abundances not only distorts our understanding of their contribution to the oceanic carbon cycle but also hampers efforts to define their ecological niches. Despite numerous studies recognizing the impact of mesh size on the observed species abundance and diversity, few have specifically addressed this bias (Berger, 1969; Peeters et al., 1999; Schiebel and Hemleben, 2000). In a pioneering study, Berger (1969) derived a set of equations to compute equivalent catches for different mesh sizes.
The equations were based on a limited range of samples, and used assumptions such as the invariance of the assemblage composition through space and time, or that the size mode of the catch corresponds to the mesh size used. Since the 1960's, a wealth of new data have become available and recently synthesized in the FORCIS database (Foraminifera Response to Climatic Stress database, Chaabane et al, 2023; de Garidel-Thoron et al. 2022), allowing us to explore the possibility to design a better-constrained normalization approach. This step is fundamental to fully compare samples collected using different mesh sizes (Fig. S2), which need to be normalized at a similar mesh size (e.g., larger than 100 µm).

In the FORCIS database, abundance data are compiled as size-fractionated abundance data, called subsample, and are available for thousands of individual samples. For instance, from a single sample split into six size fractions (100-125 µm, >125-150 µm, >150-200 µm, >200-250 µm, >250-315 µm, and >315 µm), the finer size fraction (encompassing the three size classes 100-125 µm, >125-150 µm, and >150-200 µm) presents the highest abundance (Fig. S1).

The objective of this study is to propose a model to normalize the abundance (ind/m³) of planktonic Foraminifera caught in a net having a certain mesh size to the abundance that would be measured if a net with another mesh size was used instead, allowing us to relate all counts to a standard mesh size. We assess the sensitivity of this model, normalized to sizes equal to or larger than 100 µm, to different water depths, seasons, oceanic basins and species size categories within the training set. The predictive skill of our model is validated using a large, independent, size-fractionated validation dataset from the North Atlantic Ocean (Retailleau et al. 2011). Lastly, we compare the predictions of this model with the only analogous normalization methodology developed by Berger (1969) on planktonic Foraminifera. Our normalization approach is applied to global observations on samples from plankton nets, plankton pumps, and Continuous Plankton Recorders (CPRs) to derive a size harmonized estimate of planktonic Foraminifera abundance. The correction methodology is applicable not only to planktonic Foraminifera but could be adapted to correct for mesh size bias in other plankton groups sampled using similar methods, such as pteropods or other taxa.

2. Materials and Procedures

2.1 Size-normalized catch

2.1.1 Model training dataset
For this analysis, we select a subset of samples collected by plankton nets from the FORCIS database (Chaabane et al. 2023) and fractionated into different size fractions, to characterize the relationship between abundances versus the size fractions of Foraminifera tests. A total of 1026 samples match this criterion, and were sieved and separated into six different size fractions \((k\) from 1 to 6) , where \(k_{\text{max}}\) is the number of the last size class within a single subsample (an aliquot plankton sample obtained from a specific depth range, time interval, size fraction range, and location) for each of the following mesh size fractions: >100-125 \(\mu m\) \((k=1)\), >125-150 \(\mu m\) \((k=2)\), >150-200 \(\mu m\) \((k=3)\), >200-250 \(\mu m\) \((k=4)\), >250-315 \(\mu m\) \((k=5)\), >315 \(\mu m\) \((k=6)\). These samples contain cytoplasm-filled and empty planktonic Foraminifera tests sampled from the North Atlantic, Indian and Arctic Oceans (Fig. 1), and cover a depth range of 0 to 1000 m water depth (three different casts sampled at different levels).

### 2.1.2 Multiplication factor calculation

The cumulative abundance \(C_{\text{cum}}; \text{ in individuals/m}^3\) refers to the total abundance (defined as the number of individuals per cubic meter) across a range of size fractions. For size classes ranging from size class 1 (100-125 \(\mu m\)) to class \(k\), the cumulative abundance, \(C_{\text{cum},k}\), is calculated as follows:

\[
C_{\text{cum},k} = \sum_{i=1}^{k} C_i
\]

where \(C_i\) is the abundance in size fraction \(i\).

Observations indicate that the cumulative abundance typically exhibits an asymptotic shape (Fig. 2A). This behavior is quantified through a multiplication factor \(f_k\), defined as the ratio of the cumulative abundance for a given class \(k\) to the abundance of the first class \((C_1)\):

\[
f_k = \frac{C_{\text{cum},k}}{C_1} \quad \text{Eq. (1)}
\]

The selection of a fitting function for the multiplication factor must adhere to the principle of parsimony, implying a limited number of parameters for calibration to avoid overfitting. Additionally, the function should asymptote to a maximum value, reflecting the biological reality that a population’s total abundance cannot be infinite and that there are no more individuals above a certain size. The chosen function's parameters should also correlate with the distribution patterns of size versus abundance within specific Foraminifera populations.

To accurately model the gradual increase and eventual plateau of the cumulative distribution starting from a nonzero initial class size, we sought a fitting function that depicts this pattern without overemphasizing smaller size fractions. Among the considered alternatives, the logistic
function, while commonly used, seems inadequate due to its inability to constrain the lower range effectively, particularly evident in our study where not all size spectra were covered by our mesh sizes.

Therefore, the Michaelis-Menten (MM) function (Michaelis and Menten, 1913), traditionally linked to enzyme kinetics but broadly applicable to various biological phenomena demonstrating an asymptotic behavior, was selected as the most suitable fitting function. This decision was supported by its demonstrated effectiveness in other biological contexts, such as growth modeling studies (Walters et al. 2024; Sardari et al. 2023). The MM function’s adaptability stems from its simplicity, encapsulating the dynamics of planktonic Foraminifera test size distributions with only two parameters: the maximum potential abundance ($f_{max}$) and the initial response rate or sharpness of the curve ($S_{half}$) at the beginning of the distribution.

The MM function facilitates a nuanced interpretation of Foraminifera population distributions, particularly in the context of size ranges. The $f_{max}$ parameter, indicating the maximum multiplication factor achievable between the smallest size class (100-125 µm) and the total population abundance, reflects the breadth of the size distribution. Higher $f_{max}$ values suggest a significant contribution from larger specimens to the population's total abundance. Conversely, the half-saturation constant ($S_{half}$), quantifies the distribution curve's steepness. Lower $S_{half}$ values indicate a rapid attainment of the maximum abundance, emphasizing the dominance of smaller size fractions within the population.

By applying a scale translation to the initial size range, we derive a model for the multiplication factor as follows (Fig. 3):

$$f_k = 1 + (f_{max} - 1) \frac{(S_{sup,k} - S_{sup,1})}{(S_{sup,k} - S_{sup,1}) + (S_{half} - S_{sup,1})}$$

Eq. (2)

where $S_{sup,1}$ and $S_{sup,k}$ are the upper size limits of size class 1 and $k$, respectively, and $S_{half}$ and $f_{max}$ are the fitted parameters.

To derive $S_{half}$ and $f_{max}$, we fit $f_k$ to $S_{sup,k}$ through a MM equation, using the “MM.2()” function of the drc package in R (Ritz et al. 2015).

2.1.3 Size-normalized catch model
Finally, to scale the measured abundance from any given size range \((C_{sz, sup}^{sz, inf})\), with \(sz\_inf\) and \(sz\_sup\) being the lower and upper limits of the size fraction, given in micrometers, µm) to the theoretical abundance at a size range extending from a chosen normalizing value \((C_{sz, norm}^{∞})\), the following equation is applied:

\[
C_{sz, norm}^{∞} = C_{sz, sup}^{sz, inf} \frac{f_{max} - f_{sz, norm}}{f_{sz, sup} - f_{sz, inf}} \tag{Eq. (3)}
\]

where \(sz\_inf\) and \(sz\_sup\) are the lower and upper size limits, respectively, of the measured size fraction, \(sz\_norm\) is the normalization size and \(f_{Sz}\) is the multiplication factor associated to the \(Sz\), computed as follows:

\[
\begin{aligned}
\text{if } Sz &= 100 \mu m & f_{100} &= 0 \\
\text{if } Sz \in [125; +∞] & f_{Sz} = 1 + \left( f_{max} - 1 \right) \frac{(Sz - S_{125})}{(S_{sz} - S_{125}) + (S_{half} - S_{125})} & \tag{Eq. (4)} \\
\text{if } Sz &≡ +∞ & f_{∞} &= f_{max}
\end{aligned}
\]

with \(S_{125} = 125 \mu m\)

The size fractions of 100 µm and 125 µm used in this formula are dictated by the finer size class selected for calibration in the database collection, so that the size-normalization (Eq. 3) can be applied to 100 µm or any size above 125 µm.

As shown in figure 3C, for the distribution of the abundance between 100 µm and 125 µm, normalization is possible by extending the calculation of the multiplication factor as follows:

\[
\text{if } Sz \in [100; 125] & f_{Sz} = \frac{Sz - S_{100}}{S_{125} - S_{100}} \tag{Eq. (5)}
\]

Before determining the parameters \(S_{half}\) and \(f_{max}\), we removed outliers (10% of the data below and above the high-density curves), detected on the multiplication factor of each size fraction, considering values beyond 1.5 times the interquartile range from the median as outliers. The Median +1.5 * IQR (Interquartile Range) method, commonly used in box plots to spot outliers, sets a boundary indicating where most data points should lie within a box plot. This boundary is calculated by adding 1.5 times the Interquartile Range (IQR) to the median. If data points exceed this boundary, they are flagged as potential outliers. This method is preferred for its ability to handle extreme values and skewed distributions effectively, offering a balanced way to identify outliers while safeguarding the integrity of the data. We used the function
"boxplot.stat()" in R (version 2021.09.1; R Core Team, 2021) for the detection. The very low abundances in all size classes (flat fit), unrealistically high values in fine size fractions (likely due to technical failure or sampling effects such as clogging of the plankton net) followed by very low abundances in the coarser size fraction, were detected as outliers (Figs. 2B). Both parameters ($S_{\text{half}}$ and $f_{\text{max}}$) were determined for each ocean basin where we have the selected data to conceptualize the model (Arctic, North Atlantic, and Indian Oceans), season and water depth interval (between 0 m and 50 m, 50 m to 100 m, 100 m to 300 m and 300 m to 1000 m), based on the mean depth of the sample to check whether the MM parameters are time and/or space dependent (Tab. S2; Fig. S3). It was also assessed for three size classes based on adult test size, categorized as small (100 to 200 µm), medium (>200 to 300 µm), and large (>300 µm) species (Tab. 1; Schiebel and Hemleben, 2017).

2.1.4 Validation

To assess the robustness of the proposed model, the obtained Eq. 5 was applied to an independent validation dataset (included in the FORCIS database), that is not part of the training dataset (Fig. 1). This validation set includes 552 subsamples analyzed in three different size fractions, i.e., 100-150 µm, >150-315 µm, and >315 µm, where abundances in the >100 µm net samples were counted from 0 to 700 meters (Retailleau, 2009; Retailleau et al. 2011).

2.2 Berger’s “equivalent catch”

Since the use of different mesh sizes could yield different faunal compositions for the same sample, Berger (1969) defined the “equivalent catch” to compare the abundances of living planktonic Foraminifera collected with plankton nets of unequal mesh size. The standardized abundance of planktonic Foraminifera would then be obtained using the following empirical equation (Eq. 6):

$$\text{Abundance}_{\text{standardized}} = \text{Abundance}_{\text{actual}} \times \left( \frac{S_{\text{actual}}}{S_{\text{standardized}}} \right)^a$$  Eq. (6)

where $S_{\text{actual}}$ is the observed (net or sieve) mesh size used, $S_{\text{standardized}}$ is the standard mesh size, and $a$ is a constant number set to 3 for plankton tow samples (Berger (1969) and Peeters et al. (1999)).

To compare the equivalent catch using the Berger (1969) equation to the size-normalized catch defined in this study, we used the size-fractionated training dataset (5117 subsamples extracted from the FORCIS database) separated into six different size fractions (>100-125 µm, >125-150
µm, >150-200 µm, >200-250 µm, >250-315 µm, >315 µm). Then, we calculated the abundance at different equivalent mesh sizes (>100 µm, >125 µm, >150 µm, >200 µm, >250 µm, and >315 µm), as computed by the equivalent catch formula in Berger (1969). The obtained abundances were then standardized to 100 µm and therefore used “100” as $S_{(standardized)}$ in Eq. 6 and to a known abundance (>100 µm).

The coefficient of determination ($r^2$) of the linear regression models were calculated through the R scripts using the ‘ggpmisc’ packages, to investigate the similarity between the values calculated using the Berger (1969) equation (Eq. 6) and our approach (Eq. 4).

2.3 Application of the size-normalized catch model

We extracted more than 175 000 subsamples of species counts from the FORCIS database (Chaabane et al., 2023), i.e., single species aliquots collected within a depth range, time interval, and size fraction range at a given location. Subsamples include counts on the total of cytoplasm-filled and empty tests in a defined water parcel. The subsamples include data published between 1950 and 2018 and collected in different oceanographic environments by continuous plankton recorders (CPR; 157 000 samples having 157 000 subsamples), plankton nets (6 000 samples having 19 000 subsamples) and plankton pump (300 samples having 400 subsamples) in the Atlantic, Pacific, Indian, and Arctic Oceans (Fig. S2). The various types of counts (raw data or relative abundance) for total planktonic Foraminifera and species-specific count data within these samples were extracted from the FORCIS database (Fig. S3). Only samples containing information regarding both the total number of individuals and the filtered volume were considered for conversion to abundances (ind/m$^3$). All CPR data have only one subsample per sample and were collected only between 5 m and 10 m water depth using a mesh size of 270 µm.

The corresponding season at the time of the collection of each sample was extracted from the FORCIS database. Seasons were distinguished between the Northern and Southern Hemispheres. In the Northern Hemisphere, autumn comprises September, October, and November; winter includes December, January, and February; spring consists of March, April, and May; and summer spans June, July, and August. Conversely, in the Southern Hemisphere, the seasons are reversed: spring occurs in September, October, and November; summer in December, January, and February; autumn in March, April, and May; and winter in June, July, and August.
3. Results

3.1. FORCIS size-normalized catch

3.1.1. Sieve size and abundance

The relationships between abundances of planktonic Foraminifera, water depth, and size fraction have been assessed using the cumulative abundances as a function of size fraction in each sample (Fig. 2A). Despite the large differences between samples, they show similar and parallel patterns over a wide range of abundances (Fig. 2A), with a notable feature: the cumulative abundance is higher for the samples from shallow waters compared to their deep-water counterparts (Fig. 4A). The pattern of the cumulative curves closely resembles that of a logarithmic function, exhibiting the most pronounced increases in the finer fractions (from 100-125 µm and 125-150 µm), which then level off as the size fractions become coarser. The cumulative abundances range between 0.2 and 600 ind/m$^3$ for the finer size fractions (from 100-125 µm and 125-150 µm) and between 4 and 900 ind/m$^3$ for the coarser ones.

The multiplication factors (by which the abundances increase in each size fraction) derived from the cumulative abundance curves (see methods), show an overall asymptotic pattern across all samples (Fig. 2B). The latter is similar to the Michaelis-Menten curve with constants $S_{\text{half}}$ and $f_{\text{max}}$ obtained for the general fit of about 178 ± 3 µm and 2.48 ± 0.02, respectively (Fig. 4B and Tab. 2).

A steep slope in multiplication factor is observed between 100 µm and 150-200 µm for all curves, and levels off for size fractions larger than 250 µm. Generally, the mean value of the factor ($f_k$) is about 1.47 for the >125-150 µm size fraction and reaches 1.85 for the size fraction >150-200 µm, increasing to 2.21 for the size fraction larger than 315 µm. The increase of $f_k$ between the size fractions >200-250 µm and >250-315 µm is low and close to 0.11.

The obtained asymptotic curves (multiplication factors vs. size fractions) also show similar patterns for each water depth (0-50 m, >50-100 m, >100-300 m, >300-1000 m, and >1000 m) with a steep slope between 100 µm and 150 µm, and a decrease with depth (Fig. 4B). The constants of these MM fits range between 167-186 µm and 2.18-2.75 for $S_{\text{half}}$ and $f_{\text{max}}$, respectively. Highest values are observed at the shallower depths (from 2.54 to 2.75 and from 184 µm and 186 µm for $f_{\text{max}}$ and $S_{\text{half}}$, respectively), between 0 m and 100 m, especially in the coarser size fractions (Tab. 2).
Between 100 µm and 150 µm, the multiplication factor is the same at all depths (Tab. 3), while the difference between two successive multiplication factors at all depths is decreasing at coarser sizes. For instance, the difference between two successive multiplication factors is low for the size fractions >250 µm, and particularly low for the size fractions >315 µm, where the abundance is also close to 0.

In deeper samples, the coarse fraction shows lower abundances than in shallow waters (Fig. 4). For example, for the size fraction >315 µm the standard deviation among the multiplication factors is 0.16, while it is only 0.03 for the size fraction >125-150 µm (Tab. 3).

The relationships between the multiplication factors derived from the cumulative abundance curves and the size fractions were assessed per season, ocean basin, and water depth interval (Fig. 5). Overall, the resulting envelopes (confidence intervals at the limits of 2.5 % and 97.5 %) show similar parallel patterns with higher uncertainty at larger sizes, especially for the North Atlantic Ocean data in summer, spring, and autumn, where the number of data points is high. However, these envelopes overlap for the Indian and Arctic Oceans in all seasons, where the sampling coverage is low (for example, Indian Ocean summer observations are 80 % lower than in the North Atlantic Ocean). Overall, if MM parameters vary slightly between ocean basins and seasons (Figs. 5B, 5C and 5D; Tab. S2), the differences cannot be thoroughly confirmed in this study due to the sampling coverage. The depth dependance of the MM parameters is the factor that is shown to be robust in our analyses. Consequently, our analyses rely on globally derived MM parameters across depth intervals.

### 3.1.2. Species test size and abundance

Test sizes of adult specimens (Tab. 1) differ significantly among species and thus impact the distribution of abundances with sieve size. Typically, small-sized species, such as *Berggrenia pumilio*, *Dentigloborotalia anfracta*, *Globigerinita glutinata*, *Globigerinita minuta*, *Globoturborotalita rubescens*, *Globoturborotalita tenellus*, *Globigerinita uvula*, and *Gallitellia vivans* may not be found at all in medium to coarse mesh sizes. In contrast, large-sized species occur in all size fractions, such as *Globigerinella adamsi*, *Globorotalia cultrata*, *Beella digitata*, *Globorotalia truncatulinoides*, *Hastigerinella digitata*, and *Hastigerina pelagica*. Medium sized species like *Candeina nitida*, *Globigerina bulloides*, and *Globorotalia scitula* occur in small (100 to 200 µm), medium (>200 to 300 µm), but are often not found in the very large size fractions (>300 µm). Assessing the $f_{max}$ and $S_{half}$ on the different species groups, all species show constants that decrease with depth except for the large species where the $f_{max}$ and $S_{half}$...
is high in the 50 - 300 m depth range (from 5.15 to 4.56 and from 257.99 to 253.80 for \( f_{\text{max}} \) and \( S_{\text{hal}} \), respectively) regardless of their final size (Tabs. 2 and 4A-C). For the small species, the \( f_{\text{max}} \) varies between 1.41 and 1.57, and \( S_{\text{hal}} \) ranges between 143.32 µm and 149.08 µm, while the large species show higher \( f_{\text{max}} \) from 3.88 to 5.40, at \( S_{\text{hal}} \) of 189.53 µm to 262.36 µm, respectively. For the medium sized species, the \( S_{\text{hal}} \) and \( f_{\text{max}} \) constants have a range between the ones observed in small to large species, with \( S_{\text{hal}} \) varying from 178.86 to 196.44 and \( f_{\text{max}} \) from 3.02 to 3.30.

### 3.2. Abundance estimation using the size-normalized catch model and Berger’s “equivalent catch”

#### 3.2.1. Abundance estimation using the size-normalized catch model

To evaluate the performance and reliability of the size-normalized catch model correction scheme proposed in the study, the observed abundances planktonic Foraminifera data used to train the model were plotted vs. predicted total abundances of total specimens (all size categories; on the total of cytoplasm-filled and empty tests). These latter were calculated using Equation 3 at each depth interval proposed by the size-normalized catch model in this study (Tab. 2) and yield slopes equal to unity for the global fit (1:1) (Fig. 6A). As discussed above (see Section 3.1.1), the only robust variable is the water depth. Both seasonality and regional variability between ocean basins were not taken into account for the following analyses. Given the limited sampling across ocean basins and seasons, the robustness of observed differences in MM parameters between these categories remains inconclusive, necessitating reliance on globally derived MM parameters across depth intervals, with depth-dependent trends emerging as the sole reliable factor. The linear regression slopes are equal to one (1±0.006) for samples obtained from depths between 0-50, >50-100, and >300 to 1000 m, and change to slopes of 0.96±0.01 for a depth between 100 to 300 m (Fig. 6A and Tab. S3). The coefficient of determination ranges between 0.90 and 0.94 for the depth intervals between 0 m and 1000 m (Fig. 6A). The skill of the model was also tested by using training dataset samples (including the outliers) and the results show that the obtained slopes are higher than the fit obtained without the outliers (Tab. S3). To validate our model calibrated without including the outliers, this size-normalized catch model was applied on the validation dataset (Retailleau, 2009), with different size fractions (100-150 µm, >150-315 µm, and >315 µm) containing cytoplasm-filled and empty planktonic Foraminifera species tests. The slopes of the obtained regression lines from the observed vs. predicted abundances (i.e., our model estimates) range between 1 and 2.1 (Fig.
The coefficient of determination is higher at depth intervals between 100 m and 1000 m (between 0.79 and 0.83) than at shallower depth above 100 m (between 0.62 and 0.74). This correlation is attributed to the prevalence of large-shelled species at the surface rather than at depths as stated in Retailleau et al. 2011.

3.2.2. Abundance estimation using Berger’s “equivalent catch”

The correction factor coined by Berger (1969) - Eq. 6 – and termed “equivalent catch” was tested to the training dataset, and includes samples with known abundances for test sizes >100 µm and split into subsamples (see Section 2.1.1). The obtained regression lines from the predicted versus the observed abundances show a positive relationship (Fig. 6B). Normalizing the size class >125-200 µm to 100 µm results in a slope ranging between 1 and 1.3 and a coefficient of determination between 0.67 to 0.98, indicating that the predicted abundances are overestimated compared to the ones produced by the size-normalized catch model proposed in this study. However, normalizing the size class >200 µm to 100 µm results in a slope smaller than 1 (from 0.22 to 0.92) and a low coefficient of determination that ranges between 0.26 and 0.77, i.e., the predicted abundances are underestimated when using the Berger (1969) formula (Fig. 6B).

The equivalent catch approach by Berger (1969) was applied to the validation dataset from Retailleau, (2009) and Retailleau et al. (2011), wherein subsample abundances were aggregated into distinct size classes (>150 µm and >315 µm). Subsequently, the cumulative sample abundances were normalized to >100 µm using the Berger (1969) equation. Normalizing the size class >150-315 µm to 100 µm yielded a slope ranging from 1.5 to 8.8 and a coefficient of determination between 0.06 to 0.94, suggesting an overestimation of predicted abundances compared to those obtained via the FORCIS size-normalized catch model (Fig. 6B).

4. Discussion

4.1. Importance of depth-dependent corrections

Our study underscores the intimate connection between MM parameters and the size and distribution of planktonic Foraminifera throughout the water column, elucidating their ecological dynamics. Our analysis reveals that foraminifera assemblages from deeper water depths have a lower multiplication factor across all size classes compared to those from shallower depths. This finding aligns with the observation that planktonic foraminifera
abundances decrease with increasing water depth. On the other hand, large specimens are less abundant in surface than subsurface assemblages (Tab. 3). The MM parameters $f_{\text{max}}$ (maximum growth potential) and $S_{\text{half}}$ (sharpness of growth) are test-size dependent. Multiplication factors vs. size class rapidly increases for small species occurring only in the two smallest size classes, and flattens out toward the larger size classes. This explains the low $S_{\text{half}}$ values assessed for the different depth ranges of the small species. Whereas the cumulative abundance of the small species is slightly decreasing with depth, the $S_{\text{half}}$ does not change across the depth ranges (Fig. 4A and tab. 4A), likely because small size species exist at all water depth ranges at their size limit (maximum size). For the large species, the assessed $f_{\text{max}}$ and $S_{\text{half}}$ are high and increase with depth compared to small and medium-sized species. This implies that the larger, thus heavier specimens of a species, sink below the average depth of habitat of the entire population. This has been shown, for example, for *G. truncatulinoides*, which reach adult size below 100 m while their average living depth ranges above 100 m (Schiebel et al., 2002; Rebotim et al., 2017).

Overall, small specimens dominate the total planktonic Foraminifera assemblage present in the water column, confirming previous observations (Brummer and Kroon, 1988, Meilland et al., 2021, Schiebel et al. 2002). Generally, the spatial and vertical distribution of the planktonic Foraminifera of varying size and life stages (juvenile to mature specimens) in the surface ocean is changing due to mixing and passive vertical migration (less than 50 % of all species that migrate, Meillard et al., 2021). Furthermore, reproduction of different species, partly responsible for their shell sinking, may not happen at the same depth (Hemleben et al., 1989; Schiebel and Hemleben, 2005; Meillard et al., 2021). Also, planktonic Foraminifera species’ depth habitats are not the same across the different oceanic basins (Fairbanks et al., 1982, 1980). After reproduction, empty tests rapidly descend to the seafloor (Takahashi and Bé, 1984; Schiebel and Hemleben 2000). Tests of small and preadult individuals that died without reproducing slowly sink through the water column due to their low weight, being susceptible to dissolution while their cytoplasm lowers buoyancy (Bé et al., 1980; Bé and Hemleben, 1970; Erez, 2003; Iwasaki et al., 2019; Ofstad et al., 2021; Schiebel, 2002; Schiebel et al., 2007). Mass mortality of offspring shortly after reproduction may be caused by lack of food or/predation (Brummer and Kroon, 1988), and may not affect assemblages discussed here. Both juvenile mortality and reproduction at random depth for a part of the population explains the presence of small specimens (100-150 µm) in the entire surface mixed layer of the ocean (Meillard et al., 2021;2022).
The medium size classes (200-300 μm) encompass various ontogenetic stages, including both mature adults and specimens that have not yet reached full maturity. The abundance within these classes exhibits slight variations with depth. Furthermore, it seems that the decline of abundance with depth is related to dissolution (Schiebel et al., 2007), habitat preference (Jonkers and Kucera, 2015; Mortyn and Charles, 2003), or specimens that reduce in number from the system, and, for example, end up in predator guts (Bradbury et al. 1971; Brand and Lipps 1982). In waters deeper than 300 m, almost all planktonic Foraminifera tests are empty (Berger 1969; Schiebel, 2002). Our findings highlight the importance of using selected depth normalization MM parameters to account for the differences in planktonic Foraminifera populations.

4.2. The FORCIS correction – application, limitation and benefits for the entire plankton community

The FORCIS size-normalized catch model was applied to normalize the total abundance of the CPR and plankton tow abundances data of the FORCIS database to a 100 μm mesh size (Fig. 8). The constants $f_{max}$ and $S_{half}$ for the FORCIS abundance-size corrections were used according to the sample depth and for the total number of specimens (cytoplasm-filled and empty specimens) from all size classes (Tab. 2). When this size-normalized catch model is applied to the FORCIS database at all depths, using the MM parameters generated for each depth interval, and binned at each 4.3° latitude by 8.6° longitude, the resulting map (Fig. 8) shows that the standardized abundances normalized to 100 μm mesh size is between 1 and $10^3$ ind/m$^3$. In addition, applying the size-normalized catch model proposed in this study on independent dataset that have not been used to generate the applied model, the predicted abundances are close to the observed ones (Fig. 7A). The linear regression between the observed and predicted abundances presents slopes very close to 1 (from 1 to 1.4), and coefficients of determination higher than 0.46. This precision is somewhat reduced at the surface due to larger specimens, particularly near continental shelves, which may reflect inhibited reproduction in neritic conditions (Retailleau et al., 2011).

When applying the Berger (1969) correction (Eq. 6), the estimated abundances are generally overestimated compared to the FORCIS abundance-size correction (Figs. 6B and 7B). The FORCIS abundance-size correction method, developed in this study for planktonic Foraminifera offers a reliable ubiquitous approach to estimate the abundance of foraminifera
>100 µm across water depth ranges to normalize abundances of planktonic Foraminifera retrieved using different sieve or mesh sizes. The applicability of this correction method can be successfully transferred to a wide range of other planktonic groups, including radiolaria, diatoms, acantharia, and dinoflagellates, that follow similar cumulative abundance vs. size distribution as the planktonic Foraminifera in the water column, whose sampling techniques are also based on a variety of plankton nets mesh sizes. However, the FORCIS correction method would need a calibration step on a size-fractionated subset for each new plankton group, and may provide accurate and standardized estimates of abundances. This may enable better comparison and analyses across studies and contribute to an improved understanding of the ecological significance in marine ecosystems.

However, the technique has some limitations. For instance, if the abundance is null in one of the size fractions >100µm (i.e. \( C_k = 0 \)), it is not possible to derive the abundance using Equation S2 in the adjacent size-fraction bin and impede estimation of the total abundance (\( C_{\text{tot}} \)). When deriving the FORCIS size-normalized catch model for a specific taxon, it is essential to carefully select the suitable model. Particularly, the inclusion of taxa with diverse size spectra and adequate occurrence becomes imperative. Complications may arise when dealing with rare species or taxa that exhibit narrow size ranges. For instance, the species *Orcadia riedeli* is a rare planktonic Foraminifera species (Brummer and Kroon, 1988; Holmes, 1984; Meilland et al., 2018) with a narrow size spectrum between 100 µm and 150 µm in the FORCIS dataset. When applying the size-normalized catch model to *O. riedeli* to reconstruct the coefficients of multiplication \( f_k \) for each size fraction (100-125 µm and >125-150 µm) and water depth interval, the obtained curve is flat with a very low multiplication factor. The challenge lies in accurately predicting the abundance of small species when working with data primarily focused on large sizes. Applying a generic equation for small species may not suffice due to potential limitations in observing them within the larger size classes. Therefore, the model proposed here produces best results for size-abundance on taxa with a wide size spectrum, ranging from juveniles to adult individuals. By including data across different life stages, a more comprehensive understanding of the size-number relationship may be achieved.

**Final comments and recommendations**

Planktonic Foraminifera test size varies according to different parameters such as species, ontogenetic stage, and ecological factors including trophic conditions and water depth habitat. The FORCIS database, used in this study, assembles abundances of specimens collected with
different devices and mesh sizes over the past 125 years, which affects the estimation of their
abundance. To mitigate discrepancies arising from these varied sampling methods, a correction
factor was derived from abundance data of size-fractionated planktonic Foraminifera samples
that underwent sieve-splitting post-collection. The resulting relationships present a general
asymptotic shape, which we model using Michaelis-Menten-like fits across different water
depth intervals (0-50 m, >50-100 m, >100-300 m, >300-1000 m, and >1000 m). Using these
relationships, the abundance of planktonic Foraminifera larger than 100 µm can be assessed to
normalize the global abundance dataset of the FORCIS database.

The methodology proposed here proves to be more accurate than the initial correction scheme
proposed by Berger (1969), leveraging a much larger calibration dataset. Unlike Berger’s
method, which tends to overestimate abundance in the 125 µm to 200 µm size fraction, and
underestimate in the larger than 200 µm size fractions, our comprehensive training dataset
enables the development of a size-normalized catch model applicable not only to Foraminifera
but potentially to other planktonic groups.

Emphasizing the adoption of a standard 100 µm size criterion for future planktonic
Foraminifera research could eliminate the need for such corrections, promoting consistency and
comparability across studies. This standardization is critical for advancing our understanding
of planktonic Foraminifera populations within marine ecosystems.

Authors contributions

The study was designed by S.C., X.G., T.G.-T., R.S., J.M, G.B., L.J., and G.M. The data
analysis was carried out by S.C. and X.G. generated the equations. All authors contributed to
the interpretation and discussion of the results. S.C. wrote the paper with contributions from

Acknowledgements

The FORCIS project is supported by the French Foundation for Biodiversity Research (FRB)
[https://www.fondationbiodiversite.fr/] within the Centre for the Synthesis and Analysis of
Biodiversity (CESAB) [https://www.fondationbiodiversite.fr/la-fondation/le-cesab/] and co-funded by INSU LEFE program, and the Max Planck Institute for Chemistry (MPIC) in Mainz.

**Data Availability Statement**

All data associated with this article and codes to generate the FORCIS size-normalized catch model are available Zenodo through this link: https://doi.org/10.5281/zenodo.7437719

**Competing interests**

The authors declare that they have no conflict of interest.

**References**


https://doi.org/10.2307/1485406


King, J.E., and Demond, J. 1953. Zooplankton abundance in the Central Pacific. Fish. Bull. 82,


Figure captions

Fig. 1. Sampling location of the FORCIS data, training dataset and Retailleau et al. 2011 data.

Fig. 2. (A) The cumulative abundance frequency vs. the size fractions (100-125 µm, >125-150 µm, >150-200 µm, >200-250 µm, >250-315 µm, >315-400 µm, >400-500 µm, and >500 µm). Each black line represents a sample. (B) Box plot representing the multiplication factor curves calculated using Equation (1) vs. each size fraction (>100-125 µm, >125-150 µm, >150-200 µm, >200-250 µm, >250-315 µm, >315-400 µm, >400-500 µm, and >500 µm). Black dots are the outliers.

Fig. 3. Schematic illustration of the size-normalized catch model. (A) Cumulative abundances vs. size fractions. (B) Multiplication factors vs. size fractions. (C) Multiplication factors and translations to determine the parameters of the Michaelis-Menten fit.

Fig. 4. (A) Cumulative abundance per plankton tow sample of the training dataset calculated at each size fraction (100-125 µm, >125-150 µm, >150-200 µm, >200-250 µm, >250-315 µm, >315-400 µm, >400-500 µm, and >500 µm) per each sample (black lines). Line colors indicate depth intervals. (B) Multiplication factor curves calculated using Equation (1) vs. size fraction in each sample (black lines). The general Michaelis Menten fit was obtained on the global data (dashed blue line), and for each depth range.

Fig. 5. (A) Multiplication factors with confidence interval derived from the cumulative abundance curves plotted versus the different size fractions (100-125 µm, >125-150 µm, >150-200 µm, >200-250 µm, >250-315 µm, and >315 µm) at each depth interval, season, and ocean
basin (Arctic, North Atlantic, and Indian Oceans), (B) in the North Atlantic Ocean at two different depths and during summer, (C) in the North Atlantic Ocean during autumn and summer at 0 to 50 m depth, and (D) in the North Atlantic and Indian Oceans in autumn at 0 m to 50 m water depth. Envelopes represent confidence intervals at the limits of 2.5 % and 97.5 %.

**Fig. 6.** Comparison of observed abundances from the training datasets at various size fractions (>125-150 µm, >150-200 µm, >200-250 µm, >250-315 µm, and >315 µm) against estimated abundances obtained using size-normalized catch model proposed for different water depth intervals (0-50 m, 50-100 m, 100-300 m, and 300-1000 m) (A) and, linear regressions between observed abundance the training dataset data in the size fraction above 100 µm and abundances obtained with mesh sizes above 125 µm (dark blue line), 150 µm (purple line), 200 µm (pink line), 250 µm (green line), and 315 µm (light blue line) across different depth ranges (0-50 m, >50-100 m, >100-300 m, and >300-1000 m), standardized to 100 µm using Equation (6) (Berger, 1969) (B). Red lines represent the linear regressions of all fractions from the training datasets at the different water depth intervals.

**Fig. 7.** Observed abundances from an independent dataset (Retailleau et al. 2011) at each size fraction (>150-315 µm, and >315 µm) vs. the estimated ones obtained using the size-normalized catch model proposed for each depth interval (from 0-50 m, 50-100 m, >100-300 m, and >300-1000 m) (A) and, linear regressions between observed abundance from an independent dataset (Retailleau et al. 2011) in the size fraction above 100 µm and abundances obtained with mesh sizes above 150 µm (purple line) and 315 µm (light blue line) across different depth ranges (0-50 m, >50-100 m, >100-300 m, and >300-1000 m), standardized to 100 µm using Equation (6)
Red lines represent the linear regressions of all fractions from the independent dataset (Retailleau et al. 2011) at the different water depth intervals.

**Fig. 8.** Global (FORCIS data) planktonic Foraminifera abundance averaged over 4.3° latitude by 8.6° longitude bands over the entire water column (0-1000 m) obtained from plankton nets and CPRs, standardized to 100 µm using the size-normalized catch model proposed in this study.

**Supplementary material**

**Figure captions**

**Fig. S1.** Foraminifera abundance (A) and cumulative abundance (B) of a fractionated single sample “M10-2_218B1_S1” by sieving into six size classes (100-125 µm, >125-150 µm, >150-200 µm, >200-250 µm, >250-315 µm, and >315) and collected from the North Atlantic Ocean at a depth between 80 and 100 m (Schiebel et al., 2002).

**Fig. S2.** Sampling locations of all the data included in FORCIS (Chaabane et al., 2023) collected using plankton tows, pumps, and CPRs. Different mesh sizes are given in different colors (see legend).

**Fig. S4.** A scheme describing the workflow of the model conception, validation and test with Berger approach and application of the size-normalized catch model.
Supplementary material

Appendix #1: Size-normalized catch model

Equations presented in Section 2 are generic formulas, which can be applied for any size-fraction normalization and abundance in any size-class fraction. Here, we present how these size-normalization catch formulas were derived from a subset of the FORCIS database.

Each sample is subdivided in \( k_{\text{max}} = 6 \) subsamples of contiguous size classes. The cumulative abundance of a class \( k \) is the sum of abundances of all subsamples of finer size classes: \( C_{\text{cum},k} = \sum_{i=1}^{k} C_i \).

As defined in Section 2, the multiplication factor of a class \( k, f_k \), is the cumulative abundance divided by the abundance of the first class, so that

\[
f_k = \frac{\sum_{i=1}^{k} C_i}{C_1} ,
\]

where \( C_i \) is the abundance of size class \( i \), and the difference between two consecutive multiplication factors is

\[
f_k - f_{k-1} = \frac{\sum_{i=1}^{k} C_i - \sum_{i=1}^{k-1} C_i}{C_1} = \frac{C_k}{C_1} 
\]

Eq. S1

and

\[
f_1 = \frac{C_k}{f_k - f_{k-1}} 
\]

Eq. S2

Applying Eq. 1 to the last size class (\( k = k_{\text{max}} \)), we can express the total abundance over the entire size fraction range as

\[
C_{\text{tot}} = C_{\text{cum},k_{\text{max}}} = C_1 f_{k_{\text{max}}} = C_1 f_{\text{max}} 
\]

Eq. S3

Finally, combining Eqs. 2 and 3, the total abundance can be expressed based on the abundance of any class \( k \), and the multiplication factors framing this class (\( f_k \) and \( k_{k-1} \)):

\[
C_{\text{tot}} = \frac{C_k}{f_k - f_{k-1}} f_{\text{max}} 
\]

Eq. S4

The generalization of this formula is illustrated in Figure 3 and presented in Eq. 3.

Tables

Tab. 1. Planktonic Foraminifera species included in the FORCIS dataset and categorized by their typical adult size: <200 \( \mu \text{m} \) classified as small, 200 - 300 \( \mu \text{m} \) as medium, and > 300 \( \mu \text{m} \) as large." (Brummer & Kuccera, 2022, Schiebel and Hemleben, 2017, Meilland et al. 2021, Meillard et al. 2022).

<table>
<thead>
<tr>
<th>Species name</th>
<th>Average adult size (( \mu \text{m} ))</th>
<th>Adult size category</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. pumilio</td>
<td>110</td>
<td>Small</td>
</tr>
<tr>
<td>Species</td>
<td>Size</td>
<td>Size Group</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------</td>
<td>------------</td>
</tr>
<tr>
<td>D. anfracta</td>
<td>150</td>
<td>Small</td>
</tr>
<tr>
<td>G. cavernula</td>
<td>170</td>
<td>Small</td>
</tr>
<tr>
<td>G. minuta</td>
<td>110</td>
<td>Small</td>
</tr>
<tr>
<td>G. rubescens</td>
<td>150</td>
<td>Small</td>
</tr>
<tr>
<td>G. tenellus</td>
<td>160</td>
<td>Small</td>
</tr>
<tr>
<td>G. uvula</td>
<td>160</td>
<td>Small</td>
</tr>
<tr>
<td>G. vivans</td>
<td>150</td>
<td>Small</td>
</tr>
<tr>
<td>O. riedeli</td>
<td>110</td>
<td>Small</td>
</tr>
<tr>
<td>T. clarkei</td>
<td>110</td>
<td>Small</td>
</tr>
<tr>
<td>T. fleisheri</td>
<td>130</td>
<td>Small</td>
</tr>
<tr>
<td>T. humilis</td>
<td>140</td>
<td>Small</td>
</tr>
<tr>
<td>T. iota</td>
<td>140</td>
<td>Small</td>
</tr>
<tr>
<td>T. parkerae</td>
<td>140</td>
<td>Small</td>
</tr>
<tr>
<td>T. quinqueloba</td>
<td>180</td>
<td>Small</td>
</tr>
<tr>
<td>C. nitida</td>
<td>300</td>
<td>Medium</td>
</tr>
<tr>
<td>G. bulloides</td>
<td>300</td>
<td>Medium</td>
</tr>
<tr>
<td>G. conglobatus</td>
<td>300</td>
<td>Medium</td>
</tr>
<tr>
<td>G. elongatus</td>
<td>250</td>
<td>Medium</td>
</tr>
<tr>
<td>G. falconensis</td>
<td>280</td>
<td>Medium</td>
</tr>
<tr>
<td>G. glutinata</td>
<td>220</td>
<td>Medium</td>
</tr>
<tr>
<td>G. hexagona</td>
<td>250</td>
<td>Medium</td>
</tr>
<tr>
<td>G. hirsuta</td>
<td>300</td>
<td>Medium</td>
</tr>
<tr>
<td>G. inflata</td>
<td>300</td>
<td>Medium</td>
</tr>
<tr>
<td>G. ruber any</td>
<td>250</td>
<td>Medium</td>
</tr>
<tr>
<td>Species</td>
<td>Parameter</td>
<td>Size</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------</td>
<td>-------</td>
</tr>
<tr>
<td><em>G. ruber ruber</em></td>
<td>250</td>
<td>Medium</td>
</tr>
<tr>
<td><em>G. scitula</em></td>
<td>230</td>
<td>Medium</td>
</tr>
<tr>
<td><em>G. theyeri</em></td>
<td>280</td>
<td>Medium</td>
</tr>
<tr>
<td><em>P. obliquiloculata</em></td>
<td>300</td>
<td>Medium</td>
</tr>
<tr>
<td><em>T. sacculifer</em></td>
<td>300</td>
<td>Medium</td>
</tr>
<tr>
<td><em>N. incompta</em></td>
<td>200</td>
<td>Medium</td>
</tr>
<tr>
<td><em>N. pachyderma</em></td>
<td>200</td>
<td>Medium</td>
</tr>
<tr>
<td><em>G. conglomerata</em></td>
<td>300</td>
<td>Medium</td>
</tr>
<tr>
<td><em>G. crassaformis</em></td>
<td>300</td>
<td>Medium</td>
</tr>
<tr>
<td><em>N. dutertrei</em></td>
<td>300</td>
<td>Medium</td>
</tr>
<tr>
<td><em>G. tumida</em></td>
<td>280</td>
<td>Medium</td>
</tr>
<tr>
<td><em>G. ungulata</em></td>
<td>300</td>
<td>Medium</td>
</tr>
<tr>
<td><em>G. cultrata</em></td>
<td>320</td>
<td>Large</td>
</tr>
<tr>
<td><em>G. truncatulinoides</em></td>
<td>400</td>
<td>Large</td>
</tr>
<tr>
<td><em>B. digitata</em></td>
<td>400</td>
<td>Large</td>
</tr>
<tr>
<td><em>G. adamsi</em></td>
<td>350</td>
<td>Large</td>
</tr>
<tr>
<td><em>G. siphonifera</em></td>
<td>320</td>
<td>Large</td>
</tr>
<tr>
<td><em>H. digitata</em></td>
<td>800</td>
<td>Large</td>
</tr>
<tr>
<td><em>H. pelagica</em></td>
<td>800</td>
<td>Large</td>
</tr>
<tr>
<td><em>O. universa</em></td>
<td>600</td>
<td>Large</td>
</tr>
<tr>
<td><em>G. calida</em></td>
<td>320</td>
<td>Large</td>
</tr>
<tr>
<td><em>S. dehiscens</em></td>
<td>320</td>
<td>Large</td>
</tr>
</tbody>
</table>

**Tab. 2.** Estimation parameters $f_{\text{max}}$ and $S_{\text{half}}$ generated for each depth interval using the training dataset.
<table>
<thead>
<tr>
<th>Depth interval (m)</th>
<th>$S_{half}$ (µm)</th>
<th>$f_{max}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-1000</td>
<td>178 ± 3</td>
<td>2.48 ± 0.02</td>
</tr>
<tr>
<td>0-50</td>
<td>184 ± 5</td>
<td>2.75 ± 0.05</td>
</tr>
<tr>
<td>50-100</td>
<td>186 ± 7</td>
<td>2.54 ± 0.06</td>
</tr>
<tr>
<td>100-300</td>
<td>171 ± 6</td>
<td>2.30 ± 0.05</td>
</tr>
<tr>
<td>300-1000</td>
<td>167 ± 6</td>
<td>2.18 ± 0.05</td>
</tr>
</tbody>
</table>

**Tab. 3** Coefficients of multiplication for each size fraction and depth, and obtained using the Eq. 4 on the training dataset.

<table>
<thead>
<tr>
<th>Size fraction max (µm)</th>
<th>$f_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-50 m</td>
</tr>
<tr>
<td>&gt;100-125</td>
<td>1</td>
</tr>
<tr>
<td>&gt;125-150</td>
<td>1.52</td>
</tr>
<tr>
<td>&gt;150-200</td>
<td>1.97</td>
</tr>
<tr>
<td>&gt;200-250</td>
<td>2.19</td>
</tr>
<tr>
<td>&gt;250-315</td>
<td>2.33</td>
</tr>
<tr>
<td>&gt;315</td>
<td>2.44</td>
</tr>
</tbody>
</table>

**Tab. 4.** Michaelis Menten constants $f_{max}$ and $S_{half}$ generated number of subsamples used for each water depth interval and species adult size for small (A), medium (B), and large (C) species.

A. Small species

<table>
<thead>
<tr>
<th>Depth interval (m)</th>
<th>$S_{half}$ (µm)</th>
<th>$f_{max}$</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-50</td>
<td>149.08±4.65</td>
<td>1.57±0.03</td>
<td>595</td>
</tr>
<tr>
<td>&gt;50-100</td>
<td>150.01±7.49</td>
<td>1.55±0.05</td>
<td>320</td>
</tr>
<tr>
<td>&gt;100-300</td>
<td>144.6±6.11</td>
<td>1.53±0.05</td>
<td>351</td>
</tr>
</tbody>
</table>
B. Medium species

<table>
<thead>
<tr>
<th>Depth interval (m)</th>
<th>( S_{\text{half}} ) (µm)</th>
<th>( f_{\text{max}} )</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-50</td>
<td>187.43±5.85</td>
<td>3.30±0.07</td>
<td>1444</td>
</tr>
<tr>
<td>&gt;50-100</td>
<td>196.44±9.40</td>
<td>3.27±0.11</td>
<td>831</td>
</tr>
<tr>
<td>&gt;100-300</td>
<td>178.86±7.08</td>
<td>3.10±0.09</td>
<td>793</td>
</tr>
<tr>
<td>&gt;300-1000</td>
<td>180.11±6.88</td>
<td>3.02±0.08</td>
<td>024</td>
</tr>
</tbody>
</table>

C. Large species

<table>
<thead>
<tr>
<th>Depth interval (m)</th>
<th>( S_{\text{half}} ) (µm)</th>
<th>( f_{\text{max}} )</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-50</td>
<td>251.16±17.87</td>
<td>4.87±0.24</td>
<td>375</td>
</tr>
<tr>
<td>&gt;50-100</td>
<td>257.99±27.43</td>
<td>5.15±0.39</td>
<td>144</td>
</tr>
<tr>
<td>&gt;100-300</td>
<td>253.80±35.74</td>
<td>4.56±0.44</td>
<td>142</td>
</tr>
<tr>
<td>&gt;300-1000</td>
<td>244.12±25.13</td>
<td>4.59±0.33</td>
<td>175</td>
</tr>
</tbody>
</table>
**Supplementary Material**

**Tables**

**Tab. S1.** Summary of data counts used for the training dataset at each ocean basin, season, water depth interval, and size fraction.

**Tab. S2.** Michaelis Menten constants $f_{\text{max}}$ and $S_{\text{half}}$ generated for each water depth interval, season and ocean basin.

<table>
<thead>
<tr>
<th>Ocean basin</th>
<th>Season</th>
<th>Depth interval (m)</th>
<th>$f_{\text{max}}$</th>
<th>$S_{\text{half}}$ (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arctic</td>
<td>summer</td>
<td>0-50</td>
<td>1.83±0.07</td>
<td>151.60±9.31</td>
</tr>
<tr>
<td>Arctic</td>
<td>summer</td>
<td>&gt;100-300</td>
<td>2.24±0.04</td>
<td>138.45±2.41</td>
</tr>
<tr>
<td>Arctic</td>
<td>summer</td>
<td>&gt;300-1000</td>
<td>1.95±0.27</td>
<td>141.16±25.34</td>
</tr>
<tr>
<td>Indian</td>
<td>autumn</td>
<td>0-50</td>
<td>1.58±0.11</td>
<td>191.07±0.58</td>
</tr>
<tr>
<td>Indian</td>
<td>autumn</td>
<td>&gt;50-100</td>
<td>1.38±0.68</td>
<td>142.67±74.61</td>
</tr>
<tr>
<td>Indian</td>
<td>autumn</td>
<td>&gt;100-300</td>
<td>1.56±0.02</td>
<td>134.49±3.30</td>
</tr>
<tr>
<td>Indian</td>
<td>autumn</td>
<td>&gt;300-1000</td>
<td>2.39±0.35</td>
<td>152.24±46.47</td>
</tr>
<tr>
<td>Indian</td>
<td>spring</td>
<td>0-50</td>
<td>2.56±0.11</td>
<td>200.27±16.47</td>
</tr>
<tr>
<td>Indian</td>
<td>spring</td>
<td>&gt;50-100</td>
<td>2.50±0.14</td>
<td>181.13±16.73</td>
</tr>
<tr>
<td>Indian</td>
<td>spring</td>
<td>&gt;100-300</td>
<td>2.32±0.21</td>
<td>205.08±35.91</td>
</tr>
<tr>
<td>Indian</td>
<td>summer</td>
<td>0-50</td>
<td>2.38±0.08</td>
<td>181.62±10.81</td>
</tr>
<tr>
<td>Indian</td>
<td>summer</td>
<td>&gt;50-100</td>
<td>2.68±0.28</td>
<td>190.96±33.28</td>
</tr>
<tr>
<td>Indian</td>
<td>summer</td>
<td>&gt;100-300</td>
<td>2.10±0.05</td>
<td>149.13±5.86</td>
</tr>
<tr>
<td>Indian</td>
<td>winter</td>
<td>0-50</td>
<td>2.16±0.15</td>
<td>227.92±33.64</td>
</tr>
<tr>
<td>Indian</td>
<td>winter</td>
<td>&gt;50-100</td>
<td>1.97±0.10</td>
<td>195.45±22.32</td>
</tr>
<tr>
<td>Indian</td>
<td>winter</td>
<td>&gt;100-300</td>
<td>1.97±0.09</td>
<td>179.37±17.23</td>
</tr>
<tr>
<td>Indian</td>
<td>winter</td>
<td>&gt;300-1000</td>
<td>1.83±0.14</td>
<td>163.57±24.79</td>
</tr>
<tr>
<td>North Atlantic</td>
<td>autumn</td>
<td>&gt;300-1000</td>
<td>1.78±15.88</td>
<td>170.76±0.08</td>
</tr>
<tr>
<td>Depth interval (m)</td>
<td>Slope</td>
<td>std error</td>
<td>r^2</td>
<td>p value</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------</td>
<td>-----------</td>
<td>------</td>
<td>---------</td>
</tr>
<tr>
<td>0-50</td>
<td>1.1</td>
<td>0.007</td>
<td>0.90</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Class</td>
<td>Value</td>
<td>p-value</td>
<td>r value</td>
<td>p value</td>
</tr>
<tr>
<td>----------------</td>
<td>-------</td>
<td>---------</td>
<td>---------</td>
<td>----------</td>
</tr>
<tr>
<td>&gt;50-100</td>
<td>1</td>
<td>0.01</td>
<td>0.86</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>&gt;100-300</td>
<td>4.5</td>
<td>0.05</td>
<td>0.88</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>&gt;300-1000</td>
<td>3.4</td>
<td>0.04</td>
<td>0.85</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Without outliers

<table>
<thead>
<tr>
<th>Class</th>
<th>Value</th>
<th>p-value</th>
<th>r value</th>
<th>p value</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-50</td>
<td>1</td>
<td>0.006</td>
<td>0.92</td>
<td>&lt;0.001</td>
<td>1480</td>
</tr>
<tr>
<td>&gt;50-100</td>
<td>1</td>
<td>0.007</td>
<td>0.94</td>
<td>&lt;0.001</td>
<td>961</td>
</tr>
<tr>
<td>&gt;100-300</td>
<td>0.96</td>
<td>0.01</td>
<td>0.90</td>
<td>&lt;0.001</td>
<td>821</td>
</tr>
<tr>
<td>&gt;300-1000</td>
<td>1</td>
<td>0.01</td>
<td>0.92</td>
<td>&lt;0.001</td>
<td>969</td>
</tr>
</tbody>
</table>
Fig. 2
A  
Cumulative abundance (ind.m$^{-3}$)

$C_{sz\_inf}^{\infty} = C_{sz\_sup}^{125}$

B  
Multiplication factor

$f_{max}$

$C_{sz\_norm}^{\infty} = C_{sz\_inf}^{125} \frac{f_{max} - f_{sz\_norm}}{f_{sz\_sup} - f_{sz\_inf}}$

C  
Translated multiplication factor

$f' = f - 1$

$f'_{max} = f_{max} / 2$

$f' = f'_{max} \frac{Sz'}{Sz' + S'_{half}}$

$S'_{half} = Sz - 125$

$Sz' = Sz - 125$

$S'_{half} = S_{half} - 125$

$S_{half} = 125$

$Sz = Sz - 125$

$S'_{half} = S_{half} - 125$

$Sz = Sz - 125$

$S_{half} = 125$

$S'_{half} = S_{half} - 125$

$Sz = Sz - 125$

$S_{half} = 125$

$S'_{half} = S_{half} - 125$

$Sz = Sz - 125$

$S_{half} = 125$

$S'_{half} = S_{half} - 125$

$Sz = Sz - 125$

$S_{half} = 125$

$S'_{half} = S_{half} - 125$

$Sz = Sz - 125$

$S_{half} = 125$

$S'_{half} = S_{half} - 125$

$Sz = Sz - 125$

$S_{half} = 125$

$S'_{half} = S_{half} - 125$

$Sz = Sz - 125$

$S_{half} = 125$

$S'_{half} = S_{half} - 125$

$Sz = Sz - 125$

$S_{half} = 125$

$S'_{half} = S_{half} - 125$

$Sz = Sz - 125$

$S_{half} = 125$

$S'_{half} = S_{half} - 125$

$Sz = Sz - 125$

$S_{half} = 125$

$S'_{half} = S_{half} - 125$

$Sz = Sz - 125$

$S_{half} = 125$

$S'_{half} = S_{half} - 125$

$Sz = Sz - 125$

$S_{half} = 125$

$S'_{half} = S_{half} - 125$

$Sz = Sz - 125$

$S_{half} = 125$

$S'_{half} = S_{half} - 125$

$Sz = Sz - 125$

$S_{half} = 125$

$S'_{half} = S_{half} - 125$

$Sz = Sz - 125$

$S_{half} = 125$

$S'_{half} = S_{half} - 125$

$Sz = Sz - 125$

$S_{half} = 125$

$S'_{half} = S_{half} - 125$

$Sz = Sz - 125$

$S_{half} = 125$

$S'_{half} = S_{half} - 125$

$Sz = Sz - 125$

$S_{half} = 125$

$S'_{half} = S_{half} - 125$

$Sz = Sz - 125$

$S_{half} = 125$

$S'_{half} = S_{half} - 125$
Fig. 4
Fig. 5
Fig. 6

This study

\[ y = x \]
\[ r^2 = 0.92, \ p < 0.001 \]

Berger 1969 approach

\[ y = x \]
\[ r^2 = 0.94, \ p < 0.001 \]

\[ y = 0.96x \]
\[ r^2 = 0.90, \ p < 0.001 \]

Observed abundance (ind/m$^3$)

Predicted abundance based on size-normalized catch model (ind/m$^3$)

Predicted abundance based on Berger 1969 Approach (ind/m$^3$)

Observed abundance (ind/m$^3$)

Depth (m)

B >125 µm

B >150 µm

B >200 µm

B >250 µm

B >315 µm
This study

\[ y = 1.1 \times \]

\[ r^2 = 0.74, p < 0.001 \]

Berger 1969 approach

\[ y = x \]

\[ r^2 = 0.62, p < 0.001 \]

\[ y = 2.1 \times \]

\[ r^2 = 0.79, p < 0.001 \]

\[ y = 1.7 \times \]

\[ r^2 = 0.83, p < 0.001 \]
Fig. S1

A. Histogram showing the abundance of entities of different sizes (in µm) from 100 to 315 µm. The vertical axis represents the abundance (ind/m³), and the horizontal axis represents the size (µm).

B. Cumulative abundance plot showing a gradual increase in cumulative abundance with size. The size categories are 100-125, 125-150, 150-200, 200-250, 250-315, and >315 µm.
Fig. S2
Training dataset:
- Counts of living and dead specimens
- Multinet: from 0 to 1000 m
- 6 size fractions

Size-normalized catch model

Michaelis-Menten fitting function
- $F_{\text{max}}$: Maximum potential abundance
- $S_{\text{half}}$: Initial response rate or sharpness

Validation
- Counts of living and dead specimens
- Collected using multinet
- 3 size fractions
- Different depths

Test with Berger's Scheme
- Standardize samples size from >100 um to infinity to 100 um to infinity

Application of size-normalized catch model
- Counts of living and dead specimens
- Collected using multinet, pump and CPR
- Different depths

FORCIS database

Fig. S3