Response to reviewers for "What controls planktic foraminiferal calcification?" by Barrett et al., submitted to Biogeosciences.

We thank the reviewers for their insightful and constructive comments on our manuscript. Please see our response to these comments in the following, which we hope you agree, strengthens our paper. Reviewer comments are shown in bold, our responses below in normal text, and actioned responses in red text.

Reviewer #1: Brian Huber

The paper is very well written and it provides excellent observations that demonstrate that planktonic foraminiferal size normalized weight is highly variable among species, in different regions, and in response to different environmental variables. Results clearly demonstrate that size normalized should not be used as a pCO2 proxy. The paper is very well written, well organized and well-illustrated, and it merits publication and wide distribution among the research community. My edits are only minor grammatical corrections, and I have no concerns about the methodology, results, interpretations or conclusions.

We thank the reviewer for their comments, and have updated the manuscript throughout following their suggestions for improved grammar. Regarding the use of size-normalised weight as a pCO_2 proxy, we would like to clarify that although its use is not as straight forward as initially postulated as there are clearly other factors affecting weight, we do not suggest that carbonate chemistry has no impact.

Reviewer 1: I'm not sure what is meant by services [at line 34 of the original manuscript: "The unprecedented rise in CO_2 and temperature is altering our oceans and impacting marine ecosystems and their services."]

As per feedback from the reviewer and the editor we have changed the wording slightly and expanded by giving an example of an ocean ecosystem service at line 35-36 of the revised manuscript: *"…marine ecosystems and their functioning (such as marine biogeochemical cycles)"*

Reviewer 1: is there supposed to be an x axis label here [figure 2a of original manuscript] or just go by the colours on the legend?

We have added x axis labels to figure 2a make it clear to the reader that the boxplots refer to the different size fractions.

Reviewer #2 : Anonymous

This study by Barrett et al. examines the use of size-normalized weight (SNW) of planktic foraminifera shells as a proxy for reconstructing past environmental conditions, particularly seawater CO_2 levels (pCO_2). Using global data and Bayesian regression modeling, the authors find that no single environmental factor explains SNW variability across species and regions. Instead, species-specific and regionally variable responses suggest that cryptic species and phenotypic plasticity, such as changes in calcite thickness during reproduction, may influence shell weight. The study emphasizes the importance of regional calibration and careful species selection when using SNW as a pCO_2 proxy.

I find this study important and timely, as it effectively reminds the community that foraminifera shell weights are not an absolute indicator of seawater carbonate chemistry in the open ocean's natural conditions. However, I do have two major concerns about the study's methodology. First, although the authors, in their introduction and methodology, consider most of the factors that have been proposed over the years to explain variations in foraminifera shell weight, they do not account for recent studies that suggest changes in seawater density as a potential driver of these variations. I understand that seawater density and salinity may covary (in the surface ocean), but I still believe this factor is worth investigating. Second, the authors compare the SNW of deep-dwelling foraminifera species to surface ocean properties, which raises some concerns.

We thank the reviewer for their feedback and their acknowledgement of the importance and timeliness of our research. First we address their two primary concerns before responding to minor comments.

1. More specifically, Zarkogiannis et al. (2019) suggest that foraminifera utilize their shells for buoyancy regulation, adjusting their shell weight to maintain their position in the water column. In a subsequent study, Zarkogiannis et al. (2022) also discuss gametogenic calcite as a mechanism for buoyancy regulation. They observe that if foraminifera shell weights were primarily governed by CO_3^{2-} concentrations, deep-dwelling species residing in more acidic waters would exhibit the lightest shells. However, this is not the case. Their findings align with both the current study and that of Béjard et al. (2023), in which *G. truncatulinoides*, a deep-dwelling species, is observed to possess the heaviest shells. Moreover, Zarkogiannis et al. (2022) provide an additional 16 core-top samples that could be incorporated into the present study. As both CT and volumetric data are provided, size normalization to a linear dimension should be feasible.

We agree with Reviewer 2 that we should include reference to a study that suggests a seawater density control on SNW. As such, we add text to the introduction (line 64 of revised manuscript) *"....and seawater density* (Zarkogiannis et al., 2019)*"*

We stand by our decision not to include seawater density in our Bayesian modelling as although seawater density and shell weight will (in part) control the position of foraminifera in the water column, lipids and the shape of the test also impact buoyancy (Caromel et al., 2014; Schiebel and Hemleben, 2005). While the link to buoyancy is an intriguing suggestion, there are many open questions, such as why an adult would need to add the weight if this is a regulation i.e., a driven process, or could the weight in *G. truncatulinoides* be linked to its unique life cycle. Due to this complexity, it is beyond the scope of this study to investigate seawater density.

At the time of revising the manuscript, we contacted the author of the suggested paper requesting this additional data but did not receive a response.

2. My second concern pertains to the environmental data used for comparison with the SNWs. How surface were they? 0, 2,5, 5 or 10m? Additionally, do the authors believe that comparing surface ocean conditions with the SNWs of *G. truncatulinoides*, a species typically found at depths below 400 meters, is appropriate? What is the underlying assumption in comparing deep-dwelling individuals with surface ocean conditions? If not at bibliographic species-specific calcification depths, I would have expected to see at least a comparison with the averaged conditions between 0 and 100 meters depth, extracted from the models, to be used for comparison with the SNWs.

We agree with Reviewer 2 that ideally our environmental data would be from the exact living depth for each species. However, the reason we have not been able to compare SNW with this idealised environment data at depth is in part due to the challenge of estimating exact habitat depth. This changes through a foraminifera's life time hence it would be difficult to know which depth is most suitable (Schiebel and Hemleben, 2017). Even if depth was known, there is uncertainty in how much calcification happens at which depth and high resolution in situ analysis of proxies would be needed, though they would need to be calibrated for ontogenetic signals and not bulk calcite which is the current practice. Ideally, we would dissolve the foraminifera to extract oxygen isotope values and back calculate the correct habitat depth. However, this was beyond the scope of the current analysis. Furthermore, while several taxa are deep dwelling species and live below the thermocline, this depth is different in different parts of the ocean (Mulitza et al., 1997) and varies with the seasons (Waterson et al., 2017). Hence while we appreciate that species-specific calcification depths and their variation across locations would be the ideal approach to understand the multiple, competing drivers of SNW, it was beyond the scope of this manuscript to determine these exact depths for specimens.

We consider the use of sea surface data (≤ 20 m) for shallow dwelling species reasonable because the niche of shallow dwellers is largely dependent on sea surface temperature (Waterson et al., 2017). However, we acknowledge that our manuscript would benefit from clearly stating our choice in methodology regarding use of sea surface data. As such, we have added text to the manuscript that provides more information on the depth of the environmental data and a paragraph which summarises the above text on problems associated with getting at exact habitat depth (lines 166-175). We recognise the importance of this comment and as such have done some additional investigation to get at this problem of depth for deeper dwellers (Text S3).

<u>**Minor**</u>

Line 27: Change "are a plankton" to "are a type of plankton." Done.

Line 62: What about seawater density? We have added a seawater density reference, Zarkogiannis et al. (2019), to the introduction (line 64 revised manuscript).

Line 108: What are ecogroups? Consider mentioning the Aze et al. (2011) classification here, where ecogroups are first introduced. Are these ecogroups the same as those used in the group-level comparison? If so, why not refer to this comparison as ecogroup-level? Ecogroup is defined on line 70. As per your suggestion we have added Aze et al. (2011) to the definition here and have added this reference to the table at line 156 of the revised manuscript.

The group-level analysis does not refer to ecogroup-level. Instead it refers to all foraminifers being pooled together to examine whether across-species there is a universal driver of calcification. This is first defined at line 243, but we recognise the need for further clarification. As such, throughout the revised manuscript we occasionally remind the reader that group-level refers to the across species analysis that is all foraminifers in this study pooled together.

Lines 125-126: I think Marshall et al. (2013) should be cited here, as they introduced area density as a normalization method against silhouette area. We have added this reference.

Line 150: Define ESMs at this point in the text. Done.

Line 171: A reference for phosphate is missing. Thank you for bringing this to our attention. We have added Demes et al., 2009; Kinsey and Davies, 1979; Lin and Singer, 2006; Paasche and Brubak, 1994.

Line 214: I am unsure if salinity is appropriate, as its depth profile varies with latitude. In the halocline (within the first 1000 m, relevant to foraminifera), salinity increases with depth at high latitudes but decreases with depth at low latitudes. Furthermore, there is a salinity inversion in subtropical regions. If SNWs followed the salinity profile of the water column in the subtropics, there would likely be a decline in SNWs with depth, but this is not observed.

We are confused about the point the reviewer is making here regarding the methods part of the paper. We have treated all parameters as equally possible. For the water depth we are considering, the reviewer is correct that in some locations in the high latitudes salinity is inverted due to ice melt and there can be small reductions in salinity in the surface waters of the subtropics due to freshwater injections but we do not understand how this is relevant here in the testing of drivers. As we have measurements at every location, we would consider the reduced salinity in the surface at these places.

Lines 247-248: This explanation belongs in the methods section. It is the first time that the rationale behind using ESM data is addressed. We have moved the sentence to the methods (line 161-163 of the revised manuscript) which introduces this rational.

Line 261: What happened to the merging of the 250–350 µm sieve fractions? It is unclear why the sieve fractions are separated in one instance and merged in another. Please clarify this point.

We wanted to give the highest level of granularity possible here, and hence separated the size fractions when possible. We were able to do the separation in this part of the analysis as for this qualitative analysis, the number of data points is not so important. In contrast, in the statistical Bayesian analysis we had to combine size fractions to increase the size of the dataset being modelled.

To clarify our reasoning, we have added text to section 3.2 lines 312-315: "*To take out size fraction bias, all size fractions other than 250-300 µm and 300-350 µm have been removed and these two remaining size fractions have been merged to create a dataset sufficient for statistical analysis. Unless stated otherwise, the following statistics have been performed on this reduced dataset.*"

Lines 264-267: *G. truncatulinoides* has variants like *excelsa*, which are hardly differentiated, while de Vargas et al. (2001) and Quillévéré et al. (2013) identify four distinct types. Therefore, it is challenging to make assumptions based on data from a single study. Zarkogiannis et al. (2022) also present *G. truncatulinoides* shell weights, volumetric data for size normalization, and CT images for variant identification, which could be useful for the present study.

The four distinct types the reviewer mentioned are genetic types not morphologically expressed as explored in de Vargas et al (2001). Therefore, the genetic knowledge of ecological adaptation cannot be used on the morphological specimen alone. The two types which can be separated are subtropical and high latitude. The Bejard study is based on specimens from the western Mediterranean which contains only one of the morphotypes (II) as per Quillévéré et al. (2013). Therefore the reference to different genetic types is not relevant here. It is further unclear to us why the genetic to morphological differentiation is mentioned here exclusively for *G. truncatulinoides* given that this is the case for many taxa (e.g. Morard et al. (2024) for a recent summary). A key theme in our manuscript is reminding future SNW researchers to consider cryptic species, so we feel that a comment is already acknowledged.

Regarding the use of the data presented in the 2022 study, as stated above, the author has not provided sufficient meta data to include their study into our analysis and at the time of revising the manuscript had not responded to the data request. Though we appreciate the usefulness of variant identification in this dataset, its addition would not change the statement made at lines 264-267 in the original manuscript regarding the variability in the size fraction 400-500 um, as the foraminifers in (Zarkogiannis et al., 2022) were sieved at the 300-350 um size fraction.

Line 281: The sentence appears incomplete; something seems to be missing at the end.

We have reworded the sentence to clarify (lines 307-310 of revised manuscript): *"As such, although the smaller size fractions are meaningful in polar and subpolar areas (as foraminifers are smaller at the poles), they must be interpreted with caution in warm, high calcite saturation regions where including smaller size fractions might result in the selection of species which have not undergone a full developmental cycle…"*

Line 345: Do you mean "ecogroup level"? If so, please revise for consistency.

No – see above comment on model definitions related to line 108 in the original manuscript. For clarity, we have added: *"Though in our group-level model (i.e., all foraminifers)"* to line 411-412.

Line 400: Is there any collinearity in the current dataset between CO32- and phosphate, as suggested by Marshall et al. (2013)?

Assessment of collinearity is covered in section 2.3.2 Model Specification. There is no collinearity in the current dataset between CO_3^{2-} and phosphate. Furthermore, the Bayesian models have been run with QR decomposition analysis which reduces impacts of correlation between variables within the models. To check for any remaining collinearity pairs plots were visually assessed, and variance inflation factors (VIF) were verified using the package 'performance' which passes the brms model to its frequentist counterpart. A structureless i.e., "blob" like output for the visual pairwise plots indicate no collinearity between carbonate and phosphate (figure 1) and the VIF scores are under 10 (1.08 for carbonate and 1.09 for phosphate) which indicates that collinearity is not problematic (Marcoulides and Raykov, 2019).

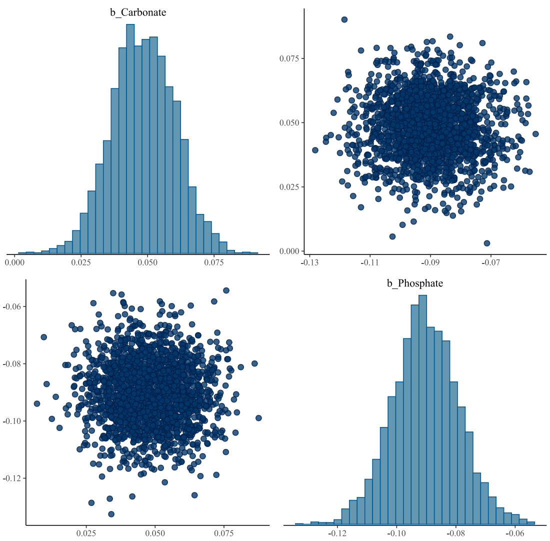

Figure 1 Pairwise plot showing no collinearity between carbonate and phosphate

Line 452: Also strongly agreed. However, I did not see any raw shell weights presented in this study. The Pangea links are missing.

Apologies, the Pangaea links are currently missing as the data is awaiting a DOI. Until then, reviewers can access raw data and model code using a temporary link to the University of Bristol's research data storage facility which we share with the editor.

One aspect not discussed in this section or elsewhere in the paper is the cleaning protocols used for foraminifera shell weight measurements in different studies. This aspect should eventually be standardized. As shown in the HyPerCal cleaning protocol, different treatments incorporate varying amounts of sedimentary contamination, which can affect SNW, particularly for specimens with large and multiple apertures, such as *G. ruber* plexus.

Yes, we agree this is important when measuring SNW. Based on CT data, the contamination is not important in specimens with larger apertures such as *G. ruber* as sediment can easily be removed. A larger amount of sediment is found in specimen of the globorotaliids, as it is harder for the sediment to be removed when washing.

Unfortunately, we were unable to add this level of complexity (i.e., cleaning protocol) to the modelling, as we are currently limited in how complex the models can be due to the limited sample size. Our hope is that once community agreed protocols for SNW data collection are widely used and data management improves, the size of the useable dataset will increase to support such analysis.

For now, we have added a comment on the cleaning protocol to the discussion where we discuss other methodological limitations (lines 521-524): "*It would also be useful for authors to report their foraminifera cleaning protocol, or even better the community agree on a standardised cleaning method as different methods can result in variable sedimentary contamination, which impacts the weight of specimens…*"

Furthermore, in a future submission please change planktic to planktonic. The correct adjective form of plankton is planktonic. The adjectives of Greek nouns ending in -on get the suffix -ic in the end like plankton – planktonic, bion – bionic, lacon – laconic. This is different to nouns ending in -os, which lose the ending -os to the previous consonant by replacing it with -ic, like bentos – benthic, cosmos – cosmic or chronos – chronic.

While we appreciate the clear knowledge of Greek by the reviewer, it is also clear that the root of the word has been lost and both terms are used by the community.

Reviewer #3: Pincelli Hull

I agree with the first two readers that this is an excellent, well-written manuscript that describes the results of a careful study that works to synthesize new and existing results. The current state of the literature on shell normalized weight can be generously described as a confusing and contradictory, so this is a refreshing paper to read indeed.

I have four suggestions that I do think are important to incorporate to make this current contribution clear, in terms of its findings and implications and in allowing this study to be useful for future research.

We thank the reviewer for their feedback. In the following we response to the reviewers four suggestions.

1. **Data and Code Availability: Like Reviewer 2, I went looking for the data and couldn't find the raw data, something they explicitly implored other studies to make available. I suspect this is embargoed at present, but as a reviewer, this was unclear. Making the code available as well, would make this study truly replicable and should be done (and maybe the authors already have!).**

   Apologies, the Pangaea links are currently missing as the data is awaiting a DOI. Until then, reviewers can access raw data and model code using a temporary link to the University of Bristol's research data storage facility which we share with the editor.

2. **Table of Statistical models and diagnostics: My apologies if I am missing this, but I was surprised not to see a table of the different models with diagnostics on relative model fit. The authors describe their models in words, and some of the results, but do not provide a table of these models and results. It is fine if model results are included in the supplement only, but they should be included.**

   Thank you for this suggestion. We include extensive diagnostics as part of our R Markdown supplement, but the link to this was missing in the preprint. This is now available via a link to the University of Bristol's research data storage facility which we share with the editor. However, for clarity we have also added a table in the supplementary material (Table S3). Please see a description of this table below.

   <u>Model structure:</u> We have added detail on the structure of the models in table 1. The reader can also see the supplementary R script which contains the code needed to reproduce our analyses.

   <u>Model diagnostics:</u>

   1) Model convergence: We do not add the Rhat diagnostics to the table as this would be space inefficient. Instead, for the sake of brevity we have added text to the methodology (lines 238-239): *"An R-hat value close to 1 (i.e., less than 1.1) indicates the chains have converged. All models had an R-hat of 1.01 or 1…"*
   2) Collinearity: In table 1 we have included the variance inflation factors (VIF) scores for each model and their associated tolerance intervals (TI). A VIF of less than 10 and TI of >0.1 indicate no collinearity.

   <u>Model fit:</u>

   1) Kernel density estimate plots. Figure S6 in the supplementary information include the kernel-density estimates, which shows the goodness of fit for the individual models. The closer that "yrep" is to "y" means the better the model was able to reproduce the original data distribution. This indicates that all models have a reasonable fit.
   2) Table S4 in the supplementary information reports the effect size and 95% credible interval, which is equivalent to the data shown in figure 3 and 4 of the original manuscript.
   3) We have added Bayes R2 values for all models in table S3.

3. **Handling and reporting of the fixed effect 'sampling type': Because I was involved in one study that found that the most important predictor of shell normalized weight was whether the test was from a sediment trap or core top, I was heartened to see this included as a fixed effect (and indeed, you discuss the literature regarding the potential importance of this factor). I was thus confused as to why the importance of this condition was not reported, nor shown, in any of the figures. I appreciate that you would want to consider the effect of this factor relative to other fixed effects, but to not report it at all nor show its effects seems like a missed opportunity. Without knowing its relative importance, it is hard to know what to do with the results that a framed in terms of the conditions present in the surface ocean around the time that the individual was alive. Here is where more information is needed on this fixed effect:**

a) **Figure 1: symbol type could show whether the same was a trap or from the ocean bottom**

Thanks for this suggestion. We have made this change to figure 1. We also add a supplementary figure (S1) which details the number of samples from each data type.

b) **Results: please report the variance explained by the null model (i.e., fixed effects including sampling type). I assumed from the methods and from the use of the phrase 'environment-only' model that the reported model with 23% of variance explained excluded the sampling method fixed effect. Did the model with random effects include sampling method? It is confusing without a table of model results to refer to (see pt #2).**

The reviewer is correct that the environment-only model in the original manuscript (i.e., the 23%) excluded the sampling method fixed effect. In the original manuscript, the model with random effects (i.e., the 86%) did not include sampling method. We wanted to disentangle the impact of adding species, but we acknowledge that assessing the importance of sampling method is of interest and hence have expanded the analyses by including the following text to the results (lines 318-324):

"A model that is "environment only" explains 20% of the variability in SNW (Bayes R2; Table S3; Gelman et al., 2019). The addition of sampling method (i.e., the "null model") improves model performance ($\widehat{elpd}_{loo}$ improved by 114.4 [±23.7]) and explained variance increases to 60% (Table S3). The "full" model (i.e., environment, sampling method and species) performs better than the "null" model ($\widehat{elpd}_{loo}$ improved by 247.5 [±19.4]) and explained variance increases to 90% (Table S3). Together, this shows that the choice of sampling method can influence the SNW recorded and that species-specific responses are important in determining SNW."

c) **Figures: Please add this fixed-effect to Figure 2,3,4, so the effect and directionality of the effect can be understood.**

1. **Update figure 2:** In the interest of figure readability, we have added this updated figure to the supplementary information (figure S9) and referenced it in the main text (line 303).

2. **Update figure 3 and 4:** Because the fixed effect output for sampling method is categorical, the Bayesian model output presents the impact of sampling method as relative to another sampling method. Therefore, we do not think it is appropriate to add to figure 3 and 4. Instead we create a separate figure (figure S10) and add corresponding text to lines 467-479.

d) **Discussion: given that the model includes this factor, you are in the position to discuss the relative importance (and overprinting) of diagenesis in the sediment versus conditions during life. I know that I, and a few others, would love to hear your thoughts on this, given the results—once presented!**

Thanks to the reviewers comments, we have been able to discuss the impact of sampling method on SNW in more depth (section 3.5, lines 466 -487 and figures S10 and S11):

4. **Depth of Environmental factors: it is unclear from the methods (but could be clear in a data table) at what depth(s) environmental factors were extracted and considered from the models. Were these all from the 'sea surface'?**

   In the methodology, we have provided clarity about what depth is extracted. *"we use surface ocean environmental data (≤ 20 m depth)"* (line 167). Please see the earlier discussion in response to comment #2 from Reviewer 2 on limitations associated with exploring conditions in subthermocline habitats.

   Are bottom water conditions considered for the sediment samples?

   We do not use bottom water conditions for sediment samples. Please see lines 480-487 for discussion related to this.

   Are deeper depth conditions considered for those species that live at deeper depths? Would this change the coherence and direction of results for those taxa?

   Please see our response to Reviewer 2 regarding the depth of environmental data (comment #2).

If these general concerns could be addressed, as I suspect they might readily be, this will make an excellent contribution in my opinion! I list a minor note below and I look forward to seeing the final version!

Minor: Line 170: change 'inhibits' to 'inhibit'

We have changed this.

Associate editor comments:

I think that the authors should move the text S1 to the main text.

Done. Please see lines 137-147.

Line 34: maybe you can rephrase this into "…marine ecosystem and their functioning.." to improve clarity.

We have reworded so line 36 in the revised manuscript reads as suggested.

Figure 1a: as shown it is the only plot in the figure with a title. Please add "sieve size fraction" as an axis label.

Done.

A note on collinearity

During this review because we now consider the impact of sampling method more in depth, we identified collinearity in two species-level models: *G. truncatulinoides* and *G. elongatus*. Additionally, as stated in the original manuscript we previously removed *N. incompta* from analyses due to collinearity. However, in the interest of sharing these data we have included these species in our analyses by adapting our methodology to use principle component analysis (PCA) for these three species. By using PCA to reduce the dimensionality of the environmental data, we are able to eliminate the problem of collinearity and investigate the impact of sampling method on these species, and can in-part understand the response of SNW to environmental drivers. We add text on the PCA to the supplementary (Text S2), and included Table 3 and relevant text to the manuscript.

Reference list

Aze, T., Ezard, T. H. G., Purvis, A., Coxall, H. K., Stewart, D. R. M., Wade, B. S., and Pearson, P. N.: A phylogeny of Cenozoic macroperforate planktonic foraminifera from fossil data, Biol. Rev. Camb. Philos. Soc., 86, 900–927, https://doi.org/10.1111/J.1469-185X.2011.00178.X, 2011.

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

Demes, K. W., Bell, S. S., and Dawes, C. J.: The effects of phosphate on the biomineralization of the green alga, Halimeda incrassata (Ellis) Lam., J. Exp. Mar. Bio. Ecol., 374, 123–127, https://doi.org/10.1016/J.JEMBE.2009.04.013, 2009.

Gelman, A., Goodrich, B., Gabry, J., and Vehtari, A.: R-squared for Bayesian Regression Models, Am. Stat., 73, 307–309, https://doi.org/10.1080/00031305.2018.1549100, 2019.

Kinsey, D. W. and Davies, P. J.: Effects of elevated nitrogen and phosphorus on coral reef growth, Limnol. Oceanogr., 24, 935–940, https://doi.org/10.4319/LO.1979.24.5.0935, 1979.

Lin, Y. P. and Singer, P. C.: Inhibition of calcite precipitation by orthophosphate: Speciation and thermodynamic considerations, Geochim. Cosmochim. Acta, 70, 2530–2539, https://doi.org/10.1016/J.GCA.2006.03.002, 2006.

Marcoulides, K. M. and Raykov, T.: Evaluation of Variance Inflation Factors in Regression Models Using Latent Variable Modeling Methods, Educ. Psychol. Meas., 79, 874, https://doi.org/10.1177/0013164418817803, 2019.

Morard, R., Darling, K. F., Weiner, A. K. M., Hassenrück, C., Vanni, C., Cordier, T., Henry, N., Greco, M., Vollmar, N. M., Milivojevic, T., Rahman, S. N., Siccha, M., Meilland, J., Jonkers, L., Quillévéré, F., Escarguel, G., Douady, C. J., Garidel-Thoron, T. de, Vargas, C. de, and Kucera, M.: The global genetic diversity of planktonic foraminifera reveals the structure of cryptic speciation in plankton, Biol. Rev., https://doi.org/10.1111/BRV.13065, 2024.

Mulitza, S., Dürkoop, A., Hale, W., Wefer, G., and Niebler, H. S.: Planktonic foraminifera as recorders of past surface-water stratification, Geology, 25, 1997.

Paasche, E. and Brubak, S.: Enhanced calcification in the coccolithophorid Emiliania huxleyi (Haptophyceae) under phosphorus limitation, Phycologia, 33, 324–330, https://doi.org/10.2216/I0031-8884-33-5-324.1, 1994.

Quillévéré, F., Morard, R., Escarguel, G., Douady, C. J., Ujiié, Y., de Garidel-Thoron, T., and de Vargas, C.: Global scale same-specimen morpho-genetic analysis of Truncorotalia truncatulinoides: A perspective on the morphological species concept in planktonic foraminifera, Palaeogeogr. Palaeoclimatol. Palaeoecol., 391, 2–12, https://doi.org/10.1016/J.PALAEO.2011.03.013, 2013.

Schiebel, R. and Hemleben, C.: Modern planktic foraminifera, Paläontologische Zeitschrift, 79, 135–148, https://doi.org/10.1007/bf03021758, 2005.

Schiebel, R. and Hemleben, C.: Planktic Foraminifers in the Modern Ocean, Springer, Berlin, 154 pp., https://doi.org/10.1007/978-3-662-50297-6_1, 2017.

de Vargas, C., Renaud, S., Hilbrecht, H., and Pawlowski, J.: Pleistocene adaptive radiation in Globorotalia truncatulinoides: genetic, morphologic, and environmental evidence, Paleobiology, 27, 104–125, 2001.

Waterson, A. M., Edgar, K. M., Schmidt, D. N., and Valdes, P. J.: Quantifying the stability of planktic foraminiferal physical niches between the Holocene and Last Glacial Maximum, Paleoceanography, 32, 74–89, https://doi.org/10.1002/2016PA002964, 2017.

Zarkogiannis, S. D., Antonarakou, A., Tripati, A., Kontakiotis, G., Mortyn, P. G., Drinia, H., and Greaves, M.: Influence of surface ocean density on planktonic foraminifera calcification, Sci. Rep., 9, 1–10, https://doi.org/10.1038/s41598-018-36935-7, 2019.

Zarkogiannis, S. D., Iwasaki, S., Rae, J. W. B., Schmidt, M. W., Mortyn, P. G., Kontakiotis, G., Hertzberg, J. E., and Rickaby, R. E. M.: Calcification, Dissolution and Test Properties of Modern Planktonic Foraminifera From the Central Atlantic Ocean, Front. Mar. Sci., 9, 864801, https://doi.org/10.3389/FMARS.2022.864801/BIBTEX, 2022.