Review for Authors
Modeling seasonal and vertical habitats of planktonic foraminifera on a global scale

Kretschmer et al.

This paper builds upon preexisting work modeling planktonic foram distributions in the global oceans via a coupling to CESM's ocean model. The goal is to better understand how the vertical distribution of foramnifera species varies seasonally and throughout larger climatic changes in the ocean. The paper is generally well written, clear, and broadly does a fine job demonstrating the usefulness of the model. It is also very thorough in its examination of the model's performance against available data. The methods seem robust and I can recommend that with some minor revisions (mostly grammar and clarity) the paper be published in Biogeosciences.

I must acknowledge that I am not an expert on the biogeochemistry of planktonic forams in any way and hope the other reviewers can address the methods and paramterizaitons employed in this paper in particular. I can instead comment on the benefit of this work and the need for such proxy system models for the robust interpretation of paleoceanographic records via the use of PLAFOM2.0 + CESM1.2. To that end, my first major comment is that the authors can focus more in the introduction and conclusion on the body of literature developing forward models, or proxy system models, for understanding paleoclimate proxies and introduce this work as a part of this group of literature. A major effort has been underway to build proxy system models, link them with GCMs, and make these models publicly available, and this paper is absolutely in this category and should make as much clear.

See for example:

- Dee, S., et al. "PRYSM: An open-source framework for PRoxY System Modeling, with applications to oxygen-isotope systems." Journal of Advances in Modeling Earth Systems 7.3 (2015): 1220-1247.

- Evans, Michael N., et al. "Applications of proxy system modeling in high resolution paleoclimatology." *Quaternary Science Reviews* 76 (2013): 16-28.

- Schmidt, Gavin A. "Forward modeling of carbonate proxy data from planktonic foraminifera using oxygen isotope tracers in a global ocean model." *Paleoceanography* 14.4 (1999): 482-497.

You might also consider mentioning (in the intro or discussion) the potential for PLAFOM to assist in data assimilation exercises for periods extending back further than the last millennium, for example. A number of papers look at the impacts of using process-based models in the DA framework and this is another application of your model. See work of Hugues Goosse's lab (e.g. Goosse, Hugues, et al. "Reconstructing surface temperature changes over the past 600 years using climate model simulations with data assimilation." *Journal of Geophysical Research: Atmospheres* 115.D9 (2010)) , as well as:

- Steiger, Nathan J., et al. "Assimilation of time-averaged pseudoproxies for climate reconstruction." *Journal of Climate*27.1 (2014): 426-441.
- Dee, Sylvia G., et al. "On the utility of proxy system models for estimating climate states over the common era." *Journal of Advances in Modeling Earth Systems* 8.3 (2016):

- Hakim, Gregory J., et al. "The last millennium climate reanalysis project: Framework and first results." *Journal of Geophysical Research: Atmospheres* 121.12 (2016): 6745-6764

In Section 4, it would be nice if the authors could provide a more quantitative data-model comparison technique—you identify areas where the model does not well simulate the observations and Figure 2 summarizes this to some extent, but perhaps you could include an additional table or figure or even compute something like the RMSE for each oceanic province? Or the mean RMSE for each species over all of the locations where core-top data exist?

Finally, in the discussion, you assert (correctly) that your new model is a powerful tool for separating the independent influences of habitat and climate on foram reconstructions. I think this paper would be greatly strengthened by a demonstration of this. Can you take a well-known and vetted reconstruction and apply this model in a meaningful way to reassess the climatic interpretation? I think this would show the power of forward modeling in this field to make more robust assessments of uncertainties in oceanic climate changes… And I think having this demonstration would add weight to the assertions you make in your Discussion section.

Minor / Line by Line comments: (Page-Line)

2-10 awkward paragraph break, consider revising
2-13 comma after perspective,
2-20 Have you investigated/reviewed Schmidt et al., 1998, 1999? These papers I believe address vertical migration of foram species in the water column—worth checking/citing if appropriate.

@article{Schmidt1998,
        Author = {Schmidt, Gavin A},
        Title = {{Oxygen-18 variations in a global ocean model Relationships between $\delta^{18}O$ and S}},
        Volume = {25},
        Year = {1998}}

@article{Schmidt1999,
        Author = {Schmidt, Gavin a.},
        Journal = {Paleoceanography},
        Number = {4},
        Pages = {482},
        Title = {{Forward modeling of carbonate proxy data from planktonic foraminifera using oxygen isotope tracers in a global ocean model}},
        Url = {http://www.agu.org/pubs/crossref/1999/1999PA900025.shtml},
        Volume = {14},
        Year = {1999},

2-26 need comma after behavior. Check for needed commas and small grammatical errors throughout text.

3-6 comma after estimate,
3-13 this phrase is awkward, revise ("with the biogeochemical model being enabled")

3-15 change "aimed for" to 'aimed to'

3-16 change "at geologic timescales" to "ON geologic timescales"

Check for similar awkward language throughout.

3-23 comma after configuration,

3-30 no paragraph break.

4-9 what do you mean by 'data models' for the atmosphere, etc.? Are you not using the fully coupled simulations and using some kind of statistical representation of the other components?

Heading 2.4 consider changing this to "Coupled GCM Setup" ?

7-15 missing space before new sentence.

8-21 comma after 'life cycle,'

Throughout section 3, be extremely clear about whether you are referring to observations vs. the model simulation of foram distributions/abundances etc. The reader gets a bit lost in the data-model comparison here unless that's super clear.

16-29 no comma after 'data'

16-30 this is a run-on sentence—consider shortening/rewriting

I appreciate the thorough discussion of the model – data comparison limitations on page 17.


Figure 5 has some strange cropping issues along top margin.