

Response to Referee #3:

Ref.: Ms. No. bg-2017-429

Title: Modeling seasonal and vertical habitats of planktonic foraminifera on a global scale

We would like to thank reviewer Jelle Bijma for his constructive comments and suggestions, which will help us to greatly improve our manuscript. Based on the comments of all four reviewers we will prepare a new version of our manuscript as outlined below.

However, during the review process, we discovered an error in the underlying ocean model. Unfortunately, the ocean circulation is not correctly represented in the used coarse resolution (i.e., $\sim 3^\circ$) model configuration. For a correct representation of the ocean and to yield scientifically consistent results, we had to perform a new model run with a higher horizontal resolution (i.e., $\sim 1^\circ$) on a supercomputing system. This model run takes ca. 5 weeks and is currently in the final production phase. At a first glance, the new results will not differ that much from our previous results as the representation of the upper ocean, where the analyzed foraminiferal species live, was actually reasonably well simulated in the coarse resolution model configuration compared to, e.g., the World Ocean Atlas 2013. We expect that the distribution of only a few species might be affected, when using the higher resolution model configuration with a more realistic representation of the ocean physics. Since we have not yet obtained the final results, we were not always able to provide detailed answers to your comments and had to keep our responses rather general.

Please find, in the following, the original comments in black and our responses in light blue; the indicated page and line numbers refer to the previously submitted manuscript.

Referee #3 comments:

Scientific significance: Excellent

The manuscript by Kretschmer et al. represents a substantial contribution to scientific progress within the scope of Biogeosciences. It is the latest one in a series of “foram-flux modelling” papers from the Bremen group. In 2006, Zaric et al. Developed the first empirical model that described globally the fluxes of planktonic foraminifera at species level in dependence of sea-surface temperature, mixed-layer depth and export production. Over the years, the foram model itself, its parameterization, and its implementation and coupling to other models has evolved (e.g. Fraile et al., 2008; 2009; Kretschmer et al., 2016). The aim of all of these papers has always been to project the effect of changing environmental conditions on species distributional patterns in time and space. The current paper adds a vertical dimension to the existing foram model by applying the previously used spatial parameterization of biomass as a function of temperature, light, nutrition, and competition on depth-resolved parameter fields.

Scientific quality: good

The scientific approach and methods are valid. The results are discussed appropriately but the discussion lacks a critical analysis of the model-data comparison beyond the caveats mentioned in section 4.2 “Comparison with local observations”.

Please refer to our response to reviewer 2 for proposed additions to the discussion regarding this point.

Even though the model-data-comparison revealed several discrepancies and is subject to caveats, the model produces nonetheless seasonal and vertical abundance patterns that are consistent with our current understanding and which emerge without any explicit parameterization of abundance in time and space. These patterns emerge from the model itself.

In addition, each model component (i.e., POP2, BEC, PLAFOM2.0) of the used model configuration consists of a rather complex model structure itself and rendering sensitivity experiments will be very time-consuming, expensive and non-trivial. We find that for a first try we obtain very good results.

The authors write on p. 17 line 22-23: “This vertical migration of planktonic foraminifera during their ontogeny cannot be reproduced by PLAFOM2.0 as the model parameterizations do not include the individual species’ life cycles.”. It is quite understandable that implementing true reproduction cycles of cohorts of foraminifera, including “real” population dynamics and ontogenetic migration is beyond the present manuscript. Hence, the model does not calculate absolute or relative numbers of a certain species within a certain ontogenetic size class based on reproductive success and size specific growth- and mortality-rates, but rather calculates changes in species specific carbon concentration (in mmol C m⁻³), which can be converted to numbers afterwards.

There is nothing wrong with this approach but it means that the parameterization of PLAFOM2.0 is based on practical “sum” or “composite” parameters. These are then used to tune the model outcome to the overall data. For instance, growth of all species is approximated using a modified form of Michaelis-Menton kinetics in dependence of species specific food availability and temperature sensitivity (Fraile et al., 2008). To account for the light dependence with depth, influencing the growth of only symbiont bearing foraminifera, the authors included a “photosynthetic growth rate”. They use “.....a similar approach as Doney et al. (1996) and Geider et al. (1998), who determined phytoplankton growth rates by available light and nutrients..... (p.5 line 15-17)”. Such a parameterization is normally used for phytoplankton, that has orders of magnitude higher densities and cell division rates that respond very fast (within a day) and directly

to light and nutrients. The symbiont bearing forams in this manuscript obey a (semi) lunar reproduction cycle and occur in densities that are very much lower, such that a “phytoplankton” kind of response cannot be expected. The authors use it as an additional tuning parameter for symbiont bearing forams next to food preference and temperature to develop species specific depth (light/nutrient) habitat preferences. Although it is a valid approach, the authors should clearly state that it is artificial.

Here, we applied a similar approach as Doney et al. (1996) and Geider et al. (1998) as a first approximation to account for a photosynthetic growth rate for the symbiont-bearing species. We are aware that a phytoplankton kind of response to light is not transferable one to one to the response of planktonic foraminifera. We will make this more clear in the manuscript and we will state that this approach is a first approximation, and in that way it should be considered as rather artificial. Nevertheless, we also think that this is a valid approach, given that the photosynthetic growth rate accounts in numerical terms most likely only for the smallest proportion of the total growth.

Growth is balanced by mortality, which is not a formulation for “real” mortality but another tuning parameter: “we adjusted parts of the mortality rate equation to improve the model accuracy (p. 5 line 8-9).”.

Overall, there are many factors that allow tuning, e.g. “p% represents the fraction of photosynthesis contributing to growth (p.5 line 31)”. Interestingly, the authors have a higher p% for *T. sacculifer* (0.4) than for *G. ruber* (0.3), where I would have done it the other way around (see my comments on these species further below).

Here, we followed Lombard et al. (2011), who also used a somewhat higher p% for *T. sacculifer* (0.40) than for *G. ruber* (0.37). We performed a few short preliminary test runs using different p%-values but obtained the best results on a first glance by using the given parameter values. We were not able to perform a suite of sensitivity experiments with changing the p%-values due to the long runtime of the used complex model configuration.

Another tuning factor is the temperature dependence of the predation term: “.....we followed Moore et al. (2004) and adjusted the temperature dependence of the predation term (MLpred in mmolCm⁻³s⁻¹) (p.6 line 3-4). Also “...we included a stronger competitive behavior of *G. bulloides* by adjusting the free parameters in the competition term. (p.6 line 10-11). Having collected planktonic foraminifera by SCUBA diving for many, many years and looking at average typical blue water densities of ca. 10 specimens per m³ per species, and 3 dominant species in an assemblage, it is hard to believe that they compete with each other for resources as each of

them occupies a space of only a few mm³ and they are stationary in the water column.

A good point indeed. Whether or not planktonic foraminifera compete directly is a field of active research. However, we would like to point out that even though foraminifera occur at very low densities and may never directly meet, they are still likely to compete for scarce resources. It is therefore reasonable to include a competition term in the model.

Certain boundary conditions also correct model misfits, e.g. "...zero fluxes have been replaced by half of the observed minimum flux. (p.7 line 25-26)".

All of these parameters were introduced to allow a good fit between model output and data but maybe not for the right reason. As such, we do not know how realistic this parameterization represents real planktonic foraminiferal population dynamics which is more complex (including lunar based reproduction cycles, ontogenetic migration, etc.).

This is true, but nevertheless we are able to simulate the seasonal and vertical habitat of the five considered foraminiferal species remotely realistic using our approach. However, for a more realistic representation of planktonic foraminiferal population dynamics, PLAFOM2.0 needs to be extended by, e.g., considering the ontogenetic migration, reproduction cycles as well as additional foraminiferal species. Thus, PLAFOM2.0 will become more complex and more parameters have to be introduced. In addition, using, e.g., reanalysis data as forcing instead of a climatological forcing could also lead to a more realistic representation of the modern foraminiferal population dynamics when considering a point-by-point comparison with present-day data.

However, even if our understanding of foraminiferal population dynamics will be largely improved in the future due to, e.g., more laboratory experiments, and if we are able to properly translate those complex processes into model code, we will still only be able to provide an approximation of the real dynamics.

Winter mixing, thermocline shoaling and annual irradiation changes are probably important parameters controlling foram population dynamics just as certain density layers may be important for gamete fusion in real foram life. I'm not sure how well these features are implemented in the models.

This is a very good point and all those processes you mentioned likely affect the dynamics of the foraminiferal population. Here we used an ocean-ice-only model configuration and applied a climatological forcing to obtain our results. Hence, there is no explicit interaction between the ocean and the atmosphere and additionally an inter-annual variability of the forcing variables can be excluded. In addition, the lower the resolution of the ocean model the less well represented are

processes such as winter mixing, thermocline shoaling, and upwelling. Since we are now going to present results of a 1° ocean model simulation, most of these processes will likely be better represented than in the previously used 3° simulation, not only because of a more realistic representation of the ocean physics, but also due to the higher resolution, which could likely improve our model results. However, some small-scale processes, oceanic fronts, river runoff areas, and coastal upwelling regions might most likely still not be well represented. In order to analyze inter-annual variability of the foraminiferal population and to investigate how annual radiation changes influence the population dynamics, the model system should be forced with reanalysis data rather than climatologies. Furthermore, using a fully coupled model configuration initialized from reanalysis data could also provide information on how annual changes in the atmosphere feed back on the foraminiferal population dynamics. This, however, was beyond the scope of this study. Here, we actually aimed for an approach that is as simple and general as possible, such that we specifically avoided an explicit parameterization of depth. This way our approach is also easier to follow and we can more easily ensure the reproducibility of our study.

The bottom line is that, even though I appreciate the model and the manuscript a lot, I would like to see a discussion on these issues and if possible a statistical verification of the model performance. The description of the results and the discussion on modeled geographical ranges, seasonal and vertical distribution, as well as on the modeled seasonal variability of depth habitat, lacks a statistical treatment of the data. How good is the model performance and how sensitive is it to each of the model parameters?

Here, we did not perform a sensitivity study in regard of the different model parameters, first, because the runtime of this new model configuration is too long (with a model throughput of ~11-20 simulated years/day for the 3° model configuration and/or a model throughput of ~9.5 simulated years/day for the 1° model configuration depending on the machine capacities) to yield scientifically reasonable results and, second, because Fraile et al. (2008), who introduced PLAFOM (which is the base of PLAFOM2.0), already performed a sensitivity study of the free parameters. Fraile et al. (2008) modified the values chosen for the foraminifera module and quantified the sensitivity by calculating the change in the root mean square error between each sensitivity experiment and the standard run. They found that none of the parameters led to a uniform change for all species and that not surprisingly the parameter controlling the temperature tolerance range (i.e., σ) seems to be the most sensitive parameter (see Table 3 of Fraile et al., 2008). Since PLAFOM2.0 is in its base form identical to PLAFOM, we did not feel the need to perform another sensitivity analysis and also due to the high computational costs. However we will briefly discuss the sensitivity analysis of Fraile et al. (2008) in section 2 to assure that our applied approach is valid and that our results are in general reliable.

I would appreciate a more quantitative treatment of the model performance instead of statements like “The predicted global distribution patterns of the five considered planktonic foraminiferal species are in good agreement with the core-top data (Figure 2) (p. 11 line 14-15)?

To perform a more quantitative model-data-comparison and to provide some measure of confidence, we will now calculate the Bray-Curtis index of similarity between the model and the core-top data. For this calculation, we will account for the different sizes of each species by using a mean size for each species based on the results of Schmidt et al. (2004) and recalculate the modeled relative abundances accordingly. We will add this analysis to sections 3.1 and 4.1.1 to provide a more thorough model-data-comparison.

The discussion on the global distribution patterns is mostly related to temperature. What about the other parameters: food, nutrients, productivity, light, etc.?

Our results indicate that the habitat variability and the foraminiferal distributions are primarily driven by temperature and for the colder water species (*N. pachyderma*, *N. incompta*, *G. bulloides*) also by food supply. This was also shown by Fraile et al. (2008) and Kretschmer et al. (2016). Fraile et al. (2008) demonstrated that the foraminiferal distribution patterns respond most sensitively to changes in the temperature tolerance ranges of the individual species, indicating the strong temperature dependence of the foraminiferal population dynamics. Therefore, we mainly relate our results to temperature, but also discuss the food dependency extensively (see sections 4.1.3 and 4.2); the other parameters, however, seem to be less effective.

How does it compare to the “Longhurst Biogeographical Provinces”. He partitioned the world oceans into provinces (“Ecological Geography of the Sea”) based on the prevailing physical factors as a regulator of phytoplankton distribution, including temperature, photic depth, mixed layer depth etc. (e.g. Longhurst 1995; 1998).

The underlying parameterizations used in PLAFOM itself are based on the parameterizations used in the ecosystem model of Moore et al. (2002a) and do not include a spatial parameterization. Since Longhurst's partitioning of the ocean is more or less only descriptive, a comparison with our model results is in our understanding not appropriate. In addition, to properly compare the simulated global distribution patterns with Longhurst (1995, 1998), we would have to take into account the characteristics of each biogeographical province in the model parameterizations, which would most likely result in an overfitting.

Having “fixed” model parameters simulates so called “habitat tracking” of the forams through the seasons (but also on timescales of climate change or on glacial/interglacial cycles). This is a very important aspect to verify and would call for a section/paragraph by itself (see also Rebotim et al.,

2017). For instance, on p15 line 23-25 you write “Rebotim et al. (2017) identified an annual cycle in the habitat of *T. sacculifer* and *N. incompta* in the subtropical eastern North Atlantic. Both species appear to descend in the water column from winter to spring and reach their deepest habitat in spring to summer before ascending again to a shallower depth towards winter (Rebotim et al., 2017).”. How does this fit the “habitat tracking” picture? The authors could probably use observations on *G. ruber* and *T. sacculifer* for that as well. I may be wrong but I always thought that *G. ruber* lives closer to the surface than *T. sacculifer* (see also table 3 in Rebotim et al., 2017)? From laboratory experiments I know that *T. sacculifer* can handle living prey such as copepods much better than *G. ruber* while the latter seems to rely more on symbiont carbon, i.e. shows a more “autotrophic” lifestyle. Is it possible to see this in the data based on a more rigorous model-data comparison?

Our results reveal that outside their preferred habitat, where they naturally have to face a changing environment, the seasonal occurrence of both *G. ruber* (white) and *T. sacculifer* is limited to the warm surface layer, whereas in the low latitudes both species exhibit a weak seasonal cycle in their depth habitat (see Figure 5 of the manuscript). This indicates that both species adapt to changing environmental conditions by adjusting their habitat to local circumstances, which is consistent with the concept of habitat tracking. We will add this to section 4.1.3.

In addition, we are not able to derive dietary preferences from the model, as those are prescribed for the underlying model parameterizations. In the model parameterizations, we do not distinguish between the heterotrophic and/or autotrophic lifestyle of the species, just the parameters determining the preference for a food source differ slightly among the species (see Table 1 in Fraile et al., 2008). Additionally, those parameters introduced to account for the light sensitivity of *G. ruber* (white) and *T. sacculifer* with depth differ also among them (see Table 1 of the manuscript). So by prescribing light sensitivity and food preferences a similar depth ranking compared to observations already emerges from the model. Nevertheless, for a more rigorous model-data-comparison a sensitivity study regarding the species-specific food preferences should be performed. We will bear this in mind for a future model development.

The results of the point-by-point comparative analysis for each site and species as provided in the Supplement (Figures S3 and S4) are very helpful but also show that the model is far from perfect and sometimes there is a complete mismatch. I would have appreciated a sensitivity study to determine the hierarchy of factors for the different species controlling the shell export fluxes regional and seasonal (including e.g. bimodal patterns) as well as the vertical distribution (including ALD). This would probably be a paper by itself but in my view a very important one.

This is a real good and true point and we also think that such a sensitivity study would improve PLAFOM2.0. Therefore, we agree that such a study would be very important and should be

considered in the future. However, due to the high computational costs it is at present not feasible to perform this analysis. In addition, such a study would require using observational data with realistic year to year variability as forcing, but also for the model validation, which would in turn require a sensitivity study for each sediment trap/plankton tow by itself.

Based on the sensitivity analysis of Fraile et al. (2008) and also on our own results it seems that temperature has the strongest influence on the foraminiferal distribution regarding both the seasonal and vertical habitat. In particular, the distribution of each individual foraminiferal species seems to react most sensitively to changes in the individual temperature tolerance ranges (see Fraile et al., 2008). However, to further assess the sensitivity of the model to the chosen parameters especially in regard to the vertical distribution of the foraminiferal species a thorough sensitivity analysis should be performed in an independent study, which we will bear in mind for the future. Nevertheless, even after a further tuning based on such a sensitivity analysis the model will be far from perfect and discrepancies between the model data and the observations may always be present, as the caveats mentioned in section 4.2 will still be valid.

Presentation quality: good/fair

Although the scientific results and conclusions are presented in a relatively clear and well-structured way it is not easy to grasp why the model underestimates e.g. peak amplitude. What would happen if growth in the equation is increased or mortality is decreased? I sometimes wondered why the authors didn't play more with the model or used statistical techniques to quantify data-model mismatch (this is the reason for the "fair" mark).

As already mentioned, due to the long runtime of the model and, hence, the high computational costs we were not able to perform a thorough sensitivity analysis and just performed some very preliminary and short test runs to evaluate the model performance. In addition, since PLAFOM2.0 is based on PLAFOM, which has been tested and validated thoroughly (e.g., Fraile et al., 2008; Kretschmer et al., 2016), and since our aim was to demonstrate the applicability and the usability of PLAFOM to simulate the vertical distribution of individual foraminiferal species when combined with a complex 3D model configuration (such as CESM1.2(BGC)) without explicitly parameterizing the vertical dimension, we on purpose decided to not test what would happen if we change the given parameter setting. However, to better quantify the model-data-mismatch, we will calculate the Bray-Curtis index of similarity between the model and the core-top data. This way, we can provide some measure of confidence regarding the general model performance. In addition, we will also extend the discussion regarding the model-data-comparison by also considering potential mismatches due to the ocean model. Nevertheless, we will not entirely be able to unequivocally differentiate between the different model components (i.e., POP2, BEC or PLAFOM2.0) and their individual share likely leading to the model-data-mismatch.

The number and quality of figures/tables is good and the supplementary material is very appropriate. The English language is very good.

Thank you!

Minor corrections:

On page 2 line 18-20: “.....the lunar cycle and/or the structure of the water column), which influence the species-specific depth habitats (including their mean living depth and vertical migration) (e.g., Fairbanks and Wiebe, 1980; Fairbanks et al., 1982; Schiebel et al., 2001; Simstich et al., 2003; Field, 2004; Salmon et al., 2015; Rebotim et al., 2017), the only attempt to model the vertical habitat is by Lombard et al. (2011).”, and on page 17 line 20-23: “Several studies from different areas also showed that the main habitat depth of some species increases from the surface to deeper water layers during shell growth (Peeters and Brummer, 2002; Field, 2004; Iwasaki et al., 2017). Although I appreciate all the references that you list for ontogenetic migration and lunar cycle, there are only a few papers that specifically deal with very detailed population dynamics, lunar cyclicity and ontogenetic migration of planktonic forams that could/should be mentioned here (it was one of the first topics I studied when starting to work on planktonic foraminifera): Bijma et al., 1990; Bijma, 1991; Bijma and Hemleben, 1994; Bijma et al., 1994; Hemleben and Bijma, 1994; Schiebel et al., 1997. In my opinion, these references would fit best on p. 19 line 32-34: “.....and by explicitly parameterizing the ontogeny of each individual planktonic foraminifera, thus, by considering the changes in the species’ life cycles with depth, could considerably improve the model.”.

Thank you for pointing this out. We will add the mentioned references accordingly.

P. 9 line 27-30: “Although seasonal changes in the modeled foraminiferal peak fluxes with temperature are evident, all five species exhibit an almost constant peak amplitude (i.e., the maximum concentration divided by the annual mean) in their preferred habitat, which is, i.a., limited by temperature. Outside their preferred living conditions the peak amplitudes increase for most of the species considerably (Figure 3).”. It has not become clear to me what it means when “peak amplitude” is large or small in terms of real population dynamics (“bloom”?) and what it means in terms of model performance?

The maximum seasonal abundance or flux in itself is not a very useful parameter that can be compared among different regions/studies. Population dynamics can be much better (if not only) described in terms of deviations from the mean conditions, for instance, it would be impossible to recognize a bloom event in the absence of knowledge about the mean conditions. Moreover, any seasonal or vertical weighting of the proxy signal – and constraining this was the main motivation to develop the model – varies as a function of the relative departure from the mean. To be more

clear about this, we will rewrite this paragraph on page 9 in section 3.2 and will provide a better explanation.

P. 14 line 26-28: “This would explain why the highest modeled concentrations of *T. sacculifer* occur at shallower depths compared to *G. ruber* (white) (see Figures 4d-e and 5d-e).”. Strictly speaking this doesn’t explain it because this is what you put into the model in the first place (see my comments above).

Actually, this is a perfect example how the habitat emerges from the model. We only prescribe the light sensitivity and still obtain the right depth ranking. Throughout the model code, we specifically did not specify the depth ranking. We will, however, rewrite this sentence to avoid confusion:

“This is to some degree also indicated in our results, as the highest modeled concentrations of T. sacculifer occur at shallower depths compared to G. ruber (white) (see Figures 4d-e and 5d-e).”

P. 16 line 18: “*G. bulloides*, however, is found year-round close to the surface along the.....”. Write the genus name full at the beginning of a sentence.

Done and applied throughout the manuscript.

References:

- Bijma, J., Erez, J. and Hemleben, C. (1990) Lunar and semi-lunar reproductive cycles in some spinose planktonic foraminifers. *Journal of Foraminiferal Research* 20, 117-127.
- Bijma, J. (1991) Lunar pulses of carbonate output by spinose planktonic Foraminifera, in: Reid, P.C., Turley, C.M., Burkill, P.H. (Eds.), *Protozoa and Their Role in Marine Processes*. NATO ASI Series G: Ecological Sciences. Elsevier, Plymouth, pp. 353-354.
- Bijma, J. and Hemleben, C. (1994) Population dynamics of the planktic foraminifer *Globigerinoides sacculifer* (Brady) from the central red sea. *Deep-sea research part I: oceanographic research papers* 41, 485-510.
- Bijma, J., Hemleben, C. and Wellnitz, K. (1994) Lunar-influenced carbonate flux of the planktic foraminifer *Globigerinoides sacculifer* (Brady) from the central red sea. *Deep-sea research part I: oceanographic research papers* 41, 511-530.
- Hemleben, C. and Bijma, J. (1994) Foraminiferal population dynamics and stable carbon isotopes., in: Zahn, R., Pedersen, T.F., Kaminski, M., Labeyrie, L. (Eds.), *Carbon Cycling in the Glacial Ocean: Constraints on the Ocean’s Role in Global Change*. Elsevier, Fellhorst, pp. 145-166.
- Longhurst, A. (1995) Seasonal cycles of pelagic production and consumption. *Progress in Oceanography* 36, 77-167.

- Longhurst, A. (1998) *Ecological Geography of the Sea* ACADEMIC PRESS
- Schiebel, R., Bijma, J. and Hemleben, C. (1997) Population dynamics of the planktic foraminifer *Globigerina bulloides* from the North Atlantic. *Deep Sea Research* 44, 1701-1713.
- Fraile, I., M. Schulz, S. Mulitza, and M. Kucera (2008), Predicting the global distribution of planktonic foraminifera using a dynamic ecosystem model, *Biogeosciences*, 5, 891-911.
- Kretschmer, K., M. Kucera, and M. Schulz (2016), Modeling the distribution and seasonality of *Neogloboquadrina pachyderma* in the North Atlantic Ocean during Heinrich Stadial 1, *Paleoceanography*, 31, 986-1010.
- Lombard, F., L. Labeyrie, E. Michel, L. Bopp, E. Cortijo, S. Retailleau, H. Howa, and F. Jorissen (2011), Modelling planktic foraminifer growth and distribution using an ecophysiological multi-species approach, *Biogeosciences*, 8, 853-873.
- Moore, J. K., S. C. Doney, J. A. Kleypas, D. M. Glover, I. Y. Fung (2002a), An intermediate complexity marine ecosystem model for the global domain, *DSR II*, 49, 403-462.
- Schmidt, D. N., S. Renaud, J. Bollmann, R. Schiebel, and H. R. Thierstein (2004), Size distribution of Holocene planktic foraminifer assemblages: biogeography, ecology and adaptation, *Marine Micropaleontology*, 50, 319-338.