Ying et al present a development of the model ForamEcoGENIE (Grigoratou et al. 2021) to resolve different functional types of foraminifera at the global scale based on two key traits: spines and symbiosis. The authors perform a critical ensemble analysis and compare the model predictions against extensive datasets on the relative abundance of different foraminifera types and their carbon export fluxes. While I think this study addresses an important aspect of foraminifera ecology and has a great potential, I have some serious concerns regarding the current manuscript.

Major comments:

1) While I agree on the importance of spines and symbiosis to better understand the role of foraminifera in plankton dynamics and key nutrient fluxes in the ocean, I found that the introduction was not strongly motivated to address this gap in knowledge. I think the introduction should be reformulated, stating primary questions and providing more feedback on why spines and symbiosis are important, and importantly, clearly describing the different functional types of foraminifera (instead of focusing on trait-based models; see next comment). In addition, I think the main application of this work goes beyond predicting foraminifer flux. If one wants only to predict nutrient fluxes, statistical models might do a good (and even better) job than mechanistic models. Clearly stating why mechanistic models are important would be helpful. In addition, I think the results should be revisited to provide more mechanistic explanation for the observed patterns: how the tradeoffs implemented here help to explain the model predictions?

2) The authors put their model forward as a trait-based model. While I think that there is some room for interpretation for what a trait-based model is, I think the authors should be more careful (and specific) here. The models by Follows et al. (2007) and Ward and Follows (2016) and other models such as Chakraborty et al. (2017, Am Natur) are trait-based models in the sense that they represent plankton communities in a continuum based on key traits. They thus advocate that we can model diverse plankton communities
without describing several different functional types. Such diversity is commonly obtained by assuming size as a master trait and allowing other traits to vary as a function of size. The DARWIN model, for instance, would fall in between, because it can describe diverse size classes by incorporating allometric relationships but it does this within different functional types. I initially thought that this was the case of ForamEcoGENIE 2.0 and thus would recommend the authors to highlight the importance of the combination of both approaches: allometric relationships and functional type modeling. However, it seems that foraminifera were assumed to have a fixed size in the model and only phytoplankton and zooplankton were assumed to have different size classes. Moreover, neither spines nor symbiosis is implemented using a trait-based approach in the strict sense. Therefore, I don’t think the authors should rely too much of their motivation on trait-based modeling as their implementation of foraminifera ecology is mainly based on functional type modeling (and that is fine!).

3) The model description requires a through revision as in its current form it is very hard to understand how the plankton ecosystem is being simulated (what are the tracers (i.e. Carbon, nitrogen, phosphorus..), what are the different plankton groups, how they interact, what are the mechanisms modelled and how these are implemented). A schematic could really help the reader here. Even if the model has been published elsewhere, many changes were made here and so a full description of the equations should be given (either in the main text or in the supplement. Please be careful with providing units and descriptions for all model abbreviations that appear in the text. It is also especially hard to follow the functional type modeling for foraminifera and how the new tradeoffs related to spines and symbiosis were implemented.

4) I am not sure I agree with the way authors implement the spine and the symbiosis traits. The implementation of the “spine” trait is done so that spines incur an extra metabolic cost but offers a protection against grazing by decreasing the palatability of foraminifera and also allow them to be more efficient grazers. I am not entirely convinced by this approach for two reasons: i) how do authors calibrate the relative gain from lower palatability and the higher metabolic cost to build spines? A careful analysis to calibrate this tradeoff in a simple model would be very helpful and insightful. Was this provided previously in ForamEcoGENIE?; ii) spines widen the prey availability as explained in the manuscript. Since the authors are using a trait-based model with allometric relationships, why this was not mechanistically represented in the model? Instead, the authors simply decrease the half-saturation constant for grazing, which I don’t think translates into the mechanism they are saying they want to model. Meanwhile, symbiosis is represented using the mixotrophy model by Ward and Follows (2016) so that foraminifera can both photosynthesize and consume prey but at a cost of not being as efficient as their specialized competitors. The model by Ward and Follows was developed mainly to describe constitutive mixotrophs, i.e. mixotrophs that possess their own photosystems but can also feed on planktonic prey. Symbiosis is a much more complex process that involves the host to maintain an entire population of their prey inside their cells. Do the authors think their simplified representation of mixotrophy is appropriate to represent symbiosis? What are the limitations of this approach? I suggest to critically consider these points and address throughout.

5) Before implementing the four different foraminifera types into a global model, it would have been very helpful to just analyze model behavior in a simpler model, a 0D model that consider idealized environmental conditions. The reader could then better understand
how the tradeoffs related to spines and symbiosis play out.

6) The authors use an extensive dataset to compare their model predictions but I was surprised that the list of foraminifera sp used on this study and the respective functional type classification was not provided in the supplement.

7) A strong point of this study is the number of observations that the authors could access to run their ensemble and compare their model predictions against. Although they can represent very well the relative abundance of most of the functional types at a global scale, the absolute biomass is not well predicted by the model. It is also hard to visualize model predictions against observations in Figures 5 and 6 but especially in Fig. 7 (not possible to see the observations in panel a) and in the seasonal plots (Figs. 9 and 10). Also, many of the observations in Figs 9 and 10 do not seem to align with the model. I acknowledge the challenge of comparing model predictions against observations, and that observations are also subjected to error, but I think authors must provide a better way to visualize seasonal patterns and acknowledge model limitations. Perhaps authors can start by comparing model predictions for total nutrients and total foraminifera biomass first since these tend to be easier to simulate than the biomass of different types?

Minor comments

1) Acronyms are used throughout the manuscript and in most of the cases with no previous description, please be sure to provide their definitions.
2) Model description: start with physics, then describe the biogeochemical tracers and then the plankton components.
3) The tradeoffs description requires a reformulation and more detail.
4) Section 2.4.1 is about respiration, but mortality and palatability are also described in there, confusing.
5) Line 245: Foraminifer predation cost: this description is weak, this is not a good way to frame this, for example, what if everyone is doing this but is doing it wrong?
6) Explain up front what the rain ratio means.
7) Perhaps table 2 could be reformulated to make it clear the differences between the different types, for example, I was surprised that the non-spinose has a Pp that is the same as the spinose type.
8) Table S1 and S2 should give list of sp and their functional classification.
9) Lines 380: I don’t think comparing different models is that informative since they differ in their formulation and goal.
10) If foram biomass if overestimated by 8 times, how off do you think the model predicts POC fluxes for each group? Can we still find these estimates robust?
11) I recommend to provide model code in a repository such as github and perhaps zenodo.