Review #1

In this opinion paper, Martin and colleagues summarize the outcome of a survey conducted in preparation for a SOOS/OCEAN:ICE Workshop, aiming to highlight research priorities by the modeling community ahead of Antarctica InSync and the IPY.

While the results of the survey are certainly useful and worth publishing, the respondent group is very likely skewed towards physical ocean and ice scientists. The authors acknowledge this, but I think this could be highlighted even more for increased transparency (e.g., in the abstract). The fact that certain marine biogeochemical and atmospheric processes are not ranked high amongst suggested research priorities does probably reflect the group of respondents more than being a true reflection of what the whole polar scientist community thinks. While the ad hoc survey design and advertisement likely affect the robustness of the results, the authors report on the survey outcomes, as much as possible, in a balanced and objective way.

Overall, the paper is well written, and I have only minor comments, see below. As a result, I can recommend publication as an opinion paper once my comments below have been addressed.

Thank you very much for your kind and helpful feedback. In revising our manuscript we will carefully implement the suggested improvements.

Main comments:

My only main suggestion to the authors is to consider adding an appendix to the paper to include the exact questions and answer options given in the survey. I acknowledge that the authors make these available in the cited Zenodo repository, but I think not every reader would go there to check. While reading the paper, I found myself wondering about the exact wording of questions and answers multiple times, and I think making these more readily available would help the reader.

We will add the original questionnaire text either as appendix or supplementary material to make it more accessible.

Specific comments:

L. 21-29: In my opinion, the abstract would benefit from more explicitly mentioning that the survey results are most robust for the physical ocean-ice sciences. While I appreciate the way in which authors acknowledge throughout the text the shortcomings related to survey design and/or expertise of the respondents, I think it would be more honest if the abstract made it clear that biogeochemists, ecologists, and atmospheric scientists were likely underrepresented. In a way, the same applies for the title, which makes the opinion piece sound more comprehensive than it truly can be, given the group of respondents.

This is a valid point. The focus on physical ocean-ice science was unintended and, in hindsight, is simply owed to the biased selection of the questionnaire authors. We acknowledge this in the abstract now: "While this initiative was tailored mostly towards physical ocean and ice modelling, its outcome specifically supports …".

We refrain from changing the study title, however, because we still think the invitation for participation in the survey was shared widely enough for other communities to participate in the survey—the results do reflect such participation though in smaller numbers—and to keep an obvious connection to the survey title.

L. 34: I suggest starting a new sentence: "Ocean heat is a major…"

agreed

L. 56: "[…] the *physical* ocean modeling community" ?

Yes, it is true that physical oceanography was the focus. Good suggestion to state this clearly here.

L. 57ff: I didn't fully understand what to make of this sentence. Why "initially"? Can you elaborate in the text on why and how you changed your approach, e.g., in advertising the survey? Or did you decide to be more inclusive once you started receiving responses to the survey, realizing that modelers using a wider variety of tools answered the questions? Some clarification and/or rephrasing could help here.

Valid remark. We decided to rearrange the entire paragraph to more clearly state the intention and focus of the survey but also its shortcomings.

"The survey was designed with primarily the physical ocean modeling community in mind and thus pre-defined answers often highlight physical oceanographic processes (see Appendix). Although leaning towards such processes, coupled interactions with other components of the climate system, such as sea ice, ice shelves, and the atmosphere were considered as well. Contributions related to biogeochemical processes and ecosystem modeling were also encouraged though not covered comprehensively. It turns out that a more careful selection of the pre-defined answers would have been advantageous and likely beneficial for a broader coverage of these coupled processes. Relatively little use was made of the free text option "Other" by the respondents. Similarly, the survey started out with an emphasis on probing modelling groups using realistic regional Southern Ocean configurations and CMIP-class global climate and Earth system models to study the historical period and maybe 21st century projections. However, it evolved into covering a much broader range of spatial and temporal scales, and model complexities, for which the respondents did make use of the free text comment fields. For pragmatic reasons, we limited the geographical region and defined the Southern Ocean as the area south of approximately 50°S in the survey context."

L. 75: delete "modeling" (or something else is missing)

good catch, thank you

L. 100: This is where I first started to wonder what the responses were offered (see main comment above): only "oceanographer" or was the specification of sub-disciplines possible? Do you have any insights into how many of the respondents do not identify as physical oceanographers but biological or chemical ones?

No, unfortunately we do not have such granular information. Again, a shortcoming of the survey design and certainly part of the experience we take away from this exercise. We have added a few lines on this to the conclusion section hopefully being of guidance to similar future initiatives.

"There are also some experiences to take away from this exercise, in particular how the pool of respondents shapes the usefulness of the survey, how to ask targeted questions without being exclusive, how pre-defined multiple-choice answers simplify the analysis but reduce the variety of responses, and which metainformation really is instrumental for interpreting the responses. All of this is well known and demonstrates the importance to involve experts on questionnaire design rather than constructing an ad-hoc survey. Nevertheless, the feedback by the community to our survey has been very positive indicating that such surveys can be a valuable tool for future international program planning."

Review #2

This is a well-written and thoughtful opinion piece. The abstract presents a nice layout of goals. I applaud the authors for this effort, as it is a useful contribution. I have only a couple of suggestions.

It is noted that the focus is primarily on climate modeling. Several caveats are given in the interpretations about the nature of the respondents and how this may influence the findings. Still, I think the piece would benefit from brief thoughts on how to engage stakeholders and how the results may vary by audience. What aspects are expected to change if the focus were shifted to seasonal to subseasonal scales? Models are tools, and it would be interesting to get broad stakeholder engagement and find out what they want from SO models. The responses depend on what the user's goals are, and it was telling that 40% of the respondents said ice sheet/shelf-ocean interaction is the key SO science topic.

Thank you very much for your overall positive response to our survey and this manuscript. While we agree that a broad stakeholder engagement is desirable, we find it very difficult to speculate what the outcome would have been if the group of respondents would have been different. However, we have included a small paragraph on lessons learned in the conclusions section. This may not match exactly what you suggest here but provides some insight for future surveys.

The discussion distinguishing in situ vs gridded products is off-putting:

We agree with your interpretation and have rephrased several sentences in this section. Maybe our interpretation has been too simple and our wording misleading.

Line 217: 'less valued' is a poor word choice. You can't have gridded products without in situ data and the respondents know this.

Following criticism and another suggestion by Reviewer #1, this sentence and the next are changed to "Data from ship-borne instruments, moorings, and floats appear to be used less often directly [10-13%]. We acknowledge that observational data of all kinds, in particular including in-situ data, feed into the gridded products though."

Line 244: 'lesser use': same thing as line 217. They are using it in gridded format, but they are using it.

We rephrased this section considerably. Most relevant here, we now state: "We acknowledge that observational data of all kinds, in particular including in-situ data, feed into the gridded products though. Nevertheless, it is important to note that modelers tend to validate their simulations against these kind of 'observations', which in fact are advanced data products and rather not viewed as actual observations by the observing, sea-going science community. Further, any in-situ data that is not included in such gridded products is likely less used or even overlooked by modelers and thus does not contribute as much to the improvement of models. " This is to on the one hand acknowledge that in-situ data feed into gridded products and are thus used (indirectly) but on the other hand also to make the point that there are observational data out there that are underused.

Line 246: Maybe I am misunderstanding your inference here because I am struggling to link this to the survey questions, but I disagree with the interpretation. People know that gridded products are derived from in situ data. For example, it seems you are distinguishing between the Roemmich and Gilson mapped Argo product and Argo data, but not between the AVISO-mapped altimetry product and altimeters. I interpret the acknowledgement of the importance of in situ data as meaning they think the gridded products are important, not that they only want to have process study data. Respondents know they can't have gridded products without in situ data.

This is a very interesting interpretation which we now have included in this section as follows:"Figure 3a shows observations desired for bias mitigation with in-situ observations clearly dominating over remote sensing data. This preference can be interpreted as a need for better process understanding in

which in-situ data are considered actual "ground truth" and often provide higher resolution in space and time, which is useful in several aspects, such as model validation, identifying of processes resolved at a given grid spacing, improving model parameterizations, etc. Another interpretation is that modellers are well aware of in-situ measurements being crucial for better quality gridded products. And while such products are preferred in the actual validation process, the dire need for more ground truth data in a changing climate is acknowledged and its collection valued. "

Line 217: I would drop 'The latter' as all data go into many of the gridded products.

Changed to "We acknowledge that observational data of all kinds, in particular including in-situ data, feed into the gridded products though."

Line 220-227: Maybe this is more of a call for publishing open-source validation packages that incorporate data and help modelers use these data.

Great thinking. We have included this in the revised text.

More specific comments:

Line 134: is the problem unique to "coarse-resolution" models?

It is not exclusive to but predominantly present in models not resolving continental shelf processes. We changed the sentence to "... in many climate models, specifically those of coarse resolution, with consequences for …"

Figure 2: The figure implies carbon uptake is in the pie explicitly, but the caption implies it is counted as 'other'. Please reconcile this.

Thank you for spotting this. The list of "other" topics in the caption did not match the latest, actually shown figure. We corrected the caption.

Line 181: Please clarify this. Did you mean 'supporting model parameterization development'?

This could be development of parameterizations or hybrid models with AI elements or else. We prefer to keep this more general by just saying "model development".

Line 200: I would change the wording here. These physical processes are fundamental prerequisites to address the 'big questions'. You can't address ice-ocean interactions without understanding ocean heat transport.

Good suggestion. We rephrased: "The results suggest that oceanic processes themselves, such as dynamics from mesoscale eddies to large-scale circulation, tides, waves and mixing are not "big questions" by themselves anymore despite remaining issues and their important role in current "grand challenges" like ice-ocean interaction, warm water intrusion onto the continental shelf and biogeochemical modeling. "

Line 253: I don't see how caring about shelf processes (i.e. the main bias) is independent of calling for year-round observations. (Maybe I am getting lost comparing Figure 1 and Figure 3b, but regardless I don't see desiring year-round obs being independent of any model bias.)

This was misleading. We meant that biases are equally connected with the year-round observations—not independent thereof; connecting lines to left side in Figure 3b have similar thickness for year-round obs. We rephrased: "The strong desire for winter observations and year-round monitoring in the Southern Ocean [29%] is linked nearly equally to all major biases identified (Figure 3b)."

Line 260: This may not be justified as these (e.g. 'sea ice' or 'radiative processes') weren't given as choices in the 'most problematic ocean model bias'. I think sea ice - ocean interactions is a huge source of model uncertainty, but my response wouldn't reflect my concern based on the questions asked. In fact, I may pick 'mixing' because that is related to sea ice - ocean interaction. I would give a

caveat here that some of your impressions here may be due to the options given and the wording in the survey. (And this is a bit in conflict with Line 282-283, which seems to say sea ice is important.)

This is a valid point. We have addressed this by changing the two sentences following this statement: "This is likely owed to the limited choice of pre-defined answers we provided and the behavioral bias of the respondents preferring to tick one of those rather than entering individual answers. Nevertheless, we take this as an opportunity to point out the need for improved fundamental understanding and acknowledgement of coupled mechanisms and feedbacks within and beyond the focus group. "

Line 289: Why only early-career?

Valid remark. We removed "early-career".

Review #3

In this opinion piece, Martin et al. summarize the outcome from a questionnaire on Southern Ocean modelling biases and needs that was answered by by of group of about 100 modellers, about half of them identifying as oceanographers, the rest modelling some other sub-field of the Antarctic and Southern Ocean climate system.

Aim of the questionnaire was to get an idea of what are considered the most pressing questions in Southern Ocean research by modellers with the purpose to bring that perspective early on into the upcoming IPY and Antarctica InSync programs, and to better align observational and modelling efforts.

With a fairly high turnout of responses to the questionnaire and a range of specific modelling sub-communities involved, the opinion piece is certainly helpful for that purpose, although the authors do mention that a few modelling communities are under-represented in the answers.

In their analysis of the answers given, the authors discuss that a large fraction of the questionnaires mentions processes occurring on the shelves, such as the representation of shelf-ice cavities, and open ocean convection as two processes requiring more attention in model development and evaluation. Many of these are connected with improving parameterizations of processes in cavities, scale-aware mixing modelling, or of topographic overflows.

Concerning the scientific focus, the authors note that many of the respondents mentioned the role of the Southern Ocean for heat, freshwater, and carbon budgets, i.e. in some way or another the role of the Southern Ocean in the climate system. They mention that processes shaping the oceanic processes themselves, such as an understanding of the circulation are not mentioned as part of the big questions anymore.

Concerning the observational needs by the community, the authors note somewhat of a disconnect between the stated use of mostly gridded data sets like the World Ocean Atlas, and a desire for more direct in-situ observations. The authors interpret this as a need for using in-situ data for better process understanding, while gridded products are preferred for model evaluation, despite their larger distance from 'real' observations. This is an interesting point.

It is important to note that the manuscript not only presents the outcome of the questionnaire but in each of the three main points adds some perspective from the authors of the opinion piece themselves. In the parts on the scientific focus this consists in adding aspects that the authors find important (so do I) but that haven't been mentioned. In other parts (especially on data requirements) it is somewhat unclear what is output from the questionnaire and what is the author's opinion. So, while overall the analysis done in this opinion piece is quite helpful for the stated aims of aligning modelling and observational efforts, I feel that there should be a somewhat better separation in writing between the presentation of and the comments and opinions of the authors. This should be fairly easy to correct.

Thank you for this generally positive assessment regarding the usefulness of our study. In revising the manuscript we have addressed your remarks and concerns. For instance, we acknowledge now in the abstract already the participation bias towards the physical ice-ocean modelling community. Regarding the "big questions" we have rephrased the respective sentence to clarify that the processes less prominently mentioned still play a role in nowadays grand challenges. We also revised and extended the discussion of in-situ observations and gridded products, noting that the former are a crucial contribution to the gridded products. We also acknowledge now that this is known to the community and that emphasizing the need for more in-situ data could point at a wish for more process understanding but could also be interpreted as an appreciation of in-situ data forming the basis of gridded data products. This alternative interpretation is now included. Overall, we have tried to include wording ("we think", "we believe", "we suggest", etc.) that helps to separate survey results from author interpretations.