

The Cryosphere Discuss., referee comment RC2
<https://doi.org/10.5194/tc-2022-110-RC2>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on tc-2022-110

Anonymous Referee #2

Referee comment on "Impact of atmospheric forcing uncertainties on Arctic and Antarctic sea ice simulations in CMIP6 OMIP models" by Xia Lin et al., The Cryosphere Discuss., <https://doi.org/10.5194/tc-2022-110-RC2>, 2022

First, I apologize to the editor and authors for the delay in writing this review.

The study by Lin et al. investigates an important aspect of stand-alone ocean and sea ice models, which is the relation between the simulated fields and the atmospheric forcings. In simple terms, the paper tries to assess whether a new and arguably better forcing (JRA-55-do) leads to better simulation compared to an older forcing (CORE2). I find this a very interesting question, with important implications also for fully coupled model setups, and I would like to see more studies on these technical but rather important aspects. Congratulations to the authors for pursuing such an interesting problem.

That said, I found the authors' methodological approach unsatisfactory in providing a convincing answer to the problem. The analyses presented here are formally correct, but several central aspects have been almost completely neglected, as illustrated in my comments below. Otherwise, the paper is very well written and structured, but I have some suggestions for improving figures and tables, which sometimes are not adequate.

In summary, I have several major concerns that in my opinion should be addressed before this manuscript is considered for publication. I hope the authors find these helpful for improving their study.

MAJOR COMMENTS:

My biggest concern is that this manuscript does not consider the tuning of the systems analyzed. Let me start by acknowledging that finding out specific details about namelist parameters and other technical information is not trivial when dealing with CMIP-type simulations. Nevertheless, this aspect is important for a study of this kind and cannot be

neglected. Specifically, I am wondering whether the CORE2 and JRA-55-do simulations were run with the same model setup, or if the model was specifically tuned for a certain forcing. In my view, tuning is a fundamental step that, given the under-constrained nature and spatiotemporal variability of the sea ice model parameters, must be performed to accommodate a model configuration to a specific forcing. I would argue that the best setup for a study like this would be one where each simulation is optimized to obtain the best compatibility with a set of observations, given a specific atmospheric forcing (i.e., different parameters for CORE2 and JRA-55-do). If this is not the case and an identical model setup is adopted for all atmospheric forcings, we should at least be made aware of whether the model parameters have been tuned under CORE2 or JRA-55-do, if tuned at all for an OMIP setup. For example, if a model configuration has been tuned under JRA-55 and this same configuration is then run under CORE2, it is not surprising that one would outperform the other. Based on the information currently in the paper, we cannot say anything regarding the previous considerations, which, in my view, is a substantial limitation that should be addressed to pursue the research question from the right angle.

I think the manuscript lacks an in-depth description of the model components used in the three systems considered, which could be helpful for a more detailed interpretation of the results. The only information available in Lin et. al. 2021 is that two systems employ different versions of CICE as sea ice model components (CMCC-CM2 and NorESM2). By digging in the MRI-ESM2 model description paper we discover the sea ice component of MRI. COM4.4 is also based on CICE. Given the modularity of CICE, the model version tells us nothing about the physical configuration used by each modeling center. A better description of the model configurations would allow linking differences in the model response to the reanalyses, to differences in the specific physics used. For example, I suspect that different radiative schemes lead to very different sea ice concentration conditions in summer.

When designing the study, the authors limit the number of models considered based on the availability of diagnostic variables: the dynamic and thermodynamic sea ice concentration tendencies. The main result of this choice is to limit the analysis to systems based on different flavors of one sea ice model (CICE). I think including more models might be interesting and more in line with the scope of CMIP6 and OMIP. The tendencies analysis, which is not central to this study, can be limited to the system with the appropriate diagnostic variables.

The paper lacks a direct comparison of the reanalysis fields. Also, the CORE2 and JRA55-do forcings have both been bias-corrected in the Arctic to avoid unrealistic model behaviors (e.g. too little sea ice in summer). The correction follows the work of Large and Yeager (2009). How does this bias correction impact your results? To what extent are the forcing converging?

I believe more observations should be included in the analysis to allow a correct interpretation of the results. We know that in the Arctic different observational datasets are not always in agreement with each other and that identifying the best product is not obvious. I am surprised that this has not been done given that SITool includes multiple observational datasets for sea ice concentration, thickness, and drift. Illustrating the

differences between reanalysis in relation to different observational products would certainly be an interesting addition to the study.

FIGURES AND TABLES:

Tables 1 and 2: I find this table very hard to read and overcrowded. It might be a personal preference, but I could read these data more easily when converted into plots and reorganized. In particular, the use of parenthesis and bold and italic font is confusing. I suggest fully rethinking this, and possibly working with colors/symbols instead of font styles.

Figure 1 and beyond: Showing the results of just one model is, in my opinion, limiting. I understand the authors are concerned about having too many display items in the manuscript but storing relevant material in the appendix is not necessarily a solution. If the space is a concern, why not report only the maps with the difference between the two forcings in the main text, while moving the bias to the appendix? Also, the addition of the observed sea ice concentration in Fig 1 is not very insightful, or at least not a priority for the panel. The same is true in the following figures. Again, this is a suggestion and I realize it depends on personal preferences.

Figures 6 and 7: The color choice of the bar plots is very unhappy. Please consider differentiating the colors of the interior vs. exterior bars.