

**Moreno-Chamarro et al.: Impact of the ice thickness distribution discretization on the sea ice concentration variability in the NEMO3.6-LIM3 global ocean–sea ice model**

The authors investigate ice thickness distribution (ITD) categories in NEMO-LIM and how they impact sea ice concentration variability. They use k-means clustering as a technique in tandem with three observational based SIC datasets. The authors do not find an optimal configuration as results in the Arctic and Antarctic have opposite responses to ITD changes, so no clear benefit to NEMO-LIM is determined from changing ITD.

Overall, I believe this will be suitable for publication with a few major/moderate changes. I felt that the scientific significance and quality were good to fair, but could be improved with some expansion in the text. The Scientific Reproducibility is also fair, which again could be improved with further clarification in the text. The Presentation quality was excellent.

**Specific Comments:**

- One of the biggest concerns I have about this paper is that it doesn't generalize to modeling in general beyond NEMO-LIM to provide insight about modeling in general. I realize that this is for the NEMO special issue, however, it currently feels a bit like a sensitivity experiment to determine optimal model configuration but not otherwise generally of interest to the community of sea ice modelers who may be setting up their own models using LIM or other sea ice models.

This begins in the introduction where there should be a brief discussion of previous work about why 5 ITD categories have been chosen in the past due to volume studies (Lipscomb 2001, Remapping the thickness distribution in sea ice models, doi: 10.1029/2000JC000518; Bitz et al. 2001, Simulating the ice-thickness distribution in a coupled climate model, doi: 10.1029/1999JC000113). In fact, in Bitz 2001 one of the conclusions is "...the concentration of open water and thin ice, which is relatively insensitive to the number of categories beyond M=5," which is directly relevant for this paper. Why weren't these cited? If anything, studies using CICE that agree with these results should strengthen your results because they become more robust across models.

In the discussion and conclusions section you should add more information about how these results might be directly relevant in coupled models. This is brought up briefly but could be fleshed out and suggestions for how to test this would be useful. Additionally, you mention that parameterizations and parameter values are tuned for 5 categories (line 359). Can you specify which of these might be directly affected or changed? Are similar parameterizations present in other sea ice models? How can this be generalized for the community?

- The methods need clarification, particularly for replicability purposes. In particular, I found these sections to need to be expanded. 1) At line 145/Figure 2 the Arctic "winter" cluster was defined but didn't include April. What threshold values for these groups were used? Are you results sensitive to including different months? These things should be tested. 2) The % values in Fig.5/6 refer to occurrence, can you translate these values to number of months or something to better indicate what this means? 3) Section 3.3.1 – are these correlation differences statistically significant from one another? Can you clarify what you mean by these are significant? 4) Line 263 – how was the polynomial determined? Can you provide information about this?

- If there is not a lot of information gleaned from the de-trended Arctic analysis, then why is it presented? Can this be condensed somehow since the variability analysis primarily shows the forced trend without being de-trended?
- I think that if possible you should consider including Supplemental Figures 4 and 7 as regular figures since they are referred to in detail.

**Technical corrections:**

- Line 81: misspelled “concentration”
- Lines 269-275: It looks to me like patterns 2 and 3 are both dipoles but opposite patterns. Can you clarify where the quadrupole is?
- Lines 296-298: sentence is confusing. “...suggesting that this configuration poorly captures the forced variability but does capture interannual variability as well as any other configuration.”?
- The stippling markers are used to indicate significance in Fig. 11 but insignificance in Fig.5. It would be nice if they were used consistently.
- The first two paragraphs of the discussion were clear and concise. The next three are a bit confusing and all over the place. I’d suggest you rearrange in the following order: 1. One category has worst results necessitating multi-category sea ice models like LIM3 or CICE; 2. The standard configuration is 5 ITD levels; 3. Adding more thin categories decreases agreement; 4. Having 30+ categories can improve some but is significantly more expensive at double the cost, which is clearly significant for coupled models.