

Responses to Reviewer #3 comments

The authors thank Reviewer #3 for their comments.

Responses to the specific comments are below, with reviewer comments in brown.

Specific Comments:

Line 105: What is horizontal resolution of the atmospheric forcing? Please include this in the text.
Added (17 km).

Page 14 Fig 4: add 87.5 N gridline on Fig. 4a and 4b to make the area more discernable. Done

Line 357-358: "Improvement" in Atlantic sector is (5%) is small and difficult to see in plot. Can you make a difference plot between 10c and 10d to show this? The changes are also seen in the mean difference plots (10a,b) not just the RMSD plots (10c,d). In order to avoid adding two extra plots, and because it seems the previous wording in the paper was too vague, instead on line 357 we have added "specifically in the area immediately north of Svalbard and Franz Josef Land" after "in the Atlantic sector", to clarify where to look.

To augment the NASA IceBridge, BGEP ULS moorings and Air-EM observations, the paper would be more comprehensive if Dartmouth/CRREL IMB were included. Specifically 2015D and 2015F (<http://http://imb-crrel-dartmouth.org/archived-data/>). These buoys are outside the Beaufort Sea region, where little improvement was found with the assimilation of CryoSat-2 along-track freeboard data. How do these observations compare with the FOAM ice and snow depths? I would like this analysis added to the paper.

An assessment of model and CryoSat-2 observations against the Dartmouth-CRREL Ice Mass Balance buoy dataset was conducted in response to this comment. However, the results were suspicious. Unfortunately, there are concerns with the both the spatial representativeness and the quality of the IMB SIT dataset, making it unsuitable for use as a validation reference in this paper.

Firstly, producing meaningful matchups of the mean gridbox ice thickness of the model with the IMB dataset is difficult. The model is attempting to simulate the average behaviour of the ice field as an ITD (ice thickness distribution), whereas the IMB observations are point measurements moving with discrete ice floes. Sampling bias in these point measurements has been previously demonstrated in the literature, where IMB observations have been unable to capture the full range of ice thicknesses present (e.g. West et al., 2020). Consequently, the IMB measurements cannot be assumed to be representative of the wider region, and the average behaviour of the ice as represented by the model.

Additionally, it has also been demonstrated that there are issues with the quality of IMB snow depth measurements (e.g. Blanchard-Wrigglesworth et al., 2018). A poor snow depth observation would result in an unreliable SIT observation. If necessary, we are also able to provide supplementary

material showing further assessment of a large number of IMB snow depths against modelled and observed snow depth climatologies, which draws similar conclusions. The snow depths for the specific buoys 2015D and 2015F as suggested by the reviewer, and also 2015G and 2015J, were examined in response to the reviewer's comment and these also showed similar issues. Unfortunately, it must be concluded that the IMB SIT observations are not sufficiently reliable to be used as a reference for validation in this paper.

References:

West, A., Collins, M., and Blockley, E. (2020). Using Arctic ice mass balance buoys for evaluation of modelled ice energy fluxes, *Geosci. Model Dev.*, 13, 4845–4868, <https://doi.org/10.5194/gmd-13-4845-2020>

Blanchard-Wrigglesworth, E., Webster, M. A., Farrell, S. L., and Bitz, C. M. (2018). Reconstruction of snow on Arctic Sea Ice, *Journal of Geophysical Research: Oceans*, 123, 3588–3602, <https://doi.org/10.1002/2017JC013364>

The paper does not show any comparison to IABP ice drift data. Although not a fully coupled air-ocean-ice modeling system, have you examined ice drift and if you have, did you see any improvement in ice drift prediction between the Control and CryoSat-2 assimilative hindcast? If you have not, what is the level of effort to incorporate an ice drift analysis to the existing manuscript? A quick comparison of FOAM ice drift fields shows that there is some difference between the SIT assimilation run and the control. However, this is smaller than the difference between either model run compared with NSIDC v4 sea ice motion vectors (which also include IABP buoys). We have not compared FOAM to in situ ice drift observations directly. Since ice drift has not been extensively validated in FOAM, it would be a substantial amount of work to include a proper assessment in what is already quite a long paper. However, it is a useful suggestion and will hopefully be included in future work plans.