Response to referee comments - Referee 2

Referee comments are shown in black, our response in blue, and changes to the manuscript in red.

The paper describes assimilation of monthly mean sea ice thickness, monthly mean sea ice thickness distribution from CryoSat-2 and sea ice concentration from passive microwave observations in CICE Arctic sea ice model. The best result was achieved when all three data sources are assimilated. For the first time, a sub-grid scale sea ice thickness distribution was assimilated. This led to a significant improvement in estimation of the thickness of the thickest ice category. The paper can be published once the following comments are addressed.

Response: We would like to thank the reviewer for their time spent reviewing the paper and their helpful feedback.

General Comment:

My major concern is that the only data source used for model verification was CryoSat-2 ice thicknesses that were not assimilated. In general, the fact that the modelled ice thickness becomes closer to CryoSat-2 observations indicate that the assimilation of CryoSat-2 ice thicknesses has been performed correctly. However, in order to further assess if the data assimilation brings the model closer to the reality, it is important to do a comparison against additional sea ice thickness observations. I suggest that the authors conduct verification of the model results against the available independent sea ice thickness observations including (1) NASA's IceBridge, (2) Beaufort Gyre Exploration Project upper-looking sonars (ULS), and (3) Airborne EM observations ("EM-bird"). I think that such additional verification analysis will substantially increase the quality of the paper.

Response: The difficulty of using these data sets is that they represent different space (and time) scales from the model, with its 40km by 40km grid cell size. In theory such comparison can be made useful if we add so-called representation errors to the measurement errors. These representation errors account for the scale mismatch and are typically much larger than the measurement errors. The difficulty, however, is in determining these representation errors. As discussed in Van Leeuwen (2015) there are essentially two ways to solve this problem. The first relies on having many observations evenly distributed in a grid box, and the second relies on running models at the spatial resolution of the observations. The first is not available, and the second is not feasible in our study. Alternatively, statistical methods based on data assimilation can be used, but those rely heavily on unbiasedness of the model, which is questionable. However, we agree that a comparison of the thickness with an independent data set for data assimilation can be important so we have conducted a comparison with Operation IceBridge, please see the figures of scatter plots of CICE-PDAF runs and CryoSat-2 against IceBridge and table showing RMSE comparisons below. Operation IceBridge is much more suited for comparisons of ice thickness for evaluation than the other two datasets proposed, because it has a much larger amount of data available concurrently with the time period of our experiment. The EM-Bird campaigns took place for some days in March and April each year between 2012 and 2015, which is around the same time that Operation IceBridge campaigns take place, but with less data available than IceBridge for validation. The EM-Bird data also provides ice + snow thickness together, not seperately, which introduces additional uncertainty for the validation of ice thickness for this study (though the data is not necessarily less reliable than Operation IceBridge, and very reliable over level ice (Haas et al., 2009)). The BGEP ULS moorings provide ice draft which must be converted to ice thickness, either using a simple multiplication as in Rothrock et al (2003), or more accurately, by using snow depth – because the snow depth is very important in converting ice draft to ice thickness especially outside of summer. Uncertainty in snow density will also affect the conversion of draft to

thickness and comparisons. Of course the way CryoSat-2 uses the climatology to estimate snow depth is far from ideal. In Fiedler et al (2022), a paper which looked at assimilation of along-track sea ice thicknesses, found that the converted ice thicknesses from BGEP ULS moorings did not necessarily compare well to the CryoSat-2 observations of ice thickness, which meant that the assimilation of the CryoSat-2 ice thickness observations worsened the validation with the BGEP ULS data in comparison to the control. They also bring up the uncertainty in the conversion from sea ice draft to snow caused by ignoring the snow depth issue. From the analysis with Operation IceBridge, we can also conclude that comparing large area averages with point averages is erroneous and not necessarily useful for the other data sets where dealing with the snow depth could introduce additional errors. In terms of the comparison with Operation IceBridge, the assimilation of thickness (assim_conc_hi) and thickness distribution (assim_conc_hi_4hd) provide some improvement over the control towards Operation IceBridge, with an RMSE that is also similar to that between the CryoSat-2 data and the Operation IceBridge data (where we reiterate that it is unclear what the means exactly as the uncertainties in the IceBridge data related to the scale mismatch are unknown). Assimilating thickness distribution alongside the mean thickness did not improve the thickness RMSE much at all (only 0.01 m) toward the IceBridge thickness data. Assimilating the Bootstrap concentration alone (assim_conc) in our model significantly worsened the estimates of ice thickness against Operation IceBridge, as we already saw when we compared it against the CryoSat-2 estimates.

Table 1 RMSE of sea ice thickness (m) for the four CICE-PDAF experiments and the CryoSat-2 observations against Operation IceBridge. The control experiment features no assimilation, assim_conc features assimilation of ice concentration, assim_conc_hi assimilation of concentration and thickness, and assim_conc_hi_4hd features assimilation of concentration, thickness and thickness distribution. This compares daily sea ice thickness over grid cells and on days where Operation IceBridge data is available (For each year, around 15 days in March and April).

Experiment	**Ice Thickness RMSE (m)**
control	0.64
assim_conc	0.92
assim_conc_hi	0.58
assim_conc_4hd	0.57
CryoSat-2	0.53

Suggested change in manuscript: We will add a section explaining Operation IceBridge data and an evaluation against the CICE-PDAF experiments we have conducted using the figures and tables included in this response, and include a discussion of this comparison within the discussion section - including the evaluation of ice thickness in other studies.

Specific Minor Comments:

I noticed some minor inaccuracies throughout the paper:

Line 16 and throughout the paper. “Canadian Archipelago” à “Canadian Arctic Archipelago”. Abbreviation “CAA” can be also used.

Suggested change in manuscript: Thank you, changed all references to CAA
Fig. 1 In Y-axis title please add “)”. X-axis title should be “Month”?
Suggested change in manuscript: Thank you, adjusted plot

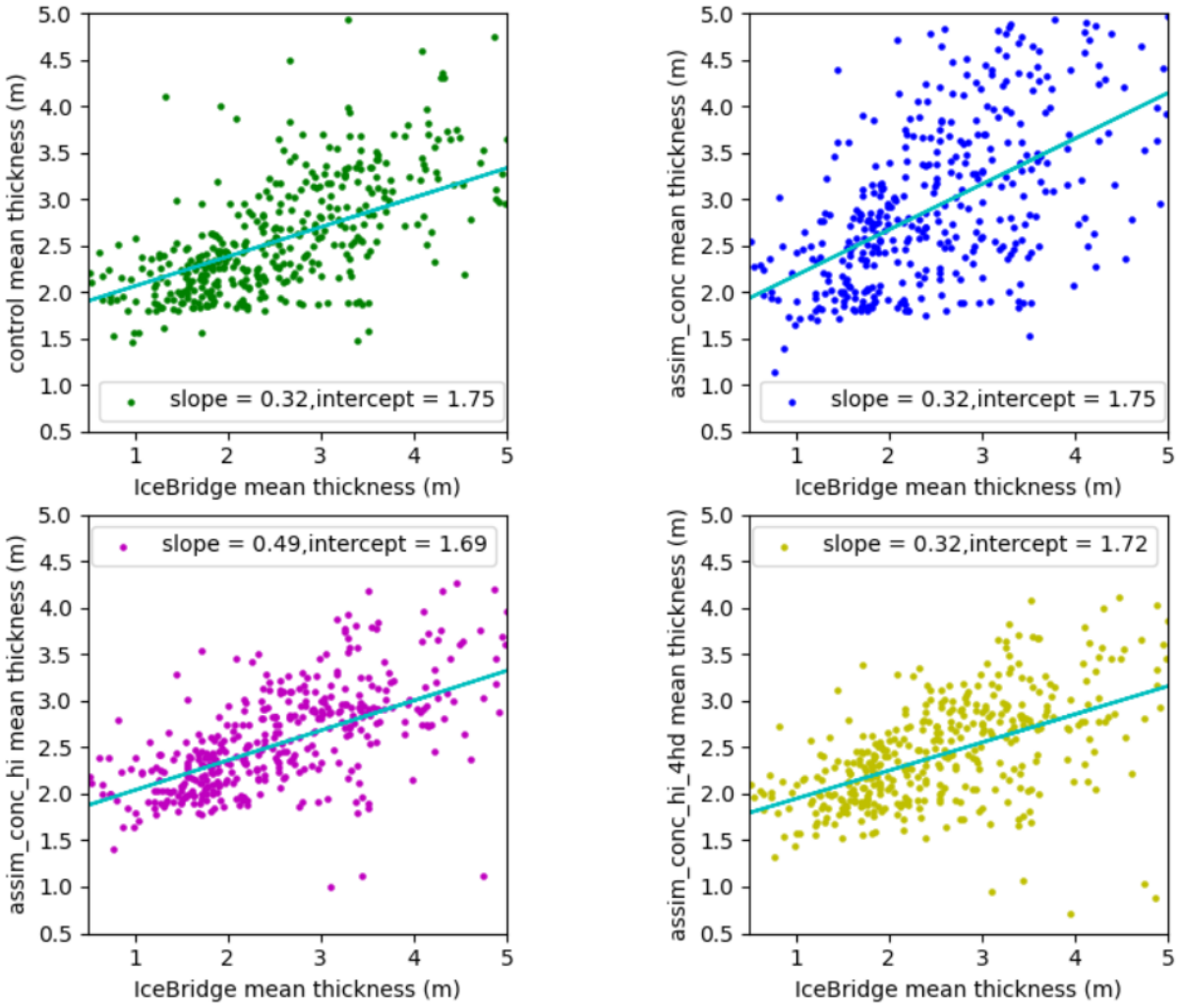

Figure 1: Scatter plots of mean sea ice thickness estimates in CICE-PDAF runs against Operation IceBridge for March and April 2012-2015

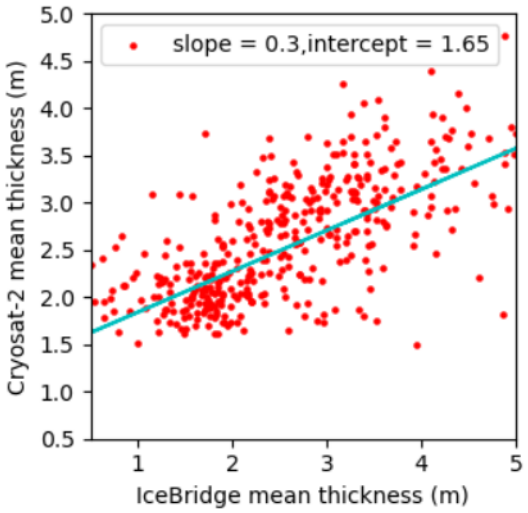

Figure 2: Scatter plots of mean sea ice thickness estimates in CryoSat-2 against Operation IceBridge for March and April 2012-2015

Line 58. Term "Optimal Interpolation" has been already introduced above.
Suggested change in manuscript: Yes, changed to introduce OI term
Line 170. Please define "(?)".
Suggested change in manuscript: reference to previous section was missing, corrected.

Line 190-191. Words "using", "use", "using" are part of the same sentence. Please rephrase.
Suggested change in manuscript: Rephrased sentence, thank you.

Line 214. Word "key" could be removed.
Suggested change in manuscript: Removed, thank you.

Line 217. "MYI", "FYI" were not defined.
Suggested change in manuscript: Defined both terms in first use of manuscript.

Line 218. "radiances" -¿ "brightness temperatures".
Suggested change in manuscript: Adjusted, thank you.

Line 226. "team" -¿ "Team".
Suggested change in manuscript: Adjusted, thank you.

Line 279. Please rephrase "Ice thinner than 0.5 is not assimilated..." $->$ "Ice thicknesses lower than 0.5 are not assimilated..."
Suggested change in manuscript: Rephrased, thank you.

Lines 399-401. "because" is used twice in the sentence.
Suggested change in manuscript: Rephrased, thank you.

Legend font on Fig. 5 is very small.
Suggested change in manuscript: Plot modified to increase legend.

Table 3. Add dimension to RMSE.
Suggested change in manuscript: Added, thank you.

Figure 8, y-axis title. Add dimension.
Suggested change in manuscript: Added, thank you.

Line 513. Should it be Fig. 13 instead of Fig. 12?
Suggested change in manuscript: Yes, corrected.

References

Fiedler, E.K., Martin, M.J., Blockley, E., Mignac, D., Fournier, N., Ridout, A., Shepherd, A. and Tilling, R., 2022. Assimilation of sea ice thickness derived from CryoSat-2 along-track freeboard measurements into the Met Office's Forecast Ocean Assimilation Model (FOAM). The Cryosphere, 16(1), pp.61-85.

Haas, C., Lobach, J., Hendricks, S., Rabenstein, L. and Pfaffling, A., 2009. Helicopter-borne measurements of sea ice thickness, using a small and lightweight, digital EM system. Journal of Applied Geophysics, 67(3), pp.234-241.

Rothrock, D.A., Zhang, J. and Yu, Y., 2003. The arctic ice thickness anomaly of the 1990s: A consistent view from observations and models. Journal of Geophysical Research: Oceans, 108(C3).

Van Leeuwen, P.J. (2015) Representation errors in Data Assimilation, Q. J. R. Meteorol. Soc., 2014, DOI: 10.1002/qj.2464