Responses to Reviewer #4 comments

The authors thank Reviewer #4 for their comments. Responses to the comments are given below, with reviewer comments shown in red.

General Comments

If I understand correctly, the modelled SIT using assimilation of CS2, and the control run (no assimilation) are compared with CS2 SIT observations that are also used for assimilation? But then it seems somehow clear that the modelled SIT using assimilation of CS2 performs better. Perhaps it would be better to evaluate with a more independent product, e.g. CS2SMOS (at least this is a different product).

Assessment of observation-minus-background statistics is standard in data assimilation, and demonstrates that the assimilation is working as expected. It is also a form of validation, since the model is brought into line with the CS2 observations and these have been independently verified (e.g. Tilling et al., 2018). It is agreed that comparison to independent data is also important, which is why the SIT analyses and forecasts are subsequently compared to various independent in situ datasets.

Note that the FOAM modelled SIT will be assessed against the CS2SMOS product in a follow-on paper on the assimilation of CS2 and SMOS together, and therefore it is not necessary to include it in this paper.

I also assume that the modelled snow depth is a very important component here and potentially adds significant uncertainty. A detailed evaluation of the modelled snow depth within FOAM would be important as well for evaluation here, but I understand if this is not within the scope of this paper, and the authors also suggest carrying out such a study in future.

It is agreed that the modelled snow depth is a very important component of the conversion of freeboard observations to SIT. Further discussion of this has been added to the paper. Future plans also include the assimilation of satellite snow depth observations in FOAM. A more comprehensive evaluation of the modelled snow depth will take place as part of that work, and is thus beyond the scope of this paper.

From line 147 now reads:

Currently, CPOM makes use of a modified snow depth climatology, based on Warren et al. (1999) and halved over first-year ice, for processing CryoSat-2 sea ice freeboard retrievals and conversion to SIT (Tilling et al., 2015). This approach is also used by other centres processing CryoSat-2 freeboard observations: Alfred Wegener Institute (AWI; Ricker et al., 2014) and NASA (Kwok and Cunningham, 2015). Instead, here the FOAM modelled snow depth is used. Modelled snow depth has a greater spatial and temporal variability than can be obtained from a climatology, as demonstrated by Mallett et al. (2021) and illustrated on Fig. 1. Using this method also maintains consistency between SIT and snow depth within the FOAM model. A preliminary validation indicates that the FOAM snow depth is somewhat thinner than the modified climatology of Warren et al. (1999), as shown on Fig. 1, particularly over multi-year ice. Tuning experiments demonstrate that simply increasing the snow depth in the model does not result in better evaluation of the SIT analysis against independent observations, owing to feedbacks in the model and between the SIT assimilation and the snow depth itself.

Snow depth uncertainty is a large source of error in radar altimetry sea ice measurements, both in the retrievals of freeboard and the subsequent conversion to SIT (e.g. Giles et al., 2007; Ricker et al., 2015). Due to the linear relationship between SIT and snow depth (Eqns. 1 and 2), an underestimation of the snow depth would lead to an underestimate in the SIT. Large uncertainties in the snow depth may apply whether it has been modelled or taken from climatology. Additional uncertainty is also introduced in Eqn. 1 through lack of knowledge of the snow and sea ice densities which, although constants in the CICE model used here, are spatially and temporally varying in reality (e.g. Alexandrov et al., 2010; Kern et al., 2015). Uncertainties due to variations in water density can be neglected (Ricker et al., 2014; Kurtz et al., 2014). In order to quantify and reduce the uncertainty in the FOAM modelled snow depth, future plans will include the assimilation of satellite snow depth observations.

Reference: Mallett et al., (2021): The Cryosphere, 15, 2429–2450. https://doi.org/10.5194/tc-15-2429-2021

Other references as already cited in the paper.

In fact, looking at the validation results using independent SIT measurements, the improvement from the assimilation is not so obvious anymore. Moreover, looking at Fig 7 (at the very beginning of each autumn), it seems that performance significantly decreases through the summer for both the assimilation and control run.

Once the daily SIT assimilation is stopped at the start of each summer, the model performance would not be expected to continue in the same way. Assimilation can't "fix" model biases, which return when the assimilation ceases. However, the results show that even after several months the model does not drift entirely back to its pre-assimilation state. This is not really a major part of the evaluation, but is included as an interesting point, demonstrating that there is some memory of the assimilation retained over the summer when SIT observations are not available for assimilation.

I think this discrepancy in the evaluation needs to be discussed in more detail… As above, a drift back to the model baseline in the absence of observations would be expected of any assimilation system.

…so it is clearer to the reader how good this assimilation really works and if it is a benefit using along track data instead of gridded products for assimilation. A comparison using gridded products for assimilation would be beneficial as well. Is it really an advantage to use along-track measurements? This should be discussed in more detail.

At this stage it is not of clear benefit to assimilate the along-track product over a gridded product, since the observation uncertainty for the along-track observations needs more work, as discussed in the paper. However, the main advantages of along-track data over gridded data are the reduction in spatially-correlated uncertainties and the determination of more accurate uncertainty estimates than is possible for data with more processing applied (this is already mentioned in the paper). As also discussed, developments towards coupled ocean-ice-atmosphere NWP at the Met Office require short assimilation time windows of the order of 6 hours, which means that along-track rather than temporally-averaged observations are required.

The paper demonstrates that the assimilation of along-track SIT data is feasible, i.e. that using the data in a different way than has been done previously is possible. This is the first step towards getting the things we need in place, and provides evidence for data producers that users are ready and waiting for along-track SIT data produced in near-real-time, and emphasises the importance of well-defined observation uncertainty estimates produced as part of the processing chain. When this is all available, the assimilation of along-track SIT data will lead to demonstrable improvements over the gridded SIT products for our systems.

Specific Comments:

L 114: It would be interesting to SIT maps also for the summer period. A figure showing SIT maps with and without assimilation throughout one season would be beneficial.

Figure 6 shows SIT difference plots for the SIT assimilation experiment minus the control, for the start and end of the 2016 summer period. Although a full set of SIT maps for the experiment and control runs would give more detail, the conclusions drawn would be the same as already included in the text. Additionally, there are already a large number of figures in the paper. Therefore, no changes have been made.

L 152-153: probably also mention dual altimetry, i.e., Ku/Ka band, e.g., Lawrence et al. Have taken this sentence out as future plans are now to assimilate snow depth observations in FOAM (which has been added to the paper here).

L 170: Why median and not mean? Using the median prevents outliers from influencing the result (have added this to paper).

L206: These are not helicopter measurements. PAM-ARCMIP measurements were carried out with a fixed wing aircraft. Replaced "helicopter" with "aircraft".

L 339-347 and Figure 7: But there is a quite high mean difference/RMS at the beginning of autumn when CS2 data for assimilation become available again. But that means that the performance decreases significantly within summer months also for the model using assimilation, see also my general comments? Yes. This is expected, see responses to comments above. Once the assimilation stops, the model begins to drift back to its baseline state.

Figures: I recommend having all relevant information included in the figure. Sometimes this information must be searched for in the caption, which slows down the flow, e.g., Figure 6: colorbar label, Figure 7, 11: dashed lines. I think it is easy to provide that information in the legend.

Added information to legend Figs. 7, 11. However, unclear what to put in a colorbar label for Figure 6? SIT difference (m) is shown in title, and is in the same format as the other figures. We will change though if it doesn't follow journal guidelines.