Review of tc-2021-185

The work presents a data fusion example, to combine CryoSat-2 altimeter data with Sentinel-1 SAR image data, for extending sea ice thickness estimates in the Barents and Kara seas. Overall it is very well written, logical, and well supported, and of interest to the Cryosphere community.

I have listed many minor comments and inconsistencies that I think would make the paper easier to read, and one major comment regarding the main scientific assumption of the work. However, I only ask for more explanation about the assumptions and limitations going into this work. I agree with the authors that this will often produce valuable results.
I will recommend major revisions, mostly because of the number of inconsistencies listed.

Major comment:

The main scientific mechanism is assuming a relation between radar brightness and sea ice thickness. That is, inferring that similar sea ice surface properties (which a SAR is sensitive too) corresponds to the same thickness. Although this may often be true, it is not necessarily true, unambiguously, and you probably cannot tell when from the SAR alone.
I think the sections explaining this may need some expansion and clarification. Consider:

- Page 6, Line 171: This sentence implies a direct relationship between level ice thickness and brightness, and that it is consistent with ice from different growth regions (i.e. land-fast ice being connected to level drifting ice).
- I thought is was well established that the radar brightness saturates rapidly for level ice greater than about 30-50cm. This means that SAR does not really "see" level sea ice thickness over a certain thickness.
- Level ice brightness is also produced by different (wavelength scale) surface roughness properties, which may be different for different growth areas, different snow packs, salinity, different moisture histories and many things.
- Brightness gets much more complicated for deformed ice, with both wavelength-scale and large-scale roughness affecting backscatter. And rougher, does not necessarily mean thicker, although it often means brighter.
- Did the fast-ice equivalent thicknesses data include deformation (e.g. ridges and rubble), or was it only level fast-ice regions?
- Is this land-fast ice regression only used in the AARI ice charts, or is this also used in your work?
- I understand that the overlapping CryoSat-2 observations are used to estimate thickness for some of the 15 segments from the mean-shift algorithm.
- You don't show a figure of these thickness versus segment properties, for the overlapping categories, to assess accuracy. Perhaps it is difficult, since you have 15 features. But some sort of regression or least squares fit line, in some dimensions may be meaningful to portray the fitting accuracy.
- Your T-distance weighted combination to get SIT for non-represented categories may work for interpolation, but how does it extrapolate? Or, is it effectively a multivariate regression?

I am not saying that this cannot be done, as it may be better than nothing, but please make some small comment about this assumption or limitations. This is your main mechanism for extrapolating the SIT measurements away from your CryoSat-2 tracks.

Minor Comments:

1. The second sentence of the Abstract might be better at the end of the Abstract, maybe the second last sentence, just before "Our results are directly applicable...". Then the Abstract follows the more logical layout of the paper, and the result comes after the lines about what you actually did and the comparisons. The Abstract is not so long that this delay should cause any loss of emphasis.

2. Section 2, page 3, line 81: It may be interesting to confirm that the "two years" of data, "2016 and 2017", are in fact caledar years, and not two whole seasons. This is, in fact, one whole season and two half seasons. Does this influence any of your statistics?

3. Page 3, Line 89: the word "figure" is missing at the end of the line. I think it should be "figure 3", and not 2, as well.

4. Page 4, Line 95: "not be used classify..." should be "not be used to classify..:".

5. Page 5, Line129: It says "...by known snow depth...". How do you know, or where do these values come from? I guess they come from the Warren climatology paper that is mentioned in the next paragraph. Instead of saying "known" at this early sentence, could you perhaps say "climatological", or, say "climatologically derived" or something similar. That will connect better to the later sentence with reference.

6. Page 5, Equation (1): Symbols for h_i and h_s are not defined. Although I can guess them, you should really define them for completeness.

7. Page 9, Line 254: The abbreviation "MS" at the end of the line is not yet defined. Please add it to the first instance of using the MeanShift algorithm, if that is what it represents.

8. Page 9, Lines 264-266: May want to include "interpolation" somewhere here, to connect it to words in your Abstract.

9. Page 9 and 10: Feature List.

   - You may want to remind readers that the two backscatter coefficients are in decibels [dB].
   - The symbols for coefficient of variation F_v is written as C_v later in the text, Page 11, Equation (7).
   - The ordering of many features in this list is different from the order they appear later in the text.

   It would help to be more consistent.

10. Page 10, Equation (3): Why is there a square on the log term? I thought Entropy was essentially p.log(p).

11. Page 11, Lines 315-318: Could you summarise what these "edge and corners" actually are here? So that we do not have to read the refs. to find out.

12. Page 11, Line 323: It looks like the figures should be 4 and 5, and not 3 and 4.

13. Page 11, Line 325: Does the L1 have a more expanded name or reference? I know its fairly well known in mathematics, but this audience may appreciate just knowing that it is a mathematical vector measure.

14. Page 11 and 12, Lines 330-345:

    - It may be interesting to emphasise that this difference function T is a weighted sum of absolute differences, rather than a sum of squared errors, which is more common.
    - Could you have used some sort of standard distance functions?
    - Presumably these weights c_t, c_d, c_k, are to account for different scaling of each feature. Do you also include any priority or enhancement too? Could you have simply normalised all features beforehand and then used a standard Euclidian distance or something?
    - You calculate the coefficients with a least-squares fit, yet the values are absolute not squared.
    - I understand that you will want to capture spatial, temporal, and spectral differences all together. Could you explain or motivate this choice of function better, as it is quite unusual.

15. Page 16, Line 494: "sqare kilometers" should be "square kilometers", and there should be a comma after Also at the beginning of the next sentence.

16. Page 17, Line 531: "melt-donw phase" should be "melt-down phase".

17. Page 18, Line 538-539: Lots of commas missing here, perhaps five of them.

18. Page 27, Figure 1: Is there a systematic bias between ascending and descending tracks. It looks like one set of tracks is mostly red coloured, while the other set are more yellow coloured. Or, is there some other factor like time, due to drift, involved? How is such bias handled in the processing, and do you mention this visible observation?

19. Page 29, Figure 4: Are these mosaics after incidence angle correction or before? There still seems to be brightness differences and edges visible. Please explain.

20. Page 31, Figure 6: The pink boxes do not fit around the text on the PDF (or the text does not fit inside the boxes). Make sure the font sizes work in the final proofs.

21. Page 31, Figure 7: It is quite natural to picture the 1:1 ideal line in such a figure, and then all the curves look quite poor. Admittedly, this could equally be reflecting a bias in the Model SIT values. At least the red, and two blue lines are pretty consistent. Could you please comment on this?

22. Page 33, Figure 10: I notice that this figure, and the caption, only talks about values of the order of 50-100 cm and yet the colour scale goes up to 600 cm. I suppose that you deliberately made all colour scales cover the full range for more consistent interpretation between figures. I now found this mentioned in the main text at the top of page 16. However, it might be nice to mention this in the caption of Figure 10 too.

23. While finding the last comment, I suspect that the figure numbers are incorrect in many places, since you inserted a new figure 3. Please re-check all references to figures to make sure that the numbers are correct.