Referee comment on "Winter Arctic sea ice thickness from ICESat-2: upgrades to freeboard and snow loading estimates and an assessment of the first three winters of data collection" by Alek Aaron Petty et al., The Cryosphere Discuss., https://doi.org/10.5194/tc-2022-39-RC1, 2022

Review Petty et al, Winter Arctic sea ice thickness from ICESat-2: upgrades to freeboard and snow loading estimates and an assessment of the first three winters of data collection.

### General comments

The paper assesses the impacts of a number of changes to ICESat-2 ATL10 processing and to the NESOSIM snow model on estimates of along-track and gridded sea ice freeboard and ice thickness. This assessment is important for users of high level sea ice products such as ATL20 gridded sea ice. Overall the paper is well conceived and written. However, there are a number of issues that need to be addressed before the paper is ready for publications. I list these below. I also have a number of specific comments.

Overall, the quality of the figures is good. However, some of them could be improved by adding descriptive titles/labels to each panel. For example, figure 7 has titles but these appear to be file variable names. Rather than "ice_thickness_unc", it would be more helpful to readers to have "Ice thickness uncertainty" spelled out. Likewise with panel (i) "ice thickness int" would be better as "Interpolated ice thickness". The authors might also want to think about a better layout and if all panels are necessary.

The Jupyter notebook is an excellent addition as is making the code available.
Different releases are used for different evaluations. The authors show that there is little difference between releases 003 through 005 but it would make for a cleaner, and more up to date, analysis to use release 005 throughout. The only exception being to show differences between releases 002 and subsequent releases.

I would like to see a map in the main paper showing the "Inner Arctic Ocean" region as the study region introduced as part of the methods. This would focus readers attention on the analysis region up front.

Figure 8 is another example of a figure that would benefit from having labels such as a) sea ice freeboard. Parameter names are on the y-axes but they are small. Panels a, b, etc should be referenced in the text.

There are a number of places in the text where important statements are put in parentheses. I think it would improve readability to rewrite these statements as part of the main text. Some of these parenthetical statements are unnecessary.

## Specific comments

L60. "is *being* developed"

L63. Suggest "collected to estimate sea ice thickness"

Section 2. I think it would be helpful to summarise upgrades to IS2 processing, NESOSIM and ATL20 gridding in a simple table.

L111 prefer "km" to be consistent.

L124 "0-3 cm freeboard changes at basin scales". Does "basin-scales" refer to the Inner Arctic region used in the current paper? Maybe say "an increase in basin average freeboard of up to 3 cm."
Suggest "New releases of ATL07 and ATL10 also reflect upgrades to the underlying ATL03 processing, such as improvements in geolocation.

ATBD for ATL07/10 use "surface reference" rather than "reference sea surface". To avoid confusion it might be better to use the same terminology as the ATBD.

Figure S1. Would it be better to have this figure in the main text? Also, the point here is that the number of reference surfaces is reduced from rel002 to rel003 because dark leads are not used. However, the count difference is positive. It makes more sense to me to have this reduction as a negative number.

Effectively the /beta and /gamma terms in equation 1 are corrections to solid precipitation. It is not clear to me what the difference is between the two terms. They could be combined into a single loss coefficient.

Do you mean "For each OIB snow depth product, snow depths are binned into 100 km grid cells using a drop-in-the-bucket averaging procedure. For each grid cell, the median snow depth of the three products is then assigned as the grid cell snow depth". So in all cases, you are taking the middle value. If the number of products was larger, I can see this as an acceptable approach to avoid outliers but for just three values, you can't really identify an outlier. It would seem that the mean is a better estimator.

"within reason" This needs some clarification. Are there limits you can set on depth or start date?

Figure 3. The left panel is busy. I suggest having a separate panel for October and April. The horizontal grid-lines should be lighter or removed.

One of the arguments for not using the Warren climatology for snow depth is that it is not representative of the present day conditions. The previous paragraph and Figure 3 have been used to argue that recent years snow depth are also lower than average and may be declining. So why would you use a climatology of NESOSIM. Wouldn't using output from an operation product or low latency reanalysis be a better option?
L266. The redistribution method needs a reference.

L296. The smoothing/gridding procedure needs more explanation. It would be helpful to say why each of the steps are done. Why use Delaunay triangulation - generally this method is used to interpolate unstructured data? Presumably the KDTree algorithm is to speed up the search for neighboring cells.

Figure 4, L343. How do these look for other months and for other years? No need to show them but a comment in the text would be helpful.

L354. Significant or major?

L356. Prefer peak rather than mode. Mode could be confused with operating mode.

Figure 4. How many segments are used to generate these plots? Are dark leads more common in November?

L370. The name NESOSIMv1.1clim has not been introduced yet.

L389. Suggest "In Figure 6, we show the correlation coefficients, mean bias and standard deviations of ICESat-2 monthly gridded ice thickness from rel002 and rel003 compared with ESAs CryoSat-2."

What are the standard deviations of?
Why mask data less than 0.25 m?

Figure 6. I suggest removing the shading and, for each month, plot release 002 and release 003 as separate columns. That way you can see the overlap. The shading suggests the data is continuous rather than discrete monthly data.

L445. NESOSIM prescribes snow density for new and old snow. The
bulk density is a weighted average of these two values. How much can be read into variations in density?

Figure 8. Why is sea ice concentration lowest in October? Is this an artifact of averaging.

Figure 9. The flow vectors obscure the thickness data. They are not really discussed. Are they necessary? Could they be relegated to supplemental material?

Line 535. Care needs to be taken with ERA5 (or any reanalysis) near-surface variables over snow. ERA5 snow parameterisation is still a single layer, which does not produce realistic surface fluxes (Arduini et al 2019).

L540. Are three years of data enough to make a statement about strength of coupling?

L591. This seems to contradict what is shown in Figure 4.

Figure 12 and 13. The multi-year ice fraction panel is not needed.