EGUsphere, referee comment RC1
https://doi.org/10.5194/egusphere-2022-501-RC1, 2022

Comment on egusphere-2022-501

Anonymous Referee #1

Referee comment on "Impact of the sampling procedure on the specific surface area of snow measurements with the IceCube" by Julia Martin and Martin Schneebeli, EGUsphere, https://doi.org/10.5194/egusphere-2022-501-RC1, 2022

Paper # https://doi.org/10.5194/egusphere-2022-501

Impact of the sampling procedure on the specific surface area of snow measurements with the IceCube

Martin and Schneebeli

This manuscript is potentially valuable as it contains an important headline result – that the manner in which snow is sampled impacts up to approximately 50% the SSA values less than 30 m^2 kg^{-1} from the IceCube (IC) instrument in comparison to micro-CT (CT). Potentially, this has big implications for 1310 nm snow reflectance measurements as the IceCube (n.b. also potentially DUFISSS in pre-production, or IRIS in non-commercial form), is an increasingly common and robust in-field instrument for objective measurements of snow microstructure. Micro-CT, a 'gold standard', is as good a direct measure of snow microstructure as we currently have. Behind this headline result, there are a number of issues that need to be addressed for the community to have confidence in the currently proposed message. It may mean the scope of the message needs to be refined with greater detail, and the implications limited to particular snow types.

The 5-25 m^2 kg^{-1} range of snow SSA that shows significant different between IC and CT are often associated with important snow types, e.g. depth hoar or wind slab in Arctic and sub-Arctic snowpacks, that are not part of the experiment. At best this paper needs to be explicitly limited to Alpine snow, otherwise unintended mis-interpretation could occur. Could more detail be provided to describe the Alpine snow types measured, e.g. densities from volumetric sampling? More details of the snowpack from which the samples were extracted would be highly beneficial so the reader can get a feel for snow types which the interpretation is both relevant for and limited to.

The sample preparation process is a key conclusion explaining the difference between IC and CT. However, there is ambiguity in the description of this method. It seems the sample was reduced to size through cutting of unsuitable material and then brushed gently to remove loose particles and measured by IC. Secondly, the 'default method following Gallet et al. (2009)' was followed, then any remaining loose particles were knocked off, then the sample was remeasured using IC. The default method in section 2 of Gallet et al. (2009) refers to the sample measurement face being shaved off with a spatula, in which they state it was difficult to obtain a 'perfect surface'. Hence more needs to be included about the shaving process and how it was applied in this experiment. I got the impression that from Gallet et al. (2009) the shaving/smearing of the surface grains by the spatula (especially when close to freezing) could have had an impact on surface optical reflectance. I expect this not to be the case in cold labs at -15 degrees Celsius, but it requires a more detailed discussion about how preparation of IC sample surfaces effect SSA. Discussions at the Davos Grain Size Measurement workshop in 2014 and my own experience of making IC measurements suggest that the SSA from IC is (thankfully) not very sensitive to sample preparation. The pressure required to cause sintering as part of the sampling process is highly unlikely to be achieved. Rather, making sure the sample container is completely full by addition of snow to fill any gaps in the extracted sample, and light compaction of snow to be flush with the container surface is preferable so that reflectance is less likely to come from the edges of the sample container. This negligible impact of sample preparation appears to be shown in the comparison of distributions of IC + particles and IC – particles in Figure 2, where distributions overlap. As both field experience and results in Figure 2 contradict the message that sample surface preparation is crucial, this message needs to be revisited.

Serious consideration needs to be made as to whether relative percentage difference is a fair way to present the results, particularly when the mean or median values range from <10 to >50 m^2 kg^{-1}. I suggest presenting the measurement uncertainty in m^2 kg^{-1} is more appropriate, e.g. a bias or RMSE. This is illustrated by Figure 2, where the actual difference between extents of upper and lower quartiles between CT out of IC / CT reference and IC respectively, either overlap for type C and E, or are approximately 2-3 m^2 kg^{-1} apart for type A and D. And when the four distributions of CT are considered against the two IC distributions, overlap of distributions is more common than not. Some discussion about what level of natural SSA variability might be expected within a sample (CT or IC) needs to be added here. Depending on the orientation of the sample in IC measurement I would expect variability in spectral reflectance, particularly in snow types that are not highly homogenous in structure, size and orientation. Hence SSA variability of the order 2-3 m^2 kg^{-1} may well be within measurement noise. As an exemplar, Fig 2 shows that differences between distributions of CT surface and CT mid are on the same order of similarity to the difference between IC + particles and CT out of IC. While I expect CT mid to be the best measurement to compare other measurements to, the fact that there is such spatial variability within a CT sample, suggests that the comparison between CT and IC is not drastically worse than the within CT measurements. Can this be discussed in further detail as it appears to add sensible uncertainty caveats to one of the headline conclusions, which is there is a SSA difference of 20-52% in the 5-25 m^2 kg^{-1} range when measured by IC and CT.

The visual and statistical comparison of distributions (Fig 1 and 2) is good, but this raises a concern at the low number of sample values (Table 1 shows n=1-8) which make up these distributions. I appreciate the time required to make CT measurements, so this not being a high n-value is understandable. However, how was the n value calculated for IC? Was it a single sample measured in different orientations? Or were there a number of

samples in the same snow layer? Considering that the IC is designed for field use and implications of results increasingly tend to be considered in recent literature when using larger distributions (n >10) of measurements in similar layers, these are very low sample numbers to be making robust conclusions. However, there is a balance to be struck here. These initial results are useful for the community to see, but I think that it points the way for future work, rather than being definitive about the applicability of IC measurements at SSA <30 m^2 kg^{-1}. It might be a suggestion for a revised version of this manuscript to be a brief communication rather than a full article in TC?

The application of TARTES to provide an explanation of potential impact of the surface grain size on optical reflectance is good. However, as a non-expert, does the choice of a 1m thick snow substrate matter, rather than a sample thickness of 300 mm which is a common depth of the IC sample container?

What is nature-identical new snow (largely >30 m^2 kg^{-1}) in the context it is presented? Figure 1 looks like the distribution of IC + particles (more realistically how samples would be measured in the field) are very similar to the CT distribution, so we need more detail on what is 'nature-identical'. It suggests the rest of the samples are dissimilar to nature, which is a worry when drawing implications from this experiment. This may just be a terminology issue, but it needs to be addressed.

Rather than a list of minor comments, at this stage I would encourage the authors to:

- Expand on the introduction to contrast IC and CT to a broader range of microstructural measurements and implications for their use.
- Increase the clarity of the 'hypothesis statement' at the end of the introduction (i.e. this paper does, this, and this, and this…)
- Add clarity on what is being presented in the box plots (are they median and IQR, or mean and standard deviation - as in Table 1).
- Add more details throughout (as per issues raised above), and in the discussion section add more on the implications of these results for a) field measurement using IC, and b) what using measurements from IC may mean for applications where SSA is crucial.
- Check the cross references are correct (e.g. 'Tab 3' is cross-referenced, but is not in the main manuscript).