Author responses below are in italic.

The authors present an application of a change-detection algorithm to estimate SWE in the Alps using Sentinel-1 C-band SAR. They explore the effect of spatial resolution on their retrievals. This is an important and timely contribution, and should be of great interest to the community. The paper is well-written so I have very few minor comments. Instead, I'll focus on a really key point which is that I think there is a great chance for readers to misunderstand the maturity level of the algorithm, based on how the paper is presented. This review is five related major comments that unpack this idea.

We are grateful for the assessment of our work by the reviewer. Please find below our response to the five posted comments.

Major Comments

First, I do not think that the paper adequately reflects the fact that we still do not understand why this method works, even at a basic level. The manuscript instead makes it sound clear that the mechanisms are understood: e.g. in the introduction, page 2, lines 32-page 3, line 2. Taking their points one by one: to their first point (page 2 line 33), no reference was given, and no reason why having lower ground backscatter would change sensitivity to depth; to their second point (page 2, line 33), Chang et al. 2014 do not make this point, that I could see. Readers will assume after reading the introduction that it is obvious why the C-band cross-pol is correlated with snow depth, which is not true. In fact, the authors of this study only introduce the idea that the “physical mechanisms that cause this increase are still uncertain” in the Results & Discussion section (page 10, line 13). Please, bring this critical point into the abstract, introduction and conclusion!

We fully agree that we need to better inform the reader about the current limitations in physical understanding of C-band sensitivity to snow, upfront in the paper. We will modify this in the abstract, introduction and conclusion of the revised version.

Regarding the statements on page 2, lines 32 to page 3, line 2: The first statement on page 2 line 33 (“surface scattering from the ground is significantly weaker in cross-polarization”) refers to the common understanding that cross-polarized backscatter is typically several dB lower than co-polarized backscatter. This is especially the case in
regions with limited vegetation and for smoother surfaces (vegetation and surface roughness increase depolarization and thus the cross-polarized backscatter, which, however, will generally still remain lower than the co-polarized backscatter). It is also common understanding that in a logarithmic (dB) scale as used in the retrieval algorithm, an increase in scatter intensity (in linear scale) will have a relatively larger impact when the prior intensity is low; hence the statement that a lower ground backscatter can be beneficial for the sensitivity to snow.

Regarding the reference to Chang et al. (2014): They mention the following statements in their introduction that support our quote (i.e., "dry snow represents a dense medium of irregularly-shaped and clustered ice crystals that primarily causes volume scattering in cross-polarization"): "In snow, the ice particles are packed closely together", "ice grains in snow do not scatter independently", "shapes are irregular and there are clustering effects", "In conventional scattering models, there is no cross-polarization in scattering when particles are spheres. In the dense media model, the electric dipole interactions of closely packed ice grains result in strong cross-polarization in the phase matrices".

Second, I think it is critical to communicate more clearly throughout that this is an empirical algorithm with calibration parameters that require known SWE data over the domain. The word “empirical” needs to appear in the abstract, in my opinion. Please somehow get this idea into the introduction, abstract, and conclusion.

We agree and will modify the text accordingly.

Third, the authors need to point out that the algorithm only works well if you have accurate SWE data to calibrate against. Indeed, they need to just note explicitly that the accuracy of the approach they are using here is limited to the accuracy of their training data. I think this needs to be presented explicitly in the abstract and conclusions, to avoid reader misunderstanding.

The approach indeed requires reference data (snow depth) in order to estimate the scaling coefficient that translates the changes in backscatter into changes in snow depth. However, we would like to highlight that the approach works already reasonably well when using a single, constant (both in time and space) scaling coefficient. For instance, Lievens et al. (2019) apply a constant scaling factor across all mountain ranges in the Northern Hemisphere, which still leads to relatively accurate retrievals. Therefore, the need for accurate reference data is not considered to be critical.

Figure 8 shows a positive (but overall limited) impact of refining the scaling coefficient, by allowing it to vary in space (but still not in time). Such refinement of the scaling coefficient may indeed require more accurate reference snow depth data, but again, this impact is limited.

Fourth, the authors should point out that in this study, they are calibrating here against very accurate model results. Here, they are applying the algorithm in this study over a domain where (in my opinion) the most accurate model results are available anywhere in the world. There is no other mountain range, to my knowledge, with the density of observations available in the Alps. Further, globally available model results in mountain ranges are inadequate for most applications, in terms of their spatial resolution and accuracy. See e.g. Mortimer et al. 2020. I think this needs to be mentioned in the conclusions.

We will mention the potential challenges to replicate our approach (with respect to refining the scaling coefficient based on model simulations) in other mountain ranges, due to the limitations in model simulations. But as mentioned above (and shown in Figure 8), the added value of the spatially varying scaling coefficient is not critical.
Fifth, the authors need to acknowledge explicitly that the first four points mean that you could not use this approach globally, calibrated to models, and achieve the kind of results shown here; this point almost certainly will be lost on readers of the abstract alone. This is a major issue with the manuscript that needs to be addressed in the abstract and conclusions.

The approach with constant scaling factor is applicable globally (in regions with sufficient snow accumulation) and has previously been applied over the Northern Hemisphere (Lievens et al., 2019). We here show that only a slight reduction in performance is to be expected in the case that insufficient or inaccurate reference data would preclude a (spatial) refinement of the scaling coefficient. We will mention this implication in the revised version.

I hope the authors do not misinterpret any of these comments: they have done an amazing job uncovering this important new dataset. It has very important possible applications. Reworking the way the paper is presented should help the community get on board with this new dataset as quickly as possible.

Thank you for this supportive comment.
