Title: Sentinel-1 snow depth retrieval at sub-kilometer resolution over the European Alps
Authors: Hans Lievens, Isis Brangers, Hans-Peter Marshall, Tobias Jonas, Marc Olefs, Gabriëlle De Lannoy
Submitted to: The Cryosphere

Author responses below are in italic. Page and line numbers refer to the revised manuscript version with changes marked in bold font.

1. Response to editor comments

Comments to the Author:

Dear Dr. Lievens:

Thank-you for your responses to the reviews of your manuscript.

There is clearly a difference in perspectives with Reviewer 1, which need to be addressed in the revised manuscript. Reviewer 1 accepts the potential for C-band backscatter sensitivity to the dry snow volume of deep mountain snowpacks (through the probability for greater snow anisotropy and hence greater scattering and depolarization of the SAR signal in deeper snow) but recommended rejection of the manuscript based on the level of processing of the Sentinel-1 imagery ("With all the processing done to the SAR imagery, it is impossible to assess the physical interactions of the SAR signal with the snowpack since the data has been smoothed multiple times and transformed radiometrically and geometrically.") Reviewer 1 also identified a series of issues with the statistical approach used in the validation (including the treatment of outliers and zero values) which will require new calculations. I was pleased to see your response to the important comment regarding open data standards.

Reviewer 2 recommended major revisions, essentially asking for a reworking of the manuscript to clearly communicate the lack of understanding of physical mechanisms, the empirical nature of the retrieval, and the impact of calibration.

Both reviewers positively note the amount of work that has been done to mine the Sentinel-1 dataset, the overall quality of the writing, and the potential impact of this work across the snow and hydrology communities. It is the very nature of this high impact which means the Sentinel-1 data processing decisions, a clear description of the physical mechanisms (or lack thereof), robust validation, and a clear description of the empirical nature of the retrieval (including the role of calibration) should be made more explicit in the manuscript. Your response to the reviews indicates agreement on most of these points.

A public comment (to which you have already submitted a reply) also expressed reservations similar to those raised by the two reviewers. It is my interpretation that revision of the manuscript to meet the issues raised by the reviewers will also go a long way to addressing the concerns of this comment.

My assessment is that major revisions are required and the manuscript will be returned for an additional round of review, after which a final editorial decision will be taken.

Thanks very much for your contribution to The Cryosphere Discussions.

Chris Derksen

Dear editor, we'd like to thank you for the assessment of our manuscript. We are also thankful to the reviewers for providing a detailed assessment. We have modified the manuscript substantially to account for the reviewer comments as summarized below, which we believe has improved the quality of the work. The public comments by Helmut Rott were much appreciated by the authors and have also been accounted for in the associated response and/or in the revised manuscript.

Based on the recommendations of reviewer 1:

- *We fully re-processed the S1 backscatter data, in order to perform the terrain flattening and terrain correction at high resolution (similar to that of the DEM), and to output backscatter as gamma0.*
- *The number of backscatter post-processing steps has been reduced. More specifically, the rescaling of sigma0 based on the incidence angle was no longer required by the use of gamma0, and the outlier correction has been deactivated.*
- *Using the newly processed gamma0 data, all analyses steps (e.g., snow depth retrieval, validation) have been repeated.*
- *The theoretical discussion has been extended to include the potential impacts of snow anisotropy and multiple scattering on layer interfaces, and future research has been recommended to investigate the impact of snow microstructure and stratigraphy (also in response to the comments by H. Rott)*
- *We have revised the validation to clearly distinguish between spatial and temporal correlation (Fig. 11; also in response to the comment by H. Rott), and to provide validation metrics for both scenarios with and without zero snow depths included. Further, the mean absolute error and bias are stratified by the measured snow depth in Fig. 12., and presented both as absolute and relative values.*
- *The wet snow detection algorithm was extended and modified to start earlier and to allow the calculation of fractional wet snow cover. This allows the user to define the preferred level of wet snow masking.*

Based on the recommendations of reviewer 2:

- *We now explicitly mention the empirical nature of the retrieval algorithm in the abstract, introduction and conclusions.*
- *Similarly, the lack of physical understanding about why C-band SAR data are sensitive to snow depth is highlighted at various places in the manuscript, including the abstract, introduction and conclusions.*
- *We have discussed more in detail the scaling of the retrievals (calibration) based on reference data and its implications to apply the method in other regions.*
- *We highlighted the need to further test and validate the method in other regions in a dedicated paragraph at the end of the conclusions.*

As a result of the above-mentioned modifications, the performance of the retrievals and validation analysis has improved. The wet snow masking component is more evolved and allows a more flexible choice in terms of the preferred level of wet snow masking. The discussion on the physical basis of backscatter sensitivity to snow has been extended, also clearly identifying the need for further research. To allow anonymous access to the data during the manuscript review, we here provide the login details to the server:

Server: *sftp://hydras.ugent.be*
Login: *sentinel1snow*
Password: *s1!csnow*
Port: *2225*

We are hopeful that these modifications adequately address all the comments by the reviewers. Please find a point-by-point reply to each of the comments below.

Sincerely, Hans Lievens, on behalf of all co-authors.

2. Response to Reviewer 1 comments

This paper builds on the work of Lievens et al., 2019 to extract snow depth from S-1 data in the Alps. As mentioned by the editor, this work is of high relevance to the snow community but also to many other research areas such as water management, tourism, climate change and biodiversity. I appreciate the work that is done here but in its current state, I cannot recommend this paper for publication since I feel there are too many unknowns and too much processing done on the S-1 imagery to be able to retrieve some sort of good quality snow information and give a proper assessment of the results shown here. This is reflected in my comments below.

Contrary to what has been stated by the authors in their response to the editor's comments, I am not skeptical of the relationship between the C-band signal and thick alpine snowpacks. I do question the physics of the approach used in this study and am concerned about the multiple layer of data smoothing in order to get good correlations with modelled data.

If the authors are willing to provide more information on the imagery processing and modify it to make it more physically accurate, I strongly believe this work has great value to the scientific community.

We would like to thank the reviewer for the detailed assessment of our work. To address the reviewer's main concern about the processing of the S-1, we have carried out a full re-processing of the S-1 data across the Alps with a revised methodology (see details below), without any significant change in the results or conclusions.

We however strongly disagree with the statement that the good correlations with modelled data are due to the multiple layers of data smoothing. The smoothing applied here is limited, as discussed below.

We are hopeful that the revised processing of the S-1 data and a more detailed discussion of the processing and algorithm steps (including some modifications, e.g. regarding the wet snow detection) adequately address the main concerns of the reviewer.

General Comments:

As mentioned above, I do agree with the authors that the cross-pol channel of S-1 can be sensitive to a thick snowpack but I disagree with the physical explanation of the authors. The physical interaction of the microwave signal with the snowpack is very complex and is not solely related to surface/volume scattering and single/double bounce. With snow layer thicknesses close or smaller than the wavelength, you have many interference and coherence effects in the signal. Recent work has shown that volume scattering and depolarization of the SAR signal comes mostly for the snow anisotropy (Leins et al., 2016) and the vertical/horizontal structuring of the snowpack at C-band. This can be achieved by a stratified snowpack horizontally or with snow grains that are structure vertically/horizontally through metamorphic processes. I would agree that with a thicker snowpack, chances are you will get more anisotropy but this is not shown with in situ measurements, temporal analysis or snowpack stratigraphic information.

We appreciate the reviewer's comments on the physics of the signal. Although the mentioned work by Leinss et al. (2016) investigates only higher (X- and Ku-band) frequencies, we agree that the anisotropy of snow crystals and/or of crystal clusters, as well as the snow stratigraphy, can play an important role.

The manuscript has been modified as follows:

P3, L2: "More specifically, dry snow represents a multi-layered, dense medium of irregularly shaped (anisotropic) ice crystals that can form larger-scale clusters. Signal depolarization can for instance occur due to scattering within the dense anisotropic snow volume, multiple scattering on snow layer interfaces and snow-ground scattering interactions (Du et al., 2010; Chang et al., 2014; Leinss et al., 2016)."

P3., L29: "Future research is recommended to further investigate the physical scattering mechanisms in snow at C-band, including the impacts of snow microstructure and stratigraphy, …"

P5, L21: "The snow depth retrieval algorithm is based on the assumption that a snowpack is typically composed of multiple snow layers, where each of the layers represents a dense medium of clustered, irregularly-shaped (anisotropic) ice crystals within an air background. We hypothesize that the microwave signal in the snowpack depolarizes primarily by the scattering on anisotropic clusters of snow crystals within the snow volume (Chang et al., 2014), by the multiple scattering between snow layer interfaces (Du et al., 2010), and by snow-ground scattering interactions. A deeper snowpack is likely to result in more opportunities for signal depolarization and therefore stronger scattering in cross-polarization."

P13, L5: "The dry snow scattering contribution is stronger, most likely because depolarization occurs after volume scattering on anisotropic crystals or clusters of crystals, multiple scattering on snow layer interfaces, and snow-ground scattering interactions."

With all the processing done to the SAR imagery, it is impossible to assess the physical interactions of the SAR signal with the snowpack since the data has been smoothed multiple times and transformed radiometrically and geometrically. You have multi-looking (averaging 10x10 pixels), border noise removal, thermal noise removal, terrain correction and reprojection to the WGS84 projection. The multi-looking is especially concerning given the topographic complexity of the Alps. It is smoothing all the topographic information (which is crucial for snow retrievals) and emphasizing only the areas of significant snow (snow drifts) which is not representative of a 100m grid cell in the Alps. Then you add incidence angle correction using a DEM (30m) that is of lower resolution than the pixel spacing (10m) of the original image. A DEM with similar resolution should be used but also, the topographic information has already been altered from the multi-looking which is not representative of the local topography. Then there's temporal averaging (Eq.2) which alters the signal even further. Finally, outliers are replaced by a 12-day average to smooth the data once more.

We would argue for the contrary: a careful processing is a pre-requisite in order to assess the correspondence between the S-1 signal and snow depth. The processing steps included in our analysis (border noise removal, thermal noise removal, multi-looking, terrain correction and reprojection to a consistent grid) are all standard and necessary procedures, recommended by any manual or handbook on SAR processing. The multi-looking is arguably the only processing step that could be considered optional. However, this was included in order to (i) reduce the impact of radar speckle, (ii) reduce the processing time (note that more than 4000 S-1 images were processed), and (iii) reduce the data storage requirements. In this context, the multi-looking is an important step to keep the processing computationally feasible also for larger areas, not limited to the Alps.

However, to address the reviewer's comment about the correction with the DEM, we have re-processed the S-1 data over the Alps, by performing the range-Doppler terrain correction and terrain flattening at the 20 m S-1 resolution instead of at the multi-looked 100 m pixel spacing. We kept the original 30 m DEM (SRTM 1Sec HGT) because this is the standard suggested DEM for processing in the ESA SNAP toolbox and can also be applied in other regions (e.g., outside Europe, or where more detailed DEM information is lacking). However, the pixel sizes of the DEM and the S-1 data were now much more similar with the re-processing. Equation 2 (Equation 1 in the revised manuscript) is not performing temporal averaging as stated by the reviewer. It applies a bias correction (of the first two order moments, i.e., the mean and variance) to every individual backscatter observation, without averaging observations over time. The bias correction reduces the differences between observations from different orbits (e.g., caused by different incidence or azimuth angles) and we strongly recommend this step for any application that aims at combining information from different S-1 orbits. We have deactivated the outlier correction in the retrieval because we observed it was slightly interfering with the wet snow detection algorithm.

Further on the processing, I would avoid talking about sigma-nought when Eq. 1 converts the sigma-nought into a pseudo-gamma-nought multiplied by cos(40). I say pseudo here because the incidence angle used to convert sigma-nought is the 100m reprojected angle and not the gamma-nought values from the SAR imagery calibration.

At the time of the S-1 processing for our initial manuscript submission, the calculation of gamma0 was not operational in the SNAP software version 7. In the revised processing, we appropriately calculated gamma0 using SNAP version 8, by first calibrating the backscatter observations to beta0 and subsequently applying terrain flattening. The entire analysis, including the snow depth retrieval and validation, has been repeated with the use of backscatter as gamma0 and is thus more conform the state-of-the-art.

If we accept the processing chain of the SAR imagery, it is still unclear that what the correlations are showing is linked to the snow depth. The errors obtained from the SAR retrievals (Figure 11) are most of the time larger than the precision of the reference data which is the model simulations. It is very difficult to determine that the correlations are statistically significant in this case and also looking at Figure 10, most of the comparison points are grouped around 0 which tends to falsely boost the correlation.

We do not see any reason to not accept the processing chain, especially considering the re-processing discussed above, which is fully compliant with the state-of-the-art. Figure 11 shows a comparison against in situ measurements as reference data (not model simulations). We are surprised that the reviewer questions the significance of the time series correlations in Figure 11, which are typically above 0.8 (for sites that feature snow depths thicker than a meter).

With respect to Figure 10, we agree that the abundance of low snow depths impacts the correlations. To identify this impact, we have calculated and reported validation metrics (R, MAE) for scenarios both with and without inclusion of zero snow depth values in Figure 10. Even though many data are clustered around low snow depths, the density plots in our opinion still clearly demonstrate the overall agreement between the S-1 retrievals and the in situ measurements also for the high snow depths, especially for the coarser 300 m and 1 km retrievals. The revised Figure 12 supports this statement, showing that the lowest relative errors are obtained for snow depths of about 2-2.5 m.

Given that modelled data is often smoothed and often have difficulty capturing extreme snow conditions and that the SAR data has been smoothed many times and outliers replaced by temporal means, I can't say I am surprised to see a good empirical relationship.

We respectfully disagree with this comment, as the S-1 processing does not include multiple smoothing steps, unlike what the reviewer states (see above). In our opinion, the strong relationship between the S-1 retrievals and the model simulations is encouraging, and is furthermore corroborated by the strong correspondence between the S-1 retrievals and in situ measurements.

Also, asking scientists to identify themselves in order to get access to the data used in this study does not comply with the open data policy.

We understand this comment. To share the snow data over the Alps, we have used the existing platform via which we also share the corresponding retrievals across the Northern Hemisphere mountains. Please note the following login details, which allow to directly access the ftp server anonymously:

Server: *sftp://hydras.ugent.be*
Login: *sentinel1snow*
Password: *s1!csnow*
Port: *2225*

Specific comments:

P.3L.5: I would disagree with the claim that an increase snow depth automatically causes an increase in volume scattering. If their is not sufficient anisotropy in the snowpack, there will not be any volume scattering in C-band. The theory will show that even if you increase the snow depth and keep all other snowpack parameters constant, you will not have a significant increase in volume scattering

We have improved the formulation of the potential mechanisms that can enhance the scattering in cross-polarization from the snowpack (see above). Please also note that recent radiative transfer model simulations using Bic-DMRT have shown that cross-polarized backscatter at C-band can increase with an increase in SWE (or depth) while keeping other parameters (snow grain size, snow clustering) constant (personal communication with Prof. L. Tsang, University of Michigan).

P.3L.6: Again, this comment is highly dependent on the stratigraphy and anisotropy of the snowpack. This section needs to be supported by snowpit measurements of the studied area or referred to past work done in the area analyzing the snowpack properties.

The statements on P.3L.6 were general assumptions based on which the empirical change detection retrieval approach is built. We have not yet analyzed these assumptions using snowpit measurements as suggested by the reviewer, but this is foreseen in future research. However, Figures 3 and 4 (based on model simulations) support the statements that (i) an increase in snow depth generally increases (especially cross-pol) backscatter, that (ii) the snow scattering (in cross-pol) is not negligible compared to the ground scattering, and (iii) that ground surface properties remain relatively constant in time due to the insulating properties of snow, thus the main changes in backscatter over time relate to changes in the snowpack.

The need to further investigate the impacts of snow microstructure (including anisotropy) and stratigraphy has now been highlighted at several places in the revised manuscript (e.g., P3, L31; P23, L17).

P.3L.7: This comment is most likely true for the studied area but again, no reference or field measurement is provided to support this claim.

Please refer to the response above.

P.3L.9: Again here, I strongly disagree with this claim. The microstructure, anisotropy changes and stratigraphy, especially in the bottom layers of the snowpack will most likely drive the changes in sigma0.

We aim to further investigate the impact of microstructure and stratigraphy in future research, based on tower-mounted radar measurements currently being collected in the Rocky Mountains, US. For now, we have generalized the statement in the revised manuscript to "the main changes of backscatter over time can be related to changes of the snowpack".

P.3L.30: Even though this is common processing of SAR imagery, this is considerably altering the SAR signal, considerably smoothing it and making it very difficult to link to any ground snow properties.

The alternative (i.e., not performing thermal noise removal, border noise removal, radiometric calibration, and terrain correction) would lead to inferior processing results, which we believe would be far less suitable to investigate the relationship between backscatter and snow depth. The multi-looking has been adjusted and is not impacting the terrain correction and terrain flattening in the re-processing. Please also refer to our responses above.

P.3L.32: Multi-looking (or block averaging here) is a good way to reduce speckle noise in flat terrain. Here though, the topography is very complex (as mentioned by the authors) and it is emphasizing on the geometric distortions and the areas of significant snow (snow drifts) which is often not representative of a 100m grid cell in alpine areas.

Please refer to our responses above.

P.4L.10: Using "local" incidence angle correction on a multi-looked image is not an accurate method. A DEM with similar resolution as the raw image should be used to correct for local incidence angle before multi-looking.

This has been addressed by the reprocessing to gamma0 with terrain correction and flattening being applied at the 20 m S-1 resolution.

P.4L.15: This relationship was developed for areas of flat terrain and is not representative of the studied area. Proper analysis of the backscattered signal as a function of local incidence angle needs to be conducted in alpine areas in order to find the proper normalization relationship. A before and after image should show that this is not normalizing the image properly. Also, this is exactly taking σ-nought and converting it to gamma-nought and then multiplying it by cos(40).

This comment has been accounted for by processing to gamma0.

P.4Eq.2: Here again, temporal smoothing of the data. There's no way of linking the spatio-temporal snow properties of the original SAR imagery.

We disagree. Equation 2 (now Equation 1) is not performing temporal smoothing, but bias correction, which results in an improved S-1 processing quality and therefore benefits the analysis with respect to snow depth.

P.4L.27: Excluding March to July is very subjective here. First, it is removing a lot of snow properties variability which can occur in March. Anisotropy and stratigraphy is stronger in the later winter season. Second, with climate change, we know that wet snow is detected outside of this period.

We are convinced that the remaining period, i.e., from August through February for two consecutive years, includes sufficient backscatter observations to allow an accurate calculation of the temporal mean and standard deviation, which are needed for the orbit bias correction. As mentioned in the manuscript, the reason why we exclude March to July is to avoid strong impacts from wet snow on the calculation of the mean and variability, and thus on the backscatter processing for dry snow conditions that are most important for the retrieval. We fully agree that wet snow can also impact the observations earlier than March. Therefore, we have revised the wet snow detection in the retrieval algorithm, to not limit the detection only to the period from February onwards (see response below).

P.4 L.30: This is not rigorous. Removing outliers is another method to smooth out the data and get better correlation with modelled data. But here they are not only removed, they are replaced by a smoothed average.

We have deactivated the outlier removal, because we observed it was slightly interfering with the wet snow detection. More specifically, the outlier removal caused some wet snow events to be undetected, because the backscatter had been modified by the outlier correction.

P.5 Eq.5: Is A applied to the ratio or only the cross-pol channel?

A is applied only to the cross-pol channel, enhancing the sensitivity to snow depth which is primarily driven by the cross-pol observations.

P.6L.17: I appreciate this approach where the index varies in time but I feel like the threshold is still limiting. I would see a temporal analysis of the SAR signal through multiple years to try and identify the proper threshold.

We have tested a range of threshold values. The lower the threshold, the better wet snow impacts are reduced, however, at the expense of reducing the coverage. We identified a threshold of 1.5 to strike a balance between wet snow filtering and data coverage. Please refer to Figure 4 showing the effectiveness of the wet snow detection algorithm for two years at two locations (in Austria and Switzerland).

P.6L.25: Again, the February start is very subjective as wet snow conditions can be detected earlier and the September-November period is most likely to be the period where you have the highest backscatter and all the values that are 3dB below might be because of small surface moisture or percolating water which is not uncommon in Alpine snow.

In the revised version, we have activated the previous wet snow detection mechanism earlier (in January), and included an additional wet snow detection mechanism. The latter consists of (i) excluding backscatter observations (between the start of the snow season and end of December) that are a threshold (e.g., 2 dB) below the 10-percentile of backscatter observations during snow-free conditions, and (ii) excluding negative snow index values from January onwards. More specifically approach (i) improves the detection of early wet snow, often in autumn, whereas (ii) mainly improves the wet snow detection in the valleys, where a sharp decrease in backscatter during snowmelt is often lacking (e.g., due to forest cover). Furthermore, the wet snow is now provided along with the unmasked snow depth retrievals, allowing the user to choose whether or not to mask out wet snow, or to use another mask (e.g., derived from modeling or an alternative wet snow detection approach). In the 300-m and 1-km datasets, the wet snow is provided as a fraction (0-1) of wet snow pixels, which allows the user to define the level of wet snow allowed. Please refer to the manuscript (P7, L19) for details on the wet snow detection method. Results with various levels of wet snow masking are shown in the revised Figure 10.

Conversely to what the reviewer hypothesizes, the September-November period is typically the period in time with the lowest S-1 backscatter values, especially in cross-pol (if not including wet snow conditions in spring). Earlier (in summer), vegetation often contributes to higher backscatter, whereas in mid-winter, a higher backscatter is caused by snow accumulation. Part of the lower backscatter values in September-November can also be explained by the potential freezing of the soil surface, and/or by early wet snow.

P.11L.7: There is no mention of layering and anisotropy which is most likely the main reason of signal backscattering of dry snowpacks.

This has been included in the revised manuscript. Please refer to our response above.

P.11L.11-13: These comparisons do not really apply to the current studies. As was mentioned by the authors in the response to the editor these studies were conducted in shallow snow conditions in tundra/taiga landscapes.

The Alps include areas with shallow snow for which the references to literature are relevant. The literature comparison also helps to indicate that (to our best knowledge) studies with cross-pol observations in deep snow are lacking.

P.11L.20: This is a strong assumption since in alpine regions you can have strong surface roughness that will depolarize your signal.

The ratio of cross- over co-polarized backscatter is considerably lower in areas with limited vegetation. Hence, this statement is supported by S-1 observations.

P.11L.33: This is normal since most of the volume scattering and depolarization will come from the forest cover. For this study, I would have masked out the forested areas because this adds unnecessary complexity to a study that is already complex. Masking the forested areas would allow to focus on the snow retrieval without getting confused in multiple empirical relationships and heavy data processing.

One could either mask out the forested regions, or stratify the performance based on forest cover. We here opted for the stratified performance assessment (see Figure 6), which is more complete. We do not consent with the assessment of 'heavy' data processing.

3. Response to Reviewer 2 comments

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 have highlighted this in the following sections:

Abstract (P1, L14): "However, future research is recommended to further investigate the physical basis of the sensitivity of Sentinel-1 backscatter observations to snow accumulation."

Introduction (P2, L33): "Although further research is needed to improve our basic understanding of the physical scattering mechanisms, we hypothesize that …"

Introduction (P3, L29): "Future research is recommended to further investigate the physical scattering mechanisms in snow at C-band, including the impacts of snow microstructure and stratigraphy, and to extend the validation over regions with different soil, vegetation and snow conditions, also using validation data at the matching scale of the satellite retrievals."

Conclusion (P23, L15): "Further research is needed to investigate more in depth the physical basis of C-band radar backscatter sensitivity to snow, for instance based on tower radar measurements and corresponding detailed measurements of snowpack properties, including snow microstructure and stratigraphy."

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 have modified the text accordingly. The empirical nature of the retrieval algorithm has been mentioned at various places (P1, L7; P3, L10; P6, L3; P7, L3; P8, L14; P21, L15).

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 snow depth data 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.

The following text is incorporated in the manuscript to address this comment:

P7, L7: "Note that the scaling based on model simulations can propagate simulation uncertainties into the retrieval."

P7, L8: "As a first approximation, E1 is considered constant in space and time, conform with Lievens et al. (2019). This approach is applicable to mountain ranges for which no detailed information on the spatial snow *depth distribution, e.g., provided by model simulations, LiDAR data or dense networks of in situ measurements, is available."*

P7, L16: "Note that such approach may improve the spatial distribution of the snow depth retrievals, but is only feasible in regions where sufficient reference information is available."

P23, L18: "Also, future research is recommended to investigate the validation of the S1 snow depth retrievals using data at the matching scale, for instance from LiDAR, and to further investigate the performance of the approach in other regions with different soil substrate types, vegetation conditions and snow climate conditions."

P3, L29: "Future research is recommended to further investigate the physical scattering mechanisms in snow at C-band, including the impacts of snow microstructure and stratigraphy, and to extend the validation over regions with different soil, vegetation and snow conditions, also using validation data at the matching scale of the satellite retrievals."

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 agree with the reviewer that the Alps is a region for which very detailed and accurate model simulations are available, which would not be available in most other regions. However, we believe that the availability of distributed snow depth reference data is not critical for the application. Figure 8 shows that the use of distributed snow depth information in the calibration only slightly improves the results, compared to the use of a simple, constant scaling. Please also refer to our response above.

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 all mountain chains in 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.

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.

Mortimer, C., Mudryk, L., Derksen, C., Luojus, K., Brown, R., Kelly, R., & Tedesco, M. (2020). Evaluation of long-term Northern Hemisphere snow water equivalent products. The Cryosphere, 14(5), 1579–1594. https://doi.org/10.5194/tc-14-1579-2020

Lievens, H., Demuzere, M., Marshall, H.-P., Reichle, R. H., Brucker, L., Branger, I., de Rosnay, P., Dumont, M., Girotto, M., Immerzeel, W. W., Jonas, T., Kim, E. J., Koch, I., Marty, C., Saloranta, T., Schöber J., and De Lannoy, G. J. M., Snow depth variability in the Northern Hemisphere mountains observed from space, Nature Communications, 10, 4629, 2019.