Author responses are shown in blue. Page numbers and lines refer to the manuscript with tracked changes.

Dear Dr. Lievens:

Thank you for submission of responses to the initial review comments, and the revised manuscript. Both reviewers have provided their assessment of the revised manuscript, and we are in a situation with divergent recommendations. While Reviewer 2 now recommends the revised manuscript for publication, Reviewer 1 has again recommended rejection. Given this situation, I have also closely reviewed the responses to the review comments and the revised manuscript. My assessment is that two issues require further attention before the manuscript can be accepted. These are outlined below. I would like to preface these comments by affirming that both Reviewers acknowledge the level of effort that has gone into this analysis, and that the work is presented is a well-written fashion. What remains is the clarification of a number of points related to better understanding the physical mechanisms driving the C-band radar response to snow depth, including the impact of Sentinel-1 processing decisions, and the impact of shallow snow (depth < 1 m) on the validation results. Because this work has a high degree of potential impact to the snow remote sensing community and beyond, I think it is important that these issues are satisfactorily addressed. As Reviewer 2 stated in their initial review "Reworking the way the paper is presented should help the community get on board with this new dataset as quickly as possible."

Dear editor, we are very grateful for your efforts to assess our responses to the review comments and the revised manuscript. Please find below our responses to your suggestions regarding the two identified issues that require further attention.

1. Reviewer 1 remains unconvinced that snow depth is driving the C-band radar response after the processing steps that are applied to the backscatter data. It is important to consider here that the Reviewer is not skeptical of a potential relationship between the C-band signal and deep alpine snow, rather the concerns remain focused on the impacts of the radar processing, and ensuring that it is indeed the snow depth that is driving the response.

My recommendation is to add an appendix to the manuscript which clearly traces the impact of all the processing steps on transforming the backscatter values from the raw measurements, using the analysis-ready backscatter illustrated in Figure 4 (time series) as the endpoint. A step by step series of figures showing the impact of multi-looking, reprojection, orbit correction, and averaging (weekly temporal; spatial aggregation) would provide the necessary clarity on how the raw radar data are transformed before the empirical retrieval is applied. This may also help address the finding that "…skill improves with coarser scales…", the physics of which are not presently addressed nor explained.

We appreciate this suggestion, which would indeed provide additional support to our processing methodology. We have decided to go a step further by modifying our processing chain in order to specifically address two out of the four reviewer comments on the processing. More specifically, we (i) de-activated the orbit bias correction step, and (ii) modified the retrieval algorithm to (a) only trace the changes in backscatter between successive observations from the same orbit and (b) provide snow depth estimates at the times of the actual satellite acquisitions to provide daily coverage over the available satellite swaths rather than weekly coverage over the entire Alps. We agree with the reviewer that this will allow to potentially better capture the sub-weekly variability originating from processes such as accumulation, sublimation, wind compaction and re-distribution, etc.

Regarding the other two processing comments (on multi-looking and projection), we are strongly opposed to modifying these basic and standard processing steps, which are moreover indispensable. Not applying multi-looking implies a backscatter processing and snow depth retrieval at the original 10-m pixel spacing that becomes computationally infeasible, both in terms of processing time and storage requirements. Not applying a projection onto a consistent coordinate system impedes a time series analysis (and change detection) and strongly reduces the applicability of the snow depth output. We also expect by no means that either of these steps can impact the main conclusions of our work.

Please note that potential reasons for skill improvements with coarser scales are discussed in the manuscript at Page 19, Line 34 to Page 21, Line 10, at Page 23, Lines 1-6, as well as in the Conclusions (Page 24, Lines 4-6), and that the spatial aggregation is applied to the snow depth retrievals and not to the backscatter data during processing.

2. The implication of Figures 4, 8, and 10 is that the backscatter is responding to snow depth across the full range of reference snow depth values (0 to ~3 m). There is no evidence in panels a, b, and d of Figure 4, that the retrievals when snow depth is < 1 m are any more uncertain than when snow depth is > 1 m. The clear message from Figures 11 and 12, however, is that uncertainty is greater when snow depth is <1 m. At present, the snow depth dependence of the results is not clearly assessed in Section 4:

"Sentinel-1 (S1) backscatter observations, particularly in VH-polarization, correlate well with regional model simulations of snow depth over Austria and Switzerland." There is no mention of snow depth dependence.

"The main uncertainties in the S1 snow depth retrievals are expected to be caused by wet, shallow and occasional snow cover and forest cover." The influence of wet snow and forest cover on the radar signal are clear, but what is it about shallow snow that causes greater uncertainty? Is it just the ground influence?

Also:
-Do Figures 11 and 12 reflect exactly the same data shown in Figure 10? (I assume yes).
-Can you add some additional text to Section 4 which provides an underlying physical explanation for the depth threshold of ~ 1 m in influencing the retrieval skill?

Figures 8 and 10 (Figures 9 and 11 in the revised version) show a similar spread for shallow snow compared to deep snow, but this similar spread will lead to a larger relative error for shallow snow (as shown in Fig. 12 in the original or Fig. 13 in the revised manuscript). Figure 6 (in the revised version) also reveals higher time series correlations between snow depth and backscatter for areas with more snow accumulation. We expect that in shallow snow, the backscatter observations are more sensitive to the conditions of the ground surface. At the same time, areas with shallow snow are typically also more prone to wet snow conditions and to the dis- and reappearance of snow cover, which we expect to be impacting the backscatter signal at C-band.

In this context, the following statements have been added to the revised manuscript:

Page 7, Lines 21-23: "An additional wet snow detection criterion mainly addresses regions where no strong decrease in gamma0 is observed due to a lower sensitivity to snow wetness. This is for instance expected in regions with shallow or patchy snow cover where the soil scattering contribution may dominate"
Page 12, Lines 33-34 "In regions with shallow maximum snow depths (<1 m), gamma0_VV generally remains relatively constant, with changes typically within +-2 dB"
Page 16, Lines 21-22: "The spatial distribution is however similar to that of gamma0_VH, also showing slight decreases in gamma0_CR in regions with shallow snow and forest cover"
Page 17, Lines 3-7: "It shows that the inclusion of zero snow depths results in higher correlations mainly in regions with shallow and occasional snow (e.g., in eastern Austria). Nevertheless, it is important to remark that the S1 gamma0 observations in these regions only show a weak (if any) correspondence with snow depth, because of the weak scattering contributions from shallow snow, frequent wet snow and melt conditions, and the frequent disappearance and re-appearance of snow cover."
Page 24, Lines 7-14: "The main uncertainties in the S1 snow depth retrievals are expected to be caused by wet, shallow and occasional snow cover and forest cover. Wet snow is known to cause a strong decrease in radar backscatter due to signal absorption. Although a wet snow detection algorithm is implemented, undetected wet snow (for instance due to an insufficient decrease in the backscatter) may cause underestimation in the snow depth retrievals. Uncertainties can also be large in regions with shallow and

occasional snow cover, where the backscatter observations can be dominated by scattering contributions from the ground surface, resulting in a weak (or even negative) correlation with snow depth. For shallow snow conditions, backscatter observations at higher frequencies (e.g., X- or Ku-band), or potentially also using InSAR phase changes at lower (e.g., L- or P-band) frequencies, could be more suitable to detect short-term snow depth changes."

Figures 11 and 12 (Figs. 12 and 13 in the revised manuscript) are derived from exactly the same data as Fig. 10 (now Fig. 11). However, the revised Fig. 12 stratifies the performance metrics by the range in snow depth, thus combining metrics from different sites that reach a similar peak value of snow depth. On the other hand, Fig. 13 stratifies by the actual snow depth, and can thus combine metrics from different time steps at the same site or from different sites. This difference in stratification is mentioned in the caption of Fig. 13.

Thanks very much for considering these comments. My hope is that the addition of a data-focused visualization of the processing chain in an appendix will provide convincing evidence, improved traceability, and increased confidence in this analysis.

Chris Derksen

We thank you for the careful assessment of our work, and sincerely hope that our modifications to the revised manuscript and responses to the review comments comply with the requirements, in order to consider our work for publication in the Cryosphere.

Hans Lievens, on behalf of the co-authors

Reviewer #1

I truly understand the amount of work that was put into this study and the amount of time it takes to reprocess such a large volume of data. I also appreciate the effort and the amount of work and detail that went into this manuscript and revisions.

We thank the reviewer for acknowledging the amount of effort that went into our study over the past 2 years.

Nonetheless, I must recommend this paper be rejected for the same reasons as my initial review. With the different levels of processing, it is impossible for me to clearly say that what is detected is snow depth. A lot of researchers have done similar studies since the early stages of C-band SAR data and many times the relationship found with the signal was not with snow depth but often a contribution from the background signal or the heterogeneity of the snowpack but not it's depth, even in alpine environments.

We regret to see that the reviewer repeated the main previous criticism ("*With the different levels of processing, it is impossible for me to clearly say that what is detected is snow depth*") even though we substantially modified the processing of the backscatter data to address part of the reviewer comments, and despite the proven agreement of our snow depth retrievals with in situ data.

Section 3.1.1. provides a literature overview on previous work using C-band backscatter for snow depth (and SWE) retrieval. As indicated in this section, cross-polarized satellite observations over deep snow have to our best knowledge not been investigated in the past, and certainly not with Sentinel-1 data, providing frequent coverage with a fixed observation scenario. In the previous review comments, the reviewer stated that our literature overview did not apply to the current study, because of the generally different conditions being investigated in past studies (shallow snow often in tundra/taiga environments studied with co-polarized backscatter data). We are therefore somewhat surprised by this contradictory statement ("*A lot of researchers have done similar studies since the early stages of C-band SAR data and many times the relationship found with the signal was not with snow depth but often a contribution from the background* *signal or the heterogeneity of the snowpack but not it's depth, even in alpine environments.*"). We invite the reviewer to provide references to studies that are using cross-polarized C-band backscatter over deep alpine snow to support this statement, which we will gladly integrate into the literature overview.

As for smoothing steps in the processing:

1- Multi-looking is a spatial filtering that yes removes speckle but also smoothes out the images spatially and removes a lot of the high local variability of alpine landscapes.

Multi-looking reduces the grid size by combining all pixels within a certain window into one single (average) value. We are determined to keep the multi-looking as part of the processing chain for the important reason that the alternative (estimating snow depth at the original 10 m pixel spacing) is currently infeasible in terms of computation time and storage requirements.

2- The reprojection steps adds even more spatial smoothing which removes even more spatial variability in high topographic terrain. These two steps are very well known and documented issues in alpine environments.

It is impossible to apply time series analysis and to provide meaningful geospatial information without projecting the data onto a consistent coordinate system.

3- The orbit correction is a temporal filtering. It is known that there are diurnal cycles in the snow conditions especially in alpine snow (melt-refreeze for one). The authors mention in their manuscript that equation 1 uses the "temporal" mean and standard deviation.

The orbit correction is again not 'filtering' nor 'smoothing', but a bias correction applied to the first two order moments of the time series, i.e., the mean and the standard deviation. The temporal mean and standard deviation are taken over the complete time series (and not over a shorter window, which would imply smoothing).

However, to address the reviewer's lack of support for the orbit bias correction, we modified the algorithm in two ways: (i) the orbit bias correction is no longer applied in the backscatter processing, and (ii) the change detection algorithm is adjusted by tracing the differences in backscatter only between acquisitions from the same relative orbit (which are typically 6 or 12 days, or multiples of that, apart). The difference in backscatter is then be imposed onto a weighted average snow index corresponding with the previous observation (6 or 12 days ago). The latter is calculated by applying inverse distance weighting over a certain time window to also incorporate the snow index estimates from other orbits. This modification also allows to retrieve snow depth for every time step when a backscatter acquisition is available (to address point 4 below). For more information, please refer to the revised Section 2.2.

Also, the different orbits and viewing geometries will have different travel paths in the snowpack which the authors neglect to mention has a major impact on the signal intensity and thus smooths out even more the influence of the snowpack properties on the SAR signal.

We agree that the incidence angle impacts the travel path length of the signal through the snow, and that we should have mentioned this impact in the manuscript. We have investigated the impact of the local incidence angle on the sensitivity of backscatter to snow depth, and found that the sensitivity mainly reduces for angles of 60°-70° or higher. In the reprocessed retrievals, we have masked out backscatter observations with a local incidence angle above 70°. Note that a more stringent masking (excluding also lower local incidence angles) would reduce the spatial coverage of the retrievals.

We have modified Fig. 3 (in the revised version) to show the backscatter time series separately for the different orbits. For the sites in this example, similar sensitivities of backscatter to snow depth are observed for the different orbits in cross-polarization. Stronger differences are found in co-polarization. We are interested to further investigate the impact of the incidence angle on the retrieval approach in future research.

The following statements have been added to the manuscript:

Page 4, Lines 19-20: "Observations with a local incidence angle >70° were excluded to reduce radar shadowing effects."
Page 12, Lines 2-5: "The gamma0 observations in Austria are shown separately for three selected orbits, i.e., descending orbit 95 (D95), ascending orbit 117 (A117) and D168, corresponding with local incidence angles of 26.5°, 61.9° and 17.8°, respectively. The gamma0 observations in Switzerland are shown for D66 (55.8°), A88 (32.9°) and D139 (46.9°).
Page 12, Lines 7-9: "Limited differences are observed between the orbits, which can be explained by the different local incidence angles (also impacting travel path lengths through the snow), azimuth angles, and overpass times (6 am for D, 6 pm for A)."
Page 12, Lines 18-19: "while also larger differences in backscatter are observed between orbits."

4- The retrieved values are weekly values. This smoothes out all the snow and SAR signal variability over the entire week which can be a lot in alpine conditions (several cm to metes of fresh snow, sublimation, wind compaction, blowing snow, etc.).

First, we do not see any issue with producing weekly average snow depth retrievals as derived from weekly average backscatter. Note that we chose a weekly temporal resolution to (i) indeed average out short-term variability and (ii) to reduce the processing time and storage requirements by computing outputs only once a week instead of daily.

However, we acknowledge that weekly average retrievals indeed no longer contain information on sub-weekly variability caused by processes such as those mentioned by the reviewer. Therefore, we modified the algorithm to calculate the snow depth retrievals at the time of the backscatter observations. We now produced output at the daily time step with coverage according to the Sentinel-1 swaths (instead of the entire Alps being consistently covered in the weekly averages) and validated the results with model simulation and in situ measurements at the daily instead of the weekly timescale. All associated figures (i.e., Figs. 5-13) have been updated accordingly.

For these reasons, it is not possible for me to recommend this paper for publication.

We are thankful to the reviewer for the detailed assessment of our work, and for providing several pathways to improve the retrieval algorithm and the quality of the manuscript. However, in our opinion the comments above do not fully justify the final recommendation of the reviewer. Two out of the 4 processing steps that are criticised (multi-looking and georeferencing) are standard practice, are indispensable, and are by no means impacting the time series sensitivity of the data to geophysical parameters, which is at the basis of our change detection method. To the contrary, these are both necessary steps to keep the processing feasible and to provide meaningful geospatial information.

While we are also convinced that the other 2 criticised processing steps (the orbit bias correction and the weekly temporal resolution of the retrievals) were by no means impacting the main conclusions of the work, we have modified the algorithm to comply with the review comments, by (i) excluding the orbit correction and (ii) producing retrievals at the time of the satellite acquisitions. As the reviewer rightly pointed out, the latter can potentially improve the capturing of sub-weekly processes.

With the above-described modifications to the backscatter processing chain and to the retrieval algorithm, and with the responses in this letter, we are hopeful that the reviewer can settle with our proposed methodology. We would find it difficult to conceive if the remaining discordances on standard processing steps (such as multi-looking and projection onto a coordinate grid) would prevent our study from being published and would thereby also prevent a novel dataset with demonstrated accuracy from being accompanied by peer-reviewed literature support and documentation.

Reviewer #2

The author have responded adequately to my comments.

We would like to express our thanks for taking the time to review our paper and for the constructive and useful feedback.