Review

Improved workflow for customized ICESat-2 ATL06 elevations captures seasonal mountain snow depths at sub-kilometer scale

The paper explores how ICESat-2 satellite data can be used to estimate mountain snow depth more accurately by comparing satellite elevations with high-resolution snow-free terrain models. The authors show that with careful processing—such as reducing positioning errors and adjusting for terrain effects—ICESat-2 measurements can closely match ground and airborne observations. The study finds that the satellite performs best in areas with moderate slopes and deeper seasonal snow, and suggests that many mountain regions have conditions suitable for reliable ICESat-2 snow-depth observations. This approach could broaden the use of ICESat-2 for monitoring snowpack and supporting water-resource modeling.

General comments:

I find that the paper is well written and that thorough analysis has been performed, identifying limitations in using ICESat-2 for generating snow depth in mountainous areas, while also showing where it can be used. However, I would like to see more descriptions of the general way the processing is done, especially the generation of the hybrid ATL06 product. That description is currently lacking in my view but can easily be fixed. It would also be of interest to include surface classification directly from the number of return photons inside each segment. That would avoid, in my opinion, fully relying on imagery as I understand it, and instead use the inherent physics of the measurements to suppelemt the analysis.

We thank the reviewer for their detailed and constructive comments, we are glad to improve the work herein. We plan to revise and expand our description of the hybrid ATL06_SR product and to include a direct pointer to Besso et al. 2024 and Shean et al. 2025 where the product is described in greater detail, please see our response to L98 for more detail. We additionally performed new smoothing length and auto-correlation analysis to give more depth to our smoothing length discussion. We also propose to add additional detail on the return photons to better explain our decision to use the NSDI imagery as opposed to n_fit_photon as we initially attempted. Please see our detailed responses to the individual comments for the specific changes we propose to make.

Karina Zikan

Line-by-line comments:

L53: Should this not be 17 m instead of 11 m?

We chose to use the ICESat-2 mean effective laser footprint diameter of 10.9 m ± 1.2 m rounded to 11 m found by Magruder et al. 2021 instead of the 17 m diameter estimate. For clarity, we will remove reference to the footprint diameter from L53 and expand L70 to explain the use of the 11 m mean effective laser footprint diameter. Please see the L70 comment for the specific proposed rewrite.

L56: “Comparing ICESat-2 data to an independently collected snow-free DTM introduces additional geolocation errors.” Can you state more specifically what you mean and why?

To add more explanation for the geolocation errors we will rewrite L56 as follows,

Current L56 text: “Additionally, comparing ICESat-2 data to an independently collected snow-free DTM introduces additional geolocation errors (Enderlin et al., 2022; Hugonnet et al., 2022; Nuth and Kääb, 2011).”

Proposed rewrite: “Additionally, if the ICESat-2 data and independently collected snow-free DTM are not properly aligned geospatially the resulting geolocation offset between the two datasets will introduce geolocation errors. The magnitude of these geolocation errors depends on the slope and aspect of the underlying terrain relative to the direction of the geolocation offset (Enderlin et al., 2022; Hugonnet et al., 2022; Nuth and Kääb, 2011).”

L70: 17 m or 11 m?

To clarify the use of the 11 m mean effective laser footprint diameter we will rewrite L70 as follows,

Current L70 text: “Each pair of beams has a beam footprint of ~11m (Magruder et al., 2021), an intra-pair separation of 90 m, and an inter-pair separation of 3.3 km (Neumann et al., 2019)”

Proposed rewrite: “Each pair of beams has an intra-pair separation of 90 m, and an inter-pair separation of 3.3 km (Neumann et al., 2019). After launch, the mean effective laser footprint diameter of ATLAS was found to be ~11 m (Magruder et al., 2021)”

L73: “ICESat-2 returns have a geolocation uncertainty of ~4.4 m.” Add the error, which is ±6 m, and the fundamental product you are referring to.

We will add the error and the ATL03 product we are referring to.

L98: I would like some more details of the hybrid data product, as this is important for the study. I think at least a paragraph or two should be dedicated to that purpose to explain how the data is generated.

We discuss the generation of ATL06_SR in more detail in section 3.1. We will add a pointer to section 3.1 to L98 and provide more detail on the correction and filters used to the ATL06_SR paragraph in 3.1.

Proposed L98 changes: “In this study we make use of the strengths of both algorithms using a hybridized data product (ATL06_SR) (Besso et al., 2024, Fair et al., 2025) that incorporates ATL08 vegetation filtering and the ATL06 algorithm into an ATL06-like product. ATL06_SR is calculated by applying the ATL06 function to ATL08-identified ATL03 ground photon returns instead of the ATL03-identified ground photon returns. The generation of ATL06_SR is discussed in more detail in section 3.1. As applied in this paper, ATL06_SR includes ATL08’s vegetation filtering but does not include the first photon bias correction, which can result in up to ~2 cm of bias, or the transit pulse shape bias, which can result in up to ~1 cm of bias.”

Proposed 3.1 (L189-192) changes: “We use the ATL06_SR product for all available ICESat-2 data acquired from October 2018 to April 2024 within the boundaries of the four study sites. We refer the reader to Besso et al., 2024 and Shean et al. 2025 for a detailed description of the ATL06_SR product used herein. Briefly, to calculate ATL06_SR, we applied the ATL06 function to ATL08 ground-classified ATL03 photons (as in Besso et al., 2024). We calculated ATL06_SR using the SlideRule Earth data processing package which allows for rapid, cloud-based processing of the ATL03 photon cloud with customized control of the ATL06 algorithm parameters (Shean et al., 2025). For this study ATL06_SR was calculated using ATL08 ground-classified ATL03 photons using otherwise default ATL06 parameters: a 40 m segment length, a step size of 20 m, a minimum along-track spread of 20 m, a maximum of 6 iterations, and a minimum of 10 ATL08 ground classified photons. The resulting ATL06_SR product therefore has an elevation estimate every 20 m.”

L176: “The snow-free ICESat-2 height residuals, h_residual, are the difference between ICESat-2 and DTM ground elevations when and where snow was not observed in nearcoincident satellite imagery.” How were the snow-free conditions determined from the satellite imagery?

We will edit L176 to mention the NDSI filtering. More detail on the NDSI filtering is in L204-214.

Current text: “The snow-free ICESat-2 height residuals, h_residual, are the difference between ICESat-2 and DTM ground elevations when and where snow was not observed in near-coincident satellite imagery.”

Suggested rewrite: “The snow-free ICESat-2 height residuals, h_residual, are the difference between ICESat-2 and DTM ground elevations in snow-free conditions as determined by the Normalized Difference Snow Index (NDSI) calculated from in near-coincident satellite imagery.”

L79: ~11 or 17 m?

~11 m

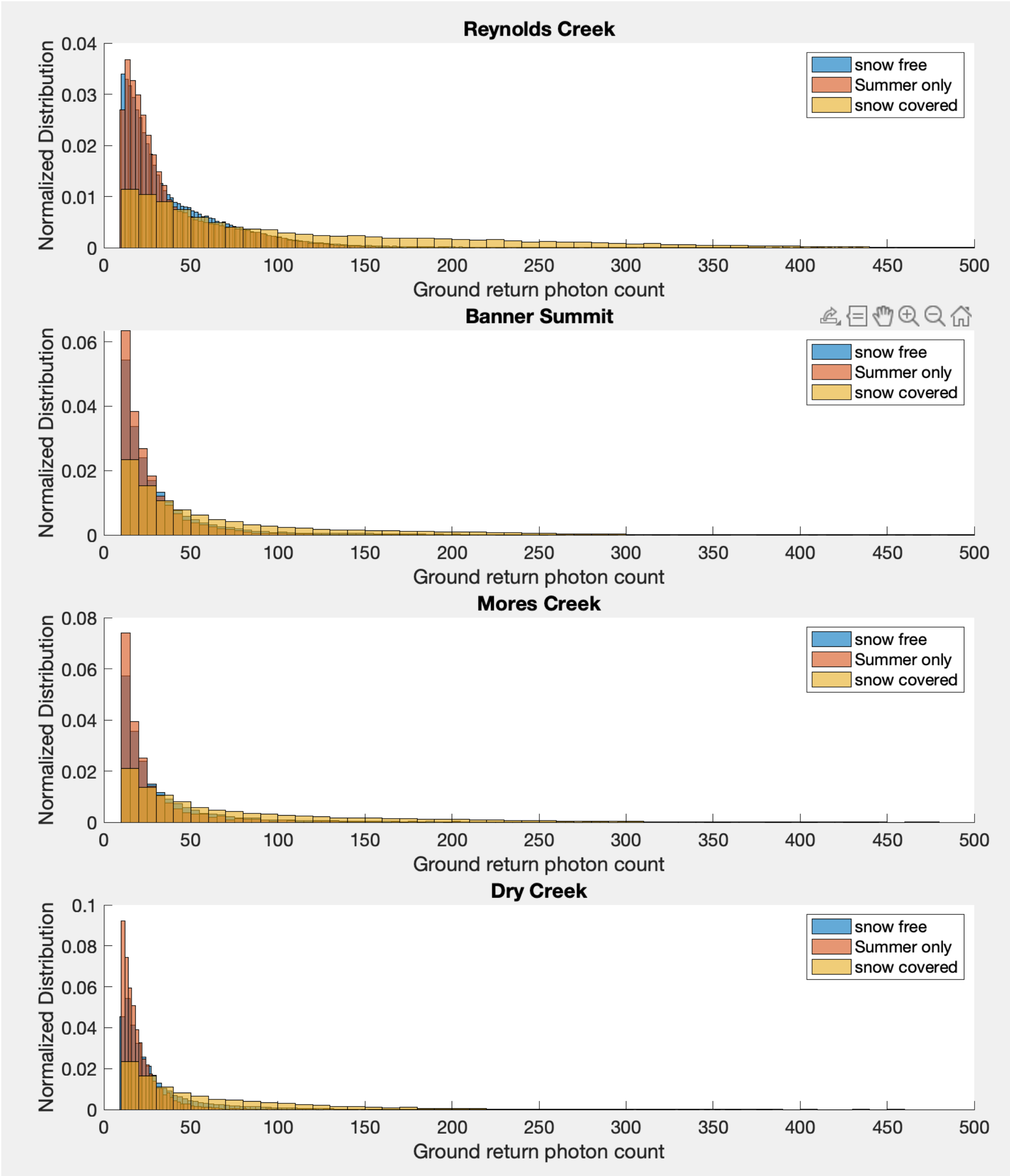
L181: Why is the “n_fit_photon” not used to calculate when you have snow or snow-free conditions, or used in combination with the imagery? The classification will be quite clear, as the number of return photons can be used to easily separate the two types of returns.

We initially tried to use the n_fit_photon to identify snow cover, however since the number of photon returns is greatly reduced by both vegetation cover and increased slope we found the n_fit_photon was not reliable for separating snow covered and snow free terrain. We chose to use the near-coincident NDSI maps as an independent snow mask unaffected by the terrain metrics we wanted to investigate.

We propose to add the following text after L182 to expand on the choice to use the NDSI maps rather than the n_fit_photon.

"Due to the effect of slope and vegetation cover, which both reduce the number of ground return photons (n_fit_photon), snow-free and snow-covered conditions could not be clearly distinguished using n_fit_photon at the sites studied. A difference in n_fit_photon distribution was observed when slopes were < 10° however we chose to use the independent near coincident NDSI maps to identify snow covered terrain so as not to impact the terrain related controls."

As you can see from these histograms of photon returns in snow-free, summer, and snow-covered conditions while there is a difference in the distribution of photon returns, there is too much overlap between surface cover to cleanly differentiate between surface conditions.



L192: Can you provide some more justification for why "h_mean" is used and not "h_li," for the reader to get a better grasp of why it's important to use it?

The choice of h_mean was a practical one as Sliderule only calculates h_mean, not h_li. We will add more detail to clarify the use of h_mean.

L206: Same question as before—can you use the photon count for each segment to determine snow-free conditions?

Please see our detailed response above.

Figure 2: The text in the figure is very small, so I suggest increasing the font size to make it more visible.

We will increase figure font sizes.

L224: Can you mention the methods that were tested, so the reader does not need to go into the supplement?

We will add the following text to list the other co-registration algorithms tested,

Proposed rewrite of L224: "In addition to the iterative grid search two other co-registration algorithms were tested and rejected during method development, the Nuth and Kääb (2011) co-registration approach (as used in Deschamps-Berger et al. 2023), and a gradient descent (as used in Enderlin et al. 2022). These are discussed further in Appendix A."

L226–L234: Are these co-registrations different from the ones in the appendix?

L226-234 and the appendix are discussing different aspects of the co-registration. L226-234 is focused on how the data is input into the co-registration algorithm (either aggregated or as individual tracks) while the appendix is focused on other co-registration algorithms we tested in addition to the iterative grid approach (L214-224). Currently we are referring to both parts of the coregistration as the "coregistration approach" which is confusing, to fix this we will refer to the coregistration algorithm as the "co-registration algorithm" and the data input into the coregistration algorithm as the "co-registration approach".

L276: "Which is more than double the expected precision (4.4 m) of ICESat-2 geolocation." The estimated standard deviation of the error is, however, 6 m, which would still fit within the 1-sigma error. I would not expect you to find an expected precision of 4.4 m, especially in regions of steep terrain.

Yes that's a good point, we will remove this statement. On reflection, since our goal of this line is to highlight that the individual coregistration finds larger offsets then the aggregated coregistration we will report the variability in the offset (interquartile range of 3.9 m) rather than the maximum shift offset.

L284: "We find that ICESat-2 snow depth has a negative bias of ~0.6 m and uncertainty of ~1 m regardless of co-registration approach." So, is there a need to apply the coregistration if these biases still exist?

We agree that there is a reasonable argument to be made that, at least with the co-registration methods we tested, co-registration does not improve ICESat-2 results, or not to an extent that it is worth the time and computation effort. We believe it is still important to report and discuss these results to inform future research and hopefully save future researchers some time. We will add the following text to the discussion in section 5.1 after L353,

"Co-registration remains an unsolved problem. Regardless of co-registration approach, ICESat-2 snow depth maintained a negative bias of ~0.6 m and uncertainty of ~1 m. There is likely a limit to the improvement possible from horizontal co-registration. While horizontal co-registration should not be inherently dismissed because there can be systematic offsets between ICESat-2 and the reference DTM depending on the georeferencing of the reference DTM, the time and computation effort required to perform the horizontal co-registration should be weighed against potential improvements."

L305: I would highly suggest that you perform a simple correlation-length analysis of the differences to get an idea of what the optimal comparison radius would be. That would better inform the maximum distance at which you can calculate statistics. Or at least provide a figure of the statistics as a function of your smoothing length (100 m, 500 m, 1000 m, and 5000 m). The optimal smoothing length would most likely be correlated with the average slope magnitude at each site.

We propose to replace table 4 with a plot of RMSE and R2 by smoothing length and movie table 4 to the supplement. We will update the text regarding smoothing length at each site based on this new figure.

Additionally we plan to add a figure plotting ICESat-2 snow depth NMAD and R2 compared to the airborne lidar snow depth by smoothing length. Will add the following text to the results around L329:

"Comparing ICESat-2 snow depths against Mores Creek airborne lidar snow depth shows that when comparing data with the same spatial coverage, ICESat-2 snow depth uncertainty and correlation both improve with smoothing length. The ICESat-2 uncertainty drops ~0.25 m across all smoothing lengths while R2 rises from ~0.48 to ~0.60. R2 is above 0.5 for all smoothing lengths > 300 m."

Comparison of MSC ICESat-2 snow depth and Helicopter lidar draft figure:

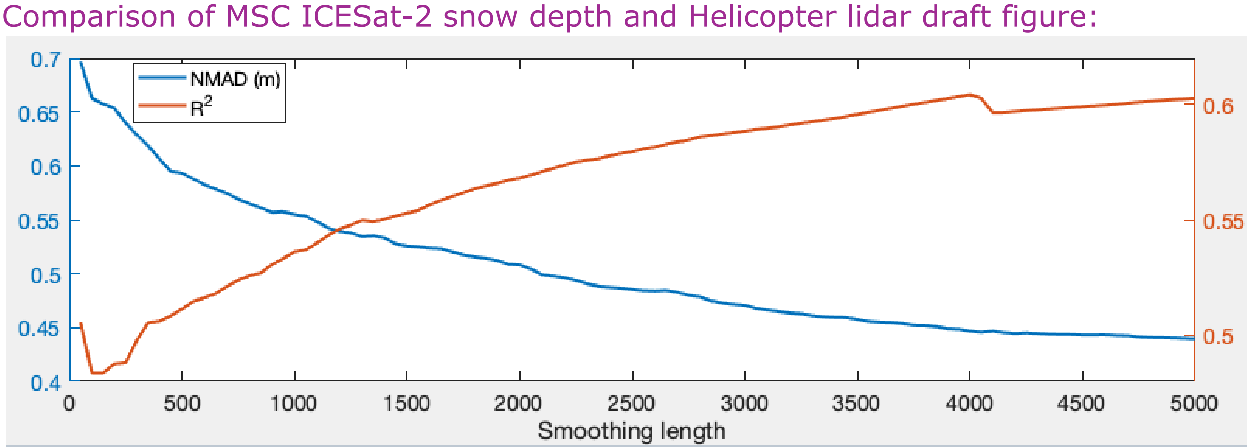
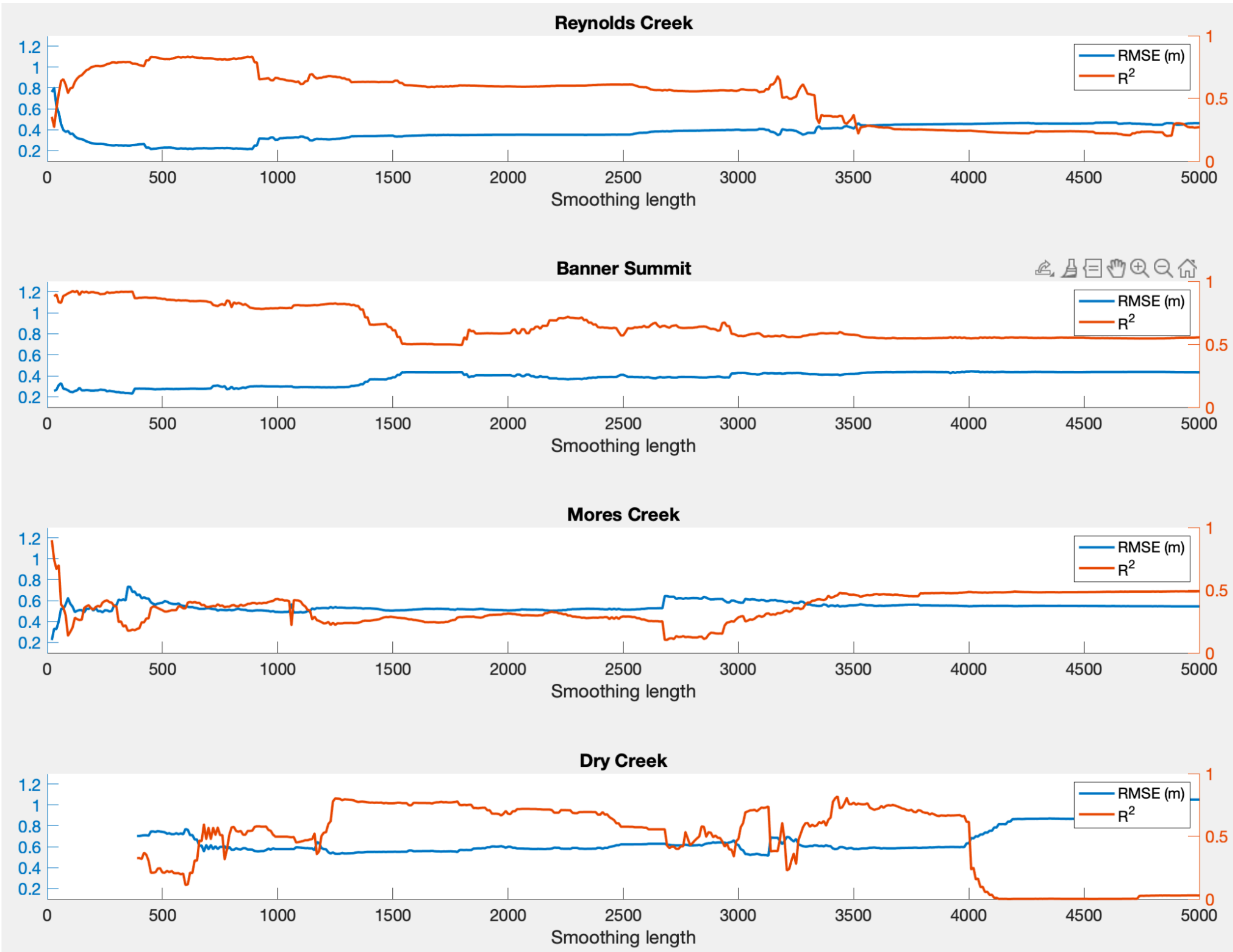


Table 4 figure replacement draft:



L360: Could it also be related to the fact that applying time-variant co-registration reduces the number of samples available and biases the dataset toward specific slope/topographical regions, increasing the noise in the registration? Maybe looking at the number of return photons can help reduce this issue by reducing the impact of mixed surface types where snow and snow-free terrain overlap.

It is definitely possible that the individual co-registrations are biased because they fall in areas with certain attributes that are not representative of the broader area. We will add text to L359-263 to highlight this point. Additions in bold.

Proposed rewrite: "Most concerningly, the application of a time-variant co-registration transform resulted in no correlation between ICESat-2 snow depths and precise independent snow depth estimates (Table 2). **The high variability in time-variant co-registration of sets may be due to the limited spatial coverage of an individual overpass. If a given overpass happens to fall over a highly sloped or densely vegetated area the increased uncertainty or bias will impact the accuracy of the co-registration.** The poor performance of individual co-registration transforms **in the winter** is likely **also** due to sparse snow-free winter terrain. Snow coverage obscures stable terrain and when <10% of the region of interest is stable the accuracy of co-registration decreases with the percent of stable terrain (Nuth and Kääb, 2011)."

L416: How large are these negative values? To reduce the risk of biasing the snow depth when removing SD < 0, could you allow for smaller negative values to be kept, perhaps within some limit or error?

The median negative snow depth is ~ -0.7 m with an interquartile range of ~1.1 m. We tested setting the negative threshold at -0.3 m instead of 0 m based on the median bias calculated from comparing ICESat-2 and airborne derived snow depths at Mores Creek, however we observed an increase in uncertainty and bias in the ICESat-2 snow depth data. We propose to add the following text after L416 to expand on this.

Proposed addition: "The overestimation of shallow snow depths could be mitigated by lowering the minimum snow depth threshold below 0 m based on ICESat-2 snow depth uncertainty, however this also increases the impact of outliers on deeper snow depth estimates."

L430: I think grouping them into elevation zones rather than horizontal distance bins would be a more effective approach, as you will increase data density. That's why I suggested calculating the spatial autocorrelation: you can use that to first get all data within that distance and then group them in elevation bands.

We agree that grouping data into elevation zones is likely a more effective approach as we can include a larger amount of data while maintaining a relatively high spatial resolution. In line with what we expect from forested mountain environments given that snow depth correlation length is typically much shorter than the 20 m ATL06_SR segment length in such environments (Trujillo et al., 2009), the auto correlation is greatest without lag and falls away precipitously as the spatial lag increases. Thus the smoothing scale must be a compromise between keeping a high spatial resolution and including sufficient data. We propose to add the following after L431 to expand the discussion of elevation zones:

"Grouping the ICESat-2 snow depth data into elevation zones may be more effective for characterizing snow depths across a landscape than directly calculating spatial distribution as it maintains a high spatial resolution while averaging over many data points to reduce variability. This assumes that the primary control on snow depth is elevation; generally this applies (fig 7), however slope and vegetation density varies greatly across this terrain therefore there can be large variation in snow depth at a given elevation. Grouping the data by elevation zones may obscure other terrain controls on snow depth. However terrain controls on snow depth will also be obscured by estimating snow depth at the larger spatial smoothing scales required to achieve similar data density."

Trujillo, E., Ramírez, J. A., and Elder, K. J.: Scaling properties and spatial organization of snow depth fields in sub-alpine forest and alpine tundra, Hydrol. Process., 23, 1575–1590, https://doi.org/10.1002/hyp.7270, 2009.