Response to Referee comments

Submission Title: *Assessing the Seasonal Evolution of Snow Depth Spatial Variability and Scaling in Complex Mountain Terrain*

Submission Number: *tc-2022-96*

Thank you again for your time providing an additional review of this manuscript. We appreciate your feedback on the manuscript and the relevance you find in the work. We addressed each comment and provided an updated version of the manuscript as well as responses to each comment below.

Reviewer 1 (Yves Bühler) Suggestions:

Dear Authors

Most of my points are satisfactorily answered by your revision. However the most critical methodological point is not.

You argue that you use the nearest neighbor resampling to "not over-smooth" the snow depth values. But resampling from high (2 cm) to low (0.5 - 20 m) with nearest neighbor is simply wrong and has assumingly a very large impact on the results and conclusions of your study.

To compare remotely sensed snow depth values at different spatial resolution you need to aggregate the HS values within the area of the larger resolution cell. Applying nearest neighbor resampling just takes the exact value of the high spatial resolution cell that is closest center point of the lower spatial resolution cell. And this is clearly wrong.

The real snow depth within a lower resolution cell is very different from the center value in particular within complex terrain (which you investigate). The deviations from the real snow depth can easily reach more than 100 %. The center value is very random if your snow depth distribution is not perfectly homogenous.

Due to this methodological flaw which probably has a very large impact on the results unfortunately, I cannot recommend the paper for publication.

Response to Reviewer 1 Suggestions:

We appreciate your perspectives and comments regarding resampling methods and completed further analysis of the effects of resampling method on spatial variability results. In addition to using the nearest neighbor (NN) method in our original analysis, we resampled our 2cm resolution DSMs to each resolution using the cubic convolution (CC) technique and completed variogram calculations on the resultant DSMs. We found that similar patterns for spatial variability exist in the experimental variograms using both NN and CC resampling methods. However, the total semivariance is consistently lower in the complex terrain of the Hourglass (Fig. A3). These subtle differences in the experimental variograms have less subtle effects on the spherically fit models which, in some cases, miss the initial short-range-distance sill (15 ~ 20 m) and instead fit to a much larger range (~ 60 m) with the accompanying increase in sill value to match (Fig. A4, A5, and A6). This reinforces our idea that the cubic convolution resampling methods over-smooths and decreases the naturally occurring variability observed in the dataset of the complex heterogenous terrain of the Hourglass. Alternatively, results are nearly identical in the homogenous terrain of the meadow when using NN and CC techniques.

To further assess whether the differences in variability using different resampling techniques are attributed to over-smoothing using CC or artificially inflated using NN (as suggested by Reviewer 1) we resampled the DSMs using NN, CC, and mean and median aggregation methods. We then completed a pairwise correlation analysis of paired points of each method to observe any systematic bias in the resampling method. We found that all of the resampling methods result in highly correlated (> 0.97) point pairs for all spatial resolutions < 1 m. Correlation values decrease markedly as the resampled resolution increases beyond 1m between nearest neighbor and all other methods decreases (Fig. A2). Therefore, resampling methods should be considered closely when resampling to resolutions greater than 1m. Additionally, while the nearest neighbor method utilizes only observed values, the cubic convolution method calculates derived values and our results suggest this method produced unreal values in highly variable areas (ex. along an avalanche crown on Feb 27).

Our original results suggest spatial resolutions < 0.5m are necessary to observe the complete picture of spatial variability at our site. Given the high correlation of paired values across all resampling methods at resolutions < 1 m, we show that the nearest neighbor technique does not artificially inflate the variance and captures the naturally occurring spatial variability of snow depth within the high-resolution context of our study. Overall, through our original analysis and this additional analysis, we are able to identify a resolution (1m) at which results start to diverge based on resampling technique, identify a resolution (50cm) at which spatial variability can be captured independent of resampling technique in the complex (and homogenous) terrain of our study site, and, finally, identify the distance at which snow depth differs across the complex mountain terrain of our study site(15 m).

We've included additional text, citations and figures regarding the resampling re-analysis in these locations of the manuscript:

- Sec 3.5: Variogram Calculation and Fit - lines 239 – 246
- Sec 4.2: Resampling Results – lines 313 – 323 (new section)
- Sec 4.3: Variogram Results – lines 326 – 331
- Sec 5.2: DSM Resampling Methods – lines 424 – 445 (new section)
- Sec 6: Conclusions – lines 524 – 525
- Fig A2 – A6 (additional figures comparing resampling results)

Reviewer 2 (Anonymous) Suggestions:

Review of the manuscript "Assessing the Seasonal Evolution of Snow Depth Spatial Variability and Scaling in Complex Mountain Terrain"

I must congratulate manuscript authors in view to the changes included in the manuscript after the major review. They have conveniently discussed all points raised by reviewers, changing the manuscript where appropriate or justifying their choices. However I have some minor points that I would like to point out:

1. Figure 3 maps: I encourage manuscript authors to include classified colors scales. The colors for snow depth and elevation can be used, but continuous color scales are difficult to interpret. 7 to 9 snow depth classes can help to more easily see the snow depth variability. For instance Figure A3 scale bar has more resolution than that of Figure 3. If possible make thicker (or darker) contour lines in Figure 3.

2. I agree that, spherical models, are commonly used in variogram analysis based on the references you provide. However there are other models with great potential. You might cite other models used in snow science with good result (log linear, exponential…) and state that you have finally used spherical. In this regard I think the comparison between experimental variograms and fitted variograms is needed. Indeed you state you have compared them (Line 304-305), but no adjustment metrics are shown (R2 or others….) or at least plot them together (in Figure A2 include right below experimental variograms the fitted ones). Please include one of these previous suggestions.

3. Manuscript authors have chosen a 1/3 length of the diagonal box and they support their decision on R "gstat" package URL documentation. This is a bit difficult to find in this documentation. Can you please help potential readers to find this with other references? This 1/3 choice deserves some discussion since, as far as I know, in snow science 1/2 distance is usually applied. Just discuss if some differences could be found if a different maximum distance is applied to semivariogram computation in view to previous works.

4. Figure 8: Circles superposition makes difficult the interpretation, Might you reduce circles size and allow some transparency?

5. Conclusions: I think it is worth to change a bit conclusion section, stating (line 471): "Despite for more than one half of our UAV acquisitions 1m sample spacing is able to capture the natural snow spatial variability, to guarantee the full capture of snow depth variability for all observation dates, 0.5 m sample spacing is required….." or a rephrase of this sentence

6. In previous figure 6 comment, there was a misunderstanding. Previous Figure 6 (now figure 7) was fine. I meant that a legend with circles color correspondence can help to interpret it (as included in figure 6 of final manuscript version).

Response to Reviewer 2 Suggestions:

Thank you for a second thorough review, and we appreciate your sentiments. We addressed your comments as follows:

1. Figure 3 maps have been updated to include discrete classified color scales and clearer elevation contours.
2. An additional comment and citation have been added regarding other variogram model fits (lines 236 – 237) and a table of RMSE and NMAD values for the spherical fits has been included in the appendix (Table A5). Comparative model fitting lies beyond the scope of the paper largely due to the additional extensive computational requirements of additional fits.

3. We have added additional discussion of these choices and citations of other works (lines 250 – 255).
4. Figure 8 has been updated with smaller and transparent points jittered around the shaded collection dates for improved interpretation.
5. We state this distinction in the results and discussion and have left it out of the Conclusion for brevity and readability.
6. Figure 7 has been updated with colored points and legend.