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 for your time providing a thorough 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) Main Points:

1. The applied nearest neighbor resampling technique is in my opinion the wrong approach to resample the different snow depth maps. If nearest neighbor is taken, the value of the coarser resolution grid is the value that is located closest to the center of the new cell. If you go to coarser resolutions (0.5 – 20 m) this makes no sense as this value can be very random. I would propose an aggregation or a cubic convolution resampling.

> We chose the nearest neighbor resampling technique to avoid over-smoothing the dataset and to maintain the naturally occurring variability found in spatially continuous snow depth. Our initial results used an aggregation method for resampling, which we found to over-smooth the inherent variability in snow depth, especially at < 1 m scales, resulting in unrealistically smooth experimental variograms with less defined ranges. Additionally, we did not want to add any additional abstraction, and therefore uncertainty, to the detrended snow depth values used in our analysis. Cubic convolution techniques may result in cell values outside the range of the input raster, especially near edges, which our vegetation masked dataset has many of and we did not want to introduce this uncertainty into our spatial variability analysis. We included additional clarification and supporting citations to the manuscript on lines 234-239.

2. There are no figures illustrating the snow depth maps or the applied detrending. It would be important for the readers to see such figures here to better understand what is done.

> We added a figure clarifying the vegetation masking and detrending steps (Figure 3) as well as figures of the timeseries of detrended snow depth maps at the two locations to the manuscript appendix (Figures A3 and A4).

3. The motivation, why semivariograms are used, is not really clear. Are there other possible methods? If yes, a comparison of the results from other methods would be very interesting.

> We set out to try to assess the appropriate scale for measuring snow depth in complex terrain. We were motivated to observe the spatial structure at different resolutions that allows us to reliably interpolate between data points while capturing the naturally occurring variability of snow depth. Variograms are the generally accepted tool for assessing the scales of variability with spatial data and have been widely used in the snow depth spatial variability literature. Therefore, we chose this method because it is well understood and makes our results comparable to previous work. We added this justification and relevant citations to the manuscript on lines 219-233.

4. The discussion is very much based on hard-to-understand metrics (Sill, Range etc.). I am missing a part where the discussion is on a level where the average reader can follow. What do these values mean discussed on examples, best illustrated with figures showing the snow depth distribution.

We added definitions of these values into the methods section (Sec. 3.5, lines 224-231).

5. The investigated site is very small and we do not know how representative this is. It is not clear if the findings that are presented are valid for further regions. We would have drone-based snow depth maps from sites in the region of Davos, Switzerland (also several dates in one winter) we could provide to check if the results are consistent in different regions.

We agree that complex terrain is challenging to effectively prove as broadly representative and have clarified our language throughout the paper to specify the complex terrain within our study site (lines 20, 89, 396, 407, 428, 456, 469, 472). It would be very relevant to see if our results are similar to those at other sites (such as Davos) and we would be interested in pursuing such collaborative research in the future. Addressing necessary computational processing power would be a key component of adding additional sites in future research, given how computationally expensive high-resolution large scale variogram processing is.