Dear Referee,

We would like to thank you for taking the time to review our paper and for all your constructive suggestions, which will help to improve the quality of the paper. For now, we would like to answer to your major comments (and some larger specific comments). Our response to the comments appears in *italic*. We will take the remaining detailed comments into account when preparing a revised version.

1. **Literature context**

   I get the impression that the authors may be newer to the snow field and so some of the statements in the introduction (see comments) need revising. Broadly there is a tremendous amount of research on processed based- energy balance snowmelt modelling that needs a clearer summary/context for this work. In terms of local-scale advection work there is a much more limited amount but can be found back to the 1970's. This needs a more complete treatment to solidify the context of this contribution as well as to distinguish these contributions from previous one. (See comments for examples).

   We will include a more elaborate discussion of the amount of research done regarding this topic. Thank you for your suggestions.

----

2. **Validation with SFM**

   I'm not exactly clear on the process by which the snow depth difference are calculated via SfM on the observed patch in order to compute expected ranges of turbulent flux contributions and model validation. Are we only considering snow depth difference for grid cells that were completely uncovered during melt (so we can have a bare ground reference) or are we also considering cells that were not fully uncovered that would have also decreased in snow elevation due to melt? What areas do these number represent? Do you observed a decrease in melt from leading edge? Many dynamics can be examined with a spatial dataset but I'm not clear on how this is also processed/what it represents and so would welcome a lot of clarification and perhaps a figure to describe this process.

   We agree that the text describing the SfM, and especially the post-processing of the DEM and orthoimages can be better formulated. Below we added a short description on how we used the DEMs and orthoimages, which we will also include in the revised version of the manuscript.

   We have 2 types of grids (DEM and orthoimages) per day (5 days) per location (upwind & downwind). So, we have 20 grids in total. The size of the grid cells is 0.04 m x 0.04 m.
   For both locations, the following post-processing is done after obtaining the grids:

   1. *Remove isolated groups of cells which are smaller than 0.05 m^2 (All grids)*
   2. *A median filter of 5 × 5 pixels is applied to diminish the influence of noise located within the areas of interest, but maintain the sharp transitions between snow and snow free surfaces (All grids)*
   3. *Compute the median height of bare ground cells per day. The conditions for the selecting the cells used during this computation:*
      a. *Bare ground on first day*
      b. *Covered by all grids*
   4. *Compute correction heights (Table 3) through comparing the daily median heights of the bare ground (step 3) with the median height of first day*
      E.g. median height bare ground cells of day 2 – median height bare ground cells of day 1 = correction value found in Table 3

5. *Remove bare ground cells out of DEM based on orthoimages of the same day → snow covered cells remain*
6. *Apply correction height (step 4) on snow covered DEM (step 5)*
7. *For snow covered grid cells that are present on each day (based on step 5), we calculate height differences between the DEMs of first day and other days (both obtained in step 6)*

The resulting height differences over time correspond to 6.7 m^2 and 30.7 m^2 for respectively the upwind and downwind edge. This might seem in contrast to what would be expected based on Figure A1 (in which the upwind edge shows a larger coverage). However, we chose to solely use grid cells that are continuously covered by snow and have a recorded height change on each day, to reduce the chance of cells being random scatter. As additional advantage this method does not include cells with relatively shallow snow depths of which the recorded melt could be affected by the presence of the bare ground below the snow and also be in the same order of magnitude as our melt error estimates. Our choice for these filters is supported by the fact that when loosening these filters, the size of the boxplots increases drastically, also to unrealistic values and variations in snow surface height, such as large increases over the course of these 5 days.

We are aware that this has an effect on the number of analyzed grid cells, especially on the upwind edge due to the varying locations of snow covered grid cells or the retreating snow line (Figure A1). For the downwind edge, the approximately constant location of the snow covered grid cells combined with the little retreat at this edge, causes this area to be significantly larger. Even though these resulting areas are relatively small, we are convinced that the obtained height changes obtained are decent estimates, also based on our error estimates.

Unfortunately, as a disadvantage of the size of the upwind area consisting of multiple separate smaller areas, we decided to treat the edge as "point" and not look further into the spatial distribution of the recorded melt (e.g. is there a decay in the melt?). The smaller areas are too far apart to do so.

----

3. Latent heat flux decay

There is an extensive treatment of the sensible heat flux decay with patch length while this is not discussed with respect to latent heat flux. Can this also be included or is there are reason it is not included. The heavily cited Harder 2017 paper suggest that latent heat is also an important contributor or at times compensatory (Harder 2018) and so would be very interested in seeing if some of those dynamics could be captured in this modelling scheme.

MicroHH does allow to include the latent heat flux (e.g. Bonenkamp et al., 2019), however we chose to only consider the sensible heat flux to explore the potential of using DNS for studying this kind of system.

We will elaborate on this when revising our paper, such that our intentions are better formulated.

----

4. Implications

There are some interesting dynamics explained but I'm not exactly clear on how those could be implemented in larger scale snowmelt prediction. There is a scaling relationship articulation for sensible heat over a patch length. Is this considered to be a parameterization that could be used in basin scale snowmelt prediction models.

Indeed, a scaling relationship would make it possible to parametrize these type of processes. However, we are not certain that our relationship is appropriate for implementation, due to some limitations of our methods, such as the use of an idealized system. Future studies would need to look further into this relationship, especially when more data is available. In a revised version, we will better emphasize our concerns about the applicability of the formulated relationship.

Line 138-140: when were these samples taken with respect to the observation interval as snow density is dynamic over melt? Was a snow tube used? Snow pit? How were 100ml samples collected? Did the melt period have a consistently ripe snowpack?

Line 160-165: How deep was the snowpack and do you have any information to say that the snowpack was ripe at the start of the melt. Were the cold content requirements satisfied at the start of the period and so all energy could be assumed to be related to melt.

Line 334-336: It seems SWE and density are being used interchangeably here which is not correct. Can this be cleared up? These are pretty high densities. Any observations from field notes about water saturation or other structural attributes. What was the overall snow depth variability? Can you report the SWE of the snow patch?

The samples were taken by digging a small snow pit and collecting 100 ml samples at 5, 25 and 45 cm below the surface at June 14 (4th day in the field). We are aware that taking these samples only on a single day does not reflect the potentially complex temporal dynamics of the snow density. However, we assume the variations occurring on these temporal scales to be relatively small compared to other uncertainties introduced to our method for computing contribution estimates of the turbulent heat fluxes to the snowmelt.

We agree that the measured snow densities are relatively high. Yet, we do think that these densities are realistic and represent a continuously ripe snowpack, given the fact that largest discharge peak had taken place already 1.5 month before the fieldwork (Figure 1) and the air temperature never decreased to freezing point during the campaign (Table 4). Additionally, during the campaign it was noted that the snow pack was relatively wet. In a revised version of the manuscript, we will add these considerations regarding the magnitude of the observed snow densities.

Lastly, indeed in L334-336 the interchangeable use of the SWE and snow density is incorrect, and will be adapted in a revised version.

Line 398-403: Granger et al., 2002 and Weisman, 1977 propose similar power law relationship to describe a sensible heat flux. Perhaps worth contrasting this formulation and the meaning of your terms with those papers?

Thank you for this suggestion, we are currently looking into this.