We thank Richard Essery for his time in providing constructive and thoughtful comments which have certainly improved the manuscript. Responses are detailed below with reviewer comment (RC) followed by an authors response (AR), in each case. Bold text indicates text sections that have been changed in the manuscript.

RC0: “Fiddes, Aalstad and Westermann present interesting results on assimilation of remote snow cover observations in an efficient model of snow accumulation and melt over complex topography. It is not entirely true, as stated in the abstract, that “grid-based models cannot be run at spatial resolutions to explicitly represent important physical processes” – there are numerous examples in literature of models representing multiple physical processes being run on high resolution grids or triangular networks – but these models certainly are not optimal and cannot be run for large areas or long periods.”
AR0: We have qualified this statement as follows:
“Spatial variability in high-relief landscapes is immense, and grid-based models cannot be practically run at spatio-temporal resolutions that explicitly represent important physical processes at scale.”

RC1: page 1, line 10 What are “surfacecheck models”?
AR1: typo, “surfacecheck” -> “surface”

RC2: The abstract should say something about what data are assimilated.
AR2: We have added the following text to abstract l.8:
“We demonstrate marked improvements in estimating snow height and snow water equivalent at various scales using this approach that assimilates retrievals from a MODIS snow-cover product.”

RC3: page 2, line 13 Data assimilation in land surface modelling schemes has been around for longer than might be suggested by citing a 2012 review. The North American Land Data Assimilation System was initiated in 1998, and the ECMWF model has had operational assimilation of snow depth observations since 1987.
AR3: Thanks for this, the sentence is currently misleading. We actually intended to refer to the high resolution surface community i.e. hydrologists or others working on surface processes such as snow deposition. LSM is obviously a term strongly connected to the climate/NWP communities but we need a suitable term for complex “surface models” such as CROCUS, SNOWPACK or GEOTOP. We have changed the text as follows:

“While DA has a long history as a tool employed in NWP (cf. ECMWF, NLDAS), only relatively recently has DA started to be utilised in high resolution surface modelling schemes (Liu et al., 2012), but it has already shown much promise in the current era of plentiful remote sensing data.”

RC4: page 5, line 32 Reference to Figure 1.3.1 should be Figure 1.
AR4: Corrected in text.

RC5: page 6, line 3 Ne is not explained. It is later described as a number of pixels on page 8 and a number of particles on page 10.
AR5: This is a mistake, to be clear Ne = N particles , Np = N MODIS pixels and Ns= N Toposub clusters. We have added the definition on first mention (p6) and corrected occurence on p8.

RC6: page 6, last line ERA5 resolution was earlier stated as 25 km.
AR6: This is a mistake, the original grid of the model is 31km. We downloaded the netcdf product which is reprojected to a regular long/lat grid according to user specification. We set this at 0.25 degrees to match the original grid resolution. We have edited this for consistency throughout the text.

RC7: page 8, line 10 I don't think that Vögeli et al. (2016) says anything about the open availability of the airborne snow height retrievals.
AR7: This is a data citation issue,therefore not always standardised. The data used in Vögeli et al. is available here: https://www.envidat.ch/dataset/10-16904-23.
However, the authors request on that landing page that the manuscript is cited if the data is used. We switch the citation for the dataset doi here as follows:
"This dataset is openly available (doi:10.16904/23)."

RC8: page 10, line 14 A positive bias of high wind velocities is not very apparent in Figure 2.
AR8: This figure was generated with development code which included a wind correction algorithm - which is intended for another paper. We have regenerated the figure with the original TopoSCALE algorithm that preserves the bias. This code development was concurrent with manuscript submission and not at that stage ready to be described or properly evaluated.

RC9: page 11, line 15 Because fSCA contains no information about HS after it reaches 100%, the method might be expected to fail for the very highest accumulations.
AR9: This is true that no info is gained if there is complete complete cover (fSCA=100), however even these extreme depths ablate to fSCA=0 in our region and are therefore suitable for DA. We mask glaciers of course. We think we miss extreme values due to averaging effects as stated - therefore we prefer to leave the text as written.

RC10: page 12, line 14 This first reference to Figure 8 is out of sequence.
AR10: corrected in text

RC11: page 16, line 23 Delete "data was obtained from"
AR11: done

RC12: Table 1 Means and variance lack units
AR12: Units have been added.

RC13: Figure 2 caption Delete "simulated" in the first sentence. The second sentence is ungrammatical and needs to be rewritten.
AR13: Caption now reads as:
"Figure 2. Experimental setup: The 9 ERA5 grid boxes were selected based on the fact that they contained GCOS SWE monitoring sites (11 stations). All IMIS stations in each box are used for evaluation (39 stations). The Weissfluhjoch research station as well as the flightpath of ADS data is located in the red outlined box, which is also shown at a larger scale in the inset."

RC14: Figure 3 Why is there a point with LW close to zero in (D)? STATION should be explained in the caption.

AR14: This erroneous point comes from the last daily mean value in the time series which has been accidently computed from a single datapoint i.e. an incomplete day, hence the low value. This applies to all plots as can also be seen in SWin. We have cut this last daily value from all plots.

The caption is inconsistent with the latest version of the plot and has been edited as follows:

“Multiyear simulations at station WFJ (WY2012-2017) in order to show baseline results for the modelling scheme. (A-E) assesses the downscaling scheme by showing downscaled ERA5 data (ERA5) compared to station measurements (OBS). (F-I) assesses the simulation of target variables SWE and HS in both time series and scatter plots. Here, ERA5 is a simulation driven either by downscaled ERA5 (ERA5) or directly by station measurements (STATION). OBS are SWE and HS measurements made at the station. WY2012 is a clear outlier in poor performing ERA5 as shown by cumulative precipitation errors and in HS and SWE time series. HS and SWE scatter plots also show this low performance in high values attributed to WY2012. Additionally, ERA5 simulated HS is increasingly biased with depth as errors accumulate over the season to max depths. The same pattern is evident with SWE. It is worth noting that in differentiating sources of error these plots are useful. OBS - STATION approximates model error whereas STATION -ERA5 approximates the forcing error.”

RC15: Figure 4 The green dots described by the caption are black crosses in the figure, and there are no red dots. Are there any HS observations that would help to resolve the disagreement between the posterior and the last fSCA observation?
AR15: This is from an earlier iteration in colour schemes where all (assimilated and non-assimilated) obs were shown. The submitted plot only shows assimilated obs for clarity. The text now reads:
“DA run at the Truebsee GCOS station, Engelberg. The left panel shows the assimilation step with the assimilated fSCA observations represented by black crosses. The shading and solid lines show the 90th percentile range and median of the prior (red) and posterior (blue) estimates. The right panel shows the target variable validation, SWE in this case. Posterior/prior are denoted in the same way. Black triangles indicate the measurements used for the validation.”

Unfortunately there are no snow depth measurements at the snowpack disappearance date - the last manual observation corresponds to the SWE data in the second panel of the plot. The closest automatic station (IMIS network) is TIT2 at 2149m asl which is significantly higher than the Truebsee station at 1769m and therefore not comparable.

RC16: Figure 7 Why do all of the distributions extend to negative snow depths? “The observed distribution is better captured by the posterior”
AR16: (a) This is an artefact of the kernel density function smoothing (R function: density Figure this out. Figure this out.). We have changed the data range parameters to constrain this to 0 in the plot. (b) Sentence corrected.

RC17: Figure 8 caption There are no vertical dashed lines in the figure. Labels on several of the figures are too small.
AR17: Vertical lines have been added and labels have been enlarged.