AMUNDSEN is a well-known snow model with multiple capabilities and applications, and this description of the open source version is a useful reference. There is little demonstration of model performance, but that is OK in this model description paper, and the model has been extensively evaluated elsewhere. There is no demonstration at all of the method for generating climate scenarios, and I wonder if so much description is warranted when it is not subsequently used in the paper. Otherwise, there are a few places where I would like to see some more detail, and I have noted some minor corrections.

Line 17

“manyfold applications” – many

This, of course, will be corrected.

64

How long are the ”longer time horizons”, and why is lateral snow distribution then especially important?

We consider „longer time horizons“ to be of decadal length or longer in the simulations, since then snow towers might build up in the upper regions where seasonal melt is smaller than the amount of previously accumulated winter snow. Such artefacts can be avoided by computing lateral snow redistribution processes, removing the snow from prominent summits and ridges, and accumulating it in the slopes and depressions below. This process is also required to correctly compute glacier mass balances in the long term; actually, a way to evaluate the process is to compare the accumulated snow masses with observed glacier mass balances as shown in Hanzer et al. (2016).

This will be re-formulated in the revised version of the manuscript.

86

“v0.9” – the title and text otherwise refer to v0.8.3. (what is it going to take to commit to v.1.0?)

We wanted to wait for all potential improvements of the model during this review process, eventually also in the model code itself, and then name the version that is described in the final revision of the manuscript with number 1.0.

Figure 4

It would be useful to see the station locations (the same as Figure 2a?). What is the resolution of the interpolated grid?

The station locations will be inserted in the revised version of the manuscript.; The resolution of the interpolated grid is 20 m; we will add this to the figure caption.

271

I guess that Liston and Elder (2006) is used for (the vast majority of) catchments that are less well gauged than Rofental. How does this compare with the dynamic lapse rates?

We will compute temperature grids for an exemplary winter season, and compare the results achieved with the two approaches. Dynamic lapse rates can only be computed if the distribution of the weather stations is well covering both the area as well as its elevational extent, and if the time series of the forcing data are mostly complete. We agree: this might often not be the case, but for the Rofental it is.

272

How are the precipitation thresholds chosen?

The threshold was chosen empirically: a value of 0.5 °C wet bulb temperature with a transition extent from 0 °C to 1 °C produced reliable results in many numerical experiments with the model, in particular for the well-gauged site Rofental. We will include this in the text.

282

The method for calculating multiple reflections from clouds and slopes is not described. These reflections contribute to measured radiation, so does the model not end up double counting?

Yes, this is correct. Multiple reflections from clouds and slopes are computed in the simulations, as they also contribute to measured radiation. In the further processing, cloudiness will be derived from the comparison of the simulation results to the observations as shown in Strasser et al. 2004 (doi:10.1029/2003JD003973).

353

It is not clear what it means that “different length scales” are used in Figure 5; none are specified.

Figures 5 and 7 (especially 7d)

The two used length scales are 50 m (this is the scale of ridges) and 5000 m (this is the scale of entire mountains). We will provide the length scales used in the revised version of the manuscript and include a better explanation of the concept of the procedure. See also the respective comment to the remarks of Referee #1.

It is counterintuitive for the areas with more snow to be darker.

We have discussed the color scheme issue for figures of this type already for quite some time, and considered different variants for this. There are both advantages and disadvantages of every color scheme. Finally, we ended up with the used scheme as being a compromise with least weakness.

The Figure 7 caption does not mention that the pink blobs are clouds (it is not a big problem, but there were better Sentinel-2 views on several other days in June 2019).

This will be added in the figure caption. Yes, there might be alternative satellite data scenes with less cloud coverage. We will search the archiv for a scene with less clouds.

422

"orby satellites" – orbital? Or just "satellites" (ones that are not orbital, such as CryoSat-1, are of limited value).

It should be two words: „or by satellites". This will be corrected.

630

I can't tell what the missing early snowfall event is in Figure 9.

This will be added. There is an early snowfall event in the observation which all model versions underestimated. We will re-formulate.

There are no metrics given, but I might judge from Figures 7 and 9 that the most sophisticated EB + Multi + SRF configuration has the worst performance in comparison with observations.

We will include a short explanation of this. In Figure 9 the most complex model formulation has a low performance, but in Figure 7 one can see that it well captures the processes. At the Proviantdepot station frequent small-scale redistribution of snow by wind occurs, leading to effects which worsen the model performance (buildup of a corniche which modifies the SPA observations). We will include a better explanation in the caption of Figure 7.

Thank you very much for supporting the improvement of our paper! All your comments are very valuable.