

Response to reviewer #2 (Marco Möller)

RC = Referee Comment

AR = Author Response

RC: Van Pelt et al. present a multi-decadal modeling study regarding snow and glacier mass balance on Svalbard that yielded results on a so-far unprecedented level of detail with respect to model resolution and captured processes. I very much congratulate the authors to this very thoroughly performed, documented and discussed modeling study which provides extremely valuable new knowledge to the field of Svalbard-wide glacier and snow research. I have no severe concerns regarding publication of this article. However, in its present form, the model description lacks a couple of important details that need to be added to the descriptions in order to make the methodology easier to follow. In this respect, three substantial issues need special and more extensive attention, including limited additional data analysis. Taken together, I recommend to accept the manuscript of Van Pelt et al for publication in *The Cryosphere* after a minor revision along the issues outlined below.

AR: We thank the reviewer for the very positive feedback! We are also grateful for the substantial and detailed comments, which have helped us to improve the manuscript.

Substantial comments:

RC: 1) P4L30f (& P9L16ff): I understand that you first linearly interpolate your 10km HIRLAM precipitation grid to the 1km resolution of your model. So far so good. However, in the next step you describe the application of a fixed linear fractional increase with elevation that you apply in addition. This step causes some concerns. I assume that the 10km elevation information in HIRLAM are not based on the S0 Terrenmodel Svalbard that you use in your mass balance model, right? This means that the average of the 1km elevations in your model across each 10km HIRLAM grid point and the elevation of this HIRLAM grid point itself do not equal each other. If this is the case, it introduces a physical inconsistency. Depending on the area altitude distribution of the 1km model grid points within each 10km HIRLAM grid point you either increase or decrease the total amount of precipitation that falls within this grid point by applying a fixed linear precipitation increase. Hence, the precipitation amounts which had been modeled by HIRLAM in a way that is physically consistent to synoptic forcing, are altered completely by your downscaling scheme. Moreover, this happens completely unstructured with respect to space, as the degree of alteration is only determined by the differences between the means of the 1km model elevations and the 10km HIRLAM elevations. I'm not sure if my interpretation above is what really happens; it could have also been a simple misunderstanding of your descriptions. In any case, I'd suggest that you comment on this issue in detail in the uncertainty discussion and/or revise your descriptions in the methods section accordingly to make them unambiguous in this respect.

AR: We believe we understand the reviewer's concern, but do not regard the precipitation - elevation correction method as physically inconsistent. The two reasons to apply the precipitation – elevation equation are to 1) account for any elevation differences between the detailed 1-km DEM and the coarser 10-km HIRLAM DEM (interpolated to 1-km resolution), and 2) to correct for any Svalbard-wide biases between modelled and observed precipitation. The former is controlled by K2, while the latter is controlled by K1. This precipitation adjustment is not 'mass-conserving', since we correct for biases between model and observations (K1 is not equal to 1). Furthermore, as suggested by the reviewer, there may be (small) elevation deviations between the mean of 1-km height values and the corresponding regional climate model elevation value. This is not at all problematic, since we in fact desire that our correction method also compensates for these height errors by altering the precipitation amount. The actual 1-km DEM will contain more detail than the interpolated HIRLAM-DEM; any positive deviation of the surface height will lead to a positive correction of the local precipitation, while a negative height deviation will lead to a negative precipitation correction. The main advantage of this downscaling is that we introduce the effect of orographic lifting on precipitation at scales

smaller than the 10-km spatial resolution of the regional climate model. To clarify our approach, we have added the following in Sect. 3.2:

“Values from the 1-km DEM (z) will contain more detail than the z_0 values interpolated from the coarser regional climate model grid; any positive deviation of the surface height ($z-z_0>0$) will lead to a positive correction of the local precipitation, while a negative height deviation ($z-z_0<0$) will lead to a negative precipitation correction. With this approach, we account for the effect of local topography on precipitation, thereby capturing the impact of orographic lifting at scales smaller than the resolution of the regional climate model. In addition to compensation for biases in modelled precipitation (by calibrating K_1) potential surface height discrepancies at spatial scales of 10-km and greater that may arise from the use of a different DEM in the regional climate model are also automatically compensated for.

RC: 2) P8L11ff: You implemented two novelties in your model. While the first one, the physically based percolation scheme, is fully referenced, the second one is not. How were the parameters of your newly incorporated albedo scheme chosen? If you use a new or updated scheme, then you need to include information about how it was calibrated or how it is justified from a physical point of view. As you have various AWS data available, I suppose that you could easily validate your new albedo scheme with shortwave radiation measurements from these stations. I do not ask for including an additional figure on this (even if it would be nice to have), but at least you should provide some comparative albedo numbers to validate the novelty in your model, especially its ability to produce a reliable course over the year. Your Svalbard-wide distribution of albedo values may also be compared to those modeled and described for 1979-2015 by Möller & Möller (2017) on the basis of MODIS data. This would yield important insights into the reliability of your so-far unreferenced albedo scheme.

AR: This is a valid point. A similar comment was given by reviewer #1, who questioned the transferability of albedo parameters (t^*) from Greenland to Svalbard. We agree that some more information about the albedo scheme is needed. In Sect. 3.1, we have now extended the description of the albedo scheme by explaining what parameter values were chosen and based on which references. First of all, we refer to Bougamont et al. (2005) for the characteristic decay time-scale (t^*) values, which were optimized on Greenland and used here as well. In the uncertainty discussion (Sect. 4.6), we have now added a discussion on uncertainty related to the potential errors in chosen albedo parameter values like t^* (in response to reviewer #1). Secondly, for three other albedo parameters (albedo of ice [0.39], albedo of firn [0.52] and characteristic snow depth [7 mm w.e.]) we used values that have been calibrated in a previous study with the model (Van Pelt and Kohler, 2015). This information is now added to Sect. 3.1. Thirdly, regarding calibration/validation of the new albedo scheme, it is worth mentioning that we optimize two key albedo parameters (fresh snow albedo & snowfall threshold to reset albedo to the fresh snow albedo) against observations of net SW radiation (see Sect. 3.2). By doing so any biases in net SW radiation resulting from the new albedo scheme are minimized. A sentence has been added to Sect. 3.1 to clarify this. We do not think it is needed to include a more detailed validation or comparison of albedo time-series in this study, although it would be useful to do this in future work with a more specific focus on albedo (this is now also discussed in Sect. 4.6).

The new description of albedo in Sect. 3.1 reads:

“Additionally, we have extended the original snow age and snow depth dependent albedo scheme (Oerlemans et al. 1998). The original fixed characteristic time-scale for exponential decay of snow albedo due to ageing has been replaced with a temperature dependent time-scale (t^). As in Bougamont et al. (2005) snow albedo decays fastest when the surface is melting ($t^* = 15$ d), and for dry snow t^* linearly increases from 30 to 100 days between 0 and -10 °C. The updated albedo scheme avoids overestimation of the albedo of melting surfaces in the early melt season. Other albedo parameters, including the albedo of ice (0.39), albedo of firn (0.52), and the characteristic snow depth for albedo decay of thin snow covers (7 mm w.e.) were taken as in (Van Pelt and Kohler, 2015). To avoid potential systematic biases resulting from the new albedo scheme, we have included the*

fresh snow albedo (α_{fs}) minimum snowfall threshold used to reset the snow albedo to the fresh snow albedo (P_{th}) in the calibration process, as described in Sect. 3.2.”

The new discussion about uncertainty in albedo parameters in Sect. 4.6 reads:

“Energy balance parameters were taken as in the aforementioned studies, with the exception of the fresh snow albedo (α_{fs}), the associated minimum snowfall threshold (P_{th}), and the background turbulent exchange coefficient (C_b), which were calibrated against observational data (Sect. 3.2). The new albedo scheme assumes that previously used values of t^ for Greenland (Bougamont et al. 2005) are also applicable to Svalbard. Potential inaccuracies in parameters like t^* will introduce uncertainty in modeled albedo values, as it introduces compensating errors in calibrated parameters; in the case of t^* , compensating errors would arise in α_{fs} and P_{th} . However, the calibration procedure assures that, despite compensating errors, net biases in relevant model output, e.g. melt and runoff, are minimized. More careful calibration of albedo parameters, including t^* , is planned for future work using a more extensive dataset of albedo measurements across Svalbard.”*

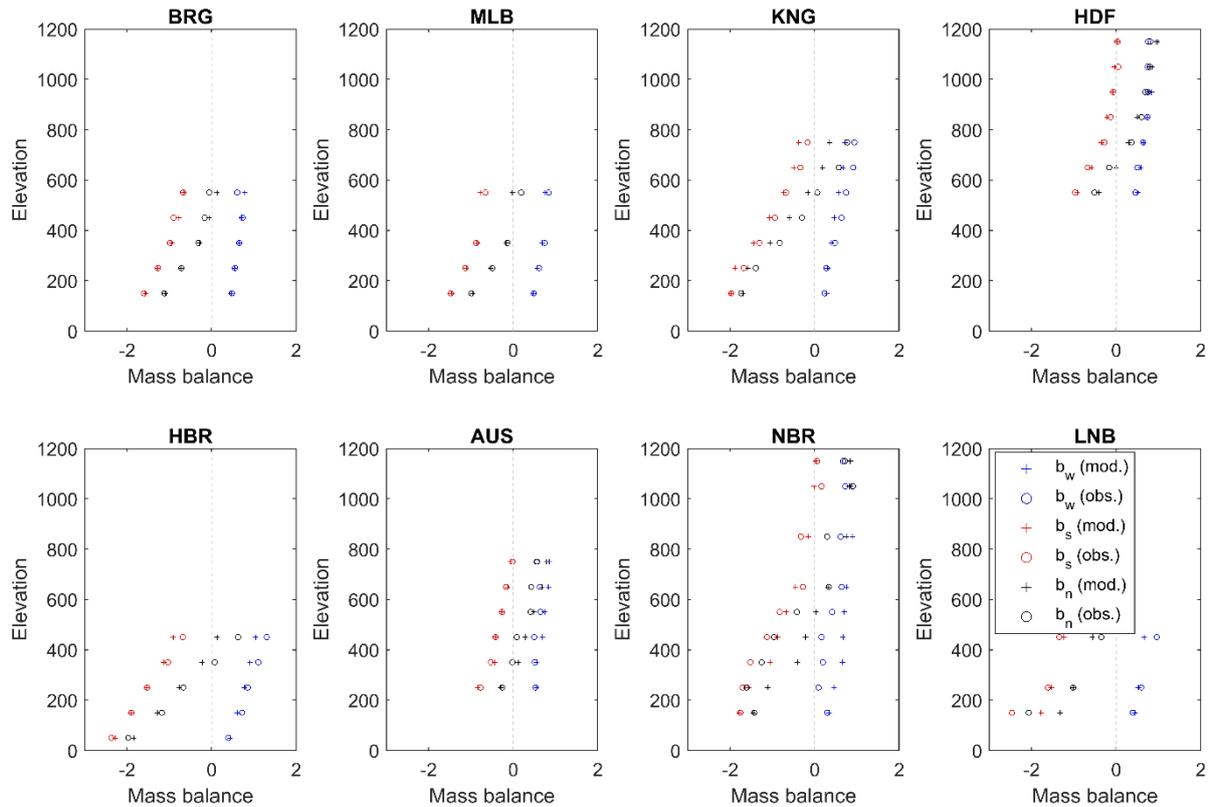
Regarding the albedo maps that were presented in Möller and Möller (2017) we are happy to share maps with the reviewer (and the data behind them) that can be used for comparison. When writing this manuscript, we have carefully selected the variables that we thought were of most value to present and albedo was in the end not selected. We decided to have more focus on the mass balance components and stay away from a detailed analysis of individual energy balance components, including albedo.

RC: 3) P8L25ff: The application of RMSE minimization for finding the optimum values for K1 and K2 is certainly valid. However, I'd like to point towards a potential problem.

In case the stake readings that are used as reference are not distributed equally with elevation, the RMSE minimization might not yield a proper combination of K1 and K2.

This is because elevations with the highest number of stake readings are overrepresented and the minimization procedure thus concentrates on getting the accumulation at this specific elevation well while paying less attention to the other, underrepresented elevations. This issue has been detailed and documented for stake-based calibrations across Svalbard before by Möller et al. (2016) and you should at least account for it in the text. However, it would be better to check it in detail in order to avoid potentially wrong gradients or scaling coefficients that might lead to substantial under and/ or overestimation of climatic mass balance towards higher elevations. My concern is backed by the fact that the winter balances in Figure 3 clearly show, that modeled values tend to systematically underestimate the measured ones the more positive they become. If you visually place a linear fit to the blue points, the line would have a slope that for sure is distinctly larger than 1. I do not think that this issue will compromise your overall results as it probably only affects the most positive mass balances. Nevertheless, it needs to be presented in the uncertainty discussions section.

AR: This is an interesting comment. We will start our response by showing a figure. To assess whether there may be a potential bias increasing with elevation we have made height profiles of b_w , b_s and b_n for the eight observed glaciers. Here data are averaged in 100 m elevation bins and over the whole observation period.



What can be seen is that there are some glaciers (particularly KNG and HBR) where b_w biases indeed increase with elevation. However, for the glaciers with the highest elevations, i.e. over 1000 m a.s.l. (HDF and NBR), we do not see this effect. What this shows is that the bias for high observed b_w values that seems apparent in Fig. 3 in the manuscript is likely to come from underestimated precipitation at the highest stakes on KNG and HBR, where long data records exist (i.e. many data points) and precipitation amounts are among the highest on Svalbard (much higher than at the much higher-located stakes on NBR and HDF). These local effects are apparently not properly captured by the model, but it cannot be concluded that precipitation amounts are off at high elevations in general. We now refer to Möller et al. (2016) and have extended the discussion on this at the end of Sect 3.2:

“On the other hand, underestimation of b_w is apparent for KNG and HBR (Fig. 3, Table 2), which results from underestimated orographic precipitation at high elevations on these glaciers. Nevertheless, high-elevation biases of b_w do not arise on the only two glaciers extending above 1000 m a.s.l., which indicates that the b_w offsets on KNG and HBR are not a systematic feature for high elevation sites in general. The relative lack of stake observations at heights above 1000 m a.s.l. implies increased uncertainty of modelled precipitation estimates at these elevations (Möller et al. 2016)”.

Detailed comments:

RC: P2L19f: The reference to Day et al. (2012) is misleading here. They only calculate changes in surface mass balance on the basis of seasonal sensitivity characteristics. They neither use a mass balance model nor do they calculate absolute Svalbard wide mass balance numbers. Instead of Day et al. and to complete the ensemble of Svalbard-wide mass balance calculations a reference to Möller et al. (2016) is missing here.

AR: We have added a reference to Möller et al. (2016) and removed the reference to Day et al. (2012).

RC: P2L34f: The references to Hagen et al. (2003) and Winther et al. (2003) are completely outdated as numerous studies related to seasonal snow coverage on Svalbard have been published since then, partly even with contributions of one or several co-authors of this study here. Hence, there is a need to include more up-to-date references here (e.g. Grabiec et al. 2011 or else).

AR: We believe the two 2003 references are still relevant, but now also include the suggested newer reference to Grabiec et al. (2011).

RC: P4L27: Include information about which ECMWF reanalysis products are used and over which periods. Even if this is partly deducible from the references it needs to be explicitly stated in the text.

AR: Fixed.

RC: P4L30f: Far more, especially quantitative, information about the applied lapse rates, increases and decays is needed here. The reader must fully understand what has been done without digging into previous literature.

AR: This is a good point. We have added the following sentences in Sect. 2.1:

“Elevation functions for temperature and air pressure were constructed per 3h time-step through respectively linear and exponential regression of the regional climate model values and their corresponding elevations; this procedure was repeated for blocks of 4x4 grid cells and regression coefficients were averaged for the whole grid to obtain a single lapse-rate for temperature and exponential decay coefficient for air pressure per time-step.”

RC: P5L1: How did you calculate the significance of the trends? Information about this needs to be included here.

AR: It is now added that significance means that a zero trend is not included in the 2-sigma confidence bounds of a trend.

RC: Figure 2: Just out of curiosity (as it is out of the scope of your study): do you have an explanation for the rather interesting pattern of precipitation trends visible in (d). I especially refer to the east coast of Wijdefjorden here.

AR: Well, the precipitation trends are generally very small; the trend values that still appear as colors (non-gray) in the map are also only just significant. In general a slight precipitation increase in Svalbard has also been observed in observational records (see also Introduction). And the somewhat stronger precipitation trends in northern Svalbard could be a result of retreating sea ice having a larger impact on moisture availability there. This effect has also been suggested to amplify in a future climate (see for example the recent Climate in Svalbard 2100 report).

RC: P7L8: This value seems to be reasonable as it is quite often used for the transition from snow to firn. However, the choice appears to be rather arbitrary in the present form of the text. Information about how this choice was made, including appropriate reference, needs to be given here. Moreover, as you give explicit information about the density applied to remaining snow you should also do so for other snow cover to ablation conversions that you consider in your model.

AR: We have removed the 550 kg m^{-3} value from the paper and instead refer to Van Pelt et al. (2018) where the summer balance calculation from stake measurements (in that case on Nordenskiöldbreen) is also described. For the conversion of stake heights to winter balance estimates, fresh snow density is needed, which is based on snow pit data. Since this is already described in Van Pelt et al. (2016) we only keep the reference to that paper for more information.

RC: Figure 5b: One might think about the color scale here. People tend to associate blue colors with cooler conditions, which means more positive mass balances in glaciological studies. However, in the

current version of this figure, blue represents a negative trend and thus a development towards more negative mass balances. Maybe it would be better skip the in the first instance ambiguous blue-red color scale here in favor of something more "uncommon". But that's just a suggestion as it reflects a rather subjective view.

RC: Figure 10: Same "problem" with the color scale as in Figure 5b. But this is still only a suggestion.

AR: Thanks for the suggestion. We understand the confusion this may cause, but we still think it is good to stick to the current choice of colormap for the sake of consistency. We like to be consistent throughout the manuscript with blue colors for a negative trend and red colors for a positive trend. For some variables this may be intuitive (e.g. temperature), whereas for others like CMB it may be counterintuitive, but this seems unavoidable.

RC: Figure 9: An additional map showing the distribution of the percentage of melt and rainwater that is refrozen should be added here and any inferable information should to be included in the discussion where appropriate.

AR: This is a useful suggestion, and we have decided to add two panels to Fig. 9 showing the fraction of melt and rainfall that refreezes (Fig. 9c) and the associated trends (Fig. 9d). These new results interestingly show that highest refrozen fraction values occur on Lomonosovfonna and lowest values in coastal regions in southern Svalbard. Furthermore, it is found that no sites experience a refrozen fraction that is close to 1 (values up to 0.8 are found), which indirectly shows the absence of cold firm in Svalbard during the simulation period. Trends of the refrozen fraction reveal most negative values in northern Svalbard. All this is now discussed in more detail in the second and third paragraph in Sect. 4.3.

RC: P22L21ff: You describe the usage of a fixed DEM and fixed glacier mask as a potential source of uncertainty and error. However, in the beginning of your paper you explicitly state that you calculate reference surface balances. Hence, your results do not suffer any "uncertainties" or "errors" due to the usage of fixed glacier extents and elevations.

They simply represent a completely different quantity that is not comparable to "real" climatic mass balances which would be based on a time-varying glacier topography. This needs to be made clear in this section. You could of course keep the given descriptions, but treat them as deviations to what really happened on the glaciers and not as "uncertainties".

AR: Good point. We now refer to 'deviations' rather than 'uncertainties' instead, and have added some text on the use of "reference" surfaces for mass balance modelling (Sect. 4.6).

RC: P23L7ff: The discussion of misestimations of precipitation fits to my substantial comment 3) in the beginning. The issue raised in this substantial comment needs to be included in this discussion, too.

AR: We have added some discussion on this now earlier on in Sect. 3.2. For more details, see our reply to substantial comment 3).

RC: P23L17ff: You might explicitly refer to the influences of wind-drifted snow, i.e. its potential to systematically increase or decrease local as well as regional accumulation rates.

This information is certainly assumed in the sentence in question here, but it should be explicitly stated and referenced (e.g. Jaedicke and Gauer 2005; Grabiec et al. 2011).

AR: We have added a notion on wind-driven snow redistribution and included the additional references (Sect. 4.6).