Response to Anonymous Referee #1

RC = Referee Comment
AR = Author Response

RC: This is an exceptionally thorough and robust modelling-based paper investigating the climate mass balance (CMB), which includes surface and subsurface processes, across Svalbard between 1957 and 2018. It builds on previous similar work by the team, esp. first-authored papers by van Pelt, but this is the first time the latest version of the model (which now includes an improved subsurface scheme based on Marchenko et al., 2017b) has been applied to the whole of Svalbard. The model has a 1 km grid and is run at a 3 hourly time step and is therefore impressive in terms of its spatial and temporal resolution. The CMB is driven by downscaled climate data from the High Resolution Limited Area Model (HIRLAM) regional climate model, which is forced by European Centre for Medium Range Weather Forecasts (ECMWF) reanalyses. This generates the meteorological forcing fields of air temperature, precipitation, cloud cover, relative humidity and air pressure. The work uses an extensive data set of measurements to calibrate / validate the model (mass balance stake measurements from 8 glaciers; weather station data from 6 sites(4 on glacier; 2 off glacier); and shallow ice cores from 4 sites). These are listed in Table 1.The calibration procedure is clearly explained and is logical and the principles have been discussed in two previous referenced papers. Here, parameters that are known to be sensitive are calibrated in sequence as described in section 3.2: 1. Two parameters affecting albedo are calibrated against net SW radiation data; 2. Two parameters in a function describing the downscaling of precipitation are calibrated against winter stake bass balance data; 3. Two parameters affecting summer melt are calibrated with observed summer balance data. The fact that this model has a good history of being used in Svalbard and the fact that the RMSEs and biases after calibration are small, mean that the results will be the best that are currently available. Results presented are quite extensive and informative and, as the authors state in the abstract, should be of value for scientists and practitioners interested in runoff to the oceans as well as ecologists interested in, for example, snow extent, duration and character (which has implications for reindeer grazing, for example).The results / discussion section is focused around a sequence of Figures showing: i)maps of mean conditions across Svalbard; ii) maps of trends over time (where significant); iii) time-series of spatially averaged trends in conditions. The consistency in the way the data are presented make the paper especially useful. The following results are shown and discussed: i) glacier CMB (Fig 5); ii) glacier ELA (Fig 6); iii) glacier firn pore space in top 14 m (Fig 7a,b); iv) firn temperatures at 14 m (Fig 7 c,d); v) refreezing on and off glaciers (Fig. 9); vi) snow onset and disappearance dates off glaciers and across glacier ablation areas (Fig 10); vii) glacier and land runoff. This represents a particularly impressive range of data sets presented and discussed from this type of modelling study. The paper discusses sources of uncertainty throughout and has a synthesis section on this towards the end (section 4.6). Where results differ from those of similar previous work (but using earlier versions of the model, calibrated in different ways, run over different time periods, and across different spatial domains) the magnitudes and reasons for the discrepancies are revealed. The results and implications of the Svalbard work are also discussed in the context of similar work where appropriate in Arctic Canada and Greenland; this is especially the case when discussing the important finding of decreasing refreezing rates over time and therefore an increase in the likelihood of firn aquifers developing around the ELA. So overall this represents excellent work by this team and shows the value of long-term monitoring but also the collection of shorter-term field measurements and their rigorous use in model development and application. The work is exceptionally well presented in terms of the overall paper structure, as well as the clarity and precision of the writing, but also in the consistency and quality of the Figures.

AR: We are very grateful for this very positive feedback! And we thank the reviewer for the useful comments, which we address below, and which have helped to improve the manuscript.

RC: As mentioned above, quoting from the paper, the meteorological forcing fields used to drive the CMB model are: air temperature, precipitation, cloud cover, relative humidity and air pressure. The answer is probably elsewhere in previous papers but a brief note on how these are used (together with other fields I assume) to calculate energy/ mass balance at the surface would be useful. For example, there is no mention of wind-speed here, and yet I assume this is required together with air temp and relative humidity to calculate the turbulent fluxes? And I assume theoretical clear sky solar radiation is used together with cloud cover to determine the incoming SW radiation?

AR: More details on the individual energy balance components and their dependence on climate fields is given in Van Pelt et al. (2012). However, we agree some more information would be helpful to include here as well, so we added the following in Sect. 3.1:

"Solving the surface energy balance requires input of near-surface meteorological conditions, including air temperature, cloudiness, relative humidity, air pressure and precipitation (Van Pelt et al. 2012). No wind information is needed since sensible and latent heat exchange depend solely on near-surface temperature and specific humidity gradients, following katabatic turbulent exchange relations by Oerlemans & Grisogono (2002)."

RC: P8 L17-19. The Bougamont et al (2015) work is for Greenland. How do you know that parameter values derived for the GrIS for t* are valid on Svalbard. The final sentence refers to the work on the GrIS I assume. Given the importance of albedo for melt and mass balance etc, some clarity is needed here about the validity of using the parameter values relevant for GrIS here in Svalbard. Is this a source of uncertainty that needs better recognition?

AR: This a good point. At present we cannot confirm or deny that the t* values from Bougamont et al. (2005) are appropriate for Svalbard glaciers. Although we did not calibrate t* values, it is worth mentioning that as part of the calibration procedure in Section 3.2 we have used SW net observations from three AWSs in Svalbard to calibrate fresh snow albedo and the minimum snowfall amount at which the snow albedo is reset to the fresh snow albedo. These are two albedo parameters to which modelled melt has in a previous study (Van Pelt et al. 2012) been shown to be highly sensitive. By using SWnet data in the calibration, we avoid substantial biases in the surface energy balance (and calculated melt rates) resulting from potentially inaccurate parameter values affecting incoming and reflected SW radiation, including t*. However, any inaccuracies in chosen values of, for example, t* would be compensated for by a potentially different value for the fresh snow albedo and/or the snowfall threshold. In future work, a more detailed comparison of modelled and observed albedo at multiple sites in Svalbard and for longer time-series would allow for more extensive calibration of albedo parameters. To acknowledge uncertainty in parameters like t*, we now include the following sentences in Sect. 4.6:

"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, is minimized. More careful calibration of albedo parameters, including t*, is planned for future work using a more extensive dataset of albedo measurements across Svalbard."*

Also, in response to reviewer #2, we have additionally extended the description of albedo in Sect. 3.1, to better introduce the albedo scheme and give more information on where the parameter values come from (if not calibrated).

RC: P9 L33&34. It's stated that the parameter Tsr has a strong impact on summer melt but most previous work has shown it's particularly important for winter accumulation. I can see it'll have an indirect impact on summer melt because of its direct impact on winter accumulation. Can you better justify why this parameter is tuned to the summer mass balance data and not the winter mass balance data?

AR: Indeed, Tsr has some impact on winter balance as well. However, we find that the sensitivity of the winter balance to Tsr changes is about twenty (!) times smaller than the sensitivity of summer balance to Tsr changes for the mass balance stake locations. Several factors play a role here, but the relatively insensitivity of winter balance to Tsr is primarily explained by the fact that rainfall during the core winter season is (still) rare in Svalbard, especially at higher elevations. And in case any rain falls during the core winter season, most of the rain water will refreeze in the snow pack thereby not inducing any runoff. The significant impact of Tsr on melt and runoff (and thereby summer balance) has previously been quantified in Van Pelt et al. (2012) for Nordenskiöldbreen. We have added a sentence to Sect. 3.2 explaining the relative insensitivity of winter balance to Tsr.

RC: P22 L1-2. There is a bit of confusion here as you seem to be discussing runoff rates due only to snow melt on land and comparing them to runoff rates due to snow and ice melt across glaciers. But, as you say later, runoff from land includes rainfall. Does runoff from glaciers also include rainfall? A better articulation of precisely how runoff is calculated for land and for glaciers is needed before the two values are compared. Can you separate out runoff from snow(ice) melt from runoff due to rainfall?

AR: We now try to avoid this confusion by first giving the definition of runoff in Sect. 4.5:

"Here, runoff refers to the amount of water originating from melt and rainfall at the surface and available at the base of the snow/firn pack (if present) or ice/soil surface after accounting for retention by refreezing and irreducible water storage."

Typos / technical issues

RC: Abstract P1 L4. Could say: "climatic mass balance (CMB) for the glaciers, snow conditions and runoff"
RC: L8. Suggest "small" not "weak"
RC: P2 L4. "reveals" not "reveal"? The Longyearbyen time-series is singular not plural?
RC: P2 L19. Could add the following reference to this list of previous studies here:
Rye, C.J., Willis, I.C., Arnold, N.S. and Kohler, J., 2012. On the need for automated multiobjective optimization and uncertainty estimation of glacier mass balance models. Journal of Geophysical Research: Earth Surface, v. 117,
RC: P4 L21 "altitudes" (i.e. plural)

AR: All fixed.

RC: P5 Table 1. Table is not quite self-contained. Suggest adding to Table Heading and referring to Fig 1 heading for abbreviation names. Also to explain variables or say they're explained in the text.

AR: We have extended the Table header with more details about the used abbreviations.

RC: P5 L10. Could add ref to Table 1 after final sentence here.
RC: P6 L8 suggest "made" not "done"
RC: P7 L5. Suggesting adding months when end of summer measurements are typically made (like April is stated earlier in the sentence for when Spring measurements are made). I'm guessing this is August or September (since 1 Sept. is stated as an average
time below)?
RC: P7 L15 Could delete "above described"

AR: All fixed.

RC: P11 L4 and Table 2. The term ‘bias’ is introduced here and referred to as “modelled minus observed”. There are different definitions of bias so it might be worth clarifying precisely how it’s defined here. Is it simply the Mean Absolute Difference (MAD)?

AR: We have added that the bias is the mean absolute difference.

RC: P11 L29. “five” should read ‘six” here I assume? There are 6 sites mentioned in Table 1 and 3.
RC: P11 L32. “…temperatures for both…”

AR: Both fixed.

RC: P13 L11. Should this say “net CMB” to distinguish it from winter or summer that are also reported? Could clarify the first time you refer to net CMB, e.g. say “net CMB, hereafter just CMB…” or some such. In Abstract you might then also add the word “net”?

AR: Good point. We have followed these suggestions.

RC: P17 L25-27. There is also some similar work to this reported recently from the Larsen C ice shelf, Antarctica that could also be compared / referenced. e.g.

Hubbard, B., Luckman, A., Ashmore, D.W., Bevan, S., Kulessa, B., Kuipers Munneke, P., Philippe, M., Jansen, D., Booth, A., Sevestre, H., Tison, J.L., O’Leary, M., and Rutt, I., 2016. Massive subsurface ice formed by refreezing of ice-shelf melt ponds. Nature Communications, 7.

Bevan, S. L., Luckman, A., Hubbard, B., Kulessa, B., Ashmore, D., Kuipers Munneke, P., O’Leary, M., Booth, A., Sevestre, H., and McGrath, D. 2017. Centuries of intense surface melt on Larsen C Ice Shelf, The Cryosphere, 11, 2743-2753.

AR: References added.

RC: P20 L3-5. There is a lack of clarity here. Here and the few sentences above need to better distinguish between a discussion of snow onset date and snow disappearance date. There’s ambiguity here as it seems as though you might be comparing the trend in onset date (+1.4 days / decade) found in this study with trends in BOTH the onset date AND the disappearance date in a previous study. There is a bigger discrepancy in the disappearance date trends in the two studies than there is between the two onset date trends, and this probably needs stressing and discussing. I wouldn’t say a disappearance date trend of +0.7 days / decade is comparable with 0 days per decade.

AR: We have reformulated the associated sentences to improve clarity.

RC: P23 L34. I think this should just read “…simulation, using the climate forcing…”.

AR: Fixed.