

We would like to thank both reviewers for their valuable comments and suggestions our manuscript. Please find our responses and relevant changes (*italic*) to comments (**bold**) below.

Referee #2

This manuscript presents a thorough evaluation of a regional atmospheric model (WRF) coupled to a mass balance/snow model (CMB) using an extensive and divers data set. The manuscript is well written and well structured. It adds to available publications in using a smaller grid spacing, resulting in a better representation of the topography and topographic related processes, and the thoroughness of the evaluation.

I do have some (minor) comments, which are listed below.

General comments

Explain the use of the term 'Climatic Mass Balance'. This paper addresses the surface mass balance and covers a period of 10 years. The term 'climatic mass balance' is not a generally used term and a bit confusing.

We use the term climatic mass balance as recommended by Cogley et al. (2011). The climatic mass balance includes subsurface processes like refreezing and melting from penetrating radiation in addition to the surface mass balance processes, while the effect of dynamics on the mass balance is not included. This is now explained in the text:

Throughout this study we distinguish between the surface mass balance (SMB) and the CMB as recommended by Cogley et al. (2011). The SMB specifies mass changes between the surface and the last summer surface, whereas the CMB also accounts for internal accumulation and ablation (i.e. below the last summer surface). We consider internal ablation as negligible and it is therefore not explicitly treated in our application.

In the title you state that the atmospheric model is coupled to a glacier model. Without reading more, this might suggest coupling to a dynamical glacier flow model, while the coupling is in fact to a mass balance/snow model.

We have now changed the title to:

The climatic mass balance of Svalbard glaciers: a 10-year simulation with a coupled atmosphere - glacier mass balance model

Specific comments

Abstract:

See comment above, add that you study the surface mass balance. You do not address the total mass balance since this setup does not provide any information about dynamics.

As described above, the term "climatic mass balance" is used instead of surface mass balance, as mass fluxes in the subsurface are accounted for in the model.

Introduction:

P5777 L17: I don't think MAR is a statistical model. Thus is the reference to Lang et al., 2015b correct?

It is correct that MAR is not a statistical model. However, Lang et al., 2015b pointed to the fact that most simulations of Svalbard glaciers were with statistical or empirical models, which is what we refer to. We have now changed this sentence to make this clear:

Regional model estimates of surface or climatic mass balance of Svalbard glaciers have so far mainly focused on individual or a few glaciers, and as noted by Lang et al. (2015b) typically been based on empirical or statistical models.

P5778 L5: Rephrase this goal. I don't see the goal of this work as to reduce the spatial and temporal gap between observations and models. I think the goal is to thoroughly evaluate the surface mass balance product from a coupled atmosphere-snow model system and to investigate what spatial resolution is sufficient to describe the observations.

The aims of this study have been clarified:

In this study, we aim to i) simulate the climatic mass balance (CMB) of Svalbard glaciers with a higher resolution than previously done with dynamical downscaling, ii) validate the model with a more extensive set of observations in this region, and iii) investigate the spatial resolution needed to describe the observations.

Methods:

P5779 L23: How is your glacier mask defined and how do you deal with grid boxes partially covered by ice and partially land?

The glacier mask is based on Nuth et al. 2013. Grid boxes are treated as either glacier, land (tundra) or sea, as WRF does currently not allow for different land types within each grid box when the NoahMP land surface model. Including the CMB model does not change this basic structure.

P5779 L24: I don't understand why 2 transient runs of 5 years instead of 1 of 10 years will save computational time. I have the impression that they are run one after the other. Furthermore, this way the second one also needs to spin up.

The two simulations were originally simulated simultaneously. This did not save CPU hours, but made it possible to complete the simulation in less time. The new simulation is however simulated as one transient 10-year simulation.

P5780 L9: I guess that 'climatic mass balance model' is the name of this model but that it is basically a surface energy/mass balance model including a subsurface snow/ice model.

See explanation of the term "climatic mass balance" above.

P5780 L16: I agree that it is not necessary to include all the details of the GMB model, nor the details of how it is coupled to MAR, but you do have to state how it is coupled in general, and what input parameters CMB needs. This coupling is what makes this research stand out compared to other efforts to determine the surface mass balance using WRF. How does this model compare to the standard WRF surface scheme? Furthermore, Collier et al., 2013 describe that there are 2 options how to couple: only one way, or two way, and that is a major difference. Which method do you use? What is calculated by WRF and what is calculated by CMB? What parameters from WRF does CMB use and what is feed back into WRF (in case of 2 way coupling).

This has now been included in the description of the CMB model:

The CMB model uses near surface temperature, humidity, pressure, winds, as well as incoming radiation and precipitation as input from WRF, and computes the column specific mass balance from solid precipitation, surface and subsurface melt, refreezing, and liquid water storage in the snowpack, and surface vapor fluxes. The model solves the surface energy balance to determine the energy available for surface melt, and resolves the glacier subsurface down to a user defined depth (here 20 m divided into 17 vertical layers). For glacier grid cells, the CMB model updates surface mass and energy fluxes in the WRF model, as well as surface temperature, roughness, and albedo, resulting in a two-way coupled model system (WRF-CMB).

P5781 L 3-11: I am not sure I understand the problem and its impact on the results. Furthermore, if you don't know what the problem is, how is the reader to judge the value of the results in general? L4 'low' mass balance values refers to negative values? L6 Why do you not exclude internal melting in all months? (I guess that internal melting results from radiation penetration?). L9 the values

stated here are not very clear. What is compared here to give these numbers? AND how can I judge whether this is big or small?

See answer to comment 3. of referee #1.

Model validation

P5785 L19: The temperature signal is dominated by the diurnal cycle, thus these high correlation coefficients are not surprising.

This correlation is calculated from daily mean values and does therefore not include the diurnal cycle. It is clear that the high correlations are partly due to the annual cycle. However, the temperature variations within especially the winter season can be larger than the difference between mean summer and winter temperatures (e.g. Aas et al. 2015). This correlation therefore mainly reflects that the WRF model forced with ERA-Interim boundary conditions is able to reproduce large scale weather in a good way.

P5785 L23: Can you explain the biases in temperature and long wave fluxes?

This is discussed in detail by Aas et al. 2015 for a similar WRF simulation of Svalbard, who found these biases to be mainly related to cloud and boundary layer processes. As mentioned in section 6.2.1 these biases are however small compared to other similar studies.

P5787 L7: Since individual years correspond less well than the mean, there are compensating errors.

This might be part of the explanation. However, Fig. 3 shows only two extreme years in addition to the mean, and it seems that the model has more problems with accurately simulating these years than the more average years, as seen from Fig. 4. This has now been clarified in the text as below: *The agreement, however, is not as good for these two years as for the mean values.*

P5787 L13: My guess is that it is not the simple ELA estimate that results in wrong ELAs, but the model being not capable of representing the surface MB correctly, resulting in wrong ELA estimates.

For the three first glaciers we agree that the differences in ELA are related to the simulated CMB. However, at Hansbreen, it is more challenging to find a single ELA, as the measured MB profiles suggest multiple ELA for some years (see figure 3). We have however modified the text to better reflect this:

To look further at temporal variations, we compare modeled ELA to that derived from stake measurements (Fig. 4). In general there is good agreement both in terms of ELA as well as inter-annual variability, with the exception again being Hansbreen. Here the model strongly overestimates ELA, in accordance with the underestimation of b_w (Fig. 3). However, at this glacier, the observed mass balance – elevation relationship shows considerable non-linearity, with the observations indicating zero b_n at several elevations during some years, rendering ambiguous ELA estimates. For the other three glaciers the model simulates ELA well, including the large difference between Kongsvegen and Høltedahlfonna, which are located close to each other (Fig. 1).

P5787 L27: Do you have any information from observations about the role of snowdrift redistribution and evaporation on Svalbard?

See reply to comment 9 by referee #1.

P5788 L9: Do you have any idea on how large the contribution of dynamics is? Without any knowledge about that the statement that the comparison is good has no real value.

We have now included the estimate of calving flux (dynamics) from Blaszczyk et al. (2009). These numbers are however still not entirely comparable as is now described. In addition we changed the statement from the comparison being good, to that the model reproduces regional differences very well:

Note that measured height changes are also a result of glacier dynamics (although the retreat or advance of the calving front was not accounted for). The model, however, does not include any glacier dynamics and the numbers are therefore not directly comparable. Still, both model and satellite data show Austfonna and northeast Spitsbergen to be the only regions with positive surface height change during these years, and northwest Spitsbergen as the region with the largest surface lowering (Table 4). The other three regions all show moderate lowering in both estimates. The model therefore seems to capture regional differences very well during this period. The mass loss from calving flux has been estimated to be 6.75 km³ yr⁻¹ (w.e.) for the years 2000-2006 (Blaszczyk et al., 2009), which corresponds to an additional lowering of about 0.2 m yr⁻¹. This suggests that the model in general simulates too much surface lowering in this period. However the time periods are different and the estimate of Moholdt et al. (2010) does not include the effect of retreat or advance of the calving front. One must therefore still be cautious when comparing these numbers.

Sensitivity to model resolution:

P5789 L6: Why did you choose an extreme case instead of an average/representative case?

Due to the high computational cost of simulating 1km resolution, we chose a period with many precipitation events in all regions of Svalbard in relatively short time. Selecting a more average year would mean that the simulation period would have to be extended to get the same number of precipitation days or the same amount of precipitation.

P5789 L9: How is this difference related to the different estimated glaciated area on 9 an 3 km resolution.

This difference does not seem to be related to different glaciated area in the two resolutions. For the BE region, where the difference in specific mass balance is largest, the glaciated area is almost the same (R9: 2592 km², R3: 2628 km²).

Climatic mass balance:

P5790 L26: Consider adding a table presenting average and standard deviations of the information presented in figures 9 and 10 in order to quantify the interannual variability and the regional variability therein.

Thank you for this suggestion. We have now included this in a new table (Table 5).

Discussion:

P5792 L4 and L21: I guess the better results are largely related to the resolution of your model run, 3 km vs 10 km for Lang et al, and 25 km for Day et al. Furthermore, Day et al do not actually calculate the surface mass balance since they do not have a snow model that calculates the melt. I am therefore not surprised by the better results. This should be stressed more.

Thank you for this comment. We have now modified the text to make both of these points more clear:

Both of these studies compare their results with ice core measurements in accumulation areas (Pinglot et al., 1999; Pinglot et al., 2001), although DA12 only includes accumulation and not melting in their simulation.

Some of the improvement found in the present study can probably be attributed to the increased model resolution. Both smaller elevation differences between stations and the model grid cell and better resolved surrounding topography likely improved the results. We also note here that our results do not cover the same time periods as DA12 and LA15b, so that the quality of the boundary conditions might differ. However, LA15b reports relatively small biases for ERA-Interim (between -

1.95 and 2.24 °C) despite similar or larger elevation differences and covering the same period as LA15b. Also, the large increase in resolution from DA12 (25 km) to LA15b (10 km) is not accompanied with a clear improvement in the mass balance simulation. It therefore seems clear that the WRF-CMB model with the setup used here offers a real improvement over DA12 and LA15b.

P5794 L14: Explain how you apply this 30%? Or perhaps phrase differently: we had to limit the stability correction of turbulent fluxes to prevent too stable conditions to occur.

Thank you for this comment. We have now rephrased this in the following way:

Based on their results, and to avoid runaway cooling of the glacier surface during stable conditions in the winter, we limited the reduction in the turbulent fluxes in stable conditions to 30%, consistent with previous studies (Martin and Lejeune, 1998; Giesen and others, 2009; Collier et al., 2015).

Conclusion:

P5795 L24: Thus the year to year variability is temperature driven? In combination with length of the melt season?

It seems that temperature is an important factor in the year to year variability. This study is however not suited to draw definite conclusions about this, which should be better reflected in the new text: The largest component in the summer surface energy balance driving this melting is the radiation imbalance, even though *temperature dependent latent and sensible heat fluxes also contribute to much of the year-to-year variability, especially during the years with anomalously large mass loss. More research is however needed to better understand the drivers of this variability.*

Tables:

T1: Is the depth scale in cm w.e. or cm snow?

This refers to the physical snow depth, which is now specified in the table. A error in this table has also been corrected: the depth scale should be 3.0 cm, not 30 cm.

Figures:

In general ALL your figures must be bigger.

All figures have now been made larger. In particular, we have increased the size of most individual figure panels. It should however be noted that the discussion paper layout makes the figures appear very small, and should appear larger in a final printed version.

F1: Increase figure size, especially axis labels of overlay figure bottom right.

This has been increased.

F2: Increase figure size, especially the bars indicating precipitation.

The four panels have now all been made bigger. We have also tested increasing the size of the bars, but this would make them overlap for Kongsvegen 2013.

F3: Increase figure size, I can't judge the differences, they are too small. Furthermore, consider adding uncertainty bars indicating interannual variability in terms of standard deviation.

Again the individual panels have been made larger, as well as the lines smaller, which should make it easier to see the differences. With data from only 10 years we considered the two extreme years to give as much information about the spread of the results as the standard deviation.

F4: Increase vertical axis in size to enhance the interannual variability.

This has now been done.

F5: Increase figure size, especially the inset figures in b and d are much too small. Furthermore, check the caption, it refers to 2005 while above the figure it states 2006.

We agree that the inset figures here became too small. We have therefore replaced figure b and d with the inset figures, as well as corrected the figure caption. Thank you for pointing this out.

F7: Check spelling of Kongsvegen in the caption.

This has now been corrected.

F9: Increase vertical axis in size to enhance the interannual variability.

The vertical axis has now been increased.