Dear Colleague,

Thank you for your helpful comments and feedback on our manuscript.

Please note that in the revised version of the manuscript we have corrected an important issue with the glacier albedo scheme, wherein the surface albedo was being reset to the value of fresh snow after only very small amounts of precipitation. Specifically, when we updated the coupled model to WRF v. 3.6.1 and the land-surface parameterization to the Noah-MP scheme, we inadvertently retained the criteria of the Noah scheme for resetting the surface albedo to the value of fresh snow, which is a frozen fraction of precipitation of 50%.

Since submitting the manuscript, we applied the model in a study of the Nepalese Himalayas, where in situ measurements revealed that a threshold based on the depth of solid precipitation, of ~1 cm, greatly improves the simulation of glacier surface albedo. This threshold is in agreement with both the default option in the Noah-MP LSM and in the latest version of the CMB model. Therefore, we adopted a 1-cm threshold and repeated the DEB and CLN simulations with an otherwise identical model configuration.

In Fig. R1, we compare altitudinal profiles of the basin-mean snow albedo in July and August 2004 from the MODIS Aqua/Terra (MOD10A1/MYD10A1; 500-m resolution) datasets with (i) the CLN simulation we showed in the discussion paper ("old") and (ii) the CLN simulation we have incorporated into the revised manuscript ("new"). The old simulations have a high glacier surface albedo (> 0.7) down to ~ 4000 m, which explains a significant part of the cold bias we struggled with in the discussion paper. Conversely, the new simulations provide a basin-mean profile that is in much closer agreement with MODIS.

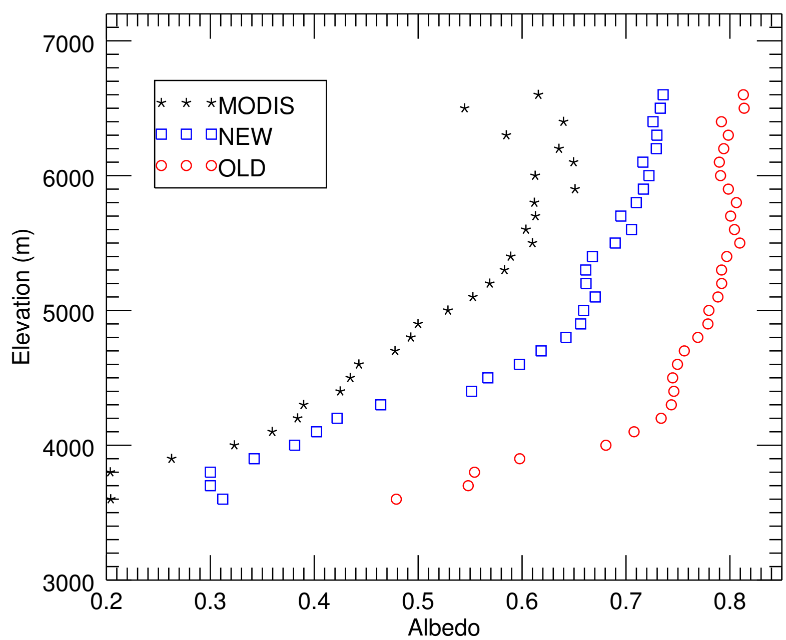

Figure R1: Elevational profiles of snow albedo in MODIS MOD10A1/MYD10A1 (black marker), the CLN simulation in the discussion paper, and the CLN simulation in the revised manuscript, averaged over glacierised grid cells in WRF D3 and the months of July and August. Valid MODIS data with the highest quality flag were used and WRF-CMB data were taken from the corresponding grid cells and time periods for comparison.

The alteration does not change any of the conclusions of the paper; however, it reduces some issues and inconsistencies that were present in the discussion paper, including:

- the maximum number of snow-free debris-covered grid cells exposed over the simulation has increased from ~ 60% in the discussion paper to more than 90% in the revised manuscript, which is more plausible.
- simulated ablation in the debris-free study of Collier et al. (2013) and the CLN simulation in this study are now in much closer agreement.

With the improved simulation of glacier surface albedo, more than 35% percent of debris-covered pixels are exposed between 1 July and 15 September 2004, giving a reduction in mass loss below the zero-balance altitude of 18% (compared with 10% in the discussion paper). Finally, considering the whole simulation period, the reduction in basin-mean ablation by 1 October 2004 in DEB compared with CLN is 14% (compared with 7% previously).

Please find our replies to your comments (**bold**) below, including modifications to the manuscript text (*italics*).

Best regards,
Dr. Emily Collier

General

Given the resulting large uncertainty in debris thickness, the present study must be regarded as a first order estimate of the impact of debris cover on glacier SMB.
We revised the beginning of the first concluding paragraph to read,
"In this study, surficial debris was introduced to the coupled atmosphere-glacier modelling system, WRF-CMB. The model provides a unique tool for investigating the influence of debris cover on both Karakoram glaciers and atmosphere-glacier interactions in an explicitly resolved framework. The first-order impact of debris was estimated, with thickness determined using a fixed gradient of 0.75 cm km^{-1} with distance down-glacier in debris-covered areas, focusing on the period of 1 July to 15 September 2004 when more than 35% of debris-covered pixels were exposed."

Major comments

Only a single summer season is simulated (2004). Show some time series from a lower-resolution atmosphere product that support the assertion that this is a representative summer.
We may not have made this clear in the discussion paper, but we did not mean to imply that 2004 is a representative summer. This year was selected due to the availability of stake measurements on the Baltoro glacier, which are described in *Mihalcea et al.*, [2006], as well as an ASTER-derived debris thickness field [*Mihalcea et al.*, 2008] that we used to evaluate our choice of the debris-thickness gradient.

In the conclusions, we now more emphatically emphasize that our results reflect only this particular ablation season: *"The alterations to the glacier energy and mass fluxes and to*

<u>atmosphere-glacier interactions presented in this study are based on the ablation season of 2004 only</u> and are sensitive to the debris thickness field, with small adjustments to the thickness gradient resulting in significant changes in basin-mean glacier CMB."

To address your comment, we checked monthly mean June-July-August-September (JJAS) 2-m air temperature and total precipitation fields over the Karakoram (~34—38N, 73—78E) in ERA Interim (Fig. R2). The resolution (~ 80 km grid spacing) and snow initialization over glacierised grid cells makes these data potentially unreliable in the Karakoram. However, as quick estimate they indicate that summer temperatures were close to the average in 2004 while precipitation was below average, which is consistent with almost all of the debris grid cells being exposed in the revised simulations.

We added to the final paragraph in Sect. 2.1 about the reanalysis forcing data: *"We note that analysis of summer (June-July-August-September) mean fields over the Karakoram in ERA-Interim indicate that near-surface air temperatures were close to the 1979--2014 mean in 2004, while precipitation was significantly below average."*

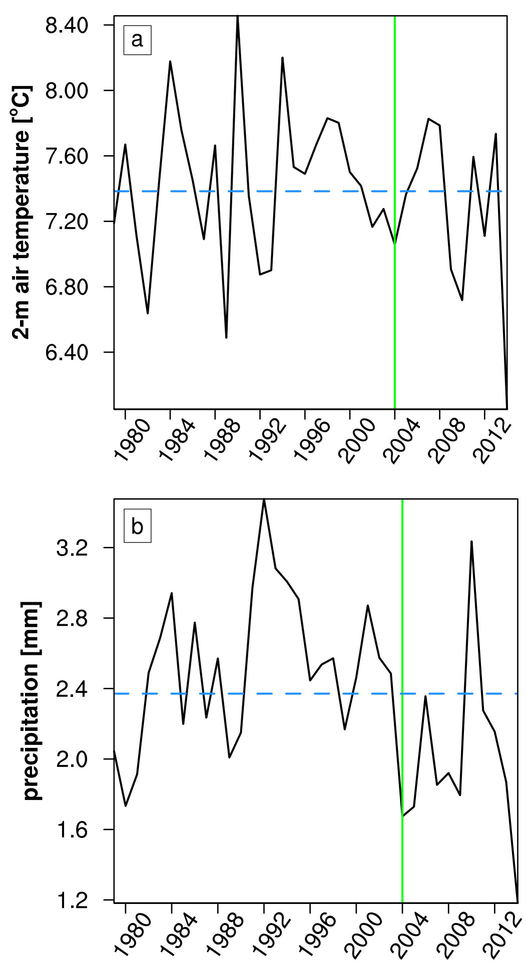

Figure R2: Summer (JJAS) mean fields of (a) 2-m air temperature [^{o}C] and (b) precipitation [mm day^{-1}] taken from the synoptic monthly mean fields in ERA Interim at the 3-hour step. The data were area averaged over ~34—38N and 73—78E. The blue dashed line indicates the mean value between 1979 and 2014, while the green solid line indicates the year of our simulation, 2004.

Page 2266: Explain why both a minimum wind speed and a maximum flux reduction must be introduced to get good results; this fix is rather blunt and I would expect one correction should suffice. And why is using different stability correction not an option, for instance Holtslag and De Bruin (1998)?
We introduced the minimum wind speed in part to be consistent with the approach in the Noah-MP LSM in neighboring non-glacierised grid cells and in part to further reduce decoupling of the surface via zero turbulent fluxes. With the model output we have saved, it's only possible to evaluate the impact of using of both limits for hourly mean (not time step) fields. In Fig. R3, we make this evaluation for the period of 1—10 April 2004. During this period, it's clear that the minimum wind speed correction has only a very small impact on the wind speed and turbulent fluxes, with mean differences (no minimum minus 1 m/s minimum) in windspeed, QS and QL of -0.02 m s^{-1}, -0.15 Wm^{-2} and 0.0 W m^{-2}, although the distribution of QS is slightly more positive with the limit.

We revised the end of the third paragraph of Sect. 2.2 to read, *"Congruent with previous modelling studies of glacier surface energy fluxes, we therefore limit the maximum amount of damping in stable conditions at 30% (Martin and Lejeune 1998; Giesen et al. 2009). In addition, we adopt a minimum wind speed of 1 m s^{-1}, to be consistent with neighboring non-glacierised grid cells simulated by the Noah-MP LSM (Niu et al. 2011). However, test simulations in early April indicate that the second correction has a minimal impact on the wind speeds and turbulent fluxes in glacier grid cells and, thus, may be unnecessary."*

Thank you for your suggestion of a different stability correction. We now conclude that the albedo issue was the main driver of our cold bias, however we plan to test alternative formulations for stability corrections in an upcoming study using eddy-covariance measurements on a glacier.

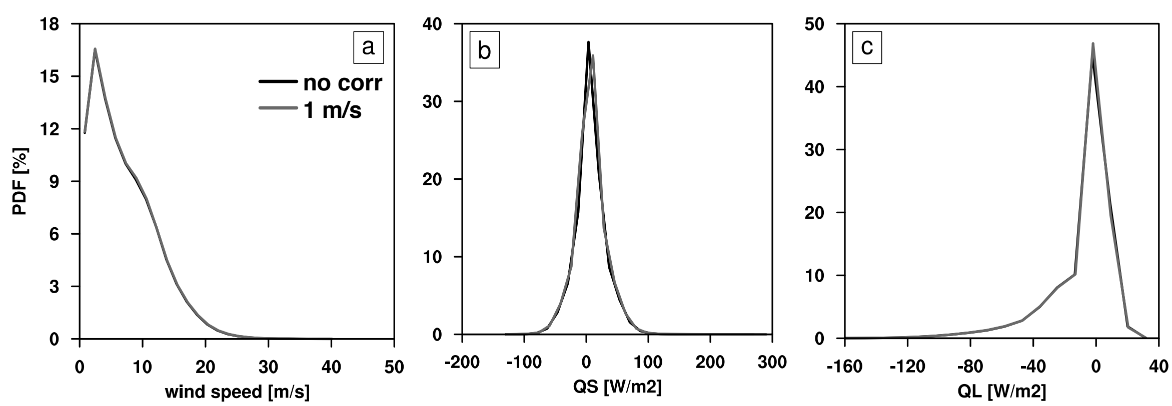

Figure R3: Histogram of hourly mean (a) wind speeds [m s^{-1}] and the turbulent fluxes of (a) sensible (QS) and (c) latent (QL) heat [W m^{-2}] between 1—10 April 2004, comparing a DEB simulation with no windspeed correction (black curve) and one with a minimum wind speed of 1 m s^{-1} (grey curve).

Minor comments

p. 2261, l. 5: suggest: "...a fraction that is approximately twice as large..."
We revised.

p. 2261, l. 12: the -> this
We revised.

p. 2263, l. 2: magnitudes? Please specify the level of agreement.
In Fig. 4a of *Collier et al.*, [2013], we show that both the stakes and model indicate a mass loss of ~O(10 cm) occurred over the observational period (~ 1—15 July 2004), with the exact level of agreement variying between sites. We think the statement that WRF-CMB "*was capable of reproducing the magnitudes of the few available observations of glacier CMB in this region*" reflects this result.

p. 2264, l. 9: suggest: "... were rasterized on a grid with a resolution that was 50-times higher than the original grid spacing of the domain."
We revised.

p. 2264, l. 28: refreeze -> refreezing
We revised.

p. 2265, l. 27: cold/warm temperatures -> low/high temperatures (please correct through MS)
We revised.

p. 2269, l. 11: LST has not yet been defined? p. 2270, l. 21: MB -> CMB
We added the definition of the acronym to the first sentence of Sect. 3.1., and corrected the second acronym.

p. 2271, l. 4: the ELA is defined over the mass balance year, so a 'focus on the ablation season' is no valid argument. The fact that you present a summer SMB profile then means that it is not allowed to call the SMB=0 elevation the ELA. Please adjust.
We replaced all references to the ELA in our simulations with the term "zero-balance altitude." The relevant paragraph was amended to,
"The basin-mean vertical balance profile indicates that between 1 July and 15 September 2004, <u>the zero-balance altitude is located at ~ 5700 m a.s.l. (Fig. 5).</u> For comparison, annual ELAs in the Karakoram are estimated to range from 4200 to 4800 m (Young and Hewitt, 1993). We note that the absence of avalanche accumulation in our simulations may contribute to an overestimate of the zero-balance altitude, as this process is regionally important and produces ELAs that are often located hundreds of meters below the climatic snowline (e.g., Benn and Lehmkuhl, 2000; Hewitt 2005, 2011)."

p. 2273, l. 5: "...in upwards of 800 additional melt hours in DEB compared with CLN." DO you mean melting at the debris-ice interface, or are you comparing surface temperatures here? In that case, the name 'melt hours' is somewhat strange, as the debris surface is warming up rather than melting.

We were comparing surface temperatures. To clarify this section, we removed all references to "melt hours" from Sect. 3.3 and the caption of Figure 9. Instead, we repeated "hours with surface temperatures above 273.15 K" in the relevant paragraph.

p. 2273, l. 27 and further: over a melting ice surface, a convective mixed layer will not develop, rather a shallow, stable (glacier wind) layer. Please reformulate to reflect this.
Thank you for bringing this point to our attention. We improved Figure 11 (included as Fig. R4 below) and revised the paragraph about this figure to,

"Figure 11 illustrates alterations to the diurnal cycles of the turbulent flux of sensible heat (QS), the planetary boundary layer (PBL) depth, and the along-glacier component of the near-surface winds over exposed debris pixels in DEB and their equivalents in CLN. Solar heating of the debris surface drives a strongly negative daytime QS in DEB (Fig.11a), which reduces the stability of the glacier surface layer and enhances turbulent mixing. Peak negative QS values in DEB exceed -200 W m^{-2}, consistent with eddy-covariance measurements of this flux over supraglacial debris (Collier et al. 2014). In comparison, QS in CLN is approximately one order of magnitude smaller and positive. As a result of energy transfer by QS, a deep convective mixed layer develops in DEB, with the mean PBL height reaching nearly 1.5 km in the afternoon compared with only a couple hundred meters in CLN (Fig. 11b). Finally, near-surface along-glacier winds in DEB are primarily anabatic during the day (directed up-glacier, which is defined as positive) and katabatic during the evening and early morning (down-glacier and negative; Fig. 11c), compared with sustained katabatic flows (glacier winds) in CLN, resulting from cooling of the air near the ice surface, which is constrained at the melting point (e.g. van den Broeke 1997). The absence of daytime katabatic flows over debris-covered areas is consistent with the findings of Brock et al. (2010)."

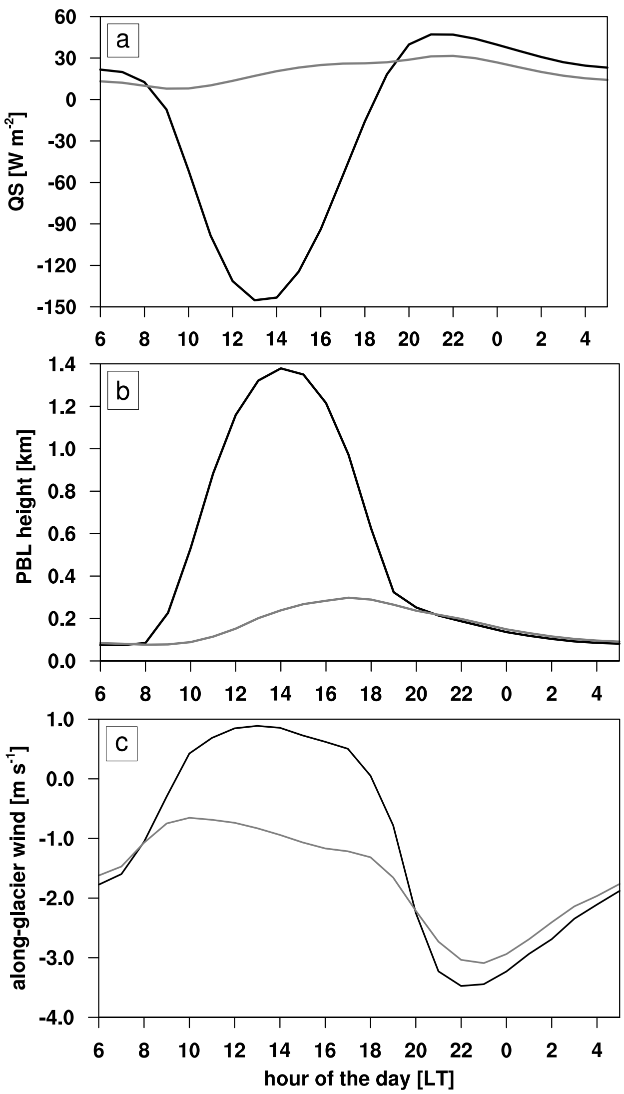

Figure R4: A comparison of the simulated diurnal cycle of (a) the turbulent flux of sensible heat, QS [W m^{-2}]; (b) the planetary boundary layer (PBL) height [km]; and (c) the along-glacier wind speed [from the lowest model level; m s^{-1}], which is positive for up-glacier flow. The data are averaged over exposed debris in DEB (black curve) and the corresponding grid cells in CLN (grey curve).

Table 3: Unit for surface roughness length should be m, not m-1. Figure 2 caption: multipled -> multiplied

We revised.