We would like to thank both reviewers for their careful and constructive reviews of our manuscript. Please find our responses (*italic*) to comments (**bold**) below. Where no response is given, the suggestion was incorporated exactly as indicated into the manuscript.

Reviewer #2

Specific Comments

1) Make it clear in the abstract and introduction that detailed evaluation of model performance using in situ data is conducted only for Baltoro glacier.

Amendments:

- *Abstract (p104, l11): "…with remote sensing data for the ablation season of 2004 as well as with in situ glaciological and meteorological measurements from the Baltoro glacier"*
- *Introduction (p107, l7): "…we aim to (1) evaluate the regional and local performance of the coupled model against available measurements"*
- *Sect. 2 Methodology (p107, l24): "…to coincide with a limited number of glaciological and meteorological measurements on the Baltoro glacier available for evaluation"*
- *Table 3: "Summary of available ablation stake measurements from the Baltoro glacier"*
- *Table 4: "Ablation rates and debris thickness on the Baltoro glacier"*

2) The fact that ablation data are all recorded on debris covered ice, while the mass balance model ignores energy-transfer processes specific to debris covered ice, needs more consideration in the interpretation of these results. Given the 2 cm critical debris thickness (p.123, l15) the model should overestimate melt at all stakes, except L2 and L3 where the model should closely match the measured rate.

Amended discussion paragraph (p123, l20) to read: "… suggesting that on the whole insulation effects should dominate over the lowering of surface albedo except at the sites L2 and L3 where debris thickness is approximately equal to the critical value. Indeed modelled ablation closely matches the measured rate at these two sites and elsewhere is overestimated by WRF-CMB (Fig. 4c), physically consistent with the exclusion of debris in this study."

3) P116, l12-15. There are too few data and too much scatter in Fig. 4d to make this interpretation. In any case, the logic is confused. If the model is working well, it should give a consistent melt overestimation, not a decreasing one over time, because it is ignoring the insulating effect of debris on these stakes. On the other hand, if you are saying the reduction in melt overestimation after 6 days is due to a snowfall event, say exactly when this event happened and which stakes are (most) affected.

Agreed. Figure 4d and the associated sentence have been removed. Unfortunately, we do not have information about the timing or magnitude of precipitation events at the stake locations with which to evaluate whether discrepancies between simulated and observed precipitation play a more important role for the stakes with the shortest observation periods.

4) P115, l14-15, more specifically, the model underestimates the amplitude of the daily temperature cycle, but this would be expected given the surface type (debris) is different from that of the model (ice).

The Urdukas meteorological data is compared with WRF-CMB data from the closest grid point to the location of the AWS (which is located on a moraine ridge adjacent to the glacier). The WRF-CMB grid cell is also non-glaciated and is assigned a soil type of 'loam' (category 6) and a vegetation type of 'barren or sparsely-vegetated' (cat. 19). The underestimation of the diurnal temperature cycle by WRF-CMB compared with the AWS data may be in part attributable to differing 'non-glaciated' surface properties. In addition, to the best of our knowledge, the AWS temperature sensor was not aspirated, which could produce a larger amplitude of the diurnal cycle.

Amended p115, l15: "Finally, the amplitude of the diurnal temperature cycle is smaller in WRF-CMB, which may be attributable to differing thermal properties of the real and modelled land surface and to the fact that the AWS sensor was not aspirated."

5) Wider significance of the results. p.122, l8-9 and elsewhere. The difference in mass balance between the offline and interactive models is quite modest for this study. However, are there situations where interactive coupling might have larger impact on simulated glacier mass balance? Conversely, are their instances where interactive coupling can be ignored?
From this initially brief application, the difference in simulated CMB between interactive and offline runs seems to be small for "clean" glaciers, since the underlying surface characteristics influencing the forcing data are sufficiently similar (≥0.5 m snow in OFF and a mixture of snow/ice in INT). However, as mentioned in the discussion section, the estimated mean debris-covered area for glaciers in the Karakoram is 18-22%. For glaciers with significant debris cover, we expect the inclusion of feedbacks to have a larger influence, as the presence of debris allows the surface to be heated above the melting point and thus drastically alters the sensible and latent heat exchanges. The interactive simulation also represents the more physically-consistent approach. Conversely, given the relatively small influence of interactive coupling for this heavily-glaciated basin, we would likely expect negligible feedbacks on the atmospheric forcing data from small ice masses or less glaciated basins.

Amended Sect. 3.4 (p122, l5): "The basin-scale influence of interactive coupling on the atmospheric forcing data, while moderate, acts to reduce the energy available for surface melt and, in concert with both reduced mass exchanges between the surface and boundary layer and increased refreezing, reduces modelled ablation during the summer of 2004. Furthermore, we demonstrate that the inclusion of additional real processes such as CMB feedbacks renders WRF-CMB capable of simulating observed magnitudes of CMB."

p123 l27: "Given the similarity of the underlying surface types in INT (snow/ice) and OFF (snow) influencing the atmospheric forcing data, the difference in simulated CMB for the clean glacier simulations is relatively small. From the results presented here, it could be expected that the inclusion of feedbacks is not essential for small glaciers or less glaciated basins. However, we would expect the interactive inclusion of the CMB model to have a larger influence for glaciers with significant debris cover, as its presence alters surface temperature and moisture properties and thus turbulent exchanges in the surface boundary layer (e.g., Takeuchi et al. 2000). To assess the role of feedbacks for debris-covered glaciers and to allow the WRF-CMB modelling system to provide long-term, accurate simulations in the Karakoram, including the effects of debris cover on surface conditions and glacier ablation represents important future work."

6) Also for consideration is that many glaciers have debris-covers, particularly in the Karakoram, with variable thicknesses, thermal properties and critical thicknesses. Is this the next challenge for the interactive modelling approach?
Our current work is focused on incorporating the effects of debris cover, with spatially-variable thickness and using typical rock values for thermal properties. As the glacier model solves the thermodynamic energy equation for subsurface temperature, the moderating influence of debris cover on climatic mass balance is prognostic.

Technical Corrections
7) P104, l22, and throughout the paper, data should be referred to in the plural, e.g. 'these data', not 'this data'.
8) P105, l28, 'Traditional approaches. . .'
9) P112, 117, should that be '. . .from longitudinal transects along. . .'?
10) P116, l2-3, was this all solid precipitation? Please clarify.

Changed to "all-phase precipitation."

11) P122, l4, '. . .reduces modelled ablation. . .'.

12) P123,l9-10, '. . .represents an important uncertainty. . .'.

13) P124, l18, replace 'paleo' with 'past'.

14) Figure 2, (b) and (c) are the wrong way around in the figure caption. Are these data for domain D3 only? Please state which domain they apply to.

Figure 2 has been relabeled and the caption amended to "Comparison between WRF-CMB D3, Noah LSM, and MODIS." The WRF-CMB data are taken from domain 3.

15) Figure 4 (d), the axes need to be reversed, as by convention the dependent variable is plotted on the y (vertical) axis.

Figure 4d has been removed. See response to #4.

References

Bryan, G.H., Wyngaard, J.C., and Fritsch, J.: Resolution requirements for simulations of deep moist convection, Mon. Weather Rev., 131, 2394-2416, 2003.

Craig, G.C., and Dörnback, A.: Entrainment in cumulus clouds: What resolution is cloud-resolving?, J. Atmos. Sci., 65, 3978-3988, 2008.

Knievel, J.C., Bryan, G.H., and Hacker, J.P.: Explicit numerical diffusion in the WRF model, Mon. Weather Rev., 135, 3808-3824, 2007.

Kusaka, H., Crook, A., Knievel, J.C., and Dudhia, J.: Sensitivity of the WRF model to advection and diffusion schemes for simulation of heavy rainfall along the Baiu Front, SOLA, 1, 177-180, 2005.

Langhans, W., Schmidl, J., and Schär, C.: Mesoscale impacts of explicit numerical diffusion in a convection-permitting model, Mon. Weather Rev., 140, 226-244, 2012.

Mölg, T. and Kaser, G.: A new approach to resolving climate-cryosphere relations: downscaling climate dynamics to glacier-scale mass and energy balance without statistical scale linking, J. Geophys. Res., 116, D16101, 2011.

Petch, J. C.: Sensitivity study of developing convection in a cloud-resolving model, Quart. J. Roy. Meteor. Soc., 132, 345– 358, 2006.

Takemi, T., and Rotunno, R.: The effects of subgrid model mixing and numerical filtering in simulations of mesoscale cloud systems, Mon. Weather Rev., 131, 2085–2101; Corrigendum, 133, 339–341, 2003.

Takeuchi, Y., Kayastha, R.J., and Nakawo, M.: Characteristics of ablation and heat balance in debris-free and debris-covered areas on Khumbu Glacier, Nepal Himalayas, in the pre-monsoon season, in: Debris-covered glaciers, IAHS Publ., 264, 53--61, 2000.

Zängl, G., Gantner, L., Hartjenstein, G., and Noppel, H.: Numerical errors above steep topography: A model intercomparison, Meteorologische Zeitschrift, 13, 69-76, 2004.