We thank the reviewer for his valuable comments and suggestions. Suggestions and comments are indicated by *blue colored italic* and our replies are denoted by a header [Reply].

1. *The manuscript is on the lengthy side and many long sentences stretch over several lines – these should be shortened or split. For example, the Results are interesting, but too long. It appears that some Discussion has been mixed into the Results. I urge the author to carefully re-read their Results, for example the last paragraph of section 4.2.2.*

[Reply] We rewrote or split some long sentences in the new manuscript. In addition, we reduced the length of Results and moved some discussions of the Results section into the Discussion section.

2. *The general approach of this work is scientifically sound, but it appears that there are shortcomings in the detail. For example, the calculated mass balances are not compared to field measurements (or DEM differences), the meteorological data used as boundary conditions for the glacial-mass balance model are sparse and/or have low spatial resolution. If the goal is to calibrate the impact of debris cover on glacial mass balance, a different study region with denser data may be more appropriate. There is little validation of existing meteorological or glacial data available.*

[Reply] The limitation of meteorological and glaciological data is the main challenge for glacier study on Mount Gongga, as well as in other glacierized regions of the world. However, our approach was calibrated by observed runoff from the HLG catchment (Fig. 1 of the manuscript). It is one of the glacierized catchments of Mount Gongga, containing three debris-covered and four debris-free glaciers. Then, the calibrated approach was verified against an independent set of runoff data and ablation data observed in different periods. In particular, the precipitation gradient, one of the important parameters of the model, was optimized by minimizing the difference between observed and modelled runoff in different years according to three assessment criteria, and the temperature lapse rate is confirmed by recent observations. Overall the approach performs well and can simulate the observations accurately. In addition, although the meteorological data used in this study have the low spatial resolution, the daily, monthly, seasonal and annual meteorological data correlated well with the corresponding observations from GAEORS (Fig. 1 of the manuscript), which is the only observation station near the glacierized regions on Mount Gongga. In order to easily understand the model performance, we added the sentences about the calibration and validation of the approach in the 3.3 section of the manuscript, which is '*The model was calibrated by the observed runoff in different years according to three assessment criteria, and then was verified against an independent set of runoff and glacier ablation data over different periods in the HLG catchment (Fig. 1), one of the glacierized catchments on Mount Gongga (Zhang et al., 2012). The model performed well and simulated the observations accurately*'. Furthermore, in the In the 5.2 section of the manuscript, we discussed the uncertainties and the overall performance of simulated mass balance using this approach and the meteorological data. Overall, we believed that the model can present a reasonable view of the general condition of the mass balance of Mount Gongga glaciers.

3. *I am little perplexed by the general conclusion (impact of debris coverage) and several statements in the Results and Discussion that highlight that debris-free glaciers are retreating about half as fast as debris-covered glaciers. Can these glaciers be easily compared? Do they have similar AAR, hypsometries, mass balances, etc?*

[Reply] We attempt to highlight the overall differences in termini retreat between debris-covered and debris-free glaciers in this region, rather than highlight that debris-free glaciers are retreating about half as fast as debris-covered glaciers. Hence, we rewrote some sentences of this section to avoid misunderstanding, and highlight '*there are significant differences in the termini retreat between debris-free and debris-covered glaciers on Mount Gongga*'.

4. *How accurate are the ELA (or snowlines?) form the CGI? The glacial areas are likely to be accurate.*

[Reply] Most of glaciers recorded in the CGI have no observed ELA or snowline (Shi et al., 2005). Hence, most of ELAs or snowlines in the CGI are estimated from aerial photographs and topographic maps (Pu, 1994; Shi et al., 2005). The method used in the CGI was evaluated by comparing ELA estimates to observed data in different regions of the Tibetan Plateau. The results indicated that there is about ±50 m difference in the ELA, and the maximum difference is less than 100m (Pu, 1994; Shi et al., 2005). In this study, the ELA from the CGI is used to determine the extent of the ablation zone. The observations and Fig. 2 of the manuscript indicated that the debris-covered area in the upper reaches of the glacier is small. Hence, the uncertainty coming from ELAs is likely to have little influence on our results.

We inserted the clarification for the accuracy of the ELAs of the CGI, which is '…, *which were estimated from aerial photographs and topographic maps (Pu, 1994). The error in ELA estimates of the CGI is about ±50 m on average (Pu, 1994; Shi et al., 2005)*'.

5. *Why are thermal resistance field measurements unrealistic (page 2422, line 1)? They are very valuable for calibration and validation purposes, especially if the field site is chosen carefully and is representative for a satellite/DEM gridded dataset. Some data are shown in Figure 3, where the authors present a validation of ground-surveyed debris thickness with ASTER-derived thermal resistances. A more meaningful comparison would by a ground-surveyed debris thickness (X-axis) vs. ASTER-derived thermal resistances (Y-axis) plot. While the pattern is convincing, the debris thickness measurements have much less variability than the thermal resistance values.*

[Reply] As the reviewer suggested, ground-surveyed the thickness and thermal properties of debris layers are very valuable for calibration and validation purposes on a typical glacier. However, it is very difficult to determine a large-scale spatial distribution of the thickness and properties of the debris cover, especially at a regional scale. We did the comparison of the ground-surveyed debris thicknesses (X-axis) and ASTER-derived thermal resistances (Y-axis) on HLG Glacier in the study of Zhang et al. (2011; Figure 5(a)). Due to a few surveyed data on another two glaciers (DGB and XGB glaciers), we cannot do such comparison. On the other hand, the debris thicknesses were surveyed by excavating the debris cover until reaching glacier ice, and the ASTER-derived thermal resistance represents the value of a grid cell of 90m, which may contains more than one surveyed points. This leads to the less variability in debris thicknesses compared to the thermal resistances, but the comparison results indicated that the ground-surveyed debris thicknesses and ASTER-derived thermal resistances correlated reasonably well over the entire ablation area. We added the sentences for the comparison results between ground-surveyed debris thicknesses and thermal resistances as discussed above.

6. *I am uncertain if the NCAR reanalysis data provide the spatial resolution to estimate net radiation at the required scale. In other words, unless some downscaling has been performed, the entire study area is covered by one NCEP/NCAR reanalysis grid cell. Maybe NCAR/NCEP data can be used as boundary conditions for a WRF model to derive sufficiently accurate net radiation measurements – but even that will require ground validation. I emphasize that the authors mention the pixel size of the ASTER TIR data, but do not refer to the spatial resolution of the NCAR data. Generally, when referring to reanalysis data, a more precise citation or reference to the NCAR version is necessary (as well as the used grid-cell size).*

[Reply] The aim of this study is to present a simple, physically-based approach that can require as less input data as possible to well understand the debris-cover effects at a large scale in the Tibetan Plateau, where meteorological and glaciological observations are scarce in glacierized regions. Hence, we used the average values of the shortwave and longwave radiation in a certain time from NCEP/NCAR reanalysis 1 (Kalnay et al., 1996). Due to the limitation of observed data, it is difficult to make the bias-correction or downscaling for NCEP/NCAR reanalysis 1, especially for solar radiation. More importantly, thermal resistances estimated from ASTER and NCEP/NCAR reanalysis data correlated reasonably well with ground-surveyed debris thicknesses at the large scale, and can reflect the spatial patterns of ground-surveyed debris thickness (Zhang et al., 2011). In addition, we calculated the thermal resistance of the debris layer in thickness of 10 cm based on the empirical equation derived from the relationship between ground-surveyed debris thicknesses and ASTER-derived thermal resistances (Zhang et al., 2011; Figure 5(a)). The thermal resistance for debris layer in thickness of 10cm is 0.0192 m^2 K W^{-1}, and the thermal conductivity (ratio of debris thickness and thermal resistance) is 5.19 W^{-1} m K^{-1}. These results are the similar order with those of Lambrecht et al. (2011), who calculated the thermal resistance in a 10 cm thick layer for different rock types and thermal conductivity of the respective material. This can partly assess the reliability of the NCEP/NCAR reanalysis data as input data.

According to the reviewer's suggestions, we inserted a more precise citation and introduction for the NCEP/NCAR version (NCEP/NCAR reanalysis 1) in the 2.2.2 section of the revised manuscript, which is '*In addition, the surface downward radiation fluxes from US National Centers for Environmental Prediction (NCEP)/US National Center for Atmospheric Research (NCAR) reanalysis 1 (Kalnay et al., 1996) were used to calculate the thermal resistance of the debris layer, which correspond to the nearest time and location of ASTER acquisition. The full details of the NCEP/NCAR project and the dataset are given by Kalnay et al. (1996)*'. Also, We inserted the discussion about the uncertainty caused by NCEP/NCAR reanalysis data in the 5.1 section as discussed above.

7. *Along the same lines, I argue that the precipitation gradient can not be interpolated into large heights. There exist several studies analyzing the spatial and elevation-distribution of rainfall that document a rainfall peak at ~4km, but less moisture above. I note that the ground-control stations to derive some of the gradients do not include stations at high altitude (i.e. > 5km). Since there are no mass-balance measurements presented and only modeling results are given, it is crucial to have reliable boundary conditions for the mass-balance model.*

[Reply] As the reviewer said, in the mountain ranges, there is a rainfall peak at ~4km, and less moisture above. However, precipitation increases linearly with altitude in Mount Gongga, even in the high elevation area (> 5000m). This pattern has been confirmed in numerous studies (Aizen and Aizen, 1994; Cheng, 1996; Li and Su, 1996; Liu et al., 2010). Based on the results of these studies, the precipitation gradient was optimized within the accepted ranges derived from precious studies (Aizen and Aizen, 1994; Cheng, 1996; Liu et al., 2010) by minimizing the difference between observed and modelled runoff in different yeas, and then was used to estimate the precipitation of each elevation band. Although very few meteorological stations are operating in this region, resulting in significant climatic uncertainties, the calibrated model was verified against an independent set of runoff and glacier ablation data over different periods, and performed well and simulated the observations accurately.

We inserted the sentences for introduce the information of precipitation, which is '*The precipitation increases linearly with increasing altitude in this region (Aizen and Aizen, 1994; Cao and Cheng, 1994; Cheng, 1996; Liu et al., 2010). We applied a precipitation gradient (% of precipitation increase per meter of elevation increase; Table 2) from the snout to the top of the glacier, which was optimized by Zhang et al. (2012)*'.

8. *One additional note regarding mass balances: The authors use a stereo-airphoto DEM from 1989. This could easily be combined with the 2000 SRTM DEM to give a first- order decadal-scale mass balance. This approach could be used to roughly verify model results.*

[Reply] Based on the DEM of 1989 and the SRTM DEM, we can easily calculate the glacier surface change during the period from 1989 to 2000, but it is difficult to assess the uncertainties of such results. For the SRTM DEM, many uncertainties emerge from the unknown radar penetration and the interpolated topography in the voids of the original data (Frey and Paul, 2012), and the increasing roughness of the terrain leads to more frequent data voids (Hall et al., 2005). Here we compared the peak elevation of Mount Gongga (glacier-free zone) between the 1989-DEM and the SRTM DEM, and found that the differences is more than 450 m. Consequently, such uncertainties for the SRTM DEM can strongly affect the accuracy of estimated surface elevation change. On the other hand, we only calculated the mass balance for the period 1998-2007. Hence, we did not use the mass change estimated from the 1989-DEM and SRTM DEM to roughly verify the model results.

9. *The mass balance setup (equation 2) is rather simple and the model of accumulation (e.g., snow/rain mixture at threshold temperature) is only moderately applicable, but certainly a first-order approach.*

[Reply] The mass balance setup in this study is identical to the study of Zhang et al. (2012), and is described only briefly here. Most of mass balance studies used the temperature threshold to determine whether precipitation falls as rain from snow. It may be a first-order approach for modeling accumulation, but gets the acceptable performance of modeling accumulation on the glaciers due to the limitation of observed precipitation data in the glacierized regions.

10. *Figures: Figure 4 is interesting, but I find it extremely cumbersome to read with the colorscale. The colorscale should be simplified and should only contain 3-5 colors (max. of 5).*

[Reply] We did it, and also revised the colorbar for Figure 5 and Figure 6.

11. *Wording: Conclusion: replace 'considerably significant' with significant.*

[Reply] We did it.

References

Aizen, V.B., Aizen, E.M.: Regime and mass-energy exchange of subtropical latitude glaciers under monsoon climatic conditions: Gongga Shan, Sichuan, China. J. Mt. Res. Dev., 14(2), 101–118, 1994.

Cheng, G.: Exploration of precipitation features on extra-high zone of Mt Gongga. Mt. Res. 14(3), 177–182, 1996 (In Chinese with English abstract).

Frey, H., Paul, F.: On the suitability of the SRTM DEM and ASTER GDEM for the compilation of topographic parameters in glacier inventories. International Journal of Applied Earth Observation and Geoinformation, 18, 480-490, 2012.

Hall, O., Falorni, G., Bras, R.L.: Characterization and quantification of data voids in the shuttle radar topography mission data. IEEE Geoscience and Remote Sensing Letters, 2 (2), 177–181, 2005.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R. and Joseph, D.: The NCEP/NCAR 40-year reanalysis project, B. Am. Meteorol. Soc., 77 (3), 437–471, 1996.

Lambrecht, A., Mayer, C., Hagg, W., Popovnin, V., Rezepkin, A., Lomidze, N., Svanadze, D.: A comparison of glacier melt on debris-covered glaciers in the northern and southern Caucasus, The Cryosphere, 5, 525-538, 2011.

Li, J. and Su, Z.: Glaciers in the Hengduan Mountains, Science Press, Beijing, 1–110, 1996 (in Chinese).

Liu, Q., Liu, S., Zhang, Y., Wang, X., Zhang, Y., Guo, W. and Xu, J.: Recent shrinkage and hydrological response of Hailuogou glacier, a monsoon temperate glacier on the east slope of Mount Gongga, China, J. Glaciol., 56, 215–224, 2010.

Pu, J.: Glacier Inventory of China VIII (the Changjiang (Yangtze) River Drainage Basin), Gansu Culture Publishing House, Lanzhou, 117–129, 1994 (in Chinese).

Shi, Y., Liu, C., Wang, Z., Liu, S., Ye, B.: A Concise China Glacier Inventory. Shanghai Science Popularization Press, Shanghai, pp. 8-9, 2005 (in Chinese).

Zhang, Y., Fujita, K., Liu, S., Liu, Q., Nuimura, T.: Distribution of debris thickness and its effect on ice melt at Hailuogou Glacier, southeastern Tibetan Plateau, using in situ surveys and ASTER imagery. J. Glaciol., 57(206), 1147–1157, 2011.

Zhang, Y., Hirabayashi, Y. and Liu, S.: Catchment-scale reconstruction of glacier mass balance using observations and global climate data: case study of the Hailuogou catchment, south-eastern Tibetan Plateau, J. Hydrol., 444–445, 146–160, 2012.