The Cryosphere Discuss., 8, C2204–C2211, 2014
www.the-cryosphere-discuss.net/8/C2204/2014/

The Cryosphere
Discussions
Open Access

Interactive comment on "Modelling glacier change in the Everest region, Nepal Himalaya" *by* J. M. Shea et al.

B. Marzeion (Referee)

ben.marzeion@uibk.ac.at

Received and published: 2 November 2014

Shea and colleagues present a model of glacier mass balance and ice redistribution and apply it to the glaciers of the Everest region. The manuscript is well organized, and the results are generally well presented. The research is clearly relevant, and the authors are following a very promising approach. However, I find three major issues that need to be addressed to ensure that the conclusions are actually supported by the results. Additionally, I have quite a few smaller suggestions that probably can be taken into account with relatively little work.

General comments:

- I do not understand how the initialization run is working (P5386, L13-19): Assuming that the applied T anomaly and the resulting mass balances are correct, how do you get the ice thicknesses and extent in 1961? Do you apply the mass balances until extents (and ice thicknesses) reach an equilibrium? How long does this take? Or if you do something else, please describe. I think this is not a trivial problem, because the lagged response of glacier extents to mass balance changes will affect the following calibration runs; i.e. how you create the state in 1961 will impact how the glaciers behave in the following decades, additionally to the different parameters you use. Of course, also the parameter set you use for the initialization run will play a role (e.g., you will see changes in area even if you keep the forcing constant, but change R). Or you could imagine a situation where the glacier extent in the calibration run is not yet adjusted to the ice thickness changes, and the resulting changes in the following calibration run are a mixture of adjustment to the new parameter set, and the equilibration between extents and ice thickness.
- The "Monte-Carlo"-type calibration: during the calibration, there are 6 degrees of freedom (i.e., 6 parameter values are being optimized). 20 parameter sets to choose from is then a very low number (e.g. if you want to cover every combination of parameters with just two values for each parameter, you would need $2^6 = 64$ parameters sets/calibration runs). My question therefore is: what part of the parameter space is covered in the 20 runs, how did you decide on the parameter sets you tried, and how do you make sure that the "best set" is within the part of the parameter space that you cover? Fixing this could be numerically very expensive (e.g., covering the parameter space with 3 parameter choices in each direction leads to the order of 1000 calibration runs) – but would it be possible to do the validation for each of the 20 calibration runs? This could give a good indication on the robustness of the calibration (i.e., one would assume the four runs with the best overall scores in the calibration to also perform well in the

validation)?

- The rationale behind the application of the CMIP5 scenarios is unclear. Why use linearized anomalies instead of the full anomalies that are readily available? Linearizing symmetrically around the year 2000 will have very significant consequences for the estimated differences between the scenarios (e.g., T change is not at all linear, and the different rates of T change from 1975 to 2035 do not reflect the difference between the RCP4.5 and RCP8.5 scenarios – this probably explains why the two scenarios are practically identical in Table 5, and hardly different in Fig. 14), but also between the ensemble members of each scenario. It should be straight forward to take the full T and P time series, and I strongly recommend to do so.

Specific comments:

- In the abstract: clarify the difference between the geodetically derived estimates of net glacier mass change you use for calibration, and the remotely-sensed observations of decadal glacier change you use for validation.
- Not sure about TC policy, but I would try to avoid "complicated" abbreviations like APHRODITE and EVK2CNR in the abstract – or at least introduce them here as you did with "ELA".
- P5376 L16: I'm not sure what you mean by calling the RCP4.5 and RCP8.5 scenarios "end members"?
- P5377 L15: Fig. 2 is referred to before Fig. 1.
- P5379 L10: abbreviation "ELA" already introduced in the abstract.

- P5379 L24: change "Results from CMIP5 ensembles..." to "Results from the CMIP5 ensemble...". Also: introduce the abbreviation CMIP5 (earliest mentioned in the abstract).
- P5380 L9: see previous comment
- P5381 L10: it is unclear here what is meant by "downscaling ensemble", as the CMIP5 data base to my knowledge does not contain downscaled ensemble members. I think this term "downscaling" is not appropriate here. It could either be used to describe dynamical downscaling as e.g. in the studies of Dimri et al. and Kulkarni et al. (both cited in the manuscript), or it could be statistical downscaling. Instead, you are simply applying the anomalies from the CMIP5 ensemble members.
- P5381 L14: change "year" to "years".
- P5382 L2: 14 x 14 grid cells?
- P5382 L4: are there any significant gaps in the SRTM DEM in the area? If so, how do you treat them?
- Fig. 3, caption, L3: delete "for".
- Fig. 3, bottom panel: I assume all points are significant? Please indicate.
- Fig. 3, upper and middle panel: add $^\circ$ to vertical axis.
- Fig. 4: add $^\circ$ to vertical axis.
- P5382 L16: introduce the abbreviation "EVK2CNR" (earliest mentioned in the abstract).

- Fig. 5: the daily precipitation almost certainly is not normally distributed, and therefore showing +/- 1 standard deviation is not that helpful. Better show percentiles (e.g., 10th, 50th (median) and 90th).
- Fig. 6, caption: change "modelled" to "predicted", both for consistency with the axis label, and since you did not really model precipitation.
- P5383 L4: Solid precipitation may (and often does) occur at temperatures slightly above freezing – have you tried threshold values of 1°C? Also, please add the unit here and in Fig. 6.
- Eq. 4: perhaps you can explain a bit the rationale behind this - why would you expect highest ablation rates on flat surfaces? Shouldn't this be a function of the day of the year (assuming that you want to capture short wave radiation effects here), i.e. a measure of exposure relative to the sun?
- Eq. 4 and/or Table 2: please ensure consistency in notation, i.e. K (Eq. 4) vs. ddf in Table 2.
- P5384 L8: How was $R_{exp} = 0.2$ determined?
- I don't think Table 1 is referred to in the text.
- Figures generally, but particularly 1 and 10: the labeling is really small, please increase the font sizes a bit.
- P5385 L3: "roughness" (typo).
- P5385 L17: "steeper cells receiving a greater share of the ice flux": how exactly is this partitioned?
- P5286 L1: "daily timesteps using a 0.2 m w.e. threshold, which represents the average seasonal snowfall depth": why are you using a threshold? I'm not sure I get what you want to say here.

- Eq. 8: please refer to Eq. 6, and add an explanation why you are using τ_0 instead of τ_b here.
- P5386 L23-27: I find the argument for a special Khumbu glacier melt factor weak: what constitutes "anomalous ice flow velocities"? The model will always show different success at different glaciers, so adapting special parameters for individual glaciers seems arbitrary (and potentially distorts the validation: what if there are more "anomalous ice flow velocities" which you don't recognize, because of lacking observations?).
- P5387 L15-18: I think this score is biased – or what would happen if all the glacier polygons are ice covered, but additionally many more? Perhaps you did it, but didn't mention it, that there also should be a penalty for ice cover outside the polygons.
- P5387 L18-20: I don't understand how this score is calculated, perhaps just because of the example numbers given are a bit ambiguous: If the mass balance during 1992-1998 had been 0, what would the score of a modeled mass balance of -0.02 be? Perhaps give equations for the scores.
- More generally, concerning the scores: why multiplying instead of adding the scores? By multiplication, a single very good score will lead to a very small total score even if the three remaining scores are disastrous.
- P5390 L3: introduce "SD".
- Fig. 8: add C to vertical axis of lower left panel.
- P5390 L29: add unit to T.
- P5391 L4&5: change "modelled" to "predicted".

- P5391 L18-19: so you actually do weight the scores (cf. P5387 L14), because you don't take run with the lowest overall score. Perhaps this would change with an additive overall score?
- P5391 L21-22: I think you can't call the parameter set of run 5 the "final calibrated degree-day factors", because you had on average less than two values to choose from (see general comment above).
- It's such a pity your model run goes to 2007 when the observation start in 2008... any chance on having an updated atmospheric data set (not necessarily now, but later)?
- P5393 L20: please mention the minimum slope already somewhere close to Eq. 8.
- P5394 L16: now I understand the threshold mentioned earlier (P5286 L1) – please add an explaining sentence there.
- P5395 L2-5: Depending on how the initialization run was made, I can also imagine this as an explanation for the differences (particularly since they are larger in the first decade, but that could of course be coincidence...).
- P5395 L7: perhaps rather say something like "extracted from members of the CMIP5 ensemble that capture the range of results in the respective scenarios".
- P5397 L5: "ensemble" (singular).
- P5397 L17: please replace "highly significant" with "good" or something similar, or "matches" with "correlations" – which would slightly change the meaning as well.
- P5397 L18: change "model" to "predict" or "downscale".

- P5398 L10-17: here would be a good place to also (qualitatively) discuss the impact of having a stationary debris cover in your model – will this lead rather to an over- or underestimation of the projected changes?
- P5398 L22-24: I don't think this is a good justification for the linearization, as (i) it would also be possible to keep natural variability constant without linearizing (e.g., apply ensemble mean anomalies); and (ii) using an anomaly coupling of the full time series will retain the characteristics of the observed monsoon, by removing the biases that the GCMs have in the present day.
- P5399 L16: b_t instead of b_a, and why > 0 if in Eq. 13 is $-H$?
- P5399 L21&22: "years" (plural).

Interactive comment on The Cryosphere Discuss., 8, 5375, 2014.