Author's response to the comments received for tc-2021-31

The following pages contain a point-by-point reply to the comments provided by the two referees that reviewed our first submission and to by the editor that commented our second submission.

*Each of the referee's comment (**RC**) and editor's comment (**EC**) are numbered. If a comment contained several points, we numbered them, and address them individually in our author replies (**AR**).*

EDITOR - ANDREAS VIELI

[EC 0.01] Comments to the author:
Dear Loris Compagno and co-authors,
Your manuscript received two very detailed and critical reviews but that both highlighted the novelty and innovation in the treatment of debris in regional/global scale glacier models and the related extensive modelling results/investigations and they confirm that the manuscript would be a very valuable contribution to TC. Although both referees were in general very positive, they also raised, besides the minor technical corrections, also a few major issues that should be improved before publication. In brief the most important points encompass
a) clarifications and more details in the methods (in particular the debris thickness data and the parametrized oestroem curve
b) issue of omitting dynamic debris thickening/thinning (velocity) effects (add at least a discussion on this)
c) better evaluate the role of debris cover inclusion (omitting, explicit, implicit)
d) more accurate reference to literature

Given the response of the authors, it seems that the authors seem to be able to address the points (or already have done so) raised by the referees well and that the revised manuscript has a good chance to be accepted. I therefore ask the authors to undertake the planned revisions and submit the revised document.

[AR 0.01] We thank the editor for the positive feedback. We will send the revised manuscript as instructed. Here below we have addressed all issues raised by the editor. The manuscript was updated accordingly.

[EC 0.02] A few more notes from me as the editor to your suggested revisions to the above more major points:

to a)
With regard to the requested more detailed explanation of the used oestroem curve parametrization by referee 1, there are still a few points unclear or missing in your suggested revisions that should maybe be already considered:
(i) - when explaining how the debris thickness maps of McCarthy were obtained it would be useful to mention the basic principle behind (e.g. inversion of mass conservation equation from remote sensing data (velocity fields, dh/dt), rather than just refer to the literature.
(ii) - clarify that parameters i_debris and k_debris are parameters that are determined FOR EACH GLACIER (I would actually also wonder what values you would get there, range, mean, median?).
(iii) - it would also be more explicit what the physical meaning of these parameters (in particular k_debris) are (in words). I believe k_debris is similar to the d_0 value of Anderson 2016 (some sort of reference debris thickness), whereas as i_debris is a pure calibration factor.

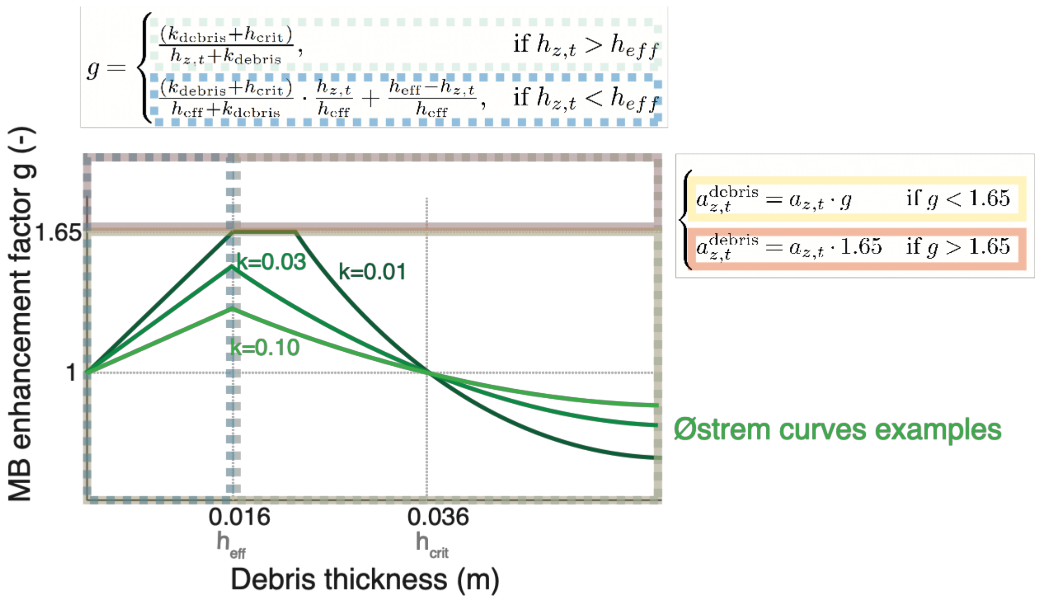
(iv) - the new Fig. S1 is very useful, but add in figure (or caption) what the parameter values are to it (I believe all green curves use the same h_eff and h_crit but a different k_debris (how variable?).
(v) - the supposedly newly added 'unceratinty' in debris thickness in fig. 1 b)-d) is not really explained in the caption, I just see a +? and a -? Number after the debris thickness value, what do they refer to? The range? The uncertainty in both directions? Calrify at least in caption.

[AR 0.02]
(i) We added an explanation for the basic principle by which the debris thickness maps were obtained:
ll. 98-100: *'In a nutshell, the inversion procedure iteratively solves for the debris thickness by using an energy balance model. The procedure uses DEMs, glacier ice thickness, surface velocity, debris proprieties, and meteorological forcing data as input, and uses them to calculate ice flux divergence and ice thinning rates. The debris thickness is then adjusted until modelled and observed ice-melt rates agree within a prescribed tolerance.'*

(ii) We added the information asked by editor:
ll. 110-112 *' […] where b is the local surface mass balance, and i_{debris} and k_{debris} are glacier-specific calibration parameters without specific physical meaning. The the mean and 95% confidence interval for i_{debris} are -1.86 and [-7.62 , -0.09], respectively. For k_{debris} the equivalent values are 0.10, and [0.01, 0.22].'*

(iii) The parameters i_{debris} and k_{debris} do not have a physical meaning. They are calibration factors used to fit the results of the energy balance model. We added this information in the manuscript (cf. AR 1.02 ii).

(iv) The editor is right, h_{eff} and h_{crit} are always the same whilst k_{debris} is glacier-specific (cf. AR1.02 ii). We added the value of k_{debris} in the Figure 3 as request by the editor.



'Figure 3: Schematic of the melt enhancement factor g (dimensionless) as a function of debris thickness for three different glaciers (green lines; "k" is the value of k_debris calibrated for each glacier). The colored, dashed boxes show regions in which the different cases of Eq. 2 and 3 apply.'

(v) The values shown in Figure 1 (b-d) define the confidence interval estimated for the mean debris thickness of each glacier. We updated the caption as follows:

Fig1: '*[…] h_{debris} is the mean debris-cover thickness with superscript and subscript values indicating its estimated confidence interval (note that the latter is not symmetric; cf. Section 2.2) […]*'

[EC 0.03] to b)
with regard to the neglectance of dynamic contribution to debris thickening/thinning (Referee 2) some discussion on additional uncertainties has been added. But this could be a bit more explicit in saying what the issue is (neglecting dynamic effects of debris thickening/thinning) and why it may be justified (indirectly included, etc… see you response to the review) to do it your way. It could just be a bit more on the point.

[AR 0.03] We slightly revised the manuscript, clarifying both what the issue is (neglecting dynamic effects of debris thickening/thinning) and that this effect is only implicitly accounted for in our approach. Similarly, we now explain why we only accounted for it implicitly, instead of explicitly. In particular, we modified the following sentences:

l 257-258: *'Combined with $b_{(z,t)}$, the long-term glacier-wide mass balance $B_{(t-9,t)}$ implicitly mimics ice-dynamical processes (see section 7.2)'*

l 504-509: *'Explicitly accounting for the dynamics of debris re-distribution (as opposed to implicitly, see Section 3.2.3) could be an alternative option. Indeed, where the ice flow velocity decreases, the local debris thickness increases and vice versa, due to the convergence of debris particles. Such an approach was followed by Anderson et al. (2021b), for example. We decided not to include such effects explicitly because (1) the absence of data to calibrate and validate a more complex parameterization at regional scales, and (2) the small sensitivity of regional-scale glacier volume and area to changes in debris cover (see Section 7.1).'*

REVIEWER 1 - BEN MARZEION

[RC 1.00] (i) Compagno and co-authors introduce parameterizations for the evolution of debris cover distribution and thickness into a glacier model, applicable on the large regional and global scale. They derive the parameterizations, calibrate parameters based on observations (some of them depending quite strongly on another model), and evaluate the parameterizations using independent observations. They apply the model for projections of glacier evolution in the 21st century in High Mountain Asia.
There is no doubt that the manuscript improves the state-of-the-art of how debris cover is represented (if at all) in glacier mass balance models applicable to the regional and the global scale. The advances presented in the manuscript clearly contribute significantly to our understanding of how relevant debris cover is in shaping the future of glaciers. The manuscript is generally well written and the results are generally presented well, but I also have a relatively large number of minor or technical comments and suggestions.
(ii) There are, however, a few issues with the manuscript that are more substantial. I believe that the authors will be able to address them, and I don't believe that the main conclusions of the manuscript will change. But since they could only be addressed in extensive revisions to the text and/or additional analyses (see below for details), I consider them major.

[AR 1.00] (i) We thank the reviewer for the very positive feedback and for the review.
(ii) Below, we have addressed all issues raised by the reviewer. The manuscript was updated accordingly.

Major comments

[RC 1.01]- The estimation of debris cover thickness (L91-99) is very unclear to me:
(i) why is "observation" in "observations-based mass balances" in quotes? **(ii)** how is the energy balance model that you apply on each glacier calibrated, ie., and how are the glacier-specific Østrem evaluated? **(iii)** What is the reasoning behind Eq. 1, and the meaning of the "free parameters" i_debris and k_debris? I appreciate that you cannot repeat the manuscript of McCarthy et al. here, but the description needs to be understandable in principle without going to the reference (even if it was accessible to readers, which it is presently not). **(iv)** Probably I just don't get it, but I am also left a bit puzzled why a temperature-index melt model is applied for the projections if the authors have an energy balance model that can deal with debris cover and is applicable for each glacier, and which the authors trust so much as to not only estimate debris thickness, but additionally glacier-specific Østrem curves.
(v) I have a similar difficulty following L154-166: may the problem is that it remains unclear whether the goal of the equations is to mimic a physical understanding of the debris effect (such that it is possible to explain the "meaning" of the different parameters), or to parameterize the shape of the Østrem curve. Could, e.g., g be called a "melt modification" parameter (or "enhancement factor", as in Fig. S1) for the temperature index parameter? And is the goal of Eq. 3 to create the shape of the Østrem curve for g? **(vi)** It might be helpful to include an example of a parameterized Østrem curve with labels for the different threshold and critical values of h (i.e., a schematic version of Fig. S1 including the names of the parameters – and potentially 2 or 3 different curves for different parameter values).
(vii) Finally, more needs to be said on the uncertainties of the estimated debris thicknesses. Otherwise, it is very hard to make sense of the relevance of e.g., the thickness differences presented in Fig. 5.

[AR1.01] (i) 'Observations' was in quotation marks because the SMB data of Miles et al. are not observations in the traditional sense, i.e. they are not derived from measurements made in the field or directly from satellite imagery. Instead, they are calculated from geodetic mass balance estimated from satellite data (Brun et al., 2017) by solving the continuity equation for each glacier. In addition to the geodetic mass balance data, this also requires ice thickness (taken from Farinotti et al., 2019) and surface velocities (taken from Gardner et al. 2019). These new estimates of altitudinal resolved glacier mass balance (which we referred to as observations) are then used in McCarthy et al. to calculate debris thickness. Stated differently: the approach by McCarthy et al. uses the estimates as observations that the inversion aims at matching. Since we recognize that the text was somewhat unclear, we reformulated it. In particular, we now avoid the wording "observation" (see iii).

(ii) The energy-balance model in the study by McCarthy et al. (in review) is not calibrated: it is a physical model with physical parameters (e.g. conductivity) for which we use values reported in the literature, and a model that requires meteorological variables as input. We recognize that both the physical parameters and the input meteorological variables have uncertainties. This uncertainty is estimated in a Monte Carlo framework in which both are perturbed within their expected uncertainty ranges. This allows for generating the mass balance data which are in turn used to determine the Østrem curves. The Østrem curves themselves are not validated since the limited observational evidence precludes it. Instead, we validate the calculated debris thicknesses using all available in-situ debris thickness data with satisfactory results (see McCarthy et al., in review). Note that the work by McCarthy et al. is now available on a preprint server (https://doi.org/10.31223/X5WW5B) thus making the full description of the methods directly accessible. We added the missing information in the manuscript at L. 97-117 (see **iii** for suggested text).

(iii) Equation 1 describes the conceptual relationship between debris thickness and sub-debris melt rates that is represented by the Østrem curve. The free parameters i_{debris} and k_{debris} are determined for each glacier by fitting a curve of the form given by Equation 1 to (a) the surface mass balance generated using the energy-balance model, and (b) the debris thicknesses used as input to the energy-balance model. We reformulated the entire paragraph, in order to make the procedure clearer.

L. 97-121: *'The debris thickness maps are based on McCarthy et al. (in review), and were obtained through a simplified surface-mass-balance inversion procedure similar to Ragettli et al. (2015) and Rounce et al. (2018). In a nutshell, the inversion procedure iteratively solves for the debris thickness by using an energy balance model. The procedure uses DEMs, glacier ice thickness, surface velocity, debris proprieties, and meteorological forcing data as input, and uses them to calculate ice flux divergence and ice thinning rates. The debris thickness is then adjusted until modelled and observed ice-melt rates agree within a prescribed tolerance. Due to the physical nature of the procedure, the energy-balance model and the Østrem curves (see below) do not need calibration. The debris thickness maps are however evaluated using a high number of available in-situ data (*148007 data points *on 13 glaciers) on debris thickness, showing good agreement (see McCarthy et al., in review).*
To generate the Østrem curves, in a first step, the energy-balance model was run at randomly chosen points on the surface of each considered glacier, and with debris thicknesses and debris properties randomly chosen within expected physical ranges. These Østrem curves are expressed as:

$$b = \frac{i_{\mathrm{debris}} \cdot k_{\mathrm{debris}}}{h + k_{\mathrm{debris}}} \qquad (1)$$

where b is the local surface mass balance, i_{debris} and k_{debris} are parameters to be determined, and h is the debris thickness (m). Note that equation 1 has similarities with the Hyper-fit model of Anderson and Anderson (2016), and Anderson et al. (2021a, b), although we note that the two approaches differ in the number of parameters and their interpretation. In a second step, the mass balances inferred by Miles et al. (2021) were used together with the fitted Østrem curves (Eq. 1) for each elevation (i.e. assuming that englacial and basal mass balance is negligible) to determine the debris thickness maps used in this study. The so-obtained information represents the supraglacial debris conditions for the period 2000-2016. With the method described above, McCarthy et al. (in review) estimate a mean debris thickness for the debris-covered part of all glaciers in HMA of

0.34m (with an uncertainty between 0.15 and 0.76 m). The uncertainties are asymmetric because surface mass balance is less sensitive to debris thickness as debris thickness increases, and are in line with other studies (e.g. Rounce et al., 2021). For our purposes, the spatially-distributed debris-cover information is divided into elevation bands of 10 m whilst the Østrem curves were directly added into our mass balance module (see section 3.1)."

(iv) There are two main reasons for why a temperature-index approach is used for the forward projections: (1) Running a distributed energy-balance model over the whole of HMA and up to the year 2100, would be computationally unfeasible; the energy balance model in McCarthy et al. was run at the point scale, with one point for each elevation band of each glacier, and this alone required a very large computation effort. (2) An energy-balance model requires additional meteorological forcing (e.g. wind, humidity, solar radiation) which may be much more uncertain in future simulations than it is in the meteorological reanalyses.

(v) The goal of Equation 2 is to mimic the relationship between debris thickness and sub-debris melt in a manner that is appropriate for an empirical, temperature index model. We sought for a simple, functional relationship that can be represented in the GloGEMflow framework. As such, the parameters do not have a strict physical meaning.

We hope that with the reformulation of L. 97-121 (cf. 1.01 iii) of our method is clearer. Further, we clarified the meaning of 'g', which is a factor enhancing ablation when debris is present.
L. 179:*'The factor g (which acts as a factor enhancing ablation due to debris) depends on […]'*

(vi) As suggested by the reviewer, we added a figure with different curves for different parameter values, and a visual explanation of Eq. 2 and 3.

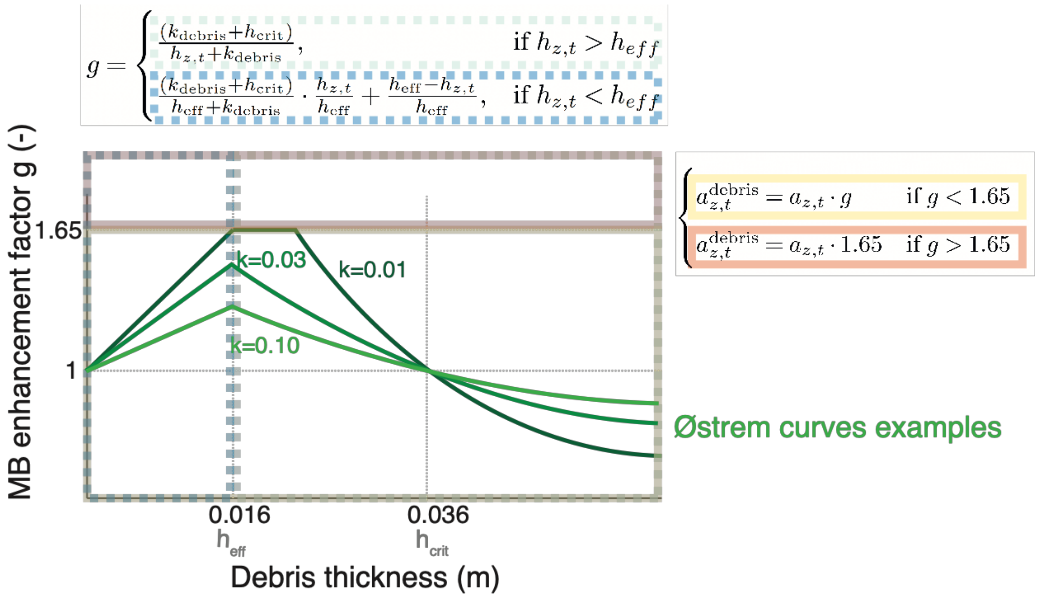


Figure 3: *'Schematic of the melt enhancement factor* g (dimensionless) *as a function of debris thickness for three different glaciers (green lines). The colored, dashed boxes show regions in which the different cases of Eq. 2 and 3 apply.'*

(vii) The uncertainty in our debris thickness estimates is moderately high. This is because of the variety of input datasets required by McCarty et al. (in review) and because of our conservative error propagation approach (in which the uncertainty accumulates).
For the whole of HMA, we estimate a mean debris thickness of 0.34 m (with an uncertainty of between 0.15 and 0.76 m). We note that similar relative uncertainties are produced by, e.g. Rounce et al, (2021).

We added the overall debris thickness uncertainty (cf. AR 1.01 iii) and the glacier-specific uncertainty in Fig. 1

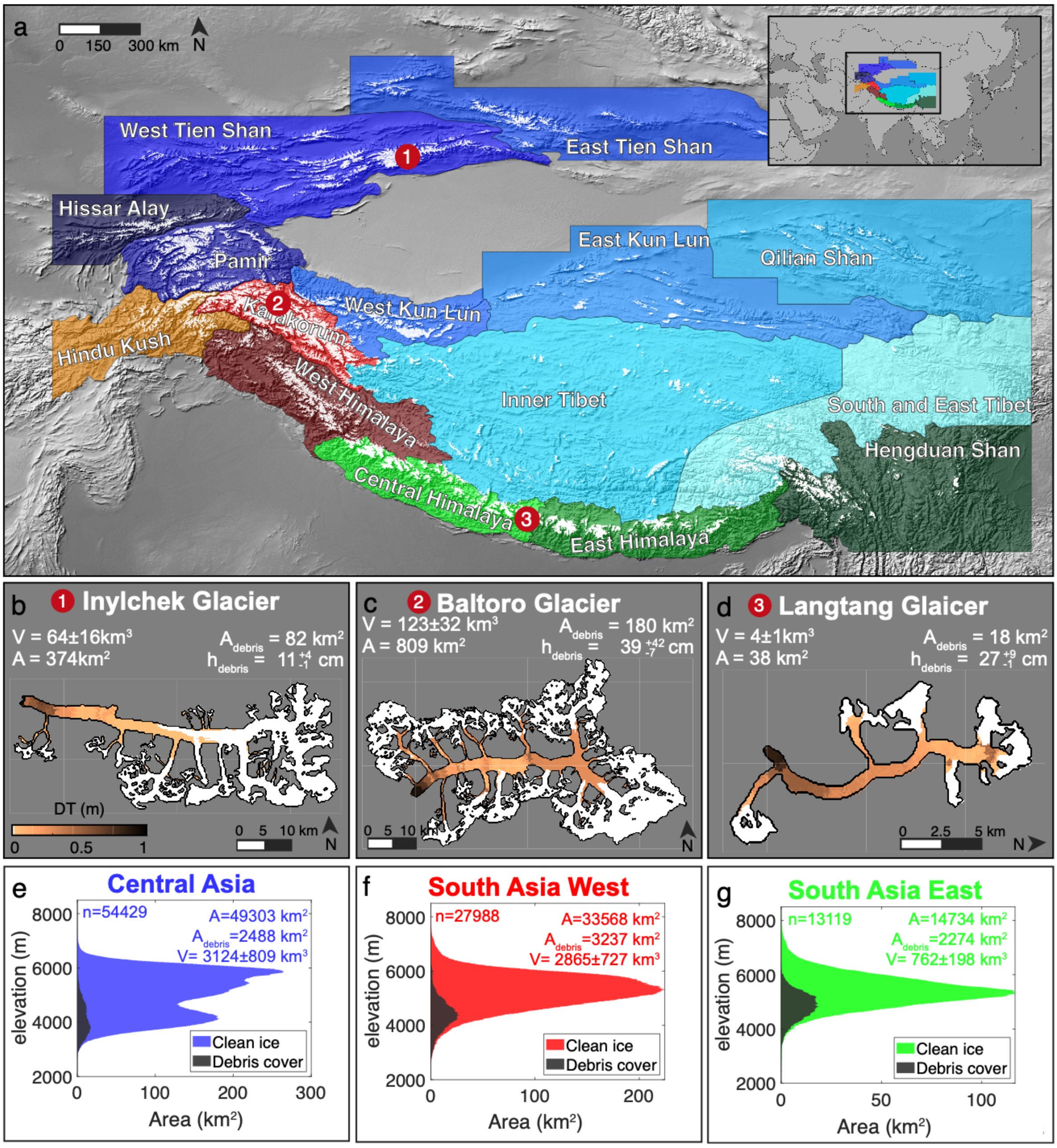


Figure 1. *(a) Extent of HMA glaciers (white) as per Randolph Glacier Inventory version 6 (RGI Consortium, 2017). The three main RGI regions (Central-Asia, South-Asia-West, and South-Asia-East) are shown by blueish, reddish, and greenish colours, respectively. RGI second order regions are labelled individually. Three glaciers are highlighted to illustrate glacier-specific model results (red spheres with numbers). (b & c & d) Map of the three highlighted glaciers with their mean 2000-2016 debris thickness given by colours (scale in panel b). Glacier outlines and debris thickness are from RGI Consortium (2017) and McCarthy et al. (sub.), respectively. For each glacier, V is the glacier ice volume according to Farinotti et al. (2019a), A is the glacier area according to RGI 6.0, A_{debris} is the debris-covered area, and h_{debris} is the mean debris-cover thickness according to McCarthy et al. (sub.). (e & f & g) Glacier hypsometry (area per 10 meter elevation band) and debris-covered area distribution at inventory date. n is the number of glaciers within each region (RGI Consortium, 2017) Map source: Natural Earth.*

[RC 1.02] (i) I don't quite understand the reasoning behind Eq. 4, specifically of including abs(b_z,t): if we assume a generally negative glacier-wide mass balance (B_(t-9,t)), an anomalous positive local mass balance would increase the fraction of debris cover in the elevation band, according to this equation. **(ii)** Similarly, if the glacier-wide mass balances were generally positive, a local negative mass balance would decrease the fraction of debris cover. I.e., the assumption that b_z,t is always negative in the elevation bands where debris is present seems a bit strong to me, **(iii)** and I'm wondering what the impacts are around the ELA altitude, where the sign could go both ways. **(iv)** The same question applies to Eq. 6. **(v)** Also, please clarify whether c_lateral and c_thickening are global or glacier-specific parameters. **(vi)** Finally, the choice of ten years as a time scale in these equations is not justified nor calibrated. It may be hard to calibrate and evaluate, but it would be good if the authors point out to which degree this is arbitrary, or whether there is a more substantial argument for ten years.

[AR 1.02] (i) In Eq. 4 (and also in Eq. 7, formerly Eq. 6), abs($b_{z,t}$) is the local mass balance of a certain elevation band, where the glacier is debris-covered. The term abs($b_{z,t}$) accounts for the fact that if we assume a constant debris concentration within the ice, the amount of released debris is proportional to the local melt: if locally there is a higher melt rate, also more debris should be released by the melted ice.

The glacier-wide mass balance B_(t-9,t), instead, decides upon the sign of the equation. E.g., if a glacier showed a negative balance during the last 10 years, then we expect that the debris fraction will increase. This is because more debris is melted out and because the ice-flow rate (which evacuates the debris from the surface) will decrease. This increase in debris fraction is observed on manly glaciers with negative mass balance (e.g. Stokes et al., 2007; Bhambri et al., 2011; Bolch et al., 2011; Shukla and Qadir, 2016; Tielidze et al., 2020).
With a neutral glacier-wide mass balance, instead, our equation results in a debris fraction that does not change in time. Indeed, the debris would be roughly in balance with the glacier mass balance in such a case, and would respond primarily to external forcings (e.g. long-term changes in debris supply and glacier dynamics' response to its geometry). Finally, if the glacier-wide balance is positive for a certain period of time, we expect a general debris-fraction decrease. This is because less debris would be melting out and because the increasingly higher ice flow speed would contribute for debris to be evacuated from the glacier.

In the reviewer's example, where we have a generally negative glacier-wide mass balance and an anomalously positive local mass balance, both the local debris cover fraction and thickness would increase if the positive anomaly occurs below the ELA. However, it is impossible to have a positive local mass balance below the ELA at the same time with negative glacier wide balance. It could only occur with anomalously high local snow accumulation, which cannot be captured with our accumulation model (thermal inversion, shaded areas or cloudy hotspots are not accounted for).
Therefore, we assume that the local positive mass balance is above the ELA, i.e. where it is not possible to have debris cover on the surface. Indeed, the debris cover in the accumulation does not matter in the model, because of snow cover.
We imagine that the reasoning behind Eq. 4 and 7 was not understood because the information that debris can not exist and therefore expand on the glacier surface where there is snow the entire year (i.e. accumulation) was lacking.

We added this information in the manuscript and in Eq.4 and Eq. 7:
L.228-229: 'Note that this equation is only applied in the glacier ablation zone. Debris cover is not permitted in the accumulation zone.'

Eq. 4:

$$\gamma_{z,t} = \gamma_{z,t-1} + abs(b_{z,t}) \cdot \overline{B_{(t-9,t)}} \cdot (-1) \cdot \gamma_{z,t-1} \cdot c_{\text{lateral}}, \quad \text{if } z < ELA$$

Eq. 7:

$$h_{z,t} = h_{z,t-1} + abs(b_{z,t}) \cdot \overline{B_{(t-9,t)}} \cdot (-1) \cdot \overline{h_0} \cdot c_{thickening}, \quad \text{if } \gamma_{z,t-1} = 0 \text{ and } z < ELA$$

(ii) This is correct, if the glacier-wide average mass balance of the last ten years is positive, a locally negative mass balance would decrease the fraction of the debris cover. $B_{z,t}$ does not change sign in the equation, because the sign of the equation is determined by $B_{(t-9,t)}$. On debris-covered glaciers, often there is a negative anomaly in local mass balance (compared to a clean ice glacier) in the upper part of the ablation area, where there is a low fraction of thin debris. Therefore, we expect the change in debris-cover fraction and thickness to be highest in the area with a low fraction of thin debris, rather than on the lower part of the accumulation area, where the ice is often close to stagnant (of course also here the debris will decrease, with a positive $B_{(t-9,t)}$).

(iii) If there is debris close to the ELA, our approach causes the debris fraction (and also the thickness) to increase when the glacier-wide mass balance is negative and to decrease when it is positive. However, remember that the model does not admit debris above the ELA.

(iv) Here the principle is exactly the same as described above (see point "i"), but is now applied for the debris thickness evolution**.**

(v) c_lateral and c_thickening are regional parameters, i.e. the same values are applied for the entire HMA. Unfortunately, there is insufficient data on debris-cover evolution to make these parameters glacier-specific. We added this information in the methods and calibration sections:

L210-211: *'c_lateral is a regional debris-cover extension parameter which is calibrated to minimize the difference between observed and computed lateral expansion of debris (see section 4.2)'*
L253-255:*'c_thickening is a regional calibration parameter for the debris-cover thickness evolution. It is constrained based on observations (see section 4.2)'*
L307-308: *'The mean value [of c_lateral] is used for all further modelling, i.e. the same value is applied for all glaciers in HMA […]'*

(vi) The time window of 10 years is meant to account for the time it takes for the debris cover to respond to a long-term change in the glacier-wide mass balance. Indeed, we do not expect the debris fraction or debris thickness of a glacier to respond to the mass balance on a year-to-year basis. Due to the lack of data for the temporal evolution of debris, it is impossible to determine the exact and optimal length of this time window. The length of 10 years is based on our gut feeling and judgment. We added this information in the manuscript.

L222-224: *'Since there is not enough observational data that would allow constraining the parameter, the time window of 10 years, which accounts for the time it takes for the debris cover of a glacier to respond to changes in the glacier-wide mass balance, is based on our own judgment.'*

[RC 1.03] Comparison between "explicit" and "implicit" treatment of debris cover: I think one of the strongest points of the study is the possibility to evaluate how "wrong" glacier models are that are not taking into account debris cover explicitly. As is pointed out in the paper – and depending on the calibration scheme of the respective model – they often do take it into account implicitly, through the calibration using observations that include the effect of debris cover. Unfortunately, this possibility is mostly left unused in the manuscript. I would like to suggest to add a focus on this, as it would be straightforward and quite informative. Some examples of simple analyses: **(i)** what is the distribution of the differences of DDFs between explicit and implicit treatment of debris cover in GloGEMflow? **(ii)** How does the performance of the mass balance model change when debris cover is treated explicitly? Does it (significantly?) reduce RMSE or increase correlation of observations, etc.? I.e., how do figures S2 and S3 change when the debris cover parameterization is introduced? I appreciate the difficulty of getting this into the (already quite long) paper, but I would argue it is quite central to address these questions when introducing a new parameterization. To some degree, this has been done (e.g., Fig. S11) – but not for the mass balance model. I can well imagine that there are no significant changes in mass balance model performance – but it would be good to point this out, because I would argue that representing more processes without significantly deteriorating the performance of a model is already a good step ahead, given the limitations of data availability for calibration and evaluation.

[AR 1.03] (i) We added the analysis requested by the reviewer in the manuscript. In line with the rest of the study, the analysis shows that the difference in degree-day factors is *relatively small* when modelling the debris cover explicitly or implicitly.
l.469-473: *'Aggregated over all of HMA, and considered in terms of glacier volume and area changes, the difference in the results between explicitly and implicitly modelling debris cover is relatively small. On average, DDF_{ice} is of 3.49 mm d^{-1} °C^{-1} when the debris cover is modelled explicitly and of 3.55 mm d^{-1} °C^{-1} when it is modelled implicitly. However, on the single glacier scale the differences in DDF_{ice} are larger (e.g. 0.47 mm d^{-1} °C^{-1} for Langtang and 0.20 mm d^{-1} °C^{-1} for Inylcheck).'*

(ii) We analysed the performance of the mass balance model when implicitly modelling debris cover (Figure R1), and the RMSE, bias and distribution are extremely similar to when the debris cover is modelled explicitly (Figure S2).
However, this is due to the fact, that the majority of the glaciers which are in the WGMS are clean ice glaciers. Therefore, we decided to not put this analysis and Figure R1 in the manuscript.

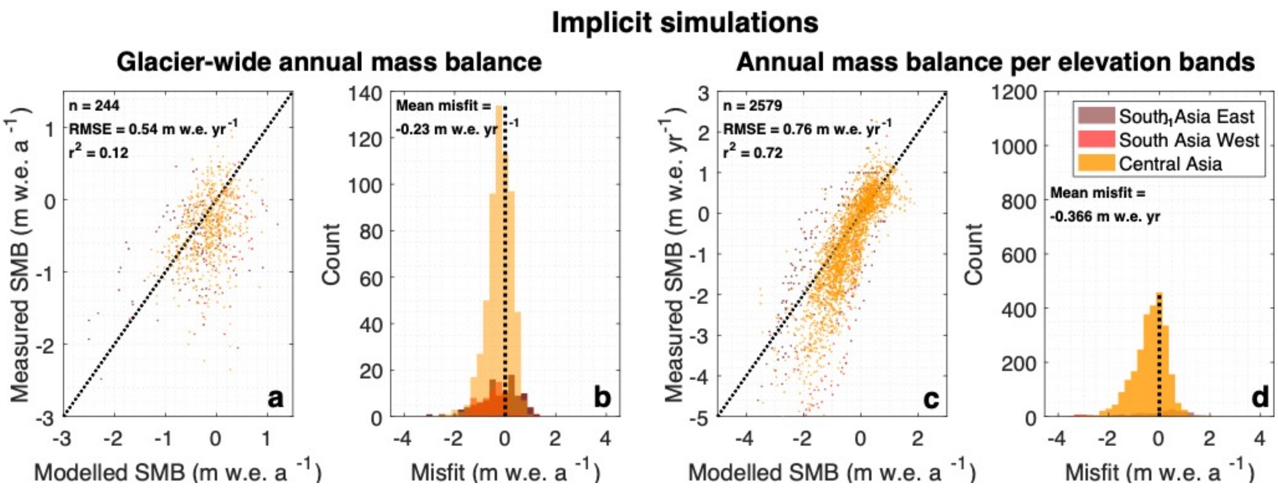


Figure R1: *'Evaluation of modelled annual glacier-wide and per elevation bands mass balance when implicitly modelling debris cover with observations from 21 glaciers provided by the World Glacier Monitoring Service (WGMS, 2020).'*

Minor comments

[RC 1.04] L10: since the focus of the manuscript is on the effect of debris cover, it would be more interesting if the impact of debris cover on projections was given instead of the totally projected mass loss. E.g., "explicitly accounting for debris cover in the projections only has a minor effect on the projected mass loss, but improves the representation of processes on the intra-glacier scale" (or something similar, based on Sect. 7.1).

[AR 1.04] We changed the sentence as suggested by the reviewer.
L10-12: *'Explicitly accounting for debris cover has only a minor effect on the projected mass loss, which is in line with previous projections. Despite this small effect, we argue that the improved process representation is of added value when aiming at capturing intra-glacier scale, i.e. spatial mass balance distribution.'*

[RC 1.05] - Throughout the manuscript: the authors often refer to the "module", but in many places the word "parameterization" would be more correct. I would suggest distinguishing between the parameterization and its implementation as a module of GloGEMflow, in order to increase clarity.

[AR 1.05] The word 'model' is used for GloGEMflow, which is divided into several 'modules' (e.g., mass balance, ice flow, and debris-cover), and each of these modules has various 'parameterizations'. We recognize that the wording was inconsistently used in in different places of the manuscript (e.g. L. 105, 316,318, and 380) and we corrected these.

[RC 1.06] - L27-30: it might be worthwhile here to be a bit more specific by explaining the impact of debris based on an energy balance perspective instead of "ice melting" (i.e., albedo, thermal conductivity, reduced turbulent fluxes of heat and water vapor, etc.).

[AR 1.06] We thank the reviewer for this suggestion. We have modified the text as follows: l.29-36: *'The presence of debris at the ice surface has the effect of reducing the surface albedo and increasing the net shortwave radiation (Owen et al., 2003; Reid and Brock, 2010). When debris is particularly thin and/or patchy, this excess energy can be readily conducted to the ice, thus enhancing melt rates (e.g. Ostrem et al 1959; Reznichenko et al, 2010; Fyffe et al, 2020). However, for thicker, continuous debris layers, the increased isolation layer allows for high debris surface temperatures (often >15 ° C), thereby increasing both the outgoing longwave radiation and the turbulent energy fluxes directed away from the surface (e.g. Nicholson et al, 2006; Steiner et al, 2018). This results in a reduced and delayed conduction of energy to the glacier ice, leading to a progressive reduction of melt with increasing debris thickness (e.g. Ostrem 1959; Reznichenko et al., 2010; Anderson and Anderson, 2016; Rounce et al, 2021).'*

[RC 1.07] - L31-32: I would argue that debris cover would evolve in response to any non-zero mass balance, also if there was no "additional" disequilibrium between glacier geometry and climate conditions.

[AR 1.07] We changed the sentence into:
L37-39: '*Since glaciers are presently far from equilibrium (Marzeion et al., 2018; Zekollari et al., 2020; Miles et al., 2021), their debris cover is evolving through time (Stokes et al., 2007; Bhambri et al., 2011; Bolch et al., 2011; Shukla and Qadir, 2016; Tielidze et al., 2020).'*

[RC 1.08] - L39-41: if the relatively constant debris cover in the Karakoram region is to be explained by neutral or slightly positive MB in the region, it is not an exception from what was explained above, which referred to negative MB.

[AR 1.08] We changed the sentence into: '*In the Karakoram region, instead, positive and negative debris-cover changes offset one another during the past 40 years.'*

[RC 1.09] L43-44: Sentence is a bit unclear: I think what you want to express is that the mass balance profile of a debris-covered glacier may have a local minimum at mid-elevation?

[AR 1.09] We reformulated the sentence into:
L. 48-49: *'Indeed, the mass balance profile of a debris-covered glacier may have a local minimum at mid-elevations, especially if the ice is clean (i.e. not covered by debris) at that elevation.'*

[RC 1.10]- L71: the references to Marzeion et al. 2020 and Edwards et al. 2021 are a bit strange here: it would make more sense to compare the results with projections obtained from the same model, but without applying the debris cover evolution parameterization.

[AR 1.10] We agree. We removed the references.

[RC 1.11]- L104-117: it is not clear here whether only the debris extents are used for calibration and validation, or whether additionally, debris thicknesses are estimated for different time periods as well, in order to evaluate the modeled thickness evolution as well (I don't think so, but please specify).

[AR 1.11] The Hexagon and the Landsat images are used for the calibration and validation of the evolution of both the debris area and thickness. Indeed, in section 4.2.2 we calibrate, and in section 5.2.3 we evaluate the debris-thickness evolution using such images (combined with the debris thickness dataset of McCarthy et al., in review). We added this missing information in the manuscript:
L.122-125: *'To calibrate and evaluate the parameterizations used for describing the evolution of both debris area and thickness (see section 3.2), we use multiple Hexagon and Landsat satellite images acquired between 1973-1976 and 1987-2019, respectively.'*

[RC 1.12]- L129-131: suggest to rephase for clarity, e.g.: "… we use 53 members of the CMIP6 ensemble (Eyring et al., 2016) from X different GCMs, covering 5 SSPs (5 members for SSP119 and 12 members for all other SSPs)."

[AR 1.12] We reformulated the sentence as suggested by the reviewer.

[RC 1.13]- Sect. 2.3: it remains unclear here whether you use the MB data for calibration and evaluation of the glacier-wide modeled mass balances, or also for the evaluation of the mass balance profiles.

[AR 1.13] We reformulated two sentences of this section, to clarify which type of data we used for calibration, and which for the evaluation.
L.140-141: *'To calibrate the mass balance module of GloGEMflow, we rely on glacier-wide geodetic volume changes available for 2000-2019 (Hugonnet et al., 2021).'*
L.146-147: 'To evaluate the mass balance module, we use independent data from in-situ observations provided by the World Glacier Monitoring Service for 21 glaciers (WGMS, 2020).'

[RC 1.14]- L131-135: would it be correct to state that you use only anomalies from the GCMs, while keeping climatology from ERA5? The phrasing in the manuscript - "a set of additive and multiplicative corrections" - seems a bit unclear to me.

[AR 1.14] In a certain way, yes. If every dataset would be perfect and the grids would be consistent, such a de-biasing would not be needed. However, since there is a bias between ERA-5 and GCM results in both temperature and precipitation, as well as significant differences in spatial grid resolution, this bias has to be adjusted. In our approach, we assume that ERA-5 should be 'closer to reality' compared to the GCMs given its higher spatial resolution and the rooting in observational data sets. Therefore, the GCMs are 'adjusted' on the ERA-5 dataset, in order to have a seamless time series from the past into the future.

Since explaining the complete procedure would need about two paragraphs, and in light of the manuscript length, we prefer to keep this shorter version (which is the same as used in Compagno et al., 2021b). Additional information about this procedure is also available in Huss & Hock, 2015, as mentioned in the manuscript.

[RC 1.15]- L172: this was mentioned before, but why actually are surging glaciers excluded from the application of the debris cover parameterization? Because of difficulties in calibration and evaluation, or because of more fundamental, physical reasons?

[AR 1.15] We deleted the repetition.
Concerning the surging glaciers: there was an error in the manuscript. In this study, surging glaciers are not excluded. We corrected the manuscript accordingly:

l.93: *'For each glacier with an area >2 km^2 with debris cover (i.e. 6115 glaciers in total, see RGI Consortium, 2017),'*

[RC 1.16]- L212-219: **(i)** it is not quite clear to me which variables are used for the regression here: do you calculate a regression of ELA over time, and then use the slope of the regression directly – or do you have a constant of proportionality (L213 says "proportional to")? **(ii)** I also would argue that claiming "no calibration parameters" is a bit bold, because some choices (e.g., new debris thickness of 1 cm, expansion by Eq. 5) seem a bit arbitrary to me (the authors point that out). It is the authors' choice not to calibrate them, and it may not be possible to do so, and I think they should say so explicitly.

[AR 1.16] (i) We agree that this was unclear. We first calculate a regression of ELA over time, and then use the slope of the regression directly. We added under which conditions Eq. 5 can be applied, and we added Eq. 6 which shows for how many elevation bands Eq. 5 is applied at time t. Here a copy of equation 5 and 6 and of the modified sentence (l.235-240):

$$\gamma_{z,t} = \frac{\gamma_{z-1,t} + \gamma_{z+1,t}}{2}, \quad \text{if } \gamma_{z,t-1} = 0 \text{ and } z < ELA \text{ and } \frac{dELA_{t-9,t}}{dt} > 0. \quad (5)$$

This process is discretized within elevation bands of 10 m. The amount of elevation bands h without debris at time t-1 that can gain debris from a nearby elevation band at time t (we use yearly time steps) applying Eq. 5, is equal to the rise of the ELA over the last ten years, determined using linear regression of the values ELA(t-9,t), i.e.

$$\#h = \frac{dELA_{t-9,t}}{dt \cdot 10} \quad (6)$$

(ii) We reformulated the sentence to make clearer that it was our choice to not calibrate this. The reason is to be found in the need of very accurate data.
L.244-245: *'The procedure was not calibrated due to the need of a considerable amount of accurate data. An evaluation of the performance is found in section 5.2.2.'*

[RC 1.17]- L294: it is unclear why two values, each with uncertainties, are given for what I would expect to be a range (two values with no uncertainty) or a number with uncertainties.

[AR 1.17] We changed the values into two ranges, like requested by the reviewer
l.324: *'[…] (on average between 1979-1983 and 1990-1996 in this study) […]'*

[RC 1.18]- Fig. 5: the Figure seems excessively complex from the perspective that the tabulated values (right part of the figure) do not seem to play any role in the analysis of the results. I would suggest to remove this part. Also: I don't understand the definition of DC_frac.

[AR 1.18] We deleted the tabulated values, and adapted the caption accordingly. By doing so, we deleted also 'DC_frac', which is the fraction in area of the debris cover. E.g. DC_frac=0.3 means that 30% of the glacier area is debris covered.

[RC 1.19]- Sect. 5.1: this should be expanded significantly, see major comment above.

[AR 1.19] We added an analysis and a plot of the validation of the mass balance model when modelling debris cover implicitly (cf. AR 1.03).

[RC 1.20]- Sect. 5.2.3: I'm wondering how the difference between activated and non-activated thickness evolution compares to the uncertainty of the estimated debris thicknesses in the "observations" used for evaluation.

[AR 1.20] The uncertainties in the estimates of debris thickness in the "observations" (results based on McCarty et al., in review) are significant (cf. AR1.01 vii). This is in line with what we showed in the manuscript, i,e. high glacier-to-glacier variance and a rather weakly contained value of $c_{thickening}$.

[RC 1.21]- Fig. 8: **(i)** I don't understand the meaning of the colored numbered circles in panels a and c, and the illustration in panel b. I guess the intention is to point out the relevant processes for different locations and times for changes in debris distribution on the glacier, but since all three parameterizations are active all times, it's hard to get a specific meaning (especially in panel c). I would suggest to remove this. **(ii)** Also: please point out how the maps (panel c) are drawn, given that the model only "knows" elevation bands. I guess the red outlines are simply based on elevation?
(iii) Finally: instead of panels f and i, please consider only showing the differences between implicit and explicit treatment; as is, they are very hard to see (which of course is relevant in itself, but maybe not the main point here).

[AR 1.21]
(i) we removed the colored numbers.
(ii) This is a result from extrapolating from 1D to 2D only used for the plot, to better illustrate the model results. The detailed method used for this is described in the supplementary material. We rephrased the sentence that provides this information.
Fig. 9 (former Fig. 8): ' *(c) Model results extrapolated to 2D (see Supplementary Material for the method used for extrapolating from one to two dimensions, and note that the extrapolation is for visualization purposes only, i.e. it does not affect the presented results)*
(iii) we changed panels f and i, into plots of differences between explicit and implicit treatment, as requested by the reviewer. We changed panels f and i also for Baltoro and Inylchek glaciers (Fig. S5 and S6). Accordingly, also the caption was changed. Fig.9: *'for every SSP, (f and i) show the difference in glacier volume and area obtained when explicitly and implicitly modelling the debris cover.'*

[RC 1.22]- **(i)** L395: whether the mass balance gradient is "erroneous" in that case is not clear, unless it is shown that the explicit treatment of debris cover actually improves the model's representation of the mass balance gradient (see major comment above). **(ii)** Also: in Fig. S7, I think the last sentence in the caption should read "higher mass balance gradient with elevation of (b)" (not "(a)").

[AR 1.22] (i) we changed the sentence into: *'In that case, however, surface mass balance gradients would not consider the effect of debris cover […]'.*
(ii) we corrected the figure caption.

[RC 1.23]- L441-443: this statement is not well backed-up by the analyses presented, at least not as categorically as it is given here: there is no evaluation of the impact of the parameterizations on model performance regarding local mass balance, glacier length, or runoff. I agree that the results including the parameterizations are more plausible – but it is not a correct/incorrect-situation, and it has not been shown that the model's performance did in fact improve.

[AR 1.23] We admit, 'captured correctly' was a bit strong. We rephrased the sentence into: l.476-479: *'Indeed, quantities such as the local mass balance, the glaciers' ice flow velocity and mass turnover, the glacier's length change or water runoff are only captured adequately when explicitly accounting for supraglacial debris and its temporal evolution.'*

[RC 1.24]- L458: I would also argue that uncertainty arises from parameters that are not calibrated here; e.g., the time scale for debris cover expansion/thickening, new debris thickness.

[AR 1.24] We agree that there are also uncertainties coming from the parameterization itself and all the parameters which are within it. We added an additional paragraph, in order to clarify that also other uncertainties exist (cf. AR2.04 v).
ll. 501-509: *' Additional uncertainties arise also from the parameterizations themselves (Eqs. 4, 5 and 7). In a simplified but realistic way, our approach parameterizes the evolution of debris cover*

on glaciers, and it is based on debris-evolution patterns observed in the last decades. We acknowledge that this is not the only way that debris evolution could be parameterized. Explicitly accounting for the dynamics of debris re-distribution (as opposed to implicitly, see Section 3.2.3) could be an alternative option. Indeed, where the ice flow velocity decreases, the local debris thickness increases and vice versa, due to the convergence of debris particles. Such an approach was followed by Anderson et al. (2021b), for example. We decided not to include such effects explicitly because (1) the absence of data to calibrate and validate a more complex parameterization at regional scales, and (2) the small sensitivity of regional-scale glacier volume and area to changes in debris cover (see Section 7.1).'

[RC 1.25]- L461: the "little impact" is, however, comparable to the magnitude of switching the parameterizations completely off.

[AR 1.25] This is true. However, for specific glaciers it is more substantial (see l. 494) and we did not expect such a small influence of debris evolution on the total regional volume evolution. Even if the magnitude is comparable, a sensitivity analysis has to be done, in order to know how sensitive, the model is to changes in $c_{lateral}$ and $c_{thickening}$.

[RC 1.26]- L464-465: unclear whether these numbers refer to the sensitivity analyses (changing c_thickening and c_lateral) of to completely switching the parameterizations off.

[AR 1.26] We added the missing information in the sentence:
L499-500: *'However, volume differences for individual large and strongly debris-covered glaciers can be as high as 18% when the debris-evolution module is disabled (e.g. Langtang Glacier 2 %, Baltoro Glacier 8 % and Inylcheck 1 %).'*

[RC 1.27] L528-529: by affecting the MB gradient, debris cover would also impact methods of ice thickness estimation based on mass continuity – wouldn't it?

[AR 1.27] Indeed. We added this information in the manuscript.
ll. 572-574: *'This results in turn in a reduced mass turnover, with consequences for the future evolution of the glacier's geometry, the modelled surface ice velocities, or the methods that use considerations of mass turnover for estimating glacier ice thickness (for overviews, see Farinotti et al., 2017, 2021).'*

Technical comments and suggestions

[RC 1.26]- L1: grammar: not the area is altering the surface mass balance, but the debris cover.

[AR 1.26] Corrected. L1 *'[…] glacier area is debris-covered, that alters its surface mass balance.'*

[RC 1.27]- L4-5: grammar: "the module" was previously called an "approach"; therefore better (instead of "the module"): "... we implement a parameterization into ..." or "We derive a parameterization and implement is as a module into...".

[AR 1.27] Corrected with reviewer's second suggestion.

[RC 1.28]- L8: replace "the model" with "GloGEMflow", in order to prevent misinterpretation with "the module".

[AR 1.28] Corrected.

[RC 1.29]- L10: replace "projections in the literature" with "previous studies".

[AR 1.29] Corrected.

[RC 1.30]- L11: replace "modelled" with "projected".

[AR 1.30] Corrected.

[RC 1.31]- L18: reference to IPCC (2019) should be replaced with reference to Lee et al. (2021), i.e. Chapter 4 of the 6th assessment report.

[AR 1.31] Corrected.

[RC 1.32]- L47: add “distribution” after “thickness”.

[AR 1.32] Added.

[RC 1.33]- L57: delete second “by”, otherwise it implies that it was the first projection generally for HMA glaciers.

[AR 1.33] Done.

[RC 1.34]- L66: I’m not sure if you can “extensively calibrate” a parameterization, since there is simply a given number of parameters; I suggest to apply the word “extensively” only to “evaluate”.

[AR 1.34] Changed.

[RC 1.35]- L81: replace “condensed” by “simplified”.

[AR 1.35] Corrected.

[RC 1.36]- L75: the number of glaciers should be referenced (e.g., “… all 95536 glaciers contained in the RGI…”

[AR 1.36] Added.

[RC 1.37]- Fig. 1: **(i)** maybe a matter of taste, but the overlapping panels make the figure look a bit messy to me; **(ii)** “red spheres” are red circles; **(iii)** please indicate uncertainties of V, h_debris and A_debris (if relevant) in panels b-g; **(iv)** “map source” is the background elevation data set?; **(v)** why is panel f highlighted with a thicker frame?

[AR 1.37] (i) We changed the configuration of the figure. Now the panels are not overlapping anymore; **(ii)** corrected; **(iii)** we added the uncertainty for the ice volume and debris thickness h_debris. We did not add the uncertainties of the debris-covered area A_debris because they are not available. **(iv)** Corrected; **(v)** No reason. We corrected the frame thickness.

[RC 1.38]- L86: grammar: “a surging glacier” (singular).

[AR 1.38] Corrected.

[RC 1.39]- L89: in reference, comma before Østrem is missing.

[AR 1.39] Added.

[RC 1.40]- L91: wrong parenthesis for references.

[AR 1.40] Corrected

[RC 1.41]- L104: “see Sect. 3.2”

[AR 1.41] Corrected

[RC 1.42]- L129: delete “Interim” (hint: the title of the reference Hersbach et al., 2019, includes the words “goodbye ERA-Interim”)

[AR 1.42] Corrected

[RC 1.43]- L138: “is illustrated”, not “are”

[AR 1.43] Corrected

[RC 1.44]- L251: fix phrasing in reference.

[AR 1.44] Corrected

[RC 1.45]- L252: delete “of”.

[AR 1.45] Corrected

[RC 1.46]- L254: please indicate the default values for the DDFs, in case step 2 of the calibration is never done.

[AR 1.46] The default value for DDF_{snow} is 3 mm d-1 °C-1. Therefore, the default value for DDF_{ice} is 6 mm d-1 °C-1. We added this information in the mansucript at L. 279: *‘If the second step is not needed, the default value of 3 and 6 mm d^{-1} K^{-1} are used for DDF_{snow} and DDF_{ice}, respectively.’*

[RC 1.47]- L278: replace “spread” with “uncertainty”

[AR 1.47] Corrected

[RC 1.48]- L278: “results’”, not “result’s”

[AR 1.48] Corrected

[RC 1.49]- L290: not the sensitivity of the value is evaluated, but the sensitivity of the results to the value.

[AR 1.49] Corrected. L. 316: *‘This value and the result’s sensitivity are evaluated in section 5.2.2.’*

[RC 1.50]- different places in the manuscript: the year 2020 is given as both using ERA5 and CMIP6 data as forcing; I guess the use of CMIP6 only starts in 2021? Or is the transition really within the year 2020?

[AR 1.50] The transition year is in 2020, i.e. the model is forced with ERA5 until 2020, and from 2021 onwards with CMIP6.

[RC 1.51]- Fig. 4 caption: typo “derbis”.

[AR 1.51] Corrected

[RC 1.52]- Fig. 4: panel a: suggest to delete “Identify” from the title; panel c: suggest to label x_axis with c_lateral for consistency.

[AR 1.52] We did the modification of the figure as requested by the reviewer

[RC 1.53]- Fig. 11: add explanation of dashed horizontal lines to caption. Also point out that this figure covers HMA, not the globe.

[AR 1.53] We added the missing information in the caption of Fig. 12 (former Fig. 11):
'Comparison of modelled volume changes with values from Marzeion et al. (2020) and Edwards et al. (2021).'
'The dashed lines correspond to the mean volume changes of this study.'

[RC 1.54]- L531: suggest to remove "appreciation".

[AR 1.54] Removed

[RC 1.55]- L537: remove "availability".

[AR 1.55] Removed

REVIEWER 2 - LEIF S. ANDERSON

Overview

[RC 2.01] (i) This is an interesting study that simulates the expansion and thickening of debris and its affect on the response of debris-covered glaciers to climate change in HMA. There are new, simple parameterizations presented to represent the expansion and thickening of debris for application on both the individual and regional scales. It is shown that based on these parameterizations of debris thickness change and the assumptions held within that debris cover plays a minor role in glacier-wide mass balance evolution for simulations of HMA glaciers run out to 2100.

I applaud the immense effort put into this work and the novel contributions made. There are a number of exciting inferences and conclusions. **(ii)** But there are some major issues that need to be addressed before it is ready for publication.

[AC 2.01] (i) We thank the reviewer for the positive feedback and for the thorough review.
(ii) Below, we have addressed all issued raised by the reviewer. The manuscript was updated accordingly.

Major comments

[RC 2.02] *Correct citing of Equation 1*
Some important citations are omitted. Some of which need to be cited or (I hate to have to say it) plagiarism is occurring. I am directly referring to Eq. (1) which was derived in its exact form by Anderson and Anderson (2016). This model is also presented in detail in the debris cover melt model intercomparison project manuscript (Pellicotti et al., in prep). This form of Ostrem's curve is also referred to as the hyper-fit model by (Anderson et al., 2021a, b). I am personally glad it is useful it but please cite it appropriately.

[AC 2.02] It was far from our intentions to commit plagiarism, and we are very sorry that the references pointed out were missing. We now make mention of the reviewer's work when introducing Eq. 1:

l. 112-114: '*Note that equation 1 has similarities with the Hyper-fit model of Anderson and Anderson (2016), and Anderson et al. (2021a, b), although we note that the two approaches differ in the number of parameters and their interpretation.'*

[RC 2.03] *Neglected role of surface velocity in thickening and expanding debris*
(i) The mass conservation equation for surface debris thickness change in time on a glacier surface includes debris melt out and dynamic re-distribution of debris:

$$\frac{\partial h_{debris}}{\partial t} = \frac{C\dot{b}}{(1-\phi)\rho_r} - \frac{\partial u h_{debris}}{\partial x} - \frac{\partial v h_{debris}}{\partial y} \qquad (3)$$

where C is near-surface englacial debris concentration, φ is the porosity of the debris, ρ_r is the density of the rock composing the debris, and u is surface velocity in the x-direction and v is surface

Debris melt out is represented by the first term on the right and dynamic re-distribution of debris is represented by the second two terms on the right. This form is taken form Anderson et al. (2021b).

The parameterizations for debris thickness change presented here (Eqs. 4 and 6) do not include the effect of ice dynamics in changing debris extent or thickness. Equation 6 in this paper really only takes into account the debris melt out term in that debris thickness change is directly related to the melt rate as it varies in time by a factor (Eq. 4 is very similar):

$$h_{z,t} = h_{z,t-1} + abs(b_{z,t}) \cdot \overline{B_{(t-9,t)}} \cdot (-1) \cdot \overline{h_0} \cdot c_{thickening}, \quad \text{if } \gamma_{z,t-1} = 0 \tag{6}$$

where $h_{z,t}$ is the debris thickness for elevation z and time t. $c_{thickening}$ is a calibration parameter for the debris-cover thickness evolution, constrained based on observations (see section 4.2). As for lateral debris expansion, the local mass balance $b_{z,t}$ relates linearly to debris thickness change. Higher melt rates will lead to faster debris thickening, thus implicitly assuming that debris concentrations within the ice are homogeneous. The long-term glacier-wide mass balance $\overline{B_{(t-9,t)}}$ mimics ice-dynamical processes. It leads to constant debris thickness for steady-state conditions ($\overline{B_{(t-9,t)}} = 0$), and to decreasing local debris thickness for consistent mass balances, thus mimicking the evacuation of debris with enhanced flow. This is in line with the few direct observations that are available (e.g. Gibson et al., 2017; Verhaegen et al., 2020). $\overline{h_0}$ is the mean debris thickness of the glacier at the inventory year. It parametrizes the effect that glaciers with a low mean debris thickness will thicken slower compared to glaciers with a high mean debris thickness. This is motivated by the assumption that glaciers with thick debris are likely to have a higher englacial debris concentration, indicative for high debris supplies from the surroundings.

(ii) Anderson and Anderson (2018) notes, using a theoretical analysis, that debris thickness patterns are strongly controlled by the inevitable decline of surface velocity down glacier. This is also basically outlined by Kirkbride (2000) and further supported by the modeling presented in Ferguson and Vieli (2021). I have not seen a compelling a reason in this manuscript why surface velocities should be neglected when considering the evolution of debris covers.

(iii) Furthermore, Anderson et al. (2021b) presents evidence that the dynamic effects of debris thickening (debris advection and debris compression) are unavoidably important for the thickening of debris in response to climate change. The dynamic affects on debris thickness are especially important where surface velocities are low and debris already tends to be thick. Meaning that debris might thicken substantially right where the model/parameterizations presented here is not accounting for it. This is because the thickening parameter is tuned with debris thickness estimates from the upglacier end of debris covers.

(iv) Anderson et al. (2021b) also shows how changes in flow patterns can change debris extent (I do believe this is mentioned breifly). I recognize that Anderson et al. (2021b) was published as this paper was coming out in TCD, but the work has direct implications for the parameterizations presented in this study.

(v) I don't think the neglect of the role of ice dynamics makes this work invalid, rather the fact that the dynamical terms in the debris conservation equation are neglected should really be discussed and stated very clearly in this paper. Right now half of the continuity equation for debris change is assumed to be negligible in this manuscript, but the affect of surface velocities on debris thickness is not negligible even where surface velocities are low.

(vi) It is difficult to evaluate this work without reading McCarthy et al., submitted and understanding how the debris thicknesses across HMA were estimated. I am not sure where to access that manuscript. How many validation data points are used to evaluate the debris thickness estimates in that study? What do the debris thickness patterns look like? In some way this pre-requisite work needs to be made available be thorough description here or elsewhere.

[AC 2.03]

(i) We thank the reviewer for providing us with the formula. From a theoretical point of view, it makes certainly sense to divide the debris thickness evolution into debris melt out and dynamic re-distribution of debris. However, note that this formula also assumes that the englacial debris concentration, the porosity of the debris, and the density of the rock composing the debris are completely known. At the regional scale (and mostly even at the local scale), such data are not available, leaving the formula with a number of unknowns. In our approach (Eq. 4 and Eq. 7), we do not account explicitly for the effect of surface velocity change. This is because during calibration it would be impossible to divide the observed debris-thickness change into the two necessary components (i.e. debris melt out and dynamic redistribution of debris). The latter would only be possible if the englacial debris concentration, the porosity of debris and the density of the rock composing the debris were known for each of the modelled glaciers. We emphasize that the physics-based analyses of debris evolution are vital for advancing understanding, but this is a reasonable first approach to represent the major changes in debris thickness patterns that should be expected, consistent with the modelling scope (climate projections).

That said, the dynamical redistribution of debris is not entirely omitted with our parameterizations. Rather, it is implicitly accounted for thanks to our calibration strategy. Indeed, the calibration of $c_{thickening}$ (and also $c_{lateral}$) with observations ensures that changes due to both debris melt out as well as dynamic re-distribution of debris are reproduced.

(ii) See above: our approach accounts for dynamical re-distribution implicitly (cf. 2.03 i). We agree that taking this into account explicitly (as done in Kirkbride (2000), Andrerson and Anderson (2018), and Ferguson and Vieli (2021), for example) would be ideal. It is impossible however at the scale of our investigation. Indeed that the mentioned studies either refer to synthetic glaciers (Andrerson and Anderson (2018), Ferguson and Vieli (2021)), or were only applied to extremely well-studied individual glaciers (Tasman Glacier in the case of Kirkbride, 2000). Applying similar methods to more than 6'000 glaciers in combination with climate forcing taken from more than 50 global circulation models would be computationally impossible, and severely limited by the lack of necessary input data. In brief, we argue that some simplifications are indispensable for regional-scale modelling, and this is exactly what our approach does.

(iii) This is not true. With our parameterization, the debris thickness will growth at a faster rate for areas with already thick debris cover compared to where debris is thinner (the debris thickness of a given elevation band is accounted for in in the second term of Equation 7). The same applies for Equation 4. Again, we note that our parameterization takes the effect of dynamic redistribution of debris into account implicitly. For an example of the results, see Fig. 9a, S6a and S6(b).

(iv) We added a reference to the mentioned paper, and we reformulated l. 255-260 to make our parameterization clearer:
'Combined with $b_{(z,t)}$, the long-term glacier-wide mass balance $B_{(t-9,t)}$ implicitly mimics ice-dynamical processes (see section 7.2). It leads to constant debris thickness for steady-state conditions ($B_{(t-9,t)}$=0), to a growth of debris thickness with negative mass balances (thus mimicking dynamic re-distribution of debris and its compression, e.g. Kirkbride, 2000; Anderson et al., 2021b; Ferguson and Vieli, 2021), and to decreasing debris thickness for positive mass balances (thus mimicking the evacuation of debris with enhanced flow).'

(v) We added a paragraph in the discussion section to discuss this topic:
l.501-509:*'Additional uncertainties arise also from the parameterizations themselves (Eqs. 4, 5 and 7). In a simplified but realistic way, our approach parameterizes the evolution of debris cover on glaciers, and it is based on debris-evolution patterns observed in the last decades. We acknowledge that this is not the only way that debris evolution could be parameterized.*

Explicitly accounting for the dynamics of debris re-distribution (as opposed to implicitly, see Section 3.2.3) could be an alternative option. Indeed, where the ice flow velocity decreases, the local debris thickness increases and vice versa, due to the convergence of debris particles. Such an approach was followed by Anderson et al. (2021b), for example. We decided not to include such effects explicitly because (1) the absence of data to calibrate and validate a more complex parameterization at regional scales, and (2) the small sensitivity of regional-scale glacier volume and area to changes in debris cover (see Section 7.1).'

(vi) We admit that it was our mistake not to provide the reviewers with the manuscript in review by McCarthy et al. The manuscript is now available at https://doi.org/10.31223/X5WW5B . Note that we reformulated the section referring to McCarthy et al. (in review) to clarify the methods (cf. AR 1.01). We used 148007 individual data points of 13 glaciers. However, the majority of these points come from closely spaced GPR measurements made on Ngozumpa Glacier. The median number of in-situ data points per glacier is 37. For the thickness pattern, there are typically an increase in debris thickness down-glacier toward the terminus, but there is also substantial variability between glaciers with different morphologies.

Minor comments

[RC 2.04] (a) There are places in the manuscript where the modelling results seem to be presented as reality but are really still just modeling results that rely on all of the assumptions inherent to the model design. More care should be taken to avoid overstatements.
(b) I also wonder: How do the debris change evaluations change if the evaluating datasets are not used in the tuning process?

[AR 2.04]
(a) We reworded many sentences in the results section, in order to clarify that we are presenting modelling results based on parameterizations, and not observations.

E.g.:

l 383-384: *'Figure 9a shows a profile view of the glacier and debris-cover evolution of Langtang Glacier according to our model results and SSP245.'*

l 397-398: '*As a result of the projected up-glacier migration, instead, the maximum elevation with supraglacial debris would increase by 280 m for the same time period.'*

l 401-402: *'With higher mass loss, our model results show that both debris-covered area and debris thickness increase, thus reducing ice melt.'*

l 413-414: *'Compared to 2020, the modelled mean debris thickness of Langtang Glacier is expected to […]'*

l 441-442: *'Locally, however, much higher debris thickness increases are modelled but these are offset at many places by the overall reduction in glacier area.'*

l 473-479: reformulated, see line by line comment AR 2.26.

(b) For calibration and validation different datasets are used. We imagine that here the reviewer is referring to the calibration and evaluation of the lateral spread out of debris cover (see **RC 2.22**).

In this case, it is correct that for the evaluation we use a subset of glaciers also used for the calibration. However, the analysis is initialized with different datasets and time periods. In the calibration, we used three to five Landsat scenes between 1989 and 2020. The model is initialized with debris extents of each of these scenes, and then compared with the older scenes. In the evaluation, we use Hexagon images of around 1974 as initial condition, and then we compare the model results with the Landsat scenes.

To investigate the sensitivity to these choices, we re-calibrated the model omitting all glaciers used in the evaluation. By doing so, the mean of the parameter $c_{lateral}$ is again found to be c_lateral=2.0 (as already used before in the manuscript). This shows that the results would not change if the glaciers used in the evaluation would not be used in the calibration, see figure below).

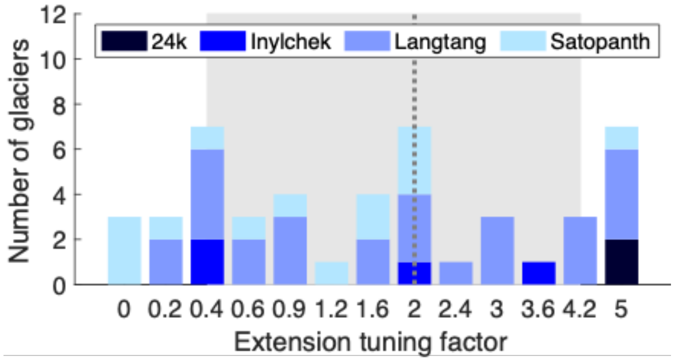


Figure R2: Distribution of the extension tuning factor $c_{lateral}$ resulting in the lowest misfit between observed and modelled debris-fraction evolution (given per number of glaciers). In contrast to Figure 5, glaciers used in the evaluation are omitted during calibration. The grey dashed line shows the mean value while the grey rectangle shows values within the 0.25 and 0.75 quartiles.

Line-by-line comments

[RC 2.05] Line 3: You could remove ‘potential’ here

[AR 2.05] Removed

[RC 2.06] Line 5-8: The sentence should be split in two as it is hard to follow as written.

[AR 2.06] We split the sentence:
L. 5-8: *‘The module is initialized with both glacier-specific observations of the debris' spatial distribution and estimates of debris thickness. These data sets account for the fact that debris can either enhance or reduce surface melt depending on thickness. Our model approach also enables representing the spatio-temporal evolution of debris extent and thickness.’*

[RC 2.07] Line 10: ‘previous projections’ would maybe be better here.

[AR 2.07] We rewrote the sentence into:
‘Explicitly accounting for debris cover has only a minor effect on the projected mass loss, which is in line with previous projections. Despite this small effect, we argue that the improved process representation is of added value when aiming at capturing intra-glacier scales, i.e. spatial mass balance distribution.’

[RC 2.08] Line 15: no need for a ‘-’ between debris and cover.

[AR 2.08] Removed

[RC 2.09] Line 48-51 see the inversion for debris thickness change by Anderson et al. (2021b) as an example of debris thickening. This paper also highlights how the change in direction in flow can lead to debris expansion.

[AR 2.09] cf. AR 2.03

[RC 2.10] Line 86. remove plural from 'glaciers'

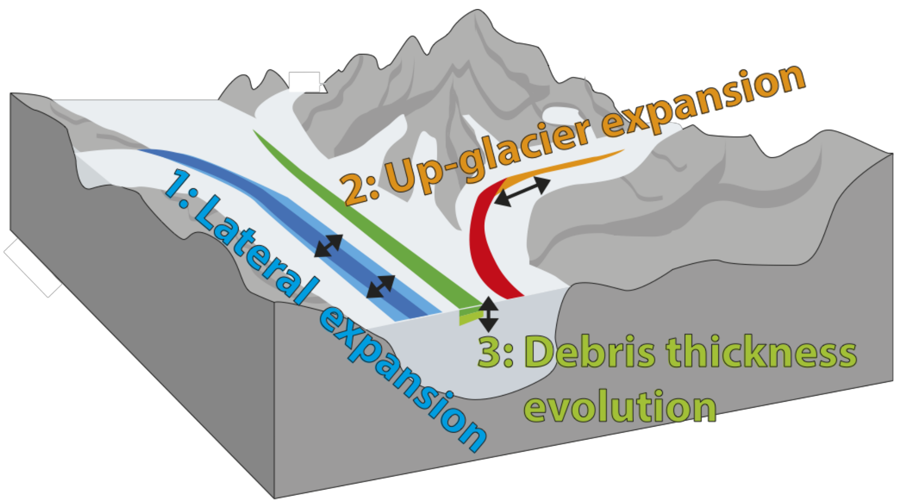
[AR 2.10] Corrected

[RC 2.11] Line 97. What do the variables represent beyond free parameters? I am surprised that Anderson and Anderson (2016) are not cited here. As that work originally derived this form of Ostrem curve and discussed what these free parameters represent in detail. The model is called 'Hyper-fit' in Anderson et al., (2021a) and Anderson et al. (2021b). I hate to have to say this by as the text is written this is plagiarism. Please cite this appropriately.

[AR 2.11] The two parameters have been fitted to results of an energy-balance model. For more details see AR 2.02. We have now acknowledged that this is similar to the 'Hyper-fit' model, although we derived, formulated, and interpreted the form of the equation independently.

[RC 2.12] Figure 3. The text in the figure is not legible in places.

[AR 2.12] We increase the font size and added a black border to the text in order to make it more legible.



[RC 2.13] Line 205. not sure what 'from the surroundings' means here.

[AR 2.13] We re-wrote the sentence into:
l. 227-228: *'Areas with abundant debris cover may grow faster due to enhanced debris supply from melt-out, or due to ice flow changes (Anderson, 2000; Anderson et al. 2021b)'*

[RC 2.14] Line 207. Deline (2005) is a valuable citation here.

[AR 2.14] We added the suggested reference in the manuscript.

[RC 2.15] (i) Section 3.2.2 this is an interesting parameterization but it should be stated that you assume that debris expansion is directly related to ELA change. **(ii)** As far as I am aware there are no datasets that show this as a direct relationship. Snowline change on glaciers is instantaneous but debris melt out does not need to be. **(iii)** This relationship is dependent on where englacial debris is present within the glacier.

[AR 2.15] (i) Yes, with this parametrization we assume that the debris expansion is directly related to the ELA change. We added this information to the manuscript:
l. 228-229: *'We assume that this is related to the rise of the ELA and to the melt-out of debris in areas that transit from the accumulation to the ablation zone (Anderson, 2000).'*

(ii) We agree that snowline change on glaciers is instantaneous but debris melt out does not need to be, therefore we compute the moving average of the ELA change over the last 10 years.

(iii) The parameterization expands the debris upglacier only from where it already exists (i.e. not in the middle of a clean ice glacier), thus accounting for the fact that englacial debris must not be present in every glacier or everywhere on the glacier.
Note that we reformulated also other parts of the section, as requested by reviewer 1 (cf. AR 1.16)

[RC 2.16] Line 221. "As for the lateral expansion of debris, the evolution of debris thickness is linked to internal debris concentration and glacier mass balance (e.g. Gibson et al., 2017; Mölg et al., 2019; Verhaegen et al., 2020)."
Please see Anderson et al. (2021b) for a detailed process-based simulation of debris thickening that shows the importance of debris advection and compression (both highly dependent on surface velocities) as well as debris melt out.

[AR 2.16] We added this information. We also discussed it in detail in AR 2.04.
l. 247-249: *'As for the lateral expansion of debris, the evolution is linked to internal debris concentration and glacier mass balance (e.g. Gibson et al., 2017; Mölg et al., 2019; Verhaegen et al., 2020), as well as to changes in ice flow velocity (e.g. Anderson et al., 2021b)'*

[RC 2.17] Line 221. Also the change in surface velocity of the glacier: see Anderson et al (2021b)

[AR 2.17] cf. AR 2.03 and cf. AR 2.16

[RC 2.18] 252. 'w.e.a –1' add a space.

[AR 2.18] We added the spaces (w.e. a^{-1})

[RC 2.19] 270. At this point I had forgotten what S1 was. Might be helpful for readers to remind them here?

[AR 2.19] We added what set S1 is:
'To determine $c_{lateral}$, we use the debris-cover observations obtained from the Landsat scenes (set S1, composed of 55 glaciers with debris. See section 2 and e.g. Fig. 4a).'

[RC 2.20] Section 4.2.2 This is a clever approach but again this is emphasizing the role of debris melt out as the only process that causes debris thickening. It is again hard to know though the validity of the McCarthy dataset without access to the errors from in situ debris thickness measurements.

[AR 2.20] Please refer to AR 2.03.

[RC 2.21] Section 5.2 It would be helpful to remind the reader where the pre-2000s climate forcings are coming from for the evaluation of the debris change. Maybe just re-state it or reference the section.

[AR 2.21] Before 2020 we always use ERA5 reanalysis. We added this information in the manuscript.
l. 343-345: *'To evaluate the parametrization for lateral debris expansion, 18 glaciers (set S1 with Hexagon satellite observations) within three sub-regions (Central Himalaya, West Himalaya and West Tien Shan) are simulated from 1974 to 2020 (forcing the model with ERA5 climate).'*

[RC 2.22] 313. So you evaluate the lateral expansion parameretrization against the data that you used to tune it? I wonder how the parameterization works on glaciers that are independent of the tuning dataset?

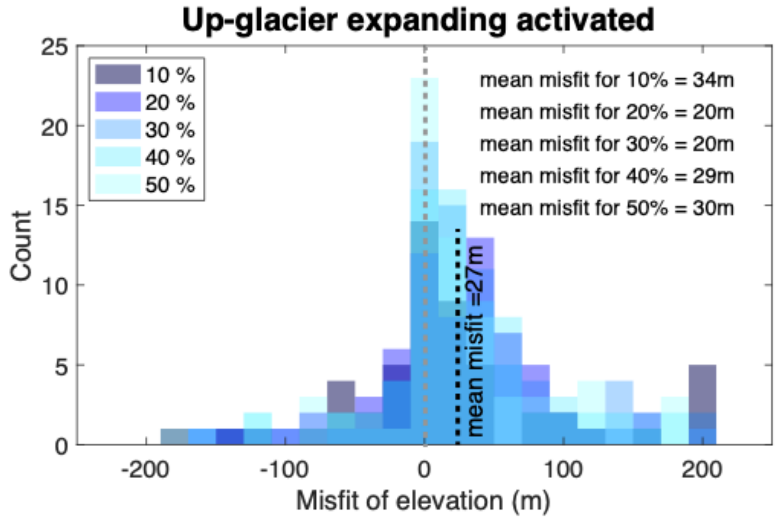
[AR 2.22] Please see AR 2.04 b

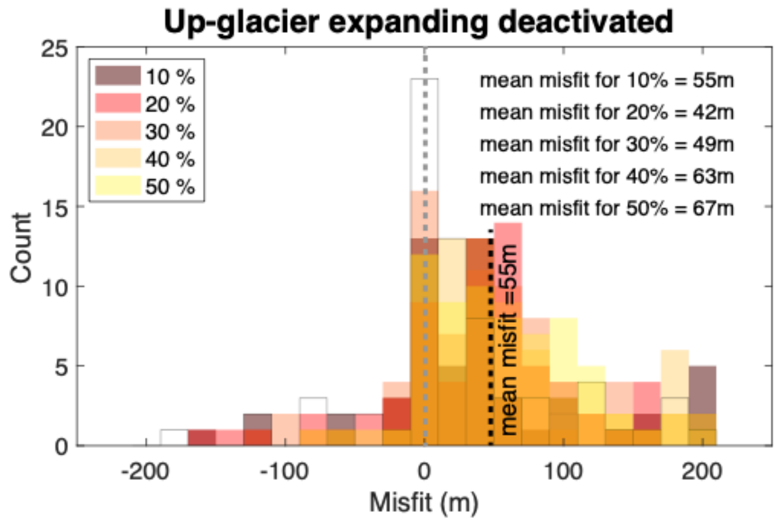
[RC 2.23] 344-346. This is an overstatement. Being off by 10 cm or ~20 cm of debris thickness, using the h_star values (same as (k_debris) in Equation 1) (the debris thickness change needed to reduce sub-debris melt rates by 50%) from A and A (2106) mean sub-debris melt rates are off by 50 to 100% or 200 to 400%. This uses h_star values = 5 cm and 10 cm. These percentages will be even bigger with melt amplification effects included. Since the evaluation of the debris thickness change estimates are coming
from the upglacier end of debris covers the errors are actually quite large.

[AR 2.23] We need evaluation data for debris thickness changes and englacial concentration that are representative for different parts of debris-covered glaciers, including the debris emergence region (Stewart et al., 2020) as well as the terminus. It is noteworthy that a single debris concentration is a key step forward for this (e.g. Anderson et al, 2021b) but insufficient given the variations in englacial debris concentration (e.g. Anderson and Anderson, 2018; Miles et al, 2021).

[RC 2.24] Figure 7. It would be helpful for the reader if the x axis was extended beyond +-200 m

[AR 2.24] We changed the limit of the axis to +-250 m, highlighting that there is no data beyond this range.





[RC 2.25] 358. In the ablation zone or across the whole glacier?

[AR 2.25] In the ablation zone. We changed the sentence into:
L. 389-390: *'This leads to a nearly homogeneous downwasting of the ablation zone (e.g. Pellicciotti et al., 2015; Ragettli et al., 2016) rather than to a retreat of the terminus (e.g. Benn et al., 2012).'*

[RC 2.26] 439-440. "In fact, accounting for the debris cover explicitly enables the model to correctly represent the driving processes, rather than compensating the lack of model capabilities through a suitable parameter choice."
I suggest that this be re-written as it is a significant overstatement from my reading. The explicit model presented here neglects the role of debris advection and compression and is evaluated with highly uncertain modelled debris thickness estimates. I would replace 'correctly' throughout this paragraph.

[AR 2.26] We adapted the sentence in order to avoid the impression of an overstatement. As suggested by the reviewer, we also adapted the rest of this paragraph:
l.473-479: *'In fact, accounting for the debris cover explicitly enables the model to mimic the driving processes, rather than compensating the lack of model capabilities through a suitable parameter choice. This is important, especially when results other than area and volume changes are of interest. Indeed, quantities such as the local mass balance, the glaciers' ice flow velocity and mass turnover, the glacier's length change or water runoff, are only captured adequately when explicitly accounting for supraglacial debris and its temporal evolution. These quantities, in turn, have to be*

modelled appropriately when aiming at anticipating other glacier-related processes, such as hazards from ice-dammed or proglacial lakes, or potential slope instabilities.'

[RC 2.27] 446. Citations would be helpful here.

[AR 2.27] We added two citations: Kraaijenbrink et al. and 2017 and Rounce et al., 2021. l.480-482: *'Compared to the static representation of supraglacial debris cover that is presently included in some regional to global glacier models (e.g. Kraaijenbrink et al., 2017; Rounce et al., 2021), the expected increase in both debris-cover fraction and local debris thickness will enhance the insulating effects of the debris cover.*

[RC 2.28] Section 7.2 Nice to have this clear statement of the sensitivity!

[AR 2.28] We thank the reviewer.

[RC 2.29] 457. Also from the assumptions held within each parameterization.

[AR 2.29] True. We added a paragraph describing this (cf. AR 2.03 v)

[RC 2.30] Section 7.3 Interesting analysis/results.

[AR 2.30] We thank the reviewer.

[RC 2.31] 519. when only debris melt out is included.

[AR 2.31] We added 'with our approach' to clarify:
l. 563-564*:'Averaged over the transient glacier area, our approach anticipates the mean debris thickness to increase only slightly'*

[RC 2.32] 528. typo

[AR 2.32] We are somewhat embarrassed but can't find any typo at this line. Possibly it is a different line than former L. 528?

REFERENCE:

Compagno, L., Zekollari, H., Huss, M. & Farinotti, D. Limited impact of climate forcing products on future glacier evolution in Scandinavia and Iceland. *J. Glaciol.* **67**, 727–743 (2021b).

Stewart, L. G., Lavers, J. L., Grant, M. L., Puskic, P. S. & Bond, A. L. Seasonal ingestion of anthropogenic debris in an urban population of gulls. *Marine Pollution Bulletin* **160**, 111549 (2020).