Thank you for taking the time to revise the manuscript. It is much improved and, from my view, close to being ready for publication. There are however some model descriptions that are not yet accurate and greatly hamper the ability of the reader to understand what the model actually does. Some comparisons to other work are also inaccurate and should be adjusted before the manuscript is ready for publication.

The first relates to the lack of clear explanation regarding the **toe parameterization** used in this study.

Grid size dependence

The toe parameterization presented here is in fact grid-size dependent but it is presented as if it is not in the main manuscript. Note that my comments reference the new updated version of the manuscript, not the marked up version. From Lines 129-131:

"We note that although the terminal boundary condition used here is similar to the one implemented in Anderson and Anderson (2016), since in both cases the glacier ends with a small debris-free terminal ice cliff, our choice of boundary condition has the subtle but important feature of being independent of grid size."

From line 570 describing the toe parameterization used in this paper:

"This boundary condition is found to be grid size independent in the sense that it results in steady state glacier extents that converge to a constant value as the grid size is decreased."

A strong contrast is drawn in the text between the toe parameterization approach of A & A (2016) and the one presented here (see lines 130-131 and 574-577). The grid size dependence is explained in A & A (2016):

"In our model, the debris thickness
h debris (x, t) represents a layer of equal thickness on any cell.
Debris thickens more slowly with a larger dx because the
debris volume advected into a cell is spread over a larger
area (due to the larger dx; dy=1; dy (in m)). There is there-
fore a timescale built into the thickening of debris in a cell
that is dependent on dx. Because ablation rates are sensi-
tive to debris-cover thickness, changing dx has an effect on
glacier evolution."

But there is virtually no difference between the model of A and A (2016) and those that follow and the new model presented here. Both models are grid size dependent and the effect will be reduced in both models as dx is reduced. Please update the text so this false distinction between the models is removed. (see text in the Line-by-line comments below as well).

Accurate explanations of the toe parameterization

The way that the sub-grid ice cliff location is found is not yet explained in Appendix A. Line 564: "Since the critical ice thickness will generally fall between two grid points, a sub-grid interpolation is performed to determine the exact location of the ice cliff"

It seems like some sort of power function is shown in Figure A2 to estimate sub-grid the terminal location. Looking to the code though it looks like the x location of the ice cliff is determined by linear interpolation between the numerical terminus of the glacier and the next cell. This does not match what is shown in Figure A2. Please update Figure A2 to show what actually happens in the code and the text to explain the actual means of interpolation used in the code.

Furthermore it is not clearly stated that the melt in the last cell is based on melt rates that are an average of the glaciated and non-glaciated parts of the cell (see equation A1). This is a non-physical representation of the melt rate at the toe of a DCG. This is not necessarily a problem, it is just not clearly stated. Tying Figure A2 to equation A1 is also needed as right now they do not reflect what actually happens in the code. This includes defining x_*, x_i and x_{i+1} in Figure A2. Without this explanation the reader cannot reproduce the toe method presented in the manuscript.

Line 563 "All the surface debris transported past the icecliff location $x*$ slides down the cliff and out of the system."

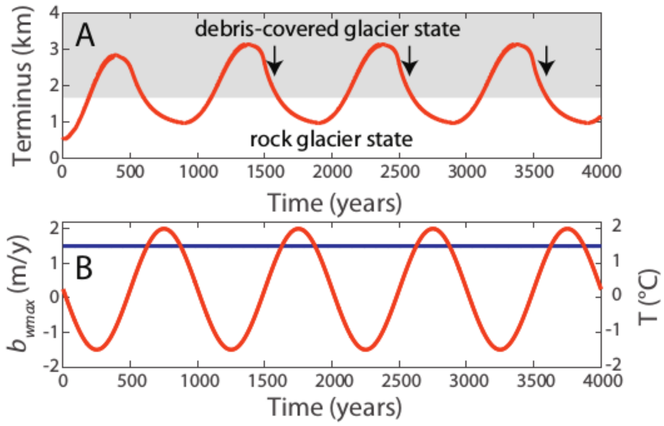
The model real does not represent debris transport at all after cell x_i and there is no ice cliff backwasting that removes debris from the glacier surface. In the actual model any debris that flows beyond cell x_i is removed from the glacier because it is set to zero in the code. Please update this description as it does not accurately represent how debris is actually removed from the modeled glacier.

There are really only two ways debris can leave a glacier 1) by ice cliff backwasting and 2) by gravitational mass wasting (A&A 2016). These processes are not represented in the model presented here.

Constant englacial debris concentration and DCG memory

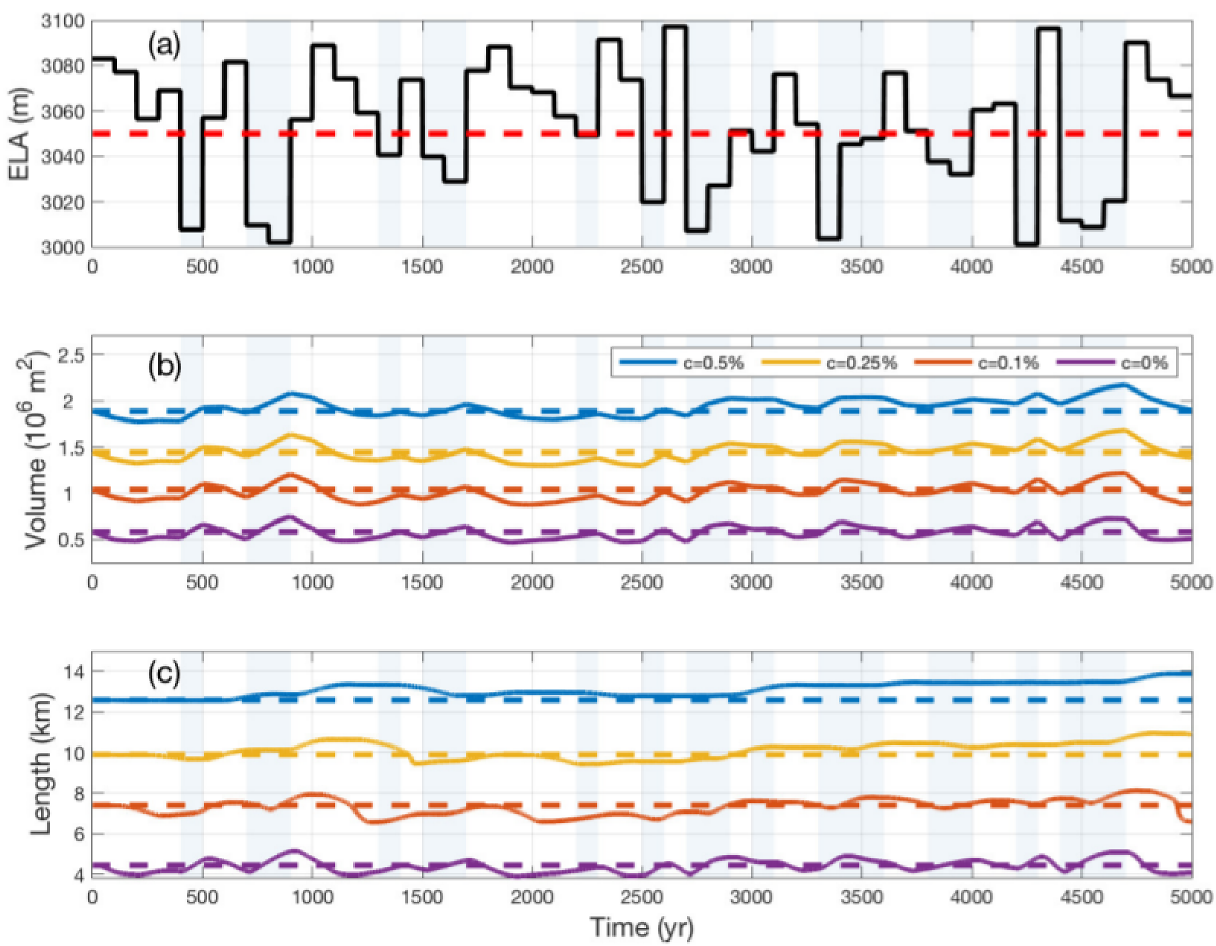
One important conclusion presented in the abstract and conclusion (that cold periods will disproportionately increase DCG length/size causing a form of hysteresis) is consistent with the effects of the assumed constant, uniform englacial debris concentration. As a glacier grows due to colder temps, more debris will be present in the glacier across entire ablation zone than should be for a steady headwall erosion rate. This will tend to increase the glacier size due to a non-physical increase in debris on the glacier surface. Because this hysteresis is consistent with the effects of constant englacial debris concentration in a glacier that is changing size I am not convinced that this hysteresis is a real effect for DCGs. The non-physical nature of the toe parameterization may also contribute to this but it is hard to evaluate this without taking into account the conservation of debris mass in the model presented here.

Here is what the model of A &A (2016) does when working through a climate cycle allowing for englacial debris advection (figure 20 from Anderson et al., 2018):



There is very little change in maximum glacier length despite cold and warm cycles. Is the amplification seen in this model is in fact dependent on the constant englacial debris assumption. I note the caveat that this model is meant to represent a small DCG transitioning into a completely DCG (rock glacier) but the enhancement effect is absent outside of the first cycle.

Compared to figure 6 from this study:



Line-by-line comments

Line 56 “has also studied the relationship between debris-covered glaciers and rock glaciers.”

Please add ‘transient’ between ‘the’ and ‘relationship’ otherwise the way this is written makes it seem like the model is only being used in a steady state sense from the text immediately before.

Line 109: “Østrem curve data representative of a medium-sized Alpine glacier”

Consider re-vising this justificatoin as it is unclear what the size of a glacier has to do with the shape of Østrem’s curve.

Line 126: “In addition, we accounted for the fact that the terminal ice velocity in SIA goes to zero, which is not physically realistic, by taking an averaged velocity over several ice thicknesses near the terminus when computing the debris transport.”

How does the ice not go to zero velocity at the toe on a real glacier? If ice thickness is 0 velocity must be zero. How physically does ice flow remove debris from a glacier? This is not the correct process for the removal of debris at the toe, which can only leave by ice cliff retreat or by geomorphic trundling down the glacier slope.

Line 130-132. “We note that although the terminal boundary condition used here is similar to the one implemented in Anderson and Anderson (2016), since in both cases the glacier ends with a small debris-free terminal ice cliff, our choice of boundary condition has the subtle but important feature of being independent of grid size.“

I see in the appendix that the terminus formulation presented in this manuscript is in fact grid size dependent.

Line 132-133: “Further details of the boundary condition, including the interpolation scheme between grid points and convergence tests for different grid sizes, is found in Appendix A.”

I did not find a reproducible explanation of the interpolation scheme and how it differs from A & A, 2016 nor the convergence tests for different grid sizes in Appendix A. Please update accordingly.

Figure 1. Nice addition!

Line 174. Helpful explanation of what is to come.

Line 354. “We showed that the memory of a debris-covered glacier is selective, exhibiting an effective hysteresis, with periods of relatively cold climate having a sustained effect on the volume and in particular on the length.”

While this is a clear result from the modeling and model set up (Fig 6), I am concerned about the effect of an assumed constant englacial debris concentration through time on this conclusion.

Perhaps one way around this issue is to state that the effect you see with glacier size enhancement at cooler temps and glacier shrinking at warmer temps is consistent with the englacial debris concentration changing in time but that it appears to be independent. I am not sure what simulations can be done with this model to show that this hysteresis is in fact not model assumption dependent.

Line 357. “Previous numerical simulations of the transient response of debris-covered glaciers focused only on the effects of sudden debris input in the form of an avalanche (Vacco et al., 2010; Menounos et al., 2013). Such a one-time debris input leads to an advance in glacier extent and foreshadows the results of our study, where a constant debris source and changing climate forcing gives rise to a more complex response.”

This is statement is not accurate and omits a number of other studies of the transient modeling of DCGs in response to climate change.

Other studies have modeled the transient response of DCGs to climate change with a steady input of debris. Not only a one time avalanche of debris. Please the Rowan et al. (2015) model of Khumbu glacier responding to a warming climate. See Anderson et al., 2018 (Glaciation of alpine valleys: The glacier-debris-covered glaciers continuum) and Crump et al., 2017 for examples of modeled glaciers that contradict this above statement.

Studies that show transient response of debris-covered glaciers to climate change:

Rowan et al. (2015). The whole paper focuses on transient responses to climate change with a steady debris deposition rate.

Anderson et al., 2018: See figures 18 and 19 (for a single climate cycle) , and figure 20 for the transient change in a debris-covered glacier over 7 climate cycles).

Crump et al., 2017: See Figures 6 and 7 show a debris covered glacier responding to step changes in ELA over a 4000 year period.

The statements in the text here need to be adjusted to reflect the contributions of these previous efforts.

Line 363. “This also means that for debris covered glaciers, no unique glacier length exists for a given climate, but rather that the length of debris covered glaciers is determined by the history of repeated cold phases. Furthermore, debris-covered glaciers under random climate forcing are expected to have a longer average length than the steady state length corresponding to the equivalent constant climate forcing.”

See comments above.

Line 439. Maybe should be up to 12% here based on Anderson et al., (2021).

“Kennicott Glacier exhibits the highest fractional area of ice cliffs (11.7 %) documented to date.”

Line 476. “Equation (14) is similar to an equation derived by Anderson and Anderson (2018) but with the important difference that we have allowed variable velocity in this derivation, rather than assuming a constant velocity along the entire glacier.”

Which equation are you referring to here from A & A (2018)?

There are two models presented and discussed in A & A (2018). And variable velocities are explicitly discussed and modeled in that paper. So I am confused as to what you are referring to. Please clarify in the text as it creates a false sense that spatially variable velocities were not considered in A & A (2018).

Appendix A

See notes in the main comments above.

Line 575. “A similar boundary condition was used in Anderson and Anderson (2016), where a debris-free region was also employed at the terminus. In their approach, the size of the terminal ice cliff depends directly on the grid spacing but the debris flux can be varied. In contrast, our terminal ice cliff is roughly independent of grid size, since it depends on the critical ice thickness $H*$, but the debris flux is fixed by the local ice velocity.”

This is not correct and should be adjusted. The ice cliff height in A & A 2016 does not vary depending on grid size and the ice cliff height is fixed. There is no debris-free region on the modeled glacier. The grid size dependence is explained in A & A (2016):

“In our model, the debris thickness
h debris (x, t) represents a layer of equal thickness on any cell.
Debris thickens more slowly with a larger dx because the
debris volume advected into a cell is spread over a larger
area (due to the larger dx; dy=1; dy (in m)). There is there-
fore a timescale built into the thickening of debris in a cell
that is dependent on dx. Because ablation rates are sensi-
tive to debris-cover thickness, changing dx has an effect on
glacier evolution.”

This effect is the same in this model as well. As dx is decreased the effect of dx on glacier evolution is reduced. It seems that the model of A &A (2016) and the model presented here have the same issue, namely that there is grid size dependence in the length that converges as *dx* is reduced. Please update the text to accurately represent the similar issues for the models as right now the text states that the model presented here is grid size independent when it is not actually.