Green = reviewer comments, black = responses

We thank the reviewer sincerely for their time in reviewing our manuscript, and for their extremely thorough assessment and detailed comments and suggestions. This is an exceptionally diligent review, and we appreciate it. We have responded with comments interspersed into their review.

SUMMARY

This study investigates the state of disequilibrium of ~5600 glaciers in Alaska using a well published "model"/equation that links mass-balance anomalies and retreat in terms of length. The analyses focus on how different response times and the shape of the climate forcing affect the state of disequilibrium. Overall, they find the glaciers are in a state of severe disequilibrium and thus a lot of retreat is already inevitable in the future.

Overall, the study is highly relevant and provides new knowledge about the state of Alaskan glaciers, thereby making it suitable for The Cryosphere. However, there are considerable areas of the manuscript that could be revised to improve readability and more clearly convey these findings. Specifically, the method descriptions were fairly hard to follow, even though the concept itself is fairly "straightforward" (i.e., there's a model that can determine the state of disequilibrium and this is stated with the f_eq metric). In fact, the "straightforwardness" of the analysis is one of the most eloquent pieces of the analysis; hence, if the methods and readability can be improved, I think it will reach a much larger audience and the key points will come across much better. I thus recommend major revisions.

MAJOR COMMENTS

Section 2 is a bit hard to read, which seems to be due to the structure of the section. For example, the first sentence refers to Figure 1, which ultimately took me to read Figure 1; however, without any context, it was very hard to understand. Another issue with this section is that Figure 1 is technically a result; however, there's no information about the model and methods given yet – that all comes later. It thus feels very confusing and forced me to read it twice in order to understand, i.e., understand the methods first, and then understand this theoretical result. It also seems to provide remarkably detailed justification related to not focusing on natural interannual climate variability without any context. If you want to keep this theoretical case (i.e., results) as a way of describing the concept within this section, I would at a minimum recommend that you restructure the section so that the text is logical to follow in a single read. I would further recommend that you focus on the key elements, and provide detailed justification as relevant later on.

Thank you for this. Both reviewers brought this up, and we agree we could have been less confusing for readers. We've restructured and rewritten this section. We do want to start off with the description/illustration of disequilibrium, rather than beginning with the model, in order to begin with a focus on the concept. The revised version:

- Goes straight to Fig. 1, and clearly walks a reader through what disequilibrium is, and introduces terms that will be used later.
- Rewrites the caption of Fig. 1, and adds a title to be much clearer about what we want a reader to get; and leads the reader through each of the panels.
- Removes extraneous material about natural variability etc. to later in the text.
- Introduces the model specifically as a tool for calculating these disequilibrium factors.

We think this improves the flow and transitions considerably. While we do use the model equations to create the curves in Fig. 1, the concepts are general and not limited to the model.

Section 2 and 3 – check on formatting requirements, but it's a bit odd to see equations that are written in both equations as well as equations written in lines (e.g., L96, 97, 105, L158, etc.).

We see past articles in this journal that include both in-line and separate enumerated equations. In our experience it is common to put minor equations and those not referenced later as in-line equations.

The heavy use of variables in place of actual text reduces readability. Even something as simple as writing out tau as "glacier's response time" or just "response time" or *H* as characteristic thickness could greatly improve accessibility and readability of the manuscript and only add one word. This is also likely part of the reason Section 2 is challenging to read.

Hopefully the layout of Section 2 is now clearer, and that symbols are now introduced without other distractions. It is our standard practice (which we hope we have adhered to consistently) to use symbols where a variable or other mathematically precise term is used repeatedly in the main body of the paper, and to use them consistently thereafter. There are dozens of places where tau and H are used. As we think is standard, we avoid using symbols in the abstract and introduction; and then make sure they are redefined in the summary and discussion. We've also aimed to use symbols that are cognitively linked to what they describe (H,tau, etc.). If the reviewer or editor feel strongly enough, we could introduce a glossary of terms used. It is our feeling that we are not over-the-top in our use of symbols, relative to other similar work.

Can Figures 3, 5, 6, 7 be merged together? This would be nice to able to see the differences. Also, it's interesting that the regional data is provided in Figure 2; however, these regional differences aren't really included within the analysis. It could be interesting to see Figures 3, 5, 6, 7 with these regional colors included within the pdfs. I recognize this would not enable the area-weighted and count pdfs to be done, so there's some discretion as to the important narrative to show. Maybe there could be a Figure 3, 5, 6, 7 pdf with just the counts in the supplement/appendix?

We understand the point here. However, because the figures follow the development of the analysis, we feel it would be more confusing for the reader to encounter information that has not been introduced yet. We've aimed to help readers make comparisons by providing all the figures in a common format.

We've chosen to do Alaska as a whole in the main paper. It would be an extra layer of analysis to show the results by region, and would not necessarily be more insightful. We would have to have separate panels for the PDFs and CDFs, etc.; and we might comment only on a small subset of the extra curves. A reader can get a broad sense of the regional breakdown already because we present a table of results broken down by glacier area (Table A1), and the distribution of areas among the various regions is presented in Figure 2. We've added a sentence in the Summary and Discussion pointing this out. So thank you for the suggestion.

There are many parts where it seems like results are provided in the methods and methods are provided in the results. See specific comments below, but these really break up the flow of the text.

We've chosen not to follow the I-M-R-D-C recipe for paper structure. It is a style preference for us to put methods close to the results that use them. Our apologies if this is jarring! We agree with (both) reviewers that Section 2 was structured in a confusing way, and hope the revisions are an improvement.

Discussion: I was surprised not to see any comparison to prior values of the state of disequilibrium or committed mass losses? It seems like referencing these studies, especially those in Alaska (e.g., Davies et al. 2024, *Nature Communications*) is needed at a minimum. I suppose that this model can only do length changes and not mass change. This may be an important distinction that should be highlighted in the discussion?

The Davies paper has a different aim from ours, we think. It has a specific process focus, and we don't see how it addresses committed retreat (or committed mass loss) in a comparable way; and a reader has to infer that there is disequilibrium rather than it being directly addressed. This comment prompted us to include a citation Mernild et al. (2013), and to Zekollari et al., (2020) which explicitly estimates committed mass loss for the Alps.

There is a general understanding that glaciers must be in disequilibrium, and obviously there are lots of studies that characterize observed rates of retreat, and project continued retreat in the future. But it is an extra step to estimate disequilibrium specifically.

The glacier-wide climatic mass balance is the response of the glacier to the climate. The equilibrium state is thus the geometry that is needed to support a given climatic mass balance gradient.

We are not quite sure that we understand the point the reviewer is making. The glacier-wide mass balance is partly a function of the glacier response, and partly a function of a changing climate. It definitely is a good complementary way to think about the state of (dis)equilibrium, and we'd say the equilibrium geometry is that which gives a glacier-wide balance of zero (assuming that's what is meant by "support"). We focus here on length disequilibrium, which lends itself to analysis when only large-scale geometry is available in the inventories.

LINE COMMENTS

L32 – consider replacing the "-" with "as" to make it easier to read.

We've tweaked this in the revised manuscript.

L42 – This sentence feels out of place as it's a result but stated as a "for example" in the introduction. I don't think it's needed and distracts more than helps.

We appreciate the reaction, but choose to keep it in. We follow the guide style that an introduction should include mention of the results. And we wanted to give a reader some quantitative guidance as to what we mean by "severe disequilibrium".

L55-56 – consider rephrasing. The "…, and for three different assumed …" doesn't make sense.

We've made this change in the revised manuscript.

L70 – this sentence describing the equilibrium length, if I understand this correctly, then this is the length if you were to hold the temperature anomaly at any given point and allow the model to run until it reaches it's equilibrium length? If so, some mention of the timescales being different (i.e., it's a seemingly infinite timescale) is important to note here; otherwise, it looks like the glacier retreats that length immediately which does not make sense.

The reviewer is correct that L'_eq is the equilibrium length change in response to the (fixed) climate change. We are not sure what the reviewer meant by different timescales here. The glacier asymptotes to its new equilibrium at a rate governed by its response time, as can be seen in Fig. 1 when the warming ceases. Hopefully this is all clearer with the restructured section 2.

L70/71 – length anomaly should specify what the anomaly is calculated with respect to, i.e., the initial length.

We replaced the word ‘anomaly’ with ‘change’, as it is easier to understand at this point of the manuscript, and is unambiguous.

L71 – consider replacing the “:” with “as” to make it easier to read.

Thanks, we switched to ‘because’.

L74 – why use years at all? The use of years implies it’s associated with a specific time period. If it’s intentional, i.e., that it reflects the change in temperature from pre-industrial, then this should be stated along with the justification. At the same time, I suspect that this would become exceptionally tricky here given Arctic amplification; hence, you have variable increasing temperatures for different areas, so the justification should make it very clear (and explicit) as to how it’s being selected.

Hopefully this will be clearer in the revised section because there is now less extraneous information. We now state we are considering an “idealized industrial-era warming scenario” starting in 1880, so we think it should be clear that our scenario is inspired by the realistic situation. The detailed amplitudes of warming are not important for the principle of disequilibrium, but give a reader a sense of the kinds of magnitudes being discussed. With the revisions we don’t think a reader is going to be disconcerted by the mention of specific years.

L78 – I don’t know if this level of detail is necessary as the concept is fairly straightforward, but for my own curiosity, what happens for overshoot scenarios? I assume the current text is assuming only scenarios of consistently increasing temperatures?

Temperature trends in this work are monotonically increasing, with the exception of the period 1940-1965 in the ETCW scenario. We’ve considered discontinuous-warming scenarios using the same model in other work (e.g., Fig. 4 of Christian et al., 2022; Roe et al., 2021). The overshoot scenario is interesting to consider and we expect response times and current disequilibrium would affect the response in such a case, but our focus here is on the dynamic state with respect to past warming, rather than potential future scenarios.

L80 – regarding “increase their rate of change until it matches that of the warming trend”, how is that possible given they’re completely different units (e.g., m/yr versus degC/yr)?

Thank you for pointing this out, the phrasing indeed implies an impossible comparison! The revised text is now dimensionally consistent.

L80 – “and so asymptote to as” doesn’t make sense. This reads as though asymptote is a verb. If I understand what’s being implied it’s referring only to the orange line? The current sentence reads quite general so I was very confused since the 75-year response time example certainly has not reached its new equilibrium in this example.

We would maintain that asymptote can be used as a verb in this mathematical context, but for clarity have changed to “approach”. We’ve changed the figure citation to refer specifically to the orange line.

L80 – this also assumes that the glacier is large enough, i.e., there’s enough mass loss (or length change remaining) to enable this.

The reviewer is right, of course. We think that readers will take this as a conceptual illustration of disequilibrium. We mention the impact on glaciers with dramatically changing geometry, and exclude the smallest glaciers from our study because of these kinds of concerns.

L81 – choose a different word than “spun up” as this implies initialization conditions of a model, which is not the case here.

Changed to “fully adjusts” to remove the ambiguity.

L90 – “circumventing uncertainty in some inputs” is clearly model specific. I would state this as other readers may be interested in applying these metrics to other models (for example) and thus this might not apply. Note: see major comment as we still don’t have any knowledge of “some inputs” yet, so this is also confusing to read.

Thanks for this. There is a general tendency for the impact of some uncertainties to be minimized, but the cancellation is only exact for a linear model. We’ve removed “input” and clarified the language.

L102 – can you add the context as to which metric (length, time scale, mass) this “estimate” refers to?

It is disequilibrium in glacier length, now stated.

L106 – “… within 5% of the asymptotic limit (e.g., Fig 1c, orange line)” – this is wonderfully clear! This type of statement following the description of the method is very valuable and I’d encourage it to be earlier.

Thank you. Hopefully the rest of the Section reads more clearly now. We get to the model much faster in the revised version.

L107-108 – This reads as an important result, but in a conceptual/methods section. Suggest moving to the results section.

Thanks for this. We included it here because it is the first time we are talking about the phases of adjustment, and it gives readers something to connect with later. We now also repeat this point in the summary and discussion, as an implication of our results. We agree it is worth emphasizing.

L109 – is this different than the "climate trend" described by L105? If different, specify this. Equation 4 – why is there a box around this?

Thanks. It is the same trend. Language tweaked for clarity. Box removed.

Figure 1 – can you show the initial length on one of these figures? It'd be nice to see the raw metric as opposed to only the anomalies/metrics.

The initial lengths are actually not defined for these calculations. We set the value of beta = 100 (included in the caption). beta = area/(width x thickness) of the initial state, so it applies to any combination of those that gives 100. For a uniform width of 1 km, and a thickness of 200 m, it would be a 20 km long glacier.

L118 – "retreat" not "retreats", no?

We prefer 'retreats' (plural) deliberately to emphasize that each glacier exhibits its own distinct retreat behavior with unique characteristics. Either would make sense in this context but 'retreats' confers a slightly different meaning.

L119-120 – again, it would be useful to show the metrics (and potentially state the differences) instead of only stating the differences here.

We're not quite sure what the reviewer means by showing the metrics. The metrics are shown in the figure. But in terms of magnitude, the fractional equilibration does not depend on beta. The magnitude of 3 km does depend on that choice for beta, but as noted above that value can apply to a family of different glacier geometries. So the intent is for those numbers just to be illustrative.

L137 – What about debris-covered glaciers? I assume their retreat behavior would complicate this?

It is a fair point. Maybe it would be possible to screen for degree of debris cover, but debris cover exists on a continuum, so it is hard to know how to filter it objectively. Debris cover would most likely impact tau via its influence on b_t, but in ways that are hard to account for. We now mention this briefly in Section 3 as a potential filtering criterion, so a reader is at least alerted to this issue.

L139 – This would read better deleting or changing "Lastly" since it refers to the last two.

Thank you, we agree. We've changed this in the revised manuscript.

Figure 2 caption – I assume the color-coding in Figure 1b and 1c refers to the subregions. It'd be good to explicitly state this.

The color schemes in Figs. 1 and 2 are unconnected. We don't think there is a substantial risk of misinterpretation. The captions of each figure are clear about what each color represents. Note that the caption for Fig. 1 has been substantially revised, reducing the chance for misinterpretation.

Figure 2 – Why not use something different besides circles for the long-term mass balance records to avoid confusion?

Thanks, we have changed the shape in the revised manuscript.

Figure 2 – I assume you're aggregating these areas based on the center lat/lon? It looks like a regular grid of some sort. Would be good to state the resolution of this grid.

We use a hexagonal grid (displayed as circular markers for visual clarity) with each marker positioned at the center coordinates of its corresponding grid cell. We have added a concise description of this as a note in the caption.

L154-157 – perhaps good reason to add a CDF (even if it's just the total) to these two plots as well as a right-hand side axis.

We appreciate the suggestion. However, taken together, along with the CDFs included in all the other results, there is a risk of information overload. These two lines of text basically serve to motivate why we present all our results as both number-weighted and area-weighted, PDFs and CDFs. The two included panels together provide intuition about the weighting of the population towards large-area glaciers. We think we'd have to add two new panels with regionally distinguished contributions to the CDFs of the numbers and areas to be consistent. We aren't clear that that would be worth it.

L157 – delete "simple". It's a glacier count. This implies glacier area is complex somehow?

We've revised the text to be more explicit about number weighted and area weighted.

Section 4.1 – Shouldn't this be "Characteristic thickness"?

Thank you! We have made this change.

Figure 3 – change "the method outlined in the main text". If you want to highlight this, state (see Section X) or something similar.

Thank you. We now refer to the specific section.

L209 – I assume you're referring to Order 2 RGI regions? If so, state this.

We mean the glacierized regions indicated in Fig. 2 (text tweaked in the revisions), indicated by the points in Fig. 2.

Figure 4 – Is there a reason that you include the accumulation areas (i.e., > 0 m/yr) when the methods state you're only using the ablation area data? Does it change the numbers you get at all?

The figure documents the available data, and does provide a sense of how constant vertical gradients in mass-balance are, across the ELA zone. Since we are picking a region-wide single number, the figure shows that number is not sensitive to those details.

L233 – no need to repeat the caption. Recommend re-writing such that you just state what the b_t values are and then refer to Figure 5a,b in parentheses.

In most cases, we prefer a direct citation style for figures, although we recognize that preferences differ. In our view, we prefer to lead a reader through a figure, and directly indicate points of focus. The indirect (parenthetical) style expects readers to make some steps by themselves, which risks ambiguity.

L250 – Figure 6 shows … is just repeating the caption.

As noted above, this is our preferred, direct citation, style.

L251 and others – I don't think you need to state (and 90% range) every time. Only at first use.

We prefer to retain this style. We recognize that it is repetitious, but want to direct a reader's attention to this range, as it carries important information about the width of the distributions.

L264 – can you give the areas or just a broad definition here of how "small" is used? < 10 km2 for example?

Thank you for this suggestion. The reviewer's suggested threshold of 10 km^2 aligns well with the glaciers we describe as 'small' in this section. We've added a note in the text to specify this.

L299-314 – this is all methods.

As noted above we are not following the I-M-R-D-C structure. We prefer to introduce methods close to the analyses and results that use them.

Figure 8 – suggest using different colors to avoid confusion with the same colors being area weighted vs. count in Figures 3, 5, etc.

Thank you for this suggestion regarding Figure 8. We agree with the reviewer's suggestion to use different colors in Figure 8 to avoid confusion with the color schemes used in Figures 3, 5, etc. We will revise Figure 8 (and Figure 9) with a distinct color palette to clearly differentiate between the three warming scenarios and the area-weighted vs. count-weighted distinction used elsewhere.

L319 – remove “see text for details” as that’s obvious. Refer to a section if essential.

We agree and will remove the redundant phrase from the figure caption.

L329 – Recommend not using highly unique acronyms like “ETCW” and “GWI” as they greatly impede readability. If writing three to four words is deemed too large, then come up with a new name to refer to them as that’s meaningful.

We choose to retain the initialisms. They are closely connected to what they describe, and they are standards used elsewhere in the literature. Different names would be confusing in that regard. There are 10 usages of ETCW and 15 of GWI. That is frequent enough to merit defining initialisms, we think. Our preferred style is to introduce initialisms if they are used frequently enough, and then to be completely consistent with their usage thereafter.

L336 – The shape of the warming is important compared to the linear shape; however, the “ETCW” and “GWI” shapes provide relatively consistent values. This seems worth mentioning earlier (noted it’s in L341-342) since it currently reads that they’re all highly different until you get there.

Thanks for this reaction, although we don’t think we share it. The two ETCW and GWI scenarios are common in the literature, and merits exploring them both. We guide a reader to Figure 8 starting at 327 (in the submitted manuscript), right after describing how we calculate f_eq. We immediately begin to describe the commonalities in f_eq of the GWI and ETCW scenarios, so we are not sure where we create the impression that the results are all highly different. The similarity of the GWI/ETCW results is a main theme for the remainder of the paper, and is noted (albeit in an indirect way) in the abstract.

L336 – Building on the impact of the shape, I would recommend analyzing this relative to the linear f_eq, i.e., reporting the differences on a per glacier basis. This would make it a lot easier to grasp the statements like L338-340. L330-334 report the 90% range, which is highlighting the glacier variability, but ultimately this whole paragraph is about the differences on a per-glacier scale.

We think the reviewer is suggesting we take the differences between the linear f_eq and the GWI/ETCW f_eq for each glacier and plot the PDF distribution of those differences? If that is right, a plot of differences is going to scramble the data in ways that are hard to interpret. For instance, from Fig. 8 it can be seen that differences are small at both low f_eq and high f_eq (long and short response times, respectively). But if the distribution of differences is plotted, that distinction will be lost. We already show and describe how the warming scenario affects f_eq in Figure 9. So a reader can take the tau of their favorite glacier and see the impact of warming scenarios on disequilibrium.

L343-345 – I don't see where this statement is supported?

Thank you! This comment caused us to reexamine what we had written, which we then realized was actually wrong. The biggest differences in f_eq occur at times when one scenario has experienced cooling (f_eq calculated after a forced cooling can exceed 1 - see attached figure). We deleted the original text, and added a short sentence with the new information.

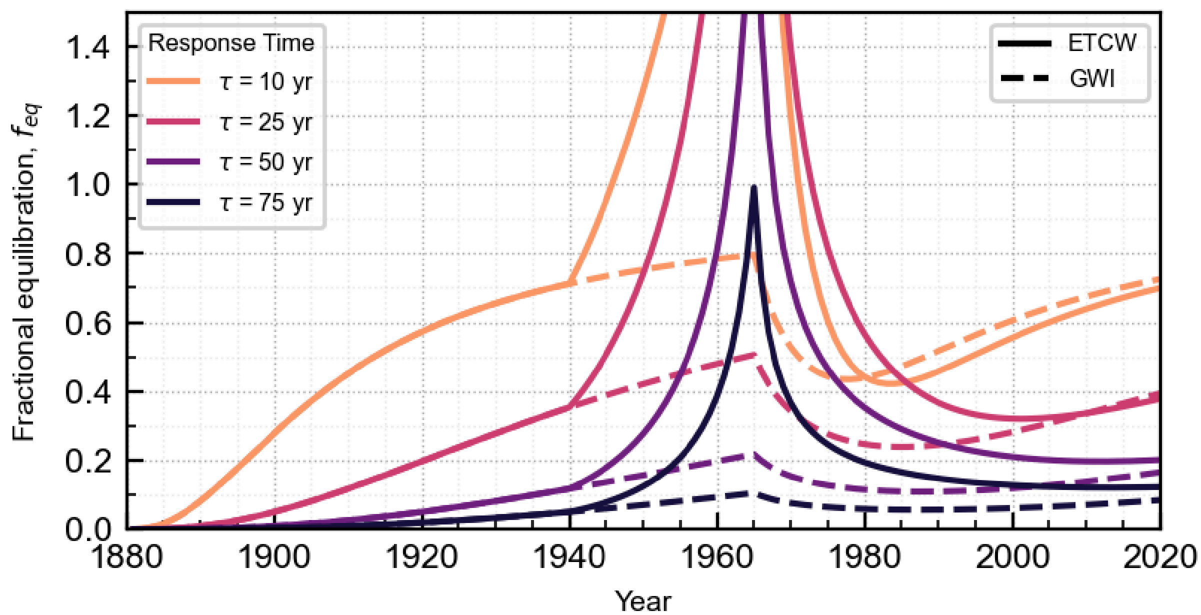

Figure R1. Comparison of f_eq over time between the ETCW (solid lines) and GWI (dashed lines) warming scenarios across an illustrative range of tau (colors). f_eq is identical between warming scenarios up until 1940 because both trends are linear and monotonic over the initial period (see Eq. (4) in the submitted manuscript).

L349 – check formatting.

Thank you! We've corrected the issue in the revised version.

Figure 9 – This looks highly smooth. I assume some function was fit through the data? Please state this in the caption.

These curves are solutions to the model equations, so no data was used. The solutions are smoothly varying functions of tau (see, e.g., eq. (4), which provides the solution for the linear scenario. We added a note in the caption pointing the reader to where we described the method.

L356 – what information is not available? This is quite vague.

We mean only that objective uncertainties do not exist for the thickness or terminus balance rates. Assumptions of what is reasonable have to be made. So we chose to instead simply apply deliberately large uncertainty bounds for both parameters to demonstrate that even under conservative assumptions, our core findings remain robust. This approach allows us to show that the conclusions are not sensitive to the availability of these specific uncertainties.

L356-359 – is this not methods?

As noted above, our style is to intersperse methods throughout the paper.

L350-367 – prior to this section, there are a lot of values provided. Then in this section, there are no values to actually accompany the uncertainty estimates. Instead the uncertainty is just qualitatively described and not shown. Can one provide some values? Maybe for some of these different examples?

We apologize if we are not understanding the point. The uncertainty is quantitatively shown in Figure 9. We are asking a reader to look at Figure 9 to see how uncertainty in tau (i.e., variations on the x axis) and uncertainty in warming scenario (variations among the three lines) affect our target f_eq (shown on the y axis). We provide a reader with our opinion of what a reasonable uncertainty on tau is; we point out specific taus at which uncertainties maximize; and discuss the impact in the small- and large-tau limits. We've added specific guidance in the caption as to how to use the figure to estimate uncertainty in f_eq.

369-375 – Like my prior comment, the interesting thing here is the difference by glacier, not the actual values. Consider changing the way this is provided. For example, are all glaciers 0.08 higher or is it highly variable? Is there not a figure that shows this that can be referenced as well?

Recall this is a population of 5200 glaciers. So a different dataset for H yields some thicker, some thinner. We do describe the general differences in the population distributions of H, and AAR from the different datasets, and then we propagate that into how it impacts the distribution of f_eq for our population. We don't know how to present this glacier-by-glacier. We've chosen to assess the whole population in this study. In the discussion we defend this approach and note that, for any single glacier, we would recommend a more thorough assessment of local information.

L380 – suggest pulling out a few important numbers from the table here to make this more quantitative. Again, it'd be interesting to know the differences, not the values for all of them (see prior comments).

We apologize for being at cross purposes here, but our reaction is the same as above. We think showing the distributions of H, bt, tau in earlier figures helps to illustrate heterogeneity between individual glaciers. But our chosen goal is to assess the population, and we report sensitivity to different datasets in the population statistics. The essential result is that glaciers are in a severe state of disequilibrium, and it is robust to these uncertainties.

L389-392 – Are these statements supported anywhere? Figure 9 does not seem to show this.

Thank you for this comment. We moved most of this material in this paragraph to the appendix. We agree that these details are not needed in the main body of the paper.

L406-411 – Consider rephrasing. It’s a bit odd to state that more detailed numerical modeling can address a problem but then state that you don’t expect it to be any different. In that case, there’s not really a point in having a more detailed numerical model.

We say it might be applied, but we are also aiming to add a note of caution. There is currently a tendency to use ever-more complicated numerical models, but in the face of uncertain inputs it does not guarantee a better answer.

L445 – “other uncertainties in the setting” – provide some examples. Otherwise, highly vague.

Thanks. We’ve been more specific.

L458 – “It seems important” consider making this line stronger. Awareness of how much retreat is already baked into the future evolution of glaciers is important for policy makers, resource managers, and the general public to know to ____. Perhaps that ____ is preserve the glaciers that remain? Reduce our greenhouse gas emissions?

We reflected on this. But we think a stronger statement would be verging on advocacy, which we don’t think is appropriate in a scientific article of this kind. What the public or policy makers choose to do with an understanding of this disequilibrium should be up to them. We see our role is to report on the information.