**Reply to the comments by the anonymous reviewer**

We thank the reviewer for the critical comments that has forced us to do our numerical analysis which, as we try to argue below, have put our results on firmer ground. These criticisms will be of immense help in improving the clarity in the subsequent versions of this manuscript. We are also thankful to the reviewer for pointing us to several important references.

Based on the reviewer's inputs, we have redone our analysis. To summarise the updated results,

1. We have now simulated the response of 810 glaciers out of the total 814 glaciers in the Ganga Basin that are larger than 2 $km^2$.
2. The transient simulation time was extended to 1000 years and the cut off on response time is set at 500 years. Only 9 glaciers, covering 2% of the total area, were omitted this time.
3. After removing glaciers with more than 50% change at the 500 year mark, our final ensemble now includes 703 glaciers.
4. The results presented in the discussion paper remain essentially even with this larger set of glaciers .
5. We checked that the biases in the scaling model for a set of idealised 1d glaciers with the same linear bedrock, mass-balance gradient and Glen's flow-law constant.

More details are provided in out point-by-point replies (in black) to the comments by the anonymous reviewer 2 (in red). For ease of reference, our replies below are marked with [1], [2], … .

In this manuscript, Banerjee and colleagues simulate the evolution of 551 glaciers from the Ganga basin (Himalaya). They start from glaciers in steady state resembling present-day glaciers and force them with a stepwise change in equilibrium line altitude until they evolve into a new steady state over a 500-yr time period. These simulations are performed with a V-A scaling method, with a linear response model and with a 2-D flow model based on the shallow ice approximation (SIA). The authors find that the V-A scaling method, when subject to a time-invariant scaling, underestimates the modelled future loss compared to the simulations performed with the SIA model. From this finding, they suggest that relying on V-A scaling is problematic for studies focusing on the future sea-level contribution from glaciers as this contribution is thus likely to be underestimated. Some of the ideas put forward in this manuscript are interesting, but the manuscript has several problems (some of which are major in my opinion).

[1] At the outset, we respond to the main issue that was raised by the reviewer throughout his report. It was suggested that the biases in scaling models as pointed out in our paper, stem from an (known) inability of these models to capture the magnitude of the change in steady-state volume in response to a step change (in other words, an underestimation of the climate sensitivity of volume/area in scaling methods). The reviewer conjectured that scaling methods do not capture the effects of typically gentler slope and thicker ice near glacier termini and thus, underestimates climate sensitivity. It was suggested that alternative scaling formulations that include slope-dependent corrections should be able to take care of such biases.

First, a similar bias is seen even for a set of glaciers with the same linear bedrock in 1-d (fig. R1). All of these glaciers have the same linear mass-balance profile, and the same rate constant as well. The pattern of the deviations is, in fact, very similar to that in the main figures 1 and 6. Therefore it is unlikely that slope dependent correction to scaling can take care of the biases that we have pointed out. Second, the reviewer did not consider a critical limitation of the scaling model that it, by construction, implies an equality of area and volume response time. This too, cannot be corrected by slope dependent correction.

As demonstrated in our manuscript, the discrepancies in scaling model results (main fig 1,6 and fig R1) are due to the following general limitations:

A) <u>An underestimation of climate sensitivity</u>

We provided numerical evidence in favour of a higher climate sensitivity of glaciers simulated with SIA in main figure 3 and fig R1 above.

(This is something that the reviewer also agrees to. However, the reviewer's explanations based on slope effect cannot be the full explanation as is clear from fig R1.)

B) <u>Area and volume response times are predicted to the same</u>

We provided corresponding theoretical proof in section 2.3 (eq. 10), and numerical evidence in main fig 4.

(Due to this limitation, the transient trajectory of glaciers simulated with scaling models shows a linear trend in V-A log-log plot. In contrast, SIA-derived trajectories are non-linear due to relatively faster changes in volume right after the ELA change. This is apparent from fig R1 and main figures 1 and 6.)

Apart from demonstrating the above limitations of scaling models, the thrust of our manuscript is to come up with an alternative linear-response method that minimises above biases. Of course, we do not claim that there are no other zero-dimensional methods that can be used to achieve the same goal.

We shall include fig R1 in supplementary, and add a clear discussion of the above points.

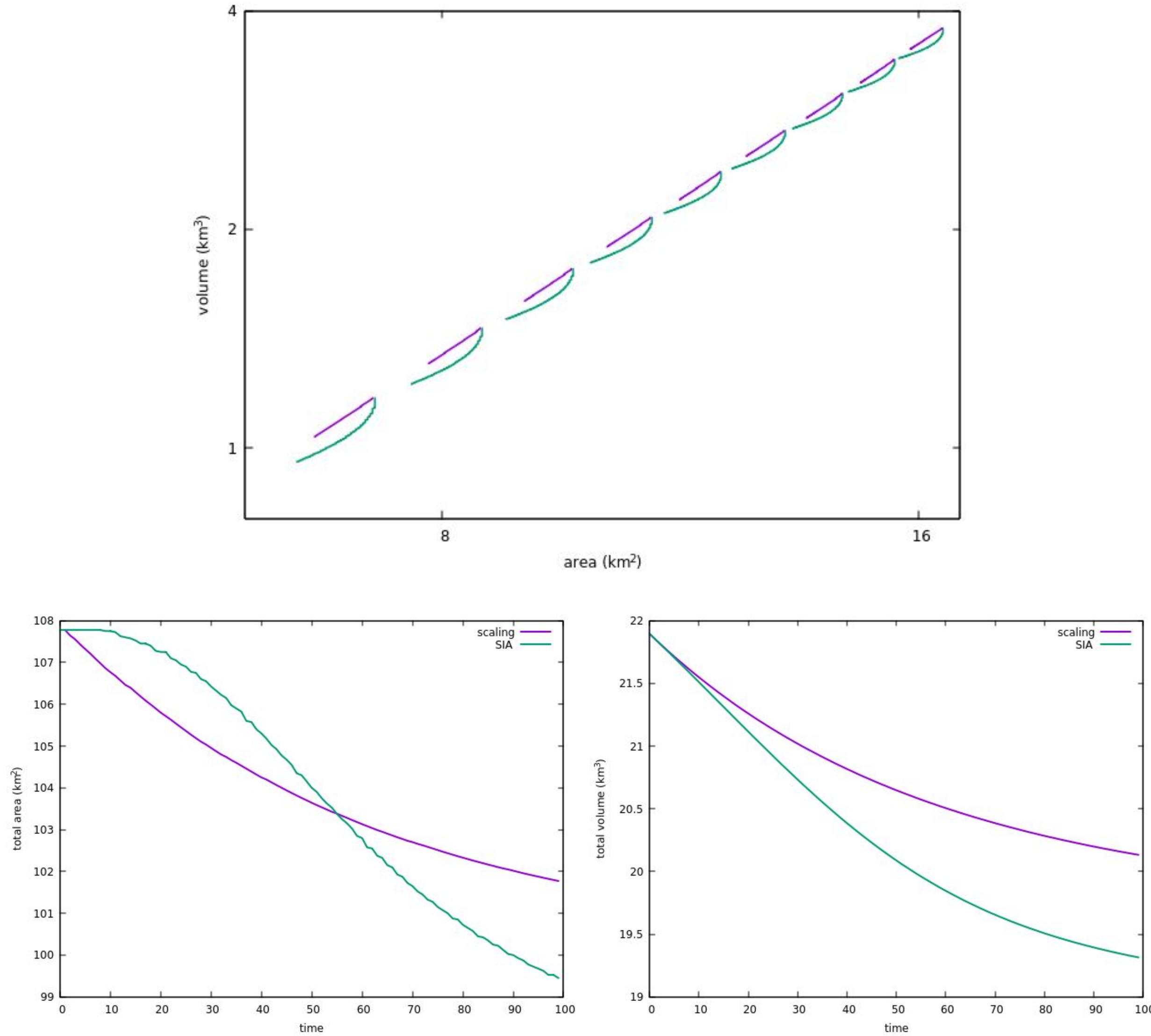



Fig. R1: (top) A comparison of the scaling-based (purple line) and SIA-based (green line) evolution of 9 idealised 1-d glaciers on linear bedrock (slope 0.1) and linear mass-balance profile for 100 years, after a step change in ELA by 50 m. (Bottom) The evolution of total area and total volume with scaling and SIA based models.

**[1] State-of-the art.**

The authors compare the outcome of V-A scaling with results from a SIA model. V-A has indeed been used in some important regional-to global studies in the past (e.g. Marzeion et al., 2012; Radić et al., 2014) due its computational efficiency. Moreover, with spatial estimates of ice thickness lacking for individual glaciers at the time, V-A methods offered a good alternative to estimate the volume of a glacier (and its changes through time). However, increasing computational performance and new glacier-specific inventories on e.g. ice thickness (Huss & Farinotti, 2012; Farinotti et al., 2019)and mass balance (e.g. Brun et al., 2017; Braun et al., 2019; Dussaillant et al., 2019; Zemp et al., 2019), now allow for far more sophisticated methods to simulate the dynamic evolution of glaciers. This includes methods based on imposing observed geometry changes in which the glacier geometry is explicitly accounted for(e.g. Huss & Hock, 2015; Rounce et al., 2020a, 2020b)and more recently also flowline models in which glacier dynamics (i.e. mass transfer within a glacier) are included when projecting glacier changes at regional to global scales (Maussion et al., 2019; Zekollari et al., 2019). When reading this manuscript, it seems like V-A scaling is a state-of-the art approach, and that you compare it to something more sophisticated (2-D SIA model). This comparison would have been very relevant a few years ago, when V-A scaling was state-of-the art (I am for instance thinking about the excellent work realized by Surendra Adhikari during his PhD; see e.g. Adhikari & Marshall, 2012), but has, in my opinion, lost some of its interest by now. With the new glacier-specific ice thickness estimates and other information derived from remote sensing becoming widely available (outlines, surface topography, ice thickness derived from this), the importance of V-A scaling methods is now strongly reducing and is likely to continue doing so so in the near future (see e.g. discussion by Haeberli, 2016). I do therefore have some reservations whether the 'The Cryosphere' is the ideal medium to share these (somewhat outdated?) findings. This concern is furthermore strengthened by my doubts about the experimental setup and the validity of your main conclusions as elaborated in the following points.

[2] We apologise for not mentioning the regional-scale glacier-change studies using various approximate descriptions of ice-flow (Rounce et al, 2020a, 2020b, Zekollari et al. 2019, Clarke et al, 2015, ...). We shall correct this error. We shall also refer to probabilistic and data-based methods of computing future sea-level rise.

As far as computation of future sea-level contribution is concerned, we stand by our statement that "The existing global-scale estimates of the mountain-glacier contribution to sea-level rise mostly rely on low-dimensional approximate parameterisations of the glacier dynamics (van de Wal and Wild, 2001; Raper and Braithwaite, 2006; Hirabayashi et al., 2010; Radi ́c and Hock,2011; Slangen and van de Wal, 2011; Marzeion et al., 2012; Giesen and Oerlemans, 2013; Huss and Hock, 2015; Hock et al, 2019). Several of these parameterisations are based on an statistical area-volume (or area-volume-length) scaling relation for any set of mountain glaciers." A majority of the available estimates of the contribution of glaciers to future sea-level rise is scaling based as of now. The state-of-the-art large-scale glacier models referred to in the reviewer's comment have not yet been employed for sea-level rise computations as far as we know.

We do agree with the reviewer that in the near future one/two/three-dimensional ice-flow models are likely to be used more often for sea-level rise predictions. However, we believe that investigations of the limitations/biases of scaling-based models are still needed, as most of the existing global estimates do rely on such models. For example, 5 out of the 6 models in the intercomparison study by Hock et al. (2019) use some form of scaling. Recent studies of the global-scale vulnerability to sea-level rise continue to utilise results from scaling-based models - e.g., Kulp and Strauss (2019) used scaling-based estimates of glaciers' contribution to sea-level rise by Marzeion et al. (2012). Studies like the present one may be useful in identifying biases in such studies. For example, our analysis indicates that the multi-model mean sea-level change by 2100 as presented in Hock et al (Table 3, 2019) is likely to have a negative bias (This specific point is discussed in detail later in this document).

We do not agree with the reviewer's point-of-view that simpler effective models go outdated as higher-order dynamics become computationally feasible. In several (if not most) branches of science, a hierarchy of models with varying degrees of complexity coexists. In fact, the low-complexity models are often useful in aiding theorecital understanding/development. The need for critical investigations and development of simple 0-d models of glacier dynamics, we believe, cannot be overemphasised. That is the motivation behind the present study, where we not only point out possible biases of scaling-based models, but also present an alternative linear-response model that reduces above biases.

**[2] The experimental setup:**

a. Comparing different methods and models is always quite complicated. This is especially the case when considering 'real' cases (glaciers with real geometries in your case). A study such as the one presented here would have greatly benefited from an idealized setup, which would have made comparisons more straightforward and allowed to disentangle differences between simulations obtained fromV-A scaling and those relying on2-D SIA modelling:see e.g. Leysinger Vieli and Gudmundsson (2004)and Adhikari and Marshall (2012). Here a 'selection' of glaciers is considered, due to some 'problems' occurring when considering all glaciers in the region (see point 2b), which makes it even questionable how representative these are for this given region. With idealized glacier geometries, you could have explored the effect of glacier size, surface slope,...on the discrepancies between V-A based results and SIA modelled results more carefully.

[3] The discrepancy between scaling based and SIA based models are quite general in nature, and therefore, are expected to be present for any slope, geometry, mass-balance profile, rate-constant etc. While the magnitude of the bias is likely to depend on some of these factors as suggested by the reviewer, our aim in this paper is not to analyse such dependencies. The objective of the present study is to demonstrate that,

*1) The above bias could be significant for a set of real glaciers when a certain type of scaling models are used for predicting long-term glacier change or sea-level rise.*

*2) A linear-response model calibrated using 2d SIA outputs can be a viable alternative.*

Our experimental design is more suited for the above purpose than one using idealised geometries as suggested by the reviewer.

Also, our paper contains three main results:

A) The central assumption of a time-dependent scale factor is violated by SIA-simulated glaciers (main fig. 1),

B) A scaling-based model underestimates the climate sensitivity and response time of glaciers in comparison with SIA (main figs. 2 and 4) and assumes an equality of area and volume response time, and

C) a suitably calibrated linear-response model is free of the above low bias (main fig. 6).

A simplified glacier geometry may be useful to highlight the differences between SIA and scaling-based simulations (i.e. results A and B above), but it does not help in obtaining parameterisations of the linear-response properties and testing the performance of the corresponding model for a set of realistic glaciers (result C above).

As shown in fig R1, the same deviations of scaling model is indeed seen for idealised 1-d glaciers having the same bedrock slope - just as suggested by the reviewer. This does prove the general nature of the limitations of scaling-based models and supports our results obtained from simulations of 2-d glaciers.

A discussion about the omitted glaciers and evidence in favour of the representativeness of the selected glaciers are presented later (replies [5] and [6]). Also, now we have repeated our computation for the 814 glaciers, and now have 703 glaciers in the final ensemble (reply [9]). In the updated version of main fig. 6 below (fig R2) that uses the present ensemble of 703 glaciers obtain similar results as that from the earlier set of 551 glaciers.

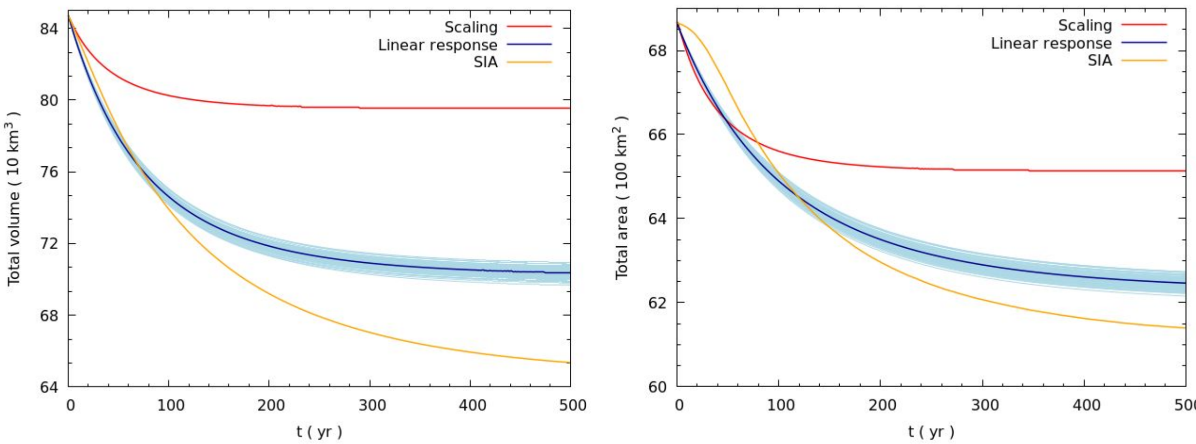



Figure R2: The evolution of the total glacier volume (A), and (B) area for the ensemble of 703 glaciers simulated with three different methods, namely, SIA, scaling and linear-response model, are shown with orange, red, and blue solid lines, respectively.

b. Several arbitrary steps and decisions are made in the manuscript. A few examples of decisions that are hard to understand / seem not well funded:

[4] We describe the rationale for each of the steps below.

ol. 181-182: you exclude glaciers with a large change in area over the 500-year time period? Why? This seems arbitrary, but you must have a reason for this. Moreover, how this this influence your results? This makes the sample less representative …

[5] By definition, a linear-response model is only applicable when fractional changes are small (e.g. Oerlamans, 2005). So to apply 1) SIA, 2) scaling, and 3) linear-response models on the same set of glaciers, we have to exclude glaciers with large changes that are not described by linear- response models.

The threshold of 50% change that is used to exclude glaciers with large change, is indeed arbitrary. However, we confirm that our results do not depend on the specific value of the threshold chosen as (Fig. R3 below).

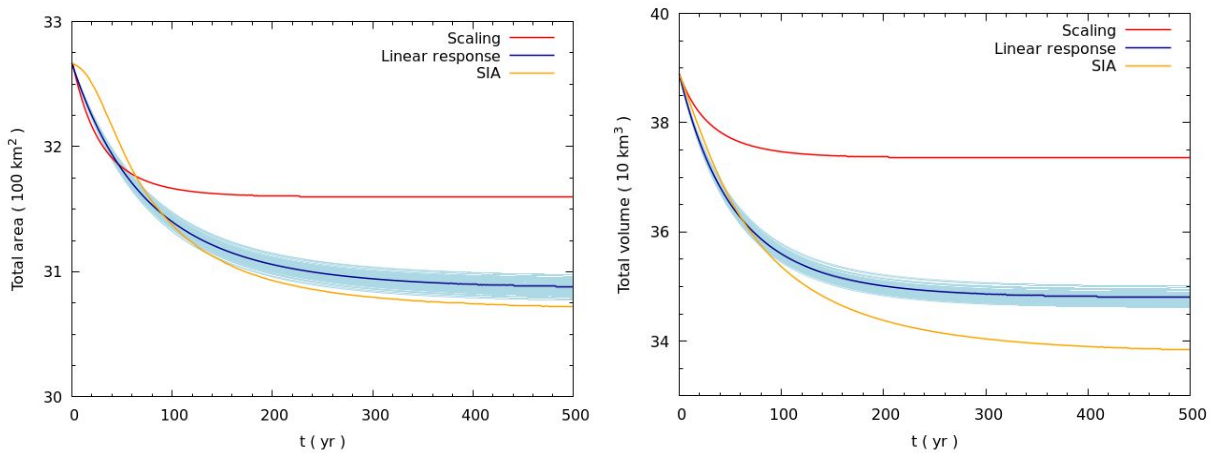
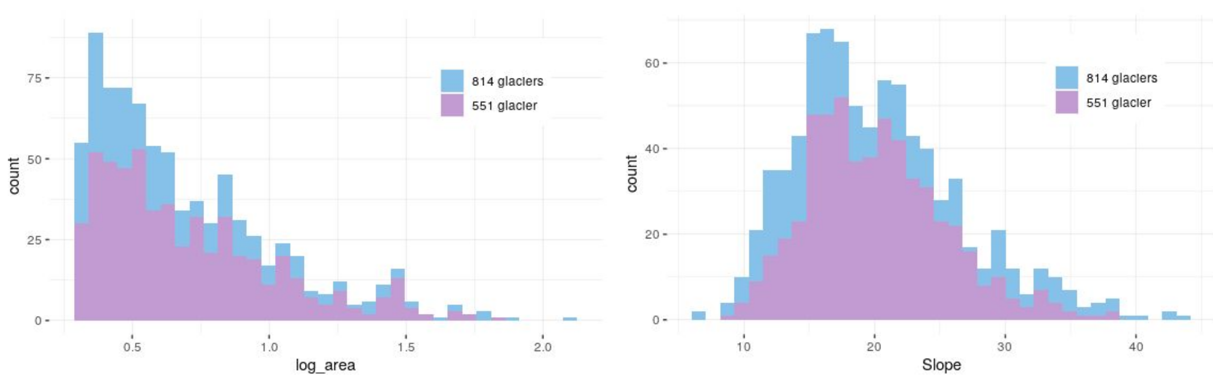



Figure R3. A comparison of the volume and area evolution computed with the three models for the set of glaciers with less than 20% change.

We had checked the representativeness of the selected 551 glaciers. As shown below, the set of 551 glaciers considered has reasonably similar distributions of area, and mean slope when compared to corresponding distributions for the full set of 814 glaciers.

Motivated by the above criticism, we have rerun our simulations, and included 703 out of the total 814 glaciers in our analysis (reply [9]).



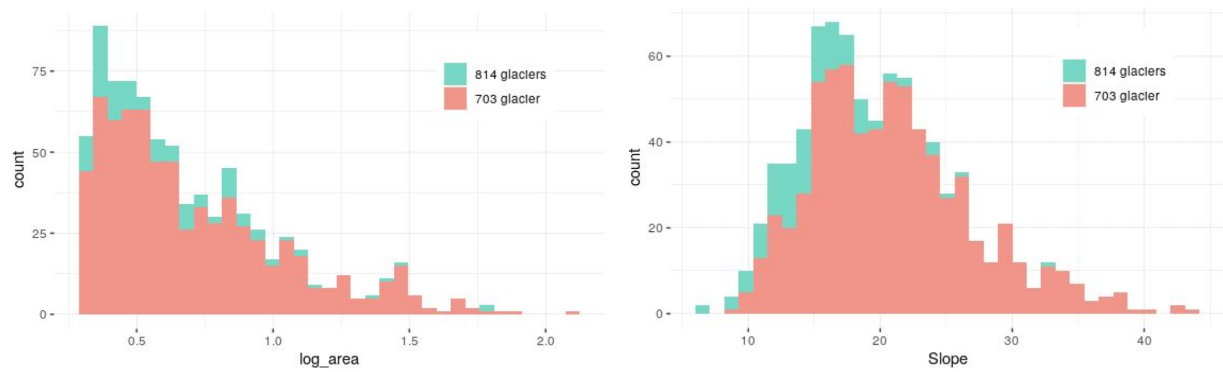



Figure R4. Frequency distribution of glacier area (left column) and slope (right column) for all the glaciers and the selected ones. The top row shows the comparison for the set of 551 selected glaciers analysed in the discussion paper. The bottom panel shows the same for the 703 glaciers that are being used now in the updated simulations.

ol. 182-183: why do you exclude glaciers with long response times? Again, this makes your sample less representative (you probably exclude a certain type of glaciers, likely those that are gently sloping: see e.g. Haeberli & Hoelzle, 1995). Is this because these glaciers are not in steady state after 500 years? If so, you should simply run your experiments for longer and not exclude these glaciers.

[6] The slope distribution of all the 814 glaciers and that of the selected 551 glaciers are quite similar as shown in fig. R4 above, indicating that the chosen glaciers do form a representative set.

We do not require the simulated glaciers to reach a steady state. However, the response time has to be smaller than the simulation period (please see fig 3 in our response to reviewer 1).

Based on the above criticism, we have now extended the runtime of the transient simulations to 1000 years. This allows for setting the response time cutoff at 500 years. This excludes 10 glaciers that cover only about 2.5% of the total area.

We shall include figs. R3 and R4 in the supplementary.

o Figure 1: you show ‘200 randomly chosen glaciers’: why? Should show them all!

[7] The figure with all the glaciers (and without the factor-of-10 scaling) is shown below (fig R5). With all the glaciers included the figure becomes somewhat cluttered. However, based on the above suggestion we shall show all the glaciers now to avoid

confusion, and shall include another figure with a few glaciers the supplementary (Fig. R5).

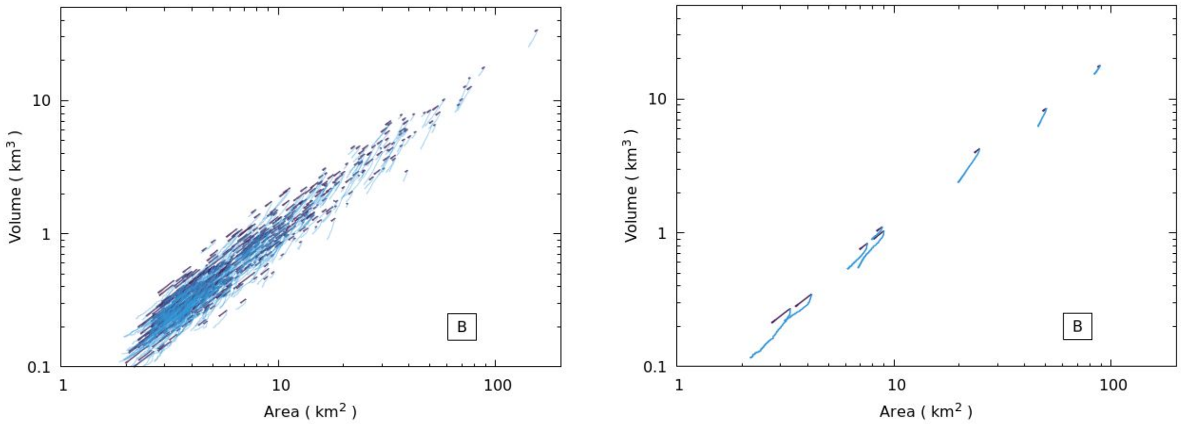



Figure R5: (left) Updated version of figure 1b with all 703 glaciers shown. (right ) the same plot, but for a random selection of 8 glaciers. In both the plots, purple (blue) line denotes scaling (SIA) results.

o l.249-250: ‘In fig. 2b, about 30 data points,...were not included in the fit’: why? You mention something about possibly creating a bias in the linear fit in the next sentence, but I do not see where this would result from / what the problem could be.

[8] Based on comments from both the reviewers, we have removed all such cut-offs used while fitting for the glacier response properties (the excluded points in main figs. 2, 3, and 5). This leads to small changes in the best-fit coefficients in the expressions for linear-response properties, but does not impact our basic conclusions. With the present set of 703 glaciers our results for linear response properties are as follows,

$$\frac{\Delta V_\infty}{V} = (1.93 \pm 0.02)\frac{\Delta A_\infty}{A}; \; \frac{\Delta V_\infty}{V} = (1.71 \pm 0.03)\alpha^*;$$

$$\tau_V = (0.687 \pm 0.004)\tau_A; \; \tau_A = (2.56 \pm 0.04)\tau^*$$

(please see the discussion paper for the notation)

c. The Setup of your SIA model is not fully clear.

oYou mention that for > 100 cases ‘our algorithm for finding a steady-state similar to present extent did not converge or the final steady state glacier geometry was not realistic’: how is this possible? How can a simple SIA solution not ‘converge’ to steady state(in fact, even analytical solutions may exist that do not even require running the

SIA model to find the steady state: see e.g. Jouvet & Bueler, 2012)? And what do you consider ‘not being realistic’?

[9] First of all, in response to the above criticism, we have now included 810 out of total 814 glaciers in our analysis as discussed below.

In several of the glaciers that were omitted in the discussion paper, the problems were due to incorrect glacier boundaries in RGI6.0. These are mostly due to mapping errors where multiple glaciers have been merged fully or partially, parts of a glacier have in excluded, or debri-covered parts have been missed. In few other cases, a noisy bedrock led to a violation of ice-conservation (fig. R5). It is a well known problem with SIA that it violates mass conservation in regions with rugged topography. We performed a 3x3 moving window smoothing of berock to avoid this issue, before starting the simulation with an ice-free initial state. A few glaciers where such non-conserving behaviour was present could be identified as we tracked ice conservation explicitly (reply [10] below). At present there are two glaciers where violation of mass conservation is seen.

For a few glaciers for which a steady-state ELA could not be found, it was a limitation of the simulation time. For each of the 814 glaciers, we vared ELA to find a steady state with extent similar to that of the RGI6.0 glacier outline. This step was numerically expensive as for each of the trial values of ELA we needed to run the model long enough to check if a steady state is reached. There is no intrinsic issue with our implementation of SIA that prevents steady state, it is only related to runtime of the ELA tuning algorithm. Earlier we did not take up the task of finding here primarily because existing algorithm was able to simulate 694 out of all the 814 glaciers, yielding an ensemble that is large enough to test the performance of scaling-based and linear response models. The fact that the final set of 551 glaciers have very similar slope and area distributions compared to the set of all 814 glaciers supports that claim (fig R3).

Responding to the above criticism, we have now rerun the simulations for all 814 glaciers, and obtained initial steady state for 811 out of them. In the remaining 3 glaciers, two showed conservation error due to steep bedrock and were removed (Fig R6). For the other one, a steady state could not be found even with extended run.

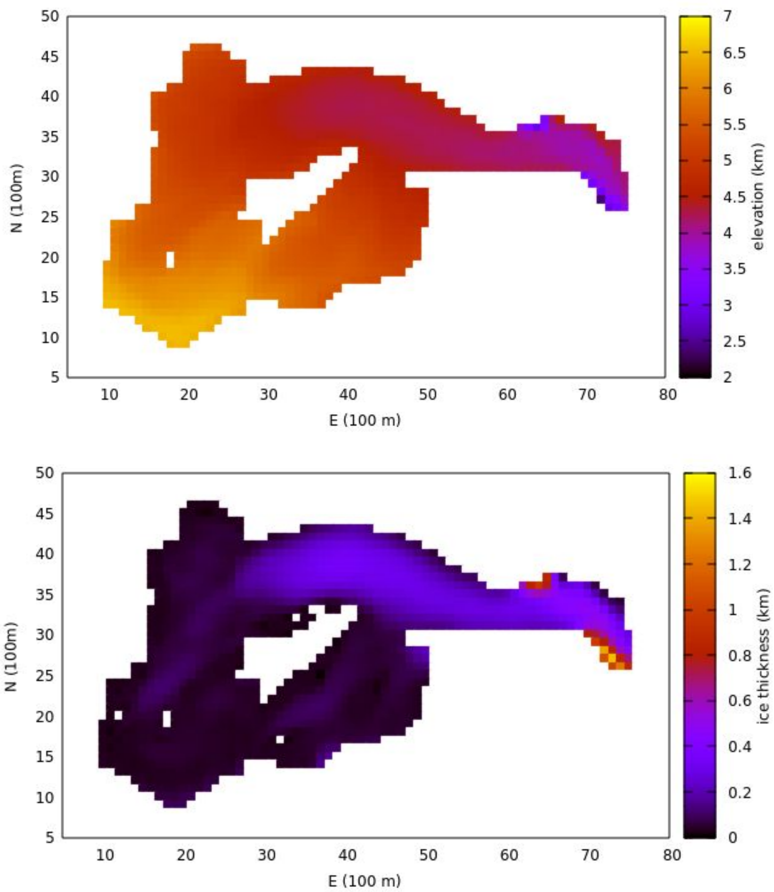



Fig R6: One of the two glaciers (RGIID_15.04060) where ice conservation was violated due to noisy bedock (top plot) near the terminus leading to thick ice there (bottom plot).

There was another glacier, where part of the debris-covered ablation zone of a tributary was absent in the RGI outline and that lead to very thick ice in the truncated tributary (fig R7). This glacier was also removed. That left us with a total of 810 glaciers where response to step-change in ELA was simulated.

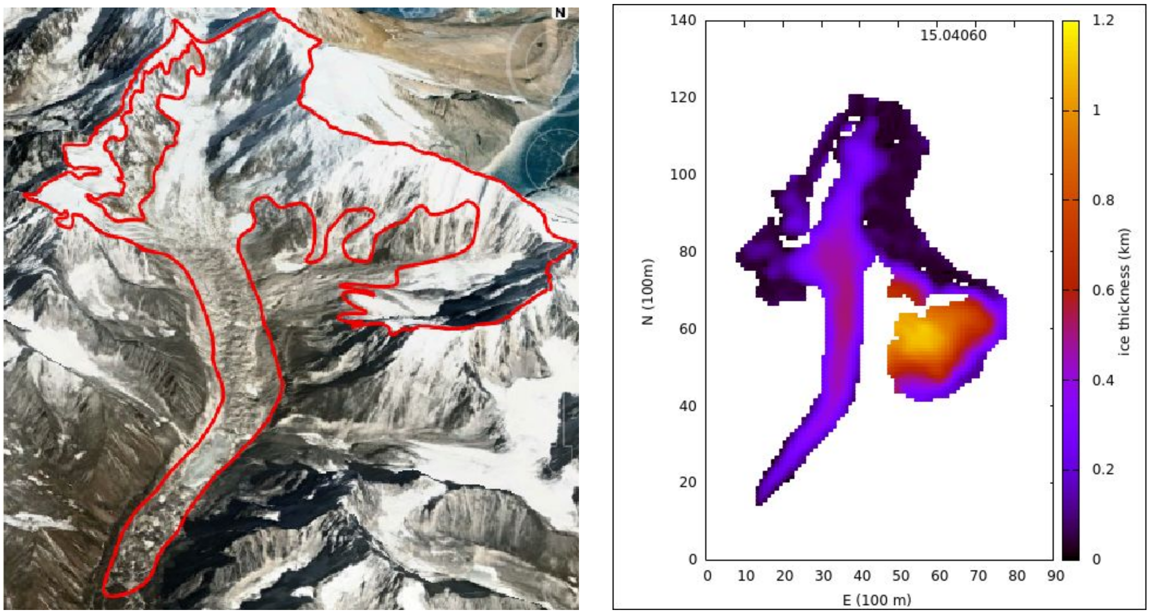



Figure R7: The outline of the glacier Rgi-15.04060 overlain on googl-earth (left) and the corresponding simulated ice thickness map. The RGI outline has incorrectly truncated debris-covered tributaries on the left. This blocks the ice-flow path and thickens the ice there in the SIA simulation. This glacier is not considered in our analysis.

After applying a criterion of less than 50% change (see reply [5]) and a 500 year cut-off on response time (reply [6]), we now have a total of 703 glaciers in our final set covering 89% of the total area.

Which boundary conditions did you use to ensure mass conservation (e.g. to ensure specific ice-free regions do not become ice-covered)? You mention that mass conservation was monitored (l. 162-163): but how do you do this (this is not so straightforward to do...)? Did you check that the integrated SMB over your glacier is zero for the steady states (which it should be)?

[10] A no-flux boundary condition was used as discussed in reply [9] above. The domain boundary for each of the glaciers using the RGI 6.0 boundary. We perform a 3x3 moving-window centrally-weighted smoothing of the berock before starting the simulation with an ice-free initial state to minimise cases where ice-conservation is violated.

We implement a straightforward algorithm to check mass conservation over the domain which implies,

*Total ice present =*

*Total accumulation over the simulation period*

*− Total melting over the glacier simulation period*

− *fluxes out of the glacier boundary into the ice-free part of the domain.*

Each of the terms in the above equation was computed numerically at every time step. It was confirmed that ice is conserved at every time step during the transient evolution up to a fractional error of ~O($10^{-9}$) for 812 out of 814 glaciers (e.g., Fig R8). Only on two glaciers conservation was violated due to noisy bedrock (e.g., Fig R6). These two were left out of our analysis.

We confirm that the total accumulation equals total ablation in steady state.

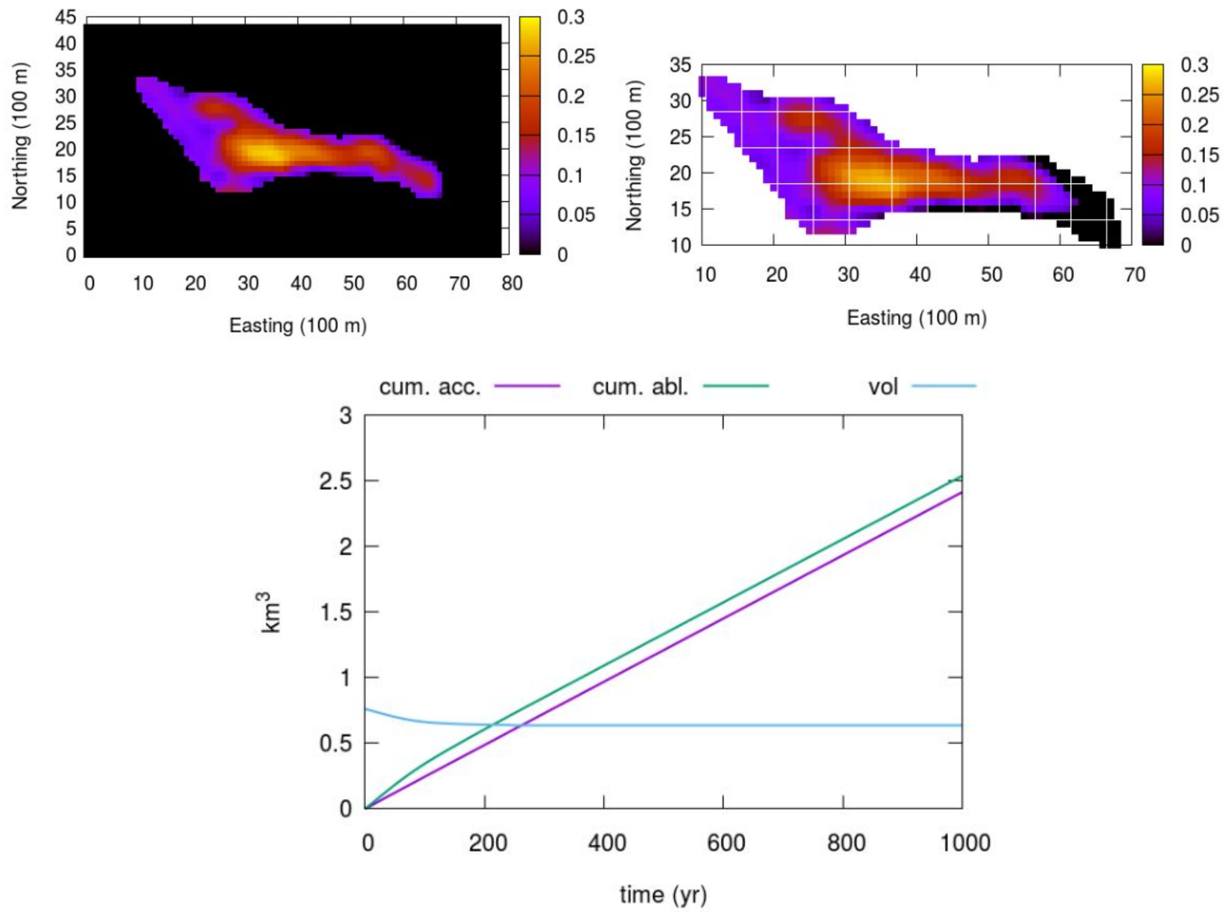



Figure R8: The top panel shows thickness maps of the initial (left) and final (right) states of a randomly chosen glacier (RGIid-15.07168). The color scale denotes ice thickness in km. The bottom panel shows the variation of the cumulative accumulation, the cumulative ablation and the total volume as a function of time after the step change in ELA.

We shall provide ice-thickness maps (eg., fig R8), ice-conservation plots and fits to transient area/volume evolution for all the 814 simulated glaciers in the supplementary.

Would also be good if you could consider some benchmark experiments (e.g. Jarosch et al., 2013)to make sure your model is mass conserving.

[11] As explained in reply [10], checking for ice conservation is relatively straightforward in our implementation, and at each time step ice was conserved up to a fractional error of the order of $10^{-9}$.

o Why do you randomly pick the values for the rate factor in Glen's flow law (not 'Glenn' + add a reference to the original studies, e.g. Glen, 1955)? The value of the rate factor will have a large influence on the local ice thickness and on thus the glacier volume. By picking this randomly: could be 'off' quite a lot from the 'reference/observed' volume of the glacier. Why do you not match this to the reference volume from every glacier that you have from Kraaijenbrink et al. (2017)?

[12] A wide range of values of rate factors has been used to model Himalayan glaciers. Typically, it is used as a tuning factor to obtain a good match with the observed velocity and thickness profiles. We have not done any tuning as our objective is limited to comparing the performance of the three models for the same set of glaciers with realistic geometries. Without such tuning our modeled volume is likely to have bias, but that does not interfere with our plan of comparing the three methods for the exact same set of model glaciers. Of course, we agree that for accurate prediction of glacier mass loss in the Himalaya such a tuning is a necessity.

Choosing the rate constant randomly from a wide range ensures that our results are not specific to a particular value of the same. The variability of rate constant, glacier geometry and mass balance profile lead to scatter in
area-volume scaling plot for the modelled glacier (main figure 1). Such scatter is expected for a set of real glaciers as well. The variability in rate factor, balance gradient and bedrock geometry allows us to test the performance of scaling and linear-response models for a more realistic situation.

Is this also not problematic when working with single values for c and g later in your analyses for all glaciers (e.g. for the best fits): you make some glaciers too thin and some too thick.

[13] As clearly explained by Bahr et al (2015), the scaling law does not apply to a single glacier. It only holds statistically for an ensemble of glaciers and over/underestimation of thickness for individual glaciers cannot be avoided. According to the authors, c may

vary with time but $\gamma$ does not for a given set of glaciers. We have followed their prescription.

d. Lack of in-depth analyses. Often you seem to be perplexed by some findings yourself and leave important questions unanswered, which is unsatisfying for the reader. This questions the thoroughness of your approach, e.g.:

[14] We have obtained several numerical results here that do not have a clear theoretical explanation in the literature. While we could explain some of those results here, we carefully pointed out each of the cases where we could not. In all the instances listed by the reviewer except the first one, we were attempting to distinguish between the numerical results that are supported by theoretical arguments, and ones that are purely numerical observations.

We shall revise the text to bring more clarity and to avoid any confusion.

ol. 186-187: '...we did not do a detailed glacier-by-glacier analysis of the reason behind the failure of the algorithm'... Well, you should do this! May be something intrinsically wrong with your setup (e.g. in terms of mass conservation, boundary conditions; see 2c). If this is the case, this is likely to have direct consequences for your results and for some of your conclusions…

[15] We apologise for omitting this important step. However, we had checked that the omission of 120 glaciers out of a total of 814 glaciers did not affect our conclusions or compromised the representativeness of the set (Fig R4).

As described in reply [9], we have now successfully simulated 810 glaciers among the total of 814.

We have provided details of boundary condition and mass conservation in reply [10]. In two glaciers violation of  mass conservation was observed due to noisy bedrock (fig R6).

ol. 247-248: 'We do not have a clear explanation of this effect as yet': ...

[16] What we mean to say here is that, while for scaling-based evolution the relation,

$$\frac{\Delta V_\infty}{V} = \gamma \frac{\Delta A_\infty}{A},$$

Is supported by both theoretical arguments and numerical evidence. However, the corresponding linear relationship for SIA glaciers is a purely empirical one.

We shall clarify this in the revised version.

ol. 256: ‘Again, we do not have a theoretical argument for such a power-law behavior and did not explore this further here’: …

[17] We meant to say here that we stayed away from a power-law fitting form which would have led to a better fit, and used linear fits as they are supported by theoretical arguments.

We shall modify the text to state this more clearly.

ol. 304-305: ‘..., it remains to be investigated if the results described here depend on the regional characteristics of glaciers to some extent’:...

[18] Mass-balance profile and glacier bedrock profile varies from one region to another. The statement is meant to acknowledge that our best-fit parameterisations need to be checked more thoroughly before applying it for a set of real glaciers on a global scale.

We shall reword the sentence.

[3] The **main conclusion** drawn your manuscript, and which also appears in the title, is that using V-A scaling methods (with ‘time-invariant scaling’) are likely to underestimate the future sea level contribution from glaciers.

a. I am not sure that the material you presented is convincing enough to support this statement and that the experimental setup is adequate (see previous point).

[19] We have already answered the criticism in our replies above and are not repeating the arguments here.

b. Another major concern that I have is: if this would be the case: why do we not see this when comparing outcomes of V-A scaling estimates compared to more sophisticated methods relying on retreat parameterizations (Huss & Hock, 2015)or flowline models (Maussion et al., 2019)? The first phase of the GlacierMIP project (Hock et al., 2019), in which future large-scale glacier simulations from the literature were compared, did not reveal a tendency for V-A scaling methods to underestimate the contribution to sea-level rise (SLR). Also in the second phase of the GlacierMIP experiments, in which several ice dynamic (vs. V-A) were included and in which coordinated experiments were performed, no clear tendency can be seen when considering V-A scaling vs. methods in which the glacier geometry (and in some cases also ice dynamics) are explicitly considered. From the material at hand, I would rather tend to believe the outcomes from GlacierMIP than the main conclusions put forward here when it comes to the implications of using V-A scaling for future sea level projections.

[20] We are thankful to the reviewer for raising this critical issue that we should have discussed in the manuscript. Despite a more realistic and therefore a more complex experimental setup employed by Hock et al. (2019), we believe their results contain signs of the biases in scaling models as pointed out in this manuscript.

Consider Table 3 of that paper which is reproduced below. The GloGem model (the only model employed by the authors that does not use scaling) predicted the largest change in both area and volume for 14 out of 16 of their experiments (shown with red arrows in their table 3 reproduced below). This may be an indication of a systematic underestimation of glacier change by the scaling-based models. The majority of the model in the ensemble being scaling based ones, the multimodel means may, thus, have a low bias. Based on our long term simulations, it is likely that if the experiments of Hock et al (2019) was to be extended over longer periods, the above differences between scaling-based and GloGem models may grow larger. Checking the outputs of the different models for a single climate forcing may also be useful here.

**Table 3.** Modeled global glacier mass and area losses by 2100 relative to 2015 (%) for four RCP emission scenarios. For each glacier model, data refer to multi-GCM means (± 1 Std dev.). *Model mean* refers to the arithmetic mean ± 1 Std dev. of all model runs for the same RCP regardless glacier model or GCM. Not all glacier models were run for all four RCPs. Results are also shown excluding the Antarctic periphery (A), and excluding the Antarctic and Greenland periphery (A + G) since some glacier models do not cover these regions

| | Volume loss (%) | | | Area loss (%) | | |
|---|---|---|---|---|---|---|
| Glacier model | Global | Global excl. A | Global excl. A + G | Global | Global excl. A | Global excl. A + G |
| **RCP2.6** | | | | | | |
| SLA2012 | 17 ± 3 | 18 ± 4 | 19 ± 4 | – | – | – |
| MAR2012 | – | 29 ± 7 | 31 ± 7 | – | 31 ± 7 | 33 ± 7 |
| GIE2013 | 14 ± 3 | 14 ± 3 | 17 ± 4 | 18 ± 4 | 19 ± 5 | 22 ± 5 |
| GloGEM | 24 ± 7 | 28 ± 9 | 29 ± 9 | 29 ± 7 | 32 ± 9 | 33 ± 9 |
| *Model mean* | *18 ± 7* | *23 ± 9* | *24 ± 9* | *22 ± 8* | *27 ± 9* | *29 ± 9* |
| **RCP4.5** | | | | | | |
| SLA2012 | 21 ± 5 | 22 ± 5 | 23 ± 6 | – | – | – |
| MAR2012 | – | 34 ± 9 | 36 ± 9 | – | 36 ± 8 | 37 ± 8 |
| RAD2014 | 28 ± 8 | 33 ± 10 | 33 ± 10 | 31 ± 10 | 34 ± 12 | 37 ± 12 |
| GloGEM | 33 ± 8 | 38 ± 11 | 39 ± 11 | 39 ± 9 | 43 ± 10 | 45 ± 10 |
| *Model mean* | *27 ± 8* | *31 ± 11* | *32 ± 11* | *35 ± 10* | *38 ± 11* | *40 ± 10* |
| **RCP6.0** | | | | | | |
| SLA2012 | 24 ± 6 | 26 ± 8 | 27 ± 8 | – | – | – |
| MAR2012 | – | 35 ± 8 | 37 ± 9 | – | 36 ± 9 | 37 ± 8 |
| *Model mean* | *24 ± 6* | *32 ± 9* | *33 ± 10* | – | *36 ± 9* | *37 ± 8* |
| **RCP8.5** | | | | | | |
| SLA2012 | 33 ± 6 | 35 ± 7 | 36 ± 8 | – | – | – |
| MAR2012 | – | 46 ± 10 | 48 ± 10 | – | 47 ± 10 | 48 ± 10 |
| GIE2013 | 27 ± 5 | 27 ± 5 | 31 ± 6 | 30 ± 9 | 33 ± 10 | 38 ± 11 |
| HYOGA2 | – | – | 17 ± 4 | – | – | 32 ± 6 |
| RAD2014 | 40 ± 8 | 46 ± 11 | 46 ± 10 | 47 ± 10 | 53 ± 13 | 55 ± 12 |
| GloGEM | 48 ± 9 | 55 ± 12 | 55 ± 12 | 54 ± 9 | 59 ± 10 | 60 ± 10 |
| *Model mean* | *36 ± 11* | *41 ± 13* | *40 ± 14* | *43 ± 14* | *47 ± 14* | *48 ± 14* |

Table 3 of Hock et al. (2019) with red arrows added to highlight the experiments where Glogem predicted the highest loss.

We do acknowledge that there are important differences between our set-up and that of Hock et al (2019) which prevents a direct comparison. We apply all the three models on the same set of steady glaciers, use the same mass-balance profiles, and consider the same idealise ELA perturbation. In contrast, the different models in Hock et al. (2019) do not use the same prescription for computing mass balance forcing, the input climate data used for mass-balance computations are also not identical. In addition, their comparison is over a relatively short period of 85 years, whereas we look at a longer period of 500 years. While our model experiments involve an idealised step change, the authors considered slower and more realistic forcing. As a result of all these differences, our setup is more sensitive to the differences in performance of the glacier-dynamics models alone. Of course, the experimental design of Hock et al. (2019) is tailor-made for the problem of predicting sea-level rise which is the main thurst of that paper.

We shall mention in the updated manuscript that the results of Hock et al. (2019) possibly support our claim of a systematic bias in scaling-based models.

c. You draw your main conclusion (that the loss from V-A scaling with time-invariant scaling is underestimated vs. SIA) from two steady states: an initial one and a final one. You present your results like transient results (e.g. in plots, when describing response times,in section 4.1. describing that cis time-dependent and decreases with time,in section 4.4.,....etc.), but in the end, **it boils down to the fact that the volume of the final steady state with time-invariant V-A scaling is 'too large'**(compared to the SIA). Due to this, the transient volume loss when evolving to this steady state is underestimated (always with respect to SIA results). The main question that you thus need to address is: **why is the V-A scaled final steady state too big?** I am not an expert in V-A scaling, but I would find it surprising that this issue has not been addressed in other V-A scaling studies and that no solutions to this problem have been formulated. In the end, from my understanding, what happens is that many glaciers that reduce in size lose their lowest part, which are often the most gently sloping parts of the glacier and where the highest ice thickness is thus found (in most ice thickness reconstructions this clearly appears, where in the end, a large part of the reconstruction results from the negative correlation between the surface slope and the local ice thickness; see Farinotti et al., 2017). It is thus to be expected that the V-A scaling that you use to create the initial steady state does not hold for the final one. This is something that would need to be explored in more detail, and for which studies in which the volume scaling also uses information from other glacier characteristics (e.g. the glacier slope) could be useful (Grinsted, 2013; Zekollari & Huybrechts, 2015; see e.g. Fig. 9a in the latter, which summarizes the main point made here).

[21] We agree with the reviewer's assertion that an important limitation of a scaling-based model is that the scaling-based model underestimates the climate sensitivity of glaciers. However, we do not agree with his/her point of view, that is the only issue. As we have already demonstrated in the manuscript, there are two more critical issues: A. the scaling-based model underestimates glacier response time (please refer to replies [1] and [3]), B. the volume and length response times are predicted to be identical under scaling.

While there are several existing investigations of the relative performance of the scaling-based models (including those cited in our discussion paper and in the reviewer's comment), the above three limitations may not have been brought out clearly.

Also, a common issue with the existing comparisons of scaling-based model with ice-flow models is that often both $\boldsymbol{c}$ and $\boldsymbol{\gamma}$ were taken to be time-dependent fitting parameters empirically. This was advised against by Bahr et al. (2015) on solid theoretical grounds.

We do agree with the reviewer's comment above that a fixed time-independent scaling form does not work. Our SIA simulation clearly shows that the scale-factor ***c*** is time-dependent (please see reply [38]). While the gently-sloping lower ablation zone, as suggested by the reviewer, may be one of the reasons behind this, it is not the only reason. For a set of 1-d glaciers on linear bedrock with the same bedrock slope, a similar bias in the scaling-based model is obtained (fig R1, reply [1]). So any modified scaling formulation involving glacier slope etc. may not be able to cure the bias entirely.

If a parameterisation of the time dependence of *c* in terms of slope and other factors is found, then that may give a clear answer to the questions why scaling predicts relatively smaller sensitivity (as pointed out by the reviewer) and response time. However, even that is not going to correct the drawback of having the area and volume response times (reply [1], [3]). Therefore, in this paper, we choose to focus on obtaining an alternative linear-response model that reduces the above biases, rather than trying to investigate how scaling models can be improved.

[4] Unclarities in the manuscript. I found the text difficult to follow and quite often had to re-read sentences several times before being able to grasp their meaning. A few examples include:

[22] We shall work on improving the readability, and appropriately rewrite the following sentences.

a. l. 8-9: '..and validate them with results from scaling-based simulation of the ensemble of glacier'

b. l.84-85: '...are then empirically extended in order to obtain accurate parameterisations the linear-response properties of the SIA-simulated glaciers'

c.l.86-87: 'The linear-response model the long-term total shrinkage of glaciers as predicted by the scaling-based method (Radićet al., 2007), and the linear-response model are compared with the corresponding response'd.....etc.
See also comments on specific sections below.This makes it tedious to go through the manuscript. Furthermore, there a substantial number of grammatical errors, some of which (but not all) have already been pointed out by the first reviewer.

[23] We shall try to minimise such errors in subsequent revision.

Also, many figures cannot be interpreted/read independently, without having to refer to the caption. It would be good if all essential information (e.g. meaning of colors used, R^2 values, equations,...etc.) could be directly included in the figure. Some other comments for specific sections(non-exhaustive list and not focusing on grammatical errors)

[24] We shall update the figures mentioning the best-fit forms and $R^2$ within each of the figures.

o1. Introduction:
'methods solving the dynamical ice-flow equations' à'numerical cost of such a computation on a global scale is prohibitive': well is not really the case anymore. In general: would be good to acknowledge regional-to global studies in which ice flow is explicitly accounted for (Clarke et al., 2015; Maussion et al., 2019; Zekollari et al., 2019).

[25] We shall reword the statement, and include relevant references including the ones mentioned here.

o1.2. Motivation for the present study: difficult to follow the first paragraph: be more specific when you refer to c and gamma not continuously mix with other terminology 'time invariant scaling-based parameterisation', '...given the known violation of the time-invariant scaling assumption'.

[26] The scaling equation,

$$V = cA^{\gamma}$$

does not require *c* to be a time-independent constant (Bahr et al., 2015). However, that assumption is necessary to use this equation to predict glacier evolution. So they are not interchangeable.

o2. Quite abstract and thus very difficult to go through for someone who is not an expert in V-A scaling. Could make it less technical by for instance adding some additional information that links the various parts.

[27] This section sets the notation, and provides the mathematical derivation of various results. So it may appear a bit technical. However, we appreciate the need for making it

more accessible, and shall add a paragraph in the end summarising the results in plain language.

o3.1.: 2-dimensional SIA model:

•l.152-154: where did you get the ice thickness from? From Kraaijenbrink et al. (2017)directly? As the ice thickness is quite crucial in your story (it determines the volume...), why did you not consider the consensus estimate of Farinotti et al. (2019), which is freely available?

[28] We had used data from Kraaijenbrink et al (2017) as it has debris-cover and debris thickness information as well - which we initially planned to incorporate. Also, when we started the project Faronotti et al. (2019) was not published. Since we do not tune our models to obtain accurate descriptions of each of the glaciers, we stick to the bedrock originally used. Any reasonable choice of bedrock serves our purpose of simulating an ensemble of glaciers with realistic geometry. We agree that accurate bedrock and tuning the parameters of the SIA model to fit the thickness and velocity maps of each of the glaciers are necessary for any computation of future glacier change or sea-level rise.

•SIA: refer to the original work by Hutter (1983)also.

[29] We shall refer to this article.

•You neglect basal sliding (l. 161). Justification? Could refer to other studies where this is done, like e.g. Gudmundsson(1999) and Clarke et al. (2015).

[30] We had neglected sliding as the theoretical results of Bahr et al. (2015) that we built upon are derived in that limit. For a set of real Himalayan glaciers sliding would surely be important as mentioned in section 4.2 of the manuscript.

•l.168-178: this is related to the SMB, which you apply in all cases(i.e. also for the linear-response model and the V-A scaling, right?). Not sure this section is correctly placed here in the ‘2-dimensional SIA model’ section.

[31] we shall add reference to eq. 13 in the later sub-sections describing the 0-d models.

•l.183: through several exclusion you keep 68% of the initial glaciers... How much does this represent in terms of glacier volume and glacier area when compared to the total glacier sample?

[32] With the present runs we keep 703 out of the total 814. This corresponds to a number fraction of 86% and an area fraction of 89% . Also see fig R4 above for the area and slope distribution of all the glaciers and the selected ones.

•l.188-196: you explain some simplifications related to debris cover, avalanche and sliding have been made and that this may influence your results. Well, you have made much larger simplifications than this: e.g. linear SMB profiles with strongly imposed max. SMB, steady state assumptions for glaciers,... à not even worth mentioning these more detailed simplifications in my opinion. With all these simplifications, would have been better to opt for idealized setup likely (see main comment 2a).

[33] We prefer to list out all the limitations of our model as depending on the glacier being considered, one or several of  these issues have proven to be quite critical in the Himalaya.

We have already discussed (replies [1] and [3]) why idealised glaciers do not serve the purpose for us.

•l. 194: 'These simplifications do not weaken our study': not sure you can judge on this yourself...

[34] the sentence will be modified.

o Section 3.2.:

•l.204: 'was fixed at... because...': don't understand the causality (i.e. link between cause and consequence).

[35] The sentence will be modified. We mean to say: the linear mass-balance profile implies m=1, and then, using the formula given by Bahr et al (2015), gamma = 1.286.

•Figure 1: SIA-derived volumes are scaled by a factor 10: why? Does not really make sense and unclear when just looking at the figure without reading the caption... Axes should be correct in the figure and not only for a part of the data you show..Also illustrates the unclarity in the figures mentioned in main comment 4 (problem that figures cannot be interpreted without referring to their caption).

[36] We have modified the figure by removing the scaling by 10 (Fig R5). We shall update the other figures following the suggestion by the reviewer.

oSection 3.3.:

•l.208-210: complicated way to say that you consider e-folding time scales. Would reformulate this and add references for this to e.g. Leysinger Vieli & Gudmundsson(2004).

[37] As our aim is to find the best fit linear-response properties. Thus instead of computing e-folding time, we directly fit the following linear-response form to obtain both climate sensitivity and response time for each glacier.

$$\Delta V(t) = \Delta V_{\infty}(1 - e^{-t/\tau_v})$$

Only for a purely exponential decay the response time would be identical to the e-folding time.

oSection 4.1.:

•l.225: $V=cA^{1.286}$: not sure I understand. Does this statement apply for the initial and/or final steady state volumes? And can all the volumes be described with this single relationship? Is the fact that quite different rate factors are used not a problem for this (see main comment, point 2c)?

[38] As shown in main figure 1 (A) both the initial and final states obey the same scaling form with different values of *c.* This is consistent with the rigorous derivation of the scaling law (Bahr et al., 2015). As long as a large enough ensemble of glaciers with a wide range of glacier area is considered, a fixed area-volume scaling relation is a good statistical description, though it may have considerable bias if any single glacier in the set is considered (Bahr et al., 2015).

The variability of rate factor, geometry (as long as width exponent q is the same), mass balance profile (as long as exponent m is the same) etc only add noise the scaling, and lead to scatter in the scaling plot. Such effects do not ruin the volume-area scaling. This is consistent with main figure 1A and fig. R1. In fig. R1, there is little scatter as geometry, rate factor, mass-balance gradient are exactly the same for each of the glaciers. However, in main fig 1A a considerable scatter is present due to variability of these factors. Of course, in a set of real glaciers significant variability of the above parameters, and consequently, a significant scatter in the corresponding scaling plots are present (Bahr et al., 2015).

•l.227 + l.230 + l.232:here you mention that c is time dependent. Not sure you can say that it is time dependent: simply results from the fact that final steady state volume for V-A scaling is ‘overestimated’ (vs. SIA). As a result the evolution to this steady state is different. See main comment 3c for this.

[39] 1. We have shown in main figure 1A that initial and final best-fit c are different.
2. We have shown in main figure 1B that fo SIA simulated glaciers time-evolution trajectories are not linear in V-A plane with log-log scale.

The above two observations prove that c depends on time (as long as the same form $V= cA^{1.286}$ is used to describe the ensemble as argued by Bahr et al. (2015) on theoretical ground).

•l.235-237: relates to main comment 3c again. If you do not modify the V-A scaling, then problems will arise when considering the same glacier that is much smaller in a warmer climate (when rising the ELA in your case): you typically lose the lower parts where most volume is and volume will thus be ‘overestimated’. Is this not accounted for in some way in future glacier evolutions based on V-A scaling? As a part of this discussion, studies in which V-A scaling is extended with other glacier characteristics (such as the surface slope; Grinsted, 2013; Zekollari & Huybrechts, 2015)would be good to include. Such relationships which could prove to remain valid over time, even without changing scaling and exponents.

[40] We agree with the reviewer that with the incorporation of appropriate additional variables like slope etc., it is possible to improve the scaling formulation. However, as shown in reply [1] above, this can not cure the bias entirely. Also, it cannot resolve the issue of equality of area and volume response time (reply [1] and [3]) that is inherent in scaling. Therefore, we took an alternative route of developing an accurate linear-response description to minimise the bias. We do not rule out that there may be other possible routes to reduce the bias.

We shall make the above point clear in our revised discussions.

oSection 4.2.:

•l. 243: ‘This is exactly what is seen in Fig. 2b, which shows...’: I cannot directly see this…

[41] We wanted to point out the fact that fractional volume change is 1.87 times the fractional change in area, while scaling theory predicts a smaller prefactor of factor of 1.286.

We shall rewrite the sentence to make it clear.

•l.244: ‘change in c to the tune of ~13%’: what does this mean?

[42] We meant to say that: SIA simulation showed a fractional change in volume that is ~50% larger than that predicted by the scaling method for a given fractional change in area. A relatively small 13% change in c is not the only factor behind this, the associated mass-balance feedback also plays a significant role.

•l.255: ‘The above figure’: will depend where your figure comes in final manuscript…

[43] We shall refer to the figure number here.

oSection 4.3:•
I was wondering what the point is that you want to make with this section? It is known from literature that volume responds faster than area (e.g. Oerlemans, 2001; Leysinger Vieli & Gudmundsson, 2004).

[44] We apologise for not referring to prior work by Schmeits and Oerlemans (1997), Oerlemans (2001) and, Vieli and Gudmundsson (2004) that had discussed that volume response is faster than the length response. We shall correct that error.

The main result in this section is that a scaling-based evolution leads to the same response time for area and volume (as shown with both theoretical arguments in section 2.3 and numerical results in main fig 4), while in SIA simulations the area response time is larger. This is again a major limitation of scaling-based methods.

We shall rewrite the paragraph to make the statement clear.

• l.260-264: relationship between volume and area response times. How does this compare to the relationship others have found in the literature?

[45] We thank the reviewer for the comment.
Oerlemans ( 2001) suggested a proportionality constant of 0.74 between volume and area response time. Vieli and Gudmundsson (2004) obtained corresponding values of 0.60, 0.70 and 0.67.
The above values are comparable to our best-fit value of 0.69.

We shall discuss this point in the revised manuscript.

oSection 4.4.:

•l.271-272: ‘with most of the changes taking place during the first couple of centuries’: this is not a result/finding.. This directly results from the e-folding time-scale when forcing a steady state glacier with an instantaneous forcing in SMB.

[46] The sentence “Starting with an initial volume (area) of 603 $km^3$ (5144 $km^2$) the 551 glaciers simulated by SIA loses 123 $km^3$ (521 $km^2$) of volume (area) in 500 years after the step-change in ELA by 50 m, with most of the changes taking place during the first couple of centuries (fig. 6).” is a fair description of the data presented in fig. 6.

•l.273: ‘underestimates the long-term change’: not about reaction/response. This is direct consequence of fact that final steady state volume is too large (see main comment 3c)

[47] Irrespective of the interpretation, the statements in this paragraph is a correct description of data presented in fig 6.

•l.279-280: ‘...suggests that there might be significant negative biases of mountain glacier contribution to sea-level rise as computed by scaling-based methods’(+ section 4.5, l.300-302): well, do not see this in GlacierMIP phase 1+2... Is a very strong statement to make and should be sure that it is well-founded.

[48] Please refer to comment [20] where give evidence that GlacierMIP contains enough hints about such a possible bias. We shall include that discussion in the revised manuscript.

O Section 4.5:6

•l.296-297: ‘More detailed studies that relaxes some of the above mentioned assumptions are needed...’: not sure what you mean by this. Would also make sense that you dig into this: e.g. by focusing on real transient response vs. comparing two steady states (what you do now and then translate into an analysis of the transient response resulting from this: see main comment 3c).

[49] We are considering transient response to step change in climate and not steady states. We are following a well established method for characterising glacier response by computing response properties of steady states (e.g. Oerlamans, 2001). If steady-state linear-response properties (eg climate sensitivity and response time) are known, the response to any arbitrary mass balance forcing can be computed as (we quote from our reply to comments by reviewer 1),

$$\Delta V(t) = \Delta V(0)e^{-t/\tau} + \frac{\Delta V_\infty}{\delta E}\int_0^t \Delta E(t')e^{-(t-t')/\tau}dt'.$$

Here, $\frac{\Delta V_\infty}{\delta E}$ is the climate sensitivity of glacier volume. $\Delta V(0)$ is the initial departure from a steady state that can be obtained from the observed rate of volume change as $\Delta V(0) = -\tau\frac{dV(0)}{dt}$. A similar expression can be written down for area evolution as well. We propose to include these details in the discussion section of the revised manuscript.

We note that a similar linear-response framework had been applied to reconstruct climate forcing from glacier length change records (Oerlmans, 2005). While we focus on idealised climate forcing to obtain linear-response properties, the response properties derived from that analysis allow predicting the response to any general forcing as long as the fractional glacier changes are small.

•l.299: ‘intruding more scatter in the fits’: what does this mean?

[50] This is a typographical error. “Intruding” will be replaced by “introducing”.

oSummary and Conclusions:

•l.309-310: scale factor reduces over time. Well, not sure the time dimension is adequate here. Boils down to having a final steady state that would require a smaller value for c: see main comment 3c.

[51] There is probably no inconsistency between our statement and that of the reviewer. Unless c is time-dependent, c cannot be different for the initial and final states. Also, the discrepancies between SIA and scaling model during the first hundred years ( see main fig 1, fig R1) when the glaciers are in a transient state, is consistent with a time-dependent c.

•l.324: computational efficiency. OK, still important, but is not really a limitation anymore, due to which V-A scaling becomes less important (and also driven by the release of new datasets with regional-to global spatial coverage at individual glacier level: see main comment 1).

[52] Here we are only referring to the “computational advantage” and not claiming that linear-response is the only way of computing long-term glacier change.

•Code availability: for which models is the code available? Seems to suggest that the SIA code is not available. Not sure if this fully agrees with the policies of The Cryosphere: see www.the-cryosphere.net/about/data_policy.html

[65] We shall make all codes available as per the policies of the journal.

**References:**

Bahr, D. B., Pfeffer, W. T., and Kaser, G. (2015). A review of volume-area scaling of glaciers. Reviews of Geophysics, 53(1), 95-140

Clarke, G. K. C., Jarosch, A. H., Anslow, F.S., Radić, V., & Menounos, B. (2015). Projected deglaciation of western Canada in the twenty-first century. Nature Geoscience, 8, 372–377. https://doi.org/10.1038/ngeo2407

Farinotti, D., Huss, M., Fürst, J. J., Landmann, J., Machguth, H., Maussion, F., & Pandit, A. (2019). A consensus estimate for the ice thickness distribution of all glaciers on Earth. Nature Geoscience. https://doi.org/10.1038/s41561-019-0300-3

Hock, R., Bliss, A.,Marzeion, B., Giesen, R. H., Hirabayashi, Y., Huss, M., et al. (2019). GlacierMIP −A model intercomparison of global-scale glacier mass-balance models and projections. Journal of Glaciology. https://doi.org/10.1017/jog.2019.22

Huss, M., & Hock, R. (2015). A new model for global glacier change and sea-level rise. Frontiers in Earth Science, 3, 1–22. https://doi.org/10.3389/feart.2015.00054

Kraaijenbrink, P. D. A., Bierkens, M. F. P., Lutz, A. F., & Immerzeel, W. W. (2017). Impact of a global temperature rise of 1.5 degrees Celsius on Asia's glaciers. Nature, 549, 257–260. https://doi.org/10.1038/nature23878

Kulp, S. A., & Strauss, B. H. (2019). New elevation data triple estimates of global vulnerability to sea-level rise and coastal flooding. *Nature communications*, *10*(1), 1-12.

Leysinger Vieli, G. J.-M. C., & Gudmundsson, G. H. (2004). On estimating length fluctuations of glaciers caused by changes in climatic forcing. Journal of Geophysical Research, 109(F01007), F01007. https://doi.org/10.1029/2003JF000027

Marzeion, B., Jarosch, A. H., & Hofer, M. (2012). Past and future sea-level change from the surface mass balance of glaciers. The Cryosphere, 6(6), 1295–1322. https://doi.org/10.5194/tc-6-1295-2012

Oerlemans, J. (2001). Glaciers and climate change. Dordrecht: A. A. Balkema.

Oerlemans, J. (2005). Extracting a climate signal from 169 glacier records. *Science*, *308*(5722), 675-677.

Rounce, D., Hock, R., & Shean, D. (2020a). Glacier mass change in High Mountain Asia through 2100 using the open-source Python Glacier Evolution Model (PyGEM). Frontiers in Earth Science, 7, 331. https://doi.org/10.3389/feart.2019.00331

Rounce, D. R., Khurana, T., Short, M., Hock, R., Shean, D., & Brinkerhoff, D. J. (2020b). Quantifying parameter uncertainty in a large-scale glacier evolution model with a Bayesian model: Application to High Mountain Asia. Journal of Glaciology.

Schmeits, M. J., & Oerlemans, J. (1997). Simulation of the historical variations in length of Unterer Grindelwaldgletscher, Switzerland. *Journal of Glaciology*, *43*(143), 152-164.

Zekollari, H., Huss, M., & Farinotti, D. (2019). Modelling the future evolution of glaciers in the European Alps under the EURO-CORDEX RCM ensemble. The Cryosphere, 13, 1125–1146. https://doi.org/10.5194/tc-13-1125-2019