

General comments

Malles & Marzeion quantify the sea-level contribution from the world's glaciers during the 20th century, and its associated uncertainty. They use a published global mass-balance model with some new developments regarding the representation of glacier response time in the mass-balance calibration. The model is applied to all glacierized regions in the world except Antarctica. A number of key model parameters are optimized against in-situ observations of glacier mass balance, and the performance of nine different meteorological forcing datasets ('ensemble members') is assessed. For each region, they also provide uncertainty estimates associated with the model and the ensemble of forcing datasets.

The authors account for total uncertainty as being the combination of the uncertainty in the forcing and the model itself. They find that uncertainty in the atmospheric boundary conditions (the forcing) used may exceed uncertainties related to model parameters. Mass-loss estimates from the different ensemble members diverge more for time periods with lesser known atmospheric conditions (e.g. early 20th century), which are found to often coincide with areas where mass-balance observations are scarce (areas outside W. Canada + USA, Scandinavia, Central Europe, Central Asia).

Generally there are few studies that reconstruct global mass loss from glaciers all the way back to the early 20th century, let alone provide robust estimates of associated model and forcing uncertainty. This is a well-written manuscript that clearly fits within the scope of TC, and is in my view a timely contribution that furthers our understanding of global glacier mass loss, and provides some excellent strategies for how to deal with uncertainties in mass loss estimates. With this said, I have a number of suggestions for improvement. My major comments are mostly related to a somewhat confusing structure of the paper, an incomplete description of the model, and a limited discussion of the results. The authors should be able to address these comments without running additional model simulations. If comments are properly addressed, I think the paper is highly likely to deserve publication in TC.

Major comments

Structure and model description

The way the Data and Methods are presented is a bit confusing. First you explain the model equations, before we have an idea of what data you use. I think this makes the model description unnecessarily hard to follow (see below). One option would be to swap 2.1 and 2.2, with some minor rewriting of cross-references. Alternatively, if you keep the current structure, you should make sure that enough information about the data is included for us to understand the model setup.

The description of the mass-balance model would also benefit from more detail. While the most important model equations are included, and you state that more details can be found in Marzeion et al. (2012), I think a bit more rationale behind the model equations and their

underlying assumptions would strengthen this section and help the reader. Right now, the model details are stated without much justification or reasoning behind the model choices. This includes the length-area-volume scalings and relaxation, which currently are only mentioned in one sentence without detail or rationale.

Section 2.1.2 is important, because here you explain how you account for glacier response time, which in my view is an exciting novelty of this study. Unfortunately I found this section a little difficult to follow. Partly this may be because we're not really sure where you're going without knowing the data you use. I think it needs to be articulated more clearly what this new model development regarding response time brings to the model, and why it is needed. It would also be useful with a concrete example illustrating your approach and the underlying logic, perhaps throwing in some typical numbers that help us to better understand what you are doing.

Finally, I think including a new flowchart-type figure of your experimental design would be something that would help the reader to get an overview of your approach. Such a figure could include for example the steps involved in your initialization, geometric scaling, calculation and interpolation of model parameters, calibration, validation and optimization steps, the datasets used and parameters optimized, and the output results.

Comparison against geodetic mass balance data

I wonder whether there is a particular reason for only comparing modelled mass balance against in-situ data, when geodetic data is available in many places and indeed also used in some of the previous studies you cite. Even a tentative comparison with some geodetic surveys would strengthen the study. Alternatively, the authors should clearly explain why this was not done, and why it would not be needed for the purposes of the current study.

Discussion

The discussion is nicely written, with interesting comparisons with previous studies. Specifically, I commend the authors for including some underlying explanations for the global mass loss evolution simulated (Fig 7). Although the focus of the present study is mainly on quantifying mass loss and estimate uncertainty, it would be interesting with an expanded discussion of the underlying causes of mass loss (relative role of temperature and precipitation, timing, and potential lag of mass loss). This could be based on what you have shown in Fig 7, and should also be put into context with previous literature. To be clear, I think a detailed discussion of the underlying explanations for mass loss over time for each region would be beyond the scope of this paper. Still, findings like “the increase in precipitation between ca. 1930 to 1950 was accompanied by a similar increase in solid precipitation, indicating that the warm anomaly at the same period was too weak reduce accumulation” (note missing ‘to’) are enticing and warrants further discussion.

Right now you present your results divided into the world's glacierized regions, which makes sense for your purpose. To understand the model performance in more detail and where

potential improvements can be found, I wonder whether it would be possible to also mention if the model performs better/worse for specific glacier types, e.g. alpine valley glaciers, ice caps, cirques etc., or at certain elevation ranges. If it is not possible to disentangle such detail, at least some discussion on what you expect based on the model equations, assumptions and limitations would be useful.

In addition, I would like to see a more thorough discussion on the steps ahead, preferably in light of the limitations of the current study, which themselves are barely discussed at all. The last paragraph (L414-416) hints at this, but could be expanded on. What recommendations do you have for future studies like the current one? How can models be improved? Do we have to live with these higher uncertainties in the early 20th century, or are there ways to improve estimates further? What recommendations do you have for scientists working with meteorological data sets (which you assessed in detail)? What about Greenland periphery, which contributes heavily to the model uncertainty AND contributes highly to global mass loss; can we improve here? And Antarctic glaciers, which you did not have meteorological data for, what should we do with them?

Minor and technical comments

Abstract

L3. would rather use “estimated” that “calculated” here, since “calculate” to me sounds too precise for what you’re talking about.

L11. “The goal is...” I think this sentence needs to come earlier, before you describe the methods.

L13. “ensemble members” not really clear at this stage what you refer to here; are you talking about the different meteorological datasets, or the model parameters? I think you mean the met. data, but you should clarify.

L14-15. “... the availability of mass balance ...” I think you can write this sentence in a more direct way, e.g. “where mass balance observations are abundant/in well-observed mass balance regions, forcing data is also better constrained/more known” or similar.

L16. “out-of-sample uncertainty” – what do you refer to here?

L17. “cross validation” vs “cross-validation” check consistency throughout the manuscript

Introduction

L21. “is constituting” -> constitutes (change to more direct writing)

L22. “constitute” again – change to are/comprise to avoid word repetition

L23. “their vanishing” – here it is unclear what the word “their” refers to, it could be referring to “regions of the world”, implying that some regions of the world may vanish – which is obviously not what you mean. Rewrite.

L26. I think you should define what you mean by mass balance the first time you mention it. Also you need to make clear whether you talk about glacier mass balance, or surface mass

balance throughout the manuscript.

L28. lacking -> lack (more direct)

L28. “comprehensive in-situ measurement densities” – I think you can write this in a less convoluted way

L33. “temperature index melt models” – missing hyphen

L33. “melting precipitation” – do you mean melting of snow?

L34-36. This is useful info, but I think you could explain here why such scalings are used, and emphasize why ice dynamics are not included in these models. I would also add in Discussion how these types of scalings may have influenced your results.

L44. I guess adding additional parameters not only makes model optimization more complex, but also increase model uncertainty?

L48-49. Not clear at this stage how the cited study using OGGM links to your work. Do you follow their approach? On L37-38 you state that you don't include ice dynamic processes, so the reader is left somewhat confused here.

L58-59. “heterogeneously distributed” – I assume you mean within and across geographical regions here? Also I think you can split this sentence into two separate sentences.

L67. I think “metric” would be more appropriate than “parameter” here.

L55-79. The words evaluation, calibration, validation, and optimization are used somewhat interchangeably in this section. Please check that you're consistent, and clarify what you mean by these terms.

Data and Methods

L94. “as before” -> as in Marzeion et al. (2012)

L97. Please clarify what you mean by “leave-one-glacier-out” in this context.

L99-100. Beta* is estimated by spatially interpolating from the ten closest glaciers with 3+ years of in-situ observations. This is fine, but because in-situ observations are very unevenly distributed (which is one of your main points in the paper), the constraints on beta* will vary greatly across different regions. In regions with plenty of observed glaciers, the interpolation will likely work better than in sparsely observed regions. I think you should explain this explicitly to the reader, and how this influences your results.

L104-114. If I understand your model correctly, you assume that the fraction of solid precipitation (snow) varies linearly with glacier elevation. If this is the case I think this could be explained in words, in addition to stating the equations. Also you should state why this assumption is sensible for your purposes, and whether this assumption may break down in some glaciological or climatic settings.

L121-125. The scaling procedure and the underlying assumptions should be explained in more detail, see Major comments above. The iterative approach finding the initial area (and volume?) of each glacier needs a bit more elaboration too. For how many glaciers did the iterative procedure work, and in what regions mostly (if any patterns can be seen)?

L138. How many glaciers have 3+ years of in-situ measurements, and are thus included in this calculation?

L147. exactly what is “not generally the case” here?

L151. Change to (a) instead of “upper panel” – and do so throughout the manuscript

L157. mid -> middle

L158. “It” – are you referring to Fig 1b ?

L164. mid -> middle

L169-170. “no longer” I assume you mean as opposed to what was done by Marzeion et al. (2012) ?

L183. they were -> full temporal coverage were

L187. of -> in

L188. the modeled time interval – I think you can state here what this exact time interval is.

L193. Please state why you ignore the uncertainties in the in-situ observations (for simplicity, because of computational constraints, because they don't matter for your purposes?)

L192-194. What about geodetic mass balance estimates? See Major comments.

L196. a -> an

L197-203. How were these ranges selected? Based on previous literature, based on first-principles?

L225. I would try to avoid starting a sentence with “E.g.” – replace with “For example” or similar. Otherwise a very useful and nicely written paragraph!

L230. a respective -> each (check use of “respective” throughout manuscript)

Cross validation and uncertainty assessment

L264. Avoid starting sentence with “E.g.”

L266. be -> by the

L266. than than (double word)

L281. different -> differs

L284-286. I think you can split this sentence in two for readability.

L306. Rewrite to “In Fig. 3a, ... “ – and check figure cross-refs (use of B vs (b) etc.) and be consistent throughout the manuscript (for example, at L309 and 313 in this section)

L316-319. Great point, well said!

L333. “and for and” – redundant word

L334. A little weird to say that something “grows backward in time” – would rather say that it starts high, then decreases. Similarly on L348, when something “shrinks going back in time”.

I would rather say that the further back in time we go, the fewer meteorological datasets (ensemble members) are available.

Global Glacier Mass Loss

Overall a very interesting section!

L350. Not sure why all words in the section heading here are capitalized, why for other headings, they aren't. Check to be consistent.

L352. Use a/b instead of upper/lower panel when you refer to figures.

L354. “are diverging” -> diverge (more direct)

L355-358. Check the sign of your mass loss estimates (and in Fig 5) – one can get the impression that mass loss decrease over time with the current figure 5 version, which obviously not is what you want to show.

L358. please state explicitly what “the most recent period” is

L359. “earlier years” – a bit vague, please be more concrete

L366. state when “the end of the modeled period” is (2018?)

L369-374. Interesting findings regarding the explanation for the found changes in mass loss. These changes regarding the temperature and precipitation and the resulting changes in mass loss over time are interesting enough to be included in the Abstract, or at least in the Conclusions.

L376. state which periods you refer to by “the more recent periods”

Discussion

L383. Gardner et al citation should be without parentheses.

L385-386. I think you can remove the parentheses around “because ... glaciers”

L402. Again please remind the reader when “the most recent time” is

L405. Please state when “earlier periods” is in this context

L410. Remind the reader that you’re looking at Greenland periphery rather than Greenland

L414. Use of e.g. here breaks the flow of reading, would rather use wording like “for example”

L415. Would add ‘(e.g. Greenland periphery)’ here as an example of such a largely glaciated but less observed region

In general a nicely written section, but missing a few key parts, see Major comments above.

Conclusions

L421. Not clear from the context here what “validation data” refers to, please be explicit on what type of data this is

L428. Be explicit on what has been decelerating and accelerating (the mass loss, not the glacier flow)

Figures

In general nicely crafted, clear figures! A great addition would be a flowchart-type figure explaining the model setup and experimental design, see Major comments.

Figure 1

I would write out what beta is in both the ylabel and the caption.

Add (a), (b) labels, and for subsequent figures.

Also I would write out what t_{tilde} and t^* is in the caption, so that the reader can understand the figure without looking up the variables in the text.

Add cross-reference to the appropriate section after “cross-validation procedure”.

Figure 2

Would write out the entire ylabel “number of observations” (can be split over two lines to fit the yaxis.

Caption: state when “the modeled period” is.

Figure 3

I think you can remove “RGI” in the title for (a), as it is not crucial to understand the figure and is just another abbreviation to check for the non-expert reader.

Figure 5

Check sign of mass loss, see Minor comments above.

I think you can also write out “Sea-level contribution” in the ylabel.

Figure 6

I would include the legend from Fig 5b also in Fig 6b.

Figure 7

You have space to write out temperature anomaly and precipitation anomaly in (a). In (b), you don’t really need a legend, as there is only one time-series to show.

Tables

Table 1

Resolution – missing unit (degrees?)