

Response to reviews for tc-2019-222: “21st century ocean forcing of the Greenland Ice Sheet for modeling of sea level contribution”

We would like to thank the editor and all three reviewers for their constructive, considered comments and for taking the time to carefully review our manuscript. We are pleased that the manuscript was received positively overall, with the three reviewers describing it respectively as ‘a valuable framework for future modeling work’, having ‘a clear, direct message and nice figures’, and ‘an important step forward for modeling ice sheet response to various warming scenarios... the ideas are well-organized and logically presented.’

We also recognize the reviewers’ major concerns, notably in four areas: (i) on the sensitivity of boundary conditions to the bias corrections and the reliability of observational datasets (reviewers 1 and 2), (ii) on the clarity of our descriptions of current process understanding and modeling that motivate the methods we have used (reviewers 2 and 3), (iii) on general quality of writing (reviewer 3), and (iv) on the importance of making clear the limitations arising from the simplifications we have made (reviewer 2).

In response, we have (i) added statistics and extended the description of the bias corrections (P8L16-24 and P11L5-7) and analysed and discussed the impact of the bias corrections on the projections in a new discussion section (P21L14-26) and a new figure in the supporting information (Fig. S10). We have also added a new supporting information figure (Fig. S2) and discussion (P10L18-22) on the reliability of the EN4 dataset used for the ocean bias correction. We have (ii) substantially rewritten and restructured the introduction so that it better represents current process understanding, its representation in models, and the motivation for the treatment we are proposing in ISMIP6 (P2L28-P4L6). We have (iii) followed all of the recommendations for improving the writing as suggested by reviewer 3 (see detailed responses below for page/line references), and thoroughly proofread and improved the full paper. We have (iv) ensured there are clear statements and extensive discussion on the assumptions and limitations of our proposed ocean forcing (e.g. P4L21, P5L24-27, P23L24-30). Please find our full responses below, including to all of the minor issues raised.

*Our responses to the comments are in **blue italics**. Where page and line numbers are given, these refer to the track changes version of the manuscript.*

Initial editor comments

Section 2.1: the whole first paragraph feels like more general background rather than being specific to Methods - would it be better integrated into the Introduction?

Agreed – this was also brought up by Reviewers 1 and 3. We have now integrated this paragraph into the revised introduction (section 1, specifically P2L28-P4L6).

There is, understandably, a lot of referring the reader to Slater'19 for details of the retreat method. I felt it would be helpful if this paper contained just a tad more information from that here – e.g. a line sketching out how equation 2 was arrived at and calibrated.

This is a good suggestion, also noted by the reviewers, and has been added at P6L4-6.

Ditto, a few words summarising the decision process that guided Barthel'19's choice of CMIP5 models would be useful context.

Agreed, this has now been added (P6L25-27).

There are summary/overview descriptions of what is done here in the abstract, Introduction, Discussion and Summary - it started to feel a bit repetitious to me, can at least the last two be condensed together?

Yes, we have condensed the discussion and summary descriptions together. The introduction and start of the methods section have also been revised and are now less repetitious (section 1, specifically P2L28-P4L6).

Reviewer 1: Surui Xie

To project Greenland Ice Sheet mass loss due to runoff and ocean property changes in different greenhouse gas emission scenarios, the authors present a modeling effort to apply ocean forcing on a continental scale. Two implementation approaches are presented, including “retreat” and “submarine melt”. Both implementations require a parameterization for submarine melting, and the authors consider local ocean velocity and ocean thermal forcing as two primary parameters for submarine melting. The former is implemented through subglacial runoff, and the latter is by ocean temperature. In the retreat implementation, glacier terminus positions are determined by estimated submarine melt rates. It is accessible to all ISMIP6 ice sheet models, but ice dynamics such as glacier advance due to motion or retreat due to calving were not considered. While the second implementation (submarine melt) takes into account of more factors affecting retreat projections, it is computationally expensive, and some of the considered factors are not currently well understood.

I think this paper is well motivated and well written. It provides a valuable framework for future modeling work of ocean forcing on Greenland Ice Sheet. I only have several fairly minor comments, listed below:

1) Atmospheric-driven runoff and ocean thermal forcing are two primary inputs for the models. While available data or models are limited, some of the assumptions made in this study need to be justified. In section 2.2.3, the runoff bias correction may be necessary to provide a continuous transition from present to future atmospheric forcing, but it may also result in spatial discontinuities, especially when applying a uniform temperature or salinity offset for the entire sector. Figure 2c shows a relatively small bias at Helheim – maybe this is a well monitored glacier so the models perform better? Many other glaciers have much larger values of runoff bias (please see your Figure S1). Could the sector-uniform offset and various bias be major contributors to the difference between different sectors in the projections? This question also applies to the ocean property correction.

We do not quite follow the reviewer's point on 'spatial discontinuities' – the runoff field is certainly spatially discontinuous (Fig. 6b) because adjacent tidewater glaciers have different runoff. It is true that the spatial discontinuity may be enhanced or reduced by applying a bias correction per glacier, but it is not introduced by the bias correction because different glaciers have different runoff in reality. In contrast, the way in which we have treated the ocean certainly introduces artificial spatial discontinuities (Fig. 6a) but we consider this necessary for the reasons outlined in section 2.3.1. The ocean sectors were chosen so that as far as possible, there is little variation in ocean properties within sectors (section 2.3.2; Slater et al., 2019) so that this sector-averaging should not strongly affect the temperature at an individual glacier.

It is true that the runoff bias correction is small for Helheim compared to other glaciers (Fig. S1). This must arise from relatively close agreement in simulated regional climate during 1995-2014 between RACMO2.3p2 forced by ERA-Interim and MAR3.9.6 forced by the CMIP5 models. Since the CMIP5 models receive no information from studies on Helheim, we expect that the small runoff bias correction is explained by the CMIP5 models doing a decent job of simulating present-day climate in SE Greenland, rather than because Helheim is well studied. Therefore we don't expect the runoff projection to be better (smaller bias) at one specific glacier over another, but there may be regional patterns whereby simulated regional climate differs strongly in the CMIP5 models compared to reanalysis. Fig. S1 suggests that the north of Greenland is one region where the models disagree strongly on simulated climate, reflected in the diverse runoff bias correction. For the ocean temperature biases, we again believe that the magnitude of the bias is controlled primarily by the representation of the regional ocean in the CMIP5 models, though clearly the present-day thermal forcing defined by EN4 is sensitive to how many oceanographic profiles are available (see answer to your second point, to the 4th major point of reviewer 2, and the new Fig. S2).

As we understand it, the main thrust of the reviewer's comment is whether the bias corrections lead to sector-by-sector differences in the projections. We have added a new figure to the supporting information showing the effect of the bias correction on the retreat projections (Fig. S10). In general, the bias correction can result in a few km of difference, ~0-20% of the retreat imposed on the ice sheet models, and is ~equally likely to increase or decrease the projected retreat. There are a few instances where the impact of the bias correction is larger; in NorESM1-M, the bias correction decreases retreat by 36% in SE Greenland and increases it by 20% in CW Greenland. These follow from the large thermal forcing bias corrections applied to this model (Fig. S3). Without this bias correction, the projected retreat for SE and CW Greenland in NorESM would be approximately equal, while with the bias correction, CW Greenland is projected to retreat much more. Therefore the bias correction can contribute to sector-by-sector differences in retreat projections, but we do not think it is a dominant control across all of our results (Fig. S10). These points have been added to a new section on the impact of bias corrections (P21L14-26).

2) Ocean temperature is a critical model input in this paper, and is detailed in section 2.3.3. I am curious about the temperature model selection. In Figure 3c, it seems to me that the MIROC5 model produces a quite different temperature profile than the observational EN4 data. Is it rational to use the EN4 data, by simply correcting the bias with a constant offset adding to the entire depth profile? I see that a depth-varying bias correction may lead to

unphysical profiles, but is there a reason to choose the mean difference between the specified 200-500 m depth range? According to the authors, this range is perhaps “most relevant to tidewater glacier grounding lines in Greenland”. But I feel a slightly different depth range (e.g., 100-400 m) can produce a significantly different offset – especially near the surface. Some discussion on the sensitivity of model to different temperature bias correction may be helpful.

We agree that the treatment of ocean thermal forcing is simplistic, but for the reasons now better outlined in the introduction (P3L21-P4L15; principally our lack of parameterisations for fjord processes, and lack of resolution in AOGCMs), we feel it is the best possible for the present purpose. Given that our ocean bias correction relies heavily on EN4, we have now added a figure to the supporting information that shows the coverage of EN4 and provides an indication of how much confidence we should have in the dataset (Fig. S2), and described this figure in the paper (P10L18-22). Please also see our response to major point 4 of reviewer 2 below, where we have undertaken a comprehensive discussion of the difference in temperature profiles between MIROC5 and EN4. Our reason for choosing the 200-500 m depth range is exactly as the reviewer states – this is the depth range most relevant to glacier grounding lines, but we agree that a different depth range could lead to a different bias correction. Following our response to your first major comment (and particularly Fig. S10), the bias correction can indeed have a significant, if not dominant impact on the retreat projections. This is however unavoidable until the representation of Greenland present-day climate improves in the CMIP AOGCMs. We hope that the addition of Fig. S10 and the substantial new discussion (P21L14-26) covers these points sufficiently.

3) For the two implementations, could the retreat history before 2014 be calculated? If this is possible and won't add too much extra work, figures illustrating the historic retreats before 2014 (and maybe comparison with available observations) would improve the integrity of modeling results. Such plots could be added to Figures 4 and 9 as positive retreats.

Thank you for this suggestion. We don't think it makes sense to use the CMIP5 forcings before 2014 because they do a poor job of capturing observed climate variability over the past 50 years. We could use the 'observational' datasets we used for the bias correction to go back in time, but we are concerned that this would add confusion to what is already quite an involved manuscript. We also don't have complete terminus position datasets going backwards in time – we do have long records for individual glaciers, but not for every single glacier as would be required to make the sector average plots in Fig. 4c and 9. Therefore we hope the reviewer understands if we don't follow this suggestion, and leave the integrity (i.e. past performance) of the parameterisation to Slater et al. (2019).

4) Page 17, line 18: Add “in” after “variability”?

Following suggestions from another reviewer, this sentence has been removed (P21L11-12).

5) Figure 5b: Maybe mark the ~350 m depth point on the dashed red profile? This may help readers understand the effective depth. I had difficulties in understanding the “deepest point” at the beginning – I thought it was rather a shallow (if not shallowest) point at a

distance of ~33 km by looking at Figure 5, then I realized that this is a point along the depth of a certain location.

Added as suggested (Figure 5).

6) Figure 11: Maybe add a vertical line in each panel to mark the largest glacier in the corresponding sector?

This is a good suggestion, but because Fig. 11 includes results from all 6 CMIP5 RCP8.5 models, there would be 6 vertical lines for the largest glacier. We think the figure would become confusing if we added all these lines, and so we hope that instead Fig. 10 shows sufficiently how melt rates change at the largest glaciers.

Reviewer 2: Neil Fraser

The paper investigates the effect of two different parameterisations for ice/ocean interaction, specifically at Greenland's glacier termini, in the context of future ocean/atmospheric conditions as predicted by a range of climate forecast models. This is part of a wider community effort to adequately couple ice sheet models with coupled (ocean/atmosphere) climate models. While both parameterisations consistently predict greatly increased mass loss from the Greenland Ice Sheet under a high greenhouse gas emission forcing regime, the spatial distribution of mass loss varies depending on which climate model is used.

The paper is mostly well written with a clear, direct message and nice figures. I like the ethos of finding a workable solution to a tough problem at hand and helping the wider community. However, as the authors acknowledge, many aspects of the physical environment are not considered. I felt that this paper really highlighted that major obstacles must yet be overcome before we can expect models to predict future mass loss accurately. I therefore think the results, while certainly valuable, should be interpreted qualitatively rather than quantitatively.

Major comments

1. Thermal forcing, TF, is very simplistic. The authors do well to flag up the shortcomings in section 4.3, but nonetheless there are major shortcomings. The empirical tuning might alleviate this to some extent, but I would still expect the omission of these processes to result in large uncertainties. I understand that sheer necessity offsets these issues to some extent, as the next generation of climate models require parameterisations such as the ones presented here. But I think that any quantitative conclusions about future sea level drawn from those models (which will undoubtedly be very high- impact results) should come with the footnote that ice-ocean parameterisation is still very basic. This is not a criticism of the authors: it's hard to see where major advances will come without much higher resolution AOGCMs.

We completely agree – given the lack of both ocean and ice model resolution, and our emerging understanding of processes, this has been a tough exercise to put together. We do

think it has significant value, because for example the retreat parameterisation has been empirically tuned and we can propagate uncertainty through to the sea level projections via the low and high cases (Fig. 9b). Similarly the CMIP5 models used have been chosen partly to bracket the range of warming we can expect (Barthel et al., 2019). Therefore even with this very simplistic approach we hope to bracket the possible range of sea level rise stemming from ocean warming. Our hope is that future work will improve on this approach; for us this seems most likely through (i) downscaling of the CMIP5 models through high resolution regional ocean models, or (ii) improved parameterisations for fjord and shelf processes. We have made the limitations of our approach clear at several points in the manuscript (e.g. P4L21-24, section 4.4, P23L24-30) and we will certainly ensure that future results drawn from this work feature prominent statements regarding the simplicity of the forcing.

2. Using annual mean temperature is inappropriate when melt is nonlinear in TF (Eq. 1). Mean melt is not equal to melt calculated from mean TF. The effect is likely small as the exponent is close to one, but it will result in a systematic error.

This is true but the systematic error is very small (as you also say). In Fig. R1 we estimate the size of this systematic error by considering MIROC5 RCP8.5 output for the SE sector during 2091-2100 at various depths. We consider monthly output (we expect that most of the sub-annual thermal forcing variability in these coarse, diffuse simulations is seasonal, so that monthly output should be sufficient to capture most of the sub-annual variability). We consider the relative difference between first raising TF to the power 1.18 and then taking an annual mean, compared to taking an annual mean and then raising to the power 1.18 (the latter is what is done in the paper). Even at 55 m depth, where there is substantial seasonality in thermal forcing, the systematic error introduced is <0.7%. For the depth layers most relevant to Greenland's tidewater glaciers (200-500 m), the systematic error introduced is <0.02%. This is very small compared to modifications made at other stages of the processing (e.g. the bias correction) and so we do not feel we need to re-run our results. We have added a statement to this effect at P14L7-9.

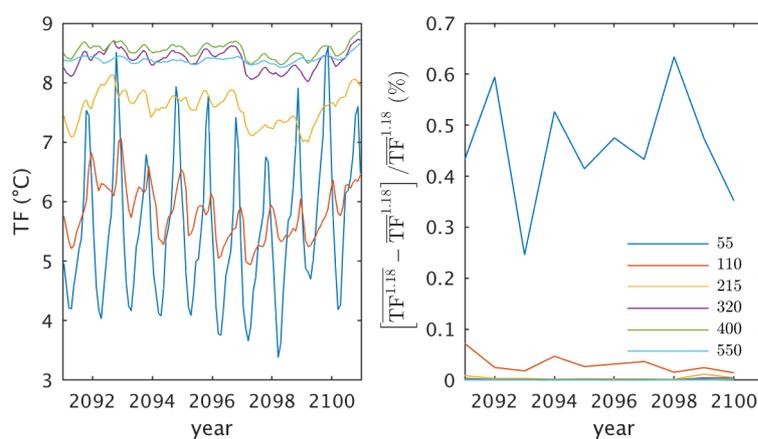


Fig. R1. Left: monthly output from MIROC5 RCP8.5 during 2091-2100 for the SE Greenland ocean sector at various depths from 55-550 m as indicated in the legend on the right. Right: systematic error (%) at the same depths introduced by using annual mean thermal forcing when melt is non-linear in TF.

3. My understand is that, if $dL = \text{melting} + \text{calving}$, retreat represents both terms while submarine melting represents only the first term. This should be made more explicit earlier on. Some of the language makes it a bit unclear what the inputs and outputs are for each parameterisations, and can seem at odds with Equations 1 and 2. I comment below on the specific instances of this.

Thank you for these suggestions. We have clarified up front that under the retreat implementation, all terms that determine glacier retreat are essentially assumed to be proportional to submarine melting, while in the submarine melt implementation, we represent only the submarine melt term and the ice sheet model does everything else (P5L24-27). We have also clarified the inputs and outputs of the parameterisations in each of the instances noted by the reviewer (see below).

4. Using EN4 for bias correction makes sense in theory, but do you have a sense of how many direct observations actually influence the EN4 gridded product for the regions/times of interest? EN4 has had issues in the Labrador Sea, and the EN4 temperature profile (Fig. 3c) is not a good representation of typical SE Greenland stratification (there should be a subsurface temperature maximum). You could add a figure in the supplementary material showing the mean EN4 confidence weightings for each ocean sector. Could bias correction be done instead using the available CTD profiles from each sector?

These are great points. We have added a figure to the supporting information (new Fig. S2) that shows how many profiles have influenced the bias correction for each ocean sector. The figure also shows a metric called the 'observation influence' that is provided along with EN4 to try to quantify the extent to which the gridded ocean properties are being driven by real observations. We would also point the reviewer to Figs. S4 & S5 of Slater et al. (2019), which show the location of CTD profiles in EN4 from 1960 to present. Together these suggest that the SE, SW, CE and NE sectors are relatively well observed, while the CW, NW and NO regions are rather sparsely sampled. We do also see from the new Fig. S2 that the shelf is sparsely sampled everywhere, and the number of profiles increases quickly beyond the shelf break. This spatial distribution of profiles was one of the motivating factors for extended our sector definitions beyond the continental shelf, as described in the text (section 2.3.2).

We believe that the EN4 SE Greenland temperature profile shown in Fig. 3c may not have a subsurface maximum because it is driven mostly by data from beyond the shelf break. Fig. R2 below shows all temperature profiles in EN4 in SE Greenland during 1995-2014 (see also new Fig. S2). There are many more profiles from beyond the shelf break ($n=6965$) than on the shelf ($n=617$). Those off the shelf do not show a subsurface maximum, while those on the shelf do, presumably because of the presence of polar water on the shelf. Since the profile in Fig. 3c of the manuscript is an average over both on and off-shelf areas, and since EN4 does some spatial interpolation which is likely dominated by the off-shelf profiles, the mean temperature profile for the SE Greenland sector ends up looking like those off the shelf, with no subsurface maximum.

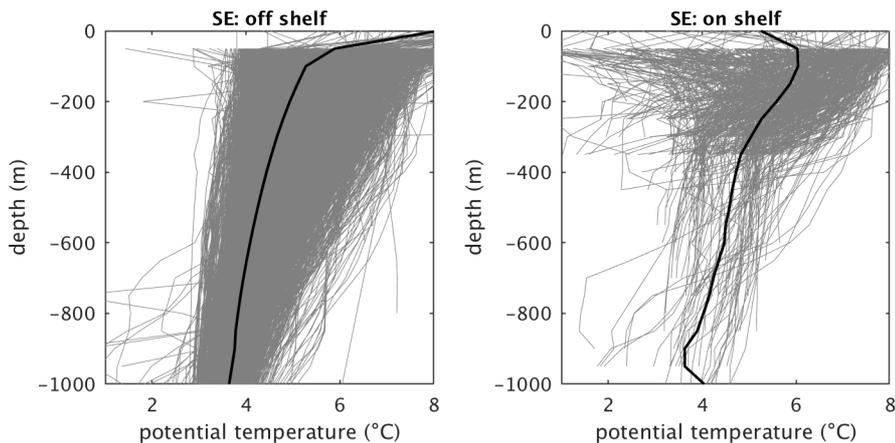


Fig. R2. Profile data behind the EN4 gridded product in SE Greenland in 1995-2014. Left plot shows all profiles beyond the shelf break (>1000 m bottom depth) while right plot shows all profiles on the shelf (<1000 m bottom depth). Thick black line shows the average profile.

A follow-up question is then why does the MIROC5 output have a subsurface temperature maximum? Fig. R3 shows the temperature stratification in SE Greenland for all 6 CMIP5 RCP8.5 simulations in the paper. Some (e.g. NorESM and HadGEM) look quite similar to EN4, while others (e.g. CSIRO and IPSLCM) are rather different. We suspect this variability will arise from differences in the relative proportions of Polar water and Atlantic water on the continental shelf.

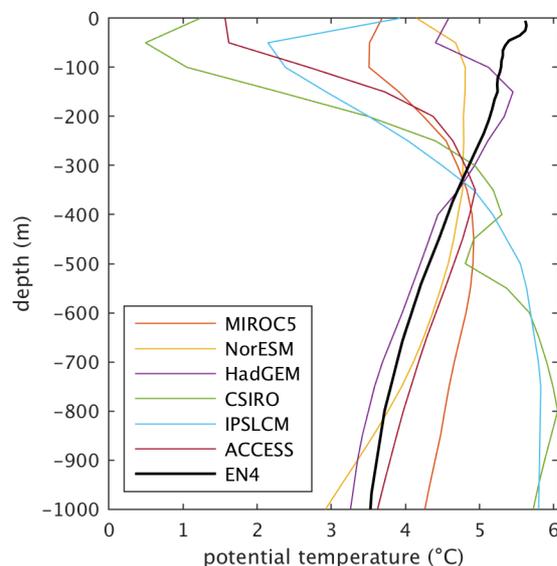


Fig. R3. 1995-2014 mean potential temperature in the SE Greenland sector in the CMIP5 RCP8.5 models and EN4.

The bias correction could perhaps be done using the actual CTD profiles, but each profile would represent only one position at one point in time, so it would not be straightforward to relate the measured properties to an annual mean over the whole sector, as is required for the projected time series. Essentially we would end up having to do our own temporal and

spatial interpolation, and without a lot of extra work that is beyond the scope of this paper, likely we could not do better than the gridded product already provided by EN4.

In spite of these difficulties, we think that the approach we have taken is the best solution available at present. It ensures that the mean temperature in the depth range relevant to tidewater glaciers is consistent with our best estimate of present day values. We agree however it is limited by the uncertainty on the observations, and perhaps it calls for a more detailed evaluation of the performance of CMIP models around Greenland, particularly in terms of how they represent different water masses. In the present work we feel the best we can do is to raise these issues in the paper (P10L18-22) and expand the discussion on the sensitivity of our results to the bias correction (new Fig. S10 and new discussion section 4.3).

Minor comments

Some of these are stylistic comments which the authors are entitled to disagree with.

P1L8: It's misleading to say that retreat is a function of submarine melting, do you mean subglacial runoff? I read this to mean that one parameterisation feeds into the other. This is related to major comment 3.

Apologies if this appeared misleading. Retreat is parameterized as a function of subglacial runoff and ocean thermal forcing (Eq. 1). But the way in which runoff (Q) and thermal forcing (TF) are combined ($Q^{0.4}TF$) is motivated by submarine melt rate parameterisations (Slater et al., 2019). That is, $AQ^{0.4}TF$ is essentially the in-plume submarine melt rate for an appropriate choice of constant A (Slater et al., 2016). This is the reason we stated that retreat is projected as a function of submarine melting. Perhaps the confusion arises because we have then described a slightly different submarine melt rate parameterisation for the submarine melt implementation (Eq. 2). To avoid confusion, we have changed the wording in the abstract (P1L10) and added clarification to the methods (P5L35).

P1L9: You should give RCP2.6 and 8.5 formal definitions, if not here then in the introduction or methods.

Good suggestion, this has been added to the methods (P6L28-30).

P2L21: Can you be more quantitative about the number of ice shelves than "a handful"?

Following comments from Reviewer 3, this has been changed to "ice shelves and floating ice tongues". Since we believe floating ice tongues could form seasonally (e.g. Moyer et al., 2019), and do not know of a quantitative assessment of their number, we don't think we can be specific here. Recognizing that "a handful" is a bit informal, we have changed to "several" (P2L30).

P2L25: Perhaps also worth mentioning here that since these regions are very poorly observed, especially in winter, large uncertainties remain with regards to Greenland fjord/shelf processes (i.e. while you correctly state that these processes are not captured in models, we still don't know exactly what we are trying to capture!).

This is a great point – we have added this to the revised introduction (P3L18).

P3L13: I would considering moving this first paragraph to the introduction. I see that it leads nicely into the second paragraph in 2.1, nonetheless when I finished reading the nice introduction it was frustrating to find myself reading what was essentially just more introduction.

Agreed, thank you for this suggestion. Following similar comments from the editor and Reviewer 3, we have rewritten much of the introduction, and merged the mentioned paragraph into the revised introduction (P2L28-P4L15).

P4L1: If submarine melt rate is denoted by \dot{m} and dL is linear in submarine melt, then should this not be make explicit in the expression for dL ? Otherwise, perhaps more careful language should be used. Again this ties into major comment 3.

*There are essentially two different submarine melt rate parameterisations being used here. The first could be expressed as $\dot{m}_{retreat} = A * Q^{0.4} * TF$. The retreat parameterisation linearly relates retreat to parameterised submarine melting: $dL = B * \dot{m}_{retreat} = B * A * Q^{0.4} * TF$. Once we write $\kappa = B * A$, we get to Eq. 1 of the paper. The second submarine melt rate parameterisation is that used in the submarine melt implementation, as written in Eq. 2. We accept that this was confusing in the initial submission because we were often referring to the implicit submarine melt parameterisation in Eq. 1. We have now explained the implicit submarine melt parameterisation in Eq. 1 (P5L35) and been more careful about language throughout (e.g. P1L10).*

P4L14: Personally I don't like multiplication signs in formulae, and I think it would read better if you dropped them.

These have been dropped, and for consistency with Slater et al. (2019) and to make the equations read better we have substituted capital delta for 'd' throughout.

P4L16: Refer to section 2.3.1 instead of "further below".

Following a rejigging of the description of the parameterisations, this sentence has been removed (P6L10).

P4L22: I'd change "even in future projections" to "particularly in future projections" since one would anticipate annual and summer means to converge as summer becomes longer.

Our manuscript was a bit confusing on this point – apologies. By 'summer', we meant June, July and August only, but later we were not consistent with this definition. With this definition, 'summer' cannot become longer, and one would anticipate annual and 'summer' means to diverge as more runoff happens outside of 'summer' (i.e. outside of June, July and August). We have clarified these points (P6L15-19).

P5L25: I really like this section on bias correction. Very clearly thought out and explained. It might be worth citing Menary et al. 2015 GRL, who explore CMIP5 temperature and salinity biases in the Labrador Sea west of Greenland, to underline your motivation.

Thank you, this citation has been added (P8L3-4).

P10L5: I understand that certain simplifications are necessary for these parameterisations to work in coarse climate models, but this paragraph completely ignores a lot of the research into fjord/shelf hydrodynamics. Much of these shortcomings are acknowledged later in section 4.3, but I think they should be made clear up front.

We are the glad the reviewer understands the need for the simplifications we have made, but we do not mean to diminish the importance of research into fjord and shelf hydrodynamics. We have now included a summary of this research in the introduction (P3L5-33) and stated very explicitly how simplified the thermal forcing is throughout (P21L29).

P10L22: If TF used in equation 1 differs from the TF used in equation 2 then perhaps they should be given different symbols or subscripts.

This is a good suggestion, but we are wary of overdoing the notation and think that the explicit statements of what TF is for each implementation (P9L17, P11L17 and section 2.5.2) should make it clear what TF means in each case. Hopefully the reviewer finds this sufficient.

P10L31: See major point 2.

Thanks – see our response above and P14L7.

P12L19: I'd remove the word "however" as it isn't necessary.

Removed as suggested (P15L25).

P13L18: This appears to be a strong argument for using more than one RCP2.6 model in your experiment.

Yes, we agree. The reason for using only a single RCP2.6 model is to reduce as far as possible the workload that is placed on the ice sheet models taking part in ISMIP6, and a decision was made much earlier in the process to focus efforts on the high emissions RCP8.5 scenario (Nowicki et al., 2016). Therefore we are unfortunately not in a position to include more RCP2.6 models at this stage.

P15L23: This seems to imply the thermal forcing is an input for the submarine melt regime only, when in fact it is an input for both. These two sentences could be rewritten to make it absolutely clear what the input and output variables are for each regime.

Following this comment, the two that follow concerning P15L25, and comments from the editor about repeating ourselves, this paragraph has been merged into the summary. We

have made clear what the input and output variables are for each implementation (P23L14 and P23L19).

P15L25: Change “. . .as they see fit” to “. . . as required” or similar, to avoid referring to a model as “they”.

Following our response just above, this sentence has been removed (P19L7).

P15L25: Sentence starting “Each implementation...”: This sentence is really jarring and frankly bizarre. If it wasn’t interesting you wouldn’t be writing a paper on it!

Following our response just above, this sentence has been removed (P19L8).

P16L31: Typo, “large uncertainty in. . .”

Corrected – thanks for spotting this (P20L15).

P18L3: Even without dense overflows, the properties of the water trapped behind the sill can (and will) be modified by downward mixing of buoyancy from the upper layers.

Agreed, we have added an acknowledgement of vertical mixing here (P22L10).

P18L14: The paragraph could also mention wind-driven heat delivery via the internal wave field, which has been found to deliver ocean heat to fjords in Greenland. Also, ideally the submarine melt parameterisation would capture (horizontal) ocean current speed adjacent to glacier termini, which we know is related to e.g. fjord width (i.e. Jackson et al. 2018) and impacts melting.

These two processes have now been described in the revised introduction (P3L11, P3L15, P3L22, P3L23). We would rather not mention them explicitly here, as then we feel we would need to mention all such processes. Instead this paragraph seeks to stay at a broader level (e.g. ‘we have neglected... the processes responsible for transporting and transforming ocean waters between the shelf and calving front’).

Fig. 3c: EN4 temperature profile looks suspect, what are the EN4 confidence weightings here? (major point 4)

Please see the response to major point 4 above. We suspect the profile looks as it does because it is being dominated by off-shelf profiles rather than on-shelf (the latter being where we would expect to see a subsurface temperature maximum due to the presence of polar water).

Fig. 6a: Is this figure saying that in 2100, ocean water will have flooded beneath the interior of the ice sheet? If so, this is a major result which should be flagged up in the text.

No, this is just showing that the submarine melt implementation defines submarine melt rate at every point under the ice sheet that is below sea level and connected to the ocean. An

ice sheet model would only use these interior melt rates if the ice sheet margin retreats all the way into the interior, which is very unlikely by 2100. This has been clarified in the figure caption to avoid confusion.

Fig. 9: To me negative retreat implies advance, so I'd change either the axis label (to "frontal position"?) or sign. Figure 10 uses positive values to denote mass loss, it'd be better if they were consistent.

We'd rather keep retreated values negative for consistency with Slater et al. (2019), in which the parameterisation is described. Fig. 10 is different because it denotes submarine melt rate rather than retreat, and we believe having increasingly positive values for increasing submarine melt is intuitive enough. We hope the reviewer understands this reasoning. We have however changed the axis labels and figure caption as suggested (and an additional instance on Fig. 4).

Fig. 10f: There's a missing dot above the m labelling the y-axis.

Now fixed – thank you for spotting this.

Overall, an important step towards the goal of coupled air-sea-ice climate models (but there is still a way to go).

Thank you – we agree there is a still a way to go!

Reviewer 3

The manuscript details the implementation and results of applying two different ocean forcing strategies (termed the retreat and submarine melt implementations) to a suite of AOGCMs contained in CMIP5 for use in the upcoming ISMIP6 to inform the next IPCC report (AR6). The authors present the model parameterizations, including their motivations and the limitations of each implementation. Then, they apply the implementations to a set of CMIP5 models and present the resulting forcing parameters (subglacial runoff and ocean thermal forcing) and model projected retreat and submarine melt rates.

The work presented in the manuscript presents an important step forward for modeling ice sheet response to various warming scenarios and consequent contributions to sea level rise. The authors do a great job presenting their model implementations and the motivations for each, and overall the ideas are well-organized and logically presented. However, as currently written the manuscript suffers from several key problems, outlined below, that must be addressed before it is suitable for publication.

1. Lack of clarity. Specifically, a clearer introduction would eliminate a lot of potential confusion later in the manuscript surrounding how ocean boundary conditions influence ice sheet mass changes and how these are both important in models. As currently written, the distinction between process understanding within the field and effective modeling of those processes is unclear. There is a constant switching between observations and modeling that leaves the reader guessing which one is currently being discussed, and descriptions of how

the processes are linked would greatly improve clarity for the modeling components of the writing. In some cases, significant detailed disciplinary knowledge is required to explain and justify the limitations and assumptions used within the model.

Thank you for these suggestions. We completely agree with these points. We have substantially rewritten the introduction (see all of section 1) to clarify the link between ocean boundary conditions and ice sheet mass loss, taking care to separate process understanding from observations (P3L5-20) and modeling efforts (P3L21-33). We hope this more detailed discussion of the processes involved provides a better background to justify the choices on boundary conditions that we subsequently make. In revising the introduction we have also integrated the former first paragraph of the methods section, and removed other sentences from the methods section that really belonged in the introduction. We hope the revised introduction clarifies the current process understanding, its representation in models, and the motivation for the treatment we are proposing in ISMIP6.

2. The discussion section is underdeveloped. Many interesting discussion points are presented and then left hanging without further exploration. Similarly, some of the assumptions and simplifications presented throughout the methods and results are not fully discussed, even where more information is available to inform a discussion (e.g. the magnitude of bias corrections and their interpretation; the influence of uncertainty in bathymetry; the potential uncertainty stemming from the assumption that submerged ice area remains constant).

We have expanded the discussion following the suggestions in the line-by-line comments below (e.g. on the atmosphere-ocean and sector-by-sector compensation at P20L9 and P20L26). We have also clarified our methods and assumptions in all the places brought up by each of the reviewers (e.g. the definition of hydrological basins at P7L21 and the definition of ocean sectors at P9L30). We have added key statistics of the bias correction at P8L22 and P11L5, a new discussion section on the bias corrections (P21L14) and a new figure to the supporting information (Fig. S10) showing the impact of the bias correction on the projections. We are not sure how we would meaningfully quantify the influence of uncertainty in bathymetry on our projections, but we have included clear statements on the importance of better bathymetric datasets (P22L14). Similarly, we cannot quantify the impact of the assumption of constant submerged surface area without knowing the evolution of submerged surface area, but have highlighted this as an area for improvement (P14L29).

3. The writing needs work both for content and grammar. As previously noted, the writing is overall unclear, with a lot of extraneous words at the expense of sufficient content in some places (e.g. p6 line 25: what is “inefficient” about current parameterizations? Are they computationally expensive or are they simply ill constrained?). Passive voice, dangling modifiers, and phrases not clearly linked to their parent idea are prevalent throughout, along with unneeded words and phrases (including connecting phrases such as thus, therefore, however, then, here). These detract and distract from the real power of the manuscript. Comma usage is quite poor (it is pointed out completely only in the abstract line comments, below), including inconsistent use of the Oxford comma and missing and incorrect comma usage, particularly around non-compound sentences. Lastly, please

proofread for subject-verb agreement, overall grammar, and consistency in formatting (e.g. capitalization of the Greenland Ice Sheet). This includes literature references, which are currently not consistently cited for the same ideas and lack cohesive formatting throughout (i.e. in-text citations are in random orders) and section references, which are an interesting mix of high level and lower-level references and do not consistently point to the most logical section (thus leading to confusion rather than clarity).

At p6 line 25 we should really have stated there are no parameterisations (efficient or inefficient) that can represent fjord-shelf exchange and fjord circulation without resorting to full hydrodynamic models. This has been clarified at P9L12. Thank you for raising the grammatical issues in the line-by-line comments below – all of these instances have been addressed and we have thoroughly proofread the manuscript to improve it throughout. We have removed many of the connecting phrases as suggested (e.g. P11L19 and P11L28 to name a few), but have retained a few where we believe they improve the readability of a passage. We have checked comma usage throughout, removing all instances of the Oxford comma. Capitalization of ‘Greenland Ice Sheet’ has now been applied throughout. We were only able to find one instance of citations being in random orders (the one noted by the reviewer below that has now been fixed), otherwise they were all ordered chronologically (as allowed by the Cryosphere guidelines). We have reviewed all the section references to ensure they point to the most logical section, and fixed the instances raised by the reviewer (e.g. P11L24). We hope that the reviewer finds the writing to be improved.

Specific (“line”) comments:

Abstract:

p1 Line 1: Please expand on what oceanic “changes” you mean

We meant changes in ocean properties (i.e. temperature and salinity). This has been clarified (P1L1).

p1 Line 3: Provide some examples of what you mean by “key physics” and make “limitations in processing understanding” less ambiguous

Added as suggested (P1L3).

p1 Lines 3 and 15: unnecessary comma

Removed as suggested (P1L11).

p1 Line 9: comma before respectively

Added as suggested (P1L12).

Introduction:

p1 Line 20: passive voice

Now fixed (P2L4).

p2 Line 2: dangling modifier

This has been removed (P2L6).

p2 Line 6: comma needed after thus

Added as suggested (P2L10).

p2 Line 8: CMIP6 is used as an acronym before it is defined

We now defer use of the acronym until after it has been defined (P2L19).

p2 line 19: it would be helpful for the reader to succinctly describe what CMIP is in this paragraph, as you have done for ISMIP.

Added as suggested (P2L19).

p2 Lines 21-22: ice shelves and floating ice tongues (and remove comma – not a compound sentence)

Corrected as suggested (P2L30).

p2 Line 22: clarify model design, or it sounds like you are designing the ocean forcing itself

Good point. This sentence has been removed in the revised introduction, but in a similar sentence (P4L7) we now clarify that we are trying to project ocean-induced mass loss, not design the ocean forcing itself.

p2 Line 32: inconsistent reference ordering

According to the manuscript preparation guidelines for The Cryosphere, the order of in-text citations “can be based on relevance, as well as chronological or alphabetical listing, depending on the author’s preference.” We ordered our references chronologically in all places except the instance noted by the reviewer, and this has now been fixed (P4L12).

p3 first full paragraph: clean up language and extra words

The first half of this paragraph no longer appears in the revised introduction. The second half has been condensed from the previous version (P4L24).

p3 Line 9: use of Oxford comma needs to be removed or added throughout manuscript

All instances of an Oxford comma have now been removed.

Methods:

Overview:

p3 first paragraph: many of these ideas are repetitive with information presented in the introduction (though with different sets of references). The temporal words (past decade, since) are misleading relative to the information presented (warming in the late 1990s) and references (2010).

Agreed. This paragraph has now been merged into the revised introduction (specifically P3L5-P3L33), where we no longer use these temporal words.

p3 line 27: the links between calving rate, glacier retreat, and ice sheet mass loss have only been tenuously drawn. Please include a clearer description of these physical processes prior to discussing their modeling.

The revised introduction better describes these links (P3L5-20), in advance of discussing their modeling (P3L21-33).

p3 line 31-32: both italics and quotations does not match the abstract formatting

We have removed the italics and quotations to ensure the formatting is consistent throughout the paper (P5L17).

p3 Line 33: taking part in ISMIP?

Yes – added as suggested (P5L18).

p4 eq 2: for the reader not intimately familiar with Slater et al 2019, another sentence about kappa (how it is calibrated, under what conditions it is applicable/scalable) would be helpful.

Agreed, this was also noted by the editor. This has now been added at P6L4.

p4 eq 1 and 2: switching the order of presentation of these two equations would provide order consistency with the presentation of the retreat, then submarine melt, implementations throughout the text

The order has now been switched (P6L2 and P6L8) – thank you for the suggestion.

p4 line 34: the “or CMIP6” is confusing here. It might be helpful to instead note above, where you are addressing your use of CMIP5 inputs, that the process would be identical for using CMIP6 inputs.

Changed as suggested (P6L24).

Atmosphere:

p5 line 3: define acronym MAR

Now defined (P7L7).

p5 line 4: the use of “physically downscaling” is confusing, especially given the later statement that the downscaling is done statistically. Removing “physically” would improve clarity.

Changed as suggested (P7L7).

p5 line 8: repetitive statement

This sentence has been removed (P7L11).

p5 section 2.2.2: If I am understanding correctly, hydrologic drainage basins are determined based on hydrologic potential (fine). Then, subglacial runoff is determined using surface runoff for those previously delineated basins. I think the authors need to better support and acknowledge the inherent assumptions here, including: 100% of surficial runoff reaches the bed and the surficial runoff reaches the bed with a similar spatial distribution, such that subglacial drainage basins with surface melt volume are appropriate for estimating subglacial melt volume. I would also like to see the use of $f=1$ substantiated.

The assumption of 100% of surficial runoff reaching the bed close to where it melted is supported by the high moulin density observed near the margins of the ice sheet during summer, so that it is unlikely meltwater travels far on the surface before reaching the bed through a moulin. For example, Yang et al. (2016) find that for a region of west Greenland below 1800 m.a.s.l., there is a moulin approximately every 25 km², limiting the mean distance travelled by meltwater on the surface to ~5 km. Meltwater may be temporarily stored in lakes, but Fitzpatrick et al. (2014) estimate this to be only 7-13% of total meltwater for a large catchment in west Greenland, a furthermore most of these lakes will drain to the bed by the end of the melt season (Fitzpatrick et al., 2014). There may be regional variability in these findings, but we have added a statement on these issues to our manuscript to support our assumptions (P7L26). Note that refreezing of meltwater within the firn is taken into account through our use of MAR (which includes a firn model).

Measurements of water pressure beneath the ice sheet to inform the choice of f are limited but have been obtained from boreholes. Meierbachtol et al. (2013) find mean values of $f = 0.8$ to 1 in boreholes at a land-terminating region of west Greenland, Andrews et al. (2014) find values ‘close to or above overburden’ (i.e. $f > \sim 1$) at a marine-terminating region of west Greenland, and Doyle et al. (2018) find $f = 0.95$ to 0.97 close to the front of Store Glacier in west Greenland. On the basis of these observations we believe that $f = 1$ is an appropriate assumption for our purposes, but clearly this is a simplification because f will have temporal and spatial variability. A summary of this discussion has been added to the manuscript (P7L21).

Lastly, a number of studies have found agreement between subglacial runoff estimated as in our manuscript (section 2.2.2) and estimates independently derived from oceanographic data (Jackson et al., 2016; Mankoff et al., 2016; Jackson et al., 2017), providing further confidence in our methodology. This has now been noted in the manuscript (P7L29).

p6 line 5: how is $Q_j(1995-2014)$ for RACMO or PROJ calculated? Is it a mean? Median? Cumulative?

Yes, this needs clarification (P8L16). It is a mean value.

p6 line 9: it would be helpful to provide some basic information on the bias corrections within the text (e.g. range and median + uncertainty).

Thank you for this suggestion. Over all glaciers and AOGCMs the bias correction is 2 ± 56 m^3/s (mean \pm standard deviation). The minimum and maximum corrections are -527 m^3/s and $+519$ m^3/s respectively. Considered as a fraction of the 1995-2014 mean runoff, the bias correction is 0.13 ± 0.47 . These statistics have been added to the manuscript (P8L22).

p6 lines 9-10: “we note that it might be thought preferable” is very wordy and passive language

This sentence has been revised (P8L26).

p6 line 11-13: this sentence could also be made stronger, particularly by quantifying the insignificance of the difference between RACMO and MAR (and noting the range of annual variability).

Agreed – this has now been added. For a given glacier and AOGCM, we define the interannual variability as the standard deviation of the 2015-2100 runoff after the runoff has been detrended by subtracting a lowess smoothed projection. The mean interannual variability over all glaciers and AOGCMs is then 74 m^3/s . This can be compared to the bias correction of 2 ± 56 m^3/s . This sentence now includes these calculations (P8L28).

Ocean:

p6 line 16: needs parenthesis

Following reviewer comments that this paragraph belonged in the introduction, this sentence has been removed (P9L3). We have checked the rest of the manuscript to ensure TF is in parentheses where appropriate.

p6 line 22: what atmospheric process? For calculating surface runoff?

Yes, this referred to the downscaling of the CMIP AOGCMs by the regional climate model MAR, but since this paragraph has been moved to the revised introduction, this sentence has been removed (P9L9).

p6 line 30-31: remove “details in”

Removed as suggested (P9L18 and P9L20).

p7 line 10: quantify “some distance”. What criteria did you use to determine the extent of the sector beyond the shelf break?

The sectors were extended to approximately the centre of the ocean basins and straits surrounding the ice sheet so that we capture the water properties of the currents at the shelf-break and cover a large ocean area to maximise the observations available within a sector (the observations are used in the bias correction and were used in the calibration of the retreat parameterisation). We did not extend to the full ocean basins because we do not wish to capture waters that may already have interacted with the ice sheet and are recirculating (e.g. the west side of Baffin Bay). The exceptions are Nares Strait, which is so narrow we included its full width, and the Arctic Ocean, Greenland Sea and Labrador Sea where extending to the centre of the ocean basin would encompass a very large ocean area and would sample less relevant water masses. In these regions the sectors extend approximately 150 km beyond the shelf-break to be consistent with the distance from the shelf-break to the centre of the ocean basins in the Irminger Sea and in Baffin Bay. A shorter description has been added to the manuscript (P9L30 onwards) and the supporting information now has a plot showing the abundance of CTD casts in the sectors (Fig. S2).

p7 line 13: quantify “coarse resolution”

Typical CMIP5 ocean model resolutions are 20 to 100 km around Greenland – this has been added (P10L5).

p8 line 7: as for the runoff, it would be helpful to present the range and some statistics (range and mean/median with uncertainty) on the applied bias corrections

Added as suggested (P11L5) – the mean correction is $+0.1^{\circ}\text{C}$ with a standard deviation of 1.5°C and a minimum and maximum correction of -3.1 and $+3.2^{\circ}\text{C}$ respectively.

Retreat Implementation:

p8 line 14: in the last equality, how are the units converted from salinity to temperature?

The units on the lambda constants have now been included (P11L15). With these units, the units of the terms in square parentheses is $^{\circ}\text{C}$.

p8 line 16: how were these constants determined, and are they valid for use here?

These constants (or values very similar) have long been used in polar oceanography and this linearised freezing point relationship is considered valid for polar water masses (e.g. Hellmer & Olbers, 1989; Jenkins, 1991). Jenkins (1991) state that this freezing point relationship is a linearized form of a full non-linear treatment given by Millero (1978). Given its very common

usage throughout ice-ocean studies, and the existing citation to Jenkins (2011), we do not feel we need to add any clarification to the manuscript.

p8 line 25: the thermal forcing itself is actually described in section 2.4.1, not section 2.3

Changed as suggested (P11L24).

p8 line 31: I'm not sure why this equation is presented independently of equation 2, which is the general form. I don't think this equation is substantially different enough to warrant a second presentation, particularly since the text notes the projection is relative to 2014.

We agree that this equation is similar to the first presentation of the retreat parameterisation (now Eq. 1), but we think it is worth showing separately because it shows more explicitly how the factors are calculated. That is, the runoff is calculated per glacier but the thermal forcing is calculated per sector. These are of course described in the text but we don't feel there is any harm in clarifying this in an equation. We hope the reviewer understands this reasoning.

p9 lines 1-5: the information on kappa presented here should be included the first time the equation is mentioned.

We completely agree – this information has been removed from this location and moved to where the retreat equation is introduced (P6L4).

Submarine melt implementation:

p10 line 10: what criteria were used (e.g. slope) to determine when a feature was large enough to be considered "blocking"?

There is no criteria to determine when a feature is large enough other than the depth of the feature (Morlighem et al., 2019). If we were beside the glacier at a certain depth and there is a clear path to the open ocean at that depth, then water at that depth is assumed to be in direct communication with the ocean. If alternatively there is no clear path, because all possible paths to the ocean are too shallow, then we would term this 'blocking'. We have tried to clarify this in the text (P13L13).

p10 line 29: no capitalization on where

Now fixed (P14L1).

p11 line 2: I'm not entirely convinced of the validity of using JUST the ocean bottom value for the thermal forcing, particularly if it's not the highest thermal forcing within the vertical profile. What rationale can be provided to suggest this won't underestimate melt rates?

The melt rate parameterised by Eq. (2) is intended to be the melt rate within a plume initiated by subglacial runoff. The plume rapidly mixes with fjord water and upwells deep water towards the fjord surface, so that the temperature profile within the plume is well approximated by the value close to the ocean bottom. Mankoff et al. (2016), their Fig. 5, is

the best observation of this effect but it has also been widely modeled (e.g. Slater et al., 2017, their Fig. 5c). In reality the melt rate will vary with depth (due to both plume dynamics and the fjord temperature profile) and laterally (due to the presence/absence of plumes and wider fjord circulation). Until continental ice sheet models are sufficiently resolved to be forced by more complex characterisations of submarine melting, we feel that in-plume melt rate estimated using the ocean bottom temperature is the best way to proceed. We have now noted the upwelling of deep waters by the plume in the manuscript (P14L5).

p11 line 14: typo

Fixed (P14L21) – thank you for spotting this.

Results:

p12 line 6: section heading misses emphasis within section on glacier runoff

The section heading has been changed to ‘Future subglacial runoff’ (P15L13).

p12 line 7: the use of “and” is confusing and suggests multiple runoff values are prescribed. Perhaps “each tidewater glacier/hydrological drainage basin” or “each tidewater glacier using its hydrological drainage basin”

Agreed – we have used your latter suggestion (P15L14) – thanks.

p12 lines 15-18: these sentences are more speculations than observations

These sentences have been removed (P15L22).

p12: The switch in referencing between largest glacier by flux and region between the text and figures is confusing. Suggestions to increase clarity are: add the glacier names to Figure 3 when the sectors are introduced and more importantly to Figure 7 (a and b) where the data shown is actually for individual glaciers and not the entire region.

We have implemented both of these suggestions (see revised Figs. 3 and 7).

p12 line 30: inappropriate semicolon

Now fixed (P16L8).

p13 line 3: section 2.4.1 refers to the section on thermal forcing, making this statement confusing. I’ve stopped noting odd section references after this point

This paragraph introduces results for the ocean thermal forcing for the retreat implementation. Since section 2.4.1 outlines the final definition of the ocean thermal forcing for the retreat implementation, we think this is an appropriate section to refer to. We have also checked and fixed other section references in the paper where necessary (e.g. P9L18 and P9L20).

p13 line 16: the supplementary panel figures are not labeled with letters

Now fixed.

p14 lines 23-24: inconsistent use of sector names (e.g. At Humboldt Glacier (NO), little increase. . .)

Now fixed (P18L3).

p15 line 3: The total count note is helpful, but confusing if you don't know offhand that 58 is the number of glaciers.

We have clarified in the text that we consider 58 glaciers in NW Greenland (P18L16).

Discussion:

p15 line 25: anthropomorphizing of ice sheet models ("they see fit")

Following comments from the editor that this paragraph repeated what is already in the introduction and summary, this line has been removed (P19L7).

p15 line 26: this statement implies that you have not already contrasted modeled ice sheet response between the two implementations

When we wrote the paper this was certainly true as the ice sheet model simulations were in progress. There is now a submitted manuscript (Goelzer et al., submitted) that does analyse the ice sheet model simulations but does not focus in detail on the difference between the simulated response to the retreat and submarine melt implementations. In any case, this line has now been removed on the basis of comments from the editor and reviewer 2 (P19L8).

p16 line 1: this suggests that averaging retreat over a population of glaciers resolves the fact that we cannot currently accurately represent calving or fully account for bathymetry in models of glacier termini. I would strongly disagree. Regional averaging may improve our modeled representation of retreat, which is of scientific import for further modeling and informing future investigations, but it does not fundamentally "ameliorate these issues".

We completely agree that regional averaging does not fundamentally solve anything. We meant to suggest that projecting retreat of a population of glaciers may be easier than projecting retreat of an individual glacier, because the projection becomes less sensitive to individual glacier specifics, such as bed topography. This sentence has therefore been removed (P19L15).

p16 lines 12-21: this reads as results, not discussion

Agreed – these sentences have been moved to relevant parts of the results (P16L4-10 and P16L30-33).

p16 line 23: subglacial runoff and ocean thermal forcing cannot be compared directly, as they are volume and temperature measures, respectively.

We intended to compare the relative increase (Figs. 7b and 8b) rather than the absolute increase (Figs. 7a and 8a). This has been clarified in the text (P19L26-P20L4).

p16 line 25-26: impact on what? retreat? mass loss?

We meant the impact on the projected retreat and submarine melting. This has been clarified (P20L7).

p16-17 lines 27-2: I would like to see this idea explored further (and clarification on why the authors switch from generalized “retreat and submarine melt projection” statements to just the “retreat implementation”). What are the implications of this compensation across model suite versus across retreat scenario comparisons?

We have expanded this idea by quantifying the spread in runoff and ocean thermal forcing versus the spread in retreat and submarine melting (P20L10). The switch to just the retreat implementation was not intentional and has been rectified (P20L17, P20L20 and P20L25). The implication is that once applied to an ice sheet model, one can expect a greater spread in sea level projections from the low versus high retreat scenarios than from the medium retreat scenarios that use different CMIP5 AOGCMs, or from the submarine melt implementation with different AOGCMs. In terms of sampling uncertainty in sea level projections, this means that it may be more worthwhile to prioritise ice sheet simulations forced by more retreat scenarios for a given AOGCM, rather than simulations forced by additional AOGCMs. We can however only make this statement for the forcings we have presented here (retreat and submarine melting) and for the 6 AOGCMs we have considered. It might be the case that improved process understanding of ocean forcing would suggest the use of different parameterisations, for which there may be more spread in forcing between AOGCMs, and it might be the case that other AOGCMs not considered here do not show this ‘atmosphere-ocean’ compensation. These points have been added to the manuscript (P20L20-25).

p17 lines 3-8: again, develop this idea further

We have highlighted a specific example of this sector-by-sector compensation, and expanded our discussion of the implications (P20L26-P21L2).

p17 lines 18-19: rewrite – currently not a sentence

Yes, this was not clear as written. In fact this sentence repeats what was already stated at P21L11 and so it has been removed.

p17-18 lines 33-6: This paragraph leaves out the problem of incomplete bathymetry observations, a key area where improved models will still be limited by lack of observations.

Agreed, this has now been added (P22L14).

p18 line 10: other processes DO play a role, not MAY play a role. Perhaps the authors mean to emphasize that the other processes MAY play a SUBSTANTIAL role?

We completely agree that other processes play a role. We have followed the suggestion of adding 'substantial' (P22L19).

p18 line 13: dash needed between ice sheet and ocean

Now added (P22L22).

p18 lines 17-20: can you make the argument that the physics of plumes are well understood if we severely lack constraints for key constants?

Our use of 'physics of plumes' here was meant to refer to the bulk dynamics of the plume, i.e., the entrainment and the evolution of bulk plume velocity, temperature and salinity. We would argue that the bulk dynamics are relatively well understood on the basis of laboratory experiments and field observations. Clearly we severely lack constraints for parameters in the submarine melt parameterisation, but this can be seen as separate from the bulk dynamics of the plume since for tidewater glaciers the bulk dynamics are highly insensitive to the induced submarine melting (Slater et al., 2016). We have made this separation between bulk dynamics and submarine melting clearer in the text (P22L26).

Figures:

Overall: It would be helpful to use a different color scheme for showing comparisons of model runs than those colors used for plotting different sectors. This would allow the reader to more readily distinguish between sector-based results versus those averaged over the entire ice sheet that show model variability.

This is a great suggestion – we have used the previous colour scheme for the sectors and tried a new colour scheme for the models. We hope this makes the plots easier to digest.

Figure 2: a zoom-in of the shaded portion of panel c (which is unlabeled but presumably indicates the time period used for the bias correction) is needed. As shown, it is difficult to see the similarities and differences between the datasets used to make the bias correction, and the zoomed out version suggests some apparently large differences between RACMO and MAR that are not substantially addressed within the text.

A zoom-in has been added – thank you for this suggestion. We have clarified in the caption that the shaded portion corresponds to the present-day period used for the bias correction. We have a statement on the discrepancy between RACMO and MAR to the manuscript at P8L19-21.

Figure 3: Add the resolution of the climate model shown in panel a for clearer comparison with panel b (which has a stated resolution).

Added to caption as suggested – the resolution is 1.4°.

Figure 5: a is missing units; the yellow and red points/lines are not labeled

Units have been added to (a). The caption now explains what the yellow and red points/lines refer to.

Figure 6: b-why is such a large portion of the ice sheet showing no-data (entire drainage basins do not have subglacial runoff values)?

Since the submarine melt implementation applies only to marine-terminating glaciers, we have not calculated runoff for land-terminating glaciers. The drainage basins on Fig. 6b that do not have runoff values are land-terminating according to our hydrological flow routing (section 2.2.2). This has been clarified in the figure caption.

Figure 7: caption – subject verb agreement; b – is this also showing the largest glacier by ice flux for that sector? Also, the figure is missing units

Subject-verb agreement fixed. Subplot b is indeed showing the largest glacier by ice flux in each sector – this has been clarified in the caption. Subplot b has no units because it is a relative anomaly, i.e., it is the absolute anomaly divided by the 1995-2014 mean. This is now stated in the figure caption.

Figure 10: The labeling with glacier name and region is quite helpful

Thank you.

Figure S1: The SE and NW colors are difficult to tell apart.

It is true that they are quite similar but given the number of colours we need to use in the paper (7 ice-ocean sectors and 7 different CMIP5 AOGCMs) it is hard to find another colour that isn't too similar to something else. If possible we would like to leave these colours as they are.

Figures S3-S8: why not utilize some of the white space where there are no subplots as adequate space for the legend (particularly where it has been separated across multiple plots)

The legends have been moved into the white space – thank you for the suggestion.

Acknowledgements: Michalea's name is spelled wrong

Now corrected (P34L25) – thank you for spotting this.

Literature cited (that is not already in the manuscript)

Fitzpatrick, A. A. W., Hubbard, A. L., Box, J. E., Quincey, D. J., van As, D., Mikkelsen, A. P. B., Doyle, S. H., Dow, C. F., Hasholt, B., and Jones, G. A., 2014. A decade (2002-2012) of supraglacial lake volume estimates across Russell Glacier, West Greenland. *The Cryosphere*, 8, 107–121, doi:10.5194/tc-8-107-2014.

Goelzer, H. and the ISMIP6 team (38 authors). The future sea level contribution of the Greenland ice sheet: a multi-model ensemble study of ISMIP6. Submitted to *The Cryosphere*.

Hellmer, H. H. and Olbers, D. J., 1989. A two-dimensional model for the thermohaline circulation under an ice shelf. *Antarctic Science*, 1, 325-336.

Jenkins, A., 1991. A one-dimensional model of ice shelf-ocean interaction. *Journal of Geophysical Research* 96 (C11), 20671-20677.

Millero, F. J., (1978). Freezing point of sea water. Eighth report of the Joint Panel of Oceanographic Tables and Standards, Appendix 6, *UNESCO Technical Papers on Marine Science*, 28, 29-31.