Comment on tc-2021-175
Sophie Nowicki

Community comment on "Brief communication: A roadmap towards credible projections of ice sheet contribution to sea-level" by Andy Aschwanden et al., The Cryosphere Discuss., https://doi.org/10.5194/tc-2021-175-CC2, 2021

We post this comment as members of the scientific steering committee of the Ice Sheet Model Intercomparison Project for CMIP6 (ISMIP6): Sophie Nowicki, Antony Payne, Eric Larour, Heiko Goelzer, William Lipscomb, Helene Seroussi, Ayako Abe-Ouchi, Andrew Shepherd.

We welcome comments and suggestions for follow-on extensions for the ISMIP6 protocol. We are pleased to see these suggestions from some of the participants of the original ISMIP6 project.

General comments
1. The commentary in this manuscript is focussed on a single multi-model comparison (ISMIP6 Greenland). Many of the suggestions would similarly apply to other initiatives in the ice sheet modelling community (e.g., ISMIP6 Antarctica, ABUMIP, and LARMIP2), outside of it (e.g., GlacierMIP or even CMIP), and equally to many individual projections. While ISMIP6-Greenland can still be used as an example, we would appreciate clarification on why the authors did not undertake a broader approach (for example, using both ice sheets).

2. The framing of how ISMIP6 results have entered the AR6 report is problematic. The authors state: "This ISMIP6 distribution has since been adopted as the foundation for the IPCC AR6 consensus estimate of sea level contribution from ice sheets." As of now, the AR6 report has not been released, so it is not yet possible to say how the AR6 authors used ISMIP6 results. IPCC guidelines (Mastrandrea et al., 2010) on uncertainty do not support the use of a single line of evidence as the basis for an assessment, and therefore the AR6 consensus estimates will likely draw from many studies.

3. The ISMIP6 protocol and goals are not always accurately represented in this commentary. For example, participants were not limited to a "single" contribution, so anyone could have contributed to ISMIP6 a series of simulations to sample uncertainty in the way the authors describe. See specific comments below.

4. The authors generally do not clearly distinguish ice sheet model uncertainty from forcing uncertainty. Their proposed method of partitioning uncertainty relies on an oversimplification of the ISMIP6 protocol as it ignores complexities not present in, for example, Aschwanden et al. (2019).
5. Can the authors more clearly link Sect. 5 to the rest of the paper? This section seems to focus on a separate issue that might be relevant to a different audience. Taken in isolation, the ISMIP6 Greenland projections do not provide a strong basis for the commentary.

**Specific comments**

*L. 3: "...few models reproduce historical mass loss accurately"

This is an important argument in the manuscript, but it requires a more nuanced treatment. Questions that need to be addressed are: Can and should the ISMIP6 projections be expected to reproduce observed mass changes? Would reproducing mass change over the relatively short observational period be a meaningful quality criterion for centennial time scale projections? We address these points in more detail below.

*L. 6: "the future sea level contribution from Greenland may well be significantly higher than reported"

Little evidence is presented for this statement in the main text. If the claim is based on the possibly larger uncertainty, it could mean both higher and lower contributions. We suggest reformulation.

*L. 8-11: “Finally, we note that tremendous government investment ...”

We agree on the need for more investment in ice-sheet and sea-level research. However, if the "significantly volunteer effort" refers to ISMIP6, the wording ("is founded on") may exaggerate the role of ISMIP6 results in government planning. See the comments above on multiple lines of evidence in IPCC assessments.

*L. 20: “accurate”

To be "accurate" implies comparison to observation. Projections are of the future, so it is impossible to assess their accuracy now. We suggest a different phrasing.

*L. 21: “defensible assessment of uncertainty”

The IPCC has robust guidelines to ensure that uncertainty is indeed assessed (e.g., Mastrandrea et al., 2010). As written, this sentence calls into doubt the IPCC methods of assessing the uncertainty. Please rephrase.

*L. 23: “Ice Sheet Model Intercomparison for CMIP Phase 6 (ISMIP6)”

The correct name is “Ice Sheet Model Intercomparison Project for CMIP6 (ISMIP6)”.

*L. 24: “(ISMIP6)”

References for ISMIP6 should include Nowicki et al. (2020) and Payne et al. (2021). Nowicki et al. (2020) contains the description of the experimental protocol which is mostly criticised here.

**Figure 1:**
This figure is presented as a key line of evidence in Sect. 1, but important choices of
selecting data, processing output, and combining results have not been motivated and described with sufficient detail. In particular:

* It is unclear how the uncertainty envelope has been derived. The original figure for the historical period (Fig. 4, Goelzer et al., 2020) does not attempt to show uncertainty in the model results, but simply reports the ensemble. Please clarify how the 90% credibility interval was derived.
* The uncertainty envelope of, e.g., IMBIE depends on the choice of assuming fully correlated or uncorrelated errors when accumulating uncertainties (compare again Fig. 4 in Goelzer et al., 2020). Can you motivate your choice for the narrower envelope and discuss the implications?

* Please discuss the conceptual difference between the historical experiments (until 2014) and the projections (2015+). While modellers were free in their simulation of the historical period, the projections were tightly constrained by CMIP model output. Combined analysis across those experiments (e.g., 2008–2020) is therefore difficult to interpret. See also the comment on L. 45 below.

L. 27-28: "ISMIP6 produced probabilistic distributions of projected sea level contribution"

ISMIP6 did not produce probabilistic results – it presented ensembles with no probabilities attached. Others (e.g., Edwards et al., 2021) have used these ensembles to make probabilistic assessments, but their analysis includes additional information and not simply the ISMIP6 results.

L. 31: “This ISMIP6 distribution has since been adopted as the foundation for the IPCC AR6 consensus estimate.”

What is the basis for this statement? See the general comment above.

L. 35: "Our skepticism regarding the ISMIP6 projections is based on the premise that accurate predictions of the cryosphere’s contribution to sea level require that models:

1. Fully characterize uncertainties in model structure, parameters, initial conditions, and boundary conditions.
2. Yield simulations that fit observations within observational uncertainty."

Although the requirements are laudable, it is nearly impossible for any study to achieve both, let alone a large multi-model project such as ISMIP6. To “fully characterize uncertainties” is demanding indeed, but this is not a problem within a single study, as long as other research addresses other uncertainties. It is also impossible for a model to fit all available observations within observational uncertainty, unless the model is overtuned. One can argue that particular observations are supremely important, but it is not obvious that recent mass loss is more important than, say, an accurate simulation of observed ice extent, thickness, and velocity.

Although this paragraph is set up to elaborate both requirements, the subsequent analysis of Fig. 1 is based only on the second requirement. See the comments above on the augmentation of ISMIP6 results.


The period 2008–2020 straddles the ISMIP6 historical period (ending in 2014), during which modellers used the forcing of their choice, and the future period (2015+), when forcing was provided by climate models. Mass loss from 2015 reflects, in part, natural variability that would not be reproduced by the climate models.

Why focus on 2008–2020? Figure 1 (and Fig. 4 in Goelzer et al., 2020) start before 2008, so the statement applies also to a longer time period.
L. 45: “Underestimating recent mass loss likely translates into underestimating mass loss at 2100 as well.”

This is not necessarily true. As pointed out above, it is important to distinguish the historical period (before 2015) from the projections. Generally speaking, in order for ice sheet mass loss to be accurately simulated for the recent past (2008–2020), two things are required: The climate forcing should be accurate, and the ice sheet model should accurately represent the processes translating this forcing into mass loss.

Modellers were free to choose their own forcing for the historical period; in most cases, they used SMB output from regional atmosphere models such as RACMO and MAR. Some ice sheet models may have applied SMB forcing that was biased positive for the period 2008–2014. Also, most models did not apply forcing to outlet glaciers before 2015.

ISMIP6 climate forcing from 2015 onward was derived from the CMIP5 and CMIP6 Earth System Model (ESM) ensembles. A known complication of this forcing is that interannual variability (known to be important in determining Greenland’s mass budget) is seldom in phase with the observed climate. This significantly complicates a model–observation comparison over a short period of 12 years.

For these reasons, models that underestimated mass loss during 2008–2020 might have been responding realistically to biased forcing. To demonstrate that they underestimate mass loss at 2100, one would need to argue that (1) the ESM-derived SMB forcing through 2100 is biased positive, and/or (2) the models underestimate recent mass loss when forced with an accurate SMB and output glacier forcing.

L. 61: “In this short communication we will not address the issue of uncertainty in the forcing F (Team et al., 2010) but concentrate on the uncertainties arising solely from ice sheet models.”

There appears to be some confusion about this separation. In ISMIP6 (unlike Aschwanden et al., 2019), the SMB and outlet glacier forcing are prescribed and are not part of the ice sheet model formulation. Much of what can be assigned to parametric uncertainty in Aschwanden et al. (2019) has no equivalent as ice sheet model uncertainty in the ISMIP6 projections, but rather is connected to uncertainties in the forcing. Thus, the uncertainty framework described here does not apply in the same way to ISMIP6. Goelzer et al. (2020) state explicitly that we did not sample RCM uncertainty (which could loosely map to some of the uncertainties in PDD factors), but these complexities are not discussed here.

Since uncertainty in the forcing (SMB and outlet glaciers) could account for the issues highlighted by Fig. 1, it seems appropriate to address that uncertainty.

L. 102: “This lack of knowledge induces parametric uncertainty, for example, different values of thermal conductivity within firn might lead to different predictions of sea level contribution.”

This is true, but the thermal conductivity within firn is relevant only for the RCMs computing the SMB. So this example pertains to MAR and RACMO, but not the ensemble of ISMs.

L. 134: “(Slater et al., 2020)”, “(Barthel et al., 2020)”

Please cite two papers by Slater et al.: Slater et al. (2019) and Slater et al. (2020). The
reference to Barthel et al. (2020) should be replaced by Nowicki et al. (2020), since only the latter paper shows atmospheric boundary conditions.

L. 134-137: “To allow a wide range of modeling groups to participate, concessions had to be made, resulting in an experimental setup that does not always reflect advances in modeling practices since the Sea-level Response to Ice Sheet Evolution (SeaRISE Bindschadler et al., 2013; Nowicki et al., 2013) project, including calving and frontal ablation.”

Please rephrase to better reflect the ISMIP6 standard and open protocols. The open protocol allowed groups to include calving and frontal ablation, and any other advances since SeaRISE.

L. 142-144: “Each group decided on the best parameter set for their simulations. This means that each model contributes a point estimate consisting of a single ‘best’ model run to the larger ensemble”

This statement is incorrect, since the protocol does not limit the number of simulations per group. Some groups submitted multiple runs with different physics options, resolutions, initial states, etc.

L. 146: “While it is difficult to gauge the magnitude of this underestimation, Aschwanden et al. (2019) suggest that the parametric uncertainty (inter-quartile range) at 2100 is 0.3 and 12.9 cm SLE for RCP 2.6 and 8.5, respectively, which is larger than the model uncertainty suggested by the ISMIP6 experiments (0.8 and 3.4 cm SLE, respectively).”

As pointed out above, some of the uncertainties in Aschwanden et al. (2019) would need to be mapped to forcing uncertainties in ISMIP6, in order to meaningfully compare the uncertainty in the two studies.

L. 148: “If one takes the Aschwanden et al. (2019) estimate of parametric uncertainty as reasonable, then the variance in ISMIP6’s predictive distribution is greatly underestimated with respect to the real variance”

There are reasons to think that the Aschwanden et al. (2019) estimate of parametric uncertainty may be too wide: (1) The range of sampled PDD factors in their simplified melt model dominates their uncertainty. That range was motivated by measurements of PDD factors in the field. However, comparison with results of similar models in GrSMBMIP (https://doi.org/10.5194/tc-14-3935-2020) shows that PDD factors that produce reasonable simulations of the historical SMB are much lower than the upper range used in Aschwanden et al. (2019). Incidentally, the way the forcing is designed in Aschwanden et al. (2019) does not allow them to constrain their PDD factors based on their own historical simulations: a problem that prevents the second premise proposed in this commentary (requiring a projection system to reproduce historical behaviour) from being applied to those results.

(2) The Aschwanden et al. (2019) projections have a shortcoming in the forcing protocol, using spatially uniform temperature forcing that likely overestimates temperatures in the ablation zone. Taken together, these points suggest that results from Aschwanden et al. (2019) may be biased in both maximum contribution and uncertainty range and are not simple to interpret in relation to other estimates.

Please provide a stronger argument as to why the Aschwanden et al. (2019) projections should be considered as reasonable, bearing in mind the above comments.

L. 155: “Initial ice sheet extent varied among models by up to 17%.”
How well does the 17% figure characterize the variance? Please comment on the entire distribution, not just the outliers.

ISMIP6 guidelines, as expressed in Goelzer et al. (2020), state that several metrics should be considered to evaluate a model’s representation of the initial state and possible biases (see, for example, Fig. 5 and its discussion in the text). This could be mentioned here as well.

L. 183-184: “with random climate and ocean forcings developed in collaboration with their respective modelling communities (cf. Robel et al., 2019”).

Please clarify what is meant by “random climate and ocean forcings.”
Also, it could be noted that some ice sheet models cannot run large parameter ensembles, which would reduce the number of models able to participate.

L. 192-193: “To address both of these problems simultaneously, we advocate for conditioning ensemble predictions on relevant observations.”

Are there other examples in the literature that could illustrate this approach?

L. 247: “International governments directly support development, maintenance, and operation of the Earth System Models that serve as the foundation for CMIP6 (Eyring et al., 2016), and this financial support has contributed to a suite of models that now convincingly reproduce observed climate variability (Jones et al., 2013). It is time to similarly bring ice sheet modeling to an operational level and support it with the funding the problem deserves.”

Here, the authors might mention the work in progress to couple ISMs within ESMs.

L. 265: Data availability


Please update to reference Goelzer et al. (2020). The data are the same as given and displayed in the paper, e.g. in Fig. 7.

References

