Calibration of a coupled ice sheet-ocean model using observations of ice dynamics and basal melt in West Antarctica

Brad Reed, Jan De Rydt, Kaitlin A. Naughten, and Daniel N. Goldberg

Replies to referee comments

Thank you to the reviewer for their helpful and thorough feedback on our manuscript. We have responded to their comments below. The referee comment is shown in black; our replies are in blue and the new text in the manuscript is in *blue italic.*

Reviewer 2

The authors present a new calibration approach for coupled ice sheet-ocean models. Instead of calibrating against observations of basal melt rates alone, they calibrate against combined datasets of basal melt rates, changes in ice thickness, and changes in ice speed. Additionally, the authors modified the 3-equation parametrization for basal melt to elevate basal melt rates while keeping overall melt low. Furthermore, a dynamic ice-shelf geometry was used during calibration. An optimal parameter set for the Amundsen Sea sector has been identified and used for an extreme-case future emission scenario. This run has been compared with a model run calibrated using the common approach: static ice shelves, a simple 3-equation formulation, and only basal melt. The authors show that calibration matters for sea-level projections, and common calibration approaches might lead to an underestimation of ice mass loss.

The paper is very well written and structured. The figures are of high quality and well-chosen to support the main text. The motivation is clear, and the experimental design is well-designed. I like the detailed description of the models and methods used. I also very much like the detailed discussion of limitations and potential next steps.

However, I have two general remarks:

You mention equifinality, and that different parameter combinations yield similar good fits during calibration. Your Figure B4, however, indicates that individual basins show different fits even when the combined calibration shows equifinality. E.g., the Pine Island sector requires a somewhat larger heat transfer coefficient than the other basins for lather transition heights to match the optimal parameter setting you have chosen. I was wondering whether you could discuss what makes Pine Island special in this regard compared to PSK and Thwaites, and whether this information could be generalized to some extent.

The difference in fits to observations (different ice responses) and optimal melt parameter values are likely related due to the uncalibrated ice-sheet model. We have now included several comments explaining this throughout the manuscript and how this will be improved in future experiments.

Now included at the end of “Sec 2.4.2 Hindcast ensemble and ocean-model calibration” with the following text:

- “*The methods presented here provide a framework for calibrating ocean-model melt parameters in a coupled system, representing a first step toward fully calibrating a coupled ice sheet-ocean model. To our knowledge, the only other study to jointly calibrate ice and ocean parameters is \citet{rosier2025-Calibrated}, who find no clear relationship between melt parameters (E_0, Γ_{TS}) and ice dynamical parameters (m, n), suggesting that the two parameter sets can be treated somewhat independently. The results we present will demonstrate the value of calibrating melt parameters using the transient-coupled system, rather than the common approach of using a static ice geometry and melt observations alone. We note that the current framework assumes a single realization of an uncalibrated ice-sheet model, meaning that optimal melt parameter values may partially compensate for structural deficiencies in the ice-sheet model; a natural extension would be to jointly calibrate ice-dynamics parameters alongside the melt parameters.*”

The other remark is about your experimental design. You nicely demonstrate the improvement in melt rates achieved by your approach compared to what is usually done. However, you apply three changes simultaneously: 1) calibration for three different datasets; 2) adaptation of the 3-eq-formulation; 3) dynamic geometry. While you show sensitivity to calibration across single datasets in your Figure B1-4 (which I like very much), it is difficult to tell how much of the improvement comes from the dynamic geometry and how much from the changes in the melt parametrization. It would be helpful to discuss that.

We agree that introducing these simultaneous improvements in the calibration may introduce confounding effects, and that a more systematic analysis allows to disentangle and quantify the improvements associated with each individual component.

So far, we have presented results for:

1. the **modified** melt parameterization with **evolving** cavity geometry; and
2. the **default** melt parameterization with **static** cavity geometry.

In the revised manuscript, we will additionally include results for the **modified** melt parameterization with a **static** ice-sheet geometry (static-melt-mod; SM). This allows us to separate the contribution of the new parameterization from that of the transient-coupled calibration.

The main conclusion from this is that the modified melt parameterization, with its wider parameter space, allows higher melt to be generated at the grounding line compared to the default case. However, this calibration still underperforms compared to the transient-coupled calibration, which better captures upstream thinning and speed changes, particularly for Pine Island Glacier. We will add the following text to the methods and results sections:

- Methods, “Sec 2.4 Experiment setup”:
    - *“For this static-melt approach, we include both the default melt parameterization (static-melt-def; SD) and the new modified parameterization (static-melt-mod; M).”*
- Results “Sec 3.2 Melt-only calibration”:
    - “*We compare our new coupled calibration approach with the common practice in the literature of using only observed \textit{melt} rates for a \textit{static} ice geometry. To make this comparison fair, we identified the highest-scoring simulation that uses the \textit{modified} melt-rate parameterization (static-melt-mod), as well as the most common approach which uses the \textit{default} parameterization (static-melt-def).*

      Unlike the transient-coupled calibration, where high likelihood scores are concentrated in a narrow range of parameter combinations, the static-melt-mod calibration produces high scores across a broad range of parameters, with generally lower melt rates than the transient case (Fig.~\ref{fig:app-heatmap-meltGstat}). For the top-ranked simulation, with $\mathcal{P}=(4,1,200)$ (SM), this results in greater slowdown and less thinning in the PIG basin leading to grounding-line advance (third column in Fig.~\ref{fig:results-spatial}). Across the remainder of the domain the thickness and speed changes are broadly similar to the transient calibration. Hence, this set of parameters is ranked ninth overall when comparing all modelled metrics against observations for a transient geometry (dashed orange outline in Fig.~\ref{fig:results-heatmap}).”

Although we have now added additional hindcast results (Figs 4 and 5) showing the optimal parameters when calibrating using the modified melt parameterization and a static geometry, it is not feasible to extend the corresponding model simulation to 2100 for additional forecast analysis. However, the hindcast results demonstrate that the most appropriate set of parameters for century-scale projections are those obtained with the transient-coupled calibration, which produces the best overall agreement with observations.

Specific comments:

Figure 1b: At first, it was difficult for me to recognize that the white area is ocean (presumably because I associate white with ice). Maybe it would be more intuitive to chose a different face color here.

The figure has been changed so that the ice domain is in white, and the ocean area is in light blue, in the separate panels.

Figure 2: The transition zone is a bit difficult to distinguish from the colormap. Maybe use a different color for it.

The shading over the transition zone has been removed, and the thickness of the contour reduced. An additional figure (Fig. C1) has been added to the appendix showing the melt-rate fields without overlaid contours, with extra subplots showing the transition zone, water column thickness and ice draft.

Equation 1: Where do parameters “Gamma_Turb” and “Gamma_Mole” come from?

The following text will now be included in section “2.2.2 Modified melt-rate parameterization”:

- *“In MITgcm, melt rates beneath the ice shelf are calculated using the ‘three-equation model’ (Holland and Jenkins, 1999; Losch, 2008), which evaluates turbulent fluxes of heat and salt across the ice-ocean boundary layer. In this parameterization, the turbulent exchange velocities can be assumed to be either constant (e.g., $\gamma_T = 1 \times 10^{-4}$ $m\ s^{-1}$ and $\gamma_S = 5.05 \times 10^{-7}$ $m\ s^{-1}$ in Hellmer and Olbers (1989); Losch (2008)), or computed using a velocity-dependent formulation (Holland and Jenkins, 1999; Dansereau et al., 2014),*

  $$\gamma_{T,S} = u_* \Gamma_{T,S} \quad (1)$$

  where u_* is the friction velocity in the boundary layer beneath the ice shelf, defined by

  $$u_*{}^2 = C_d U_M^2 \quad (2)$$

  with U_M the mixed-layer velocity and C_d a dimensionless drag coefficient, and $\Gamma_{T,S}$ are dimensionless transfer coefficients given by

  $$\Gamma_{T,S} = \frac{1}{\Gamma_{Turb} + \Gamma_{Mole}^{T,S}}. \quad (3)$$

  In Eq. (3), the turbulent exchange (Γ_{Turb}) and molecular diffusion ($\Gamma_{Mole}^{T,S}$) terms are given by

  $$\Gamma_{Turb} = \frac{1}{k} ln\left(\frac{u_* \xi_N \eta_*^2}{f h_v}\right) + \frac{1}{2\xi_N \eta_*} - \frac{1}{k} \quad (4)$$

  and

  $$\Gamma_{Mole}^{T,S} = 12.5(Pr, Sc)^{2/3} - 6, \quad (5)$$

  where h_v is the thickness of the viscous sublayer calculated as $h_v = 5\nu/u_*$, k is the von Kármán constant, ξ_N is a stability constant, η_* is a stability parameter, and f is the Coriolis parameter. Pr and Sc are the molecular Prandtl and Schmidt numbers, defined as the ratios of kinematic viscosity to thermal and salinity diffusivity, respectively. Values of all parameters above are given in Holland and Jenkins (1999), unless otherwise stated.”

L105: “Ice fronts were fixed”: This is also true for the future runs, I assume? If this is the case, how could this affect future sea level projections?

Now included in the limitations section 5.2:

- *“A further limitation is that ice fronts were fixed throughout all hindcast and forecast simulations. While calving and damage can reduce ice-shelf buttressing, potentially leading to greater mass loss* *in the Amundsen Sea Embayment \citep{joughin2021-Iceshelf,sun2023-Speedup}, fixed ice fronts are a common assumption in ice-sheet projections in the absence of a universal calving law. This means that our projections likely represent a conservative estimate of future mass loss, as the loss of buttressing through calving is not accounted for. Incorporating evolving ice fronts is not yet possible in the current coupled ice sheet-ocean model configuration, and quantifying their effect on century-scale projections is an important direction for future modelling studies."*

L280: Why two different reanalysis data sets? Could there be any inconsistencies arising at the domain boundaries?

These two datasets, used in the control case only, serve complementary roles. WOA18 is a climatological product based on gridded in-situ observations and is used for temperature and salinity at the open boundaries, as it provides reliable, observation-based stratification. B-SOSE, is a data assimilation product (state estimate) and provides additional variables required by the model (e.g., velocities, sea surface height) that are not available in WOA18. In regions where B-SOSE is poorly constrained by observations, its salinity fields can be unreliable, and tests during model development showed that stratification on the eastern boundary had a significant impact on the variability in the Amundsen Sea; WOA18 was therefore preferred for T/S specifically. Any inconsistencies arising at the boundary between the two datasets are mitigated by a sponge layer, which relaxes the boundary conditions smoothly towards the modelled interior values.

L330: "we adopt a spatially-uniform error". Why 5 m yr-1? Please justify.

This should have said 3m/yr, and it is determined by the spatial average of melt errors from the only (to our knowledge) Antarctic-wide dataset for basal melt rate errors derived from physically based uncertainty propagation (Paolo et al 2023). This detail has now been included:

- *"To avoid this, we adopt a spatially uniform error of 3~m~yr^{-1} for the melt dataset, computed as the spatial average of the uncertainty estimates from \citet{paolo2023-Widespread}, who to our knowledge provide the only Antarctic-wide dataset of melt-rate errors derived from physically based error propagation. Whilst a spatially uniform error does not capture regional variability in melt-rate uncertainty, it acts in practice as a weighting on the melt-rate metric relative to the ice-speed and thickness change metrics; decreasing this value produces a ranking closer to a melt-only calibration, whilst increasing it gives greater influence to the ice-dynamics metrics (not shown). Until more reliable and spatially varying melt-rate uncertainty estimates become available, this remains the most practical option."*

L453: "which we keep fix in our simulations". Please clarify that it is fix only in your hindcast simulations.

The surface mass balance in the ice sheet model is a spatially varying but time-invariant snow accumulation rate from RACMO2.3 climatology for the period 1979-2013, and is prescribed in all hindcast and forecast simulations. We will amend the wording in the results section and add a note in the limitations section.

Results section 3.3:

- *"This is compounded by observational uncertainties of approximately 1~mm in cumulative sea-level contribution over the hindcast period, which are large enough that all simulations fall within the observational bounds."*

Limitations section 5.2:

- *"Surface mass balance (SMB) was represented by a spatially-varying but time-invariant field throughout all hindcast and forecast simulations. Although this does not affect the relative comparison between calibration approaches, since SMB is identical across all simulations, it may influence the absolute magnitude of projected mass loss \citep{donat-magnin2021-Future}. In future work, time-varying SMB forcing should therefore be incorporated to better constrain projections."*