

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: Xylar Asay-Davis

I wish my name to be relayed to the authors, as I feel I am always a better reviewer when I am not anonymous and I encourage others to consider reviewing non-anonymously whenever they feel able.

General Comments:

This manuscript presents a new calibration strategy for a coupled ice sheet–ocean model of the Amundsen Sea sector, using spatial observations of basal melt rates together with observed changes in ice speed and thickness over 2013–2017. The authors' stated aims are to produce the first historically calibrated coupled ice sheet–ocean model using both oceanic and glaciological observations, to provide evidence for the importance of accurately simulating basal melt near grounding lines, and to assess the implications of this calibration for projections through 2100. To do so, they introduce a modified melt parameterization, run an ensemble of 68 coupled hindcast simulations spanning a range of melt-parameter values, and compare a preferred “transient-coupled” calibration against a more standard “static-melt” approach. Their main findings are that the highest-scoring hindcast simulations are those that enhance melt near deep grounding lines, that including ice-dynamical observations leads to a much better reproduction of observed thinning and speed changes over the hindcast period, and that the preferred transient-coupled calibration yields substantially larger projected sea-level contribution by 2100 than the static-melt calibration, including a 14 mm larger contribution under RCP8.5.

The manuscript has a number of important strengths. First, it makes a forceful and, to my mind, persuasive case that observed ice-sheet thickness and velocity changes should be treated as key constraints on basal-melt calibration, rather than relying on melt-rate observations alone. The paper is especially effective in showing that calibrations based only on present-day melt estimates can miss the observed dynamical response of the ice sheet, whereas including ice-dynamical constraints materially improves agreement with the historical thinning and acceleration patterns. More broadly, I think the authors are right to emphasize that this type of multi-observational calibration remains underused in the wider community, despite the fact that coupled systems are meant to reproduce both the ocean forcing and the ice-sheet response to that forcing.

Second, the manuscript makes a strong case that the combination of an improved melt parameterization and a transient calibration strategy can greatly increase the model's ability to reproduce observations over a short hindcast relative to more standard techniques. The comparison with the static-melt/default-parameterization approach is compelling in showing that the standard setup tends to favor lower grounding-line melt, leading to poorer reproduction of the observed dynamical changes and to coupling shock. Relatedly, I was impressed by the rigor of the parameter-space exploration: running 68 coupled hindcast simulations for a model of this complexity is a substantial undertaking, and it gives the calibration exercise much more credibility than a lightly sampled tuning exercise would have. I also think the paper makes an important practical contribution by showing that, at least in the Amundsen Sea sector, a roughly 5-year calibration window is already long enough to extract meaningful constraints from observed thickness and velocity change. That is an important result in its own right, because it suggests that transient coupled calibrations of this sort may be computationally demanding but still tractable for broader use.

Finally, I found the paper to be very well written and well organized. The structure is clear, the motivation is easy to follow, and the figures and experiment design do a good job of guiding the reader through what is

otherwise a fairly complex modeling framework. I also appreciated the discussion of limitations and future work. In particular, the manuscript does not present the new approach as an endpoint, but instead lays out several intriguing next steps: developing a more physically grounded treatment of grounding-line melt processes, moving toward joint calibration of ocean and ice-sheet parameters, testing alternative hindcast windows as new observations become available, and extending the calibration framework to additional forcings and regions. That forward-looking discussion makes the paper useful not only as a study in its own right, but also as a contribution that could help shape future work in this area more broadly.

My main concern is that the manuscript's central interpretation is not fully supported by the comparisons currently shown. The paper repeatedly emphasizes the importance of transient calibration of the coupled ice-sheet–ocean system using observations that constrain both components, but the main comparison between the preferred “transient-coupled” approach and the “standard” approach changes more than just the calibration strategy. As defined in the manuscript, the transient-coupled configuration combines three elements at once: a transient geometry, multiple observational constraints, and the new modified melt parameterization. By contrast, the static-melt comparison uses static geometry, melt observations alone, and the default melt parameterization. Because these factors are changed simultaneously, the relative importance of transient coupled calibration versus the new melt parameterization is not isolated by the results currently shown. This matters because the manuscript also shows that the highest-scoring hindcast simulations all use the modified parameterization, whereas the default-parameterization cases are among the poorest performers, and the low-scoring cases are explicitly linked to insufficient melt near the grounding line and the resulting coupling shock. In other words, the evidence presented seems at least equally consistent with the interpretation that much of the improvement arises from the added flexibility of the new parameterization, especially its ability to enhance grounding-line melt, rather than from the transient coupled calibration alone.

For that reason, I think the paper would be substantially stronger if the authors added simulations that explicitly disentangle these effects. Most importantly, I would like to see a calibration using the new melt parameterization but retaining the more standard setup of static ice-shelf geometry and melt-rate-only constraints. That experiment would directly test whether the modified parameterization alone can recover much of the apparent improvement now attributed to the transient coupled calibration.

Relatedly, the manuscript already shows that when the calibration is restricted to melt observations alone but transient geometry is retained, the preferred solutions shift within the family of modified-parameterization cases, and Fig. B3 appears to suggest notably better performance for those runs than for the default-parameterization cases. That makes the missing control case especially important, because it raises the real possibility that a substantial fraction of the skill improvement could also be obtained without the transient coupled calibration, provided the new parameterization is used. I am typically quite hesitant to recommend new simulations as part of a review: You designed the experiments, not me, after all. But, if these additional simulations are not feasible, then I think the authors must at minimum soften the repeated claims about the specific importance of transient coupled calibration and more clearly state that the present study demonstrates the benefit of a bundled methodological package, not of transient calibration in isolation. This may substantially change the emphasis and character of the paper.

As summarized by the reviewer, our transient-coupled calibration framework differs from commonly used approaches for basal-melt parameter calibration in three key respects: the modified melt parameterization, the inclusion of evolving cavity geometry, and validation against ice-dynamics observations. We agree that introducing these elements simultaneously 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 melt rates for a static ice geometry. To make this comparison fair, we identified the highest-scoring simulation that uses the modified melt-rate parameterization (static-melt-mod), as well as the most common approach which uses the 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}).”

A separate concern is that the manuscript may understate the extent to which the calibrated melt parameters could be compensating for structural deficiencies in the ice-sheet model, rather than uniquely identifying the physically correct basal melt field. The paper notes that damage and calving are not represented and also acknowledges that only the ocean-side melt parameters are calibrated here, while a single set of ice-sheet parameters is held fixed. The limitations section further states that, in a fully calibrated ice-sheet model, calibrating against either ice-speed change or thickness change should converge on the same optimal basal-melt parameters, whereas in the present results these metrics favor different optima precisely because the ice-sheet model parameters have not themselves been historically calibrated. To me, this is an important caveat that extends beyond the limitations section. If missing or mis-specified ice-sheet physics affect the modeled dynamical response, then the melt calibration may partly be absorbing those errors. That possibility seems especially relevant here because the preferred parameter choices are those that increase melt near grounding lines, where the coupled system is most sensitive and where omitted processes such as damage, calving, or other ice-dynamical biases could plausibly matter. I therefore encourage the authors to discuss more explicitly throughout the manuscript that the optimized basal-melt field may in part be compensating for biases elsewhere in the coupled system, and that a future joint calibration of both ocean and ice-sheet parameters is needed before the inferred melt parameters can be interpreted as uniquely constrained.

This is indeed an important point, and we will stress this 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 \cite{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.”

As a result of these concerns, particularly about better acknowledging the fact that the current simulations do not untangle the relative importance of the new melt parameterization compared with the transient calibration, I feel I must recommend major revisions. My more specific comments are detailed below, with a few formatting suggestions under “technical corrections.”

Specific Comments:

I. 122-126: Thanks for detailing these changes! These problems in BedMachine (not fixed in v4 to my knowledge) make the dataset hard for coupled ice sheet-ocean modeling. It’s really helpful that you point out the problems and give other groups a sense of how they can be mitigated.

Thank you, we thought it was important to be honest about how we approached this problem and to put this detail upfront to assist other modellers.

I. 190 Eq. (1): Could you say what Γ_{Turb} and $\Gamma_{T,S}^{Mole}$ are? Also, where is this choice of notation from? McPhee et al. (1987) uses Φ rather than Γ as far as I can tell and I am used to $\Gamma_{T,S}$ being a nondimensional heat or salt transfer coefficient in the numerator, rather than the denominator so that’s leading to a bit of confusion with Γ terms in the denominator here for me.

The notation we use is from Holland and Jenkins 1999, so the McPhee citation has now been removed to avoid confusion. 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., $\gamma_T = 1 \times 10^{-4} \text{ m s}^{-1}$ and $\gamma_S = 5.05 \times 10^{-7} \text{ m s}^{-1}$ 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.”

I. 225: “As opposed to the default melt parameterization, with C_d as its only tunable parameter...” I wonder why Γ_{Turb} is considered a fixed parameter and C_d is tunable. I think it is not unusual in ice-shelf-cavity modeling to consider both to be tunable.

Added to section “2.2.2 Modified melt-rate parameterization”:

- *“In the default melt parameterization (Eq. ~\ref{eq:smGammaTS}), C_d and Γ_T are generally considered to be tunable parameters; however, observations constrain only the product $\sqrt{C_d}\Gamma_T$, rather than the individual parameters (citep{jenkins2010-Observation}). Given that the problem is therefore underdetermined, fixing Γ_T and tuning C_d is a practical and physically justifiable way to reduce the parameter space. Indeed, previously published sensitivity experiments show that while melt rates increase with Γ_T for a given C_d , the depth distribution of melt is relatively insensitive to the specific combination of C_d and Γ_T , provided the total melt rate is similar (see Fig. ~6 in) [jourdain2017-Ocean].”*

Sec. 2.4.2: This is a section where I feel you could go further in acknowledging that the current approach runs the risk of tuning melt parameters to compensate for ice-sheet errors, e.g. inaccurate initialization, lack of calving or damage.

This has now been addressed with added text (See “Sec 2.4.2” above)

I. 330: “To avoid this, we adopt a spatially-uniform error of 5 m yr⁻¹ for the melt dataset.” Could you provide a justification for choosing this value?

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 (cite{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.”*

I. 387-389: “Some of this bias is likely attributable to ice damage and calving processes (Lhermitte et al., 2020; Joughin et al., 2021), which are not represented in the ice-sheet model.” This seems like a good opportunity to acknowledge that you are calibrating the melt model to help compensate for this missing process, possibly at the expense of more accurately representing the basal melt field itself.

We have added the following discussion to “Sec 3.1.1 Highest scoring parameters”:

- *“We note that the calibrated parameter values may in part reflect compensation for missing or poorly constrained ice dynamics in the uncalibrated ice-sheet model; nevertheless, the values and resulting melt rate patterns remain physically plausible and provide valuable constraints on the melt parameters within this coupled system. The parameter set provides the best overall agreement with the available observations and is therefore selected for the transient-coupled calibration (TR), and subsequently compared with a melt-only calibration in the forecast simulations (Sect. ~\ref{sec:forecast}). Further discussion of the limitations is provided in Sect. ~\ref{sec:disc_ice}.”*

I. 407-411: “When using only melt-rate observations in the calibration (while maintaining a transient ice-shelf geometry), the highest-scoring simulations shift towards those with velocity-independent melt over a wider region but lower γ_{T0} values (i.e., reduced turbulent exchange velocities and therefore lower melt rates) (Fig. B3). However, these same simulations are not the highest scoring when evaluated against changes in ice speed (Fig. B2).” I feel like this deserves more discussion. Here is the evidence that the new melt parameterization is lifting a lot of the weight compared with the “standard” calibration approach. It also feels like more should be said about why B2 and B3 may differ, and the likelihood that improvement in B2 come partly for the “right reasons” (better ice sheet response directly to melt) and partly for the “wrong reasons” (e.g. basal melting compensating for lack of damage and caving or biases in ice-sheet thickness).

The reviewer is right to suggest more discussion is needed about why the melt-only and velocity-only calibration approaches yield different optimal parameters. We have now amended this paragraph as follows:

- *“The results presented above are based on calibration using a combination of three different observational datasets. However, the analysis can be repeated using fewer observational constraints and different modelling approaches. For example, when using only melt-rate observations in the calibration (while maintaining a transient ice-shelf geometry), the highest-scoring simulations shift towards those with velocity-independent melt over a wider region but lower γ_{T0} values (i.e., reduced turbulent exchange velocities and therefore lower melt rates) (Fig. ~\ref{fig:app-heatmap-meltG}). Yet, these same simulations are not the highest scoring when evaluated against changes in ice speed (Fig. ~\ref{fig:app-heatmap-dv}). **The fact that different parameter combinations are produced for melt-only and velocity-only calibrations suggests that the two observational datasets are not fully compatible given the current model physics and measurement uncertainties. As noted in Sec. ~\ref{sec:methods-exp-ensemble}, we only use a single realization of an uncalibrated ice-sheet model, which might mean that poorly constrained processes such as ice damage, ice rheology and basal sliding, likely contributes to a bias in the optimised melt parameters. Without a more complete calibration that includes ice-sheet model parameters, we cannot conclusively disentangle these contributions. This is discussed further in Sec. ~\ref{sec:disc_ice}.***

Setting aside these limitations, we now 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).”

I 416-418: “Nevertheless, because this represents the top-ranked parameter set following the most common calibration approach in the literature, we extend this simulation to 2100 and compare it with the transient-coupled calibration in Sect. 4.” Again, how do you then know how much transient-coupled calibration really changes things, as opposed to the new parameterization?

As mentioned above, in the revised manuscript, we will additionally include results for the **modified** melt parameterization with a **static** ice-sheet geometry (static-melt-mod; SM). The main conclusion from the additional results is that the modified 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.

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.

I. 429-231: “Typically, tuning is done using depth-binned or spatially-integrated values of basal melt for a static ice-sheet geometry. This approach has two key limitations. First, spatial-integration does not provide information about the distribution of melt, which we show matters for reproducing ice dynamics.” I am not aware of depth-binning being a common practice in calibrating 3-equation parameterizations in ocean models with ice-shelf cavities. While some such models may use spatially averaged melt rates for calibration, I am also not aware of that being a standard method of calibration and it is not what is done in my own group – we compare the full melt field with an “observed” melt climatology wherever both have valid data. I am more aware of these types of metrics being used for analysis rather than calibration, and in parameterizations of basal melting in standalone ice-sheet models. I would suggest taking out or rewording your first point here.

Reworded:

- *“Tuning melt-rate parameters is a critical step in predicting future sea-level contribution, particularly when the ocean model is coupled to an ice-sheet model, where melt rates are a key driver of ice dynamics and grounding-line evolution. While previous studies have calibrated melt parameters or assessed model performance by comparing spatially-averaged or depth-binned basal melt rates against observations for a static ice-shelf geometry (dansereau2014-Simulation, jourdain2017-Ocean, bett2024-Coupled), our dynamically evolving hindcast simulations demonstrate that this approach can yield suboptimal parameter values when evaluated against a broader set of observations that include ice dynamics.”*

I. 434-435: “This is evident in our low-scoring coupled model simulations, where low grounding-line melt rates cause upstream thickening and slowdown, in contrast to observations (Figs. 4–5).” I know I keep harping on the same points but this feels like yet another place to acknowledge that melt tuning may be compensating ice sheet model biases.

Added detail:

- *“This is evident in our low-scoring coupled model simulations, where low grounding-line melt rates cause upstream thickening and slowdown, in contrast to observations (Figs. ~\ref{fig:results-heatmap}--\ref{fig:results-spatial}), though we note that the uncalibrated ice-sheet model may also contribute to this response.”*

I. 436-437: “The modified melt-rate parameterization compensates for low vertical and horizontal resolution in the ocean model by generating enhanced melt near the grounding line...” Given the ad hoc nature of the parameterization, it might be best to be broader about what it might be compensating for in the ocean model. My feeling is that the boundary-layer physics under ice shelves is pretty far from being a solved problem or one what a single accepted parameterization at basically any practical resolution that ocean models with ice-shelf cavities can achieve. So the new parameterization is likely compensating for a variety of factors including low horizontal and vertical resolution as well as missing physical processes.

Detail has been added to the methods section and the results:

- Methods Sec 2.2.2 Modified melt-rate parameterization:
 - *“Although this is the commonly used approach for representing ice-shelf basal melting in ocean models, the physics of the ice-ocean boundary layer is not yet fully understood, and the parameterization likely omits processes that remain unresolved or poorly constrained. Furthermore, in z -coordinate ocean models, such as MITgcm, the combination of coarse horizontal and vertical resolution can introduce multiple issues in simulating basal melt.”*
- Results Sec 3.3 Implications of different calibration approaches:
 - *“The modified melt-rate parameterization likely compensates for a combination of unresolved physical processes and the coarse vertical and horizontal resolution of the ocean model near the grounding line. By allowing enhanced melt in thin water-column regions, it produces melt distributions that better match observed melt rates and capture a wider range of ice-dynamic responses, enabling more robust parameter selection against observational constraints.”*

Although no model or parameterization will perfectly represent all physical processes, introducing stronger and time-dependent observational constraints, as in our transient coupled calibration, is a pragmatic pathway to improved sea-level projections. Indeed, it is the addition of transient ice dynamics in the calibration that produces the elevated grounding-line melt rates needed to sustain changes in ice thickness and speed consistent with observations.”

I. 450-458: I really appreciate this paragraph! It really drives home the value in the methods you are proposing here. Intuitively, I tend to feel like ice sheet models respond slowly to ocean forcing so it's valuable to have it made starkly clear how clear changes in ice sheet thickness and velocity can be on these short time scales.

Thank you.

Fig. 6: I find the contents of the insets hard to see without zooming in an unreasonable amount of. Would it make sense to make this into two figures, or break the insets out into their own panels?

This has been changed to have the zoomed hindcast in a separate panel.

Sec. 5: To me, this is a really important and valuable section. You are showing with humility a lot of the important caveats of this work and where it can be taken in the future. You show that it has been far from a trivial exercise but it's not outside the realm of possibility for other groups to do something similar. You point out the very important next step of calibrating the melt parameterization and the ice sheet together!

Thank you.

I. 629-630: “Importantly, by calibrating the ice-sheet model using ice-dynamics observations, the calibration of the ocean melt parameters would only require one type of these datasets along with melt observations.” I didn't follow this last bit. Why would basal melt now only be calibrated with “one type of these datasets”? Do you mean it would be calibrated only with ice-sheet velocity or only with ice thickness changes? If so, why?

Once the ice sheet model is calibrated, the geometry and thickness changes are related through ice physics (any change in geometry uniquely defines a corresponding change in velocity, irrespective of the history of changes). Therefore, the observational datasets for these two metrics become dependent, and they should either be treated as such through an appropriate covariance structure in the likelihood, or only one set of measurements should be used. Because I have mentioned this in the previous paragraph, I will remove this sentence and merge the paragraphs. The new text is as follows.

- *“This matters because changes in ice speed and ice thickness are physically coupled through the ice-flow equations. In a fully calibrated ice-sheet model, accurately reproducing observed changes in ice-sheet geometry will yield the corresponding observed changes in ice velocity. In particular, this means that calibrating basal melt parameters against either geometry change or velocity change alone would produce the same optimal values, making only one of these observational metrics necessary as an independent observational constraint. However, in our current results, calibrations based on individual ice-dynamics metrics yield different likelihood scores and different optimal melt parameters (Figs. ~\ref{fig:app-heatmap-dh}--\ref{fig:app-heatmap-dv}). This discrepancy may indicate that the assumed ice-sheet model parameter set does not fully reproduce the observed coupling between changes in geometry and velocity, likely because the ice-sheet model parameters themselves have not yet been comprehensively calibrated. Alternatively, temporal mismatches between datasets and observational uncertainties alone may be sufficient to explain the discrepancy. “*

Sec. 5.3: This section is a useful acknowledgement of the limited scope of the calibration. It is not really phrased as a limitation (or in the context of future work) so it was not clear to me if it belongs in Sec. 5 or could be moved elsewhere.

Now reworded so it fits more appropriately here:

- *“The likelihood scores were calculated across the three major basins of the Amundsen Sea Embayment (Pine Island, Thwaites, and Pope-Smith-Kohler), rather than the full domain. These basins contain the fastest-flowing glaciers in West Antarctica (Fig.~\ref{fig:methods-domain}a), and are experiencing the greatest mass loss \citep{smith2020-Pervasive}. This means that observational data, model validation and prior ice-sheet model initialization have been focused on these regions \citep{naughten2022-Simulated,reed2024-Melt,derydt2024-Geometric}. This focus in our current work was therefore a practical necessity given the available data and model configuration, and ensured a balance between reproducing observed mass loss across the embayment while avoiding overfitting to individual glaciers, which have their own sensitivities and optimal parameter values (Figs.~\ref{fig:app-heatmap-meltG}--\ref{fig:app-heatmap-all}). Nevertheless, extending the calibration to the full domain, which would require additional observational constraints and broader model validation, represents an avenue for future work, and would provide a more complete picture of ice-ocean interactions across the Amundsen Sea.”*

I. 654-357: “Despite the limitations, our transient-coupled calibration approach represents a significant advance in capturing the complex interactions between ice dynamics and ocean melting. Our results demonstrate that calibrations using present-day melt estimates alone systematically underestimate ice-sheet response over the historical period. By incorporating ice dynamics constraints, this bias is removed entirely, resulting in more robust projections of future mass loss.” This paragraph brings up many of the concerns I have mentioned in my general comments. I do not believe your simulation results show that transient-coupled calibration is the significant advance here, as opposed to a new melt parameterization. Your results “demonstrate that calibrations using [only] present-day melt estimates...” and MITgcm’s existing parameterization of sub-ice-shelf melt “...systematically underestimate ice-sheet response over the historical period.” But you haven’t shown that transient-coupled calibration is the reason that you get better results when you use transient-coupled calibration and a new parameterization.

Reworded:

- *“Despite the limitations, the transient-coupled calibration approach represents an important step forward in capturing the complex interactions between ice dynamics and ocean melting. Our results demonstrate that calibrations relying solely on observed melt-rate estimates and the default sub-shelf melt parameterization systematically underestimate ice-sheet response over the historical period. By incorporating a modified parameterization that better captures grounding-line melt patterns, and constraining the model with transient ice-dynamics, we obtain improved agreement with observations over the hindcast period. This provides a more physically consistent basis for projections of future mass loss from the Amundsen Sea Embayment.”*

I. 660-663: “Whereas previous calibrations have relied on melt estimates only and assumed static ice-shelf geometries, we have shown that including feedbacks between basal melting, evolving ice thickness, and cavity circulation in the calibration leads to a greatly improved representation of ice-sheet dynamics over the observational period.”

We are not sure if the reviewer intended to add something extra here or if this is related to the next comment. However, if it is related to the next comment, then we have now included an additional calibration result, to aid our conclusions, and acknowledged the role that the modified melt parameterization is having.

Sec. 6 Conclusion: I found the conclusion more measured than some of the broader framing in the manuscript. In particular, the conclusion does not appear to claim that transient-coupled calibration alone is responsible for the improved reproduction of observed ice-sheet change; rather, it presents the results as demonstrating the value of a multi-observational, transient coupled approach using the modified parameterization. However, the conclusion would still benefit from a more explicit acknowledgment that some of the inferred melt-parameter changes may be compensating for unresolved or uncalibrated ice-sheet processes, a possibility that is recognized in the limitations section but not carried through to the final synthesis.

Updated in the conclusion:

- *“To calibrate ocean-model melt parameters, we ran an ensemble of coupled simulations with varying parameter values in a modified melt parameterization. We compared modelled outputs*

against observed spatial patterns of ice-shelf melt rates and changes in ice speed and thickness between 2013 and 2017. With this approach, parameter combinations that enhanced melt in the vicinity of the grounding line produced the best match with observations of thinning and accelerating grounded ice while also constraining overall melt rates. While the calibrated parameter values may partially compensate for uncalibrated ice-sheet dynamics, they remain physically plausible and represent the best estimate of optimal melt parameters within the current coupled framework. In contrast, the common calibration approach of tuning against only melt observations for a static geometry produced optimal values with low grounding-line melt, leading to coupling shocks and widespread glacier slowdown.”

Technical Corrections:

Fig. 2: I found the pink transition zone and the bold black line for its center to make it very hard for me to compare the melt rates in panels e) and f) with those in a). My preference would be to put that transition zone in its own panel. It seems important to be able to visually see how abrupt or smooth the transition in melting is across the transition zone without the pink shading and black line obscuring it. But an alternative that still might help would be to use another color (green?) that is clearly not so close to the colormap for the transition zone.

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.