Response to the Reviewers

January 4, 2026

Response to Reviewer 1

Summary and general comments

This is a well-organized and well written paper describing the application of physics-informed neural networks (PINNS) for further improving the generation of sub-glacial bed topography datasets, building on previous "BedMachine" efforts that have been ongoing for the past decade or so. Two conservation equations – the continuity equation and a momentum balance equation for ice flow – are considered and introduced as additional constraints on the loss function, akin to their introduction as constraints in the "cost function" of PDE-constrained optimization (a good analogy to consider including for readers more familiar with the language of glaciology modeling / optimization?). The two constraints are considered on their own and in combination and the resulting bed topography datasets are compared and contrasted with previous BedMachine results for three glaciologically distinct regions. Overall, the authors argue convincingly that the new approach has merit and demonstrates potential for improving inferred bed topography in regions where the traditional BedMachine approach begins to break down. Overall, the work seems very worthy of publication and readers of The Cryosphere will find it a worthy contribution. My suggestion would be to accept for publication with minor revisions, noting that these are those suggested revisions (detailed below and identified by their line number in the submitted version) are largely editorial in nature.
Response: We appreciate the reviewer's thorough and thoughtful comments. Thank you for your encouraging feedback and for emphasizing the overall quality and relevance of our work.

My one more substantial suggestion – not necessarily for this publication but possibly for a future effort – is that I think it would be very useful to redo this exercise for a single region (I realize the computational cost could be a challenge, so pick a single region, like the most challenging one discussed here) but using L1L2 (Blatter/Pattyn) for the momentum balance model as opposed to SSA (which you've already done and could reuse as the baseline). Because the former allows for internal deformation, a possibly very different vertical velocity profile (and hence depth-averaged velocity and flux divergence) might be implied in regions of slower moving ice (noting that the modeled 2d surface velocity field could still be used as the velocity constraint, so that there should not necessarily be any significant reformulation of the loss functions discussed herein). It would be interesting to see if the more accurate stress balance constraint helped to alleviate any of the remaining problems discussed below.
Response: We appreciate the reviewer's feedback about using L1L2 (Blatter/Pattyn) for the momentum balance model to account for vertical shear. While this is currently outside the scope of this paper, we will include this in our discussion of future work.

Detailed comments

21: Are these refs now considered the definitive source for defining the amount of potential SLR locked up in the ice sheets? If not, maybe consider adding one or two more from other authors for the sake of diversity?
Response: We believe these references: IMBIE Mass Balance Intercomparison (Otosaka et al., 2023) provides the latest consensus on the current contribution to sea level rise and BedMachine (Morlighem et al., 2017; Morlighem, 2020) and Bedmap 3 (Pritchard et al., 2025) are the most up to date estimates for defining

the amount of potential sea level rise. We will include these in the manuscript.

23: "numerical ice sheet modeling"
Response: Thank you for the suggestion, we will make this change in the manuscript.

25: It seems like a summary-level reference might also be appropriate here (?), e.g. something from one of the recent IPCC reports (that integrates results from a large number of individual publications).
Response: Thank you for the suggestion. In addition to Aschwanden et al. (2021), we will also add the ISMIP6 papers as reference for multi-model ensembles.

28-29: It should probably be noted here that these experiments were assuming a marine ice sheet with a significant over-deepening inland (since this configuration would necessarily be more sensitive than say, an ice sheet grounded above sea level).
Response: We will include a note that these experiments were assuming a marine ice sheet with a significant over-deepening inland, and therefore likely more sensitive.

37-38: Is the ∼2 km limit proposed here coming from the Durand et al. (2011) paper? While I more-or-less agree with this idea, I don't know that this single reference is adequate to support the precision implied by this statement. Maybe consider softening it a little bit to something less precise, e.g. "order km-scale spatial resolution"?
Response: The sentence will be modified to state "kilometer-scale spatial resolutions" required by ice sheet models instead.

63: "three regions in Greenland"; maybe add a few words of clarification here that they are glaciologically distinct / different? E.g., presumably you mean regions where velocity occurs primarily via fast sliding, a region where it occurs via a mix of sliding and deformation, etc.
Response: The sentence will be modified to state we have chosen "three glaciologically distinct" regions.

Figure 1 caption: "The loss function is comprised of . . . " or "The loss function includes data loss . . . "
Response: Caption will be changed to state "the loss function is comprised of...".

89: "fully connected layers", maybe use "fully connected ('dense') layers . . . " ?
Response: Thank you for the suggestion, we will make this change in the manuscript.

101: Should it be "the apparent mass balance residual" ?
Response: The "mass balance residual" in this case refers to the residual of the conservation of mass. The "apparent mass balance" is defined as $\dot{a} = \dot{M}_s - \dot{M}_b - \partial H/\partial t$. To minimize confusion, we will remove the term "apparent mass balance" from the manuscript and re-phrase our explanation.

103-105: I am guessing that maybe this is discussed further below (?), but it seems like you are already potentially limiting the usefulness of this approach by restricting the momentum balance to SSA. I.e., if one of the main interests here is in improving the inference in regions of slower moving ice flow, which is presumably due to less sliding and more internal deformation, then SSA doesn't seem like the right assumption to make for the model dynamics. I know that ISSM has higher-order approximations available (e.g., L1L2 or "Blatter-Pattyn"). Has that also been explored (acknowledging the obvious additional computational burden) and compared against the approach using SSA?
Response: We decided to start with a simpler approximation of the momentum balance (i.e., SSA) to minimze computational cost and to see if this method of using two conservation laws could produce sensible results. Further, since we wanted our method to be comparable to the method in BedMachine, we needed to use the depth-averaged conservation of mass and so also needed to use a 2D approximation of the momentum balance. That being said, for the interior, slower-moving regions of the ice sheet we hope to use higher-order approximations of the momentum balance, and we will include this as an additional point in our discussion section.

116: "... from THE regional climate model RACMO..."
Response: Thank you for the suggestion, we will make this change in the manuscript.

Section 2.3: It sounds like the basal mass balance term in equation 2 is assumed to be 0? If so, it would be good to note that explicitly here in the discussion of the apparent mass balance term.
Response: We will make this change and explicitly mention that the basal mass balance term is assumed to be negligible.

152: "to prevent from taking" (omit "from"?); "or diving by zero" ("dividing by zero")
Response: Thank you for the suggestion, we will make this change in the manuscript.

176-177: Would it be worth commenting on the choice of median vs. mean? Is the median chosen because of the small number (5) of samples, such that the mean could be easily biased?
Response: Yes, we chose to use the median because since we are using an ensemble of five models for each region, and the mean could easily be biased. We will mention this explicitly in the methods section.

180: By "challenging to implement", do you mean where the traditional / previous mass conservation approach does not perform well? Implementation sounds more like the approach is challenging, but I imagine the approach is just as easy to implement in these regions, it's more the prior / baseline result that you are not happy with.
Response: By "challenging to implement", we do mean that the mass conservation inverse approach does not perform well since the ice is moving slowly. We will make this more explicit within the manuscript.

242: It's not clear here exactly what "Fig.2(1)" is referring to.
Response: Fig. 2(l) refers to the ice thickness values along ice-penetrating radar flight tracks for the Deception region. To minimize confusion, the sentence will be modified to read "ice-penetrating radar flight tracks in Fig. 2(l)".

Figure 5: In the caption for this and figure 4 it would be helpful to remind the reader which dataset is subtracted from which and shown in panels d-f (e.g., PINN minus original BedMachine product or vice versa?).
Response: We will definitely make this more explicit in the Fig. 4 and Fig. 5 captions. We will state that the difference is $\hat{\boldsymbol{b}}_{\text{BedMachine}} - \hat{\boldsymbol{b}}$ for panels d-f.

Table 4: It might also be useful to provide some percent / fractional metrics here? E.g., for the apparent mass balance RMSE, how does that number compare to the average apparent mass balance over the same area? Such a table could be added to the SI if it's not deemed important enough for the main text.
Response: We will include an additional column in Table 4 for the fractional/percent metrics.

3.2.2. – It's left hanging a bit as to the significance of the differences in u, apparent mass balance, and sfc. elevation when using the different approaches. For example, how do these differences compare to those that arise when using the original BedMachine approach? Would it make sense to include those metrics (differences in u, apparent mass balance, and sfc elevation) somewhere here for comparison? It's a bit unclear to me what the broader implications are of these secondary metrics w.r.t. using the derived datasets for modeling. If the authors have additional thoughts on this they would be welcome in the supplementary information.
Response: For the mass-conserving approach in BedMachine, the raw data (i.e., $\boldsymbol{v}_{\text{data}}, \dot{a}_{\text{data}}$) are used within the mass balance equation to directly invert for the ice thickness. With the PINN approach, in order to infer the ice thickness, the PINN must predict the $\boldsymbol{v}, \dot{a}, s, C$ fields that are used within the mass balance and momentum balance residuals. Therefore, the differences in these derived datasets arise due to the PINN architecture – while the PINN can minimize the error in these predicted fields with respect to the ground-truth data, it will never be exactly the same. We tried to explain this in the discussion section as BedMachine imposing mass conservation more 'strongly' and the PINN imposing the conservation laws more 'weakly', however we will develop this point further within our discussion to avoid confusion.

326: If the discussion starting in 4.1 is intended to be specific to Deception, then maybe that should be noted earlier in this paragraph? Alternatively, if the discussion in 4.1 up to line 326 where Deception is mentioned is supposed to be generic, then perhaps line 326 should be something more like, "… a far more realistic bed topography map, particularly for Deception."
Response: The discussion is meant to be specific to the Deception region, we will make this more clear in the manuscript.

329-330: Would "…slightly higher RMSE SUGGESTS …" be more appropriate here than "indicates"? I think the speculation in this sentence makes sense, but it seems like it is perhaps speculation as opposed to a concrete fact.
Response: Thank you for the suggestion, we will make this change in the manuscript.

336-348: W.r.t. the prediction of thinner ice – is it also possible that this could be the result of the chosen stress balance model? E.g., in order for the SSA model to match surface velocities, it would need to assume a depth-averaged velocity profile that is larger than would be assumed in a model that allowed for internal deformation (E.g., L1L2), because SSA can only accommodate velocity via a change in the sliding component (unless I'm misunderstanding the model used here). If that is indeed the case, then it seems like the optimization process might necessarily bias the ice thickness on the thin side; if the depth averaged velocity is too large, the same flux (constrained by continuity equation and the apparent mass balance terms) can only be accommodated by reducing the ice thickness.
Response: This is a good point, however the region where the PINN predicts thinner ice is in the fast-flowing region where we might expect there to be more sliding rather than internal deformation (which is perhaps more typical in the slower-moving regions). While we will add this as a note to our discussion, we do think that the prediction of thinner ice in these areas is largely because the PINN 'struggles' to predict sharp transitions in the ice velocity.

345: "These reasons imply that …" is a bit awkward. "This implies that …"? "These arguments imply that …" ?
Response: Thank you for the suggestion, we will make this change in the manuscript.

353: "… and exceeds their INDIVIDUAL limitations …" ?
Response: Thank you for the suggestion, we will make this change in the manuscript.

398: "We observe that the PINN better captures…" → "We observe that the PINN captures observable features better with …"
Response: Thank you for the suggestion, we will make this change in the manuscript.

403: "… state variable predictions". (remove plural on "variable")
Response: Thank you for the suggestion, we will make this change in the manuscript.

415: "… mass-conserving approach, AS CONFIRMED BY THE DISCOVERY OF new bed features beneath Narssap and Deception."
Response: Thank you for the suggestion, we will make this change in the manuscript.

417: "… we recommend USING this approach …"
Response: Thank you for the suggestion, we will make this change in the manuscript.

A last thought / general comment: The implied "geomorphology" of the three focus areas studied here look very different from one another. E.g., The Upernavik and Narsaap beds look very smooth when compared to Deception. In the areas where there are no troughs, they almost look like high-resolution DEMs from past, heavily glaciated regions of Canada. Is there any published work on previous Greenland glaciations that might provide some more insight into this? I'm not suggesting it should be part of this paper, but it could be interesting to look into whether or not the "smoothness" that your methods are implying about the bed in different regions is in line with current glacial geological / geomorphological understanding. It would seemingly be a further testament to the power of the methods used here if you were resolving that level of information about the bed through hundreds / thousands of meters of ice.

Response: Thank you for your comment. We agree that the geomorphology of the regions studied in this paper look quite different from each other. When we refer to "smooth" PINN predictions, we are comparing predictions directly with the data along ice penetrating radar flight tracks which we observe is far more "rough". Like BedMachine, the PINN also does capture lower-frequency approximations of the bed, however misses out on the higher-frequency details.