Review of “The transferability of adjoint inversion products between different ice flow models” by Barnes et al. for The Cryosphere

General comments

Ice sheet model initialization is a crucial step to ensure that a model is as close as possible to the current state of the ice sheet. In this paper, the authors propose to infer two poorly known parameters (the basal friction and the stiffness or viscosity of the ice) in three different models and evaluate the differences between these models right after inversion and after a prognostic simulation, when transferring an initial state from one model to another.

The paper is relatively clear and I enjoyed reading it. The problem they set out to address is well-introduced with appropriate references. I found that the methodology and results were detailed well, although some sections and technical choices were harder to follow due to back and forth between the main text and the appendices and sometimes lacking references to the appendices (see specific comments). The authors rightly recognize that the three models they use show significant differences (especially for the rate factor *B*) but that the large-scale distribution agrees well. While this is true for ISSM and STREAMICE, I am concerned by the results of Ua, which show particularly lower misfit between observed and model velocities. In this regard, the authors investigate various possibilities (in the Appendix) for explaining the variations between the models, but I would have liked to see more discussions on the L-Curve analysis. I detail this in a more specific comment but the difference in the misfit between the models brings up some questions about the regularization parameters used in the different models, e.g., are the three initial states really the minimum of each L-curve? Since an inversed state is particularly sensitive to the regularization parameters, it would be interesting to see the L-curve distribution and the location of the initial state picked for each model.

The paper then evaluates the transferability of Ua, ISSM and ICESTREAM initial state in Ua and with a coarser version of the Ua mesh. They conclude that this process is not straightforward and leads to substantial variations, but lies within the range of intercomparison experiments such as initMIP. However, in initMIP, the prognostic models are all different, which means that the differences are not only due to the initial state but also the physics of the transient models themselves (GL parametrization, etc.). Also, the models had various complexities while here, ISSM and Ua both use the SSA. In this regard, it is surprising that STREAMICE (L1L2) and ISSM (SSA) behave closer to each other than Ua (SSA) and ISSM. I think that the differences between the 3 prognostic simulations are relatively high and make it hard to believe that the transferability of initial states is a success. Comparing initial state transferability to the effect of different friction laws on sea level projections (Yu et al., 2018) is also a bit misleading since the latter involves changes in the physics of the model rather than difference in the numerical implementation (especially when Ua and ISSM both use the SSA). I would therefore recommend to temper these conclusions.

In addition, I think that studying the effect of the transferability from initial states to the other two models (Ua and ISSM to STREAMICE, Ua and STREAMICE to ISSM) could greatly benefit the study. This is a substantial effort (and the authors already did a significant number of sensitivity analysis) but it would provide a more comprehensive idea of the real transferability of initial states in the context of multi-model experiments like ISMIP6 (Serrousi et al., 2020), where one initial state could be provided to all the models.

Regardless of my concerns, the paper certainly deserves to be published (after revision) and will be useful to the community. The ability to use a similar initial state in different models for intercomparision experiments or to speed up some fastidious and repetitive initialization phases is of great interest to me. This paper shows the difficulty of the process and the remaining challenges we face in doing so.

Specific comments

- What are the boundary conditions for the different models? I assume that it could have some impact if they are different (especially the calving front). I guess that ISSM and Ua use very similar conditions, although the inside boundary in Ua is at the ice divide (***u***=0) while it is not for ISSM. Also, how is the calving front treated in STREAMICE? Some details about this could be included in the model description. I could only see a reference to the Dirichlet boundary condition at the ice divide (***u***=0) in *Sec. 5.1*. line 325.
- Section 4***:***
    - I am surprised by the relatively high-speed misfit of ISSM in Fig 3. Misfits exceeding 200 or 300 m/yr in most fast-flow regions seems very high. Especially when comparing with the other 2 models but also with other studies (e.g. Brondex et al. (2019), their Figure 4). Why is that so? One possible reason could be that in ISSM, since you also optimize the logarithm of the velocity misfit, you put a lot of weight on slow regions, limiting the optimization of fast flow regions. Regardless of the optimization, the SSA is also not particularly appropriate in these slow regions.
    - The problem you solve is ill-posed by nature, which is a common problem in glaciology but Ua seems particularly under-constrained here, giving a very nice (too nice?) velocity fit but creating a very different field of B (as mentioned by the authors in the Appendix A2).
    - For ISSM, is the inversion chosen here the real minimum of the L-curve? This minimum can be tricky to choose in a 3-D L-curve (I, R_{p1} and R_{p2}).
    - The average misfit on the entire domain is also a bit misleading here since a large part of the domain is slow-flow regions where the absolute misfit will always be small, even with a poor inversion. Could you provide an average misfit for the fast-flow region (with a threshold of, for example, 50 or 100 m/a)?
- Please also specify the prior you use for the friction coefficient. I could not get any info on the friction prior before reading Appendix A4, which is, I think, never mentioned in the main text. Would it be useful to use an approximation based on the driving stress to construct the prior (instead of a constant value $\tau_b = 80$ kPa)?

- Check that all the Appendices are referenced in the main text.
- Appendix: algorithm performance (M1QN3 vs Interior Point) is one difference in the implementation of the optimization (for the gradient descent). However, another difference could be the way the adjoint model is derived in each model. Do all the methods consider a "self-adjoint" problem or do some of them use a complete gradient (see Martin and Monier, 2014)?
- Consider zooming on areas of interest like the grounding zone and the fast flow regions instead of always showing the entire domain. For example, it is very hard to compare the GL position of the different runs in Fig 8.

Technical comments

- Line 32: "[…] inverse problem may have an infinite number of arbitrarily different solutions […]": I would avoid using "arbitrarily" since the solution is still based on the method implementation (cost functions, regularization parameters, …).
- Line 33: I also think that "direct" problem (steady state or snapshot response of the model for the value of the inferred parameters) is better phrasing than "forward" (which has a transient connotation, as used later by the authors) since the forward response can be different for two initial states giving the same velocity misfit.
- Line 82: "inversion runs" instead of "inversions run".
- Line 98: consider developing the effective strain rate (or second invariant) term
- Line 118: consider using a vector $\boldsymbol{p} = (p_1, p_2)$ here since it contains 2 parameters or components
- Line 130: "study" instead of "report"?
- Line 128: I think that the first term of the integral is a vector with *x* and *y* components and should use a norm like in Eq. (10), i.e. $\left\|\nabla\left(p_k - p_{k,prior}\right)\right\|$. How do you choose the prior values?
- Line 135: I am not sure that *b* needs to be introduced in Eq. (7) since it is directly equaled to 1. Is *b* always kept to 1 in different ISSM studies?
- Line 143: In Eq. (10), use p_k instead of *p* to keep consistency with Eq. (6)
- Line 159: $A_{i,j}$ is mentioned as the cell area but not used in the in Eqs. (11) and (12). Given that you invert for $p_1 = \sqrt{B}$, B_0 should be an initial estimate for $\sqrt{B}$ not B.
- Line 162: Are the regularization parameters also chosen with a L-curve analysis? Consider specifying it here too (since you did if for the two other models) or only mention once that you apply a L-curve analysis for the 3 models.
- Line 246: delete on of the closing parenthesis ")".
- Fig 4: first and second panels are both referred as (e), change for (d) and (e).
- Line 273: Is it only due to the fact that the regularization is conducted on the entire domain in Ua or to the prior used in Ua? This is answered in Appendix A4 but you do not refer to it in the main text.
- Line 282: Add "of" between "vicinity" and "the grounding line".

- Line 285: what do you mean by “to include peaks inside the rings of low values”? Do you directly constrain β to stay positive during the inversion, like at each iteration?
- Line 335: The term “grounding line regularization” feels unclear until we read the appendix (making the reader jumping several pages). I think it is good to keep the details in the appendix but maybe a sentence to explain what “the grounding line regularization” is would be welcome in the main text. Also, the different values for the coefficient *kH* are given in the appendix but it is never explained what it refers to in the implementation of the grounding line dynamics (or position).
- Line 346: Did you test different interpolation methods (nearest neighbor vs linear)? Also, what append to B and β^2 when directly inverting on Mesh2 and Mesh3? Are the values systematically higher than the interpolated fields, which could explain why the velocity is higher when interpolating?
- Fig. 6: Could you be more precise in the caption that this figure uses Mesh2 (in opposition to Fig. 3 using Mesh1)? Also, what is the reason for the very high misfit in panel (c) around x,y = (-1400, -600 km) ? Same for Fig B1.
- Fig. 7 and related text (line 360-368): I am confused here. Panels a, b and c display the same y-axis label (change in grounded are). Is that normal? If so, what is the difference between the panels? I understand from the text that 7c shows the change in grounded ice area but what about a and b? Is it also right that STREAMICE is ungrounding the most but that ISSM is losing more ice?
- Line 450: Missing “Pa” in the units for the friction parameter dipping below 10^2 Pa $m^{-1/3}$ $a^{1/3}$
- Line 460: I agree with the authors’ choice of capping the extreme values when calculating the correlation. Is this something you did for all the correlation values you got? I think it could be worth capping extremely low and high values. From my experience, 10^6 Pa $m^{-1/3}$ $a^{1/3}$ is a good value for capping low friction coefficient but the threshold values below and above which the flow is virtually not affected could be tested in a more systematic way to see if it could increase the Pearson correlation coefficients.

References

Martin, N. and Monnier, J.: Adjoint accuracy for the full Stokes ice flow model: limits to the transmission of basal friction variability to the surface, The Cryosphere, 8, 721–741, https://doi.org/10.5194/tc-8-721-2014, 2014.

Seroussi, H., Nowicki, S., Simon, E., Abe-Ouchi, A., Albrecht, T., Brondex, J., Cornford, S., Dumas, C., Gillet-Chaulet, F., Goelzer, H., Golledge, N. R., Gregory, J. M., Greve, R., Hoffman, M. J., Humbert, A., Huybrechts, P., Kleiner, T., Larour, E., Leguy, G., Lipscomb, W. H., Lowry, D., Mengel, M., Morlighem, M., Pattyn, F., Payne, A. J., Pollard, D., Price, S. F., Quiquet, A., Reerink, T. J., Reese, R., Rodehacke, C. B., Schlegel, N.-J., Shepherd, A., Sun, S., Sutter, J., Van Breedam, J., van de Wal, R. S. W., Winkelmann, R., and Zhang, T.: initMIP-Antarctica: an ice sheet model initialization experiment of ISMIP6, The Cryosphere, 13, 1441–1471, https://doi.org/10.5194/tc-13-1441-2019, 2019.

Seroussi, H., Nowicki, S., Payne, A. J., Goelzer, H., Lipscomb, W. H., Abe-Ouchi, A., Agosta, C., Albrecht, T., Asay-Davis, X., Barthel, A., Calov, R., Cullather, R., Dumas, C., Galton-Fenzi, B. K., Gladstone, R., Golledge, N. R., Gregory, J. M., Greve, R., Hattermann, T., Hoffman, M. J., Humbert, A., Huybrechts, P., Jourdain, N. C., Kleiner, T., Larour, E., Leguy, G. R., Lowry, D. P., Little, C. M., Morlighem, M., Pattyn, F., Pelle, T., Price, S. F., Quiquet, A., Reese, R., Schlegel, N.-J., Shepherd, A., Simon, E., Smith, R. S., Straneo, F., Sun, S., Trusel, L. D., Van Breedam, J., van de Wal, R. S. W., Winkelmann, R., Zhao, C., Zhang, T., and Zwinger, T.: ISMIP6 Antarctica: a multi-model ensemble of the Antarctic ice sheet evolution over the 21st century, The Cryosphere, 14, 3033–3070, https://doi.org/10.5194/tc-14-3033-2020, 2020.