Summary of revisions in manuscript “The transferability of adjoint inversion products between different ice flow models”

A separate document showing all changes to the manuscript is attached to this submission.

Earlier responses to the reviews are copied below for convenience.

In summary, to address the comments of the reviewers, revisions have been made as follows:

- Addition of experiments in which all inversion products are used in ISSM and STREAMICE for time-dependent simulations, and the necessary rewriting of section 5 along with new figures 6 and 7. Please note, all experiments presented in section 5 are new, including the Úa simulations. The revision process and comparison between forward runs offered us an opportunity to identify an issue which was causing results from Úa to differ significantly more than they should. The issue was due to the values of beta^2 on the ungrounded ice causing areas of regrounding after an initial advance of the grounding line. This was rectified by setting the beta^2 values in this region manually, and is discussed in the text. Thus, the factor of 2 in the differences from the original experiments is now greatly reduced.
- Addition of some numerical values to the abstract, to provide realistic expectations to the reader about the outcomes of the paper.
- Expansion of the discussion to give more detail on comparisons being made to other studies, and pointing out the limitations of some comparisons.
- Addition of more references to relevant sections of the appendix, along with summaries of the information contained within where appropriate, to increase clarity in the main body of the text.
- Addition of a new section to the appendix (now Appendix A), giving more detail on the choice of regularisation and displaying relevant L-curves.
- Addition of a new section to the appendix (now Appendix B5) addressing differences in the derivation of the adjoint.
- What was formerly section 5.1 has been merged into Appendix C, and referenced in the main text. The conclusion that the diagnostic results are not a good indicator of time-dependent performance is worthy of note in an appendix, but the details of this were confusing the narrative in their previous position.
- Information on the choice of priors added to section 3.1.
- Information on boundary conditions added to section 3.2.
- Average misfits for some defined velocity thresholds added to the analysis in section 4.1.

Response to RC1

We thank the referee for their review, which will be helpful in improving our manuscript.

In response to the point about section 4, we would argue that similarities and differences between the inversion outputs are of great interest. The nature of inversion is that infinite solutions are possible, even when using the same inversion method, due to the ill-posedness of the problem. Our three models use different methods, and so there is theoretically no reason to assume that they would produce similar results. Through inversion processes, it could be possible to produce velocity fields with low misfit compared to velocity measurements, but with fields of B or Beta^2 which do not physically represent the system. This is why we feel it is important to examine and compare the fields of B and Beta^2 before moving on to their application in transient simulations, and that the misfit alone is not necessarily an indicator of the quality of the inversions.

Regarding the transient simulations, a factor of 2 in the sea level contribution may sound large, but it is important to assess this within the right context. Several examples are listed in section 5.2, but this review makes it clear to us that significant expansion on this point, and some detail of other studies we are referencing, is required in the text to highlight the significance of our results. As one example, the control experiment in the initMIP-Antarctica comparison (Seroussi et al., 2019) shows a range of sea level contributions between -243mm and +167mm for the whole of Antarctica, with different models using different initialisation procedures. Within this range are a few examples of simulations run with the same models which differ by factors >3 due to differences in their initialisation. This, and results from other referenced studies, will be explicitly stated to aid in contextualising our results.

We agree with the referee that running diagnostic and transient simulations in all three models rather than just one is a good idea, and it is a point which the authors will act on. We hope that this will eliminate any doubt over the quality of the models we are using, reinforce our argument and help us to build a stronger case.

In general, the argument for Úa inverting for B across the entire domain is simply that we do not have definite information about the value of B. Not inverting for it assumes that there is zero uncertainty in the initial values imposed, which is not true. In the context of this work, we set out to compare the inversions produced by our models implementing their normal methods, and to show that despite the variety in the methods the results are physically robust. The difference in the B inversion is part of the variability between our methods, and not one of the controlled variables. An explanation to this effect will be added into section 2. The result of Úa turning off inversion for B is looked at in the appendix, section A2 (column c in Fig. A2). In this experiment, the speed misfit is higher, as the referee expects. The distribution of Beta^2 in this case remains consistent with the other experiments. This is mentioned in the main text (although perhaps needs to be emphasised) in section 4.4.

In response to specific comments:
There are indeed artifacts which need to be fixed in Fig. 6c. This will be done.
Line 235 refers to the misfit (Vdiff), and the wording will be updated to clarify this.

Response to RC2

We thank Cyrille Mosbeux for the detailed and constructive review, which will help us to improve our manuscript. Comments from the review are displayed in blue.

To address the general comments in the review, and those of Referee #1, actions are being taken (in fact, much has been done already) as follows:

- Addition of more references to relevant sections of the appendix, along with summaries of the information contained within where appropriate, to increase clarity in the main body of the text.
- Addition of a new section to the appendix, giving more detail on the choice of regularisation and displaying relevant L-curves.
- Addition of experiments in which all inversion products are used in ISSM and STREAMICE for time-dependent simulations. These experiments have already been run, and the results are available to be added to the manuscript.
- Expansion of the discussion to give more detail on comparisons being made to other studies, and the limitations of these comparisons, leading to a tempering of conclusions regarding the success of transferability.
- Editing of the abstract and introduction to reflect the revised tone of the conclusions, and temper the expectations of the reader.
- Merge section 5.1 into Appendix B, and summarise in the main text with clear reference to the appendix. We conclude that the diagnostic results are not a good indicator of time-dependent performance, which is worthy of note thus should be kept in an appendix, but this section in its current form seems to confuse the narrative and take some attention away from the main results (particularly based on the comments of Referee #1).

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 ($\boldsymbol{u}$=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 ($\boldsymbol{u}$=0) in *Sec. 5.1*. line 325.

ISSM has a Dirichlet boundary condition at the edge of the domain on grounded ice, given values based on the observed velocities. In STREAMICE, the grounded edges are given a no-flow boundary condition. These boundaries are sufficiently far from the area of interest to make no difference to the outcome.

All models apply an ice front stress condition. In Úa this is at the domain boundary, while in ISSM and STREAMICE the location of the ice front is fixed by an ice/ocean mask created using the geometry data.

This information has been added to section 3.2.

- 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.

It is true that ISSM produces higher misfit than the other two models in our study, and this is likely down to the equal weighting given to the absolute and logarithmic misfits. Perhaps a detailed study of these weighting choices could be the subject of some future work.

However, we do not believe the misfit is unreasonably high. There are not many areas which produce misfit >200 m/a. Regarding the comparison to Brondex et al. (2019), it appears that our misfit falls somewhere in the middle of the range of inferred states examined in that paper, similar to their Fig. 4(b).

-
    - 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).

A large amount of the difference seen in values for B, and the misfits, is a direct result of Úa inverting for B over the entire domain, whereas the other models do not. This is common practice in Úa and this project, while controlling certain input datasets, set out to compare inversions carried out under normal working practices in each model.

Relevant to both this and the point above, in Figure A2, column c, we display the result of carrying out an inversion in Úa following the practice of ISSM as closely as possible, and obtain a similar misfit to the ISSM inversion. In the revised manuscript, this will be referred to in the main text when discussing the misfit of our inversions.

-
    - 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,$Rp1$ and $Rp2$).

In the case of ISSM, the inversions for B and beta^2 are carried out separately, one after the other, as described in subsection 2.2.2. This means that the two regularisation parameters can be chosen independently. This results in two separate 2D L-curves, rather than one in 3D. The same cannot be said for the other models, however. We are adding a new appendix section to explain the regularisation choices and display the L-curves. The issue of multiple regularisation parameters and how these were approached in each model will be discussed here. The appendix section will be referred to in section 2.2.

    - 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)?

This is a good point, and average misfits for a defined velocity threshold will be added to the analysis in section 4.1.

- 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 τb=80 kPa)?

The priors will be specified in section 2.2, and references to appendices will be improved throughout the manuscript. The use of priors defined in different ways, including one based on driving stress, could be an interesting topic for further experimentation, but beyond the intended scope of this project. For the purposes of this paper we chose one spatially varying prior to contrast with a spatially uniform prior and the original priors used by each model.

- Check that all the Appendices are referenced in the main text.

This is being improved upon in revisions to the manuscript.

- 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)?

The models do use slightly different derivations of the adjoint. ISSM uses the exact adjoint described in Morlighem et al. (2013). Following that paper, we prefer to refer to the alternative as “incomplete” rather than “self-adjoint”. STREAMICE uses the method described in Goldberg et al. (2016), but with a relatively weak tolerance placing it somewhere between the exact and incomplete adjoints.

We did not consider these among the differences in Appendix A, and will mention this as a possible factor alongside the others in the revised manuscript.

- 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.

A good point! We will zoom in to display the areas of interest in Fig. 8 and make the differences much clearer to see.

Technical comments

Thank you for pointing out typographical errors, and areas which require greater clarity. These shall all be addressed. A few of these technical comments require specific responses.

- 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.

The parameters for STREAMICE were selected based on an L-curve analysis from previous experiments run in the model (Goldberg et al., 2019), rather than independently for this work. This will be made clear and the reference cited in a new appendix section discussing regularisation choices.

- 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?

There was not a constraint within the inversion. The effect is a result of post-processing of the data. This has been clarified in the text.

- 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).

Clarity on this matter is being improved by the merging of section 5.1 and Appendix B. Improvements are being made throughout on references to the appendices, and reducing the need to jump back and forth between pages.

- 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?

The repeated labels are an image processing error which will be corrected. Panel a) should read “Loss of volume above flotation (Gt)” and panel b) should read “Change in total ice mass (Gt)”.

This entire section will be expanded to include results from the other models, and to improve clarity.

- 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.

Several thresholds were tried for capping extreme values, and the chosen values are those which were deemed to strike the best balance between preserving the shape of the data and discarding anomalous spikes. This was not done for every correlation value, but only where necessary in some of the comparisons presented in Table A1. The affected values are indicated in the table and caption.

References

Brondex, J., Gillet-Chaulet, F. and Gagliardini, O., 2019. Sensitivity of centennial mass loss projections of the Amundsen basin to the friction law. *The Cryosphere*, *13*(1), pp.177-195.

Morlighem, M., Seroussi, H., Larour, E. and Rignot, E., 2013. Inversion of basal friction in Antarctica using exact and incomplete adjoints of a higher-order model. *Journal of Geophysical Research: Earth Surface*, *118*(3), pp.1746-1753.

Goldberg, D., Narayanan, S.H.K., Hascoet, L. and Utke, J., 2016. An optimized treatment for algorithmic differentiation of an important glaciological fixed-point problem.

Goldberg, D.N., Gourmelen, N., Kimura, S., Millan, R. and Snow, K., 2019. How accurately should we model ice shelf melt rates?. *Geophysical Research Letters*, *46*(1), pp.189-199.