Reconciling ice dynamics and bed topography with a versatile and fast ice thickness inversion - response to reviewer 1

Thomas Frank, Ward van Pelt, Jack Kohler

June 2023

1 Response to reviewer 1

We would like to thank reviewer Samuel Cook for his remarks that helped us to improve our manuscript. We reply below by highlighting the reviewer's comments in italics above our answers.

1.1 Summary

This paper presents a new method for inverting for ice thickness and, optionally, basal friction at a glacier by iteratively minimising misfit between observed and modelled dh/dt through adjusting an initial guess at the bed (to which the method is insensitive), and by adopting the same approach with observed and modelled surface velocity for the basal friction field. The authors test their method on a synthetic ice cap in order to quantify its performance and sensitivity, and then apply it to Kronebreen, a tidewater glacier in Svalbard, finding that the method generally reproduces the bed well, and is fairly robust when subjected to a wide range of model parameters and input data perturbations. They also show that their method allows the creation of a self-consistent model glacier that can then be used in prognostic simulations without any relaxation. I think this is a good paper that makes a useful contribution to the field of glacier thickness inversion, but I have a few major concerns around the general framing and structure of the paper that I feel need to be addressed before I can recommend it for publication. I do not think any of these are particularly difficult to solve, however, and I think the authors should be congratulated on coming up with an ingenious new solution to a difficult conundrum!

We would like to thank the reviewer for his overall positive comments.

1.2 Major comments

The title and abstract of the paper major on this being a fast method for inverting for thickness, but I think the paper doesn't necessarily back that up.

- *An easily fixable issue here is that the paper never actually says how fast the method is, beyond saying it's 'fast'. The authors present results in terms of the number of iterations they took to achieve, but never say how long it takes to run an iteration. So, somewhere, I think the paper has to provide an actual time per iteration or per experiment to allow the reader to judge how fast the method is*

- *I feel it's also difficult to say the method is fast when its speed depends on the user's choice of ice-flow model. Presumably, if you were using a full-Stokes model, the method would be considerably slower?*

- *I also think that the speed of the method is not really the important thing about it. I would say the real advantage of the method is that you end up with a self-consistent model glacier that you can stick into a prognostic simulation without having to do any relaxation. After all, you can invert for thickness very quickly if you just assume perfect plasticity and have a DEM. The resulting glacier will be a horrible mess, but it will be fast. What we don't have so much yet is a method that shortcuts all that relaxation running time, and that's where this method really innovates, I think.*

- *So, overall, I would strongly encourage the authors to reframe the paper towards the 'self- consistent model glacier' aspect and away from the 'fast' aspect. I think changing the title and the abstract would go a long way towards doing that – the discussion already points this way, so I really think this is more-or-less a cosmetic front-end change than anything else. Though I think the authors really do still need to give somewhere an idea of how long the method takes to run in actual time, rather than just giving a number of iterations, as that's a useful comparator with other methods, regardless of any claims the authors do or don't make about how fast the method is.*

We agree with the reviewer that our claim of a **fast** inversion technique has not been discussed enough in the original manuscript. At the same time, we do not think that it is very meaningful to give an actual time per iteration, as suggested, since the computational time of every inversion depends on a range of glacier-, hardware (as in computational capacity) and ice flow model characteristics as well as model resolution that will vary greatly between applications of our method. As correctly pointed out by the reviewer, running a Full-Stokes model, for instance, would make the method slower. So, instead of giving a time per iteration, the revised manuscript will contain are better discussion around speed of ice thickness inversion methods. Specifically, we will define that *fast* must be understood as *comparatively fast* considering what other approaches yielding a similar output would cost in computational time. Based on our test glacier Kronebreen, we will demonstrate that running 8000 iterations, each of which comprises a forward run of 0.01 yrs, results in a net computational time equivalent to running a forward model for 80 yrs. This can be compared to other inversion approaches, e.g. van Pelt et al. (2013) (requires 1000s of model years, Pollard and DeConto (2012) (100,000s of model years), Le clec'h et al. (2019) (1000s of model years) or expensive adjoint-based ones (e.g Goldberg and Heimbach, 2013).

On re-framing the study, we do agree that the abstract could better reflect the 'self-consistent model glacier' aspect. We will stress this part more in the abstract and instead move the focus away from the large-scale applicability of the method.

Section 3.2 (on model sensitivity) needs some rewriting. I'm very glad the authors tested the sensitivity of the model to both input errors and model parameters – that's a really useful thing to do – but I think this section needs quite a bit more work to make it really helpful for the reader. As it stands, it's much too vague and imprecise to provide a clear picture of model sensitivity – I came away from it feeling I didn't really have much more of a clue as to how sensitive the model really was compared to when I started reading it. I know the authors give a lot of the details in the Appendix, but I think there are some headline statistics the authors could add to the main body of the text to make this section useful without having to make the reader dig down into the appendix.

- *In particular, in Section 3.2.1, it would be helpful if the authors could provide some quantification of how much the perturbations to the input data impacted on the final bed/friction field, beyond just saying 'large bed responses', which is a bit vague – 10 m, 50 m, 100 m?. The authors state* *that the ice volume is pretty insensitive to all the perturbations (though I would like some idea of what the percentage change in volume across all the simulations is), and helpfully explain all the mechanisms by which the errors feed through into glaciological changes, but I didn't really get any sense from the section as written what the magnitude of those glaciological changes were. As it stands, the section leaves open the possibility that it could well be that the perturbations don't affect the overall ice volume, because they lead to the bed being 500 m too high in half the domain and 500 m too low in the other half. I would also want to know how much the authors had to increase theta to mitigate the issues – 1.5x, 3x, 5x, 10x? - and I think there needs to be a bit more detail around 'more dependent on initial conditions'. The best solution is perhaps a table showing the mean bed and dh/dt mismatch for each simulation (or some other metric of the model's performance), and then maybe a figure showing some of the final beds to show what 'more dependent on initial conditions' actually means.*

- *For Section 3.2.2, I would similarly like some idea of the range of values of beta and theta the authors tested (the authors give the range of values for T, for example, in Section 3.2.1; the same should be done here). The authors repeat their previous observation about larger theta leading to greater dependence on initial conditions here, with a little more detail as to why, but I'd still like to get some sort of sense of what that looks like or means in practice without having to dig through the appendix. If the authors do put together some sort of figure to show that, I think it fits better here than in 3.2.1, as this is the section about explicitly testing theta (in which case, the authors could just stick a 'see below' in Section 3.2.1 when they mention that larger theta lead to greater dependency).*

We appreciate the suggestions to modify section 3.2. As mentioned in the manuscript, the influence of errors in the input data on the calculated bed depend on several glacier characteristics and are thus hard to generalize - hence why we refrained from giving specific numbers. However, we follow the reviewer and will provide more quantitative metrics, both in the text and in a new table. We will include metrics for each perturbation experiment, namely the ratio between the ice volume given the perturbation and the true ice cap volume, as well as the absolute bed misfit. Together, these two metrics reflect whether or not a perturbation results in the correct ice volume and if the bed is similar in shape to the true bed. A figure on varying theta will also be included. Furthermore, we will provide the range of tested values in sec. 3.2.2 as suggested.

As a more general point, I'm wondering what would happen if the authors updated friction first, then the bed. Would the method converge towards the same friction/bed combination, or would you get something different? Fundamentally, how do the authors deal with the problem of equifinality (as in, there are lots of possible friction/bed combinations that would reproduce the observed variables – how do the authors know theirs is a sensible one?)? Some discussion of this would be a useful addition, either in Section 5, or elsewhere where the paper discusses the robustness of the model. One additional simulation changing the order of the friction and bed updates would, I think, be very instructive and I would encourage the authors to see if it makes a difference.

The question of equifinality is indeed an interesting one that has not been discussed enough in our initial submission. Recall that the concept of the apparent mass balance shows that dh/dt and mass balance together determine the total mass which is fluxed through a glacier. This mass may be transported either by a thick but slow glacier, or a thin but fast glacier. By forcing our model to reproduce observed velocities via the friction coefficient there would only be one ice thickness that matches the mass flux if ice thickness and ice speed were independent of each other. However, since ice thickness also influences ice speed by changing the driving stress, there is a small realm of overlap between two regimes, namely when observed velocities can be reproduced either by a change in ice thickness or in basal friction. Due to the different characteristics of sliding (non-local, vertically uniform velocity profile) and shearing (local, vertically dependent velocity profile) ice physics, we think that it is typically very unlikely that changing thickness or basal friction would lead to the same ice velocity and mass flux, and therefore, that equifinality is an issue. Given the fact that all of our experiments, synthetic and real-world ones, converge to a sensible bed even with different initial conditions corroborates this idea. We will include this discussion in the revised manuscript.

On the comment of changing the order between friction and bed updates: Friction updates have a near-instantaneous effect on ice velocity and therefore ice thickness, leading to a rapid convergence to a new state after a friction update. And although the effect of a friction update applied at one grid point is non-local due to the mathematical nature of the SSA, the radius of affected cells around is limited, again indicating that a new state is reached rapidly after a friction update. Bed updates, on the contrary, lead to changes in the mass 'export' of an affected cell, and thus affect all downstream grid cells which will also change their mass 'export'. Because of that, bed updates are travelling downstream through the entire glacier, implying that it takes much longer to reach a new state after bed updates. In fact, through this concept, it can be understood that the convergence time to a new bed is equivalent to the response time of the glacier to a mass balance perturbation. So, given the difference in convergence time of bed vs. friction updates, it does not make sense to swap the order between them. We did conduct experiments that show just that, but for the sake of conciseness suggest that a discussion as done here will suffice in the revised manuscript.

1.3 Minor comments

p.1, l.14: 'fieldwork' not 'field work'

Will change.

p.2, l.24: 'allow the derivation of' not 'allow to derive'

Will change.

p.2, l.30: 'former' not 'previous'

Will change.

p.2, l.47: Some examples of where the conditions/assumptions in the SIA aren't met (sliding glaciers...) would be good here - here the text assumes that people will know and don't explain, unless the reader goes back a few lines and works out that 'considering only internal shear' means 'not valid where there's something else going on'. It just makes the reader's life slightly easier.

Will change.

p.2, l.51: The text doesn't quite explain explicitly why using beds derived using standard thickness-inversion processes will require you to do model relaxation if you use them as inputs in a higher- order model. The text states previously that you'll get errors in the ice thickness using one of the standard approaches, but why do those errors then require model relaxation (the velocity and thickness fields won't match so there'll be artefacts so...)? It might also be worth explaining what model relaxation actually is. As with the previous point, the text is perhaps assuming slightly too much knowledge on the part of the reader and making them work harder than they perhaps should in the introduction!

Will change.

p.2, l.55: 'the present day'

Will change.

p.3, l.57: 'allows the inclusion of' not 'allows to include'

Will change.

p.3, l.80: 'way' not 'part'

Will change.

p.3, l.82-3: How well do we actually know the mass balance at most glaciers? This approach assumes the surface variables are well-known (usually true for the surface topography) and that the errors in the model are therefore entirely due to bed errors, but I'm not convinced that assumption is valid at most glaciers. I realise this is dealt with later on, but some acknowledgement here that it is an assumption that isn't always going to be true would be good.

What we state here is the assumption (not the requirement) under which our method is applied. We state that the surface shape and mass balance need to be *sufficiently* well represented, i.e. good enough to derive a reasonable bed estimate. How good an observation of mass balance needs to be to suffice to that end is a rather complex question. In our sensitivity experiments, we show that it depends on glacier characteristics and that the mass balance error can be quite large for non-sliding glaciers without causing significant bed errors (e.g. a 75% mass balance overestimation causes only a 12% ice volume overestimation). Due to this complexity, we do not think that sec. 2.1 is the right place to discuss mass balance errors.

p.4, l.89: 'point out' Habermann et al is plural.

Will change.

p.5, l.126: 'that gently induce'

Will change.

p.5, l.127: 'that it cannot otherwise reproduce'

Will change.

p.5, l.132: 'that could lead' I think I'll stop pointing these out now, but the general rule is that 'which' is used when introducing a new clause, otherwise use 'that'. Pretty much, if you've stuck a comma in, use 'which', if there's no comma, use 'that'.

Alright, thanks for the clarification.

p.9, l.219: 'error' not 'errors'

Will change.

p.9, l.220: 'lead to similar outcomes' not 'induce similar mechanisms'

Will change.

p.10, l.252: 'the subject'

Will change.

p.13, l.332: 'it can be higher' – you're talking about the mismatch (singular). Linked to this, how much higher is 'higher' here?

Will clarify.

Section 4.5: I'm not entirely convinced by the post-processing step here, as it seems a little counter-intuitive to me to use an equation based on the SIA to correct errors caused by using the SIA. I don't deny that it works, I just don't like it conceptually. To be fair, I'm not sure what else could be done that would be reasonable, but I suspect readers might react in the same way as me, so if the authors can just add a sentence acknowledging that using the thing that caused the problems to fix the problems is what's going on and that the reader's possible confusion is understandable, it might make it a slightly easier pill to swallow.

We will add a note on that!

p.17, l.418: 'such as' not 'such that'

Will change.

Section 5: I think the approach presented in this paper could be usefully applied to the ITMIX dataset to benchmark it against other widely-used inversion methods. Are the authors considering doing this? Either way, a statement on yes/no and why/why not towards the end of the section would be good, because it's otherwise a bit of an elephant in the room

We chose Kronebreen as a test case due to its exceptional coverage with ice thickness observations which is even better than for the ITMIX glaciers. Nevertheless, a comparison with one of the ITMIX glaciers is a reasonable suggestion which we also considered including in our original manuscript. However, there are reasons that speak against that: First, describing another setup will inflate the length of the manuscript considerably and lead to an overall confusing structure of the paper. Second, running the hybrid SIA-SSA model on valley glaciers, as are mainly found in the ITMIX setup, is challenging. In fact, using PISM for such a setup would stretch the model´s capabilities and possibly lead to poor results. A model more suited for valley glaciers (i.e. a higher-order or Full Stokes model) would hence need to be used as a forward model, but again, describing this model in the manuscript would inflate the length and divert the focus away from the key points of the paper. We, therefore, suggest to refrain from adding further test glaciers. However, we are currently preparing another manuscript where we use the Full-Stokes emulated ice-flow model IGM (Jouvet, 2022) to apply our method on a regional scale which includes ITMIX glaciers. We will add an outlook in the discussion section on that. Furthermore, to aid comparability of our method with other inversion techniques, we will more explicitly discuss how the results of Kronebreen compare to other thickness products available for that glacier in the revised manuscript.

References

Goldberg DN, Heimbach P. 2013. Parameter and state estimation with a time-dependent adjoint marine ice sheet model. The Cryosphere 7:1659–1678. doi:10.5194/tc-7-1659-2013.

Jouvet G. 2022. Inversion of a Stokes glacier flow model emulated by deep learning. Journal of Glaciology :1–14doi:10.1017/jog.2022.41. Publisher: Cambridge University Press.

Le clec'h S, Quiquet A, Charbit S, Dumas C, Kageyama M, Ritz C. 2019. A rapidly converging initialisation method to simulate the present-day Greenland ice sheet using the GRISLI ice sheet model (version 1.3). Geoscientific Model Development 12:2481–2499. doi:10.5194/gmd-12-2481-2019. Publisher: Copernicus GmbH.

Pollard D, DeConto RM. 2012. A simple inverse method for the distribution of basal sliding coefficients under ice sheets, applied to Antarctica. The Cryosphere 6:953–971. doi:10.5194/tc-6-953-2012.

van Pelt WJJ, Oerlemans J, Reijmer CH, Pettersson R, Pohjola VA, Isaksson E, Divine D. 2013. An iterative inverse method to estimate basal topography and initialize ice flow models. The Cryosphere 7:987–1006. doi:10.5194/tc-7-987-2013.

Reconciling ice dynamics and bed topography with a versatile and fast ice thickness inversion - response to reviewer 2

Thomas Frank, Ward van Pelt, Jack Kohler

June 2023

1 Response to reviewer 2

We are grateful to the reviewer for his/her remarks that helped us to improve our manuscript. We reply below by highlighting the reviewer's comments in italics above our answers.

1.1 Summary

In this manuscript, the authors present an iterative approach to simultaneously infer the poorly constrained basal topography beneath glaciers as well as to initialise a glacier system model into a self-consistent state for prognostic simulations. Input requirements are surface observations on ice geometry and ice velocity as well as surface mass balance estimates (SMB). Thickness measurements are not ingested but withheld for validation. The approach is applied to a synthetic ice-cap setup and to a real-world glacier, i.e. Kronebreen on Svalbard. The synthetic setup serves to present a performance baseline for non-slip and sliding regimes under idealised input conditions. It is also exploited to analyse the sensitivity to input uncertainties and a-priori parameter choices. The method is then put to the test on Kronebreen, for which further methodological refinements - also during post-processing - are introduced. The performance of the iterative optimisation approach convinces for the synthetic setup, as the bed is retrieved with high accuracy. In the real-world, the misfit remains elevated as input fields and model parameters are less well known. Nonetheless misfit metrics are comparable to other approaches when observations are not ingested.
When I accepted the review, I was mostly attracted by the fact that this approach allows for a simultaneous initialisation of a forward model. Although this remains a side aspect in the manuscript, I consider this initialisation a big asset of the approach presented in this manuscript. I want to congratulate the authors to their concise presentation of this initialisation strategy. During my review, I however identified some major concerns on methodological details and the experimental design. The manuscript is well written and easy to follow, yet the structure can be improved. Moreover, there is a clear need for extending the discussion, certainly with regard to an assessment of the performance. In summary, I am very positive about this manuscript and I recommend that the editor should continue to considered it for publication in The Cryosphere after my concerns have been alleviated. This will certainly imply a major revision.

We thank the reviewer for the overall positive attitude towards our manuscript.

1.2 Major comments

Benchmark Glacier

I think you should extend your setup to one of the ITMIX2 benchmark glacier, preferably a valley glacier. ITMIX2 is the reference benchmark for such approaches. As glacier, I suggest Austre Grønfjordsbreen as it is also on Svalbard and as most input is available. I ask for this because you deliberately forward regional-scale applicability on which such valley-glacier setups are regularly encountered. Moreover, sliding is less important and might be a challenge for the hybrid SIA-SSA model variant of PISM that uses an empirical function to combine sliding and creep. For a valley-glacier setting, I wonder how your initial viscosity choice will affect your performance even when allowing for sliding updates. Moreover, an ITMIX setup would provide possibilities to directly compare to the performance of other reconstruction approaches.

We chose Kronebreen as a test case due to its exceptional coverage with ice thickness observations which is even better than for the ITMIX glaciers. Nevertheless, a comparison with one of the ITMIX glaciers is a reasonable suggestion which we also considered including in our original manuscript. However, there are reasons that speak against that: First, describing another setup will inflate the length of the manuscript considerably and lead to an overall confusing structure of the paper. Second, running the hybrid SIA-SSA model on a valley glacier, as suggested, is indeed challenging. In fact, using PISM for such a setup would stretch the model´s capabilities and possibly lead to poor results. A model more suited for valley glaciers (i.e. a higher-order or Full Stokes model) would hence need to be used as a forward model, but again, describing this model in the manuscript would inflate the length and divert the focus away from the key points of the paper. We, therefore, suggest to refrain from adding another test glacier. However, we are currently preparing another manuscript where we use the fast Full-Stokes emulated ice-flow model IGM (Jouvet, 2022) to apply our method on a regional scale, thus attesting to the large-scale applicability of our method. We will add an outlook in the discussion section on that. Furthermore, to accommodate the reviewer´s comment on comparability of our method with other inversion techniques, we will more explicitly discuss how the results of Kronebreen compare to other thickness products available for that glacier.

Thickness observations

I truly appreciate that you compare your Kronebreen results to thickness observations. Yet I wondered why you did not use them during the iterative optimisation to better constrain your bedrock result. As you do know the bed in some locations, you could simply apply a 'restoring' in your bedrock update (Eq. 1) that drags the bed in each iteration towards these observations. In this way, you could further reduce the final bed misfit that currently exceeds 100m in the Kronebreen setup (Fig. 4a). I also want to emphasise that reconstruction models in the 2nd round of ITMIX were asked to ingest thickness measurements. Moreover, ITMIX2 did highlight the importance of direct measurements and gave suggestions on acquisition strategies. Therefore, your approach would highly benefit from the capability of assimilating thickness measurements.

We appreciate the suggestion from the reviewer. So far, we did not include any thickness observations in our bed inversion since we regard it as a strength of our method that it does not require any observations. Considering the lack of bed observations for most of the glaciers world-wide, this is a great advantage which will be handy when applying the method on a large scale. Nevertheless, we agree with the reviewer that it is interesting to test how good the method would work when thickness observations are available. We therefore ran simulations where we used different percentages (30%, 50% and 80%) of the bed observations and kept the bed height fixed there. However, this did not improve our results further. We think that this because our method places great weight on internal model consistency meaning that the method can only yield a result that is consistent between the boundary conditions and the model physics. When artificially fixing the bed height at a location, an inconsistency is created which the model cannot take advantage of. Instead of prescribing bed heights, we therefore suggest to tune certain parameters (e.g. ice viscosity) to assimilate ice thickness observations. We will reflect this point in the discussion of the revised manuscript.

On iterations and alternations

If I understand your alternating optimisation strategy well, the basal and surface topographies are iteratively adjusted during prescribed 1000 steps followed by a single friction update. I wonder why the frictions coefficient is not updated iteratively as well. Friction seems decisive. Did you try this? I wonder if this could speed-up the convergence. As it stands now the convergence will very much depend on the actual value for λ. I further wonder about the stopping criterion, which is set as a fixed number of iterations (which is not mentioned in the methods but during the experimental description - please adjust). In this way, much computing time is spent in negligible bed updates. An objective criterion on effective basal topography changes per iteration could help to further speed-up the optimisation. In addition, it would imply actual convergence.

To clarify: we conduct in total several thousand iterations per experiment, so friction is updated several times, unlike what the above comment seems to indicate ("*I wonder why the frictions coefficient is not updated iteratively as well.*"). If the reviewer is asking whether we could update friction field and bed topography simultaneously in each iteration, we have done such experiments. However, we find that doing so the method does not converge. As stated in the manuscript (L139-142): "Importantly, bed and friction are not updated in the same iteration. This is because the dh/dt misfit and the velocity misfit both represent a mismatch in modelled vs. observed flow dynamics and it is not clear how to disentangle which part of that mismatch is due to an erroneous friction field vs. an erroneous bed." In the revised manuscript, we will add that this statement is based on experiments we conducted.

Furthermore, if the reviewer is asking whether the friction field could be updated more frequently than the bed: Friction updates have a near-instantaneous effect on ice velocity and therefore ice thickness, leading to a rapid convergence to a new state after a friction update. And although the effect of a friction update applied at one grid point is non-local due to the mathematical nature of the SSA, the radius of affected cells around is limited, again indicating that a new state is reached rapidly after a friction update. Bed updates, on the contrary, lead to changes in the mass 'export' of an affected cell, and thus affect all downstream grid cells which will also change their mass 'export'. Because of that, bed updates are travelling downstream through the entire glacier, implying that it takes much longer to reach a new state after bed updates. In fact, through this concept, it can be understood that the convergence time to a new bed is equivalent to the response time of the glacier to a mass balance perturbation. So, given the difference in convergence time of bed vs. friction updates, it does not make sense to swap the order between them. We did conduct experiments that show just that, but for the sake of conciseness suggest that a discussion as done here will suffice in the revised manuscript.

Regarding the comment on a stopping criterion, the revised manuscript will include a remark stating that a stopping criterion could easily be set based on, e.g. Δdh, by triggering a new friction update when Δdh is smaller than a pre-defined value. For the sake of simplicity of this first application of the method, we do not explore this possibility further in our manuscript.

Sensitivity analysis

I deeply appreciate the sensitivity analysis for your synthetic ice-cap setup in the appendix. To me the most unconstrained parameters/variables are the climatic mass balance (CMB) and the ice viscosity. As I understand it, the synthetic geometry is in equilibrium, which means that the specific CMB should be zero. In Figure A1, you present relative CMB values with regard to a reference, which I suspect *should be zero in a specific sense. So how can this value be increased by 25-75%. This is certainly a misperception from my part (I might sense that you actually increase the point values by these relatvie values – which would imply that the specific values remains unchanged; in reality we do not know the specific CMB). In any case please explain and think about rather providing absolute specific perturbation values in metres ice equivalent per year. Concerning ice viscosity, I think your analysis is convincing, certainly as you also present results without friction update (Fig. A4). Finally the sensitivity test with regard to the time step dt of the forward model seems dispensable. In my view, no matter the choice, equation (1) should internally compensate the time step scaling. The only constraint is that the mass conservation implementation remains numerically stable.*

As correctly inferred by the reviewer, we modified the point values of CMB rather than uniformly adding or subtracting mass at all points in the domain. Although, as the reviewer rightfully points out, we do not know specific CMB usually, we can, prior to an inversion, ensure that the mass balance $\int_\Omega CMB d\Omega = \int_\Omega \frac{dh}{dt} d\Omega$ over a glacier is closed. While this admittedly may be harder for marine-terminating glaciers (although large-scale calving flux estimates exist (Kochtitzky et al., 2022)), it is straight-forward for land-terminating glaciers such as our ice-cap example. Because of that, it would not have made sense to perturb CMB with an absolute value. We will clarify our experimental setup for the mass balance perturbations in the revised manuscript.

On the comment of testing dt, we do think that there is a point in conducting a sensitivity experiment. Specifically, by calculating $\frac{dh_{\mathrm{mod}}}{dt}$ from the ice thickness at the start of the forward simulation minus the ice thickness at the end of the forward simulation, we assume that the change between these two points in time is linear. However, this is not necessarily the case because $\frac{dh_{\mathrm{mod}}}{dt}$ reduces as the glacier relaxes towards a new steady-state. By choosing a small dt, we may, therefore, always sample the steep part of the glacier relaxation curve, while choosing a large dt could also include points in time when the glacier is already fully relaxed. The fact that we do not find any such effects suggests that this is indeed not a significant problem, but this is not directly obvious.

Manuscript structure

In my view, you should better distinguish between methods, results and discussion. I somehow like your division by synthetic and real-world setup in terms of experimental setup and results. Yet the methodological updates in the latter setup, concerning the control parameter update (on β) and the post-processing, are confusing. I would introduce both concepts already in the methods section and apply them consistently both in the synthetic and real-world setup. Please streamline both setups in terms of methods. The manuscript would be easier to follow. Moreover, the last section on 'Discussion & Conclusions' is confusing. Please separate both aspects into dedicated sections.

We thank the reviewer for suggestions on streamlining the manuscript. Following the comment, we will completely restructure the methods section and add a new subsection 'Stabilization techniques' where we introduce several tools used later in the experiments which aid convergence, among others the 'ramp-up' for β. Doing so, we hope to accommodate the reviewers request to better separate between methods and results. In fact, since submission of the initial manuscript, we have discovered a new stabilization method which improves the results for the synthetic ice cap. Corresponding changes in the ice cap section and the appendix are therefore made, which, however, do not change any general conclusions of the paper. For Kronebreen, the results remain unchanged. We will also restructure the discussion section to improve readability.

Regarding the post-processing step, we do not consider it part of the core design of our new method, but rather an addition which generally should not be needed with our method. If using our method for spin-up, for instance, the post-processing as done in this manuscript would not be advisable. Because of that we suggest to keep that chapter at its current place.

We will add a separate conclusions section.

Discussion
As it stands, your discussion focusses on the benefits of the sliding updates, limitations from regularisation and the post-processing. I miss some comparison of how your approach performs with respect to others. You forward the mean absolute error as a measure of performance - is this quantity available for other approaches. To my knowledge the global consensus estimate by Farinotti et al. (2019) can be exploited for a direct comparison to your Kronebreen results

As mentioned in response to a previous comment, we will include a more detailed discussion on that in the revised manuscript. Specifically, we will present the mean absolute error for three other available ice thickness products and discuss them in relation to our results.

1.3 Minor comments

L1-4 *I totally understand your intrinsic excitement/motivation to raise global scale applicability in the first sentence. Yet in L4 your phrasing already moderates this applicability to local and large scales. As this study presents a new method, the abstract should rather focus on performance not so much on the outlook. This outlook is indeed exciting and should/could be an aspect for your conclusion section.*

We agree with the reviewer that the original abstract may have been focused too much on the outlook and will rephrase.

L6-7 *Here you claim that your iterative approach also serves for model initialisation into a self-consistent state. Your experimental setup does not substantiate this claim. It is taken as a fact from the methodological design. It could be worth to run the synthetic ice-cap setup forward in time after bed retrieval. Ideally the geometry would not change much in this equilibrium setup.*

This is a good point by the reviewer. To substantiate our claim of a self-consistent model state after the bed and friction inversion, we have conducted an experiment where we run the ice cap forward for 100 yrs after the inversion.The ice volume change over that period is 0.18%, suggesting that the ice cap is in a steady-state. A remark on that will be added in the revised manuscript.

L92 *You formulate that the observed elevation change is the primary target quantity for optimisation. I am a bit worried on the model capabilities to reproduce these rates near the glacier margin (no matter if land- or marine- terminating). It is known that this regions is critical in terms of flux divergence. In your synthetic ice-cap setup, you deliberately exclude the margin from the bed retrieval (L183). For the real-world glacier you even introduce strong gradients by the applied masking. Together with the SIA aspect in the ice-dynamic formulation, I wonder about any consequences for the applicability. This is certainly another point that should be picked up in the discussion.*

Unfortunately, we are not entirely sure what the reviewer refers to exactly. Indeed, ice margins can be difficult to model, hence why interpolation there may be in order (which we will describe in the newly create section 'Stabilization techniques'). However, we do not see how the masking in the Kronebreen example introduces strong gradients (see below), and we also do not generally observe issues at the ice margins.

L151 *As your iterative approach infers the bedrock topography and the friction coefficient simultaneously, you should discuss potential ambiguities. Is this problem well posed? Is there only a single* *solution and are you convinced that the target parameters can be well differentiated.*

The question of whether there are unique solutions is indeed an interesting one that has not been discussed enough in our initial submission. Recall that the concept of the apparent mass balance shows that dh/dt and mass balance together determine the total mass which is fluxed through a glacier. This mass may be transported either by a thick but slow glacier, or a thin but fast glacier. By forcing our model to reproduce observed velocities via the friction coefficient there would only be one ice thickness that matches the mass flux if ice thickness and ice speed were independent of each other. However, since ice thickness also influences ice speed by changing the driving stress, there is a small realm of overlap between two regimes, namely when observed velocities can be reproduced either by a change in ice thickness or in basal friction. Due to the different characteristics of sliding (non-local, vertically uniform velocity profile) and shearing (local, vertically dependent velocity profile) ice physics, we think that it is typically very unlikely that changing thickness or basal friction would lead to the same ice velocity and mass flux, and therefore, that non-unique solutions are common. Given the fact that all of our experiments, synthetic and real-world ones, converge to a sensible bed even with different initial conditions corroborates this idea. We will include this discussion in the revised manuscript.

L183-184 *Do you also apply this margin masking for the real-world setup.*

No, we will clarify that in the revised manuscript.

L267 *This initial Gaussian filtering seems vital for application of the SIA. Still for more complex valley-glacier geometries, this step might not remove all artefacts. I therefore wonder if you also tried an initial relaxation with a prognostic run for which the geometry is not allowed to evolve too much (by capping the elevation change rates). This strategy could be more robust and beneficial as a prior step to your reconstruction.*

It is unclear to us how this should work, since any prognostic run requires bed elevation as input, and hence any glacier surface shape resulting from the forward simulation would be directly dependent on the assumed bed shape. A bed inversion starting from a surface shape derived in that manner would likely yield a bed very similar to the one used in the initial relaxation.

L282 *The masking of ice cover to the Kronebreen outline seems a bit harsh as it will introduce extreme gradients in surface elevation for example at the divide with Kongsbreen. I suggest to rather keep the full ice geometry also outside the Kronebreen outline and only update the basal topography within the mask (and prescribe/freeze it outside). It should not be difficult. Probably this is anyway what you have done.*

If we understand correctly what the reviewer refers to here, it may indeed be that we did as the reviewer suggested. To clarify, there are no strong gradients in surface elevation at the margin of Kronebreen since the surface height in the immediate vicinity of our mask is given by the same DEM used inside the mask. When we state that outside the mask, ice thickness is forced to zero, it means that the bed elevation there corresponds to the surface elevation. If what the reviewer suggests is to prescribe a fixed non-zero ice thickness outside the mask, we do not think that this is meaningful as it would require us to make an assumption on the surrounding ice thickness which is unknown. As stated in the manuscript (L.286), with our setup some remaining boundary effects still cannot be ruled out, but we have no reason to suspect any major irregularities.

L295-303 *Many details on parameter choices of this paragraph can be added to Table 1. See below* *comment on Table 1. In this way, I sense that this paragraph can be reduced.*

We agree and will extend Table 1 accordingly.

L317 *The iterative increase of this relaxation parameter seems very fundamental in terms of methodology. I therefore urge you to include it in the main method also covering the synthetic setup.*

While being beneficial to the approach in that it aids convergence, iteratively increasing β is not vital for the method and does not improve the results for the synthetic ice cap. But following the suggestions on restructuring the manuscript, we added a paragraph on this in the new section 'Stabilization techniques'.

L342-343 *In my view membrane stresses are in general captured by the hybrid SIA-SSA ice-dynamic variant in PISM. It might be that the PISM strategy to merge SIA and SSA anyway suppresses this effect. Please be more specific and rephrase.*

In paragraph 43, Bueler and Brown (2009) state that "[in] the nonsliding case, none of the driving stress is held by the membrane strain rates. In the latter [nonsliding] case the SSA predicts no flow so we propose that the SIA should "take over" and predict flow by shear in vertical planes." In our understanding of the SIA-SSA hybrid, the SIA alone controls ice flow in the absence of basal sliding. This is explicitly stated in our manuscript L.295-296.

L348-365 *Here you introduce a post-processing correction in the middle of your results section. This comes as a surprise to the reader. I suggest that you rather introduce it as an optional post-processing filter in the methods. I am sure that, there, it can presented more concisely.*

As discussed above, we do think that the post-processing step better fits here as it is not at the heart of our new inversion methodology. However, we will integrate this paragraph better in the surrounding text in the revised manuscript so that it hopefully does not come as a surprise to the reader anymore.

L366-376 *This entire paragraph has a discussion character and does not fit into the results section. Please adjust according to my main comment on the manuscript structure.*

In the revised manuscript, the discussion section covers many aspects generally related to our inversion method (much more extensively than in the original manuscript, thanks to the reviewer´s comments) which, after all, is the novelty of our paper, while Kronebreen is only a test case. Discussing the mass balance sensitivity of Kronebreen would therefore not fit with the general focus of that section. So instead, we will pack this paragraph into a new sub-section of the Kronebreen results which we hope also contributes to an overall clearer structure of the manuscript.

L372 *'too high' -> 'too low*

Thanks for pointing that out!

1.4 Figures and Tables

Generally, the figure quality can be improved to better guide the reader to the important details by structural and visual re-formatting.

Fig.2 *I think this figure tries to serve two purposes. First it presents a schematic of the iterative approach. Second it introduces the synthetic ice-cap setup. I would split these two aspects in two individual figures. The introduction of the synthetic ice cap is the second figure and it should be formatted similarly to the present Fig. 3 (see respective comments). I would also not blend input fields and results in one figure. I therefore better like your presentation of the real-world setup in Fig. 3 and Fig. 4. Please try to present the synthetic setup analogically. For the results of the iterative figure (currently Fig. 2b), I urge you to also show iterative reduction of the mismatch between observed and modelled velocities and ice thickness. This is more easily done once you present the results as an individual figure. The velocity mismatch should decrease during the friction updates. I wonder what happens during the subsequent 1000 iterations of bed updates.*

We follow the reviewer´s suggestions and re-structure the content of Fig. 2. Specifically, we will add a schematic representation of the iterative approach to Fig. 1, and divide the inputs and results of the ice cap experiment in two separate figures. In that second figure, we will also plot the velocity mismatch. However, the thickness mismatch is exactly the same as the bed mismatch, so we do not think that it is necessary to show that.

Fig.3 *I miss the thickness observations in this figure. Please add another panel.*

Will do.

Fig.4 *Please add the requests on Fig. 2. Furthermore, can you rather show relative thickness errors instead of absolute bed errors in panel c. This will facilitate the assessment of the importance of these differences. In addition, I request that you do not focus on this central location for the velocity comparison but rather show an extra figure covering the entire Kronebreen catchment showing modelled and observed velocities. Then the reader can better assess your velocity results, which are otherwise not presented. The latter could be an extra figure.*

As above, we will also show the velocity misfit per iteration.

Showing relative thickness errors instead of absolute ones is problematic since it significantly inflates errors where the true ice thickness is small. A model that produces an ice thickness of 10 m when the true ice thickness is 20 m is quite good, but will appear very bad if a relative comparison is made. To adequately judge the capabilities of an inversion and allow the reader to identify where the results are good and bad, we therefore find it much more informative to show absolute bed errors.

On adding a figure with observed and modelled velocities, we would like to point out that the velocity results are already shown in the present figure, namely in panel b) where a comparison between modelled and observed velocities is made. In the interest of conciseness, we therefore do not think that it is necessary to add another figure.

Fig.5 *What do the colours mean in panel c? To me it would make sense to have panel c also presented for the synthetic ice cap. This could serve as a baseline for an ideal setup and help to assess your approach.*

The colors show represent point density to indicate where the majority of observations is found. Panel c will also be added for the synthetic ice cap.

Table 1 *Please extend this table to cover all experimental setups (ice cap, Kronebreen and potentially Austre Grønfjordsbreen). In this way, comparison is facilitated (also see comment to L295-302).*

Will do.

References

Bueler E, Brown J. 2009. Shallow shelf approximation as a "sliding law" in a thermomechanically coupled ice sheet model. Journal of Geophysical Research: Earth Surface 114. doi: 10.1029/2008JF001179.

Jouvet G. 2022. Inversion of a Stokes glacier flow model emulated by deep learning. Journal of Glaciology :1–14doi:10.1017/jog.2022.41.

Kochtitzky W, Copland L, Van Wychen W, Hugonnet R, Hock R, Dowdeswell JA, Benham T, Strozzi T, Glazovsky A, Lavrentiev I, Rounce DR, Millan R, Cook A, Dalton A, Jiskoot H, Cooley J, Jania J, Navarro F. 2022. The unquantified mass loss of Northern Hemisphere marine-terminating glaciers from 2000–2020. Nature Communications 13:5835. doi:10.1038/s41467-022-33231-x.