We would like to thank both reviewers for their careful work and encouraging comments. We reply to all comments below. Revised text in the manuscript is highlighted with red.

Reviewer #1

General Comments:
In previous work, Born derived and demonstrated a model of ice flow focused on the evolution of layer packages (Born, 2017). An explicit output from this model is the isochronal layer geometry – a field that is directly comparable with radar observations, which have the potential to provide a spatially and temporally comprehensive check on model performance. The formulation presented in (Born, 2017) outperformed existing models which use Eulerian velocity fields to contour the age field of the ice sheet, avoiding issues of numerical diffusion that can result in unrealistically smooth age fields. That work forms the backbone of this manuscript, which is focused on making that framework modular, such that it can be applied to existing 3D models of ice flow and allow for the use of observed englacial layers in model tuning.

At its core, this is a methodological paper, pursuing an important objective at the cutting edge of ice sheet modeling. But the authors spend most of the paper discussing the specifics of their model results – what drives model-layer/observed-layer mismatch and the interaction between specific tuning parameters in Yelmo (the underlying ice physics engine (Robinson et al., 2020)). This would be important if the tuned model (i.e. the depth-age model of Greenland since LGM) were the central product of this work, but the real scientific contribution here is what the authors have learned about the process – that (1) it is possible to apply the layer tracking scheme in (Born, 2017) to 3D models that do not explicitly track layers, that (2) the resulting layers are an improvement over results of previous methods, that (3) tuning ice flow models (or at least, this ice flow model) to the ice thickness alone can result in large errors in englacial dynamics, and (4) that it is important that future models use this method, as Eulerian tracers produce a systematic bias in model-layer age.

While I have only minor questions about the technical work done, the changes that I think are most necessary are to the writing, to maximize the paper's impact and ensure that the scientific contribution of the work is clear. Right now, the key messages are buried in extensive description of Greenland accumulation, and the large, multi-panel figures of model mismatch do little to articulate this work's core message. In the technical comments below, I provide specific changes that I think will help resolve these issues. Ultimately, the layer tracking module developed here has the potential to be a widely used tool and help constrain models across a wide range of complexities, and I want to ensure this work has the impact it deserves.

If this were simply another model of Greenland from LGM to today, the scientific contribution would be limited, as there is no articulated "experiment" here probing Greenland dynamics. The discussion of errors in model forcing provides insight into the climate parameterizations chosen, but they distract from the methodological improvements that will be this paper's legacy. To make clear the scientific contribution, I think three primary changes are required:

We are grateful for the detailed and constructive criticism of this review. We agree with the comments and implemented almost all of the suggested changes.

1. There should be a reproducible description of how the layer tracing scheme couples to the 3D model. There is extensive description of the climate spin up, and the model parameters being tuned, but no description of the implementation that translates output from (Robinson et al., 2020) to input in (Born 2017). This should (1) make clear to the reader exactly how this method avoids the pitfalls of Eulerian tracers, especially while using an ice sheet model that solves the physical equations on an Eulerian depth grid, and (2) enable future application of the method to other ice sheet models.

We extended section 2.4 with additional text and one additional figure that illustrates the algorithm during one time step, what input data it requires from the host model, and when.

2. A more succinct description of the optimal model should be provided, but primarily to highlight which model parameters are sensitive to the stratigraphic constraint (indicating which processes / boundary forcings this optimization approach is likely to capture). The extensive description of figures 4, 5, 10, and 11 can be substantially trimmed. In addition, I think the readers would benefit from a deeper explanation of the differences observed and the drivers of that difference in figure 12, which demonstrates the value of the improved parameter optimization.

We agree that the description of the composite analysis was too long. This text and the former figures 10 and 11 have now been moved to an appendix. The main findings are summarized in section 3.2 as a single paragraph. We believe that the discussion of former figures 4 and 5 is already quite compact and important as an introduction to how the simulated isochrones can be compared with the reconstructed stratigraphy. We think this is important to build intuition before discussing the RMSE of the ensemble simulations and therefore prefer to keep section 3.1 unchanged.
At present, we cannot provide additional detail on the sensitivity of individual model parameters. We were surprised by the apparently weak sensitivity of the simulated isochrones to dynamic model parameters, which appears to contradict findings from Born (2017). We suspect the optimization of basal friction to have a strong influence on this result, which is part of the learning process that Reviewer #1 refers to above. This explanation is now included at multiple locations in the manuscript and we plan to revisit the issue in a future study.
Lastly, the description of former figure 12 (now 11) has been revised to make it clearer. Although that did not make the text much longer, we hope that the now more precise description more clearly conveys the added value of our new layer advection scheme. See also our reply to the specific questions below.

3. At present, this paper avoids the discussion of an important and active area of research: fitting layer shapes in the dynamic regions of Greenland and Antarctica. This is in-part because the outlet glacier modeling done here is simplified. But layer fitting in these areas has the potential to capture spatial and temporal heterogeneity in the basal boundary condition that no other method can address, and given the high-profile nature of features like the layer draw-down in Northeast Greenland (e.g., Fahnestock et al., 2001) and the complex folding at Petermann Glacier (e.g., Bons et al., 2016), I think it would be appropriate for this work (especially given its title: “Modeling the Greenland englacial stratigraphy”) to directly address dynamically controlled folding which dominates the marginal ice. There is an extensive literature on the Weertman Effect (e.g., Hindmarsh et al., 2006; Leysinger Vieli et al., 2007; Parrenin et al, 2007; Wolovick et al., 2016), models of layer shape in Greenland (Leysinger Vieli et al., 2018), and direct comparison of models

and data (e.g., Holschuh et al., 2019) that could be used to substantiate the need and interest in developing these layer modeling methods. While you could never provide a complete review of the literature here, the repeated claim that accumulation history is the dominant variable relies on an implicit assumption that we ignore the dynamic outlet glaciers. A full description of your method should include its applicability to the ice sheet margins and how it fits within the existing literature on the subject.

We revised the last section to discuss how we can constrain the ice dynamics and limitations of the current method. This discussion includes references to most of the studies mentioned above. Note that in its present form, our isochrone advection scheme uses a vertical axis where age has to increase monotonically. This does not allow for layer folding.

Line-Item Corrections:
Page #: 1
Line #: 6-8
I find this description confusing, as mass transfer happens within Yelmo (outside of the layer evolution scheme). How is it that mass transfer between layers is avoided when the solver exists in depth, not time? (I think this is just part of a larger desire to see a clear description of the coupling).

We rephrased this part of the abstract slightly and hope that the revised description in section 2.4 clarifies this question.

Page #: 1
Line #: 10
The phrasing "... selecting simulations..." is unclear here -- selecting them for what? Perhaps rephrase to "Using an ensemble of simulations to optimize climate and ice dynamic parameter selection, we show that direct comparison with the dated radiostratigraphy data yields notably more accurate results than choosing parameters based on fit to total ice thickness alone."

We have changed the verb to “calibrating”.

Page #: 1
Line #: 16-22
I appreciate the oceanographic analogy here, it adds nice context!

Thank you.

Page #: 1
Line #: 21
"The proverbially glacial flow" -- I'm not sure what you mean here by "proverbially".

This is a failed attempt at a play of words. We were trying to emphasize the relative slowness of ice sheet dynamics with the figurative sense of the word “glacial”, while naturally also referring to “glacial flow” in the literal sense. The word “proverbial” has now been removed from the text.

Page #: 2

Line #: 28-31
There is an error in construction here, with the sentence that reads: "The ... layers could aid..., where to find..., to reconstruct..., or to determine...." Either each clause should start with an infinitive, or they should follow from a common verb. It could be "The layers could aid, find, reconstruct, or determine", or it could be "The layers could help us to select, to find, to reconstruct, or to determine".

This has been corrected.

Page #: 2
Line #: 36
Should be "finite-difference" not "finite-differences"

corrected

Page #: 3
Line #: 63-64
Given that your layer thicknesses are smaller than the vertical resolution of the solver, I am still a bit confused about how you can solve for changes in layer thickness and still guarantee no numerical diffusion? (This is where a discussion of the coupling would be helpful).

Please see section 2.4.

Page #: 0
Line #:
Section 3.2
I had difficult following the narrative through this section, especially Figure 10 and 11. If you intend to keep all of this, it would help to guide the reader through it in a more directed way -- referring to specific panels in the figures (not just a grid of 40 Greenlands), and pointing back to the motivating questions that justify the extensive description provided. Ultimately, I did not see any need for detailed description of the specific model output you provided, as there are more sophisticated modeling exercises that one could turn to for full description of dynamics in Greenland. But if you think there is value in dissecting the specifics of this model configuration and output (as opposed to just focusing on the exercise of modeling and optimizing), you need to motivate that more clearly somewhere.

We agree that 2 * 40 Greenlands is too much for the main body of the paper but would like to keep this figure as a reference for interested readers and therefore suggest moving it to the appendix. We think that a broad discussion of physical properties like isochrone depth is useful in addition to the RMSE and so we include the most important findings from the full discussion as a summary at the end of section 3.2.

Page #: 19
Line #: 315-329
This section had me thinking about a more general question -- do isochrones add value in constraining processes outside of the time range that they span? Making an explicit statement about how temporal coverage of the data impacts temporal constraint in the model could be very interesting.

Yes they do. Additional precipitation during the glacial period leads to a higher ice sheet and steeper surface gradients. Additional accumulation that falls on top will therefore experience a faster advection and hence more dynamic thinning. This is clearly seen in the reduced depth of the 11.7 ka isochrone when comparing the high f_{LGM} composite with the ensemble average (Fig. A2).
We included this explanation in the greatly shortened discussion of the ensemble composites in the main text.

Page #: 19
Line #: 345
Space between numbers and units.

corrected

Page #: 20
Line #: 350
You regularly state that the Eulerian age tracer (orange curve) in Figure 12 shows older ages, but in all situations it seems that the age of the orange curve falls below the blue curve. Am I misreading Figure 12? "Older" continually appears in your description of the Eulerian method, and I am having trouble rectifying that with the figure.

This section was not as clear as it could be and we made several changes to the text to improve it. There are two “Eulerian curves in figure 11 (new numbering)”. The dashed blue curve shows the Eulerian tracer diagnostic for the $BEST_{all}$ simulation, i.e., the Eulerian diagnostic for a simulation calibrated with the isochronal scheme. It can be used to directly compare the quality of the Eulerian age tracer with our new scheme (blue solid). Here, ages are older because of upward numerical diffusion. In addition, the orange curve shows another simulation that was calibrated using the Eulerian age tracer, but shows the age profile as simulated by the more reliable isochrone scheme. Again, this data can be compared directly with the solid blue curve.

Page #: 20
Line #: 351-353
A clear description of the coupling will certainly answer this question, but somewhere depth and age must be mapped to one another to couple the ice flow model to the layer evolution model, and I'm still not clear on how the horizontal flow speeds within a given layer are calculated (to prevent flow across boundaries).

We hope that the revised section 2.4 answers this question. The horizontal flow speeds within a given isochronal layer are interpolated from the coarser Yelmo grid. This means that they are calculated in the Eulerian host model and subject to numerical diffusion. However, the velocity field varies rather smoothly with depth and the main variable that controls ice viscosity, temperature, is strongly influenced by *physical* diffusion that significantly lessens the impact of its spurious numerical counterpart.

Page #: 20
Line #: 357-358
Okay, I think I understand the "older" comment here -- if a model were optimized

using the Eulerian scheme, the true model age (when calculated correctly using the new layer evolution scheme) would actually be older than the constraint. But that seems different than the previous statement, that the Eulerian tracer data produces older ages. I am probably just confused, but some clarity through this whole section on the nature of the bias of the Eulerian method would be useful.

This is correct. All simulations simulate both the new layer scheme and the Eulerian age tracer. We only show the latter once, as the dashed blue line in figure 11, and we discuss a single simulation that was calibrated using the Eulerian tracer, shown as orange curves in figures 9, 11, and 12. Unlike $BEST_{all}$ and $BEST_{ice}$, this simulation does not have a specific name.

Page #: 21
Line #: 378-380
I think this sentence is a bit of a tautology -- the model calibrated to the stratigraphic data fits the stratigraphic data better. It would be better to appeal to a third target variable to evaluate accuracy. Something like: "Models that are optimized to match the ice thickness require unrealistic precipitation histories, resulting in erroneous layer ages. These precipitation histories can be ruled out when constraining model parameters with both thickness and layer age."

We believe that the original text was accurate and not stating an obvious fact because it referred to the finding that the simulation calibrated with the stratigraphic constraint also fits the *ice thickness* better. However, we appreciate the suggestion to rephrase the text and would like to include an only slightly modified version.

Page #: 21
Line #: 381-386
This paragraph is only true because you exclude the dynamic regions of Greenland from your analysis. Of course accumulation matters more when dynamic vertical velocities are otherwise very small. Where the Weertman effect is large, surface mass balance history will be much less important. This is why I advocate for a broader discussion of what affects layer shapes near the margins, contextualized in the literature.

We modified the first and last sentences of this paragraph and revised the part of the discussion that follows it. Please see our reply to general comment #3 above.

Page #: 23
Line #: 423
Again, this point that "accumulation eclipses the impact of dynamics" is not universally true, and will depend on the target region of interest for future applications of this method. In the outlet glaciers, this is far less likely to be the case, so I would hesitate to use such strong language when this method could be applied beyond just the interior of ice sheets as was done here.

We agree that this statement was too strong and modified it accordingly.

Reviewer #2

This model presents a modeling of the Greenland ice sheet comprising a module for solving the age equation, which allows to compare the modeled isochrones with radar-observed isochrones. This age module is derived from the study of Born (2017), but it is now decoupled from the thermo-mechanical model used. The work of Born (2017) demonstrated how this new numerical scheme, based on the time domain, outperforms classical Eulerian schemes which show large diffusive artifacts.

This study presents itself as "the first three-dimensional ice sheet model that explicitly simulates the Greenland englacial stratigraphy". The ice sheet model used here is YELMO. The climate forcing used is based on two snapshots (Present-day and LGM) and a climate index based on paleoclimatic archives. A pseudo inverse method is used to fit a few parameters so that the model best fits either the present-day topography, the internal layering, or both. It is found that the model fitted onto the englacial stratigraphy gives a better overall fit than the model fitted onto the surface topography. So it is "easy" to have a model that fit the surface topography for wrong reasons.

The article is generally clearly written, and is a good contribution to the field of ice sheet modeling and comparison to observations. It certainly opens a new chapter of model-data comparison in ice sheet modeling.

My main comments are similar to the remarks of the first reviewer: I reckon that some sections are too detailed while others are not enough, so that the main output is too diluted to be easily accessible. I feel that the method section of the article describe in too much details the YELMO ice sheet model used and its climate forcing, instead of focusing on the new age numerical scheme and how the variable are transferred from the thermo-mechanical grid to the age grid. Then, the results section describe in details the fit of the model to the observations, while the inverse method used here is really basic and does not allow to explore the full parameter space. I would rather focus on the comparison of this new age numerical scheme with previous numerical schemes, like the Eulerian and semi-Lagrangian schemes. In short, I reckon the value of this article is more on the age numerical scheme than on the inverse method.

We thank Reviewer #2 for this evaluation of our work. We agree that the primary contribution of this study is the method itself rather than the results that we were able to obtain from it so far. It was our intent to clearly outline what we have learned to this point and how future work may improve upon ours. In our understanding, Reviewer #2 raises two major points: 1) A lack of detail in the description of the isochronal tracer advection scheme, in particular in comparison to the detailed description of the host model Yelmo, and 2) an incomplete description of how our methodology is superior to alternative numerical schemes.

We addressed the first point by extending the technical description in section 2.4, including a new figure showing the execution flow during one time step. We would like to keep the description of Yelmo in its current form, because the model is relatively new and many of the technical details have not yet been published.

With regard to the second point, we revised section 3.3 and the comparison with the Eulerian age tracer. The comments of Reviewer #1 made us realize that this part was not particularly well written and important points likely to be misunderstood. We chose not to extend the comparison with the semi-Lagrangian scheme, because we cannot assess it directly and therefore cannot say much more than what is already included in the discussion: “ Compared to Lagrangian or semi-Lagrangian tracer advection schemes, our method avoids costly interpolation or low particle densities. It is possible to use an uneven spacing of the isochronal grid to concentrate computational cost to key periods of interest.”