Response to the reviewer 1

August 17, 2020

Summary and High Level Discussion

This paper is about using adjoint calculus to determine the sensitivity of ice sheet surface velocities and elevation to perturbations in basal friction and basal topography. The ice sheet models are the full Stokes and Shallow Shelf Approximation coupled with a time-dependent advection equation for the kinematic free surface. The authors propose a few test cases with both numerical and analytic solutions to the underlying forward and adjoint equations and argue that it is necessary to include the time-dependent advection equation for the ice surface elevation into the models. The reported findings show that: 1) there is a delay in time between a perturbation at the ice base and the observation of the change in elevation, 2) a perturbation at the base in the topography has a direct effect in space at the surface above the perturbation and a perturbation in the basal friction is propagated directly to the surface in time, and 3) perturbations with long wavelength and low frequency will propagate to the surface while those of short wavelength and high frequency are damped.

The topic of the paper is very interesting and it is worth publishing. However, it needs a serious revision. The content as is presented is very difficult to digest. Below I list several specific comments/recommendation.

Comments

All the equation and section numbers refer to unrevised version of this paper.

1. Introduction

   (a) It is not entirely clear from the introduction (and abstract) what the motivation for running a sensitivity analysis is. It would be great if the authors could motivate this study and perhaps emphasize the impact of the sensitivity study results (in the intro especially) explicitly.

   Response: We have expanded on the motivation for the sensitivity analysis and the relation to the inverse problem. For example, some perturbations at the base induce small perturbations at the surface.

Such basal perturbations will be difficult to detect from observations at the surface. Consequently, these perturbation will not appear as a result of an inverse optimization based on surface data.

(b) It would be beneficial to discuss the companion paper (Cheng and Lötstedt, 2020) in more detail; in particular, what is the novelty in this paper compared to the previous one? If this companion paper would be useful for the reader to help him/her understand the (heavy) modeling part in this paper, it would be great to state this earlier or explicitely. Are some of the derivations in the Appendix also done in Cheng and Lötstedt, 2020? If so, perhaps the authors don't need to repeat these here.

Response: A short review of Cheng and Lötstedt, 2020, is written in the end of Introduction. Only the results of the derivations in Appendix are given in the companion paper with a marginal overlap of contents. It is noted in the end of Section 3.1.3 that the computed adjoint ψ is very small in Cheng and Lötstedt, 2020, as expected by theory in the section. References to the companion paper concerning the SSA and FS equations are made at Conclusions (i) and (v) and after Conclusion (vi) in Section 3.2.3. The conclusions agree in this paper and the companion paper.

(c) In lines 18-19 on page 2, I would like to suggest the following reference for the inversion for the geothermal heat flux as well: Zhu, H., Petra, N., Stadler, G., Isaac, T., Hughes, T.J.R., Ghattas, O.: "Inversion of geothermal heat flux in a thermomechanically coupled nonlinear Stokes ice sheet model". The Cryosphere 10, 1477-1494 (2016).

Response: Thanks for the reference. We refer to it now in Introduction.

2. How is $h(x, t)$ initialized, i.e., how is $h_0(x)$ defined?

   Response: h_0 could be any smooth function such that $h_0 > b$. The inequality is added.

3. Are there any constraints on C in equation (4)? For instance, does it have to be positive? From line 13 it appears so. If this is the case, how are the authors making sure that this constant stays positive during inversion?

   Response: C is non-negative (changed now). We are not solving the inversion problem but since it is an optimization problem one can add inequality constraints $C \geq 0$ to be satisfied at the optimum in an optimization algorithm. There are numerical techniques to do that.

4. How are the Dirichlet boundary conditions set/defined, i.e., how are u_u and u_d set?

   Response: The assumption in the adjoint problems is that the inflow velocity u_u and the outflow velocity u_d of the ice are known. If the domain is at the ice divide then $u_u = 0$ and u_d is the velocity at the calving front.

If the latter is unknown, it can be computed with a different boundary condition there e.g. like (9) and then use u at Γ_d in the adjoint equations.

5. What is H in equation 7? I assume this H is the height as shown in Figure 1a, please clarify.

   Response: It is defined in the beginning of Sect 2.1 but we repeat it after (7).

6. In line 15, page 6, the authors state: "friction coefficient $C(x,t) \geq 0$, just as in the FS model. For the FS model it looks like $C > 0$, please clarify the possible equality here.

   Response: We have changed the FS coefficient to $C \geq 0$.

7. Second row, page 7: It is not clear how the adjoint equations have been derived. The authors say "Lagrangian of the forward equation?" (same in line 5, page 8). Do the authors mean the Lagrangian of the optimization problem governed by this PDE? What is the optimization objective function in the Appendix?

   Response: The inverse problem is an optimization problem to find the optimal b and C. In the sensitivity problem, δb and δC are known and we want to know the effect of these perturbations at the surface. In the inverse problem, we determine δb and δC iteratively such that $u + \delta u \to u_{\mathrm{obs}}$. The relation between the inverse and sensitivity problems is discussed in an extended Section 3.1.2. The Lagrangian $\mathcal{L}$ for the sensitivity problem is defined in the appendices for SSA in (A4) and for FS in (A15). It differs from the Lagrangian of the inverse optimization problem in $F(u,t)$ in (10) (see the comment in the beginning of Section 3.1.2 and Section 2.2.5 in Cheng and Lötstedt 2020). The adjoint equations are the same except for the F_u and F_h terms. A comment about this is added after (13). Examples of different F-functions are found after (11).

8. Line 9, page 7: Need to define the topography $b(x)$.

   Response: The correction has been made.

9. Line 10, page 7: Please reformulate "its forward solution . . . ", it is not clear what solution we are talking about here. Same for the adjoint.

   Response: 'its' is replaced by 'the'

10. What do the authors mean by "The same forward and adjoint equations are solved both for the inverse problem and the sensitivity problem but with different forcing function F", does this difference is due to inversion versus sensitivity or due to the fact the the objective is different for the two? In fact it is not clear how F is chosen for inversion versus sensitivity study. The authors gave a few examples for F but did not specify if F is or must be different. Same statement is made in line 5 on page 11 and similarly in line 5, page 25.

Response: This issue is discussed also in Comment 7. A better explanation of the difference is found after (12) and in Section 3.1.2.

11. The last 2-3 lines on page 7 need to be explained more clearly. It sounds like there is an optimization/minimization problem solved, if so, what is the gradient? How is this optimization problem solved?

    Response: The optimization problem needs repeated solutions of the forward and adjoint problems to iteratively reach the minimum. The estimate of the sensitivity of perturbations is achieved by one forward and one adjoint solution. This is elaborated on before Section 3.1 and in Section 3.1.2.

12. How is the nonlinear Stokes solved?

    Response: They are solved by Elmer/Ice as we have written in a paragraph after (18).

13. It would be beneficial to state the Lagrangian somewhere in the main text in order to help the reader follow the derivations and given expressions. This seems to be given in A15 for the Full Stokes, perhaps this should be moved to the main text.

    Response: The Lagrangian of the FS model is now also defined in the main text.

14. Line 16, page 9: Why do the authors consider e^i?

    Response: We use the same unit vectors in Section 3.1.2. If we are interested in δu_i, $i = 1, 2, 3$, at (x_*, t_*) then

$$F_{\mathbf{u}} = \mathbf{e}^i \delta(\mathbf{x} - \mathbf{x}_*)\delta(t - t_*)$$

15. The effect of the perturbations seems to be local. How do the authors choose where to induce these perturbations?

    Response: The perturbations δb and δC can be local in space and time or more extended in space and time. The effect at the surface is determined e.g. by the the formula (16). Since it is a linearization, it is important that δb and δC are small. A comment about that is inserted before Section 3.1.1. The perturbations at the surface are always registered in one point in space and time but it can be chosen arbitrarily by changing $F(u, h)$.

16. In general, it is difficult to follow all the variables, it would be great if the authors would remind the reader what is what. For instance I am not sure what the 'perturbation δu_1' is (in the discussion for Fig 2 on page 10), is u_1 perturbed, or is it the effect of the perturbation in C or basal friction on the velocity component u_1?

    Response: This is clarified now.

17. Please define exactly what "variation $\delta\mathcal{F}$ of the inverse problem" means? Similarly, what does the "variation of a functional" mean (e.g., in line 3, pag. 12))? Are these directional derivatives? It would be beneficial to show the mathematical definition in general and then apply it.

    Response: This is explained in words now in Section 3.1.2. $\delta\mathcal{F}$ tells what happens to $\mathcal{F}$ when b and C change to $b + \delta b$ and $C + \delta C$.

18. It is not clear how equations 22 and 23 are related.

    Response: The paragraph has been reorganized with additional text.

19. Line 9, page 18: What do the authors mean by "The relation in (38)...can also be interpreted as a way to quantify the uncertainty in u"? Please be more precise and define mathematically what you mean by "uncertainty". Same discussion needs more details in line 6 on page 25 and also in lines 5-6, page 3.

    Response: Examples of uncertainties are given on page 3. They are known measurement errors. New text to explain uncertainty and a book on uncertainty are included on page 9. 'uncertainties' $->$ 'perturbations' on page 25.

20. In general, this paper is difficult to follow. Perhaps the authors can add some roadmap to the beginning of each section to guide the reader a bit through the research and findings. For instance I had to write out the sections to see how everything fits together because it got a bit impossible to navigate through so many setups and subsections. The structure seems to be the following:

```
1. Introduction
2. Ice Models
2.1 Full Stokes
2.2. Shallow shelf approximation.
3. Adjoint equations
3.1. Adjoint equations based on the FS model
3.1.1. Time-dependent perturbations
3.1.2. The sensitivity problem and the inverse problem.
3.1.3. Steady state solution to the adjoint elevation equation in two dimensions.
3.2. Shallow shelf approximation
3.2.1. SSA in two dimensions.
3.2.2. The two-dimensional forward steady state solution.
3.2.3. The two-dimensional adjoint steady state solution with $F_u \neq 0$.
3.2.4. The two-dimensional adjoint steady state solution with $F_h \neq 0$.
3.2.5. The two-dimensional time dependent adjoint solution.
```

    (a) Sometimes the titles are not very representative or consistent, for instance Subsections 3.2.1 and 3.2.2 focus on forward equations and solutions eventhough Section 3.2. is called “Adjoint Equations”, this is a bit confusing. Perhaps the authors should move forward problem matters to section 2.

Response: Thanks for the comments. We have moved the forward SSA solutions in 2D to Section 2 .

(b) Also, consider creating a table that summarizes all the examples and cases, shows the similarities and differences, parameter values, etc. and then refer back to this table from the sections and text. It is difficult to see the big picture with all the small subsections and various proposed scenarios.

Response: We will write more about the MISMIP example and refer to that.

(c) The description of adjoints and problem setups are mixed with results. I recommend separating these to the extent possible.

Response: We will make the separation between theory and examples more explicit.

(d) Finally, there are several modeling information and parameter values inserted in the text which makes the reading of the actual research study and findings difficult. A table that summarizes somehow all these values might help to ease the discussion.

Response: We will collect parameters in a table together with the description of the MISMIP test case.

21. Line 1 page 25, not sure what the point of the sentence "...confirm the conclusions here and are in good agreement with the analytical solutions." is here. Please add more details to explain.

    Response: There is a brief summary of the paper now in Introduction. Reference is made to it here and there in the text and in the first paragraph of Conclusions.

22. Finally, the authors talk about sensitivity analysis, however throughout the paper the authors compute the effect of some perturbation in the parameters on some quantity of interest. To do a proper sensitivity analysis (or derive the sensitivity equations) one should look at the (total) derivative of the objective with respect to the parameter (of interest). This will give the equations to compute the sensitivity of the forward solution with respect to the parameter (or in finite dimensions to all the parameter components), etc. The authors should define clearly at the begining what they mean by “sensitivities” and how are these computed.

    Response: We mean pointwise sensitivity in space and time at the ice surface. There is great freedom to choose another type of sensitivity by changing F in (10). This is remarked in Introduction and before Section 3.1 there are examples of different F.