Review of “A finite element approach to ice sheet balance velocities”

This revision is much improved, there is still a bit of a lack of rigor on the mathematical description of the method that can be easily addressed. The authors have added an entire new section to the paper, I am not sure if this adds a real value to this work, but it illustrates the use of the algorithm in the specific case of model initialization applied to the Greenland Ice Sheet. I am not really sure how relevant the new section is, as the changes in thickness are stronger than measurement errors, the depth average velocity loses important features, and the correction applied to the apparent surface mass balance does not seem to have any physical meaning.

1 Main points

The authors keep referring to the driving stress direction smoothing procedure as a way to account for “longitudinal stresses” and this is not really obvious to me. I have been scratching my head over this and I think I just don’t get it. I don’t understand, for example, why this method would account for longitudinal stresses and not lateral drag as the smoothing is done isotropically (not just along flow lines).

2 Minor points

The beginning of section 2 is a bit misleading. If it is true that $\nabla \cdot \boldsymbol{v} = 0$ is a property of incompressible fluids, the boundary conditions (eq. 2 and 3) are specific to ice sheets only. Also, $\boldsymbol{u}$ in eq. 1 is 3-dimensional, whereas only the horizontal components are present in $\boldsymbol{u}$ of eq 2 and 3 (i.e. it is 2-dimensional. It would be convenient to just use $\boldsymbol{u}_{\parallel}$.

In eq. 4, since $\bar{\boldsymbol{u}}$ is 2-dimensional, so is the divergence operator $\nabla\cdot$ that is exactly the same as the one in eq. 2 and 3 (no need to include $\parallel$ in the notation).

At the end of section 2 page 5, it is said that “no boundary condition need be specified”. This statement is confusing because there is always a boundary condition (even though it might be a zero Neumann like eq. 9). It should just be said that no Dirichlet boundary condition needs to be applied.

In the beginning of section 3, it is said that "The operator appearing in Eq. (8) is self-adjoint". This is also a confusion, I think the authors refer to the bilinear operator of the weak formulation associated to this term of the equation. Just remove "self-adjoint". The operator is not discretized in the Finite Element Method, only the space of solutions, so the sentence should say something like "Eq. (8) is solved using the Finite Element Method".

After eq. 11, the Hilbert space ${H^1}_0$ is not defined. The index 0 is not standard and generally means that $\boldsymbol{\varphi} = \mathbf{0}$ along $\partial\Omega$, and this is not what you do here according to eq. 9.

After eq. 12, I am not sure I agree with the definition of the space V, since λ appears in a divergence operator (eq. 12), its derivative should be integrable, so it should be a space $\mathcal{H}^1(\Omega)$

After eq. 13, the authors are mixing continuous and discretized forms (so far, everything is in continuous form). λ is in V, so it should just be said that triangle linear finite elements are used.

In section 4.2 I agree that we cannot compare $\bar{\boldsymbol{u}}$ to $\boldsymbol{u}_s$ but simple tests such as $\|\bar{\boldsymbol{u}}\| < \|\boldsymbol{u}_s\|$ can be performed. I would not be surprised if the results do not look good...

In the first part of section 5, I think there is a confusion between PDE-constrained optimization, Cost function, Lagrangian and Lagrange multiplier. Eq 14 is the Lagrangian associated to the minimization of $\mathcal{I} + R$. What is minimized is $\mathcal{I} + R$, not $\mathcal{I}'$! We are not looking for the minimum of $\mathcal{I}'$ but its saddle point. Given that the authors are using automatic differentiation, I would just remove this entire section as automatic differentiation provides the gradient of $\mathcal{I} + R$ without the need to introduce a Lagrangian.

Eq 16 and 18, λ should be inside the integral.

Below eq 21, these are the "spaces" of the control variables. The "state" is actually $\bar{U}$.

"$\boldsymbol{N}$ are estimated from smoothed steepest descents", steepest descent generally refers to a minimization method, I think "driving stress directions" would be more accurate.

The equation between eq 23 and eq 24 is really unclear, and again, since you are using automatic differentiation, you should just remove this part.

"the transitions between areas of no data and data are free of gradients" are seamless would be less confusing (there are lots of different gradients in this paper).