Interactive comment on "Towards a better ice sheet model initialisation and basal knowledge using data assimilation" by Cyrille Mosbeux et al.

Answer to *S. L. Cornford (Referee)*

This paper addresses a well-known issue in ice sheet modelling: the fact that even if a model can be tuned to give a very good match to observed velocity, errors in the ice thickness result in implausible thickening rates. There is widespread agreement among ice sheet modellers, especially those concerned with near term simulations, that accurate maps of ice thickness (and so bedrock) are needed. Ideally, of course, these things would be measured to the required resolution and precision, but given the distance between airborne radar tracks, the difficulty in making such measurements to coincide with measurements of ice velocity, and the inevitable approximations in ice flow models, methods of the sort described here, where ice thickness is estimated at the same time as basal traction (and also englacial viscosity), are also required.

The authors try out two approaches in the context of a synthetic inverse problem, where they attempt to recover known fields . In one (ATP) they attempt to optimize the basal traction and thickness simultaneously, using a gradient based method. In the other (ANP) they construct an iterative procedure, alternating between optimization of the the basal traction and evolution of the bedrock through the mass conservation equation with an extra a source term. Both methods appear to work well in regions of fast flow. Overall, I suggest publication with some minor alterations.

We thank the referee for his useful and constructive comments on our work.

Specific Comments

L22/23 and elsewhere: Inverse methods? Although this phrase seems to be gaining in popularity, it does not tell us much. I'd prefer more specific terms, e.g gradient-based optimization methods (used to solve an inverse problem)

We agree that the term "gradient-based optimization methods" is more accurate in the sentence lines 22/23 and it as been changed. Other occurrences of the term "inverse methods" have not been changed as they are used as a generic term to refer to our 2 methods (ATP and ANP).

L50: 'no guarantee...' true, but since models like ISSM can be tuned to give a close match to the observations of velocity, the discrepancy may not be all that great.

We agree that the discrepancy may not be great, but it is not controlled by the method itself so in general there is no guarantee that an ice flow model initialised on the new bedrock and with the optimised velocities as a target will lead to lower flow divergence anomalies. We have reformulate the sentence, we hope that it is clearer now:
"However, there is no constraint that the optimised velocities are a solution of the stress equilibrium equations, so that, in general, the method do not guarantee that the flow divergence anomalies experienced by an ice flow model initialised with the optimised fields will be reduced. »

L86: 'To our knowledge...[claim of novelty]' A number of journals specifically forbid this phrase, and I agree with them. Maybe it is OK in GMD. By the way, I am aware of at least one other paper (accepted at J. Glac.) that takes the ANP approach - obviously it is not citable at this time.

We agree. The sentence has been removed and replaced by:"Our objective is then to illustrate its ability to reconstruct the bedrock topography by comparison with the results of the more mathematically founded first algorithm".

L99: 'effective mean viscosity' ! 'vertically averaged effective viscosity'

It has been modified accordingly in the text.

L129: 'find the parameter space' ! 'find the parameter vector' (which sits inside the parameter space)

Right. It has been changed in the text.

L137-144: I think a bit more clarification - with formulae - is needed here. A good part of Martin and Monnier (2013) relates to the inclusion of terms arising from the dependence of ν on ¿ u.

We have tried to clarified as much as possible this part but without going too far in the details and formulae: "The implementation in Elmer/Ice is carried out in a way that stays as close as possible to the differentiation of the discrete implementation of the direct equations. This method should lead to a better accuracy on the gradient computation than the discretisation of the continuous equations. Elmer/Ice uses programming features that are not supported by automatic differentiation tools and the differentiation of the discrete source code crucial parts (e.g. cost function computation, matrix assembly) has been done manually. If the problem is non-linear, as here due to the dependence of the viscosity to the strain-rate, and the non-linearity solved using a Picard iterative scheme, the iterations should be reversed (at least partially) in the adjoint code to achieve a good accuracy of the computed gradient (Martin and Monnier,2013). However, as our direct solver is equipped with a Newton linearisation of the ice viscosity so that it remains self-adjoint (Petra et al., 2012), we do not reverse the newton iterations in the adjoint code and only keep the last iteration. The adjoint code has been validated on standard test by comparing the gradients with those obtained from a finite difference evaluation. The agreement is usually better than 0.1%.»

section 3.2 - perhaps these short sections could be merged (e.g section 3.2.3 is a single sentence)

We agree that these sections are short but they have the advantage to clearly distinguish the observations required by the method. This will be more difficult if we merge all the sections. However, we merged the very short "Surface elevation" section with the "Bedrock elevation" section.

L258/eq 12 : $|u|, |\theta|$ rather than u, θ ?

Changed.

L274: 25%, and similar elsewhere - in what norm?

The definition used for the calculation of the relative error on the basal drag has been formalised at line 275 in order to avoid any confusion for the reader.

L275 : 'more than 10 times less' → 'less than 10 times'

The sentence has been changed to:
"i.e. more than a tenfold decrease of the initial misfit".

L395: 'no need to inverse a shape variable...' I think a bit of rewording is needed here, along the lines of L78 (ref to Perego 2014)

The sentence has been changed to:"As already mentioned in the introduction, in this latter case, the bedrock topography is no more a state variable but affects the domain geometry making the derivation of the adjoint even more complicated (Perego et al., 2014)."

All figures : plots show the drag coefficient (ß), rather than the traction itself ($|\tau_b|=\beta|u|$. Either is fine, but I suggest that $|\tau_b|$ is the real result. If you switched to a nonlinear basal friction law, you would get the same (or similar) $|\tau_b|$ but a different ß. ß is useful in maps because it shows the difference between slippy and sticky regions, but that doesn't apply so strongly here.

We considered the suggestion and made the change in graphs and legends. However, we prefer to keep ß in figure 4c for readability. Indeed, the perturbation is done on ß and this has little impact on the basal shear stress $|\tau_b|=\beta|u|$ since the increase of ß induces a decrease of $|u|$.