Dear Editor,

Please find in this document the new version of our manuscript. We would like to thank you and referees for their work and their useful and constructive comments on our work. We carefully considered every comments in order to offer the best version as possible of our manuscript. We sincerely hope that our revisions meet the expectations of both referees and editor.

Below you will find a point-by-point reply to all comments made by the reviewers as well as the new version of the manuscript with changes highlighted in red.

The most important changes made in this new version of the manuscript are:

- title change in order to make it more consistent with the content
- changes and precisions of vocabulary according to referees' advices
- clarification of some sentences, explanations and justifications
- deeper justifications on some technical choices such as the value of some parameters

Best regards,
Cyrille Mosbeux

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 $|u|$.

Answer to *R. Arthern (Referee)*

I think this manuscript could be published in Geoscientific Model Development after a few changes. This paper provides a comparison between two different approaches for estimating the initial state and parameters of an ice sheet model. The paper provides an illustration of simultaneous inversion for bed slipperiness and bed elevation using adjoint methods. This part is not an especially novel endeavour in itself, but it is used here to provide a reference for another method, the combined adjoint/nudging method. The merits of the latter approach lie more in its ease of implementation than its theoretical justification. Nevertheless, if adjoint/nudging is shown to be competitive with more complicated approaches, as seems to be the case here, this would represent a valuable service to those ice sheet modellers that presently have the wherewithal to invert for basal drag coefficient, but have not yet considered the shape optimisation problem of recovering the basal topography.

The paper is well-structured and clearly written. The figures are useful and clear. The two parameter adjoint approach has perhaps been described better elsewhere, but I think the examination of the combined adjoint/nudging approach as described here is probably still worth publishing in GMD. The results are sufficient to support the interpretations and conclusions. The authors make clear which parts are new, and which have previously appeared in the literature. The title is OK, although the new feature of the paper is the combination of adjoint-based inversion and nudging and this is not prominent in the title. The abstract is fine. The mathematical presentation is clear enough. The number and quality of references are OK

We revised the title in order to give more details about the assimilation methods used in the paper. Therefore, the new title is : “Comparison of adjoint and nudging methods to initialise ice-sheet model basal conditions".

The simulations used to illustrate the comparison are undoubtedly highly simplified: a simplified approximation of the stress state is used, a 2D flowline rather than a 3D ice sheet is considered, and all the measurements considered in this manuscript are synthetic. These simplifications are expanded upon below. However, to my mind, these do not detract from the central purpose of the manuscript, provided that it is recognised that this paper provides a necessary test that should be passed by the adjoint/nudging method, rather than a sufficient test that would guarantee its usefulness by other models in more general circumstances. In short, this paper might motivate readers to consider the adjoint/nudging method for initialising their models, but each modeller will still need to demonstrate that the method works for their model, in 3D, not just 2D, and each modeller would preferably test the approach with real observations as well as idealised ‘twin’ experiments.

The simulations use the Shallow Shelf Approximation SSA. This is the shallow aspect ratio limit appropriate for flow over a very slippery substrate. For shearing flow over non-slippery sediment, another commonly used limiting approximation, the shallow ice approximation (SIA), which is not used here, would be more appropriate. Nowadays, the practical initialisation problem for ice sheets is more likely to be performed with a more sophisticated stress-balance using a vertically integrated ‘hybrid’ blend of SIA and SSA stress states, or a depth-resolved higher-order model, or Stokes flow. In these more sophisticated models, the transition from slippery to non-slippery substrate poses no special complications, while for the SSA approximation used here, the accuracy of the model will deteriorate whenever the assumption of extreme slipperiness is violated.The

paper would be improved significantly if similar twin experiments were performed using the adjoint/nudging approach for a hybrid model, a higher order model, or a Stokes flow model. This would be especially valuable if it turned out that the bed recovered from the inversion was shown to depend on the approximations used in the momentum equations. The chief selling point of the combined adjoint/nudging method is that it would be easy to apply to more complicated models, so I am not sure why this is not done in this paper. As it stands, the paper points to the promise of this approach for initialising more complicated models, but without a relevant example, it is hard to know whether this is real promise or false promise.

A relevant point, which arises in the upper paragraph, concerns the potential added value to the article if the adjoint/nudging approach was used with a higher order model or a Full-Stokes model. Indeed, the purpose of the adjoint/nudging approach is its easy adaptation to such models. However, we only considerate the SSA approximation since the point of the paper is to compare this approach to a more conventional one (the inversion of the two parameters using adjoint method).

Concerning the impact of the use of higher order model on the bedrock and basal drag solutions, it is clear that the solutions will depend on the direct model. However, this was not the main purpose of the paper although the idea is relevant.

Also, our case of reference is a steady-state case constructed using the SSA approximation. Using a Full-Stokes direct model (or others higher order models) would change the reference case since the reference friction or the reference ice surface velocities are adapted to SSA. These changes could make harder the comparison between the different inversions.

Moreover, the use of adjoint/nudging in full-Stokes model is one of our next goal as well as 3D applications where the results will be compared with the results using SSA.

I have fewer concerns about using a 2D flowline simulation for illustrating the two methods, but some readers will wonder whether the two methods would still perform comparably in 3D. The paper is still quiet compact, and a 3D example would make for a fuller investigation. The paper states that the methods can be applied in 3D, but it would be better to show an example. The use of exclusively synthetic observations represents a limit to the information provided by these simulations. To the authors credit, the data used are based on a real flowline, so the bed inversion at least can be checked. There are a number of regularisation parameters in the inversion(λ_α , λ_{zb} , T, k). Inevitably, these parameters represent rather vague prior information and are quite poorly constrained (Arthern R.J., J. Glaciol., 61 (229), 947-962, 2015, doi: 10.3189/2015JoG15J050). At least it would be good to include a table showing how much the inversion of the bed can vary from the ‘true’ bed when these are varied.

The purpose of the paper is to illustrate the performance of the different methods in a synthetic case. Although, the algorithms are implemented to work in 3D (pseudo-3D for ATP) we made the choice to keep a 2D example which appears more didactical from our point of view.

Concerning the regularisation parameters used in the inversion, the choice of these parameters has been further discussed in the revised version of the paper (see also the answers to the minor point below).

For ten year simulations with the forward model, it should now be possible to test the evolution of the surface against real altimetric observations. This is perhaps too much to ask of an initial demonstration paper such as this, but time series of elevation data for Jakobshavn are available and it would be interesting to see how well the different methods reproduce the actual behaviour over ten years.

In reality, the twin experiment is inspired from the true observations such as the bedrock shape in order to make it more realistic. However, the case is synthetic. The reference friction is inferred using adjoint method, optimising the misfit between observed and modelled velocities, after what we have conducted a model relaxation up to steadystate. The reference velocities and the rate of change of ice thickness (equal to zero) correspond to this steady-state.

Moreover, the flowline behaviour can be very different that the 3D behaviour of a glacier which is more complex. Therefore, although the idea is very interesting, it would not be very relevant to make a direct comparison between our surface evolution during 10 years and the current evolution of the Jakobshavn Isbrae glacier.

In summary, this paper makes one point quite nicely – that the adjoint/nudging approach can work well for the SSA, for flowline models, as judged by synthetic 'twin' experiments, but it still leaves many avenues to be explored.

Thanks for this positive comment. It leaves many avenues to be explored such as adaptation of the method to a 3D real cases. As we already mentioned, this is our next goal and we really are looking forward to show it soon.

Minor points:

Line 46: Replace 'the' with 'then'

Changed in text.

Line 49: Replace 'constrain' with 'constraint' and 'are solution' with 'are a solution'

Changed in text.

Line 125: Are unweighted least squares cost functions such as these appropriate, or should error covariance weighting be applied? Might be worth some discussion.

Actually, we can add an error covariance weighting. In this « idealised » case we do not favour some dhdt_obs with respect to others since measured dhdt_obs are considered perfectly known. Of course, in real application we could add some covariance weighting by using some error estimations on observations such as we can found in Flament and Remy. 2012. However, these estimations seems hard to get.

Note that we tested the effect of noise on dhdt_obs. Firstly, by adding Gaussian white noise to dhdt_obs (see line 370) with no significant effect on the results of the inversion. Of course systematic bias have an impact and the introduction of an error covariance weight in the cost functions could be helpful to reduce this impact (by reducing the weight of the points where bias is known to be high).

We specified in the new version of the paper that covariance error can be taken into account and why we do not use it in our case where we have perfectly known observations or vitiated with Gaussian noise (see line 135).

Line 175: Give more details of the Gaussian used to define *k*. How do results depend upon this choice?

Of course the choice of the variance of the Gaussian will impact the results. Tests show that excessive variance values induce unphysical call-back amplitudes when departing from observations. After a few cycles, the resulting bedrock induces an increase between modelled and observed velocities that the basal drag inversion is not able to overcome.

In our specific case, the threshold on the variance, between excessive and acceptable misfit on velocities is slightly above1 km.

Under this threshold, the value of the variance have little impact on the final result in term of cost functions. However, we tested a few values (0.2, 0.4, 0.6, 0.8 and 1km) among which the variance of 1km gives the best agreement between misfit on velocities and misfit on the rate of change of the ice thickness.

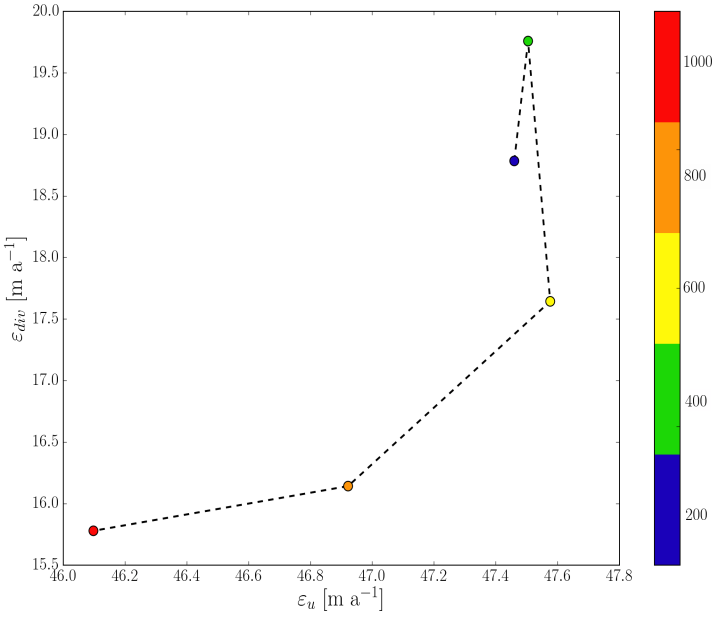

Fig. Misfit on ice surface velocities and on the rate of change of the ice thickness with respect to the variance of the parameter *k (200m, 400m, 600m, 800m, 1000m).*

Line 210: Jakobshavn is misspelt

Changed in text.

Line 267: Sometimes the ‘L’ of the ‘L-curve’ is very clear, sometimes not. It would be good to show the two cross sections through the ‘L-surface’ at the chosen values using a log-log scale.

In our case, the 'L' is visible by looking at the 3D-Lcurves (Jv, Jdiv, Jreg) and (Jv, Jdiv, Jzb). However, it is more an « optimal area » than a specific point. The 3D graphs are not so easy to read if we cannot navigate in the 3D space. It seems more easy to show the

« specific point » we selected by plotting two scatter plots : (Jv,Jdiv) and a color for each dot indicating the value of Jreg and the value of Jzb.

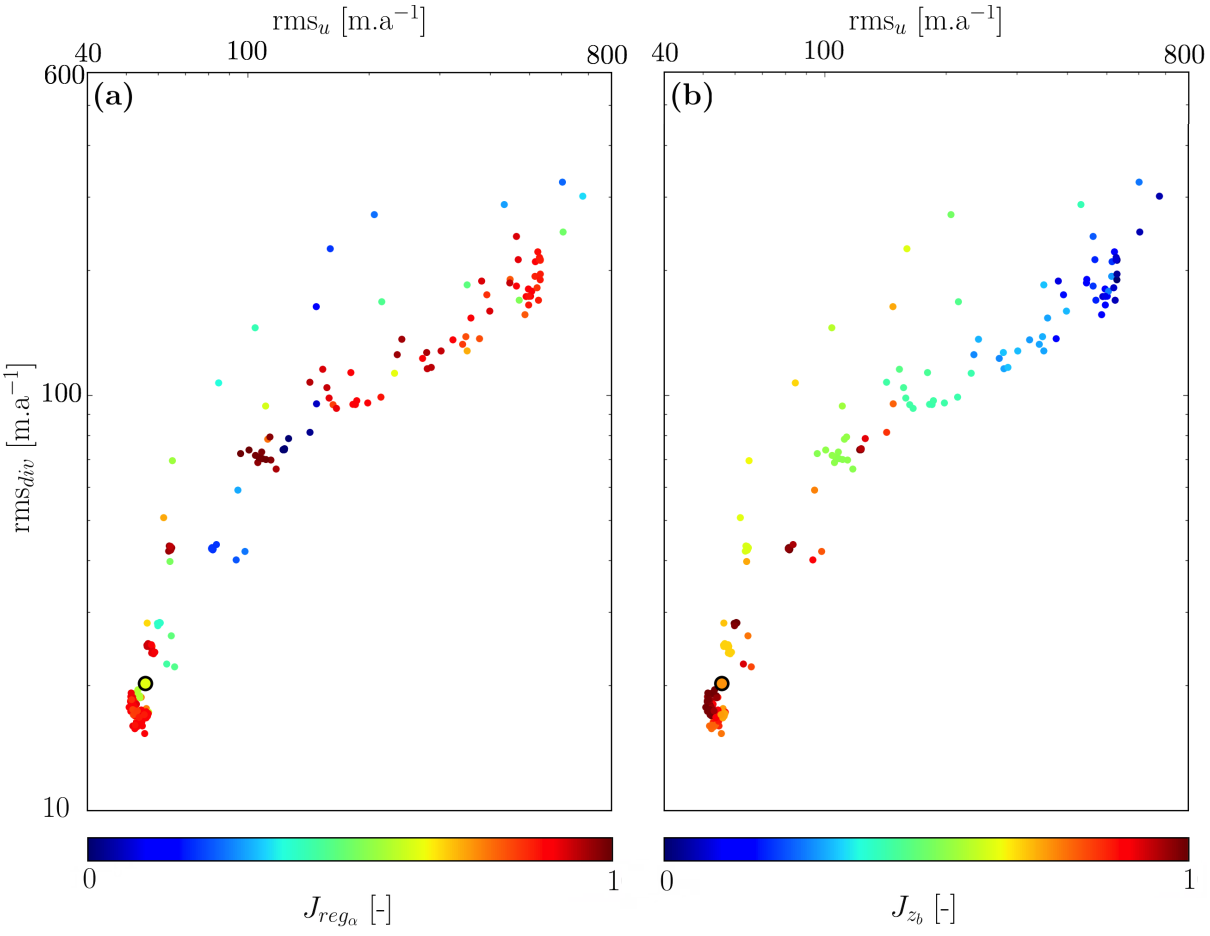

Fig. *Average velocity misfit on the horizontal axis, average misfit on the ice thickness rate of change of the divergence on the vertical axis for each of the 255 couples of regularisation. Values for (a) Jreg and (b) Jzb are represented on the color axis. The optimum is marked with a black circle.*

The graph shows that the dot marked with a black circle, corresponding to our optimum ($\lambda_\alpha = 1e11$, $\lambda_{zb} = 1e7$), is a good compromise between each cost function and allow to minimise both misfit on the divergence and ice velocities while keeping a good agreement with the *a priori* on zb and the regularisation on ß.

As suggested, the figure above has been added to the article to support the choice of the regularisation parameters. It also has been referenced and discussed in text.

Line 305: Since *T* has in effect become a regularisation parameter it would be good to comment whether this trial of a few values is consistent with the treatment of the other regularisation parameters – is there an equivalent to the 'L-curve' for choosing *T*.

Indeed, the choice of *T* can be done similarly to the choice of the regularisation parameters since there is a relation between *T* and Jv and Jdiv. Nevertheless, the choice of *T* seems less straightforward in the present case. In the Figure below, we plotted T in function of the average misfits on surface velocities and on the thickness rate of change.

We can also follow the scatter plot approach we have used for the ATP. Both graphs are given below.

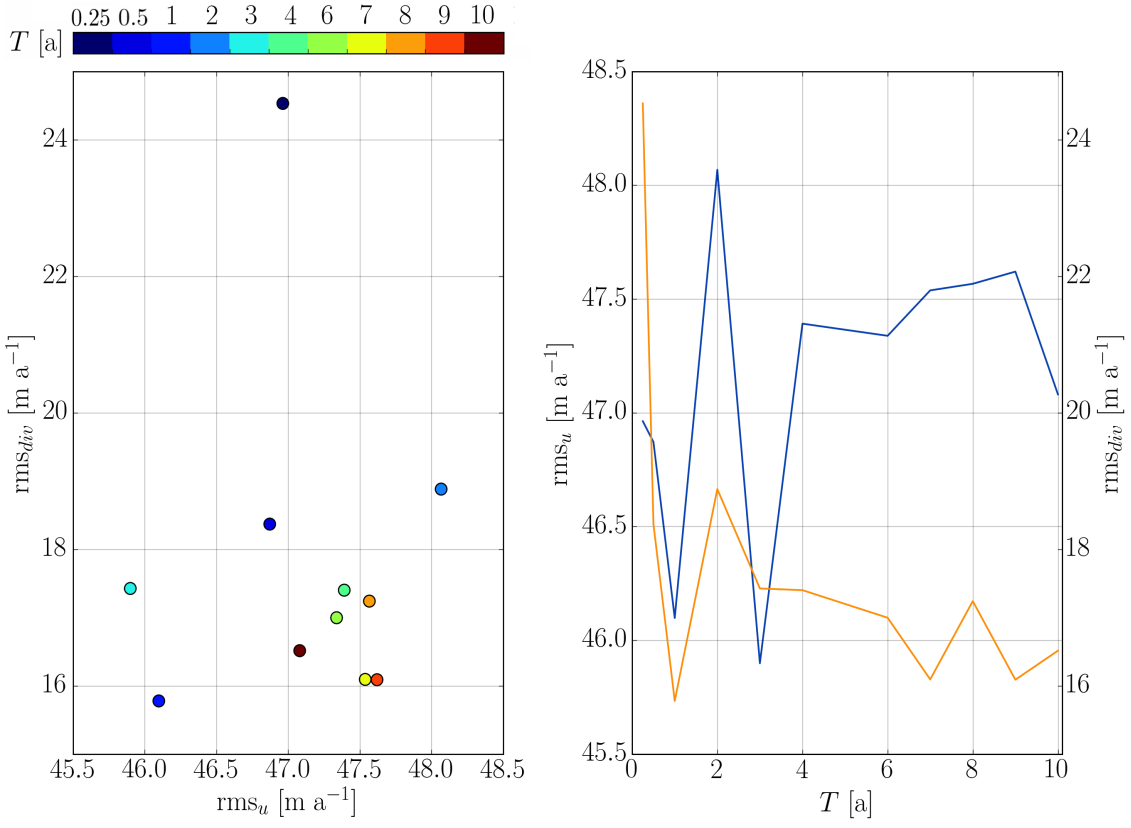

Fig. *Misfits on velocities and ice thickness rate of change with respect to T. (left) Scatter plot where the color correspond to the size of T. (right) Blue curve corresponds to the velocity misfit and orange curve correspond to the ice thickness rate of change misfit.*

Here, 10 ANC cycles are conducted. The figure shows that *T* periods over 4 years tend to largely increase the misfit on Jv since there is no control on velocities during nudging (see sec. 4.2). *T* periods under one year generate high misfit on Jdiv since *T* are not long enough to significantly reduce divergence (of course more ANC cycles would solve this last problem but at the expense of the computational cost). Note also that very long periods, such as 20 years seems, unable to converge since it induces too high ice thickness evolutions in the early cycles of ANC.

The scatter plot shows us that the 1 year T periods gives a good agreement between fitting velocities and divergence. However, the choice of periods of 1 (minimum for Jv) or 3 years (minimum for Jdiv) is still debatable but the choice of a shorter *T* period is, again, to the advantage of the computational cost, which can become a significant choice criterion in cases of larger domains (extension to 3D cases) or more refined meshes.

We decided to not add this Figure to the paper but to detail our choice and our method of selection for the *T* period in sec. 4.2.

Manuscript prepared for Geosci. Model Dev.
with version 2014/09/16 7.15 Copernicus papers of the LaTeX class copernicus.cls.
Date: 20 May 2016

Comparison of adjoint and nudging methods to initialise ice-sheet model basal conditions

Cyrille Mosbeux[1,2], Fabien Gillet-Chaulet[1,2], and Olivier Gagliardini[1,2]

[1]CNRS, LGGE, F-38041 Grenoble, France
[2]Univ. Grenoble Alpes, LGGE, F-38401 Grenoble, France

Correspondence to: C.Mosbeux (cyrille.mosbeux@lgge.obs.ujf-grenoble.fr)

Abstract. Ice flow models are now routinely used to forecast the ice-sheets contribution to 21st century sea-level rise. For such short term simulations, the model response is greatly affected by the initial conditions. Data assimilation algorithms have been developed to invert for the friction of the ice on its bedrock using observed surface velocities. A drawback of these methods is that remaining uncertainties, especially in the bedrock elevation, lead to non-physical ice flux divergence anomalies resulting in undesirable transient effects. In this study, we compare two different assimilation algorithms based on adjoints and nudging to constrain both bedrock friction and elevation. Using synthetic twin experiments with realistic observation errors, we show that the two algorithms lead to similar performances in reconstructing both variables and allow the flux divergence anomalies to be significantly reduced.

1 Introduction

Robustly reproducing the responsible mechanisms and forecasting the ice-sheets contribution to 21st century sea-level rise is one of the major challenges in ice-sheet and ice flow modelling as highlighted by community-organized efforts such as SeaRISE (Sea-level Response to Ice Sheet Evolution) (Bindschadler et al., 2013; Nowicki et al., 2013a, b) or ice2sea (e.g., Gillet-Chaulet et al., 2012; Shannon et al., 2013; Edwards et al., 2014).

Such projections on decadal timescales are sensitive to the model initial state which can account for an important source of uncertainty in the model response (Aðalgeirsdóttir et al., 2014). Improving the reliability of the model projections requires the model initial state to be better constrained from observations. The problem is that observations are often uncertain, sparse in time and space and indirect, so that the model state depends on many poorly determined physical parameters and

boundary conditions. ~~Inverse methods~~ Gradient-based optimisation methods, such as the control method (Macayeal, 1993) or the Robin inverse method (Arthern and Gudmundsson, 2010), are efficient means to constrain such model parameters and boundary conditions. These methods have been implemented and applied with success in ice flow models of different complexity in order to infer the basal drag, one of the most uncertain model parameters (e.g., Morlighem et al., 2010; Jay-Allemand et al., 2011; Schäfer et al., 2012; Gillet-Chaulet et al., 2012).

However, remaining uncertainties lead to non-physical ice flux divergence anomalies (Seroussi et al., 2011) resulting in undesirable transient effects in the free surface evolution. A solution to dissipate these transients is to conduct a surface relaxation step prior to the projections (Gillet-Chaulet et al., 2012). This allows admissible flux divergence rates to be reached but at the expense of the accuracy of the modelled surface elevation and surface velocities which can then depart significantly from observations after the relaxation step.

Among the remaining uncertainties, one of the most important is the uncertainty related to the bedrock elevation. The basal topography is derived from ice thickness measurements, mostly obtained from airborne ice penetrating radars. These measurements can have large uncertainties and are usually at a lower resolution than required model grids (Durand et al., 2011). Standard bedrock elevation maps for Antarctica and Greenland are then produced by interpolation or Kriging and report standard errors ranging from few tens of meters to several hundreds of meters depending on the distance to observations and local topographic variability (Fretwell et al., 2013; Bamber et al., 2013). For comparison the uncertainty on the surface elevation is usually one order of magnitude lower (Fretwell et al., 2013).

Because of theses large uncertainties, several methods have been proposed to consider the bedrock elevation as an optimisation variable. For example, Morlighem et al. (2011) derive the adjoint of the continuity equation for the ice thickness. The depth-averaged velocities and surface mass balance are ~~the~~ then optimised to minimise the mismatch between modelled and measured ice thicknesses. Surface velocity measurements are used as initial guess for depth-averaged velocities, and by construction the flux divergence produced by this approach is in equilibrium with the prescribed surface mass balance. ~~However, there is no constrain that the optimised velocities are solution of the stress equilibrium equations, so that there is no guarantee that a mechanical flow model can match the optimised divergence field.~~ However, there is no constraint that the optimised velocities are a solution of the stress equilibrium equations, so that, in general, the above method does not guarantee that the flow divergence anomalies resulting from an ice flow model initialised with the optimised fields will be reduced.

van Pelt et al. (2013) developed an iterative algorithm where the discrepancy between the surface elevation predicted by the model and the observations is used to correct the bedrock elevation. So, the method does not rely on the accurate computation of the derivative of a cost function as in a control method and is then more similar to nudging methods that have been widely studied in the

past decades in meteorology (e.g., Hoke and Anthes, 1976) and later in oceanography (Verron, 1992; Blayo et al., 1994). However, the method proposed by van Pelt et al. (2013) does not use observed surface velocities to control the model parameters.

Several methods have been explored to construct model states where both the basal friction and the basal topography are treated as optimisation variables. In a pioneer work, Thorsteinsson et al. (2003) developed a least-squares inversion using analytical solutions for the transmission of small scale basal perturbations to the ice surface. This method has been extended in a non-linear Bayesian framework by Raymond and Gudmundsson (2009) and applied to an Antarctic ice stream by Pralong and Gudmundsson (2011). Bonan et al. (2014) have tested the performances of an ensemble Kalman Filter on twin experiments using a shallow ice flow line model. The adjoint method has been tested by Goldberg and Heimbach (2013) and Perego et al. (2014) with models of different complexity. All these methods usually show good performances in reconstructing both basal friction and basal topography when using observations of both surface elevation and surface velocities, so that mixing between the two variables does not seem to be too problematic for realistic applications (Gudmundsson and Raymond, 2008). In addition, Pralong and Gudmundsson (2011) and Perego et al. (2014) show better performance when the rates of surface elevation change are also constrained from observations.

In this paper, we explore two different algorithms to infer both the basal friction and the basal topography and initialise the model state using simultaneous observations at a given time. The first algorithm is in line with Goldberg and Heimbach (2013) and Perego et al. (2014) since it uses the adjoint solution of the force balance equation. We use the Shallow Shelf Approximation to facilitate the derivation of the adjoint. Indeed, in this case the ice thickness appears as a state variable, while it changes the geometry of the domain for higher order approximations (Perego et al., 2014). In its simplest formulation, the algorithm minimises the misfit between model and observed surface velocities, but an additional constraint where the flux divergence is close to a given surface mass balance can be added. The second method is an algorithm combining inversion of basal friction using the adjoint method and nudging of the bedrock topography. The control from the surface velocity observations is imposed by the adjoint step while the nudging step allows to decrease the discrepancy between the flux divergence and the surface mass balance. The main motivation of this second algorithm is its ease of implementation as no inversion of the model with respect to the ice thickness is required. ~~To our knowledge, this second method has never been explored to reconstruct the bedrock topography in a systematic manner.~~ 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. Both algorithms are implemented in the finite element ice-sheet / ice flow model Elmer/Ice (Gagliardini et al., 2013). The methods and algorithms are described in details in Sec. 2. To test their performances, we design a twin experiment in Sec. 3. The results are discussed in Sec. 4.

2 Methods

2.1 Direct Model

For the force balance, we use the standard vertically integrated Shallow-Shelf Approximation (SSA) equations (MacAyeal, 1989). This approximation neglects the effects of vertical shearing and is, hence, more adapted to model the flow in areas where the friction is low, resulting in an ice motion dominated by sliding. The horizontal velocity field (u, v) is a solution of :

$$\begin{aligned}
&\frac{\partial}{\partial x}\left(2H\nu\left(2\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)\right)+\frac{\partial}{\partial y}\left(H\nu\left(\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\right)\right)-\beta u=\rho g H\frac{\partial z_s}{\partial x}\\
&\frac{\partial}{\partial x}\left(H\nu\left(\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\right)\right)+\frac{\partial}{\partial y}\left(2H\nu\left(\frac{\partial u}{\partial x}+2\frac{\partial v}{\partial y}\right)\right)-\beta v=\rho_i g H\frac{\partial z_s}{\partial y}
\end{aligned} \tag{1}$$

where β is the friction coefficient, ν the ~~effective mean viscosity~~ vertically averaged effective viscosity, ρ_i the ice density, g the gravity, and $H = z_s - z_b$ the thickness, with z_s and z_b the top and bottom surface elevations, respectively.

Natural boundaries are the calving fronts where the Neumann condition results from the difference between the ice pressure and the sea water pressure:

$$\begin{aligned}
&4H\nu\frac{\partial u}{\partial x}n_x+2H\nu\frac{\partial v}{\partial y}n_x+H\nu(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial x})n_y=(\rho_i g H-\rho_w g H_0)n_x\\
&4H\nu\frac{\partial v}{\partial y}n_y+2H\nu\frac{\partial v}{\partial x}n_y+H\nu(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial x})n_x=(\rho_i g H-\rho_w g H_0)n_y
\end{aligned} \tag{2}$$

where ρ_w is the water density, H_0 the ice thickness below sea level and n_x and n_y the two components of the horizontal unit vector normal to the calving front. Dirichlet boundary conditions are prescribed for other non-natural boundaries. The continuity equation for the ice thickness is given by:

$$\frac{\partial H}{\partial t}+\frac{\partial(uH)}{\partial x}+\frac{\partial(vH)}{\partial y}=a\,, \tag{3}$$

where a is the surface mass balance and accumulation/ablation at the bedrock interface is neglected.

2.2 Inverse methods

The objective of the methods is to produce a model state from Eq. (1) that best fits the observations of surface velocities and the rates of change of ice thickness. To minimise the discrepancy between the model and the observations, the optimisation parameter space p contains both the basal friction coefficient β and the bedrock elevation z_b.

2.2.1 Cost Functions

The misfit between the model and the corresponding observations is evaluated using cost functions. The first cost function measures the difference between modelled ($\boldsymbol{u}$) and observed ($\boldsymbol{u_{obs}}$) surface

velocities :

$$J_v = \int_\Gamma \frac{1}{2}(|\boldsymbol{u} - \boldsymbol{u_{obs}}|)^2 d\Gamma\,, \tag{4}$$

where Γ is the model domain.

The second cost function measures the misfit between modelled and observed thickness rates of change:

$$J_{div} = \int_\Gamma \frac{1}{2}\left[\left(\frac{\partial H}{\partial t}\right) - \left(\frac{\partial H}{\partial t}\right)_{obs}\right]^2 d\Gamma. \tag{5}$$

The modelled rate of change of ice thickness $\partial H/\partial t$ is evaluated from Eq. (3) as the difference between the flux divergence solution of Eq. (1) and the prescribed surface mass balance. Observed rate of change of ice thickness $(\partial H/\partial t)_{obs}$ can be estimated from surface elevation trends extracted from radar altimetry measurements (Flament and Rémy, 2012).

In general, both Eq. (4) and (5) could be weighted with error covariance estimates such as the one of Flament and Rémy (2012). However, these informations are not often available. In this paper, observed ice surface velocities ($\boldsymbol{u_{obs}}$) and observed rate of change of ice thickness $(\partial H/\partial t)_{obs}$ are considered perfectly known or perturbed with a Gaussian noise which would make unnecessary the addition of a covariance term.

The objective is then to find the parameter ~~space~~ vector p that minimises J_v and J_{div}. This can be achieved in different ways as illustrated in the following sections.

2.2.2 Adjoint method

The two cost functions have an implicit dependence on the parameter space p through the model surface velocities **u** which are solution of Eq. (1). The gradient of the cost functions with respect to p can be computed efficiently using the adjoint equations of Eq. (1). The derivation of the continuous adjoint equations and the gradient of J_v with respect to the friction coefficient β can be found in Macayeal (1993). This can be easily extended for the computation of the gradient with respect to H.

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 crucial parts of the discrete source code (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 the present direct solver is equipped with a Newton linearisation of the ice viscosity so that it remains self-adjoint (Petra et al., 2012), the

newton iterations are not reversed in the adjoint code and we 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 %.

~~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. It is similar to Martin and Monnier (2013) except that the iterative algorithm used to solve the non-linearity of the direct equations, arising from the explicit dependence of the ice viscosity on the velocity field, is not reversed in our adjoint computation. However as our direct solver is equipped with a Newton linearisation of the ice viscosity it remains self-adjoint (Petra et al., 2012). The other parts of the discrete source codes have been differentiated manually.~~

Inverse problems are often ill-posed leading to instabilities. It is then necessary to add regularisation terms to the cost function to avoid over-fitting of data. This can be done in the form of a Tikhonov regularisation. Here we define two different regularisations. The first one measures the norm of the first spatial derivative of p, thus allowing to give preference to smooth solutions:

$$J_{reg} = \frac{1}{2}\int_{\Gamma}\left(\frac{\partial p}{\partial x}\right)^2 + \left(\frac{\partial p}{\partial y}\right)^2 d\Gamma. \tag{6}$$

The second forces the optimisation variables to stay close to a certain prior or background information p_b. This background can be based on observations or on empirical knowledge. This second regularisation term is written as

$$J_b = \frac{1}{2}\int_{\Gamma}\frac{1}{\sigma_p^2}(p-p_b)^2 d\Gamma\,, \tag{7}$$

where σ_p is a spatial parameter allowing to give more or less weight to the prior information p_b.

The computation of the gradients of these two functional with respect to p is trivial. How these regularisation terms are weighted with respect to the model-data misfit functionals Eqs. (4) and (5) is described in more details with the description of the algorithms in Sec. 2.3.

This minimisation is achieved using the quasi-Newton routine M1QN3 (Gilbert and Lemaréchal, 1989) implemented in Elmer/Ice. This method uses an approximation of the second derivatives of the cost function and is therefore more efficient than a fixed-step gradient descent.

2.2.3 Nudging method

By definition, the steady-state solution of Eq. (3) where a is replaced by the apparent mass balance $a-(\partial H/\partial t)_{obs}$ is the minimum for J_{div}. Running the model forward in time with a constant forcing is then a simple way to minimise J_{div}, equivalent to a relaxation step. Here we assume that the surface elevation is known so that computed changes in ice thickness are used to correct the bedrock elevation z_b. During this process, the ice thickness can substantially deviate from obser-

vations. Nudging methods, also called Newtonian relaxation, can remedy to this problem by constraining the thickness to fit observations through an additional call-back term in Eq. (3), which now writes:

$$\frac{\partial H}{\partial t} + \frac{\partial (uH)}{\partial x} + \frac{\partial (vH)}{\partial y} = a - \left(\frac{\partial H}{\partial t}\right)_{obs} - k(H - H_{obs}), \tag{8}$$

where the coefficient k defines the amplitude of the call-back at each node of the model. These methods imply a trade-off, adjustable through k, between model physics and observations. The call-back term can depend on many different criteria such as observation accuracy or distance to observation (Hoke and Anthes, 1976). Here, we take k as a Gaussian function of the distance to the closest observation so that the call-back is maximum where an observation is available and decreases to zero far from all observations. The choice of the variance for the Gaussian function is discussed in Sec. 4.2.

2.3 Algorithms

From the methods presented in the previous section we design two algorithms to infer simultaneously the friction coefficient β and the bedrock elevation z_b. To ensure that the friction coefficient remains positive during the inversion we use the following change of variable

$$\beta = 10^{\alpha} \tag{9}$$

2.3.1 Adjoint method with two parameters (ATP)

This algorithm uses the gradients of the cost functions derived using the ajoint method to optimise both α and z_b. For the regularisation, a constraint on the smoothness is imposed for α using Eq. (6) while a constraint on the background information is imposed for z_b using Eq. (7). The total cost function then writes

$$J_{ATP}(\alpha, z_b) = J_v + \gamma J_{div} + \lambda_\alpha J_{reg_\alpha} + \lambda_{z_b} J_{b_{z_b}} \tag{10}$$

where γ is a constant fixed to give a similar weight to J_v and J_{div} while λ_α and λ_{z_b} are two constants allowing to adjust the weight given to the regularisation terms. Following Fürst et al. (2015), several pairs (λ_α,λ_{z_b}) are tested using a L-curve approach, and optimal values are taken from the combinations that avoid two extremes: overfitting of the observations or excessive regularisation.

2.3.2 Adjoint-nudging coupling (ANC)

In this algorithm, the adjoint method is first used to optimise α only by minimising the following total cost function

$$J_{ANC}(\alpha) = J_v + \gamma J_{div} + \lambda_\alpha J_{reg_\alpha}. \tag{11}$$

The bedrock elevation is then updated using the nudging method by solving Eq. (8) for a given time period T. T should be neither too short nor too long to allow to reduce J_{div} without over-fitting observations. The sensitivity of the method to T is discussed in the results section.

These two steps are then repeated iteratively until changes in J_v and J_{div} between two iterations are less than 1%.

3 Manufactured data sets

A twin experiment is design to investigate the ability of the two methods to reproduce simultaneously good estimates of the basal friction coefficient and the bedrock elevation. A flowline geometry is preferred to reduce the computational cost and easily test the method, however all the algorithms can be applied to 2D plane view simulations. A reference experiment for which all the model parameters are prescribed is produced to generate synthetic observations. These observations are then used to test the performances of the two algorithms.

3.1 Reference experiment

A flowline of ~~Jackobshavn~~ Jakobshavn Isbrae, Greenland, is used to test the two algorithms with realistic conditions. ~~Jackobshavn~~ Jakobshavn Isbrae is one of Greenland's three largest outlet glaciers and has one of the largest drainage basin on the ice sheet's western margin (Bindschadler, 1984). It is also the fastest Greenland glacier with a terminus velocity greater than 13 $\mathrm{km\,a^{-1}}$ (Joughin et al., 2008, 2014). The flowline is 550 km long and runs from the ice divide to the ice front. The surface and bedrock elevations are taken from available digital elevation models (Bamber et al., 2013). The basal friction coefficient field is first adjusted so that the model velocities fit observed velocities (Joughin et al., 2010). To have realistic thickness rates of change, the free surface is relaxed to steady state. The surface mass balance a in Eq. (3) has been calibrated so that the steady state is close to the initial geometry, and is meant to take into account the flow convergence or divergence along the flowline. The steady state solution is used as the reference of the twin experiment.

The geometry is discretised through a mesh of 500 linear elements, increasingly refined to the front of the glacier. The element size decreases from ~ 2 km in the upper part of the glacier to $\sim$ 400 m down to the front.

Results will only be presented on the first 100 km upstream of the glacier front where velocities are above 100 $\mathrm{m\,a^{-1}}$ and where the SSA is more appropriate but the inversion is done all along the flowline up to the ridge.

3.2 Synthetic observations

Synthetic observations are generated by sampling and/or adding noise to the reference simulation. Details for each required field are given below. These synthetic observations and initial fields for the inverse methods are compared to the reference in Fig. 1.

3.2.1 Surface velocities

Surface velocities are assumed to be observed at the same resolution as the reference simulation but with a white Gaussian noise with a mean $\mu = 0$ and a standard deviation $\sigma = 50\ \mathrm{m\,a^{-1}}$. This corresponds to a root mean squared error of 47.8 $\mathrm{m\,a^{-1}}$ for the entire flowline. The reference and noisy observed surface velocities are shown in Fig. 1c together with their absolute difference.

3.2.2 Surface mass balance and thickness rate of change

The surface mass balance, a, and thickness rate of change, $(\partial H/\partial t)_{obs}$, in Eq. (5) are assumed to be perfectly observed. As the reference simulation corresponds to a steady-sate, $(\partial H/\partial t)_{obs} = 0$. However, the methods are also tested in Sec. 4.3 for cases where $(\partial H/\partial t)_{obs} \neq 0$ to show their ability to initialise the model when the flux divergence is not in equilibrium with the surface mass balance.

3.2.3 Surface and bedrock elevations

The surface elevation is assumed to be perfectly observed. For the bedrock elevation z_b, we simulate observations representing airbone radar measurements crossing the flowline. Bedrock elevations are sampled every 10 km with a Gaussian noise centred on zero and with a standard deviation of $\sigma = 50\ \mathrm{m}$. This leads to a rms error of 62.4 m on the 55 observation points of the entire flowline. This error is similar to the errors given in practice on recent bedrock elevation maps (Fretwell et al., 2013; Bamber et al., 2013). For the mesh nodes between the observations, the bedrock is linearly interpolated as shown in Fig. 1a. This is used as the first guess for the inverse methods and as the background information for the regularisation, Eq. (7).

3.2.4 Model parameters

The ice viscosity is assumed to be perfectly known and corresponds to the viscosity used in the reference experiment.

Assuming that no observation of the friction coefficient is available, an initial solution has to be postulated. A good first guess for β is provided by using the driving stress to estimate the basal shear stress :

$$\beta_{ini}(x) = \frac{\rho_i g H(x) |\theta(x)|}{|u(x)|} \tag{12}$$

where $H(x)$, $\theta(x)$, and $u(x)$ are respectively the ice thickness, the surface slope and the surface velocity at position x. The reference and initial values are shown in Fig. 1b.

The rms errors on the surface velocities and the rate of change of ice thickness between the initial state and the synthetic observations are, respectively, 761 $\mathrm{m\,a^{-1}}$ and 357 $\mathrm{m\,a^{-1}}$.

The average relative error on the basal shear stress is measured as :

$$\varepsilon_{\tau_b} = \frac{1}{L}\int_{\Gamma} \frac{|\tau_b| - |\tau_{b,ref}|}{|\tau_{b,ref}|} d\Gamma, \quad (13)$$

where $\tau_{b,ref}$ is the basal shear stress in the reference experiment and L the length of the flowline. The relative error on the basal shear stress with our initial estimate of the basal friction β_{ini} is 394 %. The performances of the two algorithms in reducing these initial errors are presented in the following section and will be compared to these initial errors.

4 Results

4.1 Adjoint method with two parameters (ATP)

A set of 255 pairs (λ_α,λ_{z_b}) is tested to adjust the weighting of the regularisation terms of Eq. (10). ~~The constants $\lambda_\alpha = 10^{11}$ and $\lambda_{z_b} = 10^7$ show an optimum between data fitting and regularisation.~~ The misfits on the different cost functions of Eq. (10) for the different pairs (λ_α,λ_{z_b}) is given in Fig. 2. Both graphs show that most of the pairs fitting well the observed velocities can also adequately reproduce the observed rate of change of the ice thickness. Fig. 2b also shows that smaller misfits on J_{z_b} clearly involve higher rms misfits on the ice surface velocities (rms_u) and on the rate of change of the ice thickness (rms_{div}). On the contrary, Fig. 2a does not show a clear relation between the magnitude of J_{reg_α} and the magnitude of rms_u and rms_{div}. Both graphs also show a high density of pairs for small rms_u and rms_{div}. However, the pair ($\lambda_\alpha = 10^{11}$, $\lambda_{z_b} = 10^7$) seems to come off the others, giving a good trade off between data fitting and regularisation. Notice that the constant γ is fixed to 1 since J_v and J_{div} have the same order of magnitude.

The optimisation of both α and z_b simultaneously allows to reach a rms misfit of 49.7 $\mathrm{m\,a^{-1}}$ on velocities, very similar to the observation rms error, showing no overfitting of velocity data. The rate of change of ice thickness misfit is also largely decreased with a rms value of 19.2 $\mathrm{m\,a^{-1}}$. The resulting ~~β~~ basal traction τ_b and z_b as well as the misfit for the surface velocities are given in Fig. 3. The ~~friction~~ basal traction variability is accurately reproduced with a corresponding average relative misfit of only 25 % along the entire flowline with respect to the reference basal shear stress $\tau_{b,ref}$, i.e. ~~more than 10 times less than the misfit obtained with β_{ini}~~ more than a tenfold decrease of the initial misfit. We only notice local over-estimations of slipperiness in bedrock pits without significant impacts on the flow velocities. Indeed, under a defined value of β corresponding to a nearly perfectly sliding case, an additional reduction in friction has no impact on the flow. The same reasoning applies to a nearly perfectly sticky case, where an increased friction would not involve more decrease of the velocity. The bedrock elevation z_b is well reconstructed in the first 50 km upstream of the glacier front. The discrepancy with respect to the reference bedrock is larger upstream where the cost function J_v is less sensitive because of lower velocities. This could possibly

be improved by using a cost function measuring the logarithm of the misfit as in Morlighem et al. (2010), but with a greater risk of fitting noise since the relative observation error is higher in these regions.

In order to assess the influence of accounting for J_{div} on the method, the optimisation is repeated without the J_{div} term in the total cost function Eq. (10). The pair (λ_α, λ_{z_b}) is kept equal to the previous case since the optimum is hardly affected by the absence of J_{div} in the total cost function. The result is given and compared to the previous one in Fig. 3. The friction coefficient β is again pretty well reconstructed, with a corresponding relative average misfit of 31 % on basal shear stress τ_b to be compared to the 25 % obtained with the optimisation of the J_{div} term. However, z_b shows non-consistent high frequencies involving a higher discrepancy with respect to the reference bedrock elevation that in the case of optimising J_{div}. Therefore, the optimisation J_{div} has a clear regularisation effect on the parameter z_b, by reducing the non-consistent high frequency oscillations of the solution.

Introduction of a Gaussian noise on $(\partial H/\partial t)_{obs}$ has been experienced in order to assess its effect on the optimisation of J_{div}. Different levels of standard deviation σ have been tested. Experiences show that the optimisation is little affected by this noise even for standard deviations σ going up to the same order of magnitude than the surface accumulation a. Introduction of systematic bias on $(\partial H/\partial t)_{obs}$ in a physically acceptable range, i.e. of the same order of magnitude than surface accumulation a, also have little consequences on the optimisation.

4.2 Adjoint-nudging coupling (ANC)

The steps consisting in the optimisation of α only are conducted with a value $\lambda_\alpha = 5.10^9$ which allows a good agreement between the different cost functions and a value $\gamma = 1$.

Additionnaly to the regurlarisation parameters of Eq. (11), ANC algorithm depends on the time period for the nudging steps T and the variance of the Gaussian k in Eq. (8). The nudging period T impacts the convergence on J_v and J_{div} after each cycle. The convergence is substantially similar for T from 1 to 4 years. Longer periods mainly involve a worse minimisation of J_v since there is no control on velocities during nudging. Shorter relaxation times do not involve sufficient change of z_b inducing a lower minimisation of J_{div} for a given number of cycles. Therefore, a relaxation time $T = 1$ year is adopted, which seems sufficient to allow significant changes of z_b without too much adaptation to the previous intermediate value of the friction coefficient. The algorithm is stopped after 10 cycles, corresponding to the stopping criterion of Sec. 2.3.2. For a given T period, tests show that variance values of the Gaussian k in Eq. (8) larger than 1 km are excessive and induce non-physical call-back amplitudes when departing from observations. After a few cycles, the resulting bedrock induces an increase between modelled and observed velocities that cannot be overcome by the basal drag inversion. Variance smaller than 1 km has little impact on the final result in term of

cost functions. However, among the acceptable values, the 1 km variance gives the best agreement between misfit on the surface velocities and misfit on the rate of change of the ice thickness.

~~In order to find the optimal time period for the nudging steps, different T are tested looking at the resulting convergence on J_v and J_{div} after each cycle. The convergence is substantially similar for T from 1 to 4 years. Longer periods mainly involve a worse minimisation of J_v since there is no control on velocities during nudging. Shorter relaxation times do not involve sufficient change of z_b inducing a lower minimisation of J_{div} for a given number of cycles. Therefore, a relaxation time T of 1 year is adopted, which seems sufficient to allow significant changes of z_b without too much adaptation to the previous intermediate value of the friction coefficient. The algorithm is stopped after 10 cycles, corresponding to the stopping criterion of Sec. 2.3.2.~~

The model is in good agreement with observations with a rms misfit of 46.1 $\mathrm{m\,a^{-1}}$ in the range of observation noise, for velocities and 15.8 $\mathrm{m\,a^{-1}}$ for thickness rates of change. The ~~friction coefficient~~ β basal shear stress τ_b is close to the reference one despite exacerbated variations at some locations. The corresponding relative average misfit with respect to the reference $\tau_{b,ref}$ ~~on basal shear stress~~ is 30 % for the entire flowline. The reconstructed bedrock elevation z_b is also close to the reference on almost 100 km upstream of the front of the glacier. This reflects, especially in fast flowing region, a real improvement of the basal knowledge with respect to the first guess. Moreover, the use of nudging, instead of adjoint method, does not show the same problem of non-sensitivity in region of slower flow velocities, as mentioned in Sec. 4.1. Note however that z_b significantly departures from the reference bedrock elevation from 80 to 100 km to the front, strongly linked to the poorer fit of β (see Fig. 4).

As for ATP, introduction of a Gaussian noise in the observed thickness rate of change $(\partial H/\partial t)_{obs}$ has also been tested. Results show no significant impacts on the optimisation. Nevertheless, introduction of systematic bias in $(\partial H/\partial t)_{obs}$ has direct consequences on the nudging steps inducing an offset of z_b of the range of the systematic bias cumulated on the nudging period T. ANC is therefore more sensitive to systematic bias than ATP.

4.3 Further sensitivity experiments

In order to evaluate the efficiency of both algorithms in transient states, we construct new reference cases where $(\partial H/\partial t)_{obs} \neq 0$. This is achieved by multiplying β_{ref} by a factor 2, 3, 4, 5 and 10. As a consequence, increasing ~~Therefore, these increases in~~ the basal friction involves a disequilibrium of the glacier, an ice thickening and a decrease of ice flow velocities.

The time period for the glacier to come back to equilibrium, after this change of friction parameter, depends on the amplitude of the perturbation. Here, the perturbation is only applied during 5 years in order to keep the five cases in disequilibrium. Resulting thickness rates of change $(\partial H/\partial t)_{obs} \neq 0$ are in the same order of magnitude than the tuned surface mass balance a. The five new reference cases are presented in Fig. 5.

The results of the optimisations for the five cases of perturbation are shown in Fig. 6 for ATP and Fig. 7 for ANC . The velocity misfit for ATP increases with the amplitude of perturbation with rms values between 47.8 $\mathrm{m\,a^{-1}}$ and 52.3 $\mathrm{m\,a^{-1}}$ while the rms misfit for thickness rate of change increase from 12.7 $\mathrm{m\,a^{-1}}$ to 21.8 $\mathrm{m\,a^{-1}}$. ANC reaches rms misfits from 45.9 $\mathrm{m\,a^{-1}}$ to 47.2 $\mathrm{m\,a^{-1}}$ for velocities and 12.5 $\mathrm{m\,a^{-1}}$ to 21.5 $\mathrm{m\,a^{-1}}$ for thickness rate of change. The friction coefficient β is well reconstructed for both methods. The corresponding average relative error (with respect to each reference) on basal basal shear stress varies from 22 to 30 % for ANC and 20 to 28 % for ATP , still according to the amplitude of the perturbation. Both algorithms also allow to improve the knowledge of the bedrock elevation z_b with regard to the first guess. We notice a tendency to overestimate the amplitude of bumps and pits in some locations which generally corresponds to an under estimation in the amplitude of variations of β. This last behaviour highlights the limits of the algorithms and the difficulty of distinguishing the effects of two basal parameters as closely linked as the friction and the bedrock topography. This behaviour had been already highlighted in Goldberg and Heimbach (2013) and Gudmundsson and Raymond (2008) where a higher ice thickness with respect to the reference is compensated by a higher basal friction and conversely.

4.4 Flow divergence in transient model

In this section, we assess the impact of our initialisation algorithms on the prognostic response of the model forward in time assuming the same constant forcing used to build the reference state. By doing so, if the initialisation was perfect, one would expect no change of the geometry and ice flow during this prognostic simulation. The experiment is performed from ATP and ANC initial states. A third initialisation state is constructed for which only the friction coefficient has been optimised, keeping z_b equals to the $a\ priori\ z_{b_b}$. This third initialisation, called "β only" involves a rms misfit on velocities of 43.3 $\mathrm{m\,a^{-1}}$ and an average relative error $\varepsilon_{\tau_{b,ref}}$ of 36 % on basal shear stress, similar to the ATP and ANC initial states. However, the rms misfit on the thickness rate of change is significantly higher, 147.8$\mathrm{m\,a^{-1}}$.

The prognostic simulations are conducted on a 10 years period in order to see how the initial thickness rate evolves during this time and how it impacts the final ice thickness and ice surface. The thickness rate of change after 1 and 10 years of simulation are shown in Figs. 8a and b respectively, while the mismatch on the surface elevation after 10 years is shown in Fig 8c.

ANC and ATP initial states involve thickness rates of change much closer to zero than the optimisation of "β only". This also leads to a lower mismatch ε_s on the surface elevation with respect to the reference after 10 years of simulation. Indeed, this mismatch is well below 20 m for both ANC and ATP, except on a few kilometres in the upstream region, whereas the optimisation of β only gives rise to a mismatch globally above 20 m with some regions exceeding 50 m.

In that way, the two algorithms implemented in this study show substantial improvements compared to the optimisation of "β only". We especially notice a better reproduction of low scale varia-

tions of the surface elevation due to the transfer of similar variations from the bedrock elevation z_b (Fig. 9). These variations tend to disappear with the optimisation of the friction only, giving rise to a lower resolution of the surface. However, we should point out that this direct transfer of bedrock variations to the surface is a consequence of the SSA ice-flow model used and that a full-Stokes model would produce a more diffusive transfer response .

5 Conclusions

The presented algorithms allow the reconstruction of two poorly known parameters: the bedrock topography z_b and the friction coefficient β at the same time.

The optimisation of these two parameters mainly relies on the knowledge of some other data easier to measure: ice surface velocities and thickness rates of change. Some local measurements of bedrock elevation and associated errors are necessary in order to define a background z_b. The two algorithms aim to infer the set of parameters which minimises the misfit between the model and the corresponding observations of ice surface velocities and thickness rates of change. If the optimisation of ice surface velocities is usually sufficient to infer β, the inference of a second parameter requires more information to distinguish the effects of each parameters on the flow. Observations of rates of change of ice thickness are necessary to allow optimising z_b as well.

The two algorithms are based on the optimisation of the friction coefficient β with the adjoint method. The bedrock geometry z_b is reconstructed in two different ways, again with adjoint method for the first algorithm (ATP) and with a nudging method based on mass conservation equation for the second one (ANC).

We have shown that ATP algorithm reproduces β well and the corresponding basal shear stresses while z_b is particularly well reproduced in high velocities regions but does not depart from the background when the velocities become lower. The iterative algorithm coupling adjoint method and nudging (ANC) gives as good results. Moreover, ANC allows a better reconstruction of the bedrock geometry z_b in most regions. This is a very good sign for an adaptation of the method to non depth-integrated flow models such as full-Stokes models where the bedrock topography is no more a state variable but affects the domain geometry making the derivation of the adjoint even more demanding (Perego et al., 2014). ~~Indeed, there is no need to inverse a shape variable like bedrock topography which is a usual obstacle to adjoint-based methods.~~

Furthermore, the transient simulations over 10 years from initial states reconstructed with the two algorithms developed give very encouraging results. The model divergence is clearly decreased with respect to usual inversion methods of the friction coefficient only. The integration of observations like thickness rates variation through an optimisation of the divergence during inversion or nudging steps, allows to regularise the solution in a physical way and also clearly improves the results.

Finally, the sensitivity experiments shows that the different algorithms can take into account the disequilibrium of mass balance, which is particularly interesting considering that a large amount of outlet glaciers in both Greenland and Antarctica present this feature.

References

Arthern, R. J. and Gudmundsson, G. H.: Initialization of ice-sheet forecasts viewed as an inverse Robin problem, Journal of Glaciology, 56, 527–533, doi:10.3189/002214310792447699, 2010.

Aðalgeirsdóttir, G., Aschwanden, A., Khroulev, C., Boberg, F., Mottram, R., Lucas-Picher, P., and Christensen, J.: Role of model initialization for projections of 21st-century Greenland ice sheet mass loss, Biocontrol Science and Technology, 60, 782–794, doi:10.3189/2014JoG13J202, 2014.

Bamber, J. L., Griggs, J. A., Hurkmans, R., Dowdeswell, J. A., Gogineni, S. P., Howat, I., Mouginot, J., Paden, J., Palmer, S., and Rignot, E.: A new bed elevation dataset for Greenland., Cryosphere, 7, 2013.

Bindschadler, R. A.: Jakobshavns Glacier drainage basin: A balance assessment, Journal of Geophysical Research: Oceans, 89, 2066–2072, doi:10.1029/JC089iC02p02066, 1984.

Bindschadler, R. A., Nowicki, S., Abe-Ouchi, A., Aschwanden, A., Choi, H., Fastook, J., Granzow, G., Greve, R., Gutowski, G., Herzfeld, U., Jackson, C., Johnson, J., Khroulev, C., Levermann, A., Lipscomb, W. H., Martin, M. A., Morlighem, M., Parizek, B. R., Pollard, D., Price, S. F., Ren, D., Saito, F., Sato, T., Seddik, H., Seroussi, H., Takahashi, K., Walker, R., and Wang, W. L.: Ice-sheet model sensitivities to environmental forcing and their use in projecting future sea level (the SeaRISE project), Journal of Glaciology, 59, 195–224, doi:10.3189/2013JoG12J125, 2013.

Blayo, E., Verron, J., and Molines, J. M.: Assimilation of TOPEX/POSEIDON altimeter data into a circulation model of the North Atlantic, Journal of Geophysical Research, 99, 24 691, doi:10.1029/94JC01644, 1994.

Bonan, B., Nodet, M., Ritz, C., and Peyaud, V.: An ETKF approach for initial state and parameter estimation in ice sheet modelling, Nonlin. Processes Geophys., 21, 569–582, doi:10.5194/npg-21-569-2014, 2014.

Durand, G., Gagliardini, O., Favier, L., Zwinger, T., and le Meur, E.: Impact of bedrock description on modeling ice sheet dynamics, Geophysical Research Letters, 38, L20 501, doi:10.1029/2011GL048892, 2011.

Edwards, T. L., Fettweis, X., Gagliardini, O., Gillet-Chaulet, F., Goelzer, H., Gregory, J. M., Hoffman, M., Huybrechts, P., Payne, A. J., Perego, M., Price, S., Quiquet, A., and Ritz, C.: Effect of uncertainty in surface mass balance–elevation feedback on projections of the future sea level contribution of the Greenland ice sheet, The Cryosphere, 8, 195–208, doi:10.5194/tc-8-195-2014, 2014.

Flament, T. and Rémy, F.: Dynamic thinning of Antarctic glaciers from along-track repeat radar altimetry, Journal of Glaciology, 58, 830–840, doi:10.3189/2012JoG11J118, 2012.

Fretwell, P., Pritchard, H. D., Vaughan, D. G., Bamber, J. L., Barrand, N. E., Bell, R., Bianchi, C., Bingham, R. G., Blankenship, D. D., Casassa, G., Catania, G., Callens, D., Conway, H., Cook, A. J., Corr, H. F. J., Damaske, D., Damm, V., Ferraccioli, F., Forsberg, R., Fujita, S., Gim, Y., Gogineni, P., Griggs, J. A., Hindmarsh, R. C. A., Holmlund, P., Holt, J. W., Jacobel, R. W., Jenkins, A., Jokat, W., Jordan, T., King, E. C., Kohler, J., Krabill, W., Riger-Kusk, M., Langley, K. A., Leitchenkov, G., Leuschen, C., Luyendyk, B. P., Matsuoka, K., Mouginot, J., Nitsche, F. O., Nogi, Y., Nost, O. A., Popov, S. V., Rignot, E., Rippin, D. M., Rivera, A., Roberts, J., Ross, N., Siegert, M. J., Smith, A. M., Steinhage, D., Studinger, M., Sun, B., Tinto, B. K., Welch, B. C., Wilson, D., Young, D. A., Xiangbin, C., and Zirizzotti, A.: Bedmap2: improved ice bed, surface and thickness datasets for Antarctica, The Cryosphere, 7, 375–393, doi:10.5194/tc-7-375-2013, 2013.

Fürst, J. J., Durand, G., Gillet-Chaulet, F., Merino, N., Tavard, L., Mouginot, J., Gourmelen, N., and Gagliardini, O.: Assimilation of Antarctic velocity observations provides evidence for uncharted pinning points,

The Cryosphere, 9, 1427–1443, doi:10.5194/tc-9-1427-2015, http://www.the-cryosphere.net/9/1427/2015/, 2015.

Gagliardini, O., Zwinger, T., Gillet-Chaulet, F., Durand, G., Favier, L., de Fleurian, B., Greve, R., Malinen, M., Martín, C., Råback, P., Ruokolainen, J., Sacchettini, M., Schäfer, M., Seddik, H., and Thies, J.: Capabilities and performance of Elmer/Ice, a new-generation ice sheet model, Geosci. Model Dev., 6, 1299–1318, doi:10.5194/gmd-6-1299-2013, 2013.

Gilbert, J. C. and Lemaréchal, C.: Some numerical experiments with variable-storage quasi-Newton algorithms, Mathematical Programming, 45, 407–435, doi:10.1007/BF01589113, 1989.

Gillet-Chaulet, F., Gagliardini, O., Seddik, H., Nodet, M., Durand, G., Ritz, C., Zwinger, T., Greve, R., and Vaughan, D. G.: Greenland ice sheet contribution to sea-level rise from a new-generation ice-sheet model, The Cryosphere, 6, 1561–1576, doi:10.5194/tc-6-1561-2012, 2012.

Goldberg, D. N. and Heimbach, P.: 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, 2013.

Gudmundsson, G. H. and Raymond, M.: On the limit to resolution and information on basal properties obtainable from surface data on ice streams, The Cryosphere Discuss., 2, 413–445, doi:10.5194/tcd-2-413-2008, 2008.

Hoke, J. E. and Anthes, R. A.: The Initialization of Numerical Models by a Dynamic-Initialization Technique, Monthly Weather Review, 104, 1551–1556, doi:10.1175/1520-0493(1976)104<1551:TIONMB>2.0.CO;2, 1976.

Jay-Allemand, M., Gillet-Chaulet, F., Gagliardini, O., and Nodet, M.: Investigating changes in basal conditions of Variegated Glacier prior to and during its 1982–1983 surge, The Cryosphere, 5, 659–672, doi:10.5194/tc-5-659-2011, 2011.

Joughin, I., Howat, I. M., Fahnestock, M., Smith, B., Krabill, W., Alley, R. B., Stern, H., and Truffer, M.: Continued evolution of Jakobshavn Isbrae following its rapid speedup, Journal of Geophysical Research: Earth Surface, 113, F04 006, doi:10.1029/2008JF001023, 2008.

Joughin, I., Smith, B. E., Howat, I. M., Scambos, T., and Moon, T.: Greenland flow variability from ice-sheet-wide velocity mapping, J. Glaciol., 56, 415–430, 2010.

Joughin, I., Smith, B. E., Shean, D. E., and Floricioiu, D.: Brief Communication: Further summer speedup of Jakobshavn Isbræ, The Cryosphere, 8, 209–214, doi:10.5194/tc-8-209-2014, 2014.

MacAyeal, D. R.: Large-scale ice flow over a viscous basal sediment: Theory and application to ice stream B, Antarctica, Journal of Geophysical Research: Solid Earth, 94, 4071–4087, doi:10.1029/JB094iB04p04071, 1989.

Macayeal, D. R.: A tutorial on the use of control methods in ice-sheet modeling, Journal of Glaciology, 39, 91–98, 1993.

Martin, N. and Monnier, J.: Of the gradient accuracy in Full-Stokes ice flow model: basal slipperiness inference, The Cryosphere Discussions, 7, 3853–3897, doi:10.5194/tcd-7-3853-2013, 2013.

Morlighem, M., Rignot, E., Seroussi, H., Larour, E., Ben Dhia, H., and Aubry, D.: Spatial patterns of basal drag inferred using control methods from a full-Stokes and simpler models for Pine Island Glacier, West Antarctica, Geophysical Research Letters, 37, L14 502, doi:10.1029/2010GL043853, 2010.

Morlighem, M., Rignot, E., Seroussi, H., Larour, E., Ben Dhia, H., and Aubry, D.: A mass conservation approach for mapping glacier ice thickness, Geophysical Research Letters, 38, L19 503, doi:10.1029/2011GL048659, 2011.

Nowicki, S., Bindschadler, R. A., Abe-Ouchi, A., Aschwanden, A., Bueler, E., Choi, H., Fastook, J., Granzow, G., Greve, R., Gutowski, G., Herzfeld, U., Jackson, C., Johnson, J., Khroulev, C., Larour, E., Levermann, A., Lipscomb, W. H., Martin, M. A., Morlighem, M., Parizek, B. R., Pollard, D., Price, S. F., Ren, D., Rignot, E., Saito, F., Sato, T., Seddik, H., Seroussi, H., Takahashi, K., Walker, R., and Wang, W. L.: Insights into spatial sensitivities of ice mass response to environmental change from the SeaRISE ice sheet modeling project I: Antarctica, Journal of Geophysical Research: Earth Surface, 118, 1002–1024, doi:10.1002/jgrf.20081, 2013a.

Nowicki, S., Bindschadler, R. A., Abe-Ouchi, A., Aschwanden, A., Bueler, E., Choi, H., Fastook, J., Granzow, G., Greve, R., Gutowski, G., Herzfeld, U., Jackson, C., Johnson, J., Khroulev, C., Larour, E., Levermann, A., Lipscomb, W. H., Martin, M. A., Morlighem, M., Parizek, B. R., Pollard, D., Price, S. F., Ren, D., Rignot, E., Saito, F., Sato, T., Seddik, H., Seroussi, H., Takahashi, K., Walker, R., and Wang, W. L.: Insights into spatial sensitivities of ice mass response to environmental change from the SeaRISE ice sheet modeling project I: Antarctica, Journal of Geophysical Research: Earth Surface, 118, 1002–1024, doi:10.1002/jgrf.20081, 2013b.

Perego, M., Price, S., and Stadler, G.: Optimal initial conditions for coupling ice sheet models to Earth system models, Journal of Geophysical Research: Earth Surface, 119, 1894–1917, doi:10.1002/2014JF003181, 2014.

Petra, N., Zhu, H., Stadler, G., Hughes, T. J., and Ghattas, O.: An inexact Gauss–Newton method for inversion of basal sliding and rheology parameters in a nonlinear Stokes ice sheet model, Journal of Glaciology, 58, 889–903, doi:10.3189/2012JoG11J182, 2012.

Pralong, M. R. and Gudmundsson, G. H.: Bayesian estimation of basal conditions on Rutford Ice Stream, West Antarctica, from surface data, Journal of Glaciology, 57, 315–324, doi:10.3189/002214311796406004, 2011.

Raymond, M. J. and Gudmundsson, G. H.: Estimating basal properties of ice streams from surface measurements: a non-linear Bayesian inverse approach applied to synthetic data, The Cryosphere, 3, 265–278, doi:10.5194/tc-3-265-2009, 2009.

Schäfer, M., Zwinger, T., Christoffersen, P., Gillet-Chaulet, F., Laakso, K., Pettersson, R., Pohjola, V. A., Strozzi, T., and Moore, J. C.: Sensitivity of basal conditions in an inverse model: Vestfonna ice cap, Nordaustlandet/Svalbard, The Cryosphere, 6, 771–783, doi:10.5194/tc-6-771-2012, 2012.

Seroussi, H., Morlighem, M., Rignot, E., Larour, E., Aubry, D., Ben Dhia, H., and Kristensen, S. S.: Ice flux divergence anomalies on 79north Glacier, Greenland, Geophysical Research Letters, 38, L09 501, doi:10.1029/2011GL047338, 2011.

Shannon, S. R., Payne, A. J., Bartholomew, I. D., Broeke, M. R. v. d., Edwards, T. L., Fettweis, X., Gagliardini, O., Gillet-Chaulet, F., Goelzer, H., Hoffman, M. J., Huybrechts, P., Mair, D. W. F., Nienow, P. W., Perego, M., Price, S. F., Smeets, C. J. P. P., Sole, A. J., Wal, R. S. W. v. d., and Zwinger, T.: Enhanced basal lubrication and the contribution of the Greenland ice sheet to future sea-level rise, Proceedings of the National Academy of Sciences, 110, 14 156–14 161, doi:10.1073/pnas.1212647110, 2013.

Thorsteinsson, T., Raymond, C. F., Gudmundsson, G. H., Bindschadler, R. A., Vornberger, P., and Joughin, I.: Bed topography and lubrication inferred from surface measurements on fast-flowing ice streams, Journal of Glaciology, 49, 481–490, doi:10.3189/172756503781830502, 2003.

van Pelt, W. J. J., Oerlemans, J., Reijmer, C. H., Pettersson, R., Pohjola, V. A., Isaksson, E., and Divine, D.: 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, 2013.

Verron, J.: Nudging satellite altimeter data into quasi-geostrophic ocean models, Journal of Geophysical Research: Oceans, 97, 7479–7491, doi:10.1029/92JC00200, 1992.

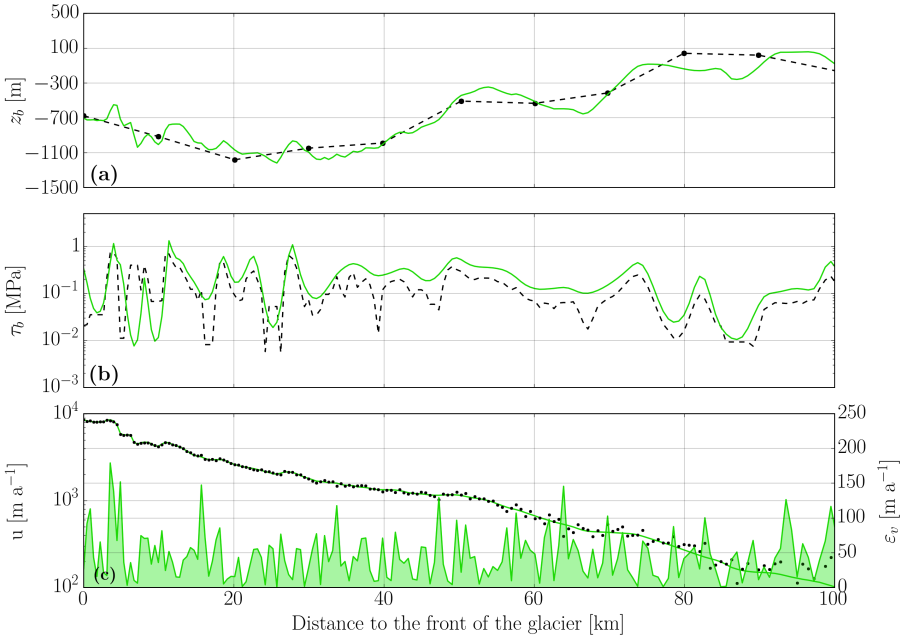

Figure 1. Reference (solid lines) and initial (dashed lines) state for **(a)** the bedrock elevations z_b, **(b)** the estimated ~~friction coefficient β~~ basal traction $\tau_b = \beta u$ and **(c)** the surface velocities. In **(a)**, synthetic observations every 10 km are the plain black circles. In **(c)** the "observed" velocities are depicted by the circles and the shaded green curve is the absolute difference between observed and reference surface velocities (right axis).

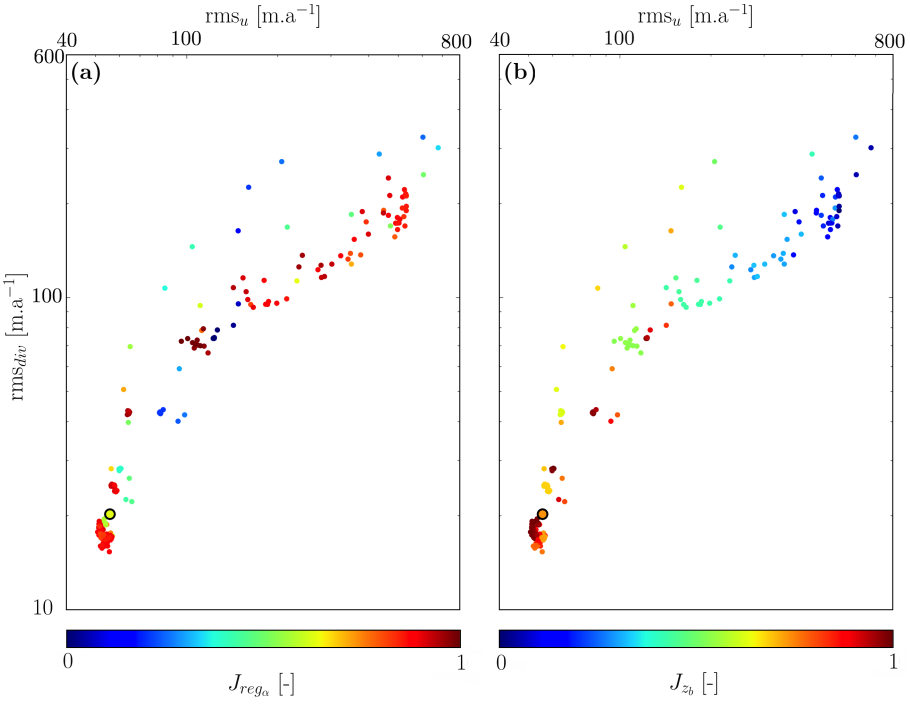

Figure 2. Mean error on the thickness rate of change (rms$_{div}$) as a function of the mean error on velocity (rms$_u$) for the 255 pairs of regularisation parameters (λ_α, λ_{z_b}). Colors scales show the normalised regularisation terms **(a)** J$_{reg_\alpha}$ and **(b)** J$_{z_b}$ (0 corresponds to the lowest value and 1 to the highest value obtained with the 255 pairs). The chosen value ($\lambda_\alpha = 10^{11}$, $\lambda_{z_b} = 10^7$) is shown with a black circle.

Figure 3. Results of the ATP algorithm with (orange) and without (red) optimisation of J_{div}, i.e. $\gamma = 1$ or $\gamma = 0$, respectively, in Eq. (10): **(a)** absolute difference between observed and model velocities, **(b)** estimated ~~friction parameter~~ basal traction and **(c)** estimated bedrock elevation. The green shaded area is the difference between the noisy reference velocities and the true velocities. The green solid lines are the reference values and the black dashed line is the initial guess for the bedrock elevation.

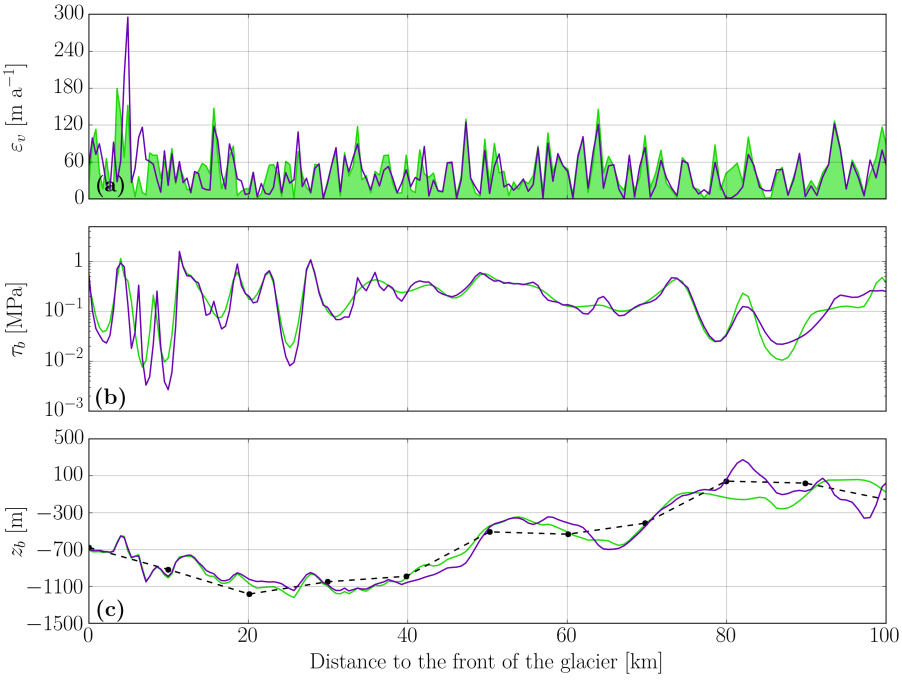

Figure 4. Results of the ANC algorithm (purple): **(a)** absolute difference between observed and model velocities, **(b)** estimated ~~friction parameter~~ basal traction and **(c)** estimated bedrock elevation. The green shaded area is the difference between the noisy reference velocities and the true velocities. The green solid lines are the reference values and the black dashed line is the initial guess for the bedrock elevation.

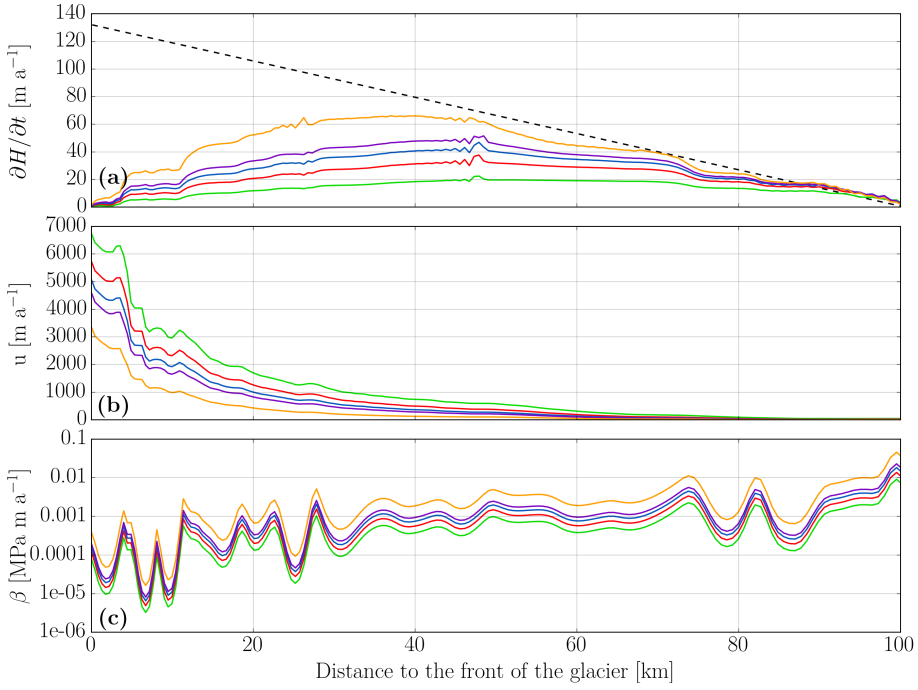

Figure 5. The five new references build from a 5-years perturbation of the initial reference by an increase of the friction parameter: $\beta = 2\beta_{ref}$ (green), $\beta = 3\beta_{ref}$ (red), $\beta = 4\beta_{ref}$ (blue lines), $\beta = 5\beta_{ref}$ (purple) and $\beta = 10\beta_{ref}$ (orange). New references for **(a)** the thickness rate of change for the different perturbations, **(b)** velocities (without observation noise) and **(c)** friction coefficients β.

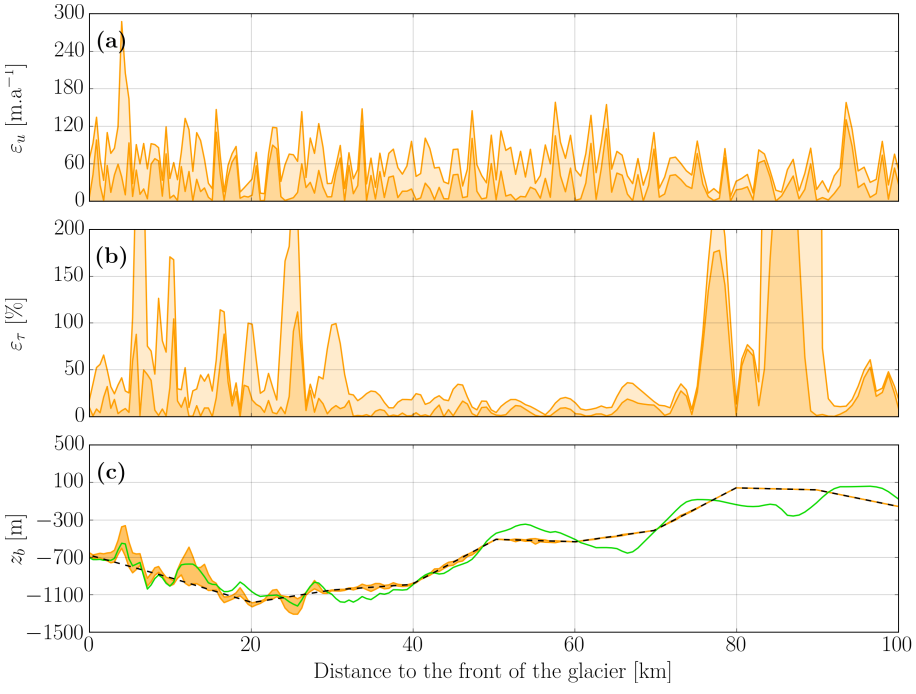

Figure 6. Range of values for ATP algorithm for the 5 perturbations of the friction coefficient β. **(a)** and **(b)** minimum (dark orange shade) and maximum (light orange shade) of absolute difference between observed and model velocities and relative error for ~~β~~ τ_b respectively. **(c)** range of values for bedrock elevation z_b (orange shade). The green solid line is the reference value and the black dashed line is the initial guess for the bedrock elevation.

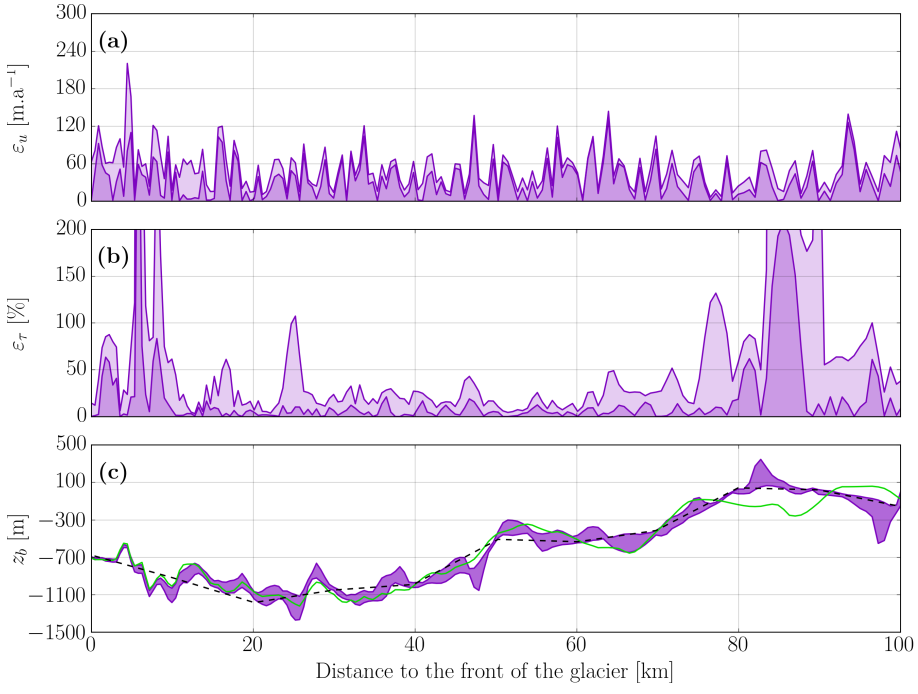

Figure 7. Same as Fig. 5 but for the ANC algorithm.

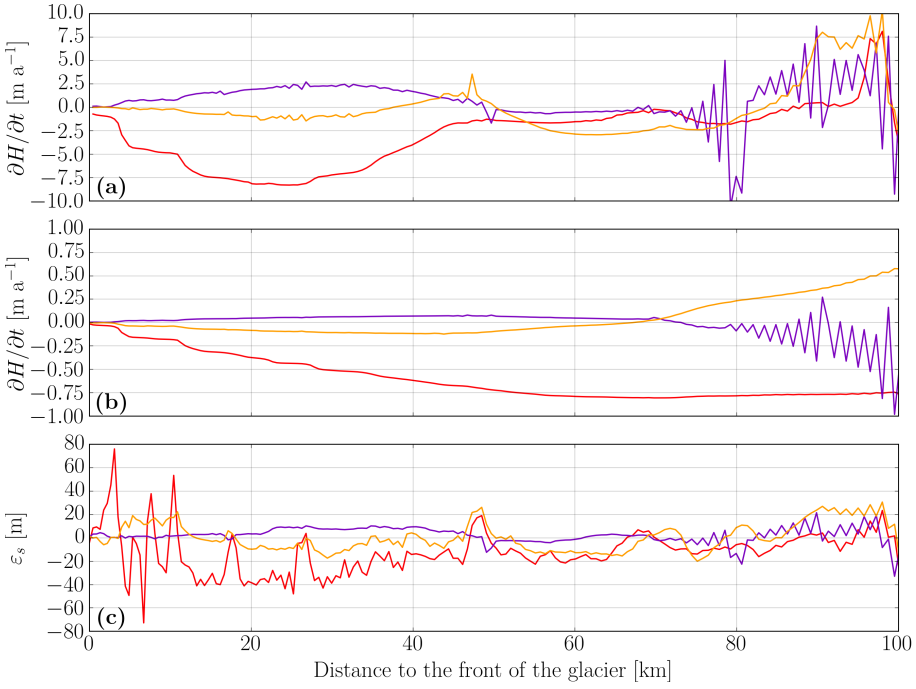

Figure 8. Evolution of $\partial H/\partial t$ after 1 year **(a)** and 10 years **(b)** of prognostic simulation and the resulting missmatch after 10 years between surfaces obtained with three different initial states and reference surface **(c)**. The orange and purple lines give the results for ATP and ANC . The red line gives the result for inversion of β only.

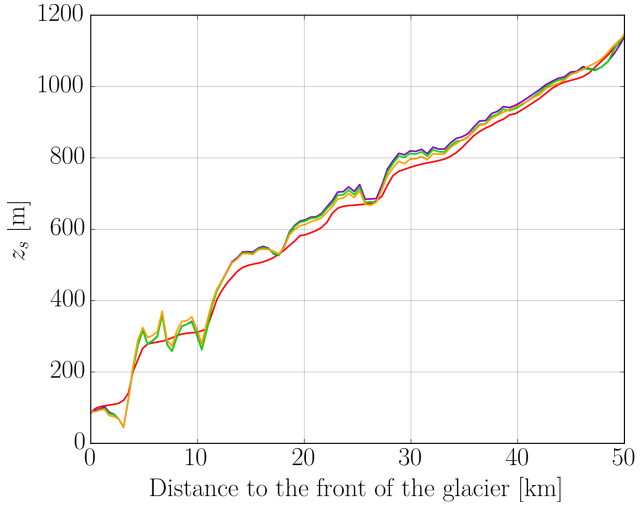

Figure 9. Ice surface elevation z_s after 10 years of prognostic simulation for 3 different initial states : initialisation with ATP algorithm (orange line), with ANC algorithm (purple line) and with the inversion of β only (red line). The green line is the reference surface elevation. The figure focuses on the first 50 km next to the front of the glacier.