Inference of the bottom topography in anisothermal mildly-sheared shallow ice flows

Jerome Monnier, Jiamin Zhu

Inference of the bottom topography in anisothermal mildly-sheared shallow ice flows

Jérôme Monnier [*]   Jiamin Zhu [*]

Abstract

This study proposes a new inverse method to estimate the bed topography elevation beneath glaciers flows from surface observations (altimetry elevations and InSAR velocities) and (very) sparse depth measurements (e.g. acquired during airborne campaigns). To do so an original form of depth-integrated flow equations (long-wave assumption) is derived. The latter are valid for highly to mildly-sheared regimes hence including moderately fast flows; varying internal thermal profiles are taken into account. The inverse problem is particularly challenging since the surface signatures integrate the bottom features (bed elevation and friction-slip amount) plus the internal deformation. The first key ingredient of the inverse method is the derivation of this non-isothermal Reduced Uncertainty (RU) version of the classical SIA equation (lubrication type model for generalised Newtonian fluids) by intrinsically integrating the surface measurements in the formulation. This resulting multi-physics RU-SIA equation contains a unique uncertain dimensionless parameter only (the parameter γ). The next key ingredient is an advanced Variational Data Assimilation (VDA) formulation combined with a purely data-driven extension of γ based on the trend observed in the (sparse) depth measurements (e.g. along the flight tracks). The resulting inverse method provides the first physical-based depth estimations in mildly-sheared mildly-slippery shallow flows. In poorly monitored ice-sheet areas (e.g. in Antarctica), the resulting estimations are noticeably less uncertain than the current ones (in particular compared to those obtained by gravimetry inversions). The present numerical experiments and experimental sensitivity analyses demonstrate the reliability of this new RU-SIA equation and the robustness of the inverse method. In other respect, the RU-SIA equation may provide a-posteriori estimations of the thermal boundary layer at bottom.

1 Introduction

Bed (bottom) topography elevation is a basic data to set up any numerical geophysical flow model. In glaciology this data is often very poorly known. Inverse methods to infer the topography beneath the flows are the only alternative. In ice-sheets (Antarctica, Greenland), ice thickness measurements are available along airborne radio-echo sounding tracks (e.g. data products from CReSIS, Univ. of Kansas and NASA Operation IceBridge) providing bed elevation measurements. The latter are dense in fast streams costal areas but very sparse elsewhere. In other respects numerous satellites provide accurate measurements of the ice sheets surfaces: altimeters provide surface elevations H at $\approx \pm 10 - 30$ cm for 1 km^2 pixels see e.g. [20], while radar interferometers (InSar) provide accurate surface velocity fields $\mathbf{u}_H$ as soon as $|\mathbf{u}_H| \gtrsim 5$ m/y, see e.g. [40].

Out of highly measured areas (that is out of fast streams costal areas) the current bed topography estimations are based on an extension of the direct airborne measurements (along the relatively sparse flight tracks) by Kriging. The resulting bed topography maps are presented in [2] for Greenland and in [13] for Antarctica. In poorly measured areas (e.g. deep inland Antarctica) at distance greater than 50 km from thickness measurements, the estimations are based on gravity field inversions hence presenting very large uncertainties [13]. On the contrary, in fast streams nearshore areas, the inversion of

---

regularized versions of the depth-averaged mass equation combined with altimetry data enables to fill up more accurately the gaps. Indeed by following the streamlines downstream or upstream the measurements, the inversion of a regularized mass equation (linear transport) is possible, see [45, 37, 38]. Up to now, no physical-based inversions have been performed out of these densely measured fast streams areas. To do so, the considered model needs to be physically consistent and the inversions need to be stable, robust. This is a real challenge, see e.g. [31, 32] and [48]. Moreover, inland where the estimated bed elevations are the most uncertain (since less monitored), the measured surface velocities are due in part to the internal ice deformation (sheared flow) and in part to slipperiness at the base; that is mildly-sheared mildly-slippery flows. Then inverting the surface data becomes much more challenging compared to fast streams - pure slipping flows since the bed topography surface signature needs to be separated from the basal slip one (and the internal deformation).

An adequate physical model class to consider for the inversions are shallow flow models since rich enough (mass and momentum conservations may be taken into account) but simpler than the fully 3D free surface model. Complete 3D free surface models are extremely complex to invert in a geophysical context since the lack of data and since representing multi-scale multi-physics phenomena. Their inversions would very likely lead to severe equifinally issues (due to the underlying ill-posed inverse problems). Basically, more the model is complex with numerous uncertain unknown parameters, more its inversion is challenging.

In glaciers, the vertical internal deformation (sheared viscoplastic fluid) is highly dependent on the vertical profile of the creep parameter (sometimes called rate factor in glaciology), therefore highly dependent on the vertical temperature profile. This additional phenomena makes the inversion of ice flows even more challenging. Finally the unknowns of the considered flow models to be inferred are : the bed (bottom) topography elevation, the basal slipperiness (or equivalently the friction parameter at bottom) and non uniform creep parameter (due to varying vertical thermal profiles). This is the challenging inverse problem which is addressed in the present study.

It should be noted that effective bed topography profiles only can be infered since the flows act as low-band filters. Indeed, the bed variations are filtered by the flow; the filtering features depending on the flow regime, see [16, 17, 29, 30] for detailed analysis applied to generalised Newtonian fluids including glaciers.

The inference of the bed topography, but not of the composite unknown (bed topography, friction coefficient) has been addressed in numerous studies by inverting ice models, see e.g. [45, 55, 52, 37, 38, 19, 49, 31, 32]. These studies consider restricted flow regimes only since considering either purely slipping ice-streams (fast plug-like flows in coastal areas) or fully sheared flows (extremely slow flows, no slip at bottom). These studies are sometimes based on flow models but always incomplete ones. Typically the mass equation of plug-like flows (providing good estimations if the airborne measurements are dense and cross-lines, see e.g. [38]), or including the momentum equations too but isothermal and with no-slip at bottom, see e.g. [32] and the detailed review made in [34]. [34] has the same goal as the present study one, that is inferring the bed topography in non isothermal mildly-sheared mildly-slippery ice flows. However in [34] the vertical thermal profile was supposed to be given hence simplifying the inverse problem.

In the present study, firstly a dedicated shallow flow model is derived: the so-called RU-SIA model (Reduced Uncertainty - SIA). This model is a reformulation of the extended Shallow Ice Approximation (xSIA) model by natively integrating the surface measurements (elevation and velocity) and taking into account varying vertical temperature profiles. Recall that the SIA model derives from the classical lubrication theory applied to power-law rheology fluids by neglecting the inertial terms (creeping flows), see [12, 22, 36, 4]. [5] demonstrates the validity of xSIA model for moderate slipping amounts at bottom (i.e. the present targeted flow regimes) by stating formal error estimates. These mildly-sheared mildly-slippery regimes correspond to ice flows with surface velocity $\approx 10 - 50$ m/y, that is the targeted ice-sheet interior sectors and ice-sheds upstreams. Recall that these areas are poorly measured areas (airborne data are currently non-existent or extremely sparse).

In [34], the classical xSIA model has been enriched by taking into account a varying vertical creep

parameter profile. The present new RU-SIA equation is a diffusive equation like the classical SIA equation or the non-isothermal version derived in [34] but containing a single dimensionless parameter denoted γ (hence the Reduced Uncertainty feature). The latter contains all the unknown-uncertain multi-physics terms: basal friction - slip ratio, varying creep parameter vertical profile and rheology power-law exponent. An explicit expression of this single parameter is obtained.

Next the inversion algorithm relies on this Reduced Uncertainty physical model and two Variational Data Assimilation (VDA) processes, see e.g. [46, 25] and [6, 33, 7]. The latter aim at optimally fusing the RU-SIA model and the altimetry data (surface elevation measurements) plus sparse airborne measurements. These two VDA process are based on gradient-based minimisation and employ adjoint equations [26]. Given the measured values of depth along flight tracks, a reliable value γ is infered by one of the VDA process above. Next, its extension to the entire domain is performed by an universal Kriging estimator, see e.g. [8]. The surface velocity information (derived from InSAR measurements) is contained in a RU-SIA equation coefficient.

The VDA formulations rely on linear-quadratic optimal control problems (hence robust) with prior covariance operators and changes of the control variables. These latter are an unknown effective diffusivity, the uncertain source mass balance and the ice depth (thickness).

The novelties and strengths of these first inversions valid for anisothermal creeping shallow flows, from slow to mildly rapid, are numerous. In particular they provide: a) robust estimations independently of the in-situ (airborne) measurements locations contrarily to the intrinsically unstable inversions of the transport mass equation; b) reduced uncertainty depth estimations in very poorly monitored areas (e.g. deep inland Antarctica); c) an a-posteriori estimation of the basal thermal layer (given an a-priori vertical profile). Moreover since based on 2D (shallow) flow equations, these estimations remains affordable even for large computational domains.

The elaborated algorithms have been implemented in Python using the Fenics library [1, 27] and optimization libraries; it is part of DassFlow computational code [35]. A complete real data set of an ~ 200 km $\times 250$ km inland Antarctica area is considered; the elaborated inverse method provides a new bed topography estimation. The data set includes: the current bed estimation (Bedmap2, [13]) which provides the first guess value of the VDA processes (and a reference value), surface elevations H [20], surface velocities magnitudes $|\mathbf{u}_H|$ [40] and the climatic forcing term $\dot{a}$ [41]. The robustness of the inversion method is analysed into details through numerous numerical experiments. The area has been randomly chosen in function of its surface velocities ($|\mathbf{u}_H|$ varies from ≈ 10 to 30 m/y hence in the model validity range) and since it is sufficiently surveyed (~ 10 flights tracks) to assess the accuracy and robustness the algorithms.

The outline of the article is as follows. In Section 2, the non-isothermal xSIA model is recalled and the Reduced Uncertainty (RU) version is derived, providing the so-called RU-SIA model (model containing the single dimensionless parameter γ). A typical uncertainty estimation of γ is derived and the global inversion method is sketched. In Section 3, values of γ are estimated along the flight tracks by a VDA process. Next from an observed clear trend, γ is extended in the whole domain by an universal Kriging method. In Section 4, given γ, the ice depth (thickness) h and the RHS $\dot{a}$ are simultaneously estimated by VDA, therefore providing new estimations of the bed topography elevations. In Section 5, numerical results with numerous numerical sensitivity experiments are proposed for an inland Antarctica area. A conclusion is proposed in Section 6.

2   The direct model and the inversion algorithm

In this section the extended SIA (xSIA) model (that is the classical SIA equation with non vanishing basal velocity see e.g. [11]) is recalled. Next we derive the so-called Reduced-Uncertainty SIA (RU-SIA) model. The latter is obtained by reformulating the equations by gathering the few uncertain terms into a single dimensionless parameter γ. Finally the global inversion algorithm is presented.

2.1 The xSIA equation

Recall that the classical SIA model is derived from asymptotic calculations of the free surface Stokes equations with respect to the geometrical ratio $\varepsilon = \frac{H^*}{L^*}$, where H^* and L^* are characteristic flow depth and length respectively, see [22, 36] also e.g. [11] Chapter 10.2. The basic assumption states that the flow, thin geometry, is sheared; in other words, normal stress components are negligible. The SIA equations are first order in ε. In [5], it is formally demonstrated that the xSIA model remains valid for a slip coefficient $C \simeq \mathcal{O}(1)$ ($C = \beta^{-1}$, β the friction coefficient), hence clarifying the xSIA domain of validity. (Notice that in [5] the xSIA equation is derived in a much more general coordinate system: the Prandtl coordinate system [42]; this form which is not considered in the present derivations may be useful in Alpine glaciers).

In all these shallow flow SIA like equations, the slip parameter C and the creep parameter A (including the thermal effects), see e.g. [15] (chapters 4 and 6), are highly uncertain. The aim of this section is to re-write the equations in a form enabling to reduce the modeling uncertainty (due to C and A). This is done by taking advantage of measurements of the surface characteristics: elevation and velocity. This new derivation leads the so-called RU-SIA model which contain a single unknown (dimensionless) parameter γ.

Depth-averaged mass equation Let us denote the ice surface elevation by H, the topography elevation by b, the ice depth by $h = (H - b)$ and the 3D ice velocity by $\mathbf{u}$. The ice depth satisfies the depth-averaged mass equation:

$$\partial_t h + div\mathbf{q} = a, \tag{1}$$

where $\mathbf{q} = h\bar{\mathbf{u}}$ is the discharge and $\bar{\mathbf{u}} = \frac{1}{h}\int_b^H \mathbf{u}(z)dz$ is the depth-averaged velocity. The source term a is the mass balance: the difference between accumulation and ablation, see e.g. [15, 11].
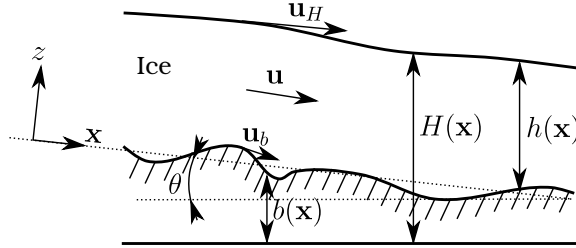

Figure 1: Schematic vertical view of the gravitational ice flow and notations

Velocity expression We define $\mathcal{S} = |\nabla H|$ the slope value, the parameter $\bar{\rho} = (\rho g \cos \theta)^q$ with ρ the ice density, g the gravity amplitude, q the power-law exponent of the rheology law, and θ the mean slope value in the (x, y)-plane, see Fig. 1. In ice-sheet modelling, it is adequate to set θ to 0. If the flow is sheared, the SIA equations may apply. The SIA velocity $\mathbf{u}(\mathbf{x}, z)$ expression reads, see e.g. [11] Chapter 10.2 or [15] Chapter 5.4:

$$\mathbf{u}(\mathbf{x}, z) = (u, v)(\mathbf{x}, z) = \mathbf{u}_b(\mathbf{x}) - 2\bar{\rho}\mathcal{S}^{q-1}(\mathbf{x})\nabla H(\mathbf{x}) \int_b^z A(\mathbf{x}, \xi)(H(\mathbf{x}) - \xi)^q d\xi \tag{2}$$

where $A(z)$ is the creep parameter provided by the constitutive law; it depends on the vertical temperature profile in the ice $A(z) \equiv A(T(z))$, see e.g. [15] Chapter 4.3. The basal velocity $\mathbf{u}_b(\mathbf{x}) \equiv \mathbf{u}(\mathbf{x}, z = b)$ reads:

$$\mathbf{u}_b(\mathbf{x}) = -C\bar{\rho}h^q\mathcal{S}^{q-1}\nabla H(\mathbf{x}) \tag{3}$$

4

with $C > 0$ the slip coefficient. For glaciers, the usual exponent value is $q = 3$, see e.g. [15, 11]. Following [34], we introduce the parameter:

$$\bar{A}(\mathbf{x}) = \frac{(q+2)}{h^{q+2}(\mathbf{x})} \int_b^H \int_b^z A(\mathbf{x}, \xi)(H(\mathbf{x}) - \xi)^q d\xi dz. \tag{4}$$

Finally the depth-averaged velocity reads:

$$\bar{\mathbf{u}}(\mathbf{x}) = -\bar{\rho} \left(C(\mathbf{x}) + \frac{2\bar{A}(\mathbf{x})}{(q+2)} h(\mathbf{x}) \right) h^q(\mathbf{x}) \mathcal{S}^{q-1}(\mathbf{x}) \nabla H(\mathbf{x}) \tag{5}$$

In the isothermal case or if the vertical profile is constant then: $\bar{A}(\mathbf{x}) = A(\mathbf{x}) \; \forall \mathbf{x}$.

The xSIA equation By injecting the velocity expression (5) into the mass equation (1) (lubrication type model), the xSIA equation valid for non isothermal flows reads, see [34]:

$$-\bar{\rho} \, div \left(\left[C + \frac{2\bar{A}}{(q+2)} h \right] h^{q+1} \mathcal{S}^{q-1} \nabla H \right) = \dot{a} \tag{6}$$

It is a non-linear diffusive equation in h (recall that $H = (b + h)$). To be solved, the values of the slip coefficient C and the depth-integrated parameter $\underline{A}$ have to be given; however they are a-priori unknown.

Notice that if $C \to 0$, the no-slip condition (adherence) is imposed. On the contrary if $C \to \infty$ a pure slip condition (vanishing friction) is imposed. To remain within the SIA model validity, the slip coefficient C has to vary from 0 to $\mathcal{O}(1)$ at most, see [5] for a detailed analysis.

2.2 The RU-SIA equation

A new form of the xSIA model (6) is derived: the so-called Reduced Uncertainty (RU-SIA) equation in which the multi-physics and unknown parameters C and $\underline{A}$ are mathematically gathered to a single dimensionless parameter γ. Next a-priori uncertainty estimations on γ are derived.

Velocity expressions including surface data measurements. Still by following [34], we introduce the parameter:

$$\underline{A}(\mathbf{x}) = \frac{(q+1)}{h^{q+1}(\mathbf{x})} \int_b^H A(\mathbf{x}, \xi)(H(\mathbf{x}) - \xi)^q d\xi \tag{7}$$

Again, in the isothermal case or in presence of a constant vertical profile, $\underline{A}(\mathbf{x}) = A(\mathbf{x}) \; \forall \mathbf{x}$. Then the surface velocity norm reads:

$$|\mathbf{u}_H| = \bar{\rho} \left(C(\mathbf{x}) + \frac{2\underline{A}(\mathbf{x})}{(q+1)} h(\mathbf{x}) \right) h^q(\mathbf{x}) \mathcal{S}^q(\mathbf{x}) \tag{8}$$

Let us introduce the observational term $\mathcal{Q}_H = \frac{|\mathbf{u}_H|}{\mathcal{S}^q}$. By re-writing the slip parameter as $C(\mathbf{x}) = \frac{\mathcal{Q}_H}{\bar{\rho}(\mathbf{x}) h^q(\mathbf{x})} - \frac{2}{(q+1)} \underline{A}(\mathbf{x}) h(\mathbf{x})$, the depth-averaged velocity (5) reads as:

$$\bar{\mathbf{u}}(\mathbf{x}) = -\frac{|\mathbf{u}_H|}{\mathcal{S}} \left(1 - \frac{2\bar{\rho} \, R_A \underline{A}}{\mathcal{Q}_H (q+1)(q+2)} h^{q+1} \right) \nabla H(\mathbf{x}) \tag{9}$$

with:

$$R_A = \frac{\bar{A}}{\underline{A}} \text{ and } c_A = [(q+2) - (q+1)R_A] \tag{10}$$

In the isothermal case, $R_A = 1 = c_A$.

Let us define the slip ratio as follows:

$$R_s = \frac{|\mathbf{u}_H| - |\mathbf{u}_b|}{|\mathbf{u}_H|} = 1 - \frac{|\mathbf{u}_b|}{|\mathbf{u}_H|} \tag{11}$$

Observe that by using (3), (8) and (11), it follows the expression: $R_s = \frac{2\bar{\rho}A}{\mathcal{Q}_H(q+1)}h^{q+1}$. (In particular the slip ratio and the flow depth are related).

Finally the depth-averaged velocity (9) re-reads as:

$$\bar{\mathbf{u}}(\mathbf{x}) = -\frac{|\mathbf{u}_H|}{\mathcal{S}} \gamma \nabla H(\mathbf{x}) \tag{12}$$

with:

$$\gamma = \left(1 - \frac{c_A}{(q+2)}R_s\right) \tag{13}$$

The depth-averaged velocity uncertainty depending on the uncertain multi-physics parameters C and $A(z)$, see (9), has been reduced in the sense it depends now to a single uncertain parameter $\gamma(q, c_A, R_s)$, see (12).

The RU-SIA equation. By injecting the depth-averaged velocity expression (12) into the depth-averaged mass equation (1), the so-called RU-SIA equation follows:

$$-div\left(\frac{|\mathbf{u}_H|}{\mathcal{S}}\gamma h \nabla H\right) = \dot{a} \tag{14}$$

Recall that the RHS reads: $\dot{a} = (a - \partial_t h)$.

The surface velocity norm $|\mathbf{u}_H|$ and the surface slope $\mathcal{S}$ may be deduced from the available surface measurements. Assuming that the depth h (or equivalently the bed elevation b) is given, γ is the single unknown parameter of this equation in variable H.

The RU-SIA equation (14) is a linear elliptic equation in H, assuming that the observational term $\frac{|\mathbf{u}_H|}{\mathcal{S}}$ and the "effective diffusivity" $\eta = \gamma h$ are given. Values of H at the boundary (Dirichlet boundary conditions) close the equation.

As a consequence, assuming that the bedrock elevation b is known, γ may be estimated by an inverse method. This is what is done in next section along the flight tracks where values of b are given (from the measurements).

As already discussed for xSIA equation, the RU-SIA equation (14) is valid if the slip ratio R_s ranges from ~ 0.3 to 1. (In this case, the basic scalings done to derive the shallow flow equations (sheared flows) remain valid, see [47, 5] for detailed discussions).

The dimensionless parameter γ and the thickness h may be considered as independent fields; γ depends on (uncertain) physical parameters and the flow regime (through the slip ratio R_s) only. However as already pointed out R_s depends implicitly on h through the flow model, see (11). Actually all the quantities of the present a-priori ill-posed inverse problem are fully coupled, either explicitly or implicitly through the flow model. However the RU-SIA equation (containing the diffusivity term $\eta = \gamma h$) enables to separate the two quantities if not considering their correlation through the flow model (14). The separation is made in next section by a purely data-driven model (i.e. without physical constraints); this is Step 2) of the inverse algorithm presented in next section. The present approach and equations are new; they constitute a step forward to obtain robust inversions.

A link between the inversion of the mass equation and the inversion of RU-SIA equation. As already mentioned, [45, 37, 38] aim at infering h by inverting the mass equation $div(h\bar{\mathbf{u}}) = \dot{a}$ with the depth-averaged velocity $\bar{\mathbf{u}}$ related to $\mathbf{u}_H$ as: $\bar{\mathbf{u}} = \alpha\mathbf{u}_H$. Next the coefficient α is empirically set to values $\lesssim 1$. Recall that the divergence operator may be inverted by introducing artificial diffusion in

highly covered areas with cross-lines flight tracks; in practice in fast streams areas where R_s vanishes therefore $|\bar{\mathbf{u}}| \lesssim |\mathbf{u}_H|$.

The xSIA equation derived in [34] is valid for moderately slip regime (including with varying vertical parameter $A(z)$ hence non-isothermal flows) where we may have $|\bar{\mathbf{u}}|$ differing from $|\mathbf{u}_H|$ by a dozen of %. The derivation of xSIA (and RU-SIA) shows that: $\bar{\mathbf{u}} = -\frac{|\nabla H|}{\mathcal{S}} \gamma \nabla H$. Therefore γ enables to explicit the expression of the empirical parameter α introduced in [45, 37, 38]. (However it should be noticed that in fast streams the uncertainty on the internal deformation represented by c_A and q in (13) is negligible).

2.3   Typical uncertainty on the single parameter γ

In this section, a-priori estimations on the single diùensionless parameter γ of RU-SIA equation (14) are derived.

Let us set $q = 3$, which is the usual exponent value employed for glaciers. Following the Arrhenius law, see e.g. [15, p.54], and typical temperature vertical profiles in Antarctica, see e.g. [44, 43, 50]), the following typical vertical profile of $A(z)$ is considered:

$$A(z) = \begin{cases} A_a & \text{for } z \in [B, H] \\ \frac{A_a}{B-b} \left((1-k)z + kB - b \right) & \text{for } z \in [b, B] \end{cases} \tag{15}$$

where A_a and k are constants, see Fig. 2. We define:

$$B = mh + b, \quad m \in [0, 1] \tag{16}$$

From (4), (7) and (10) it follows that for $q = 3$: $R_A = \frac{5m^2(1-k)\left(m^3 - 6m^2 + 15m - 20\right) + 60}{12(m(1-k)(m^3 - 5m^2 + 10m - 10) + 5)}$.

For $k \in [10, 1000]$, it follows that: $c_A \in [0.5, 5.04]$ (still for $q = 3$). In Fig. 2, the parameter c_A vs m is plotted for different values of k and q. If considering the typical value $q = 3$, the following numerical estimation holds: $\gamma \in [\approx 0.1, 1[$.
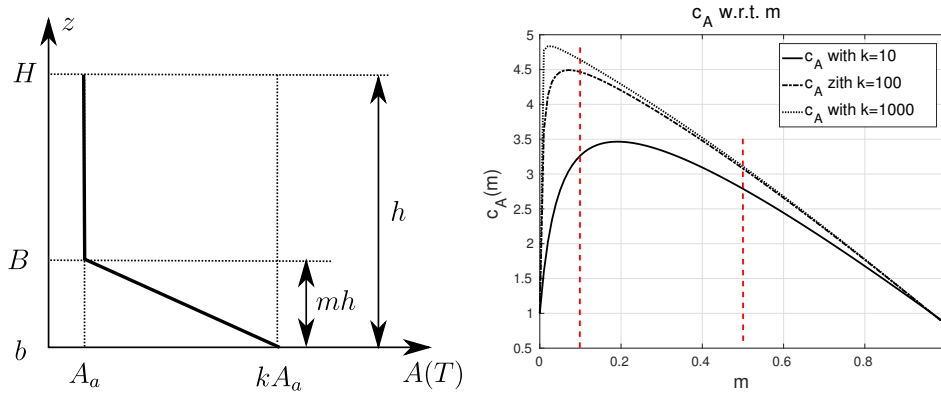

Figure 2: Left: typical vertical profile of the creep parameter $A(z)$, see (15). Right: c_A vs m (see (16) and (10)) with $k = 10$ and $k = 1000$.

In the isothermal case, $c_A = 1$, $\gamma = (1 - 0.2\,R_s)$ hence $\gamma \in [0.8, \approx 0.9]$ for xSIA comptaible regimes. Therefore in the isothermal case the uncertainty on γ is relatively small.

Both in Antarctica and in Greenland, bedrock temperatures may be assumed to be close to 0. In the Antarctica case (resp. Greenland case), the surface temperature may be equal to $-40\,C°$ (resp.

$\approx -20\,C°$), corresponding to $A_a \approx 10^{-26}$ (resp. $A_a \approx 10^{-25}$), hence $k \approx 1000$ (resp. $k \approx 10$), see Fig. 2 (Right). Moreover, assuming that the thermal boundary layer satisfies $m \in [0.1, 0.5]$, then $c_A \in [3.11, 4.64]$ (resp. $c_A \in [3.79, 3.46]$), see Fig. 2. It follows from (13) that: $\gamma \approx (1 - (0.78 \pm 0.15)R_s)$ (resp. $\gamma \approx (1 - (0.73 \pm 0.03)R_s)$).

The present basic analysis highlight orders of magnitudes of the uncertainty on γ. Moreover it shows that the uncertainty on γ arising from the vertical thermal profile through the term $(c_A/(q+2))$ is a-priori smaller than those arising from the slip ratio value R_s (in particular if $q = 3$ or greater).

2.4   The global inversion algorithm

The goal of the present study is to infer the depth (ice thickness) h. It is done by combining Variational Data Assimilation (VDA) processes into the RU-SIA equation (14) and statistical estimations. More precisely the following global inversion algorithm is considered.

- Step 1) Estimation by VDA of the effective diffusivity $\eta = \gamma h$ of the RU-SIA equation (14). The formulation is detailed in next section. Given the optimal value η^*, the value of $\gamma^* = \frac{\eta^*}{h_b}$ along the flights tracks (where depth measurements h_b are available) are kept for the next step.
  At this stage, values of the RHS $\dot{a}$ provided by Racmo2 [41, 53] are assumed to be exact.

- Step 2) Extension (interpolation / extrapolation) of γ out the flights tracks by a classical universal Kriging process.

- Step 3) Estimation of the pair $(h, \dot{a})$ by VDA.
  Both h only and $\dot{a}$ are infered; this enables to adjust the measurements of $\dot{a}$ too. According to [41, 53], the uncertainty on $\dot{a}$ is $\pm \sim 20\%$. This bound is imposed as inequality constraints in the VDA formulation.

3   Inference of the parameter γ

In this section, the method to estimate the multi-physics parameter γ is detailed; this corresponds to Step 1) and Step 2) of the global inversion algorithm sketched above.

3.1   Identification of the parameter γ along the flight tracks (Step 1)

The method to compute values of γ along the flight tracks is as follows. This computation corresponds to Step 1) of the inversion algorithm.

Estimation by VDA of the effective diffusivity $\eta = (\gamma h)$.   Given the surface measurements, the effective diffusivity $\eta = (\gamma \cdot h)$ of the RU-SIA equation (14) is infered by VDA. The following optimal control problem is solved:

$$\min_{\eta} \; j(\eta) \quad \text{with } j(\eta) = j_{obs}(\eta) + \alpha_\eta j_{reg}(\eta) \tag{17}$$

$$j_{obs}(\eta) = \frac{1}{2} \int_\Omega \chi_{tr}(\mathbf{x})|H(\eta)(\mathbf{x}) - H^{obs}(\mathbf{x})|^2 d\mathbf{x}, \quad j_{reg}(\eta) = \frac{1}{2} \int_\Omega |\nabla\eta(\mathbf{x})|^2 d\mathbf{x}$$

The control variable is η, $\eta \in \mathcal{U}_{ad}^\eta = \{\eta \in L^2(\Omega),\ \eta_{min} \le \eta \le \eta_{max}\}$; the bounds η_* are defined from the numerical estimations presented in Section 2.3: $\gamma \in\,]0, 1[$. It turns out in the numerical results (see next section) that these imposed bounds on η are not reached. The airborne data along the flights tracks are assimilated. To do so, in j_{obs} the restriction operator $\chi_{tr}(\mathbf{x})$ equals 1 if $\text{dist}(\mathbf{x}, \Gamma_{tr}) < 3\,km$;

it equals 0 otherwise. This definition is related to the length scale validity of the shallow flow model. The elevation value H_η is the solution of (14) with Dirichlet boundary conditions (which are set from the surface measurements). H is the state of the system. If the elevation H^{obs} would be assimilated everywhere in the domain then the inverse problem would be a Linear-Quadratic optimal control problem therefore admitting a unique solution η^*. However the observations are available along the flights tracks only. Although the target value is the optimal value γ^* along the flight tracks only, this inverse problem is a-priori ill-posed. Then a classical Tikhonov regularization term j_{reg} is added, see e.g. [23] and references therein. The minimisation problem (17) is numerically solved using the classical first order minimisation algorithm L-BFGS (we use the Python routine scipy.optimize.minimize). The scalar weight coefficient α_η can be chosen according to numerous rules, see e.g. [23]. In the next optimisation problem (next Section), an iterative regularisation procedure with iterative values of α_η is considered. In the present problem (17), this strategy turned out to be useless. Then various computations are performed with empirical values of α_η. (Typically these values satisfies: $\frac{j_{reg}}{j_{obs}}(\eta^*) \approx 10^{-p}$ with $p \in [2,3]$).

The VDA formulation is as follows. Given a control η, the direct model (14) is solved by a standard Lagrange Finite Element Method (FEM) order 2. Given this unique solution H_η, the following adjoint equation is solved using the same FEM :

$$-div(\frac{|\mathbf{u}_H|}{\mathcal{S}} \eta \nabla P) = \chi_{tr}(H - H^{obs}), \ \mathbf{x} \in \Omega; \quad P(\mathbf{x}) = 0, \ \mathbf{x} \in \partial\Omega.$$

The gradient of the cost functional is computed from the state H_η and the adjoint state P_η as:

$$j'(\eta) \cdot \delta\eta = \int_\Omega \left(\frac{|\mathbf{u}_H|}{\mathcal{S}} \nabla H_\eta \nabla P_\eta + \alpha_\eta \nabla\eta \nabla(\delta\eta) \right) d\mathbf{x}.$$

Finally this gradient is used in the minimisation algorithm L-BFGS to obtain a better control variable η i.e. an updated value providing a lower value of $j(\eta)$. This iterative process is performed until convergence. In practice the convergence turns out to be very robust: the optimal solution does not significantly depend on the length scale of the data smoothing nor on the restriction operator definition ξ_{tr}. Detailed sensitivity analysis are proposed in next section.

Resulting value of γ along the flight tracks. The computed optimal solution of (17) is denoted by η^*. Assuming that the depth values along the flight tracks correspond to the correct values at the scale imposed in the definition of χ_{tr}, $h^*(\mathbf{x}) = h_b(\mathbf{x})$ for $\mathbf{x} \in \Gamma_{tr}$. Then the value of γ is straightforwardly obtained as:

$$\gamma_{tr}^*(\mathbf{x}) = \frac{\eta^*(\mathbf{x})}{h_b(\mathbf{x})} \quad \text{for } \mathbf{x} \in \Gamma_{tr} \tag{18}$$

Given $\gamma_{tr}^*(\mathbf{x})$, next step of the inversion algorithm consists to estimate the value of $\gamma(\mathbf{x})$ out of the flights tracks.

Observe that the measurements along the flight tracks should be be considered as exact. Indeed these measurements are averages; also the characteristic function χ_{tr} introduces a a-priori length scale. Following [13], the uncertainty on these measurements equal approximatively $\pm\delta h_{tr}$ with $\delta h_{tr} = 140\,m$. For a characteristic value $\bar{h} = 2.7$ (this is the mean value of the forthcoming test area), this uncertainty represents an uncertainty of $\pm \approx 5\%$ of the "true effective" depth. This uncertainty will be considered at Step 3) of the algorithm (see next section).

3.2 Extension of γ out of the flights tracks (Step 2)

Given the computed values $\gamma_{tr}^*(\mathbf{x})$ along the flights tracks, an extension (interpolation-extrapolation) of γ is performed by employing a classical universal Kriging. The computational method is described below.

The universal Kriging extension based on the locally observed trend. The field γ is decomposed as follows:

$$\gamma(\mathbf{x}) = \bar{\gamma}(\mathbf{x}) + \gamma_r(\mathbf{x}) \quad \text{for } \mathbf{x} \in \Omega \tag{19}$$

That is $\gamma(\mathbf{x})$ is decomposed as a deterministic trend function $\bar{\gamma}(\mathbf{x})$ plus a real-valued residual random function $\gamma_r(\mathbf{x})$. $\gamma_r(\mathbf{x})$ is supposed to be intrinsically stationary with zero mean and variogram function $v_Y(|x - x'|)$ (the residual variogram function of $\gamma(\mathbf{x})$). $\forall \mathbf{x}, \mathbf{x}' \in \Omega$,

$$\mathbb{E}[\gamma(\mathbf{x})] = \bar{\gamma}(|\mathbf{u}_H(\mathbf{x})|), \quad v_Y(\mathbf{x} - \mathbf{x}') = \frac{1}{2}Var[\gamma_r(\mathbf{x}) - \gamma_r(\mathbf{x}')] = \frac{1}{2}\mathbb{E}[(\gamma_r(\mathbf{x}) - \gamma_r(\mathbf{x}'))^2].$$

At a point $\mathbf{x}_0$, $\mathbf{x}_0 \notin \Gamma_{tr}$, the "predictor" is given by:

$$\hat{\gamma}(\mathbf{x}_0) = \langle \mathbf{c}, \mathbf{z} \rangle \tag{20}$$

with $\langle \cdot, \cdot \rangle$ the inner product, $\mathbf{z}$ the vector of sampled points, $\mathbf{z} = (\gamma(\mathbf{x}_1), \cdots, \gamma(\mathbf{x}_n)) \in \mathbb{R}^n$, $\mathbf{x}_i \in \Gamma_{tr}$, $\mathbf{c} = (c_1, \cdots, c_n) \in \mathbb{R}^n$, $c_i \in \mathbb{R}$, $i = 1, \cdots, n$ the weight corresponding to each evaluation of the random function $\gamma(\mathbf{x})$ at the sample point $\mathbf{x}_i$. The weight vector $\mathbf{c}$ is computed such that it minimises the error variance: $\text{argmin}_c \left(\sigma_{krig}^2(c) \right)$ with

$$\sigma_{krig}^2(c) = Var[\hat{\gamma}(\mathbf{x}_0) - \gamma(\mathbf{x}_0)] = -\langle \mathbf{c}, V_Y \mathbf{c} \rangle + 2\langle \mathbf{c}, \mathbf{v}_{Y,0} \rangle, \tag{21}$$

while the unbiasedness condition $\mathbb{E}[\hat{\gamma}(\mathbf{x}_0) - \gamma(\mathbf{x}_0)] = 0$ is satisfied.
Here: $V_Y \in \mathbb{R}^{n \times n}$, $(V_Y)_{i,j} = v_Y(\mathbf{x}_i - \mathbf{x}_j)$, $i, j = 1, \cdots, n$, and $\mathbf{v}_{Y,0} = (v_Y(\mathbf{x}_1 - \mathbf{x}_0), \cdots, v_Y(\mathbf{x}_n - \mathbf{x}_0)) \in \mathbb{R}^n$.
The function $\bar{\gamma}(\mathbf{x})$ is the deterministic function supposed to model the (true) mean value $\mathbb{E}[\gamma(\mathbf{x})]$. In the present real-world modelling problem, the trend $\bar{\gamma}(\mathbf{x})$ of γ is unknown. However given a domain (e.g. a portion of Antarctica), a trend may be determined from the measurements (along the tracks) and the corresponding "exact" values $\gamma_b^*(\mathbf{x})$, $\mathbf{x} \in \Gamma_{tr}$ computed at Step 1). Doing so is equivalent to consider that the measurements are sufficiently representative of the entire domain.

On the uncertainty of the statistical estimation of γ. Recall that even along the tracks, the measured depth values cannot be assumed to be exact. Indeed, for $\mathbf{x} \in \Gamma_{tr}$, $h_{tr}^* = h_b \pm \delta h_{tr}$ with $\delta h_{tr} = 140\,m$, see [13] and the discussion in the previous section.
Let us set: $\gamma^* = (\bar{\gamma} + \delta\gamma)$ with γ^* the exact value and $\bar{\gamma}$ the considered estimation. The computation made for $\mathbf{x} \in \Gamma_{tr}$ is the following, see (18): $\eta^* = \gamma^* h_b$. Hence: $\gamma^* h_b = (\bar{\gamma} \pm \delta\gamma)(h_b \pm \delta h_{tr})$. Therefore given η^* and h_b along the tracks, given the uncertainty on h (which equals $\approx 5\%$ in the forthcoming numerical tests) generates an uncertainty on $\bar{\gamma}$ of the same order of magnitude. In other respect this uncertainty due to δh_{tr} may be compared to the observed residual γ_r in the statistical estimation of γ, see e.g. Fig. 7 (Right) for the present numerical test.

4 Inference of the pair $(h, \dot{a})$ with γ given

Given the surface measurements, given the parameter γ estimated at Step 2), the inference by VDA of h and the RHS $\dot{a}$ in the RU-SIA equation (14) is performed.

The optimal control problem. Here the unknown parameter (control variable of the RU-SIA equation) is: $k = (h, \dot{a})$. The optimal estimation of k is obtained by solving the following optimal control problem:

$$\min_{k \in \mathcal{U}_{ad}} j(k) \quad \text{with}: \quad j(k) = j_{obs}(k) + \alpha_k j_{reg}(k), \tag{22}$$

$j_{obs}(k) = \frac{1}{2}\|H_k - H^{obs}\|_2^2$, $j_{reg}(k) = j_{reg}(h) + j_{reg}(\dot{a})$,

$$j_{reg}(h) = \frac{1}{2}\|(h - h_b)\|_{C_h^{-1}} \text{ and } j_{reg}(\dot{a}) = \frac{1}{2}\|\dot{a} - \dot{a}_b\|_{C_a^{-1}} \tag{23}$$

The scalar coefficient α_k is given. The norms C_h^{-1} and C_a^{-1} are defined as the inverse of covariance operators (hence symmetric, positive). The state H_k satisfies the RU-SIA equation (14) plus Dirichlet boundary conditions. The background value (first guess) is $k_b = (h_b, \dot{a}_b)$ with h_b (resp. $\dot{a}_b$) the first guess of h (resp. $\dot{a}$) provided by international databases: Racmo2 [41] for $\dot{a}$ and e.g. Bedmap2 [13] for h in Antarctica.

The admissible control set is: $\mathcal{U}_{ad} = \{(h, \dot{a})(\mathbf{x}), \ h \in [h_{min}, h_{max}](\mathbf{x}), \ \dot{a} \in [\dot{a}_{min}, \dot{a}_{max}](\mathbf{x})\}$. The bounds h_{min} and h_{max} depend whether the point $\mathbf{x}$ belongs to Γ_{tr} or not. The bounds $\dot{a}_{min}$ and $\dot{a}_{max}$ are provided by the Racmo2 database [41] (see next section for more details).

The gradient of the cost functional reads: $\forall \delta k = (\delta h, \delta \dot{a})$, $j'(k) \cdot \delta k = (\partial_h j(k) \cdot \delta h, \partial_{\dot{a}} j(k) \cdot \delta \dot{a})$, with

$$\begin{cases} \partial_h j(k) &= -\int_\Omega \frac{|\mathbf{u}_H|}{\mathcal{S}}(\mathbf{x})\gamma(\mathbf{x})\nabla H(\mathbf{x})\nabla P(\mathbf{x})d\mathbf{x} + \alpha_k \partial_h j_{reg}(k) \\ \partial_{\dot{a}} j(k) &= -\int_\Omega d\mathbf{x} + \alpha_k \partial_{\dot{a}} j_{reg}(k) \end{cases} \tag{24}$$

P is the adjoint state, solution of the adjoint equation: $-div\left(\frac{|\mathbf{u}_H|}{\mathcal{S}}\gamma h \ \nabla P\right) = (H - H^{obs})$ in Ω; plus homogeneous Dirichlet boundary conditions $P(\mathbf{x}) = 0$ on $\partial\Omega$.

Change of control variable. Following [28, 56, 9, 3] (where the control variable is the initial state of an atmospheric model), the following change of variable is made. The covariance operator C_h (resp. C_a) are supposed to be bounded symmetric positive, hence it exists $C_h^{1/2}$ (resp. $C_a^{1/2}$) such that: $C_h = C_h^{1/2} C_h^{1/2}$ (resp. $C_a = C_a^{1/2} C_a^{1/2}$). Then the following new control variable is considered:

$$w = (w_1, w_2) \text{ with } w_1 = C_h^{-1/2}(h - h_b), \ w_2 = C_a^{-1/2}(\dot{a} - \dot{a}_b). \tag{25}$$

In variable w, the optimisation problem (22) reads :

$$\min_{w \in \tilde{\mathcal{U}}_{ad}} j(w) \quad \text{with } j(w) = \frac{1}{2}\|H_w - H^{obs}\|_2^2 + \frac{\alpha_w}{2}\left(\|w_1\|_2^2 + \|w_2\|_2^2\right) \tag{26}$$

Given the new variable $w = (w_1, w_2)$, it is straightforward to calculate the original variable k: $k = (h, \dot{a}) = \left(C_h^{1/2}w_1, C_a^{1/2}w_2\right) + (h_b, \dot{a}_b)$.

On the covariance operators and regularization terms. The optimal solution w^* depends on the a-priori covariance operators C_h and C_a. These operators may be viewed as prior information on the modeling problem. However the exact covariances operators are of course unknown. To simply impose correct physical length scales of variations (recall that the physical model is a shallow flow model), the following classical covariance operators are considered. For $e \in L^2(\Omega)$,

$$C_\square \ e = \int_\Omega \sigma_\square(\mathbf{x})\sigma_\square(\mathbf{x}')c(\mathbf{x}, \mathbf{x}'; L_\square) \ e(\mathbf{x}) \ d\mathbf{x} \tag{27}$$

where $\square$ denotes either h or $\dot{a}$; $c(\cdot, \cdot; L)$ is the correlation kernel function (also called Green's function). The latter is set as the classical second order auto-regressive correlation kernel:

$$c(\mathbf{x}, \mathbf{x}'; L) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|_1}{L}\right) \tag{28}$$

The standard deviation of $(h - h_b)$ (resp. $(\dot{a} - \dot{a}_b)$) is σ_h, $\sigma_h > 0$ (resp. $\sigma_a > 0$). The scalar values $L_h > 0$ (resp. $L_a > 0$) defines a length scale.

Our numerous numerical experiments have demonstrated that the change of variable (25) based on the present covariance operators improve the robustness and the convergence speed of the VDA process. For some correlation kernels - Green's functions $c(\mathbf{x}, \mathbf{x}'; L)$, it is possible to make a link with a regularisation term in the functional to be minimised i.e. the terms $j_{reg}(h)$ and $j_{reg}(\dot{a})$ in (23). Following the calculation presented e.g. in [51] Section 7 (in their case, calculations are valid in 1D and 3D only), it can be proved that the introduction of the present covariance operators (27)(28) implies that:

$$j_{reg}(\square) \sim \sum_{i=1,2} \frac{1}{8L_\square} \int_\Omega \|e_\square(\mathbf{x})\|^2 + L_\square^2 \|\nabla e_\square(\mathbf{x})\|^2 + L_\square^4 \partial_{11}^2 e_\square(\mathbf{x}) \partial_{22}^2 e_\square(\mathbf{x}) \, d\mathbf{x} \tag{29}$$

where $\square$ denotes either h or $\dot{a}$, $e_\square = (\square - \square_b)/\sigma_\square$. The calculations are detailed in the appendix.

It follows from (29) that larger the length scale $L_\square$ is, larger the regularisation effect is.

In other respect, the change of variable may be read as a preconditioning of the optimisation problem. Indeed, a simple calculation shows that:

$$\nabla_w j(w) = \left(C_h^{1/2} \partial_h j(k), C_a^{1/2} \partial_{\dot{a}} j(k)\right)$$

where $\partial_h j(k)$ and $\partial_{\dot{a}} j(k)$ are given in (24). Therefore the optimal necessary condition $\nabla_k j(k) = 0$ may be viewed as preconditioned by $(C_h^{1/2}, C_a^{1/2})$.

After discretisation (e.g. by the standard order 2 Lagrange finite element method employed here), the covariance operators C_h and C_a are symmetric positive (covariances) matrices. The (i, j)-th element, $i, j = 1, \cdots, N$, reads, see (28): $\sigma_\square(\mathbf{x}_i)\sigma_\square(\mathbf{x}_j) \exp\left(-\frac{|\mathbf{x}_i - \mathbf{x}_j|}{L_\square}\right) \triangle|\mathbf{x}_i| \triangle|\mathbf{x}_j|$.

Therefore the positivity of these matrices $C_\square$ depend on the a-priori imposed length scales $L_\square$. According to the Gerschgörin-Hadamard theorem, a sufficient condition to guarantee the positivity is to choose $L_\square$ such that :

$$\log\left(\frac{\sum_{j \neq i} \sigma_\square(\mathbf{x}_j)}{\sigma_\square(\mathbf{x}_i)}\right) L_\square < \min_{i \in \{1, \cdots, N\}} \triangle|\mathbf{x}_i|, \quad i, j = 1, \cdots, N.$$

This condition shows that the value of $L_\square$ should be chosen small enough to keep the covariance matrix positive. However, it has been shown above that larger $L_\square$ is, higher the regularisation is. In conclusion, the length scales $L_\square$ have to be set as a balance between the regularisation effects and the preconditioning effects. A similar phenomena is analysed in detail in a different context in [18].

Remark 1. *If considering the exact covariance operators and under the assumption that $((h - h_b), (\dot{a} - \dot{a}_b))$ are mean 0 random fields, the new control variable (w_1, w_2) components are uncorrelated. In this case the change of variable (25) is a whitening transformation (the covariance matrix after change of variable equals the identity). In the present context, the exact-real covariance matrices are unknown. However given the a-priori covariance matrices above, the new variables are expected to be more physically correlated. These covariance matrices (C_h, C_a) represent prior information. The reader may refer to e.g. [3, 24] for similar investigations but in a different physical context than the present one.*

Iterative regularisation strategy. The weight parameter α_w of the regularization term in (26) is set as a decreasing sequence with $\alpha_w^{(n)} > 0$, $n = 1, \cdots, n^*$. The reader may refer e.g. to [10, 23] and references therein for descriptions of various regularisation strategies.

Let us denote by F the operator that maps the control $w \in \mathcal{U}_{ad} \subset \mathcal{X}$ onto the surface elevation H, $H \in \mathcal{Y}$ ($\mathcal{X}$ and $\mathcal{Y}$ are Hilbert spaces). According to [23] Chapter 4, if using an iteratively regularized Gauss-Newton method, the stop iteration number $n^* = n^*(\delta)$ can be chosen through the Morozov' discrepancy principle [39] such that:

$$\|H^\delta - F(w_{n^*}^\delta)\| \leq \tau\delta \leq \|H^\delta - F(w_n^\delta)\|, \quad 0 \leq k \leq n^* \tag{30}$$

with $\tau > 1$ large enough and δ the a-priori error amplitude. δ is such that: $\|F(w^*) - H^\delta\| \leq \delta$, H^δ satisfying $H^\delta = F(w^*) + \delta$.

Following [23], the weight parameter sequence $\alpha_w^{(n)}$ is defined as:

$$\alpha_w^{(n)} = \alpha_w^{(0)} q^{[n/n_0]}, \quad n = 1, \cdots, n^*. \tag{31}$$

With: $n_0 > 1$ the number of iterations for each $\alpha_w^{(n)}$, $[m]$ the maximum integer smaller than m, $\alpha_w^{(0)}$ and q given constants, $\alpha_w^{(0)} > 0$, $0 < q < 1$. Values of $\alpha_w^{(0)}$, q, n_0 are experimentally set; typically: $q = 0.5$, $n_0 = 5$ and $\alpha_0 = 1$. The stop iteration n^* is set according to (30).

5 Inversions in an East Antarctica area

In this section, the bed topography elevation in a poorly monitored East Antarctica area is estimated following the algorithm previously described.

5.1 Data description & first guesses

Some information on the domain are indicated in Fig. 3 and Table 1. This test area is compatible with the flow model since $|\mathbf{u}_H|$ ranges approximatively within $[10, 30]\, m/y$.

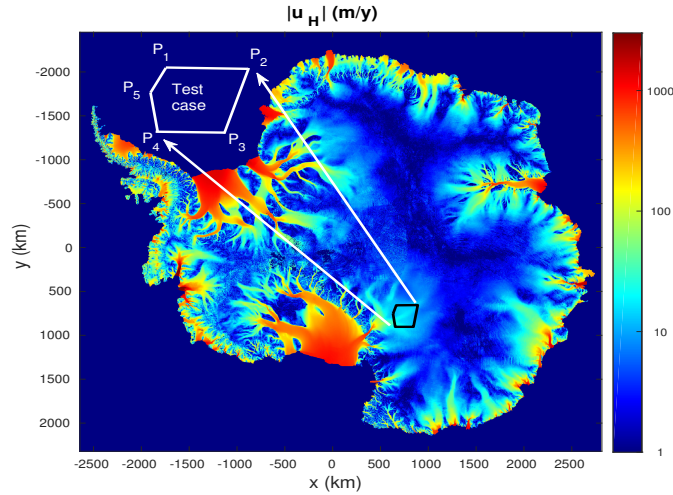Figure 3: InSAR-Based Antarctica surface velocity from [40] and the test case location.

Antarctica test case					
Points	P_1	P_2	P_3	P_4	P_5
(x, y) (km)	$(631250, 755000)$	$(665000, 658250)$	$(890000, 656000)$	$(845000, 903500)$	$(653750, 903500)$
Surface $|\Omega| = 54840$ km^2		Mean ice depth - thickness (from BedMap2 [13]): 2718 m			

Table 1: Test case location and information.

The area location points coordinates are provided in Polar stereographic coordinates with true scale at 71° S The Cartesian coordinates $\mathbf{x} = (x, y)$ (in km) are transformed from these polar stereographic coordinates, Tab. 1. The ice depth (thickness) mean value computed from BedMap2 values [13] is: $\bar{h}_b \approx 2.7\,km$. The shallow flow model (RU-SIA equation) is valid as soon as the geometrical ratio satisfies $\varepsilon = \frac{[H]}{[L]} \lesssim 0.1$. (The upper bound 0.1 is the classically admitted upper bound in the lubrication modeling community). Then the surface data $|\mathbf{u}_H|$ and H need to be smoothed at the model scale i.e. at ≈ 27 km length scale. Notice that in [57], glaciers in Antarctica presenting surface velocity ranging in $\approx [5 - 100]$ m/y are accurately modelled by the SIA model as soon as the minimal wave length equals $\approx 10 - 12$ km in mean; this detailed study confirms the validity of the present upper bound $\varepsilon \lesssim 1/10$.

To smooth the surface data, the following Gaussian function is employed:

$$G(x, y) = \frac{1}{2\pi\sigma_s^2} e^{-\frac{(x^2 + y^2)}{2\sigma_s^2}} \tag{32}$$

When applying the smoothing based on this Gaussian, values of pixels located at a larger distance than $3\sigma_s$ are unchanged (since the Gaussian values vanish). Then all surface data are smoothed with $\sigma_s = 4$ km (since $4 \times 6 = 24 \approx 27$ km). The smoothed values of $|\mathbf{u}_H|$ and H are plotted in Fig. 4. The observational term $\frac{|\mathbf{u}_H|}{S}$, factor of the effective diffusivity in (14), is plotted in Fig. 5 (Left). It can be noticed that this observational term varies by a factor ≈ 6, ranging from ≈ 0.4 to $2.4 \cdot 10^4$.
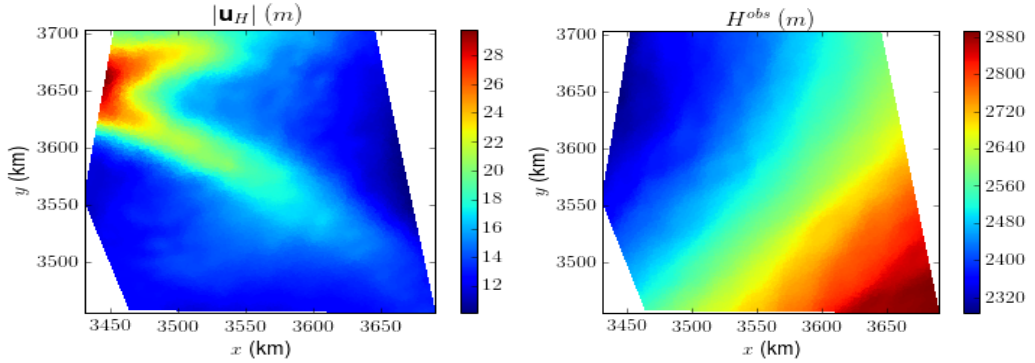

Figure 4: Surface data smoothed with $\sigma_s = 4\,km$: (Left) Surface velocity module $|\mathbf{u}_H(\mathbf{x})|$; (Right) Surface elevation $H(\mathbf{x})$.

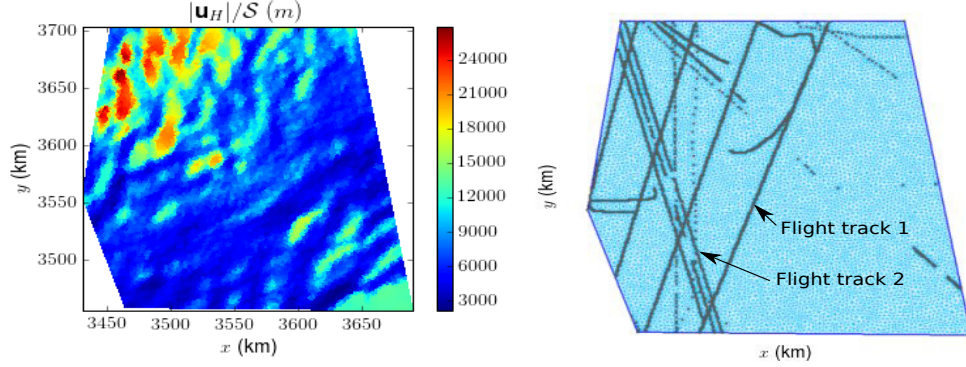Figure 5: (Left) The single observational term $\frac{|\mathbf{u}_H|}{S}$ in the RU-SIA equation smoothed with $\sigma_s = 4$ km, see (32). (Right) The finite element mesh with $\triangle \mathbf{x} \approx 3$ km and the depth measurements locations (flight tracks). Flight tracks 1 and 2 may be removed from the dataset for sensitivity analyses.

All numerical results have been performed on a medium size mesh with $\delta x \approx 3\,km$ and on a finer one with $\delta x \approx 1$ km to confirm the insensitivity of the estimations with respect to the mesh size. Indeed $\delta x \approx 3$ km provides ≈ 10 points per wave length (which equals ≈ 27 km) therefore nodes enough to properly approximate all fields. For the ≈ 3 km mesh, Fig. 5 (Right), the total number of vertices equals 8226. Meshes are generated by employing Gmsh software [14]. Vertices are imposed to be along the flight tracks, moreover with a finer mesh: $\delta x \approx 1$ km along the tracks. Flights tracks locations are provided by Bedmap2 database [13] (with the corresponding measured thickness).

First guesses of h, $\dot{a}$ and a-priori uncertainties. Natural first guess values for the VDA process, Section 4, are the ice thickness h_b provided by Bedmap2 [13] and $\dot{a}_b$ provided by Racmo2 [41]. This two first guess fields are plotted in Fig. 6.
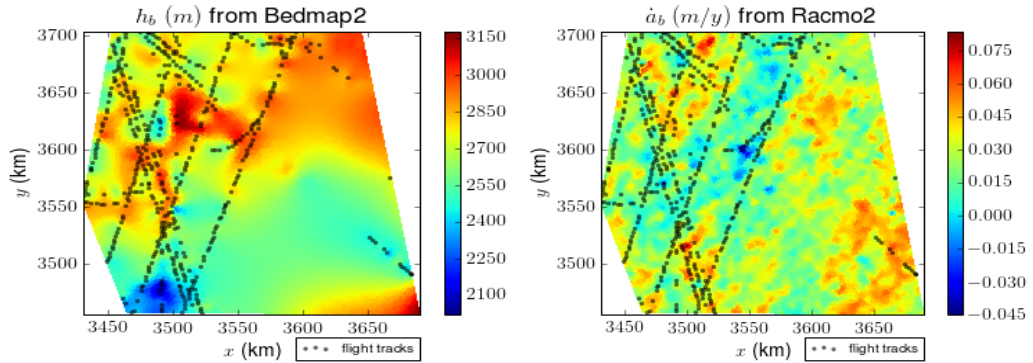

Figure 6: (Left) The depth (ice thickness) $h_b(\mathbf{x})$ from Bedmap2 [13]. (Right) The SMB $\dot{a}(\mathbf{x})$ from Racmo2 [41].

The depth estimations obtained with present inverse method are compared to the present reference values that is Bedmap2 dataset [13]. Let us recall Bedmap2 values uncertainty sources and their order of magnitudes. In [13], the depth measurements are split in two datasets (D1) and (D2). Dataset (D1) is used to build up an interpolation (see below) in the whole domain including at Dataset (D2) points (measurement values of (D2) are not employed at this stage). Next, Dataset (D2) is used to quantify the accuracy of the interpolated field. Next, basic statistics on the estimated values are derived. The a-priori uncertainties presented in [13] derive from this experimental procedure.

The interpolation is performed by employing the ArcGIS Topogrid routine (ESRI Ltd, ArcGIS 9) which is based on the ANUDEM algorithm [21]. This algorithm uses an optimised iterative finite difference interpolation technique which is essentially a thin plate spline technique, see [54].

In the present two VDA processes, a-priori bounds are imposed to the "control variable" h. More precisely:

- if $\mathbf{x} \in \Gamma_{tr}$ (i.e. along the flight tracks) then inequality constraints are imposed on h with the bounds: $h_b(\mathbf{x}) \pm \delta_{h,tracks}$, $\delta_{h,tracks} = 140\,m$ i.e. the value indicated in Bedmap2 [13].

- if $\mathbf{x} \notin \Gamma_{tr}$ (i.e. out of the tracks) then the imposed lower and upper bounds are: $\delta_{hb}(\mathbf{x}) = h_b(\mathbf{x})(1 \pm 0.6)$.

Concerning the climatic-dynamic term $\dot{a} = (a - \partial_t h)$ and according to [41], the uncertainty on $\dot{a}$ equals $\approx \pm 20\,\%$. Therefore, we impose the following lower and upper bounds: $\dot{a}(\mathbf{x})(1 \pm \delta_a)$ with $\delta_a = 0.2$. Moreover, these a-priori uncertainties $\delta_{h,tr}$, δ_{hb}, and δ_a are introduced in the covariance operators $C_\square$, see (27), as the standard deviations σ_h and σ_a. We set: $\sigma_h(\mathbf{x}) \equiv \delta_{h,tr}$ for $\mathbf{x} \in \Gamma_{tr}$, $\sigma_h(\mathbf{x}) = \delta_{hb}(\mathbf{x})$ for $\mathbf{x} \notin \Gamma_{tr}$ and $\sigma_a = \delta_a$. Finally σ_h and σ_a are empirically tuned to obtain a reasonable balance between the regularisation terms of h and $\dot{a}$, see (22).

5.2 Estimation of the effective diffusivity η and parameter γ

The effective diffusivity $\eta = (\gamma h)$ of (14) is computed by solving (17). The iterative process is stopped following the usual criteria : the cost function does not decrease anymore, the gradient norm and the control variable variations vanish. Next, following the method described in Section 3, the optimal value γ^* along the flights tracks is straightforwardly deduced, see (18) and Fig 7 (Left).
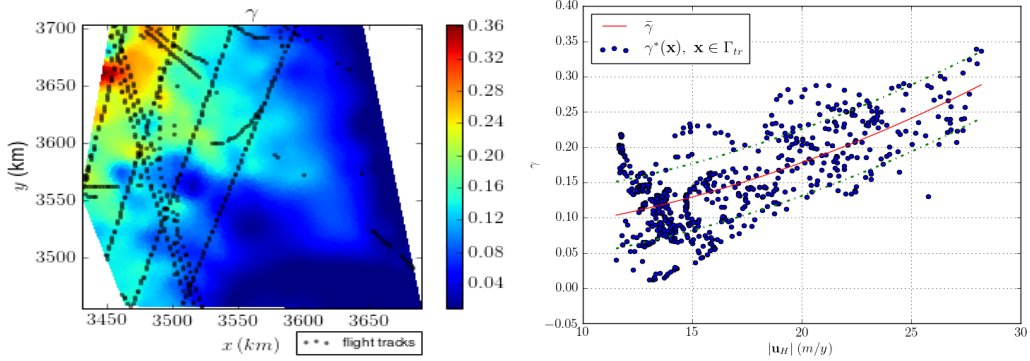

Figure 7: (Left) The parameter γ in the whole domain obtained from the relation $\gamma = \eta^*/h_b$, η^* solution of (17). (Right) The targeted values $\gamma_{tr}^*(\mathbf{x})$ vs $|\mathbf{u}_H(\mathbf{x})|$, $\mathbf{x} \in \Gamma_{tr}$ i.e. along the flights tracks (blue points), the resulting deterministic trend $\bar{\gamma}(|\mathbf{u}_H|)$ (red line) and $\bar{\gamma}(|\mathbf{u}_H|) \pm \sigma_u$ (green lines).

Following the method described in Section 3, the deterministic trend behavior $\bar{\gamma}$, see (19), is infered as a quadratic least-square optimal function: $\bar{\gamma}(\mathbf{x}) = \bar{\gamma}(|\mathbf{u}_H(\mathbf{x})|) = b_1|\mathbf{u}_H(\mathbf{x})|^2 + b_2|\mathbf{u}_H(\mathbf{x})| + b_3$. In the present case, the optimal coefficients are: $b_1 = 2.79 10^{-4}$, $b_2 = 2.78 10^{-10}$, $b_3 = 6.66 10^{-2}$. The corresponding curve is plotted in Fig. 7 (Right). Assuming that $\delta\gamma = (\gamma - \bar{\gamma}) \sim \mathcal{N}(0, \sigma_u)$, it follows that $\sigma_u = 0.05$, see Fig. 7 (Right).

Following the method described in Section 3.2, the optimal values $\gamma_{tr}^*(\mathbf{x})$ (along the tracks) provide the basic data-driven model for γ elsewhere. The extension (interpolation-extrapolation) of γ in the

whole domain is obtained by an universal Kriging algorithm (values of γ along the tracks are assumed to be representative of the whole domain). The predicted values $\hat{\gamma}(\mathbf{x})$ defined by (20) are plotted in Fig. 8 (Left). The variance σ_{krig} defined by (21) is plotted in Fig. 8 (Right). Recall that σ_{krig} measures the variance between the Kriging predictor $\hat{\gamma}(\mathbf{x})$ and the Kriging model defined by (19). The confidence interval can be defined as $\hat{\gamma} \pm 3\,\sigma_{krig}$.
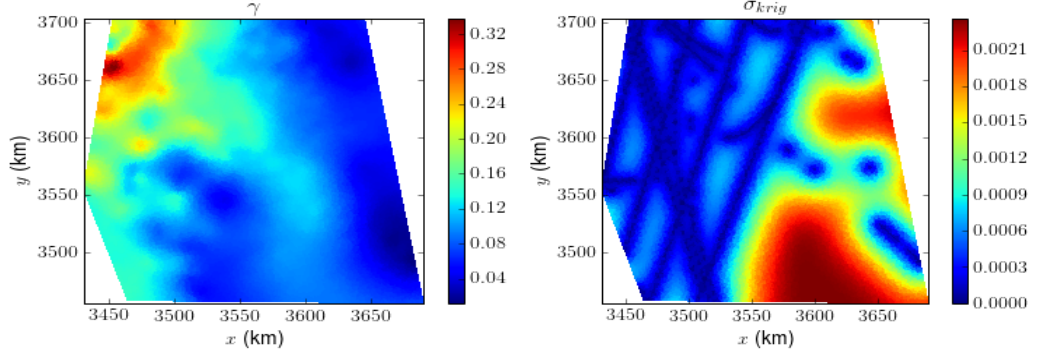

Figure 8: (Left) $\hat{\gamma}$ infered from universal Kriging. (Right) The corresponding variance σ_{krig}, see (21). (Obviously σ_{krig} vanishes along the tracks).

5.3 Inversion of $(h, \dot{a})$ with γ given: estimation of the bed topography elevation

Given the parameter γ estimated by Kriging as previously described and considering Bedmap2 values for h, the RU-SIA equation is solved to compare its solution to the surface elevation measurements. This preliminary step demonstrates the validity of the physical-based numerical model (based on the RU-SIA equation). Next, the VDA process in variables $(h, \dot{a})$ is performed; it provides the new physical-based estimations of the bed topography elevation.

RU-SIA model assessment.   Given $\gamma = \hat{\gamma}$ and $h = h_b$ (Bedmap2 values), the surface elevation $H(\mathbf{x}; h_b, \hat{\gamma})$ solution of the RU-SIA equation (14) is computed. This model output is compared to the altimetry values $H^{obs}(\mathbf{x})$; differences are plotted in Fig. 9 (Left). Basic statistics on the differences are indicated in Table 2 ("Direct model validation"). This simple direct run (without inversion process excepted for γ) fits very well the surface elevation measurements. Such a simple direct run based on the RU-SIA equation, Bedmap2 bed topography and the current estimation of γ is new. Moreover it demonstrates the reliability and the accuracy of the present physical-based numerical model, in particular the relevance of the RU-SIA equation derived in this study.
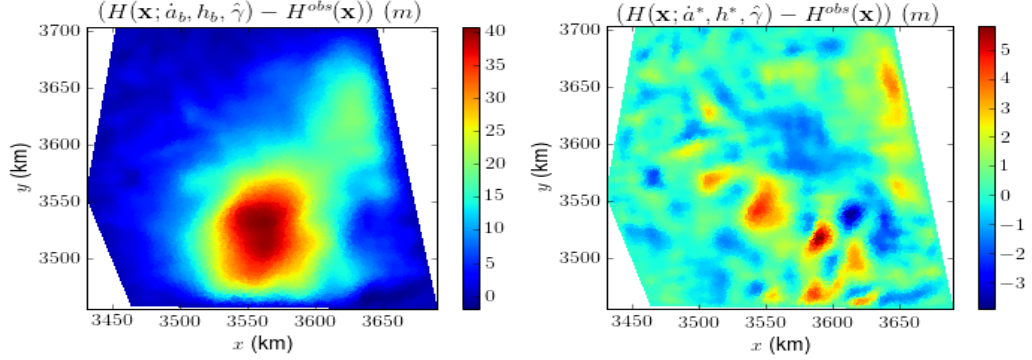
17

Figure 9: (Left) Direct model assessment: difference between $H(\mathbf{x}; h_b, \hat{\gamma})$ and $H^{obs}(\mathbf{x})$. (Right) Misfit after VDA in variables $(h, \dot{a})$: difference between $H(\mathbf{x}; h^*, \hat{\gamma})$ and $H^{obs}(\mathbf{x})$.

VDA in variables $(h, \dot{a})$. Given $\gamma = \hat{\gamma}$ (the Kriging predictor), the VDA problem (22) is solved. The cost function terms evolutions (j_{reg} and j_{obs}) are presented in Fig. 12 (red lines). The convergence is relatively fast thanks to the change of variables (w vs k). The step-like behavior is due to the iterative regularization strategy, see (31). In the present example, $n_0 = 3$ (internal iteration number for each α_n, $n = 1, \cdots, n^*$).

Antarctica test case	Median	Mean	Max
Direct model output (both with $\gamma = \hat{\gamma}$)			
$\lvert H(\mathbf{x}; h_b, \hat{\gamma}) - H^{obs}(\mathbf{x}) \rvert$	$8.3\,m$	$11.0\,m$	$46.7\,m$
$\lvert H(\mathbf{x}; h^*, \hat{\gamma}) - H^{obs}(\mathbf{x}) \rvert$	$2.6\,m$	$3.4\,m$	$21.4\,m$
Inversion results			
$\lvert h^*(\mathbf{x}, \hat{\gamma}) - h_b(\mathbf{x}) \rvert, \mathbf{x} \notin \Gamma_{tr}$	$218.6\,m$	$313.9\,m$	$1777.3\,m$
$\lvert h^*(\mathbf{x}, \hat{\gamma}) - h_b(\mathbf{x}) \rvert / \lvert h_b(\mathbf{x}) \rvert$	8.1%	11.6%	63.8%
$\lvert h^*(\mathbf{x}, \hat{\gamma}) - h_b(\mathbf{x}) \rvert, \mathbf{x} \in \Gamma_{tr}$	$0\,m$	$2.9\,m$	$80.2\,m$
$\lvert h^*(\mathbf{x}, \hat{\gamma}) - h_b(\mathbf{x}) \rvert / \lvert h_b(\mathbf{x}) \rvert$	0%	1.1%	3.5%
$\lvert \dot{a}^*(\mathbf{x}, \hat{\gamma}) - \dot{a}_b(\mathbf{x}) \rvert$	$0.4\,cm/y$	$0.5\,cm/y$	$1.8\,cm/y$
$\lvert \dot{a}^*(\mathbf{x}, \hat{\gamma}) - \dot{a}_b(\mathbf{x}) \rvert / \lvert \dot{a}_b(\mathbf{x}) \rvert$	15.8%	14.2%	20%
Ice volume change : $\lvert \int_\Omega (h^* - h_b) d\mathbf{x} \rvert / \int_\Omega h_b d\mathbf{x}$	1.7%

Table 2: Method performances

In Table 2, basic statistics on the numerical results are presented. As already mentioned, RU-SIA equation (14) set up from h_b fits already very well the surface measurements H^{obs} (see $\lvert H(\mathbf{x}; h_b, \hat{\gamma}) - H^{obs}(\mathbf{x}) \rvert$ in Tab. 2). This confirms the relevance of the present numerical model and the equation. Next if performing RU-SIA equation from the optimal depth estimation $(h^*, \dot{a}^*)$ obtained by VDA, the misfit with the altimetry measurements $\lvert H(\mathbf{x}; h^*, \hat{\gamma} - H^{obs}(\mathbf{x}) \rvert$ decreases to $2.6m$ (median) only. Concerning the RHS $\dot{a}$ and following the measurements uncertainty indicated in [41], the difference $\lvert \dot{a}^*(\mathbf{x}; \hat{\gamma}) - \dot{a}_b(\mathbf{x}) \rvert$ is imposed to be lower than 20%. The obtained differences are presented in Fig. 11 (Right). Notice that RACMO values are very slightly corrected only by the VDA process: less than 1 cm/y in the great majority of locations, see Fig. 11 and Tab. 2. The imposed maximal variation $\pm 20\%$ is reached at few locations only. Again these results confirm the reliability of the present modeling approach.

The difference between the present depth estimation h^* and Bedmap2 value h_b equals 11.6% in mean (8.1% median) out of the flight tracks ($\mathbf{x} \notin \Gamma_{tr}$), see Tab. 2. Of course, the difference remains very small along the tracks ($\mathbf{x} \in \Gamma_{tr}$) since satisfying smaller inequality constraints along the tracks (see

details in the previous section). The obtained variations along the tracks may correspond to the measurements uncertainty also it may be due to the flow model scale. The obtained differences of depth are presented in Fig. 10 (Right). Clearly the present estimation h^* is a good candidate as a new estimation of the bed elevation beneath these glaciers.

The obtained ice volume change compared to Bedmap2 is very small, see Tab. 2. Notice that in other areas (numerical experiments not shown in the present article), larger variations of the total volume have been obtained e.g. 6%.
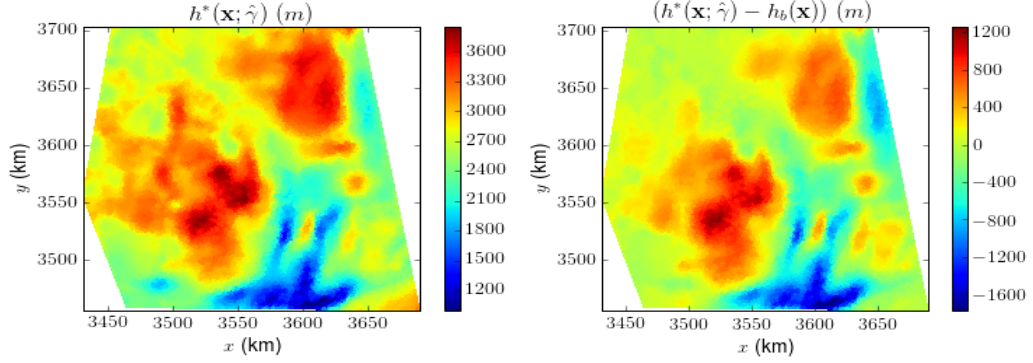
Figure 10: (Left) The depth estimation $h^*(\mathbf{x})$ (obtained from $\hat{\gamma}(\mathbf{x})$ estimation). (Right) Difference between $h^*(\mathbf{x})$ and Bedmap2 $h_b(\mathbf{x})$.
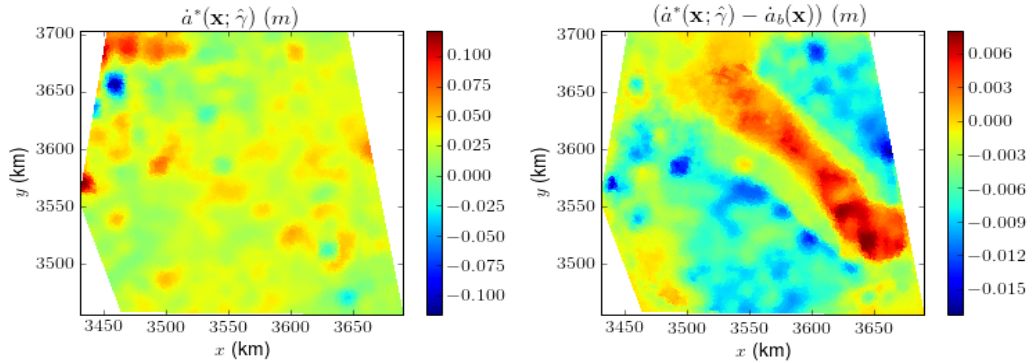

Figure 11: (Left) Infered RHS $\dot{a}^*(\mathbf{x})$ (obtained from $\hat{\gamma}(\mathbf{x})$ estimation). (Right) Difference between $\dot{a}^*$ and $\dot{a}^{obs}(\mathbf{x})$ from [41].

5.4 Sensitivity analyses on the depth estimation h^*, robustness of the inversions

The inversion process to obtain the depth estimation indicated in Fig. 10 is quite sophisticated; it combines two main steps: the estimation of the dimensionless multi-physics parameter γ and the estimation of the pair $(h, \dot{a})$ (within a-priori uncertainty bounds). Numerous different numerical experiments have been performed to assess the robustness of the complete inversion process. Moreover these empirical sensitivity analyses enable to guess basic uncertainty estimations on the results. Below various inversions are presented by considering: different estimations of γ, different density of in-situ depth measurements (by removing some flights tracks from the dataset), different smoothing length

scales of the surface data H^{obs} and $|\mathbf{u}_H|$ and different first guesses too. A summary of the results are indicated in Table 3.

With different values of γ. The estimations of the bed elevation (or equivalently of the depth h), solution of the VDA problem (22), obtained with different values of γ are compared. The considered values of γ are : $\gamma = \hat{\gamma}$, $\gamma = (\hat{\gamma} + 3\sigma_{krig})$ and $\gamma = (\hat{\gamma} - 3\sigma_{krig})$ where σ_{krig} is indicated in Fig. 8. To illustrate the good convergence behavior of these three VDA experiments, the evolution of the cost functional terms are presented in Fig. 12.
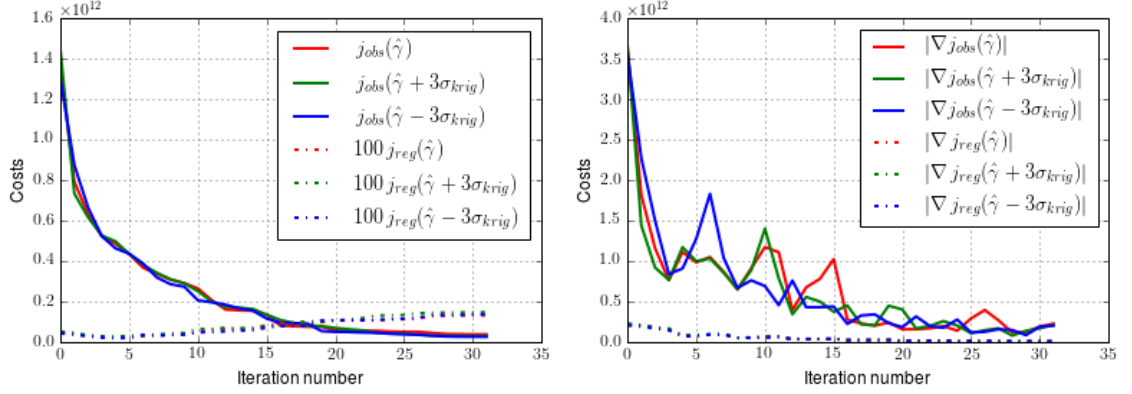

Figure 12: (Left) The cost function terms vs iterations if solving (22) with $\hat{\gamma}$, $(\hat{\gamma} + 3\sigma_{krig})$ and $(\hat{\gamma} - 3\sigma_{krig})$ respectively. (Right) The gradients of the corresponding cost function terms vs iterations.

The differences between h^* obtained with $\gamma = \hat{\gamma}$ and h^* obtained with $\gamma = (\hat{\gamma} \pm 3\sigma_{krig})$ are presented in Fig. 13. The performances are summarised in Tab. 3. The mean difference between the different infered depth are $\approx 2.8\%$. These results show the high robustness (relatively low sensitivity) of the estimation h^* with respect to the uncertain parameter γ.
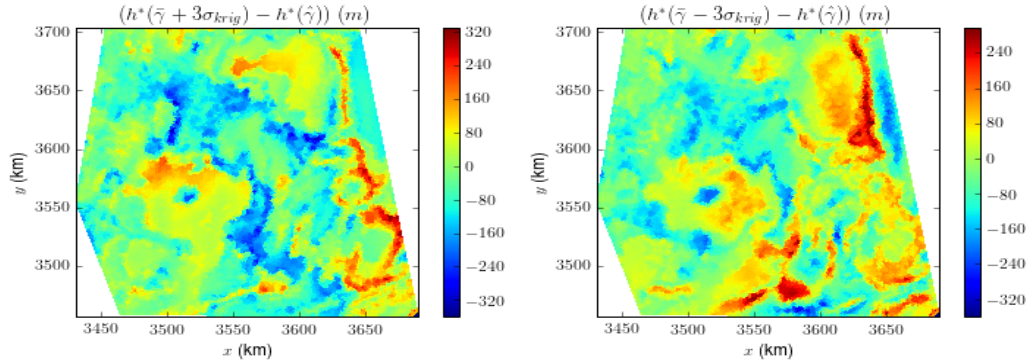

Figure 13: (Left) Difference between $h^*(\mathbf{x}, \hat{\gamma} + 3\sigma_{krig})$ and $h^*(\mathbf{x}; \hat{\gamma})$ (in m). (Right) Difference between $h^*(\mathbf{x}, \bar{\gamma} - 3\sigma_{krig})$ and $h^*(\mathbf{x}; \hat{\gamma})$ (in m). Considering that $\bar{h}_b = 2.7$km (resp. $(h_b)_{min} = 2.1$ and $(h_b)_{max} = 3.1$), a difference of 100 m corresponds to a difference of 3.7% (resp. 3.2% and 4.8%).

With less flight tracks. Other estimations of h are performed if omitting measured values along a few flight tracks; the obtained estimations are compared to the original one $h^*(\mathbf{x})$ (if considering the complete flight tracks dataset). A first "incomplete estimation" (resp. a second one) is obtained by

20

assimilating the same flight tracks set *minus* an internal one that is Flight track 1 (resp. Flight track 2) indicated in Fig. 5 (Right). The differences of the obtained estimations are plotted in Fig. 14.
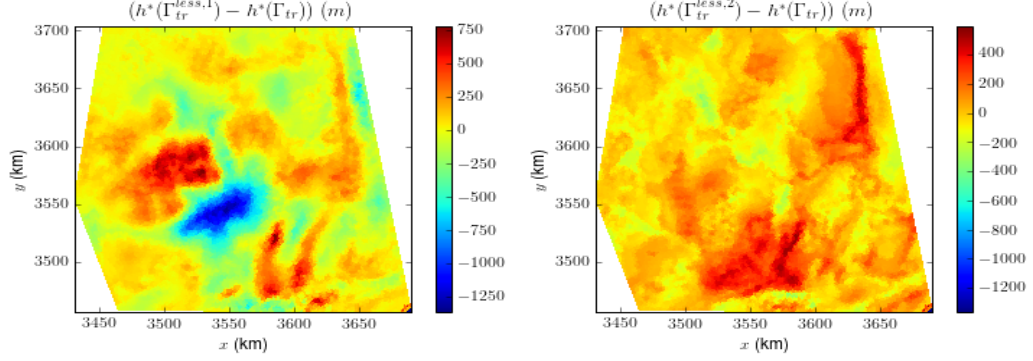

Figure 14: Difference between $h^*(\mathbf{x})$ (in m) obtained with all flight tracks dataset and the estimation with one track less : (Left) Flight track 1 (set $\Gamma_{tr}^{less,1}$); (Right) Flight track 2 (set $\Gamma_{tr}^{less,2}$). Considering that $\bar{h}_b = 2.7$km (resp. $(h_b)_{min} = 2.1$ and $(h_b)_{max} = 3.1$), a difference of 100 m corresponds to a difference of 3.7% (resp. 3.2% and 4.8%).

If removing Flight track 1 or Flight track 2, the resulting trend function $\bar{\gamma}$ remains quite close, see Fig. 7 (Right). Therefore differences between the computed depth are relatively small, lower than 6% (mean value), see Table 3. This is particularly true if removing Flight track 2: differences are only 3.9% in mean. Let us recall that the trend function $\bar{\gamma}$ fully depends on the (airborne) measurements datasets (along the flight tracks). In other respect this result highlights a great feature of the present method: the direct model is a diffusive equation therefore the inversions do not depend on the flight tracks location. This is a very important feature of the method. On the contrary, since the depth-integrated mass equation is hyperbolic (it is the linear transport equation), its inversion fully depends on the flights tracks locations and/or density (moreover with propagation of errors), see e.g. [38, 31] and the discussion in the present general introduction.

With different smoothing length scales (through the parameter σ_s). Likely the most important limitation of the present method is the large scale of the estimations due to the shallow flow assumption (long wave assumption). Thus, to be compatible with a shallow flow model, the surface data have been smoothed at the length scale $\sigma_s = 4$ km, see (32). This corresponds to the minimal length scale the physical model should be apply; the largest scale one should consider would be $\sigma_s \approx 8$ km. Below the depth estimations computed from the surface data smoothed at $\sigma_s = 4$, $\sigma_s = 6$ and $\sigma_s = 8$ km are compared, see Tab. 3 and Fig. 15. The difference $|h_{s6}^*(\mathbf{x}) - h_{s4}^*(\mathbf{x})|/|h_{s4}^*(\mathbf{x})|$ equals 7.3% (mean value), while the difference $|h_{s8}^*(\mathbf{x}) - h_{s4}^*(\mathbf{x})|/|h_{s4}^*(\mathbf{x})|$ equals 8.6% (mean). This experiment shows that the sensitivity with respect to the smoothing surface data scale is non negligible however the resulting uncertainty is smaller than the correction made on Bedmap2 values h_b. (Indeed differences with h_b equals $\approx 12\%$ (mean), see Tab. 2).
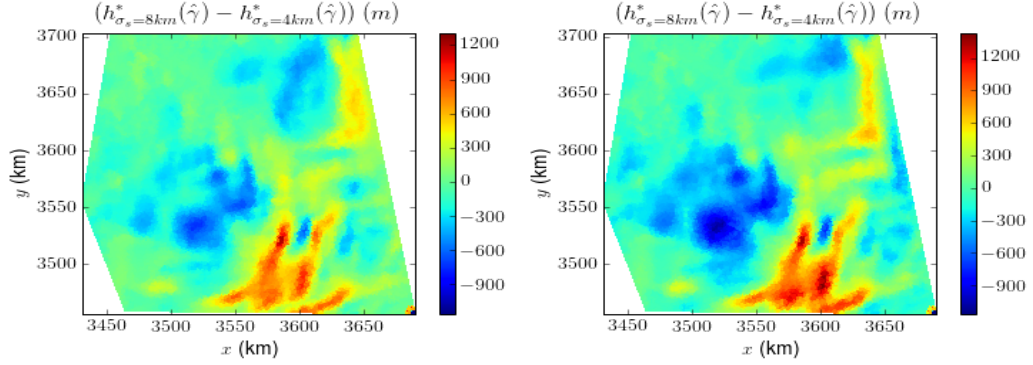
Figure 15: Robustness with respect to the smoothing length scale σ_s, see (32). Difference (in m) between $h^*_{\sigma_s=4km}(\mathbf{x})$ and : (Left) $h^*_{\sigma_s=6km}(\mathbf{x})$ (Right) $h^*_{\sigma_s=8km}(\mathbf{x})$. Considering that $\bar{h}_b = 2.7$km (resp. $(h_b)_{min} = 2.1$ and $(h_b)_{max} = 3.1$), a difference of 100m corresponds to a difference of 3.7% (resp. 3.2% and 4.8%).

With different first guesses of h.   The natural first guess of h in the VDA processes is the reference bed elevation Bedmap2 h_b, [13]. However to assess the robustness of the VDA algorithms, two other first guesses are considered: 1) h'_b built up by spline interpolation of the (airborne) measurements along the tracks; 2) h''_b built up by adding a perturbation to h_b out of the flight tracks.

After computations, differences between the different estimations h^* (obtained from the different first guesses) are lower than 3.7% (mean). The differences are plotted in Fig. 15 and statistics are presented in Tab. 3. This experiment demonstrates that the present inversions are robust with respect to the first guess. Again, the uncertainty of the present physical-based estimation is lower than the obtained correction to Bedmap2 values.
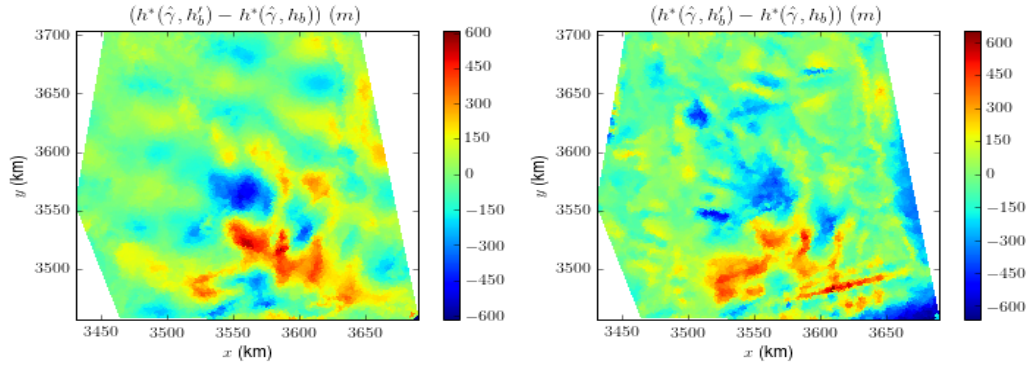

Figure 16: Robustness with respect to the first guess. Difference (in m) between $h^*(\mathbf{x}; h_b)$ and : (Left) $h^*(\mathbf{x}; h'_b)$ (Right) $h^*(\mathbf{x}; h''_b)$. Considering that $\bar{h}_b = 2.7$km (resp. $(h_b)_{min} = 2.1$ and $(h_b)_{max} = 3.1$), a difference of 100m corresponds to a difference of 3.7% (resp. 3.2% and 4.8%).

Comparison of h^* obtained with ...	Median	Mean	Max
... different γ			
$|h^*(\mathbf{x}; \hat{\gamma} + 3\sigma_{krig}) - h^*(\mathbf{x}; \hat{\gamma})|$	$59\,m$	$74\,m$	$377\,m$
$|h^*(\mathbf{x}; \hat{\gamma} + 3\sigma_{krig}) - h^*(\mathbf{x}; \hat{\gamma})|/|h^*(\mathbf{x}; \hat{\gamma})|$	2.1%	2.8%	21.5%
$|h^*(\mathbf{x}; \hat{\gamma} - 3\sigma_{krig}) - h^*(\mathbf{x}; \hat{\gamma})|$	$61\,m$	$73\,m$	$357\,m$
$|h^*(\mathbf{x}; \hat{\gamma} - 3\sigma_{krig}) - h^*(\mathbf{x}; \hat{\gamma})|/|h^*(\mathbf{x}; \hat{\gamma})|$	2.2%	2.8%	35.1%
... different flight tracks densities			
$|h^*(\mathbf{x}; \Gamma_{tr}) - h^*(\mathbf{x}; \Gamma_{tr}^{less,1})|$	$95\,m$	$161\,m$	$1389\,m$
$|h^*(\mathbf{x}; \Gamma_{tr}) - h^*(\mathbf{x}; \Gamma_{tr}^{less,1})|/|h^*(\mathbf{x}; \Gamma_{tr})|$	3.5%	5.9%	70.4%
$|h^*(\mathbf{x}; \Gamma_{tr}) - h^*(\mathbf{x}; \Gamma_{tr}^{less,2})|$	$152\,m$	$198\,m$	$1425\,m$
$|h^*(\mathbf{x}; \Gamma_{tr}) - h^*(\mathbf{x}; \Gamma_{tr}^{less,2})|/|h^*(\mathbf{x}; \Gamma_{tr})|$	2.7%	3.9%	51.4%
... different data smoothing σ_s			
$|h^*_{s4}(\mathbf{x}) - h^*_{s6}(\mathbf{x})|$	$151\,m$	$198\,m$	$1425\,m$
$|h^*_{s4}(\mathbf{x}) - h^*_{s6}(\mathbf{x})|/|h^*_{s4}(\mathbf{x})|$	5.4%	7.3%	55.9%
$|h^*_{s4}(\mathbf{x}) - h^*_{s8}(\mathbf{x})|$	$162\,m$	$233\,m$	$1560\,m$
$|h^*_{s4}(\mathbf{x}) - h^*_{s8}(\mathbf{x})|/|h^*_{s4}(\mathbf{x})|$	5.8%	8.6%	60.3%
... different first guesses			
$|h'_b(\mathbf{x}) - h_b(\mathbf{x})|$	$50\,m$	$64\,m$	$200\,m$
$|h'_b(\mathbf{x}) - h_b(\mathbf{x})|/|h_b(\mathbf{x})|$	$1.8\,\%$	$2.4\,\%$	$10.2\,\%$
$|h^*(\mathbf{x}; h'_b) - h^*(\mathbf{x}; h_b)|$	$57\,m$	$69\,m$	$377\,m$
$|h^*(\mathbf{x}; h'_b) - h^*(\mathbf{x}; h_b)|/|h^*(\mathbf{x}; h_b)|$	$2.1\,\%$	$2.6\,\%$	$30.7\,\%$
$|h''_b(\mathbf{x}) - h_b(\mathbf{x})|$	$46\,m$	$85\,m$	$734\,m$
$|h''_b(\mathbf{x}) - h_b(\mathbf{x})|/|h_b(\mathbf{x})|$	$1.7\,\%$	$3.2\,\%$	$31.8\,\%$
$|h^*(\mathbf{x}; h''_b) - h^*(\mathbf{x}; h_b)|$	$71\,m$	$96\,m$	$801\,m$
$|h^*(\mathbf{x}; h''_b) - h^*(\mathbf{x}; h_b)|/|h^*(\mathbf{x}; h_b)|$	$2.6\,\%$	$3.7\,\%$	$46.4\,\%$

Table 3: Robustness tests: statistics on the results. The subscript $4s$ (resp. $6s$ and $8s$) of H^{obs} means that the original dataset of $|\mathbf{u}_H|$ and H^{obs} are smoothed to $\sigma_s = 4km$ (resp. $\sigma_s = 6km$ and $\sigma_s = 8km$), see (32). In the last comparison, h'_b and h''_b represent different first guesses than h_b (Bedmap2); of course excepted along the flight tracks ($\mathbf{x} \in \Gamma_{tr}$).

5.5   A-posteriori estimation of the thermal boundary layer

As highlighted in Section 2.3, the uncertainty on γ is due to the slip ratio R_s uncertainty and due to the thermal - rheology parameter c_A uncertainty, see (10) and (13). c_A highly depends on $A(z)$ therefore on the thermal vertical profile, see Fig. 2. It follows from (13) that : $c_A = (q+2)(1-\gamma)/R_s$. In the present test case $\gamma^* \in (0.01, 0.34)$, see Fig. 7, $c_A \in \approx (0.8, 5)$, see Fig. 2 (Right). Considering the most employed power-law exponent value $q = 3$, these ranges imply that the slip ratio R_s ranges from ≈ 0.65 to 1. This a-posteriori estimation of R_s is consistent with the surface velocity magnitudes. This original a-posteriori analysis confirms differently the global consistency of the present flow model.

Given an a-priori vertical thermal profile e.g. the typical one providing (15) and Fig. 2 (Left), the RU-SIA equation provides an estimation of the effective thermal boundary layer $(B - b)$, see Fig. 17 (Left). Next, this thermal boundary layer can be plotted vs different fields e.g. vs $|\mathbf{u}_H|$, see Fig. 17 (Right). Such a-posteriori model analysis may be useful for ice-sheet modellers. Also the profiles may be constrained by the (very sparse) in-situ measurements of internal temperature profiles.
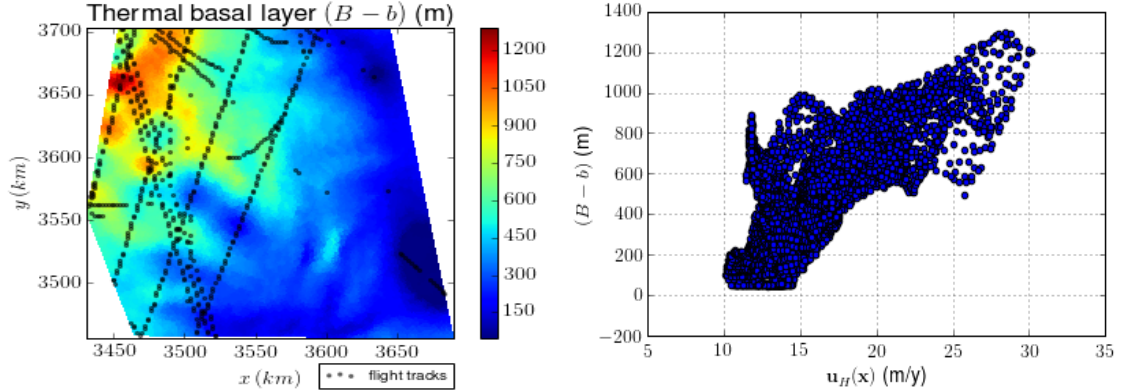
Figure 17: A-posteriori estimation of the thermal boundary layer: $(B-b)$ resulting from the inversion of RU-SUA, see (16) and Fig. 2. (Left) $(B-b)(\mathbf{x}) = m(\mathbf{x})h(\mathbf{x})$ with $R_s = 1$; (Right) $(B-b)(\mathbf{x})$ vs $|\mathbf{u}_H(\mathbf{x})|$.

6   Conclusion

This study proposes a new inverse method to infer the bed topography elevation beneath ice flows from surface observations (elevation and velocity) and sparse depth measurements. This hybrid physical-based data-driven inverse method may provide depth (thickness) estimations in areas without any in-situ measurements. It is based on the so-called Reduced Uncertainty equation (RU-SIA) which models non isothermal shallow flows; the flows may be from highly to mildly sheared therefore from slow to moderately fast. This new RU-SIA equation naturally integrates the surface measurements, also the multi-physics uncertainties within a single dimensionless parameter γ. In mildly-sheared flows, the inversion of surface measurements is very challenging since the bed topography surface signature has to be separated from the basal sliding signature and from the internal thermal profile signature (acting on the vertical velocity profile). Numerous numerical results demonstrate the reliability of this newly derived flow equation and the robustness of the inverse method therefore the depth estimations. The method provides the first physical-based depth estimations inland ice-sheets were glaciers are from slow to moderately fast. Moreover the RU-SIA equation may be interesting to provide a-posteriori estimations of the thermal basal boundary layer too (given an a-priori vertical profile). This inverse method can be applied to the great majority of the ice sheets surfaces : inland and ice-sheet upstreams also to many high mountain glaciers or even to any shallow creeping flows of generalized-Newtonian fluids (with a power-law behaviour) if sufficiently sheared. This may concern lava flows (with the thermal field given), mud flows and various polymer flows.

In the ice-sheet modelling context, the method presents many advantages such as a robustness independent of the in-situ measurements locations (here airborne ones). It may provide highly reduced uncertainty estimations in particular in unmonitored areas where the current estimations are based on highly uncertain gravimetry inversions. The method limitations are: 1) An increase of the uncertainty if the in-situ measurements are not sufficiently representative of the entire domain. Indeed the measurements provide the purely data-driven estimation of the dimensionless multi-physics parameter γ. 2) The inversions scale. Indeed the flow model is based on the long wave assumption (shallow flows) with the geometrical ratio $\varepsilon = H^*/L^* \lesssim 0.1$. (This corresponds to a length scale $L \sim 30$ km in Antarctica ice-sheet).

This original inverse method can be straightforwardly extended to unsteady flows if the provided surface observations are time-dependent (assuming that the initial condition is either not important in the considered time scale or assuming it is more or less known). This inverse method is promising and may be employed for numerous others ice-sheets areas.

24

References

[1] M. Alnæs, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M. Rognes, E. Wells, and N. Garth. The fenics project version 1.5. *Archive of Numerical Software*, 3(100):9–23, 2015.

[2] J. Bamber, J. Griggs, R. Hurkmans, J. Dowdeswell, S. Gogineni, I. Howat, J. Mouginot, J. Paden, S. Palmer, E. Rignot, et al. A new bed elevation dataset for greenland. *The Cryosphere*, 7(2):499–510, 2013.

[3] R. N. Bannister. A review of forecast error covariance statistics in atmospheric variational data assimilation. i: Characteristics and measurements of forecast error covariances. *Quarterly Journal of the Royal Meteorological Society*, 134(637):1951–1970, 2008.

[4] D.R. Baral and K. Hutter. Asymptotic theories of ice sheets and ice shelves. In *Geomorphological Fluid Mechanics*, pages 227–278. Springer, 2001.

[5] M. Boutounet, J. Monnier, and J.-P. Vila. Multi-regime shallow free surface laminar flow models for quasi-newtonian fluids. *European Journal of Mechanics-B/Fluids*, 55:182–206, 2016.

[6] F. Bouttier and P. Courtier. Data assimilation concepts and methods march 1999. *Meteorological training course lecture series. ECMWF*, 2002.

[7] D.G. Cacuci, I. M. Navon, and M. Ionescu-Bujor. *Computational methods for data evaluation and assimilation*. Chapman and Hall/CRC, 2016.

[8] J.P. Chilès and P. Delfiner. Geostatistics: modeling spatial uncertainty. *John Wiley & Sons Inc., New York*, page 695, 1999.

[9] M.J.P. Cullen. Four-dimensional variational data assimilation: A new formulation of the background-error covariance matrix based on a potential-vorticity representation. *Quarterly Journal of the Royal Meteorological Society*, 129(593):2777–2796, 2003.

[10] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems*, volume 375. Springer Science & Business Media, 1996.

[11] A. Fowler. *Mathematical geoscience*, volume 36. Springer, 2011.

[12] A. Fowler and D. Larson. On the flow of polythermal glaciers. i. model and preliminary analysis. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 363, pages 217–242. The Royal Society, 1978.

[13] P. Fretwell, H. Pritchard, D. Vaughan, J. Bamber, N. Barrand, R. Bell, C. Bianchi, et al. Bedmap2: improved ice bed, surface and thickness datasets for antarctica. *The Cryosphere*, 7(1), 2013.

[14] C. Geuzaine and J.-F. Remacle. Gmsh: A 3-d finite element mesh generator with built-in pre-and post-processing facilities. *International journal for numerical methods in engineering*, 79(11):1309–1331, 2009.

[15] R. Greve and H. Blater. *Dynamics of Ice Sheets and Glaciers*. Advances in Geophysical and Environmental Mechanics and Mathematics. Springer-Verlag, 2009.

[16] G. Gudmundsson. Transmission of basal variability to a glacier surface. *Journal of Geophysical Research: Solid Earth*, 108(B5), 2003.

[17] G. Gudmundsson. Analytical solutions for the surface response to small amplitude perturbations in boundary data in the shallow-ice-stream approximation. *The Cryosphere*, 2(2):77–93, 2008.

[18] S. A. Haben, A. S. Lawless, and N. K. Nichols. Conditioning of incremental variational data assimilation, with application to the met office system. *Tellus A: Dynamic Meteorology and Oceanography*, 63(4):782–792, 2011.

[19] C. Heining and M. Sellier. Direct reconstruction of three-dimensional glacier bedrock and surface elevation from free surface velocity. *AIMS Geosciences*, 2:45–63, 2016.

[20] I. M. Howat, A. Negrete, and B. E. Smith. The greenland ice mapping project (gimp) land classification and surface elevation data sets. *The Cryosphere*, 8(4):1509–1518, 2014.

[21] M. F. Hutchinson. A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. *Journal of Hydrology*, 106(3-4):211–232, 1989.

[22] K. Hutter. *Theoritical Glaciology*. Norwell Mass., 1983.

[23] B. Kaltenbacher, A. Neubauer, and O. Scherzer. *Iterative regularization methods for nonlinear ill-posed problems*, volume 6. Walter de Gruyter, 2008.

[24] D. Katz, A. S. Lawless, N. K. Nichols, M. J. P. Cullen, and R. N. Bannister. Correlations of control variables in variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 137(656):620–630, 2011.

[25] F.-X. LeDimet and O. Talagrand. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus A*, 38(2):97–110, 1986.

[26] J.-L. Lions. *Optimal control of systems governed by partial differential equations*, volume 170. Springer Verlag, 1971.

[27] A. Logg, G. Wells, and J. Hake. Dolfin: A c++/python finite element library. In *Automated Solution of Differential Equations by the Finite Element Method*, pages 173–225. Springer, 2012.

[28] A.C. Lorenc, S.P. Ballard, R.S. Bell, N.B. Ingleby, P.L.F. Andrews, D.M. Barker, J.R. Bray, A.M. Clayton, T. Dalby, D. Li, et al. The met. office global three-dimensional variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, 126(570):2991–3012, 2000.

[29] N. Martin and J. Monnier. Adjoint accuracy for the full-stokes ice flow model: limits to the transmission of basal friction variability to the surface. *The Cryosphere*, 8:721–741, 2014.

[30] N. Martin and J. Monnier. Inverse rheometry and basal properties inference for pseudoplastic geophysical flows. *European Journal of Mechanics-B/Fluids*, 50:110–126, 2015.

[31] L. Michel, M. Picasso, D. Farinotti, A. Bauder, M. Funk, and H. Blatter. Estimating the ice thickness of mountain glaciers with an inverse approach using surface topography and mass-balance. *Inverse Problems*, 29(3):035002, 2013.

[32] L. Michel, M. Picasso, D. Farinotti, M. Funk, and H. Blatter. Estimating the ice thickness of shallow glaciers from surface topography and mass-balance data with a shape optimization algorithm. *Computers & Geosciences*, 66:182–199, 2014.

[33] J. Monnier. *Variational Data Assimilation: from least-square solutions and optimal control to the assimilation of databases into physical-based models*. Open Online Course (videos with the accompanying manuscript -143 pages). INSA - University of Toulouse. https://www.math.univ-toulouse.fr/ jmonnie/Enseignement/VDA.html, 2018.

[34] J. Monnier and P.-E. des Boscs. Inference of the bottom properties in shallow ice approximation models. *Inverse Problems*, 33(11):115001, 2017.

[35] J. Monnier and J. Zhu. *DassFlow: Data Assimilation for Free Surface Flows.* http://www.math.univ-toulouse.fr/DassFlow. INSA - University of Toulouse, CNES.

[36] L. Morland. Thermo-mechanical balances of ice sheet flow. *Geophysical and Astrophysical Fluid Dynamics*, 29:237–266, 1984.

[37] M. Morlighem. *Ice sheet properties inferred by combining numerical modeling and remote sensing data.* PhD thesis, PhD Ecole Centrale de Paris, 2011.

[38] M. Morlighem, E. Rignot, J. Mouginot, H. Seroussi, and E. Larour. High-resolution ice-thickness mapping in south greenland. *Annals of Glaciology*, 55(67):64–70, 2014.

[39] V.A. Morozov and M. Stessin. *Regularization methods for ill-posed problems.* CRC press Boca Raton, FL:, 1993.

[40] J. Mouginot, E. Rignot, B. Scheuchl, and R. Millan. Comprehensive annual ice sheet velocity mapping using landsat-8, sentinel-1, and radarsat-2 data. *Remote Sensing*, 9(4):364, 2017.

[41] B. Noël, W.J. van de Berg, V. Wessem, J. Melchior, E. Van Meijgaard, D. Van As, J. Lenaerts, S. Lhermitte, P.K. Munneke, C.J.P. Smeets, et al. Modelling the climate and surface mass balance of polar ice sheets using racmo2-part 1: Greenland (1958-2016). *Cryosphere*, 12(3):811–831, 2018.

[42] L. Prandtl. Zur berechnung der grenzschichten. *ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift fÃijr Angewandte Mathematik und Mechanik*, 18(1):77–82, 1938. (Translated as 'Note on the calculation of boundary layers', *Tech. Memor. nat. adv. Comm. Aero., Wash.* **959.**).

[43] P.B. Price, O.V. Nagornov, R. Bay, D. Chirkin, Y. He, P. Miocinovic, A. Richards, K. Woschnagg, B. Koci, and V. Zagorodnov. Temperature profile for glacial ice at the south pole: Implications for life in a nearby subglacial lake. *Proceedings of the National Academy of Sciences*, 99(12):7844–7847, 2002.

[44] U. Radok, D. Jenssen, and W. Budd. Steady-state temperature profiles in ice sheets. *Bull. Int. Assoc. Scient. Hydrol*, 8(1):36, 1970.

[45] L. Rasmussen. Bed topography and mass-balance distribution of columbia glacier, alaska, usa, determined from sequential aerial photography. *Journal of Glaciology*, 34(117):208–216, 1988.

[46] Y. Sasaki. *Some basic formalisms in numerical variational analysis.* Citeseer, 1970.

[47] C. Schoof and C. Hindmarsh. Thin-film flows with wall slip: an asymptotic analysis of higher order glacier flow models. *Q. Jl Mech. Appl. Math.*, 63(1):73–114, 2010.

[48] M. Sellier. Inverse problems in free surface flows: a review. *Acta Mechanica*, 227(3):913–935, 2016.

[49] M. Sellier, A. Gessese, and C. Heining. Analytical and numerical bedrock reconstruction in glacier flows from free surface elevation data. *Research report*, 2012.

[50] H. Seroussi, M. Morlighem, E. Rignot, A. Khazendar, E. Larour, and J. Mouginot. Dependence of century-scale projections of the greenland ice sheet on its thermal regime. *Int. Glacio. Soc.*, 59:1024–1034, 2013.

[51] A. Tarantola. *Inverse problem theory and methods for model parameter estimation*, volume 89. siam, 2005.

[52] T. Thorsteinsson, C. F. Raymond, H. Gudmundsson, R. A. Bindschadler, P. Vornberger, and I. Joughin. Bed topography and lubrication inferred from surface measurements on fast-flowing ice streams. *Journal of Glaciology*, 49(167):481–490, 2003.

[53] J. van Wessem, W. Jan Van De Berg, B. Noël, E. Van Meijgaard, C. Amory, G. Birnbaum, C. Jakobs, K. Krüger, J. Lenaerts, S. Lhermitte, et al. Modelling the climate and surface mass balance of polar ice sheets using racmo2: Part 2: Antarctica (1979-2016). *Cryosphere*, 12(4):1479–1498, 2018.

[54] G. Wahba. *Spline models for observational data*, volume 59. Siam, 1990.

[55] R. Warner and W. Budd. Derivation of ice thickness and bedrock topography in data-gap regions over antarctica. *Annals of Glaciology*, 31(1):191–197, 2000.

[56] A. Weaver and P. Courtier. Correlation modelling on the sphere using a generalized diffusion equation. *Quarterly Journal of the Royal Meteorological Society*, 127(575):1815–1846, 2001.

[57] C. R. Williams, R. Hindmarsh, and R. Arthern. Calculating balance velocities with a membrane stress correction. *Journal of Glaciology*, 60(220):294–304, 2014.

Appendix

In this section, the equivalence between the covariance operators (27)(28) and the regularisation term j_{reg} defined by (29) is proved. The calculation is similar than those presented in [51] Section 7, however the latter is valid in 1D and 3D only. On the contrary the present calculation is valid in 2D and with non-constant standard derivations $\sigma_\square$; therefore it may be applied to 2D shallow flow models like RU-SIA equation.

The calculation related to the variable h only is presented; the calculation for $\dot{a}$ is the same. For a sake of simplicity, the subscripts $_h$ are skipped. We set: $\delta h = (h - h_b)$. Then:

$$\|\delta h\|^2_{\mathbf{C_h}^{-1}} = \langle \delta h, \mathbf{C_h}^{-1} \delta h \rangle = \int_\Omega \delta h(\mathbf{x}) \tilde{\delta h}(\mathbf{x}) d\mathbf{x}, \tag{33}$$

where $\tilde{\delta h}(\mathbf{x}) = \mathbf{C_h}^{-1} \delta h = \int_\Omega c^{-1}(\mathbf{x}, \mathbf{x}') \delta h(\mathbf{x}') d\mathbf{x}'$. To calculate $\|\delta h\|^2_{\mathbf{C}^{-1}}$, the following equation has to be solved:

$$\delta h(\mathbf{x}) = \int_\Omega c(\mathbf{x}, \mathbf{x}') \tilde{\delta h}(\mathbf{x}') d\mathbf{x}' = \int_\Omega \sigma(\mathbf{x}) \sigma(\mathbf{x}') \exp(-\frac{|\mathbf{x} - \mathbf{x}'|_1}{L}) \tilde{\delta h}(\mathbf{x}') d\mathbf{x}'$$

We set: $g(\mathbf{x}) = \exp(-\frac{|\mathbf{x}|_1}{L})$. The equation above can be written as a convolution product: $\frac{\delta h(\mathbf{x})}{\sigma(\mathbf{x})} = g(\mathbf{x}) * \left(\sigma(\mathbf{x}) \tilde{\delta h}\right)$. In the Fourier space, it reads:

$$\mathcal{F}\left(\frac{\delta h(\mathbf{x})}{\sigma(\mathbf{x})}\right) = G(\zeta) \mathcal{F}\left(\sigma(\mathbf{x}) \tilde{\delta h}\right), \quad G(\zeta) = \mathcal{F}(g(\mathbf{x})), \quad \zeta = (\zeta_1, \zeta_2).$$

It follows:

$$\tilde{\delta h}(\mathbf{x}) = \frac{1}{\sigma(\mathbf{x})} \left[\mathcal{F}^{-1}\left(\frac{1}{G(\zeta)}\right) * \left(\frac{\delta h(\mathbf{x})}{\sigma(\mathbf{x})}\right) \right]. \tag{34}$$

By applying the inverse Fourier transform $\mathcal{F}^{-1}$, it follows:

$$\mathcal{F}^{-1}\left(\frac{1}{G(\zeta)}\right) = \frac{1}{4L}\left(\delta(x_1) - L^2 \delta^{(2)}(x_1)\right)\left(\delta(x_2) - L^2 \delta^{(2)}(x_2)\right)$$

where δ denotes the Dirac distribution. According to (34) it follows:

$$\tilde{\delta h}(\mathbf{x}) \sim \frac{1}{4L\sigma(\mathbf{x})} \left[\frac{\delta h(\mathbf{x})}{\sigma(\mathbf{x})} - L^2 \triangle \left(\frac{\delta h(\mathbf{x})}{\sigma(\mathbf{x})}\right) + L^4 \frac{\partial^4}{\partial^2 x_1 \partial^2 x_2}\left(\frac{\delta h(\mathbf{x})}{\sigma(\mathbf{x})}\right) \right]$$

Finally according to (33) it follows:

$$\|\delta h\|^2_{\mathbf{C_h}^{-1}} \sim \frac{1}{4L} \int_\Omega \left[\left(\frac{\delta h(\mathbf{x})}{\sigma(\mathbf{x})}\right)^2 + L^2 \left|\nabla\left(\frac{\delta h(\mathbf{x})}{\sigma(\mathbf{x})}\right)\right|^2 + L^4 \frac{\partial^2}{\partial^2 x_1^2}\left(\frac{\delta h(\mathbf{x})}{\sigma(\mathbf{x})}\right) \frac{\partial^2}{\partial^2 x_2^2}\left(\frac{\delta h(\mathbf{x})}{\sigma(\mathbf{x})}\right) \right] d\mathbf{x}$$