



**HAL**  
open science

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

Jerome Monnier, Jiamin Zhu

► **To cite this version:**

Jerome Monnier, Jiamin Zhu. Inference of the bottom topography in anisothermal mildly-sheared shallow ice flows. 2018. hal-01827991v1

**HAL Id: hal-01827991**

**<https://hal.science/hal-01827991v1>**

Preprint submitted on 2 Jul 2018 (v1), last revised 21 Feb 2019 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

Jérôme Monnier \*      Jiamin Zhu \*

## Abstract

This study proposes an inverse method to infer the bed topography beneath ice flows from the surface observations (e.g. altimetry elevations and InSAR velocities) and sparse depth measurements (e.g. from airborne campaigns). The flow model is valid for highly to mildly-sheared regimes (hence mildly-rapid) and takes into account varying vertical thermal profiles; it is depth-integrated (long-wave assumption). The inverse problem is particularly challenging since the assimilated surface signatures integrate the bottom features (bed elevation and friction-slip amount) and the internal deformation due to non constant rate factor vertical profile. The first key step of this multi-physics flow inversion is a re-derivation of the anisothermal xSIA model (lubrication type model for generalized Newtonian fluids) leading to a Reduced Uncertainty (RU) version presenting a single uncertain multi-physic parameter  $\gamma$ ; that is the so-called RU-SIA equation. The next key steps are advanced Variational Data Assimilation (VDA) formulations combined with a stochastic extension of  $\gamma$  based on the trend observed in the in-situ measurements (e.g. along the flight tracks). The resulting method provides the first physical-based depth (ice thickness) inversions in mildly-sheared mildly-slippery shallow flows. Numerical results are presented in a poorly monitored inland Antarctica area. The uncertainty of the estimated bedrock elevation is noticeably reduced compared to the current estimations uncertainties. The robustness of the inversion process is demonstrated through numerous numerical experiments and empirical sensitivity analyses. The new RU-SIA model may provide a-posteriori estimations of the thermal basal boundary layer too.

## 1 Introduction

The knowledge of the bottom topography 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 CREGIS, Univ. of Kansas and NASA Operation IceBridge) providing bed elevation measurements. These latter are dense in fast stream costal areas but very sparse inland. Moreover numerous satellites provide accurate measurements of the ice sheet surfaces: altimeters provide surface elevations  $H$  at  $\approx \pm 10 \sim 30$  cm for  $1 \text{ km}^2$  pixels, see e.g. [?], while radar interferometers (InSar) provide surface velocity fields  $\mathbf{u}_H$  see e.g. [?] (these measurements are accurate as soon as  $|\mathbf{u}_H| \geq \sim 10 \text{ m/y}$ ).

Outside of highly measured areas (fast stream costal areas) the current bed topography estimations are based on the direct airborne measurements (along the relatively sparse flight tracks) and standard Kriging interpolation between. The resulting bed topography maps are presented in [?] for Greenland and in [?] 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, [?]. On the contrary, in fast stream nearshore areas, the inversion of the (regularized) depth-averaged mass equation combined with altimetry data makes possible to fill up more accurately the gaps; more precisely downstream and upstream the measurements following

---

\*INSA & Institut de Mathématiques de Toulouse (IMT), France. (Cor.: [jerome.monnier@insa-toulouse.fr](mailto:jerome.monnier@insa-toulouse.fr)).

the streamlines since the equation is the linear transport equation, see [?, ?] and the pioneer study [?].

Up to now, no physical-based inversions have been performed outside 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; that is a real challenge, see e.g. [?, ?] and [?]. 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.

An adequate physical model class to consider for such inversions are shallow flow models since rich enough (mass and momentum conservation are taken into account) but simpler than the fully 3D free surface model. Indeed the complete 3D model would be extremely very complex to invert in this context since very likely presenting severe equifinality issues. (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 rated factor, therefore on the vertical temperature profile. This additional physical phenomena makes the inversion of the flow model even more challenging. Finally the unknowns of the flow model to be inferred are : the bed topography, the basal slipperiness (equivalently the friction coefficient at bottom) and the thermal vertical profile. This is the challenging inverse problem which is addressed in the present study.

It is worth to notice that effective bed topography only can be inferred 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 [?, ?, ?, ?] 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 [?, ?, ?, ?, ?, ?, ?, ?, ?]. These studies consider restricted flow regimes only since considering either pure 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, [?]) or including the momentum equations too but isothermal and with no-slip at bottom, see e.g. [?] and the detailed review in [?]. [?] 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 [?] 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 taking into account the surface measurements (elevation and velocity) and space varying vertical temperature profiles. Recall that the SIA model derives from the classical lubrication theory applied to power-law rheology fluids and if neglecting the inertial terms (creeping flows), see [?, ?, ?, ?]. Moreover [?] demonstrates the validity of xSIA model for moderate basal slip (the present targeted flow regime) by stating formal error estimates. Such mildly-sheared mildly-slippery regimes correspond to ice flows with surface velocity  $\sim 5 - 50^+$  m/y, that is the targeted ice-sheet interior sectors and ice-sheds upstreams. Recall that these areas are poorly measured areas (airborne data). In [?] the xSIA model is enriched by taking into account varying vertical rate factor profile.

The present new RU-SIA equation is a diffusive equation like the classical SIA model and the non-isothermal version derived in [?], but containing a single parameter denoted  $\gamma$  (RU version). The latter contains all the unknown-uncertain multi-physics terms: basal friction - slip ratio, varying rate factor vertical profile and rheology power-law exponent. An explicit expression of this single parameter is obtained.

Secondly and based on this complete shallow flow model (taking into account all basic physical phenomena), the inversion algorithm relies on two Variational Data Assimilation (VDA) processes, see

e.g. [?, ?] and [?, ?, ?], aiming at fusing at best the RU-SIA model and the altimetry data (surface elevation measurements) plus sparse airborne measurements. These two VDA process uses the adjoint equations [?]. Given reliable estimated values of  $\gamma$  along the flight tracks, its extension to the entire domain is performed by an universal Kriging estimator see e.g. [?]. It is worth to notice that the surface velocity information (InSAR satellite 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 sough ice thickness - depth.

Novelties and strengths of these first inversions valid for anisothermal creeping shallow flows, from slow to mildly rapid, are numerous. In particular they provide: robust estimations independently of the airborne measurements locations (contrarily to the intrinsically unstable inversions of the mass - transport equation); reduced uncertainty depth estimations in the very poorly monitored areas (e.g. deep inland Antarctica); an estimation of the thermal basal 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 [?, ?] and optimization libraries; it is part of DassFlow computational code [?].

A complete real data set of an  $\sim 200 \text{ km} \times 250 \text{ km}$  inland Antarctica area is considered; the elaborated inverse method provides a new bed topography estimation. The data set includes: the current bed estimation (Bedmap2, [?]) (it provides the first guess value of the VDA processes), surface elevations  $H$  [?], surface velocities magnitudes  $|\mathbf{u}_H|$  [?] and the climatic forcing term  $\dot{a}$  [?]. 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 (ranging from  $\approx 10$  to  $30$ ] m/y hence in the model validity range), also it is sufficiently surveyed ( $\sim 10$  flights tracks) to assess the robustness and discuss into details the elaborated inverse method.

The outline of the article is as follows. In Section ??, 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 multi-physics parameter  $\gamma$ ). Also a typical uncertainty estimation of  $\gamma$  is derived and the global inversion method is sketched. In Section ??,  $\gamma$  values are estimated along the flight tracks by a VDA process, next  $\gamma$  is extended in the whole domain by an universal Kriging method. In Section ??, given  $\gamma$ , the ice thickness  $h$  and the RHS  $\dot{a}$  are simultaneously estimated by VDA. In Section ??, the numerical results obtained in an inland Antarctica area are presented and discussed into details.

## 2 The direct model and the global inversion algorithm

In this section first the xSIA model (extended SIA model), that is the classical SIA equation with non vanishing basal velocity see e.g. [?], is recalled. Next we derive the so-called Reduced-Uncertainty SIA (RU-SIA) model. The latter is obtained by reformulating the equations and gathering the few uncertain terms into a single parameter  $\gamma$ . 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 [?, ?] also e.g. [?] Chapter 10.2. The basic assumption states that the flow, thin geometry, is sheared; in other words, the normal stress components are negligible. The SIA equations are first order in  $\varepsilon$ . In [?], it is formally demonstrated that the xSIA model remains valid for a friction coefficient  $C \simeq \mathcal{O}(1)$ , hence clarifying the xSIA domain of validity. Moreover the xSIA equation are derived not only in the mean slope coordinate system (it is classically derived in the horizontal-vertical coordinate system) but in a much more general coordinate system, the Prandtl

coordinate system [?]. More precisely it is demonstrated in [?] that the same expressions and equation remain valid, hence making possible their applications to any bottom shape, i.e. without any clear mean slope, like those observed in some high mountain glaciers, for example.

In the xSIA model, the slip parameter  $C$  and the rate factor  $A$  (including the thermal effects) are unknown. The aim of this section is to re-write the model equations to reduce these two highly uncertain parameters to a single unknown parameter  $\gamma$  and by taking advantage of the surface observational term (elevation and velocity). This gives the so-called RU-SIA model.

**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:

$$\partial_t h + \text{div} \mathbf{q} = a, \quad (1)$$

where  $\mathbf{q} = h \bar{\mathbf{u}}$  is the discharge,  $\bar{\mathbf{u}} = \frac{1}{h} \int_b^H \mathbf{u}(z) dz$  is the depth-averaged velocity,  $a$  is the mass balance source term.

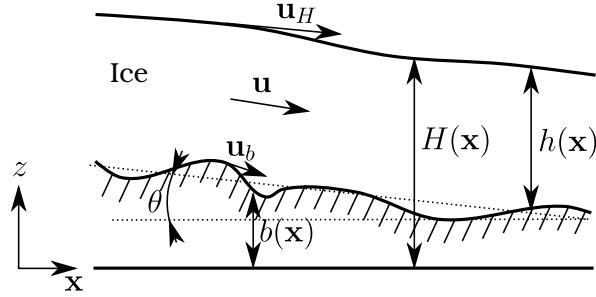


Figure 1: Schematic vertical view of the gravitational ice flow and notations

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

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

where  $A(z)$  is the rate factor in the constitutive fluid law depending on the vertical temperature profile in the ice:  $A(z) \equiv A(T(z))$ . The basal velocity  $\mathbf{u}_b \equiv \mathbf{u}(z = b)$  reads:

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

with  $C > 0$  the slip coefficient. For glaciers, the usual exponent value is  $q = 3$ . Following [?], we introduce the parameter

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

then the depth-averaged velocity reads:

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

In the isothermal case or constant vertical profile,  $\bar{A}(\mathbf{x}) = A(\mathbf{x}) \forall \mathbf{x}$ .

**The xSIA equation** By injecting the velocity expression (??) into the mass equation (??), the xSIA equation valid for non isothermal flows reads, see [?]:

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

It is a non-linear diffusive equation in  $h$ . To be solved, the values of the slip coefficient  $C$  and the depth-integrated rate factor parameter  $\underline{A}$  have to be set; however they are a-priori unknown. Note that if  $C \rightarrow 0$ , the no-slip condition (adherence) is imposed. On the contrary if  $C \rightarrow \infty$  a pure slip condition (vanishing friction) is imposed. Nevertheless, to remain within the SIA model validity regime, the slip coefficient  $C$  has to vary from 0 to  $\mathcal{O}(1)$  at most, see [?] for a detailed discussion and analysis.

## 2.2 The RU-SIA equation

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

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

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

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

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

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 (??) reads as:

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

with:

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

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

Let us define the slip ratio as follows:

$$R_s = \frac{|\mathbf{u}_H| - |\mathbf{u}_b|}{|\mathbf{u}_H|} = 1 - \frac{|\mathbf{u}_b|}{|\mathbf{u}_H|} \quad (11)$$

Observe that by using (??), (??) and the slip coefficient expression above, it follows the expression:  $R_s = \frac{2\bar{\rho}\underline{A}}{\mathcal{Q}_H(q+1)}h^{q+1}$ .

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

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

with:

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

In (??), the uncertain multi-physics parameters  $C$  and  $A(z)$  have been reduced to the single unknown parameter  $\gamma(q, c_A, R_s)$ .

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

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

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 provided from surface measurements. Assuming that the depth  $h$  (or equivalently the bedrock elevation  $b$ ) is given,  $\gamma$  is the single unknown parameter of this equation in variable  $H$ . Then the RU-SIA equation (??) is elliptic, linear 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 this elliptic equation.

The depth-integrated RU-SIA model (??) is valid if the slip ratio  $R_s$  ranges from  $\sim 0.3$  to 1. Indeed in this case, the basic scaling done to derive the SIA equation (sheared flows) remains valid, see [?, ?] for detailed discussions.

### 2.3 Typical uncertainty on the single parameter $\gamma$

In this section, a-priori estimations on the single unknown parameter  $\gamma$  of the RU-SIA model are derived.

Let us set  $q = 3$ , which is the usual exponent value in glaciers. According to the Arrhenius law and values, see e.g. [?, p.54], and typical ice-sheet vertical temperature profiles, see e.g. [?, ?, ?]), the following typical vertical profile of  $A(z)$  is considered:

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

where  $A_a$  and  $k$  are constants, see Fig. ???. Let us define

$$B = mh + b, \quad m \in [0, 1] \quad (16)$$

From (??), (??), (??), it follows:

$$R_A = \frac{5m^2(1-k)(m^3 - 6m^2 + 15m - 20) + 60}{12(m(1-k)(m^3 - 5m^2 + 10m - 10) + 5)} \quad (\text{for } q = 3) \quad (17)$$

For  $k \in [10, 1000]$ , it follows that  $c_A \in (0.5, 5.04)$  (for  $q = 3$ ). In Fig. ??, the parameter  $c_A$  vs  $m$  is plotted for different values of  $k$  and  $q$ .

Considering that  $R_s \in ]0, 1]$ , the following numerical estimation holds:

$$\gamma \in (-0.08, 1) \quad (\text{for } q = 3) \quad (18)$$

In the isothermal case,  $c_A = 1$ ,  $\gamma = (1 - 0.2R_s)$  hence  $\gamma \in (0.8, 1)$ . In the isothermal case the uncertainty on the parameter  $\gamma$  is relatively small.

Considering non-isothermal flows, the Antarctica case and the Greenland case have to be distinguished. In both case, the bed temperature can be assumed to be close to zero. In the Antarctica case (resp. Greenland case), the surface temperature can be equal to  $-40 C^\circ$  (resp.  $\approx -20 C^\circ$ ), corresponding to  $A_a \approx 10^{-26}$  (resp.  $A_a \approx 10^{-25}$ ), hence  $k \approx 1000$  (resp.  $k \approx 10$ ), see Fig. ?? (right). Moreover, If we assume 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. ?. It follows from (??) that:

$$\gamma \approx (1 - (0.78 \pm 0.15)R_s) \quad (\text{resp. } \gamma \approx (1 - (0.73 \pm 0.03)R_s)). \quad (19)$$

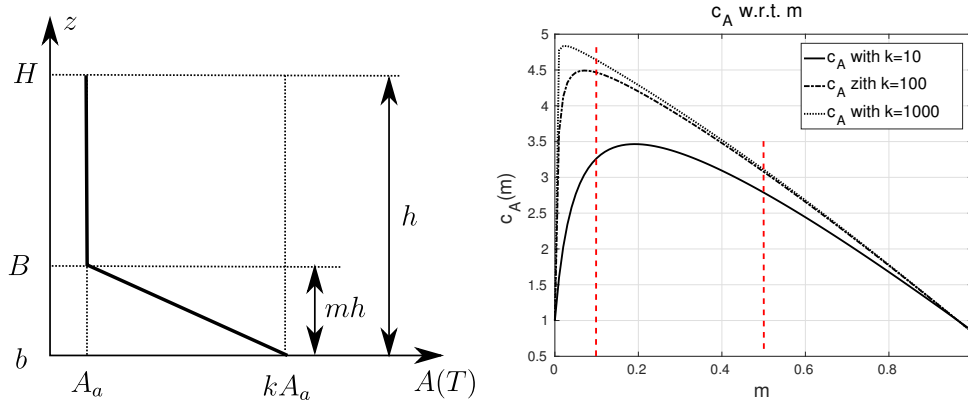


Figure 2: Left: typical vertical profile of the rate factor  $A(z)$ , see (??). Right:  $c_A$  vs  $m$ , see (??), with  $k = 10$  and  $k = 1000$ .

The present basic analysis shows the uncertainty orders of magnitudes; also it shows that the uncertainty on  $\gamma$  coming from the vertical thermal profile uncertainty (represented by the term  $(c_A/(q+2))$ ) is much smaller than those coming from the slip ratio  $R_s$  uncertainty.

## 2.4 The global inversion algorithm

The final goal of the study is to infer the depth (ice thickness)  $h$  by Variational Data Assimilation (VDA) from the RU-SIA equation (??). To do so, the following global inversion algorithm is considered.

- Step 1) Estimation of the effective diffusivity  $\eta = \gamma h$ , see (??), by VDA (formulation detailed in next section). Given the optimal value  $\eta^*$ , 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, the values of the RHS  $\dot{a}$  provided by Racmo2 [?, ?] are assumed to be exact.
- Step 2) Extension (interpolation / extrapolation) of  $\gamma$  outside the flights tracks by a classical universal Kriging process.
- Step 3) Estimation of the pair  $(h, \dot{a})$  by VDA. Not  $h$  only is inferred; that is to adjust the measurements of  $\dot{a}$  too (within the given uncertainty range). According to [?, ?], the uncertainty on  $\dot{a}$  is  $\pm \sim 20\%$ . This bound is imposed as control constraints of  $\dot{a}$  in the VDA process (optimal control problem).

## 3 Inference of the parameter $\gamma$

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

### 3.1 Identification of the parameter $\gamma$ along the flight tracks (Step 1)

The method to compute the values of  $\gamma$  along the flight tracks is presented; this corresponds to Step 1) of the global inversion algorithm.



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

$$\min_{\eta} j(\eta) \quad \text{with } j(\eta) = j_{obs}(\eta) + \alpha_{\eta} j_{reg}(\eta) \quad (20)$$

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

The control variable is  $\eta$ ,  $\eta \in \mathcal{U}_{ad}^{\eta} = \{\eta \in L^2(\Omega), \eta_{min} \leq \eta \leq \eta_{max}\}$ ; the bounds being defined from the numerical estimations (??). 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  $\text{dist}(\mathbf{x}, \Gamma_{tr}) < 3 \text{ km}$  and equals 0 otherwise. This definition is related to the length scale validity of the shallow flow model. The elevation value  $H_{\eta}$  is the solution of (??) with Dirichlet boundary conditions (set from the surface measurements); it is the state of the system. The scalar value  $\alpha_{\eta}$  is a positive constant to be set.

If the inverse problem would be well posed, that is the inverse control-to-state operator " $H_{\eta}^{-1}$ " exists and depends continuously on the data  $H^{obs}$ , if the elevation  $H^{obs}$  would be assimilated everywhere in the domain then the inverse problem is a Linear-Quadratic optimal control problem which would have a unique solution. However the observations are along the flights tracks only (the restriction operator  $\chi_{tr}$ ), although the target value is the optimal value  $\gamma^*$  along the flight tracks only, this inverse problem is a-priori ill-posed. Then, the classical Tikhonov regularization term  $j_{reg}$  is added, see e.g. [?] and references therein for regularising optimization problems. The minimisation problem (??) is numerically solved using the classical first order minimisation algorithm L-BFGS (Python routine `scipy.optimize.minimize`). The scalar weight coefficient  $\alpha_{\eta}$  can be chosen according to numerous rules, see e.g. [?]. In the next optimisation problem (next Section), an iterative regularisation procedure with iterative values of  $\alpha_{\eta}$  is considered. In the present optimisation problem (??), this strategy turned out be useless. Then various computations are performed with empirically set values of  $\alpha_{\eta}$ ; typically these values satisfies:  $\frac{j_{reg}(\eta^*)}{j_{obs}(\eta^*)} \approx 10^{-p}$  with  $p \in [2, 3]$ .

The complete VDA process is as follows. Given a control  $\eta$ , the direct model (??) is solved by a standard Lagrange Finite Element Method (FEM) order 2. Given this unique solution  $H_{\eta}$  the following adjoint equation is solved using the same FEM :

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

The gradient of the cost functional is computed from the state  $H_{\eta}$  and adjoint state  $P_{\eta}$  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  $\eta$  making decrease the cost function  $j$ . This iterative process is performed until convergence. In practice the convergence turns out to very robust; the optimal solution does not significantly depend on the length scale of the data smoothing nor on the restriction operator definition (more details are provided in the numerical results section).

**Resulting value of  $\gamma$  along the flight tracks.** The computed optimal solution of (??) is denoted by  $\eta^*$ . Assuming that the depth values over flight tracks (the airborne campaigns measurements) correspond to the correct values at the scale imposed in the definition of  $\chi_{tr}$ , i.e.  $h^*(\mathbf{x}) = h_b(\mathbf{x})$  for  $\mathbf{x} \in \Gamma_{tr}$ , the true value of  $\gamma$  is straightforwardly obtained as:

$$\gamma_{tr}^*(\mathbf{x}) = \frac{\eta^*(\mathbf{x})}{h_b(\mathbf{x})} \quad \text{for } \mathbf{x} \in \Gamma_{tr} \quad (21)$$

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

### 3.2 Extension of $\gamma$ outside the flights tracks (Step 2)

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

**The universal Kriging extension based on the locally observed trend** The field  $\gamma$  is decomposed as follows:

$$\gamma(\mathbf{x}) = \bar{\gamma}(\mathbf{x}) + \gamma_r(\mathbf{x}) \quad \text{for } \mathbf{x} \in \Omega \quad (22)$$

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  $\gamma(\mathbf{x})$ ).  $\forall \mathbf{x}, \mathbf{x}' \in \Omega$ ,

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

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

$$\hat{\gamma}(\mathbf{x}_0) = \langle \mathbf{c}, \mathbf{z} \rangle \quad (23)$$

with:  $\langle \cdot, \cdot \rangle$  the inner product,  $\mathbf{z}$  the vector of sampled points  $\mathbf{z} = (\gamma(\mathbf{x}_1), \dots, \gamma(\mathbf{x}_n)) \in \mathbb{R}^n$ ,  $\mathbf{x}_i \in \Gamma_{tr}$ ,  $\mathbf{c} = (c_1, \dots, c_n) \in \mathbb{R}^n$ ,  $c_i \in \mathbb{R}$ ,  $i = 1, \dots, n$  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:  $\text{argmin}_c \left( \sigma_{krig}^2(c) \right)$ ,

$$\sigma_{krig}^2 = \text{Var}[\hat{\gamma}(\mathbf{x}_0) - \gamma(\mathbf{x}_0)] = -\langle \mathbf{c}, V_Y \mathbf{c} \rangle + 2\langle \mathbf{c}, \mathbf{v}_{Y,0} \rangle \quad (24)$$

while the unbiasedness condition  $\mathbb{E}[\hat{\gamma}(\mathbf{x}_0) - \gamma(\mathbf{x}_0)] = 0$  is satisfied.

Here,  $V_Y \in \mathbb{R}^{n \times n}$ ,  $(V_Y)_{i,j} = v_Y(\mathbf{x}_i - \mathbf{x}_j)$ ,  $i, j = 1, \dots, n$ , and  $\mathbf{v}_{Y,0} = (v_Y(\mathbf{x}_1 - \mathbf{x}_0), \dots, v_Y(\mathbf{x}_n - \mathbf{x}_0)) \in \mathbb{R}^n$ .

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 context, the trend  $\bar{\gamma}(\mathbf{x})$  of the multi-physics parameter  $\gamma$  is unknown. However given a domain (e.g. portion of inland Antarctica), a trend can be determined from the flights track measurements and the corresponding values  $\gamma_b^*(\mathbf{x})$ ,  $\mathbf{x} \in \Gamma_{tr}$ . Doing so is equivalent to consider that the flight track measurements are sufficiently representative of the considered domain.

**Typical uncertainty on the interpolated-extrapolated value  $\gamma$**  Despite being measured, the depth value  $h^*(\mathbf{x}) \equiv h_b(\mathbf{x})$ ,  $\mathbf{x} \in \Gamma_{tr}$  in (??) is not exact, since averaged value, length scale definition of  $\chi_{tr}$  etc. Then let us set the true thickness  $h^*$  as :  $h^* = h_b + \delta h$ . It has been shown previously that  $\delta h \approx (0.05 \sim 0.1) h_b$ . If we set  $\gamma^* = (\bar{\gamma} + \delta \gamma)$  with  $\delta \gamma$  induced by the variation  $\delta h$ , it follows that:  $\eta^* = \bar{\gamma}^* h^* = (\bar{\gamma} + \delta \gamma)(h_b + \delta h)$ , hence:

$$\frac{\delta \gamma}{\bar{\gamma}} = -\frac{\delta h}{h_b + \delta h} \approx (0.047, 0.091).$$

In other words, in the forthcoming numerical results if the residual value  $\gamma_r$  in (??) turns out to be  $\approx (0.047, 0.091)\bar{\gamma}$  then it represents a similar uncertainty to the uncertainties on the measured value  $h_b$ .

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

Given the surface measurements, given the single multi-physics parameter  $\gamma$  estimated as described in the previous section, the inference of the depth (ice-thickness)  $h$  and the RHS  $\dot{a}$  in the RU-SIA equation (??) is obtained by VDA.

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

$$\min_{k \in \mathcal{U}_{ad}} j(k), \quad \text{with } j(k) = j_{obs}(k) + \alpha_k j_{reg}(k) \quad (25)$$

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

$$j_{reg}(k) = \frac{1}{2} \|(h - h_b)\|_{C_h^{-1}} + \frac{1}{2} \|\dot{a} - \dot{a}_b\|_{C_a^{-1}}$$

The scalar coefficient  $\alpha_k$  is given. The norms  $C_h^{-1}$  and  $C_a^{-1}$  are defined as the inverse of symmetric, positive covariance operators. The state  $H_k$  satisfies the RU-SIA equation (??) 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 the publicly available databases, e.g. Bedmap2 [?] for  $h$  in Antarctica and Racmo2 [?] for the RHS.

The admissible control set is:

$$\mathcal{U}_{ad} = \{(h, \dot{a}) \mid h(\mathbf{x}) \in [h_{min}(\mathbf{x}), h_{max}(\mathbf{x})], \dot{a} \in [\dot{a}_{min}(\mathbf{x}), \dot{a}_{max}(\mathbf{x})], \mathbf{x} \in \Omega\}.$$

The bounds  $h_{min}$  and  $h_{max}$  depend whether the point  $\mathbf{x}$  belongs to  $\Gamma_{tr}$  or not. The bounds  $\dot{a}_{min}$  and  $\dot{a}_{max}$  are provided by the Racmo2 database [?] (more details in the numerical results section).

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

$$\begin{cases} \partial_h j(k) &= - \int_{\Omega} \frac{|\mathbf{u}_H|}{S}(\mathbf{x}) \gamma(\mathbf{x}) \nabla H(\mathbf{x}) \nabla P(\mathbf{x}) d\mathbf{x} + \alpha_k \frac{\partial j_{reg}}{\partial h}(k) \\ \partial_{\dot{a}} j(k) &= - \int_{\Omega} d\mathbf{x} + \alpha_k \frac{\partial j_{reg}}{\partial \dot{a}}(k) \end{cases} \quad (26)$$

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

**Change of control variable.** Following e.g. [?, ?, ?, ?] (where the control variable is the initial state of an atmospheric model), a change of the control 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:

$$w = (w_1, w_2), \quad \text{with } w_1 = C_h^{-1/2}(h - h_b), \quad w_2 = C_a^{-1/2}(\dot{a} - \dot{a}_b) \quad (27)$$

In variable  $w$ , the optimisation problem (??) reads :

$$\min_{z \in \mathcal{U}_{ad}} j(w), \quad \text{with } j(w) = \frac{1}{2} \|H_w - H^{obs}\|_2^2 + \frac{\alpha_w}{2} (\|w_1\|_2^2 + \|w_2\|_2^2) \quad (28)$$

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. Obviously the exact covariances operators are unknown. To simply impose correct physical length scales of variations (the model is a shallow flow model), the following classical covariance operators are considered. For  $e \in L^2(\Omega)$ ,

$$C_{\square}e = \int_{\Omega} \sigma_{\square}(\mathbf{x})\sigma_{\square}(\mathbf{x}')c(\mathbf{x}, \mathbf{x}'; L_{\square})e(\mathbf{x})d\mathbf{x}, \quad (29)$$

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

$$c(\mathbf{x}, \mathbf{x}'; L) = \exp\left(-\frac{|\mathbf{x} - \mathbf{x}'|_1}{L}\right) \quad (30)$$

The standard deviation of  $(h - h_b)$  (resp.  $(\dot{a} - \dot{a}_b)$ ) is  $\sigma_h$ ,  $\sigma_h > 0$  (resp.  $\sigma_a > 0$ ); the length-scale is  $L_h > 0$  (resp.  $L_a > 0$ ) and  $|\cdot|_1$  denotes the  $L_1$  norm.

Numerous numerical experiments have demonstrated that this change of variable (??) and the present choice of covariance operators improve the robustness and the convergence speed of the VDA process.

It is worth to notice that for some correlation kernels - Green's functions  $c(\mathbf{x}, \mathbf{x}'; L)$ , it is possible to make a relationship with (higher-order) regularisation terms in the functional to be minimised, that is the term  $j_{reg}(k)$  in (??). Following the calculation method presented e.g. in [?] Section 7 (calculations available in 1D and 3D only in their case), it can be proved in the present case that the introduction of the covariance operator (??)(??) implies that:

$$j_{reg}(k) \sim \sum_{i=1,2} \frac{1}{8L_{ki}} \int_{\Omega} |e_{ki}(\mathbf{x})|^2 + L_{ki}^2 |\nabla e_{ki}(\mathbf{x})|^2 + L_{ki}^4 \frac{\partial^2}{\partial^2 x_1} e_{ki}(\mathbf{x}) \frac{\partial^2}{\partial^2 x_2} e_{ki}(\mathbf{x}) d\mathbf{x} \quad (31)$$

where  $e_{ki}(\mathbf{x}) = (k_i(\mathbf{x}) - k_{bi}(\mathbf{x}))/\sigma_{ki}(\mathbf{x})$  and  $k_1 = h$ ,  $k_2 = \dot{a}$ . The calculations are detailed in the appendix.

It follows from (??) that larger the length scale  $L_{ki}$  is, larger the regularisation effect is too. In other respect, it is easy to notice that the change of variable can be seen as a preconditioning to the optimisation procedure too. Indeed, a simple calculation shows that:

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

where  $\partial_h j(k)$  and  $\partial_a j(k)$  are given in (??). Therefore the optimal necessary condition  $\nabla_k j(k) = 0$  may be viewed "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,  $i, j = 1, \dots, N$ , reads, see (??):

$$\sigma_{k_m}(\mathbf{x}_i)\sigma_{k_m}(\mathbf{x}_j) \exp\left(-\frac{|\mathbf{x}_i - \mathbf{x}_j|}{L_{k_m}}\right) \Delta|\mathbf{x}_i| \Delta|\mathbf{x}_j|$$

Then the positivity of these matrix  $C_{k_m}$ ,  $m = 1, 2$  depend on the a-priori imposed length-scale  $L_{k_m}$ ,  $m = 1, 2$ . According to the Gerschgorin-Hadamard theorem, a sufficient condition to guarantee the positivity is to choose  $L_{k_m}$  such that :

$$\log\left(\frac{\sum_{j \neq i} \sigma_{k_m}(\mathbf{x}_j)}{\sigma_{k_m}(\mathbf{x}_i)}\right) L_{k_m} < \min_{i \in \{1, \dots, N\}} \Delta|\mathbf{x}_i|, \quad m = 1, 2, \quad i, j = 1, \dots, N.$$

This condition indicates that the values of  $L_{k_m}$ ,  $m = 1, 2$  should be chosen small enough such that the covariance matrix remain positive. However, it has been shown above that larger length-scales  $L_{k_m}$ ,  $m = 1, 2$  lead to a higher regularisation, see (??). That is to say, the choice of these length-scales  $L_{\square}$  has to respect a balance between the regularisation and the preconditioning effects. A similar phenomena is analysed into details e.g. in [?, ?] in a different context.

**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 (??) 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 matrix, the new variables are expected to be more physically correlated. The covariance matrices  $(C_h, C_a)$  represent prior information. The reader may refer to e.g. [?, ?] for investigations in VDA but with different physics than the present one.*

**Iterative regularisation strategy.** The weight parameter  $\alpha_w$  of the regularization term in (??) is set as a decreasing sequence  $\alpha_w^{(n)} > 0$ ,  $n = 1, \dots, n^*$ . We refer to [?, ?] and references therein for a description of various regularisation strategies.

Let us denote by  $F : \mathcal{U}_{ad} \mapsto \mathcal{Y}$  the operator that maps the control  $w \in \mathcal{U}_{ad} \subset \mathcal{X}$  to the observed data  $H \in \mathcal{Y}$ ;  $\mathcal{X}$ ,  $\mathcal{Y}$  are Hilbert spaces. According to [?] chapter 4, if using an iteratively regularized Gauss-Newton method, the stop iteration number  $n^* = n^*(\delta)$  can be chosen through the so-called Morozov' discrepancy principle [?], i.e.  $n^*$  is chosen such that :

$$\|H^\delta - F(w_{n^*}^\delta)\| \leq \tau\delta \leq \|H^\delta - F(w_n^\delta)\|, \quad 0 \leq k \leq n^* \quad (32)$$

with  $\tau > 1$  large enough. Here,  $\delta$  denotes the noise level i.e.  $\delta \geq \|F(w^*) - H^\delta\|$  where  $H = F(w^*)$  represents the exact data and  $H^\delta$  is the noisy data,  $H^\delta = F(w^*) + \delta H$  ( $\delta H$  the noise) .

In the present study, following [?] (Section 4.3), the weight parameter sequence  $\alpha_w^{(n)}$  is defined as:

$$\alpha_w^{(n)} = \alpha_w^{(0)} q^{[n/n_0]}, \quad n = 1, \dots, n^* \quad (33)$$

where  $n_0 > 1$  is the number of iteration for each  $\alpha_w^{(n)}$ ,  $[m]$  returns the maximum integer smaller than  $m$ ,  $\alpha_w^{(0)}$  and  $q$  are given constants,  $\alpha_w^{(0)} > 0$ ,  $0 < q < 1$ . The values of  $\alpha_w^{(0)}$ ,  $q$ ,  $n_0$  are chosen experimentally (e.g.  $q = 0.5$ ,  $n_0 = 5$  and  $\alpha_0 = 1$ ). The stop iteration  $n^*$  is chosen according to (??).

## 5 Inversion in an inland area of Antarctica

In this section, real data in a poorly monitored Antarctica area are considered. Estimations of the bed topography elevation are performed following the algorithm previously described, see Section ??.

### 5.1 Data description & first guesses

The considered area is indicated in Fig. ?? and Table ??. This test area corresponds to the flow model validity since  $|\mathbf{u}_H| \in \approx [1, 50^+] m/y$ . Moreover it presents minimal surface velocity amplitude large enough to consider the measurements accurate, see e.g. [?].

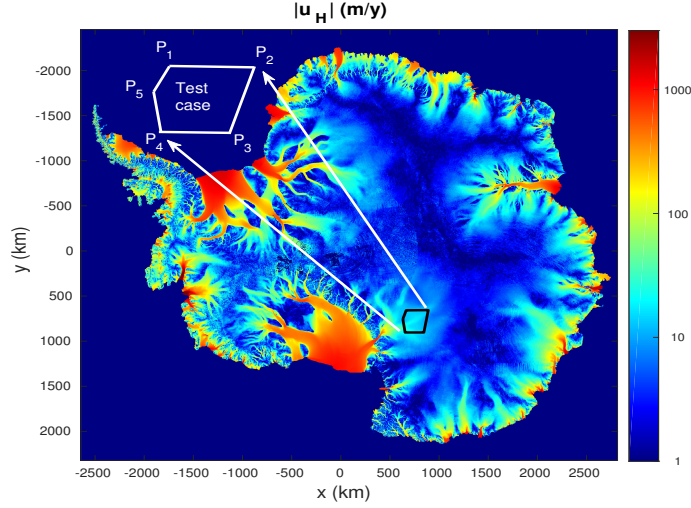


Figure 3: InSAR-Based Antarctica surface velocity from [?] and the test case location.

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

Table 1: Test case location and information.

The Antarctica location data are provided in Polar stereographic coordinate with true scale at  $71^\circ$  S. The Cartesian coordinates  $\mathbf{x} = (x, y)$  are transformed from these polar stereographic coordinates. The ice thickness mean value provided by BedMap2, [?], is:  $\bar{h}_b \approx 2.7 \text{ km}$ . The shallow flow model RU-SIA is valid as soon as  $\varepsilon = \frac{[H]}{[L]} \lesssim 0.1$ . Then the surface data  $|\mathbf{u}_H|$  and  $H$  need to be smoothed at the model scale i.e. at the  $\approx 27 \text{ km}$  length scale. To do this, the following Gaussian function is used:

$$G(x, y) = \frac{1}{2\pi\sigma_s^2} e^{-\frac{x^2+y^2}{2\sigma_s^2}} \quad (34)$$

When numerically computing this Gaussian function, the values for pixels located at a larger distance than  $3\sigma_s$  are close to 0. Therefore, all the surface data are smoothed with  $\sigma_s = 4 \text{ km}$  (since  $4 \times 6 = 24 \approx 27 \text{ km}$ ). The smoothed values of  $|\mathbf{u}_H|$  and  $H$  are plotted in Fig. ??; the effective diffusivity factor in the RU-SIA equation  $|\mathbf{u}_H|/\mathcal{S}$ , see (??), is plotted in Fig. ?? (Left). It can be noticed that this observational term varies by a factor  $\approx 6$  (ranging from  $\approx 0.4$  to  $2.4 \cdot 10^4$ ).

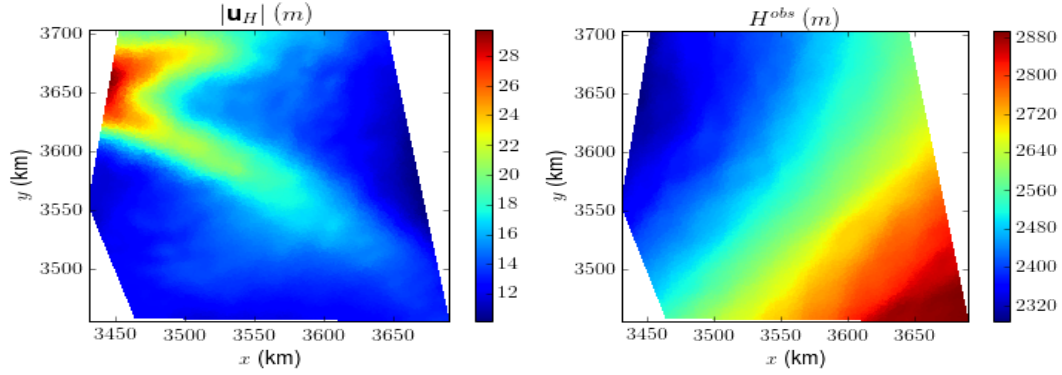


Figure 4: Surface data smoothed with  $\sigma_s = 4 \text{ km}$ : (Left) Surface velocity module  $|\mathbf{u}_H(\mathbf{x})|$ ; (Right) Surface elevation  $H(\mathbf{x})$ .

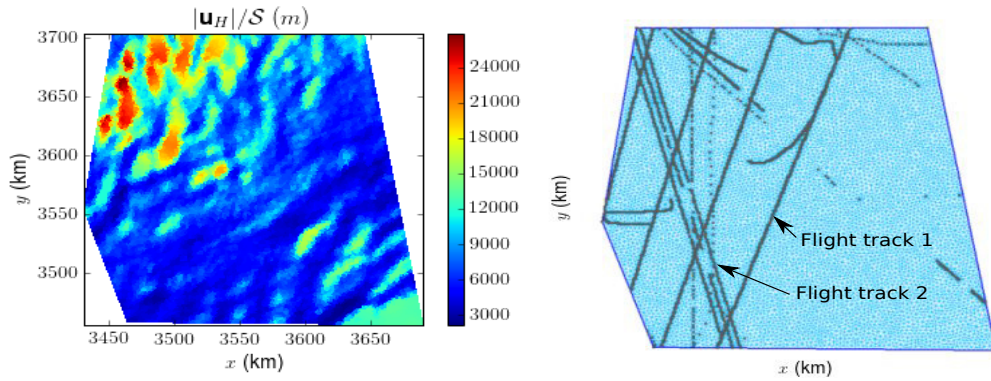


Figure 5: (Left) The single observational term  $\frac{|\mathbf{u}_H|}{S}$  in the RU-SIA equation, see (??), smoothed with  $\sigma_s = 4 \text{ km}$ . (Right) The finite element mesh with  $\Delta \mathbf{x} \approx 3 \text{ km}$  and the airborne measurements locations that are the flight tracks (with the particular flight track 1 and flight track 2 potentially removed from the data set for discussions).

The numerical experiments are systematically performed on a medium size mesh with  $\delta x \approx 3 \text{ km}$  and a finer one with  $\delta x \approx 1 \text{ km}$  to confirm the non sensitivity of the estimations with respect to the mesh size. Indeed  $\delta x \approx 3 \text{ km}$  provides  $\approx 10$  points per wave length (which equals  $\approx 27 \text{ km}$ ) hence nodes enough to properly approximate the equations. For the  $\approx 3 \text{ km}$  mesh, Fig. ?? (Right), the total number of vertices equals 8226. The meshes are generated by Gmsh software. Vertices are imposed to be on the flight tracks (the flight tracks constitutes the lines on the figure) and with a finer along the tracks ( $\delta x \approx 1 \text{ km}$  along the tracks). The airborne measurements locations are available in Bedmap2 [?] (with the corresponding measured depth).

**First guesses of  $h$ ,  $\dot{a}$  and a-priori uncertainties.** The natural first guess values for the VDA process, see Section ??, are the ice thickness  $h_b$  provided by Bedmap2 [?] and  $\dot{a}_b$  provided by Racmo2 [?]. This two first guess fields are presented in Fig. ??.

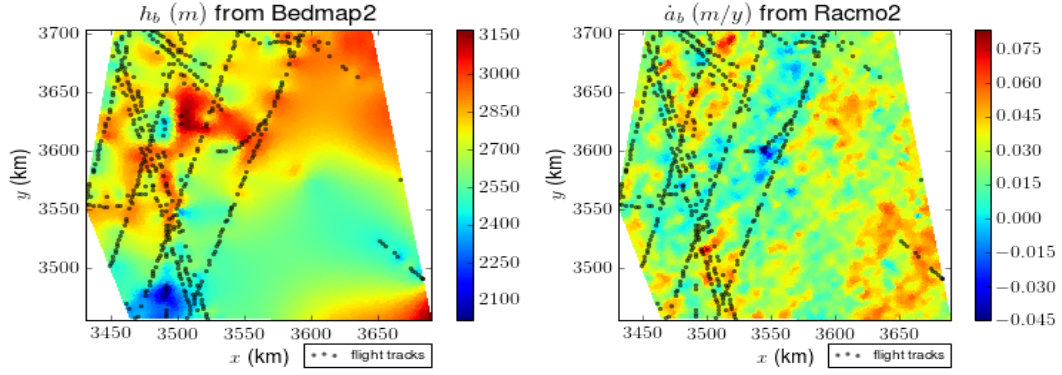


Figure 6: (Left) The depth (ice thickness)  $h_b(\mathbf{x})$  from Bedmap2 [?]. (Right) The SMB  $\dot{a}(\mathbf{x})$  from Racmo2 [?].

The bed estimations obtained in the present study are compared with the current reference bed topography, that is Bedmap2 [?]. Let us recall the Bedmap2 uncertainty source and its order of magnitudes. In [?], the depth measurements are splitted into two datasets (D1) and (D2). Dataset (D1) is used to build up an interpolation in the whole domain, including the domain where there is no measurement and at Dataset (D2) location points (the measurement of (D2) are not employed at this stage). Next, Dataset (D2) is used to quantify the accuracy of the interpolated field; basic statistics on the result are deduced. This experimental procedure provides the a-priori uncertainty of Bedmap2 estimations, [?].

Note that the interpolation is performed by employing the ArcGIS Topogrid routine (ESRI Ltd, ArcGIS 9) which is based on the ANUDEM algorithm [?]. This algorithm uses an optimised iterative finite difference interpolation technique which is essentially a thin plate spline technique [?]. In the present VDA process, a-priori boundaries of the "control variable"  $h$  are set from the Bedmap2 uncertainties described above; it gives the following bounds:

- if  $\mathbf{x} \in \Gamma_{tr}$  then inequality constraints on  $h$  are set with the bounds  $h_b(\mathbf{x}) \pm \delta_{h,tracks}$  with  $\delta_{h,tracks} = 140\text{ m}$ , see [?].
- if  $\mathbf{x} \notin \Gamma_{tr}$  then the bounds of the inequality constraints on  $h$  are  $\delta_{hb}(\mathbf{x}) = h_b(\mathbf{x})(1 \pm 0.6)$ .

While inferring  $\dot{a} = (a - \partial_t h)$ , inequality constraints are required too. According to [?], the uncertainty on this climatic-dynamic term equals approximately 20%. Therefore, we set lower and upper bounds of  $\dot{a}$  to be  $\dot{a}(\mathbf{x})(1 \pm \delta_a)$  with  $\delta_a = 0.2$ .

Moreover, these uncertainty estimations  $\delta_{h,tracks}$ ,  $\delta_{hb}$ , and  $\delta_a$  are also used in (??) as the standard deviations  $\sigma_h$  and  $\sigma_a$ , i.e.,  $\sigma_h(\mathbf{x}) \equiv \delta_{h,tracks}$  for  $\mathbf{x} \in \Gamma_{tr}$ ,  $\sigma_h(\mathbf{x}) = \delta_{hb}(\mathbf{x})$  for  $\mathbf{x} \notin \Gamma_{tr}$ , and  $\sigma_a = \delta_a$ . Note also that  $\sigma_h$  and  $\sigma_a$  are tuned numerically to have a desired balance between regularization terms of  $h$  and  $\dot{a}$ .

## 5.2 Estimation of the effective diffusivity $\eta$ and parameter $\gamma$

The effective diffusivity  $\eta$  of (??) is computed by solving (??). The iterative process is stopped following the usual criteria (that are the cost function does not decrease anymore, the gradient norm and the control variable variations vanish). Next, following the method described in Section ??, the parameter  $\gamma^*$  is immediately obtained from (??), in particular the targeted values along the flights tracks, see Fig ?? (Left).



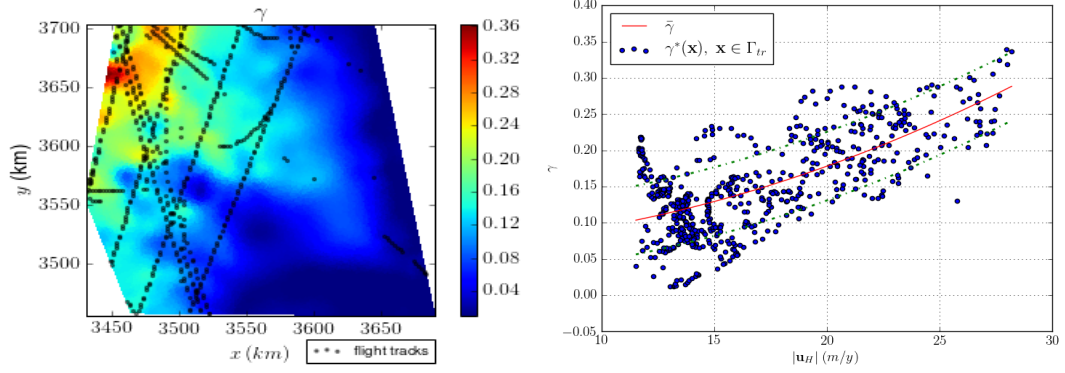


Figure 7: (Left) The parameter  $\gamma$  in the whole domain obtained from the relation  $\gamma = \eta^*/h_b$ ,  $\eta^*$  solution of (??). (Right) The targeted values along the flights tracks  $\gamma^*(\mathbf{x})$  vs  $|\mathbf{u}_H(\mathbf{x})|$ ,  $\mathbf{x} \in \Gamma_{tr}$  (blue points),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

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

Given the computed values  $\gamma_{tr}^*(\mathbf{x})$  along the flights tracks, an extension (interpolation-extrapolation) of  $\gamma$  is performed in the whole domain by the universal Kriging method described in Section ??; the values of  $\gamma$  on the flight tracks are assumed to be a realisation of the stochastic process.

The predicted  $\hat{\gamma}(\mathbf{x})$  defined by (??) is presented in Fig. ?? (Left), the variance  $\sigma_{krig}$  defined by (??) is presented in Fig. ?? (Right). Recall that  $\sigma_{krig}$  measures the variance between the Kriging predictor  $\hat{\gamma}(\mathbf{x})$  and the Kriging model defined by (??). The confidence interval can be defined as  $\hat{\gamma} \pm 3\sigma_{krig}$ .

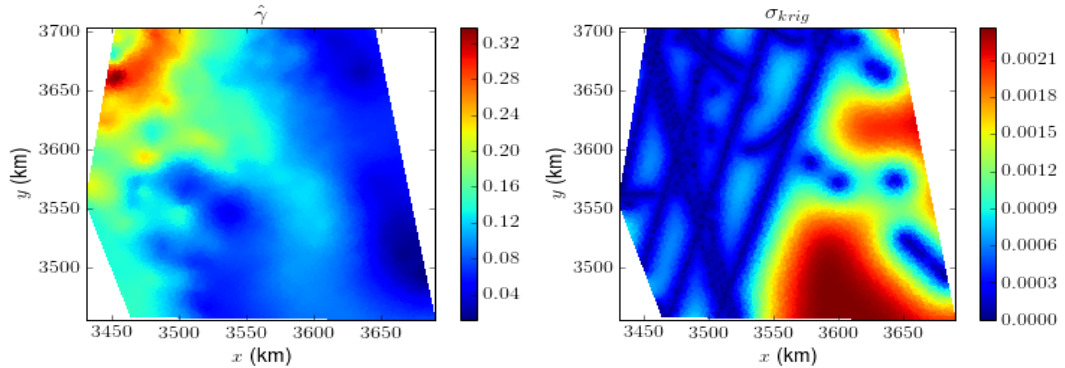


Figure 8: (Left)  $\hat{\gamma}$  inferred from universal Kriging. (Right) The corresponding variance  $\sigma_{krig}$ .

### 5.3 Inversion of $(h, \dot{a})$ with $\gamma$ given

Given the parameter  $\gamma$  estimated by Kriging as previously described, that is  $\gamma = \hat{\gamma}$ , first the RU-SIA equation is solved to compare its solution to the observations. Second the VDA process in variables  $(h, \dot{a})$  is performed, providing the sough bed topography.

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

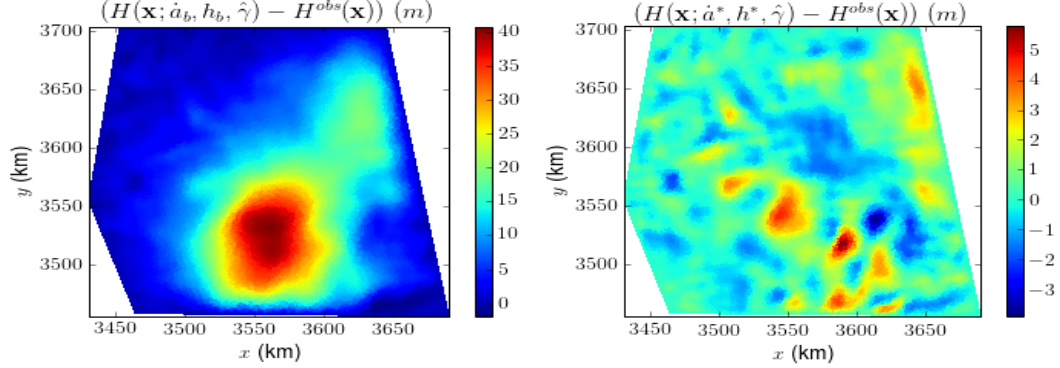


Figure 9: (Left) Direct model assessment: difference between  $H(\mathbf{x}; h_b, \hat{\gamma})$  and  $H^{obs}(\mathbf{x})$ . (Right) Misfit after VDA in variables  $(h, \dot{a})$ : difference between  $H(\mathbf{x}; h^*, \hat{\gamma})$  and  $H^{obs}(\mathbf{x})$ .

**VDA in variables  $(h, \dot{a})$ .** Given  $\gamma = \hat{\gamma}$  (the Kriging predictor), the VDA problem (??) is solved. The cost function terms evolutions ( $j_{reg}$  and  $j_{obs}$ ) are presented in Fig. ?? (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 (??). In the present example,  $n_0 = 3$  (internal iteration number for each  $\alpha_n$ ,  $n = 1, \dots, n^*$ ).

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

Table 2: Method performances

In Table ??, basic statistics on the numerical results are presented. As already mentioned, the RU-SIA model (??) fits already well the measurements  $H^{obs}$ . Next when performing the direct model from the optimal solution  $(h^*, \dot{a}^*)$  obtained by VDA, the misfit with the altimetry measurements ( $|H(\mathbf{x}; h^*, \hat{\gamma}) - H^{obs}(\mathbf{x})|$ ) decreases to 1 m (in average). The relative difference between  $h^*$  and  $h_b$  equals  $\approx 12\%$  (mean), outside the flight tracks i.e. for  $\mathbf{x} \notin \Gamma_{tr}$ . Of course, the same relative difference

remains very small along the flight tracks ( $\mathbf{x} \in \Gamma_{tr}$ ) since satisfying the imposed inequality constraints (corresponding to the measurements uncertainty and taking into account the uncertainty due to the model scale too). The difference of depth is presented in Fig. ?? (Right).

Concerning the RHS, the difference  $|\dot{a}^*(\mathbf{x}; \hat{\gamma}) - \dot{a}_b(\mathbf{x})|$  (mean) equals or is lower than the 20% corresponding to the measurements uncertainty indicated in [?]. The difference of depth is presented in Fig. ?? (Right). The ice volume change is very small, see Table ?. In some other areas (numerical experiments not shown in the present article), variations of the total volume of  $\approx 6\%$  have been obtained.

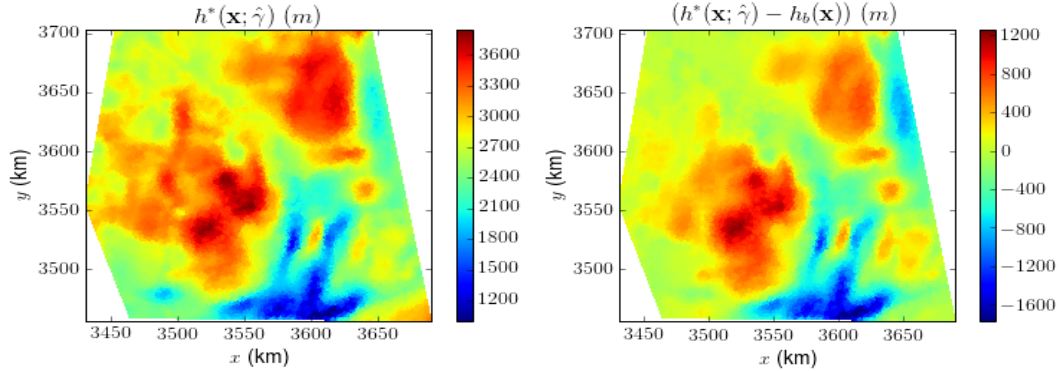


Figure 10: (Left) Inferred depth  $h^*(\mathbf{x})$  with  $\hat{\gamma}(\mathbf{x})$ . (Right) Difference between  $h^*(\mathbf{x})$  and Bedmap2  $h_b(\mathbf{x})$ .

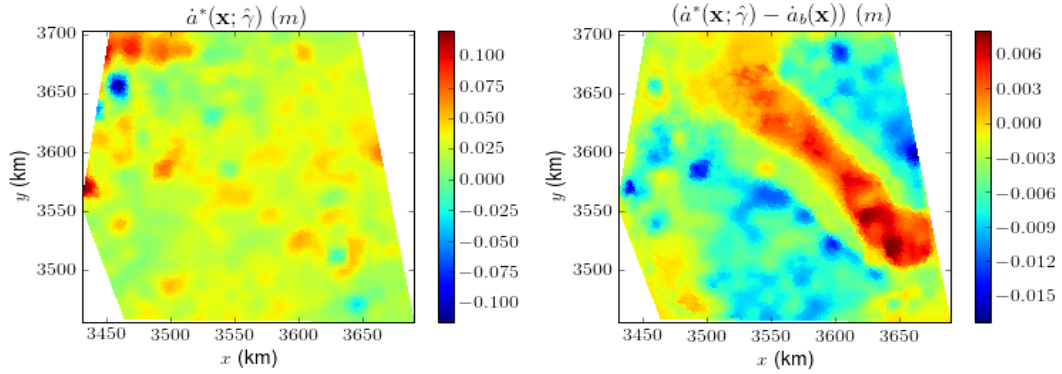


Figure 11: (Left) Inferred RHS  $\dot{a}^*(\mathbf{x})$  with  $\hat{\gamma}(\mathbf{x})$ . (Right) Difference between  $\dot{a}^*$  and  $\dot{a}^{obs}(\mathbf{x})$  from [?].

#### 5.4 Robustness of the inversions, uncertainties on the estimated depth

The inversion process to obtain the depth estimation indicated in Fig. ?? is quite complex and combine two main steps: estimation of the single parameter  $\gamma$  of the RU-SIA equation and estimation of  $(h, \dot{a})$ . Then numerous numerical experiments are performed to assess the robustness of the complete inversion process, and deduce basic uncertainty estimations on the results. The inversions are performed by considering: different estimations of  $\gamma$ , different density of flight tracks (e.g. by removing some), 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 ?.

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

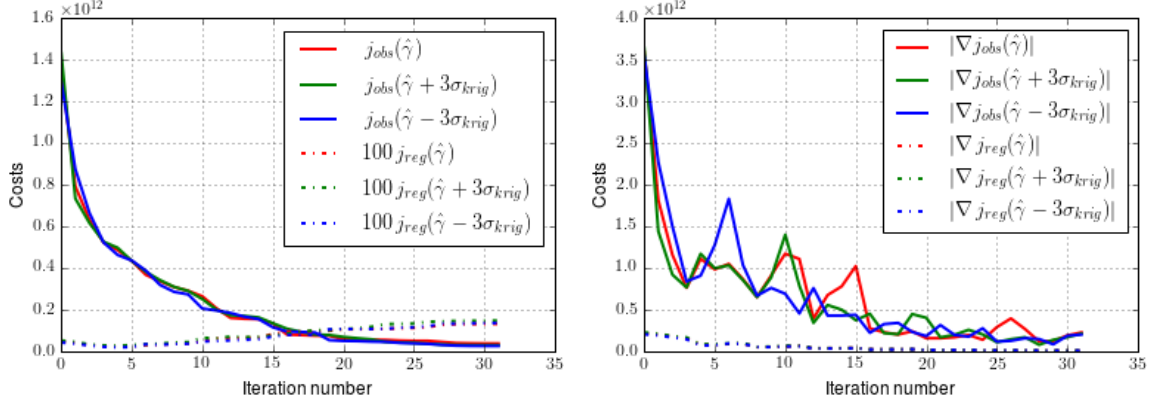


Figure 12: (Left) The cost function terms vs iterations solving (??) with  $\hat{\gamma}$ ,  $\hat{\gamma} + 3\sigma_{krig}$  and  $\hat{\gamma} - 3\sigma_{krig}$  respectively. (Right) The gradients of the corresponding cost function terms vs iterations.

The differences between  $h^*$  obtained with  $\gamma = \hat{\gamma}$  and  $h^*$  obtained with  $\gamma = \hat{\gamma} \pm 3\sigma_{krig}$  and  $\gamma = \hat{\gamma}$  are presented in Fig. ??. Also the performances are summarized in Table ??. The mean difference between the different inferred depth are  $\approx 2.8\%$ . These results shows the high robustness / low sensitivity of the depth inversions with respect to the single parameter  $\gamma$  uncertainty.

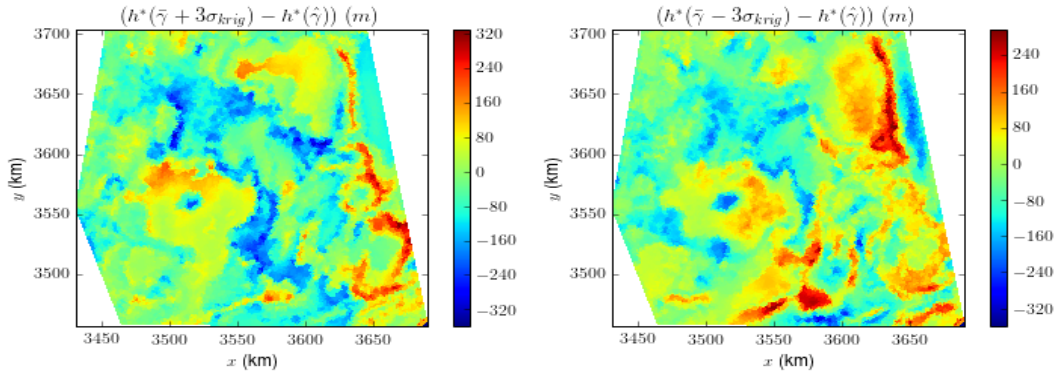


Figure 13: (Left) Difference between  $h^*(\mathbf{x}; \hat{\gamma})$  and  $h^*(\mathbf{x}, \hat{\gamma} + 3\sigma_{krig})$ . (Right) Difference between  $h^*(\mathbf{x}; \hat{\gamma})$  and  $h^*(\mathbf{x}, \hat{\gamma} - 3\sigma_{krig})$ .

**With less flight tracks.** Estimations of the bed are compared with  $h^*(\mathbf{x})$  obtained when assimilating the complete flight tracks set. The first "incomplete estimation" is obtained by assimilation the same flight tracks set *minus* the most internal one that is Flight track 1, see Fig. ?? (Right); the second "incomplete estimation" is obtained by assimilation the same flight tracks set *minus* Flight track 2, Fig. ?? (Right). The differences of the depth estimations are plotted in Fig. ??.

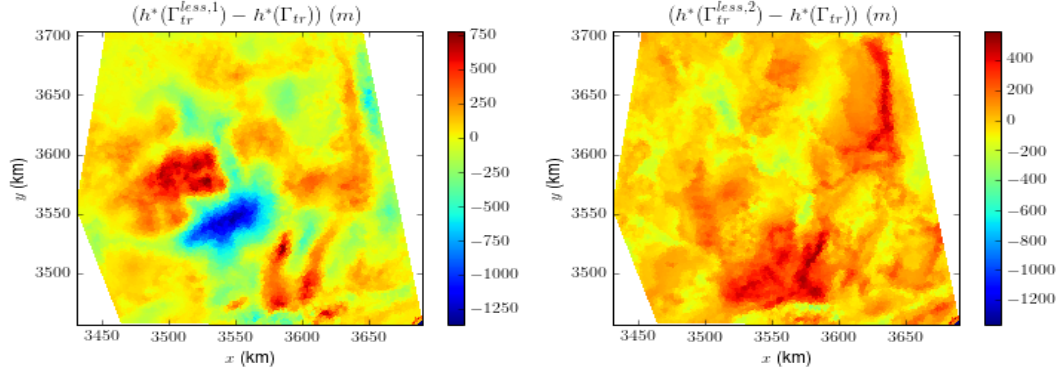


Figure 14: (Left) Difference between  $h^*(\mathbf{x})$  obtained with all flight tracks and with one less: Flight track 1 (set  $\Gamma_{tr}^{less,1}$ ); (Right) Difference between  $h^*(\mathbf{x})$  obtained all flight tracks and with one less: Flight track 2 (set  $\Gamma_{tr}^{less,2}$ ).

In the present case, the resulting trend function  $\bar{\gamma}$  remains similar, see Fig. ?? (Right) if removing Flight track 1 or Flight track 2. Therefore the differences between the computed depth are relatively small, lower than 6% (mean value), see Table ??; especially when removing Flight track 2, the difference is only 3.9% (mean). It is worth to recall that the trend function  $\bar{\gamma}$  fully depends on the airborne measurements along the flight tracks. However this result highlights a great feature of the present method: the direct model is a diffusive equation hence the inversions do not depend on the flight tracks location. This is a very important feature in practice. On the contrary, the depth-integrated mass equation is hyperbolic (it is the linear transport equation) then its inversion fully depend on the flights tracks locations and/or density (moreover with propagation of errors), see e.g. [?, ?] and the discussion in the present general introduction.

**With different smoothing parameter  $\sigma_s$ .** Let us recall that an important feature and limitation of the present method is the large scale inversion due to the shallow flow assumption. To be compatible with a shallow flow model, the surface data have been smoothed at the length scale  $\sigma_s = 4$  km, see (??). This corresponds to the minimal scale the physical model should be apply; and the largest scale one should consider would be  $\sigma_s \approx 8$  km. Then depth estimations computed from the surface data smoothed at  $\sigma_s = 4$ ,  $\sigma_s = 6$  and  $\sigma_s = 8$  km are compared, see Table ?? and Fig. ?. The mean relative difference  $|h_{s6}^*(\mathbf{x}) - h_{s4}^*(\mathbf{x})|/|h_{s4}^*(\mathbf{x})|$  equals 7.3%, while mean relative difference  $|h_{s8}^*(\mathbf{x}) - h_{s4}^*(\mathbf{x})|/|h_{s4}^*(\mathbf{x})|$  equals 8.6%. 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 the Bedmap2 bed elevation (which equals  $\approx 12\%$ , see Table ??).

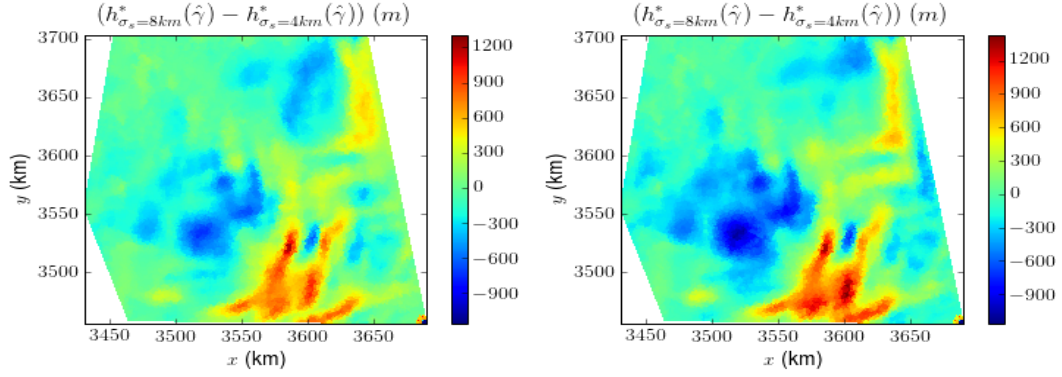


Figure 15: Left: difference between  $h_{\sigma_s=4km}^*(\mathbf{x})$  and  $h_{\sigma_s=6km}^*(\mathbf{x})$ , where  $\sigma_s$  is defined in Eq. (see (??)); Right: difference between  $h_{\sigma_s=4km}^*(\mathbf{x})$  and  $h_{\sigma_s=8km}^*(\mathbf{x})$ .

**With different first guesses of  $h$ .** The natural first guess for  $h$  is the current bedrock estimation Bedmap2. However to test the present VDA method with respect to the first guess value, two different first guesses are built up: 1)  $h'_b$  built up by spline interpolation of the airborne measurements (data along the flight tracks); 2)  $h''_b$  built up by adding a perturbation to  $h_b$  outside the flight tracks. After computation of  $h^*$ , the mean differences between  $h^*$  obtained from different initial guesses are less than 3.7%; the differences are plotted in Fig. ?? and statistics are presented in Table ?. 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 of the Bedmap2 bed elevation.

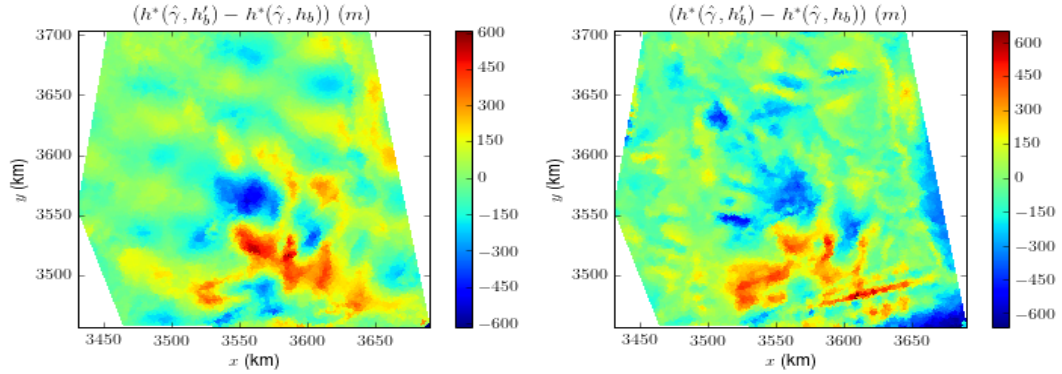


Figure 16: Robustness with respect to the first guess. (Left) Difference between  $h^*(\mathbf{x}; h_b)$  and  $h^*(\mathbf{x}; h'_b)$ . (Right) Difference between  $h^*(\mathbf{x}; h_b)$  and  $h^*(\mathbf{x}; h''_b)$ .

Comparison of $h^*$ obtained with ...	Median	Mean	Max
<b>... different <math>\gamma</math></b>			
$ h^*(\mathbf{x}; \hat{\gamma} + 3\sigma_{krig}) - h^*(\mathbf{x}; \hat{\gamma}) $	59 <i>m</i>	74 <i>m</i>	377 <i>m</i>
$ 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 <i>m</i>	73 <i>m</i>	357 <i>m</i>
$ h^*(\mathbf{x}; \hat{\gamma} - 3\sigma_{krig}) - h^*(\mathbf{x}; \hat{\gamma}) / h^*(\mathbf{x}; \hat{\gamma}) $	2.2%	2.8%	35.1%
<b>... different flight tracks densities</b>			
$ h^*(\mathbf{x}; \Gamma_{tr}) - h^*(\mathbf{x}; \Gamma_{tr}^{less,1}) $	95 <i>m</i>	161 <i>m</i>	1389 <i>m</i>
$ 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 <i>m</i>	198 <i>m</i>	1425 <i>m</i>
$ h^*(\mathbf{x}; \Gamma_{tr}) - h^*(\mathbf{x}; \Gamma_{tr}^{less,2}) / h^*(\mathbf{x}; \Gamma_{tr}) $	2.7%	3.9%	51.4%
<b>... different data smoothing <math>\sigma_s</math></b>			
$ h_{s4}^*(\mathbf{x}) - h_{s6}^*(\mathbf{x}) $	151 <i>m</i>	198 <i>m</i>	1425 <i>m</i>
$ 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 <i>m</i>	233 <i>m</i>	1560 <i>m</i>
$ h_{s4}^*(\mathbf{x}) - h_{s8}^*(\mathbf{x}) / h_{s4}^*(\mathbf{x}) $	5.8%	8.6%	60.3%
<b>... different first guesses</b>			
$ h'_b(\mathbf{x}) - h_b(\mathbf{x}) $	50 <i>m</i>	64 <i>m</i>	200 <i>m</i>
$ 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 <i>m</i>	69 <i>m</i>	377 <i>m</i>
$ 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 <i>m</i>	85 <i>m</i>	734 <i>m</i>
$ 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 <i>m</i>	96 <i>m</i>	801 <i>m</i>
$ 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 4*s* (resp. 6*s* and 8*s*) 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 (??). In the last comparison,  $h'_b$  and  $h''_b$  represent other first guesses than  $h_b$  (Bedmap); except for points along the flight tracks (i.e. for  $\mathbf{x} \in \Gamma_{tr}$ ).

## 5.5 Estimation of the thermal boundary layer

As highlighted in Section ??, the uncertainty on  $\gamma$  is due to the slip ratio  $R_s$  uncertainty and to the thermal - rheology parameter  $c_A$  uncertainty, see (??) and (??).  $c_A$  highly depends on  $A(z)$  hence on the thermal vertical profile, see Fig. ?. It follows from (??) that

$$c_A = (q + 2)(1 - \gamma)/R_s$$

In the present test case  $\gamma^* \in (0, 0.35)$ , see Fig. ??, and  $c_A$  belongs approximately to (0.8, 5), see Fig. ?? (Right). With  $q = 3$ , this implies that the slip ratio  $R_s$  ranges from  $\approx 0.65$  to 1. This (unusual) a-posteriori estimation of  $R_s$  is coherent with the surface velocity magnitudes, hence comforting on the flow model consistency.

Given an a-priori vertical thermal profile e.g. the typical one providing (??) and Fig. ?? (Left), the RU-SIA equation provides an estimation of the effective thermal boundary layer ( $B - b$ ), see Fig. ?? (Left). Next, this thermal boundary layer can be plotted vs different fields e.g. vs  $|\mathbf{u}_H|$ , see Fig. ?? (Right). Such a-posteriori model analysis may be useful for ice-sheet modellers and related studies (e.g. these profiles may be constrained by the very sparse measured in-situ temperature data).

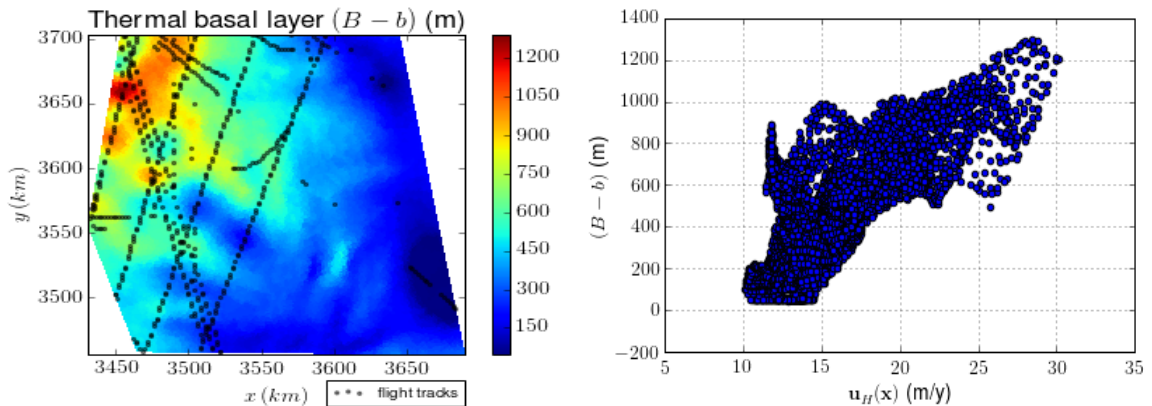


Figure 17: (Left) The thermal basal layer  $(B - b)$  resulting of the inversion of RU-SUA equation, see (??) and Fig. ??. (Left)  $(B - b)(\mathbf{x}) = m(\mathbf{x})h(\mathbf{x})$  with  $R_s = 1$ . (Right)  $(B - b)(\mathbf{x})$  vs  $|\mathbf{u}_H(\mathbf{x})|$

## 6 Conclusion

This study proposes a new inverse method to infer the bed topography beneath ice flows from surface observations (elevation and velocity) and sparse depth measurements. The flow model is valid for anisothermal highly and mildly-sheared flows (hence up to mildly-rapid). The inverse problem is particularly challenging since the bed topography surface signature has to be separated from the basal slip one and the thermal dependent vertical velocity profile too. The numerical results demonstrate the robustness of the method which provides the first physical-based depth inversions inland ice-sheets areas. The dedicated RU-SIA model (containing a single multi-physics parameter  $\gamma$ ) may be interesting for other purposes than depth estimations; for example to provide a-posteriori estimations of the thermal basal boundary layer. The 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 any shallow generalized-Newtonian creeping flows as soon as highly or mildly sheared e.g. lava flows (with the thermal field given), mud flows and many polymer flows. In the ice-sheet modelling context, the method presents many advantages such as a robustness independent of the airborne measurements locations and a reduced uncertainty of the depth estimations (particularly compared to the gravimetry inversions). The method limitations are: a) the increase of the uncertainty when the airborne measurements density decreases (the latter providing the a-priori model of  $\gamma$ ); b) the inversions scale. Indeed the direct model validity is limited by the long wave assumption that is the geometrical ratio  $\varepsilon = H^*/L^* \lesssim 0.1$  corresponding to a length scale  $L \sim 30$  km in ice-sheets.

This new inverse method can be straightforwardly extended to unsteady flows if the surface observations are given in time (assuming that the initial condition is either not important in the considered time scale or assuming it is more or less given). The elaborated inverse method is under exploitation and assessments for numerous others inland Antarctica areas.

**Acknowledgements.** The second author has been funded by a CNES TOSCA grant (oct. 2017-june 2018) during her postdoctoral contract at IMT-INSA. The authors acknowledge Jonas Verley from INSA-IMT who has nicely contributed to the understanding of the regularisation techniques introduced into the variational data assimilation process. The first author acknowledges E. Larour, M. Morlighem and H. Seroussi from JPL/Caltech-NASA and Univ. California of Irvine for co-funding him in 2014. The discussions the author had in this group greatly contributed to his understanding of the remote-sensed data.



## References

- [1] Alnæs M. *et al* 2015. The Fenics project version 1.5 *Arch. Numer. Softw.* 3 9-23
- [2] Bamber J. *et al* 2013. A new bed elevation dataset for Greenland *Cryosphere* 7 499-510
- [3] Bannister R.N. 2008. A review of forecast error covariance statistics in atmospheric variational data assimilation I: Characteristics and measurements of forecast error covariances *Q J R Meteorol Soc* 134.637 1951-1970
- [4] Baral, D. R., and K. Hutter. "Asymptotic theories of ice sheets and ice shelves." *Geomorphological Fluid Mechanics*. Springer, Berlin, Heidelberg, 2001. 227-278.
- [5] F. Bouttier and P. Courtier. "Data assimilation concepts and methods, March 1999". Meteorological training course lecture series. ECMWF, 2002.
- [6] Boutounet M. Monnier J. Vila J.P. 2016. Multi-regime shallow free surface laminar flow models for quasi-newtonian fluids *Eur. J. Mech.* 55 182-206
- [7] Cacuci, D. G., Navon, I. M., Ionescu-Bujor, M. *Computational methods for data evaluation and assimilation*. Chapman and Hall/CRC. (2016).
- [8] Chiles, J.-P., Delfiner P. *Geostatistics, Modeling Spatial uncertainty*. Wiley Series in Probability and statistics. (1999)
- [9] Cullen M.J.P. 2003. Four-dimensional variational data assimilation: A new formulation of the background-error covariance matrix based on a potential-vorticity representation. *Q J R Meteorol Soc* 129.593 2777-2796
- [10] Engl H.W. Hanke M. Neubauer A. 1996. *Regularization of Inverse Problems* Springer Science & Business Media
- [11] Fowler A. 2011. *Mathematical Geoscience* Berlin: Springer
- [12] Fowler, A. C., Larson, D. A. (1978). On the flow of polythermal glaciers-I. Model and preliminary analysis. *Proc. R. Soc. Lond. A*, 363 (1713), 217-242.
- [13] Fretwell P. *et al* 2013. Bedmap2: improved ice bed, surface and thickness datasets for Antarctica. *Cryosphere*
- [14] Greve R. and Blater H. 2009. *Dynamics of Ice Sheets and Glaciers* Berlin: Springer
- [15] Gudmundsson G. 2003. Transmission of basal variability to a glacier surface *J. Geophys. Res.: Solid Earth* 108
- [16] Gudmundsson G. 2008. Analytical solutions for the surface response to small amplitude perturbations in boundary data in the shallow-ice-stream approximation *Cryosphere* 2 77-93
- [17] Haben S.A. Lawless A.S. Nichols N.K. 2009. Conditioning of the 3DVar data assimilation problem *University of Reading, Dept. of Mathematics, Math Report Series 3*
- [18] Haben S.A. Lawless A.S. Nichols N.K. 2011. Conditioning of incremental variational data assimilation, with application to the Met Office system. *Tellus A* 63(4) 782-792.
- [19] Howat, I. M., A. Negrete, and B. E. Smith. "The Greenland Ice Mapping Project (GIMP) land classification and surface elevation data sets." *The Cryosphere* 8.4 (2014): 1509-1518.
- [20] Heining C. Sellier M. 2016. Direct reconstruction of three-dimensional glacier bedrock and surface elevation from free surface velocity *AIMS Geosci.* 2 45-63

- [21] Hutter, K. "Theoretical Glaciology: Material Science of Ice and the Mechanics of Glaciers and Ice Sheets (Mathematical Approaches to Geophysics)." (1983).
- [22] Hutchinson M.F. 1989. A new procedure for gridding elevation and stream line data with automatic removal of spurious pits *J. of hydrology* 106.3-4 211-232.
- [23] Kaltenbacher B. Neubauer A. Scherzer O. 2008. *Iterative regularization methods for nonlinear ill-posed problems* Walter de Gruyter
- [24] Katz D. *et al* 2011. Correlations of control variables in variational data assimilation *Q J R Meteorol Soc* 137.656 620-630
- [25] LeDimet F.X. Talagrand O. 1986. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects *Tellus A* 38 97-110
- [26] Lions, J. L. "Optimal control of systems governed by partial differential equations" Vol. 170. Springer. 1971.
- [27] Logg A. Wells G. Hake J. 2012. Dofin: A c++python finite element library Automated Solution of Differential Equations by the Finite Element Method *Berlin: Springer* 173-225
- [28] Lorenc A.C. *et al.* 2000. The Met. Office global three-dimensional variational data assimilation scheme *Q J R Meteorol Soc* 126.570 2991-3012
- [29] Martin N. Monnier J. 2014. Adjoint accuracy for the full-stokes ice flow model: limits to the transmission of basal friction variability to the surface *Cryosphere* 8 721-41
- [30] Martin N. Monnier J. 2015. Inverse rheometry and basal properties inference for pseudoplastic geophysical flows *Eur. J. Mech.* 50 110-26
- [31] Michel L. *et al* 2013. Estimating the ice thickness of mountain glaciers with an inverse approach using surface topography and mass-balance *Inverse Problems* 29 035002
- [32] Michel L. *et al* 2014. Estimating the ice thickness of shallow glaciers from surface topography and mass-balance data with a shape optimization algorithm *Comput. Geosci.* 66 182-99
- [33] Monnier J. des Bosc P.E. 2017. Inference of the Bottom Properties in Shallow Ice Approximation Models *Inverse Problems* 33
- [34] Monnier J., Zhu J. *et al* DassFlow (Data Assimilation for Free Surface Flows) computational software. INSA - University of Toulouse, CNES, CNRS. [www.math.univ-toulouse.fr/DassFlow](http://www.math.univ-toulouse.fr/DassFlow)
- [35] J. Monnier. "Variational data assimilation: from optimal control to large scale data assimilation." Open Online Course, INSA Toulouse. 2014. <https://www.math.univ-toulouse.fr/jmonnie/Enseignement/VDA.html>.
- [36] Morland L. 1984. Thermo-mechanical balances of ice sheet flow *Geophys. Astrophys. Fluid Dyn.* 29 237-66
- [37] Morlighem M. 2011. Ice sheet properties inferred by combining numerical modeling and remote sensing data *PhD Thesis Ecole Centrale Paris*
- [38] Morlighem M. *et al* 2014. High-resolution ice-thickness mapping in south greenland *Ann. Glaciol.* 55 64-70
- [39] Morozov V.A. Stessin M. 1993. Regularization methods for ill-posed problems. *Boca Raton, FL: CRC press.*

- [40] Mouginot, J., Rignot, E., Scheuchl, B., Millan, R. (2017). Comprehensive annual ice sheet velocity mapping using Landsat-8, Sentinel-1, and RADARSAT-2 data. *Remote Sensing*, 9(4), 364.
- [41] Noël B. *et al* 2018. Modelling the climate and surface mass balance of polar ice sheets using RACMO2-Part 1 : Greenland (1958-2016) *The Cryosphere* 12(3) 811
- [42] Prandtl L. 1938. Zur berechnung der grenzschichten *ZAMM J. Appl. Math. Mech.* 18 77-82 (Translated as “Note on the calculation of boundary layers” Tech. Memor nat. adv. Comm. Aero., Wash. 959)
- [43] Price P.B. *et al* 2002. 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
- [44] Radok U. Jenssen D. Budd W. 1970. Steady-state temperature profiles in ice sheets *Bull. Int. Assoc. Scient. Hydrol* 8(1) 36
- [45] Rasmussen L. 1988. Bed topography and mass-balance distribution of Columbia Glacier, Alaska, USA, determined from sequential aerial photography *J. Glaciol.* 34 208-16
- [46] Rignot E. Mouginot J. Scheuchl B. 2017. MEaSURES InSAR-Based Antarctica Ice Velocity Map, Version 2
- [47] Sasaki, Y. 1970. Some basic formalisms in numerical variational analysis. *Mon. Wen.Rev.* 98, 875-883.
- [48] Sellier M. 2016. Inverse problems in free surface flows : a review *Acta Mech.* 227 913-35
- [49] Schoof C. Hindmarsh C. 2010. Thin-film flows with wall slip : an asymptotic analysis of higher order glacier flow models *Q. J. Mech. Appl. Math.* 63 73-114
- [50] Seroussi H. *et al* 2013. Dependence of century-scale projections of the Greenland ice sheet on its thermal regime. *Journal of Glaciology*, 59(218), 1024-1034
- [51] Sellier M. Gessese A. Heining C. 2012. Analytical and numerical bedrock reconstruction in glacier flows from free surface elevation data
- [52] Tarantola A. 2005. *Inverse problem theory and methods for model parameter estimation* SIAM
- [53] Thorsteinsson T. *et al* 2003. Bed topography and lubrication inferred from surface measurements on fast-flowing ice streams *J. Glaciol.* 49 481-90
- [54] Van Wessem J.M. *et al* 2017. Modelling the climate and surface mass balance of polar ice sheets using RACMO2, part 2: Antarctica (1979-2016). *Cryosph. Discuss.* 1-35
- [55] Wahba G. 1990. *Spline models for Observational data* Siam, 1990.
- [56] Warner R. Budd W. 2000. Derivation of ice thickness and bedrock topography in data-gap regions over antarctica *Ann. Glaciol.* 31 191-7
- [57] Weaver A. Courtier P. 2001. Correlation modelling on the sphere using a generalized diffusion equation *Q J R Meteorol Soc* 127(575) 1815-1846

## Appendix

In this section, the relationship between the covariance operators (??)(??) introduced in the control variable and a higher-order regularisation term in  $j_{reg}$  (??) is detailed, that is the calculation of (??). The calculation is similar than those presented in [?] Section 7; however these latter are valid in 1D and 3D only. The present calculation is valid in 2D and with non-constant standard derivations  $\sigma_{\square}$ . The calculation related to the variable  $h$  only is presented, the calculation for  $\dot{a}$  being the same. For a sake of simplicity, the subscripts  $_h$  are skipped.

Let  $\delta h = (h - h_b)$ , then

$$\|\delta h\|_{\mathbf{C}_h^{-1}}^2 = \langle \delta h, \mathbf{C}_h^{-1} \delta h \rangle = \int_{\Omega} \delta h(\mathbf{x}) \tilde{h}(\mathbf{x}) d\mathbf{x}, \quad (35)$$

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

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

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

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

It follows:

$$\tilde{h}(\mathbf{x}) = \frac{1}{\sigma(\mathbf{x})} \left[ \mathcal{F}^{-1}\left(\frac{1}{G(\zeta)}\right) * \left(\frac{\delta h(\mathbf{x})}{\sigma(\mathbf{x})}\right) \right]. \quad (36)$$

By applying the Fourier inverse  $\mathcal{F}^{-1}\left(\frac{1}{G(\zeta)}\right)$ , 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  $\delta$  denotes the Dirac distribution. Therefore, according to (??), it follows:

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

Finally, according to (??), it follows:

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