

1 Physically-constrained data-driven inversions to infer the bed topography beneath 2 glaciers flows. Application to East Antarctica

3 Jérôme Monnier · Jiamin Zhu

4
5 Received: date / Accepted: date

6 **Abstract** A method for inferring bed topography beneath glaciers from surface measurements (elevation from altimetry
7 and velocity from InSAR) and sparse thickness measurements is developed and evaluated. The method is based on
8 an original non-isothermal Reduced Uncertainty (RU) version of the Shallow Ice Approximation (SIA) equation that
9 natively incorporates the surface measurements. The flow model has a single dimensionless multi-physics parameter γ .
10 This parameter takes into account the basal slipperiness and the variable vertical rate factor profiles, thus the vertical
11 thermal variations. The inversions are based on three steps involving: an Artificial Neural Network (ANN) and two
12 Variational Data Assimilation (VDA) processes. The ANN-based stage aims at estimating the multi-physics number γ
13 from the thickness measurements; the resulting estimator is remarkably robust. The full inversion method is valid for
14 half-sheared flows (presenting a moderate basal slipperiness): it can be applied to inland ice-sheets areas. Also these
15 estimates connect continuously with estimates from mass conservation only, i.e. with areas of sliding flows. Numerical
16 results are presented for areas of the East Antarctica Ice Sheet where bed elevation can be very uncertain (Bedmap2
17 values). Estimates are valid for wavelengths longer than $\sim 10\bar{h}$ (due to the long wave assumption, shallow flow model)
18 with resolution at $\sim \bar{h}$ (\bar{h} a characteristic thickness value).

19 **Keywords** Variational data assimilation · reduced flow model · deep learning · inference · topography · glaciers ·
20 Antarctica.

21 1 Introduction

22 Bed topography elevation is a necessary input data to set up ice flow models. Also, when combined with surface to-
23 pography measurements (e.g. acquired by altimetry), these data directly provide the ice volume. For the Antarctica and
24 Greenland ice-sheets, ice thickness measurements are available along airborne radio-echo sounding tracks e.g. in the
25 CReSIS RDS database ¹. These measurements are particularly dense in fast flowing areas along the coast. However, the
26 measurements are very sparse or non-existent in the deep land, see [4, 14] and references therein.

27 Many satellites have provided (and still provide) accurate measurements of the surface of ice sheets: altimeters provide
28 the surface elevation H at $\approx \pm 20$ cm for 1 km² pixels see e.g. [22, 4], radar interferometers (InSar) provide the surface
29 velocity \mathbf{u}_H see e.g. [43].

30 To obtain bed elevation values everywhere beneath the glaciers, the challenge is to infer its value between the thick-
31 ness measurement locations, given surface measurements everywhere. A simple method to try to solve this challenge is
32 to apply ordinary Kriging (the interpolation method classically used in geostatistics). This interpolation method can be
33 interpreted as a Gaussian process regression providing the best linear unbiased prediction. This is the method used to
34 obtain bed elevations for Greenland and Antarctica in [3, 14]. For the reference Bedmap2 estimates in Antarctica, [14],
35 the authors suggest that for cells within 20km of a measurement, the error is an increasing function of distance. Beyond
36 20 km, the error would be uncorrelated with distance. For cells located more than 50 km from any measurement, the
37 Bedmap2 estimate is based on the gravity field inversion. Therefore, the estimates far from any measurement have very
38 large uncertainties, up to $\pm \sim 1000$ m according to [14].

39 To reduce uncertainties in the current bed elevation beneath ice-sheets, combining physically-informed models and

J. Monnier
INSA & Institut de Mathématiques de Toulouse (IMT), France
E-mail: jerome.monnier@insa-toulouse.fr

J. Zhu
INSA & Institut de Mathématiques de Toulouse (IMT), France

¹ CReSIS RDS Data <http://data.cresis.ku.edu> (NASA, NSF, Kansas fundings).

40 datasets seems to be the right direction. To do this, a key point is to employ a flow model that is complex enough
 41 to be sufficiently representative but simple enough to lead to stable and well-posed inverse problems (i.e. not leading to
 42 severe equifinality issues). In inverse modeling of geophysical flows, equifinality issues are the common pitfall to avoid,
 43 see e.g. [5].

44
 45 In fast ice streams ($\approx 1 - 10$ km/y in Antarctica), the flows are plug-like (glaciers slide). In this case, the momentum
 46 equation becomes trivial and inverting the (depth-integrated) mass equation allows to fill in the gaps downstream (and
 47 upstream) of the measurements, see [46,40]. Because of the nature of the depth-integrated mass equation, measurements
 48 (e.g. flight tracks during airborne campaigns) must be acquired cross-lines and relatively dense manner. Indeed, it is well
 49 known that this transport equation is intrinsically unstable when inverted, moreover it propagates errors, see e.g. [36] and
 50 references therein. To locally blur this feature, one must introduce artificial diffusion which regularizes the equation. The
 51 inversion of the mass equation was proposed in [8,46]; then it was combined with surface measurements by Variational
 52 Data Assimilation (VDA) in [41,42]. This approach has used to estimate the bed elevation under ice streams, particularly
 53 in Greenland [40,42].

54 In [7], a Bayesian framework is applied to provide probability distributions of the thickness assuming Gaussian covari-
 55 ance structures of the input data. The algorithm is equivalent to ordinary Kriging if no prior data are available. Bayesian
 56 approaches are particularly rich, however, in this case they again relied on mass conservation.

57 For fully sheared flows, the isothermal Shallow Ice Approximation (SIA) flow model with no slip at bottom, was inverted
 58 in [37,21]. The obtained estimates are robust but are only relevant for a very restricted flow regime: fully sheared flow
 59 areas, thus very slow flows ($\approx 1 - 10$ m/y in Antarctica).

60 In moderately sliding flows, equivalently moderately sheared flows (medium slip, medium shear), the slipperiness at bot-
 61 tom must be considered in the inversion: it is an additional unknown parameter in addition to the rate factor modeling
 62 the internal deformation. Inversions in this flow regime are much more difficult than those in the two regimes described
 63 above. The inverse problem is here a-priori ill-posed.

64 Let us cite other studies related to bedrock estimations. [25] inverts the 1D depth-integrated SIA equation with slip and a
 65 shape factor that models the 3D characteristics of the flow. This leads to an ill-posed inverse problem, but the inversions
 66 are performed by imposing empirical constant values for the unknown parameters. Based on the inversion of the com-
 67 plete hybrid SIA-SSA system called PISM [55] (SSA for Shallow Shelf Approximation), [51] uses an empirical iterative
 68 method to calibrate the bed elevation. In [38], the SIA flow model with slip at bottom is inverted by distinguishing differ-
 69 ent subregimes. The developed half-analytical half-computational inversion methods lead to well-posed inverse problems
 70 and thus stable inversions. The only weakness of the study lies in the ice rate factor which is supposed to be constant.
 71 Unfortunately this assumption is not realistic for the Greenland and Antarctica ice-sheets.

72 Inland areas of ice-sheets with medium surface velocities ($\approx [10 - 100]$ m/y in Antarctica) correspond to moderately slip-
 73 ping / moderately sheared flows. These flows cannot be accurately modeled by either mass conservation alone, plug-like
 74 flow models (e.g. the SSA model), or fully sheared flow models (e.g. the classical SIA model with no slip at bottom).
 75 For these half-slip half-shear flows, the measured surface characteristics (elevation and velocity) are the signature of the
 76 bottom slip *and* internal deformation. Furthermore, the internal deformation depends on the constitutive behavior of the
 77 ice *and* the vertical thermal profiles. Therefore, the inverse problem in this case is very complex.

78 In addition, moderately sheared interior flows have received poorly coverage in airborne campaigns because they are
 79 large and distant areas, and therefore to fly over. The flow model to invert must be stable and robust, even in the absence
 80 of local in-situ data. This mathematical property is anything but trivial to obtain, see e.g. [2,36,48,38]. In addition, it
 81 would be very helpful if the flow model inversion were as insensitive as possible to the measurements locations.

82 Note that an *effective* bed topography only can be inferred from the surface signature. Indeed, glacier flows act as low-
 83 pass filters: bed variations are filtered by the flow with filtering characteristics that depend on the flow regime, see [18,
 84 34,35] for a detailed discussion. Given a flow regime and a flow model, this then leads to a notion of minimal inferable
 85 wavelength, see [19,34,35].

86 Finally, a comparison of various inverse methods for estimating bed elevation under glaciers (or equivalently ice thick-
 87 ness) is presented in [12]. The comparison is based on numerous test cases representing a broad spectrum of ice flow
 88 regimes. For all test cases, no thickness values are assumed to be known. The 15 methods compared are categorized by
 89 type of resolution and not by the range of validity of the method (e.g. based on the flow regime). Numerical comparisons
 90 are presented, but no analysis of the equifinality problems present in these inversions is offered.

91
 92 The present study aims to solve the following inverse problem: estimating ice thickness (equivalent to bed eleva-
 93 tion) in half-sheared half-sliding flows with a physically-informed data-driven hybrid method. The targeted flow regimes
 94 require consideration for the full physics of the flows. To do so, a key ingredient is the RU-SIA (RU for Reduced Uncer-
 95 tainties) model derived in [39]. This flow model is dedicated to the present inverse problem by intrinsically ("natively")
 96 incorporating surface measurements into its coefficients. It is a complete multi-physics flow model, depth-integrated
 97 (shallow flow, long-wave assumption). It respects a "balanced complexity" for its inversion. It takes into account mass
 98 conservations and momentum conservation, with a temperature-dependent rate factor: the internal deformation is non-

uniform, depending in particular on the vertical temperature profiles. In the RU-SIA model, all the complex multi-physics phenomena are consistently represented by a new *dimensionless parameter* denoted by γ .

The inversion method developed here is based on a combination of VDA algorithms and a purely data-driven inversion, as in our previous study [39]. However in the one test case studied in [39], a clear correlation existed between γ and one of the observable fields ($|\mathbf{u}_H|$), see [39] Fig. 7. Under these conditions, the estimation of γ by a simple Kriging method was possible. However, upon subsequent application of the method to other areas (e.g. those considered in the present study), it was found that no such correlations existed. Therefore, another method to try to estimate this dimensionless multi-physics parameter γ had to be investigated. This is what we do here using a Neural Network Residual Kriging (NNRK) algorithm, see [11,30]. Therefore a new inversion method to deduce the bed elevation has been developed; this is the method presented here. The data employed remain the same: surface measurements (elevation, velocity) plus some in-situ thickness values. The inversion method relies on the RU-SIA equation derived in [39], a first advanced VDA process, a deep Artificial Neural Network (ANN) and finally a last VDA process that allows to conclude. The ANN aims to estimate the dimensionless multi-physics parameter γ of the RU-SIA model from in-situ ice thickness measurements. These in-situ measurements are available along flight tracks of airborne campaigns. The full inversion method is shown to be mathematically and computationally robust. The areas considered are the East Antarctica Ice Sheet (EAIS) regions that a-priori respect the validity domain of the flow model. This corresponds to six large interior regions of EAIS. Recall that it is of interest to accurately estimate the elevation of the EAIS bed because global warming may threaten its stability, particularly around some of the considered areas here, see [13].

In addition a remarkable relationship between current inversions based on the RU-SIA flow model and the widely employed mass conservation method, see [46,40,42], is presented. This relationship opens promising prospects for completing estimates based on mass conservation alone in deeper regions of ice-sheets. Indeed, we show mathematically that the present and the mass conservation-only estimates meet at the interface of their respective domains of validity, i.e. at the boundaries of plug-like flows (e.g. at and beyond sheared margins).

The present inversion method may be applied to all glacier flows as long as the flow model assumption is satisfied i.e. from strongly sheared flows to half-sheared half-sliding flows. Due to the long wave assumption of the flow model, the thickness estimates are valid at a wavelength of $\approx 10 \times \bar{h}$, \bar{h} a characteristic thickness value.

The outline of the article is as follows. In Section 2, the RU equation developed in [38,39] is recalled; its domain of validity is highlighted; the uncertainty of the dimensionless parameter γ is analyzed. Then, the inversion method is detailed: Step 1) and Step 3) of the overall algorithm aim at inverting the RU-SIA equation by VDA (physically-informed inversions); Step 2) aims at estimating γ using a NNRK algorithm (deep learning, purely data-driven). In Section 4, the six large EAIS areas considered (named Ant_p , $p = 1, \dots, 6$) are presented. The methods for obtaining the reference Bedmap2 estimates [14] are briefly recalled. Each computational step of the present inversion method is analyzed in detail. In Section 5, the robustness of the estimations is evaluated in detail for the Ant_1 and Ant_3 cases, in particular their sensitivities with respect to the presence or not of additional flight tracks. A conclusion is proposed in Section 6. As a complement, the computed thickness estimations for four other areas Ant_p , $p = 2, 4, \dots, 6$, are presented in Appendix.

2 Method

In this section the inversion method to estimate the ice thickness h (equivalently the bed topography elevation b) is detailed. It is done in three steps. Step 1) aims at estimating the product (γh) by assimilating all surface data (altimetry, InSAR and climatic term Source Mass Balance) plus the in-situ thickness measurements in the RU-SIA flow model (model presented below). Step 2) aims at estimating the dimensionless multi-physics parameter γ from the in-situ measurements only (measurements available along the flight tracks of airborne campaigns). Step 3) aims at estimating the thickness h (and adjusting a climatic term) by assimilating all available data again. The final output of the inversion method is the ice thickness h , therefore the bed topography elevation b .

2.1 The RU-SIA flow model

The RU-SIA equation is obtained by reformulating the depth-integrated SIA model with basal slipperiness (see e.g. [17] Chapter 5) but with a non constant rate factor and by natively integrating the surface measurements (elevation and velocity). The resulting 2D depth-integrated flow model is original; it is relevant for large scale sheared flows with moderate slipperiness at bottom and with non constant vertical temperature profile. The various uncertain multi-physics parameters (constitutive law exponent, flow regime, temperature dependent term) are gathered into the single dimensionless parameter γ . As a consequence, all the physical parametrisation uncertainty is represented by this original dimensionless

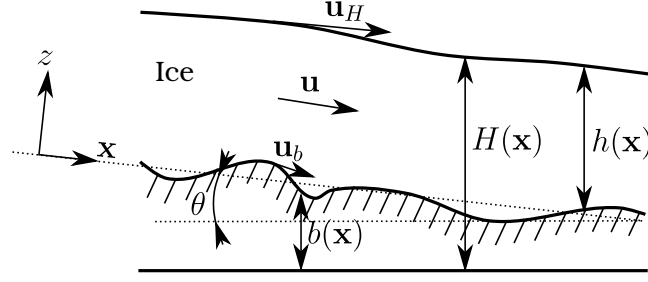


Fig. 1 Schematic vertical view of the gravitational ice flow and notations

151 parameter. The basic RU-SIA model assumptions are the same as those for the SIA model (classical lubrication the-
 152 ory) with basal slipperiness that is: the flow is necessarily sheared (normal stress components are negligible) and it is
 153 "shallow" (long wave assumption).

154 2.1.1 The equations and the dimensionless parameter γ

155 The surface slope is denoted by $\mathcal{S} = |\nabla H|$; \mathbf{u}_H is the surface velocity and \mathbf{u}_b is the velocity at bottom (basal velocity).
 156 The depth (ice thickness) is denoted by h , $h = (H - b)$ with H the ice surface elevation and b the bed elevation, see Fig.
 157 1. θ is a potential mean slope value in the (x, y) -plane ($\theta = 0$ in the forthcoming test areas).

158

159 *The flow equation.* The depth-integrated flow model SIA model with slipperiness at bottom (see e.g. [17] Chapter 5) is
 160 derived in a non isothermal case in [38] providing the so-called xSIA (x for extended) equation. Next in [39], by taking
 161 advantage of the measured surface features (elevation and velocity norm), the xSIA model is re-formulated to obtain the
 162 RU-SIA model. The RU-SIA equation reads as follows:

$$-\operatorname{div} \left(\frac{|\mathbf{u}_H|}{\mathcal{S}} \gamma h \nabla H \right) (\mathbf{x}) = \dot{a}(\mathbf{x}) \text{ in } \Omega \quad (1)$$

163 \mathbf{x} denotes the space variable and Ω the considered domain (an open set of \mathbb{R}^2). The RHS \dot{a} is the classical one defined
 164 by: $\dot{a}(\mathbf{x}) = (\partial_t h - a)(\mathbf{x})$ with a the mass balance (accumulation and ablation), see e.g. [17].

165
 166 In (1), the term $\frac{|\mathbf{u}_H|}{\mathcal{S}}$ is the observational term; it may provided by InSAR and altimetry surface measurements.
 167 The considered unknown of (1) is the surface elevation H . (1) is a linear diffusive equation in H , it is closed by Dirichlet
 168 condition at boundaries: H given e.g. by altimetry.
 169 Assuming that h (or equivalently b) is given, (1) contains γ as the single uncertain parameter.

170

171 *The dimensionless parameter:* γ is the dimensionless parameter of the equation, its expression is, [39]:

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

172 R_s is the slip ratio describing the flow regime; it is defined as: $R_s(\mathbf{x}) = 1 - \frac{|\mathbf{u}_b|}{|\mathbf{u}_H|}(\mathbf{x})$.

173 The parameter c_A is defined by:

$$c_A(\mathbf{x}) = [(q+2) - (q+1)R_A(\mathbf{x})] \quad (3)$$

174 where q is the constitutive power-law exponent ($q = 3$ in the classical Glen's law, see e.g. [17] Chapter 5, [38]) and
 175 $R_A(\mathbf{x}) = \frac{\bar{A}(\mathbf{x})}{\underline{A}(\mathbf{x})}$.

176 The parameters $\bar{A}(\mathbf{x})$ and $\underline{A}(\mathbf{x})$ are the depth-integrated quantities naturally appearing if the rate factor A depends on
 177 (\mathbf{x}, z) . Their expressions are as follows, see [38]:

$$\bar{A}(\mathbf{x}) = \frac{(q+2)}{h^{q+2}(\mathbf{x})} \left(\int_b^H \int_b^z A(\mathbf{x}, \xi) (H(\mathbf{x}) - \xi)^q d\xi dz \right)^{-1} \quad (4)$$

178

$$\underline{A}(\mathbf{x}) = \frac{(q+1)}{h^{q+1}(\mathbf{x})} \int_b^H A(\mathbf{x}, z) (H(\mathbf{x}) - z)^q dz \quad (5)$$

If the vertical profile of A is constant (A constant in z) then: $\bar{A}(\mathbf{x}) = A(\mathbf{x}) = \underline{A}(\mathbf{x}) \forall \mathbf{x}$.

Recall that the rate factor A models internal structure properties of the ice. A depends on ice temperature, crystal orientation, debris content, etc. It may be represented by the Arrhenius law, see e.g. [17] Chapter 4 and references therein. This parameter A highly depends on the temperature, [17], therefore in particular on z in ice sheets.

Isothermal case. In the isothermal case, A is classically supposed to be a constant, see e.g. [17] Chapter 5. As a consequence, in this case we obtain: $\bar{A}(\mathbf{x}) = A = \underline{A}(\mathbf{x}) \forall \mathbf{x}$. It follows that: $R_A(\mathbf{x}) = 1 = c_A(\mathbf{x}) \forall \mathbf{x}$. Next for the classically employed value $q = 3$, it follows: $\gamma(\mathbf{x}) = (1 - \frac{1}{3}R_s(\mathbf{x}))$.

2.1.2 Domain of validity of the model

The shallowness of the flow is estimated through the geometrical ratio $\varepsilon = \frac{H^*}{L^*}$, where H^* and L^* are characteristic flow depth and length respectively. In these depth-integrated asymptotic models, ε has to be small enough, $\varepsilon \lesssim 1/10$ at least, see e.g. [29]. As a consequence this flow model is valid for a minimal wave length $L^* \gtrsim 10H^*$. The flow regime is characterised by the slip ratio R_s . By construction, the SIA-like models (including xSIA and RU-SIA equations) are valid for R_s ranging from ≈ 0.3 to 1, see [23,47,6] for detailed analysis. This estimation in terms of R_s is numerically quantified in real world cases (including EAIS) in [54]. This study is based on the so-called MCL criteria (criteria proposed in [24] and defined as the length scale over which the terms of driving stress and drag are comparable). In particular it can be noticed that the ice-sheet areas 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. The six test areas Antp considered in the next sections, see Fig. 4, have been defined from the surface velocities norm values: $|\mathbf{u}_H| \in \approx [10 - 80]$ m/y. In these cases $H^* \approx 2 - 3$ km, then the RU-SIA equation is accurate for minimal wave lengths $L^* \approx 20 - 30$ km.

2.1.3 Relationship with the mass equation & its inversion in fast plug-like flows

As already mentioned, nice bed elevation estimations are obtained in fast streams by simply inverting the depth-averaged mass equation $div(h\bar{\mathbf{u}}) = \dot{a}$, $\bar{\mathbf{u}}$ the depth-averaged velocity, see [42]. In [46,41,42], $\bar{\mathbf{u}}$ is related to \mathbf{u}_H as $\bar{\mathbf{u}} = \tilde{\alpha}\mathbf{u}_H$, with $\tilde{\alpha}$ set empirically. In fast streams (actually plug like flows), we have $R_s \sim 0$ and $\tilde{\alpha} \lesssim 1$. Therefore, for such cases the uncertainty on the internal deformation (represented in the RU-SIA equation by the parameter c_A) is negligible. As a consequence $\tilde{\alpha}$ may be set close to 1 with a few percent error only. This is what is done in these aforementioned studies. In other respect, we can show that: $\bar{\mathbf{u}} = -\frac{|\mathbf{u}_H|}{\mathcal{S}}\gamma\nabla H$, see [39]. Therefore if the slopes $\mathcal{S} = |\nabla H|$ and the velocity are co-linear (this is a commonly admitted assumption) than the parameter $\tilde{\alpha}$ empirically defined in [46,41] is nothing else than the dimensionless parameter γ defined by (2).

This equality $\gamma = \tilde{\alpha}$ enables building up continuous estimations between fast plug-flows (obtained by inverting the mass equation like in [41,42]) and the moderately sheared / moderately sliding flows (obtained by inverting the present RU-SIA equation).

On the contrary to plug like flows, in moderately sheared / moderately sliding flows, γ varies importantly therefore setting its value empirically is not reasonable anymore. That is why an actual estimation of γ is required.

2.2 On the uncertainty range of parameter γ

The dimensionless parameter γ defined by (2) depends on various physics parameters: the constitutive law exponent q , the vertical temperature profile through the rate factor $A(z)$ and the flow regime (slip ratio R_s). Recall that the RU-SIA model domain of validity corresponds to $R_s \in [\approx 1./3., 1]$, see [23,47,6]. For sake of simplicity, q is supposed to be set to the widely employed value for glaciers flows, that is $q = 3$ (Glen's law). In isothermal cases, it follows that: $\gamma(\mathbf{x}) = (1 - 0.2R_s(\mathbf{x}))$. Therefore in isothermal cases, the uncertainty on γ is relatively small, $\approx 10\%$ only. The large majority of glaciers are not isothermal in particular those in ice-sheets. Following the Arrhenius law, see e.g. [17, p.54], and by considering typical ice-sheet vertical temperature profiles in ice-sheets, see [44,45] and e.g. [27,49], we consider the following vertical profile of $A(\mathbf{x}, z)$, Fig. 2 (Left):

$$A(\cdot, z) = \begin{cases} A_a & \text{for } z \in [B(\cdot), H(\cdot)] \\ \frac{A_a}{(B-b)(\cdot, z)} ((1-k)z + kB(\cdot) - b(\cdot)) & \text{for } z \in [b(\cdot), B(\cdot)] \end{cases} \quad (6)$$

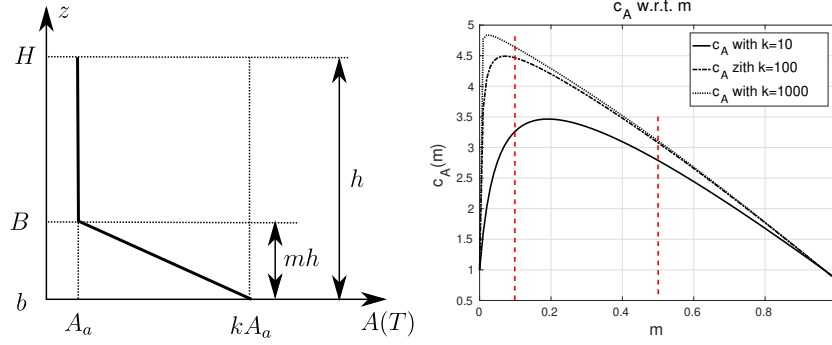


Fig. 2 (Left) A typical vertical profile of rate factor $A(\cdot, z)$, see (6). (Right) The parameter c_A vs m , see (3), with $k = 10, 100$ and 1000 .

226 with A_a and k given constants. We define: $B(\mathbf{x}) = (mh(\mathbf{x}) + b(\mathbf{x})) \forall \mathbf{x}$ with $m \in [0, 1]$. Then $A(\mathbf{x}, z)$ presents a boundary
 227 layer at bottom of thickness $(B - b)(\mathbf{x}) = mh(\mathbf{x})$, Fig. 2 (Left). The value of $c_A(\mathbf{x})$ vs m for different values of k is pre-
 228 sented in Fig. 2 (Right). The case $m = 0$ corresponds to the isothermal case: $c_A = 1$. For thin thermal boundary layers c_A
 229 increases with mh ; for thicker layers c_A decreases to a minimal value $c_A^{(min)} \lesssim 1$. (This minimal value is reached for the
 230 purely linear vertical profile: $m = 1$).

231
 232 Let us consider typical temperature values in EAIS: the bed is at $0C^\circ$ and the surface at $-40C^\circ$. These values
 233 correspond to $A_a \approx 10^{-26}$ therefore $k \approx 1000$, see Fig. 2 (Right). (The value $k \approx 10$ would correspond to typical inland
 234 Greenland cases, see [39]). Assuming a boundary layer corresponding to $m \in [0.1, 0.5]$, it follows: $c_A \in (3.11, 4.64)$ (Fig.
 235 2). Finally it follows from (2) that: $\gamma(\mathbf{x}) \approx [1 - (0.78 \pm 0.15)R_s(\mathbf{x})]$.

236 This rough uncertainty analysis based on typical values in the targeted areas shows that the uncertainty on γ comes
 237 similarly from the vertical thermal profile uncertainty (represented by the term $(c_A/(q+2))$) and the slip ratio R_s . In the
 238 targeted regimes and EAIS areas with the vertical profile (6), this corresponds to γ varying within the interval $]0, \approx 0.7]$.
 239 If relaxing the assumption on the vertical profile as defined in (6), one may estimate the upper bound of γ by setting
 240 $c_A = 1$ and $R_s = 0.5$ which gives: $\gamma \in]0, \approx 0.9]$. (Recall that in fast plug-like flows, γ is close to 1).
 241 In the forthcoming numerical results, the estimations of γ by the NNRK algorithm are within intervals $]0, \approx 0.9]$, see e.g.
 242 figs. 6 and 9(Up)(Right).

233 2.3 The inversion method

244 The inversion method to estimate the ice thickness h is developed in three steps. Step 1) and Step 3) are physical-based
 245 inversions: the RU-SIA equation (1) is inverted with respect to the product (γh) at Step 1), with respect to (h, \dot{a}) at Step
 246 3). Step 2) is based on an Artificial Neural Network (ANN) aiming at estimating γ ; it is a purely data-driven inversion.

247 2.3.1 Sketch of the complete inversion method

248 The estimations of thickness are performed in three steps as follows.

Step 1) Estimation of the effective diffusivity $\eta = (\gamma h)$ in RU-SIA equation (1) by VDA.

250 Given the surface measurements (observations) H^{obs} and $|\mathbf{u}_H^{obs}|$, the effective diffusivity $\eta = (\gamma h)$ in (1) is inferred by
 251 solving the following optimal control problem:

$$\min_{\mathbf{k}} g(\mathbf{k}) \quad \text{with } g(\mathbf{k}) = g_{obs}(\mathbf{k}) + \alpha g_{reg}(\mathbf{k}) \quad (7)$$

252 with $\mathbf{k} \equiv \eta = (\gamma h)$ and

$$g_{obs}(\cdot) = \frac{1}{2} \int_{\Omega} |H(\cdot)(\mathbf{x}) - H^{obs}(\mathbf{x})|^2 \chi_{tr}(\mathbf{x}) d\mathbf{x}, \quad (8)$$

253 χ_{tr} is the spatial restriction operator to the flight tracks, $g_{reg}(\cdot)$ a Tykhonov's regularization term, see e.g. [28]. In
 254 this step, it is defined as: $g_{reg}(\eta) = \frac{1}{2} \int_{\Omega} |\nabla \eta(\mathbf{x})|^2 d\mathbf{x}$. The weight coefficient α is classically set such that it provides
 255 a reasonable balance between the physical misfit and the regularization (regularization should be greatly lower than
 256 the physical misfit term). The surface elevation $H(\eta)$ corresponds to the solution of the RU-SIA equation (1) (with
 257 Dirichlet boundary conditions) with η given. The gradient of the cost functional is computed by introducing the ad-
 258 joint equation. The minimisation algorithm is a quasi-Newton method (the L-BFGS algorithm of the Python routine
 259 `scipy.optimize.minimize`). The iterative minimization process is performed until convergence. Numerous numerical

experiments have demonstrated robust convergences. In particular the optimal solution does not significantly depend on the smoothing length scale of the surface data (done in the present $Ant - p$ test cases at ≈ 24 km, see next Section), nor on the first guess (set to h_b the Bedmap2 value in EAIS).

Past this computational VDA step providing the optimal value η^* , the value of γ along the flights tracks where depth measurements h_b are available are straightforwardly deduced: $\gamma_{tr}^* = \frac{\eta^*}{h_b} \chi_{tr}(\mathbf{x})$. These values are inputs of the next algorithm, Step 2).

These obtained values are representative at the flow model scale, that is at $10\bar{h} \approx 25$ km minimal wave length (with ≈ 2 km mesh cells), see the investigation presented in [39].

Step 2) Extension of γ in the whole domain by NNRK.

Given γ_{tr}^* along the flights tracks (result of Step 1)), a NNRK algorithm ([11, 30]) is applied to extend values of γ to the whole area. This statistical learning algorithm is done in two steps: 1) an ANN estimator (deep learning) is built up; 2) an ordinary Kriging of the residuals is added. Details are presented in the next paragraph.

Step 3) Estimation of the pair (h, \dot{a}) in RU-SIA equation (1) by VDA.

Given γ all over the domain (result of Step 2)), the thickness h is inferred simultaneously with the RHS \dot{a} in (1) by another VDA process. Let us recall them briefly.

Similarly to Step 1), the pair (h, \dot{a}) in (1) is inferred by solving the optimal control problem (7) with g_{obs} defined by (8) but minimizing with respect to $\mathbf{k} = (h, \dot{a})$ (and not w.r.t. $\eta = \gamma h$ like in Step 1)). In this VDA process, the regularization term reads:

$$g_{reg}(h, \dot{a}) = \frac{1}{2} \|(h - h_b)\|_{C_h^{-1}} + \frac{1}{2} \|\dot{a} - \dot{a}_b\|_{C_a^{-1}} \quad (9)$$

with C_h^{-1} and C_a^{-1} covariance operators defining metrics, (h_b, \dot{a}_b) prior background values (equal to the current classical estimations). The latter are classically defined as the second order auto-regressive correlation matrices with length scale respecting a balance between the regularisation and the preconditioning effects of the VDA algorithm, see [20, 39]. Next following [33, 20], a change of the control variable is made. The numerous numerical experiments have demonstrated that this choice of covariance operators combined with the change of variable improves greatly the robustness and the convergence speed of the VDA algorithm. In (7) the weight coefficient α is defined as a decreasing sequence following an iterative regularisation strategy, see [28] for an analysis. This iterative regularisation strategy improves the convergence speed of the VDA algorithm too.

Numerous assessments of the VDA steps are presented in [39], in particular the sensitivity of the inversions with respect to: i) the uncertainties on γ ; ii) the density of flights tracks (by removing some of them); iii) the smoothing length scale of the surface data (altimetry, InSar) from ≈ 24 to 48 km; iv) the first guess (chosen here as the Bedmap 2 value h_b).

Remark 1 Using the explicit expression (2) of γ one can compute a-posteriori estimations of the (spatially distributed) slip ratio value R_s . This is an interesting feature to analyse a-posteriori the degree of the RU-SIA flow model consistency. The few a-posteriori analysis made (not presented here) remarkably confirms the good consistency of RU-SIA model.

Remark 2 Based on a-priori vertical thermal profile(s) (e.g. the one defined by (6)), the RU-SIA equation (1) provides a-posteriori estimations of the effective thermal boundary layer thickness $(B - b)$, see Fig. 2 (Left). Such a-posteriori estimations may be interesting for various analyses. Moreover the vertical profiles could be adjusted by constraining them with (the very few) in-situ measurements.

2.3.2 Details of Step 2): the Neural Network Residual Kriging (NNRK) algorithm

The employed NNRK algorithm is decomposed in three steps as follows.

Step 2a) Considering the surface data $(H, \mathcal{S}, |\mathbf{u}_H|, \dot{a}_b)$ at all in-situ measurements locations (e.g. along the flights tracks of all areas $Antp$) plus the values of γ_{tr}^* computed at Step 1), an estimator of γ is built up by training a ANN. This estimator is denoted by $\bar{\gamma}$.

The training dataset is denoted by \mathcal{D} ; it contains "examples" (I_i, O_i) , $i = 1, \dots, N_{ft}$, where $I_i = (H, \mathcal{S}, |\mathbf{u}_H|, \dot{a}_b)(\mathbf{x}_i)$ is the i -th input and $O_i = \gamma_{tr}^*(\mathbf{x}_i)$ is the corresponding output. \mathbf{x}_i , $i = 1, \dots, N_{ft}$ denote the i -th in-situ measurements coordinates (e.g. along the flight tracks).

The estimator $\bar{\gamma}$ is computed as the minimizer of the mean square misfit: $\frac{1}{N_{ft}} \sum_{i=1}^{N_{ft}} (O_i - \bar{\gamma}(I_i))^2$. This misfit function (also called loss function) reads:

$$j(\mathcal{D}; \cdot) = \frac{1}{N_{ft}} \sum_{i=1}^{N_{ft}} [\gamma_{tr}^*(\mathbf{x}_i) - \bar{\gamma}((H, \mathcal{S}, |\mathbf{u}_H|, \dot{a}_b)(\mathbf{x}_i; \cdot))]^2 \quad \text{for } \mathbf{x}_i \in \Gamma_{tr} \quad (10)$$

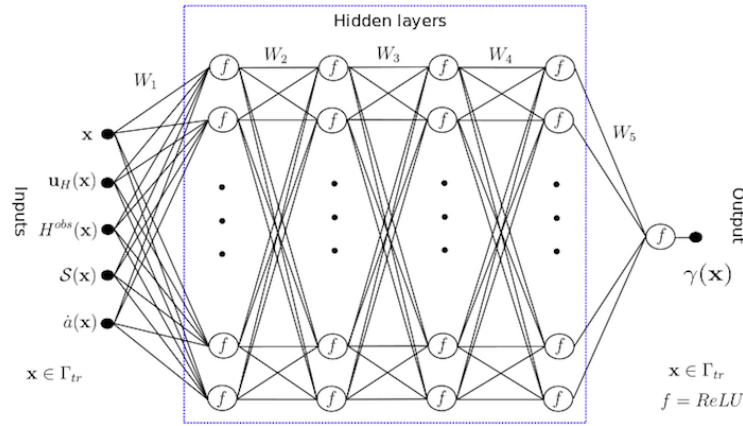


Fig. 3 The Artificial Neural Network (ANN) with four hidden layers. First step of the NNRK algorithm.

To solve this large dimensional data-based optimisation problem, the currently most efficient methods are ANN with few hidden layers (deep learning). Here, 5 layers are considered, see Fig. 3; each of the 4 hidden layers contains 50 neurons. The most efficient activation function is chosen: the rectified linear unit (ReLU) function, see e.g. [16, 32]. The ANN is determined by its architecture and the weight parameters (W_1, \dots, W_5) , Fig. 3.

The training step consists of identifying the optimal values of these parameters W_j , $j = 1, \dots, 5$. Each W_j is a matrix of dimension $n_{out} \times n_{in}$. Here, W_1 has $5 \times 50 = 250$ parameters, W_j for $j = 2, 3, 4$ have $50 \times 50 = 2500$ parameters each, W_5 has $50 \times 1 = 50$ parameters. The ANN has been coded in Python using the PyTorch and Mpi4Py libraries [10].

To minimize $j(\mathcal{D}; (W_1, \dots, W_5))$ with respect to $\{W_j\}_j$, the classical Adam method [31], a first-order gradient-based stochastic optimization, is employed. The learning rate (the gradient descent step size) is classically adjusted during the optimization procedure.

The input variables are heterogeneous and of different order of magnitude e.g. the elevation H (m) and the slope \mathcal{S} (radian). Therefore each input variable v , v an element of $\{H, \mathcal{S}, |\mathbf{u}_H|, \hat{a}_b\}$, are reduced centred as follows: $\bar{v}_i = (v_i - \text{mean}(v))/\sigma(v)$, for all i , $1 \leq i \leq N_{fi}$. The normalisation is applied to mini-batches in hidden layers; this technique is supposed to improve the stability of the model; see e.g. [30] for more details and know-hows on ANN and NNRK algorithms.

Also to avoid overfitting, the dropout method [50] is adopted. (This technique may help to prevent overfitting). As usual, the hyper-parameters of the algorithm (learning rate, decay rate, dropout probability) are experimentally chosen; the selected values are those providing the minimal value of j .

The ANN has been trained by including h_b as an input parameter or not. Both estimators (considering h_b as an input or not) turned out to have similar accuracies; thus confirming the strategy to predict the dimensionless parameter γ of the flow model from the surface data only.

Step 2b) The K-fold cross-validation method, see e.g. [1], is employed to assess the ANN accuracy and to confirm if the ANN can be used as a predictor. Let us recall that K-fold cross-validation method is as follows, see e.g. [1]:

- Divide randomly the original training data set \mathcal{D} into K (roughly equal) subsets;
- For each subset \mathcal{D}_k , $k = 1 \dots K$, the ANN is trained from the other $(K - 1)$ subsets \mathcal{D}_i , $i \neq k$.
We denote by $\mathcal{D}_{i_{est}} = \mathcal{D}_i$ and $\mathcal{D}_{i_{train}} = \cup_{j \neq i} \mathcal{D}_j$, $i = 1, \dots, K$.
- Compute the loss function $j(\mathcal{D}_{i_{est}})$ for each case.

Finally, choose the best ANN i.e. those providing the smallest total loss function $j(\mathcal{D}_{i_{est}}) + j(\mathcal{D}_{i_{train}})$.

Step 2c) The residual at the measurements locations is computed: $\varepsilon_\gamma = (\gamma_{tr}^* - \bar{\gamma})$ with $\bar{\gamma}$ computed by the (best) ANN.

Next an ordinary Kriging (with a spherical semi-variogram model) is used to extend ε_γ all over the domain. The obtained estimator is denoted by $\hat{\varepsilon}_\gamma$. By construction this residual satisfies: $\mathbb{E}(\varepsilon_\gamma) \approx 0$. Moreover the correlation between two points depends on the distance between them and not on their location. Performing an ordinary Kriging on the residual after ANN is known to be particularly efficient, see e.g. [30] Chapter 3.

The final estimator in the whole domain is denoted by $\hat{\gamma}$. It is obtained as the sum of the ANN estimator and the ordinary Kriging estimator of residuals:

$$\hat{\gamma}(\mathbf{x}) = \bar{\gamma}(\mathbf{x}) + \hat{\varepsilon}_\gamma(\mathbf{x}) \text{ for } \mathbf{x} \in \Omega \quad (11)$$

The forthcoming numerical results show that the estimator $\hat{\gamma}(\mathbf{x})$ provides (surprisingly) very accurate values of the parameter γ from the surface data only (altimetry, InSAR and \dot{a}).

2.3.3 On the linked uncertainty between γ and h

In the inversion algorithm previously described, after Step 1), one has to separate the effects of the two unknown fields: the physical-based dimensionless parameter γ and the ice thickness h . The accuracy and robustness of each VDA process are demonstrated by the numerical experiments presented later, see also [39]. It will be demonstrated in next section that the NNRK algorithm is robust and accurate too. Then it can be assumed that: $\gamma^t h^t \approx \gamma^* h^*$; where the superscript $*$ denotes the optimal computed values while the superscript t denotes the (effective) true value. Let us denote: $e_\phi = (\phi^* - \phi^t)/\phi^*$. At order 1, one has:

$$e_h \approx -e_\gamma \text{ (where } \gamma \text{ does not tend to } 0) \quad (12)$$

In other words, Step 2) and Step 3) of the inversion algorithm propagates the error made on γ to h in the same order of magnitudes (in %).

Remark 3 *It would be straightforward to apply the same NNRK algorithm to directly estimate the thickness h all over the domain. However it seems definitively more consistent to estimate a dimensionless parameter of a flow model able to represent accurately the surface data, than to estimate the thickness data partially responsible only of the employed surface data. Following this idea of purely data-driven estimations, [9] had proposed an ANN trained and assessed on synthetic data generated by an ice flow model and geomorphic premises to estimate the bedrock elevation of four mountain glaciers.*

3 Data pre-processing

In the next sections, the algorithm is applied to 6 large areas in EAIS (ranging from 250268 to 439045 km^2); they are denoted by $Antp$, $p = 1, \dots, 6$, see Fig. 4. These areas include the upstream areas of major ice-sheds; all of them respect the flow model domain of validity. The mean thickness value of Bedmap2 ([14]) ranges from 1822 to 2745 m, see tables 2-11 for details. The coordinates of the characteristic points defining each area ($\approx 100 - 150$ points per area, see Fig. 4) are available on the open source computational software DassFlow webpage².

Estimating more accurately the bed elevation in these areas may be interesting since global warming may threaten EAIS stability as mentioned e.g. in [13].

The correlations between the given variables (H^{obs} , $\|\mathbf{u}_H\|$, \mathcal{S} , \dot{a} , h_b) for all areas have been computed. It turns out that no significant linear correlation between the variables have been observed.

In this section, the method to smooth the surface data accordingly with the flow model domain of validity is presented; the definition of adequate numerical grids follows. Moreover since Bedmap2 values are considered as the reference values, the method to obtain these values is recalled.

3.1 Minimal wave length, surface data smoothing and numerical grids

The surface data $|\mathbf{u}_H|$ and H have to be defined at an adequate scale to be consistent with the shallow (long wave assumption) flow model (1); next providing a minimal wave length L^* (km) of the inversions. The RU-SIA equation is accurate as soon as $\varepsilon = \frac{[H]}{[L]} \lesssim 0.1$. (In other respect it is shown in [54] that the ice-sheet areas 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).

The mean value of the Bedmap2 ice thickness (denoted by \bar{h}_b) [14] in the 6 $Antp$ areas equals ≈ 2.7 km. In the considered regimes, the velocity field is co-linear to the slopes \mathcal{S} ; therefore to be fully consistent, the smoothing of surface data should be done non-isotropically by defining a streamline minimal wave length and a cross-line one. For a sake of simplicity, here an isotropic smoothing is performed. To do so, a Gaussian with standard deviation $\sigma = 4$ km is convoluted with each given surface field: elevation and velocity norm. Then the smoothing effects are sensitive in disks of diameter $\approx 2 \times (3\sigma) \approx 10\bar{h}_b$ km.

² Open-source computational software DassFlow: Data Assimilation for Free Surface Flows. Python version for 2D shallow generalised Newtonian fluids. INSA, University of Toulouse, CNES. www.math.univ-toulouse.fr/DassFlow.

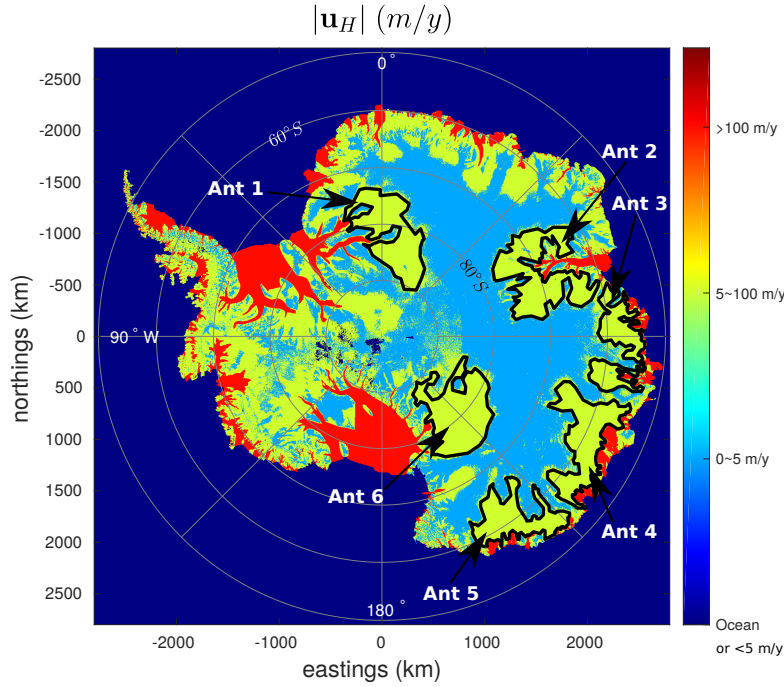


Fig. 4 Location of the 6 test areas Ant_p (east Antarctica) with InSAR-based surface velocity values in m/y (from [43]).

393 Next a finite element mesh is built up using Gmsh software [15] with a grid size $\delta x \approx 3$ km. Indeed $\delta x = 3$ km
 394 provides ≈ 10 points per minimal wave length L^* , therefore respecting the minimal number of points to properly approx-
 395 imate all fields. The flight tracks are meshed with cells of $\delta x \approx 2$ km. The given thickness measurements (provided in the
 396 Bedmap2 database [14]) are interpolated along the flight tracks .
 397

398 3.2 Recalls of the origins of Bedmap2 values

399 Bedmap2 values h_b are considered as reference values; moreover they are employed to set the first guesses of the VDA
 400 processes at Step 1) and Step 3) (see paragraph 2.3.2). h_b and its a-priori uncertainty as derived in [14] are plotted for each
 401 test area Ant_p in figures 6, 9, 12, 13, 14 and 15 (Middle)(Left). In Bedmap2 database [14], the interpolation - extrapola-
 402 tion of airborne measurements are performed throughout the domain by the ArcGIS Topogrid routine (ESRI Ltd, ArcGIS
 403 9); the latter is based on the ANUDEM algorithm [26]. This algorithm uses an iterative finite difference interpolation
 404 technique which is essentially a thin plate spline technique [53]. Next, empirical uncertainty values are stated as follows,
 405 see [14]. The thickness measurements are split into two datasets ($D1$) and ($D2$). Dataset ($D1$) is used to build up an inter-
 406 polation including at Dataset ($D2$) location points; values of ($D2$) being not used at this stage. Next, Dataset ($D2$) is used
 407 to quantify the misfit with the "predicted" - interpolated values; and basic statistics on the results with dependence on the
 408 distance to data are deduced. For cells located between 5 and 20km from any data, [14] suggests that the interpolation
 409 error is an increasing function of distance from the closest data; beyond it would be not correlated. (Observe that this
 410 distance corresponds approximatively to the minimal wave length of the RU-SIA model). For cells that are more than 50
 411 km from airborne measurements, the thickness estimation is based on gravity-field inversion (gravity-derived thickness);
 412 the proposed related uncertainty equals $\pm 1000m$. This is how the uncertainty values on Bedmap2 values h_b are defined;
 413 see figures 6, 9, 12, 13, 14 and 15 (Middle)(Right).
 414

415 4 Analysis of each inversion algorithm step

416 4.1 Step 1: Estimation of η by VDA

The effective diffusivity η defined in (1) is estimated in each area Ant_p by VDA following Step 1) described in Paragraph 2.3.1. The convergence of this iterative VDA process is very slow (a few hundreds of iterations) but very robust in particular with respect to the first guess value. This step has been thoroughly assessed in [39]. The stopping criteria is the

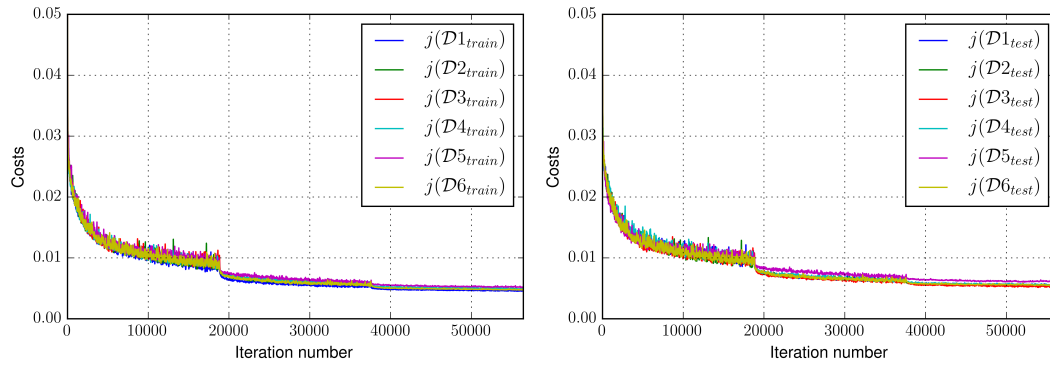


Fig. 5 Misfit functional $j(\mathcal{D})$ defined by (10), vs minimization iterations for: (Left) different train sets $\mathcal{D}_{i_{train}}$; (Right) different test sets $\mathcal{D}_{i_{test}}$. The learning rate (gradient descent step size) is decayed by 0.2 every 30 epochs (1 epoch = 1 forward pass and 1 backward pass of all the training examples), see e.g. [30] for details on this classical method.

stationarity of $\|\eta\|$. Here the RHS \dot{a} provided by [52] is supposed to be exact. After convergence of the VDA process, given the ice thickness along the flights tracks, the computed optimal value γ^* is saved for Step 2), that is:

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

4.2 Step 2: Estimation of γ by NNRK

The ANN algorithm input values, see Fig. 3, are datasets along all flight tracks in all test areas Antp, $p = 1, \dots, 6$ plus the values of γ_{ir}^* computed at Step 1). This dataset is denoted by \mathcal{D} . Following Step 2b) (see paragraph 2.3.2), the K-fold cross-validations are performed with $K = 6$. The results are presented in Fig. 5. (The value $K = 6$ is completely independent of the total test cases number). Here, the training sets $\mathcal{D}_{i_{train}}$ contains 16774 examples and a test set $\mathcal{D}_{i_{test}}$ contains 3354 examples; each example being an (input, output) pair of ANN.

It can be read on Fig. 5 (Left) that all ANN models trained from the different data sets $\mathcal{D}_{i_{train}}$ provide very close cost values $j(\mathcal{D}_{i_{train}})$ (see (10)). Moreover, the cost values of all test sets $j(\mathcal{D}_{i_{test}})$ are almost all equal. This shows that all ANN have very close prediction capability, all being excellent. Indeed after optimisation, $j(\mathcal{D}_{i_{test}}) \approx 5 \cdot 10^{-3}$ (Fig. 5 (Right)); this corresponding to $\approx 1\%$ of the mean value of $\bar{\gamma}$. The ANN trained from $\mathcal{D}_{2_{train}}$ is selected since having a slightly smaller misfit value ($j(\mathcal{D}_{1_{train}}) + j(\mathcal{D}_{1_{test}})$). These tests of prediction capabilities demonstrate the robustness and accuracy of the trained ANN.

Next, values of $\bar{\gamma}$ are predicted in the whole domains Ω_p , $p = 1, \dots, 6$ by performing the selected ANN.

Next by performing the Kriging step (Step 2c) described in Paragraph 2.3.2), the predictor $\hat{\gamma}$ is obtained.

Following the a-priori estimation derived in Section 2.2, the upper bound 0.9 is imposed to the estimation; this upper bound is active at very few locations only; moreover it is in great majority where the uncertainty on h_b is low.

For each test area, the predicted values of γ are plotted; see figures 6, 9, 12, 13, 14 and 15 (Up) (R).

4.3 Step 3: Estimation of h (and adjustment of \dot{a}) by VDA

After Step 2), the dimensionless parameter γ in RU-SIA equation (1) is given. Then the thickness h is inferred by VDA following the method presented in paragraph 2.3.1. The optimisation algorithm converges in $\sim 20 - 50$ iterations depending on the test case. Its convergence is robust; this point has been thoroughly analysed in [39].

At this step, h is simultaneously inferred with \dot{a} . Indeed this enables to adjust the value of \dot{a} within its uncertainty range which is here $\sim \pm 20\%$, see [52]. It can be noticed in tables 2, 5, 8, 9, 10 and 11 that the corrections made remains in great part within this a-priori uncertainty range. Indeed the upper bound is active at few locations only. In other words, this adjustment based on the physical-based model RU-SIA is consistent with the (totally independent) estimations derived in [52].

Of course, given the surface data, any change of \dot{a} in RU-SIA equation (1) has an effect on the inferred value of h . As an example let us compute the response H of RU-SIA equation (1) in Ant1 area with: i) the RACMO2 value \dot{a}_b in the RHS (providing $H(\dot{a}_b)$); ii) the inferred value \dot{a}^* by VDA in the RHS (providing $H(\dot{a}^*)$). The obtained difference are the ones indicated in Tab. 1. Therefore the correction made on the RHS $\dot{a} = (\partial_t h - a)$ implies a correction on the ice thickness h negligible ($\sim 1\%$) compared to the one obtained by the complete inversion method, see Tab. 2 in next section. This

Table 1 Difference between values of H the response of RU-SIA equation (1) if changing the RHS value: \hat{a}_b the Racmo2 value, \hat{a}^* the inferred value at Step 3). Domain $\Omega_p = \text{Ant1}$.

Difference of H	Median	Mean	Max
$ H(\hat{a}_b) - H(\hat{a}^*) $	3.3m	6.0m	26.3m

449 remark holds for all the domains Ω_p .

450

451 4.4 On the RU-SIA model accuracy

452 For each test area Ant_p , domain information and basic statistics on the results are presented, see tables 2, 8, 5, 9, 10, 11.
 453 Statistics on the computed surface elevation H , output of RU-SIA model, are indicated. It can be noticed that the RU-SIA
 454 equation solved from Bedmap2 value h_b and the (purely data-driven) estimation of γ obtained at Step 2) already fits very
 455 well the measured surface elevation H^{obs} , see "Direct model validation" lines in the tables. This very good accuracy
 456 (based on the Bedmap2 bed elevation h_b i.e. without any additional calibration of h) demonstrates the validity and the
 457 relevance of the RU-SIA model. After Step 3) of the inversion algorithm i.e. after the identification of h and \hat{a} by VDA,
 458 of course the RU-SIA model fits even better the measured surface elevation H^{obs} , see " $|H(h^*) - H^{obs}|$ (after h -inversion)"
 459 in the tables.

460

461 5 Results and sensitivity tests (Ant1 and Ant3 areas)

462 In this section, the bed elevation b (equivalently the ice thickness h) is inferred by the inverse method described in Section
 463 2.3 for the two areas Ant1 and Ant3, see Fig. 4, tables 2 and 5. Different estimations of h are compared, depending if:

- 464 a) isolated flight tracks (hence locally highly constraining) are considered or not;
 465 b) the learning method at Step 2) is the NNRK algorithm described in Paragraph 2.3.2, or an ordinary Kriging method
 466 like it is done in [39].

467 These comparisons aim at analysing the robustness and accuracy of the present inverse method.

468 Ant1 is a 370809 km² area north-east upstream of Bailey, Slessor and Recovery ice-streams; Ant3 is a 250268 km² area
 469 in Wilhelm and Queen Mary lands, upstream of Shackleton ice shelf and Davis sea.

470 Among the considered six areas, Ant1 and Ant3 are those presenting the largest uncovered parts during airborne cam-
 471 paigns. As a consequence they contain large areas where Bedmap2 estimation h_b is based on gravity field inversions,
 472 therefore presenting very large uncertainties.

473 For each case, the domain information and basic statistics on the numerical results are presented in tables 2 and 5. For
 474 each case, the most relevant fields are plotted, see figures 6 and 9: the surface velocity norm $|\mathbf{u}_H|$ and the flight tracks lo-
 475 cations (Up)(L), the NNRK estimation $\hat{\gamma}$ defined by (11) (Up) (R), the Bedmap2 value h_b (Middle)(L) with its uncertainty
 476 as presented in [14] (Middle)(R), the present thickness estimation h^* (Down)(L) and its difference with h_b (Down)(R).

477 5.1 Results for Ant1 area

478 This domain presents large unexplored areas during the airborne campaigns therefore huge uncertainty on h_b values, see
 479 Fig. 6 (Middle).

480 5.1.1 The ice thickness estimation h^*

481 Recall that $\hat{\gamma}$ is the NNRK estimation of the dimensionless parameter γ defined by (2). No correlation is observed between
 482 h and γ ; the only clearly observed correlation is : γ is small where $|\mathbf{u}_H|$ is small, see Fig. 6 (Up). This observation is fully
 483 consistent with the a-priori analysis done in Section 2.2, see (2) and Fig. 2.

484 Recall that h_b values are thin plate spline based estimations (see [14] and paragraph 3.1) hence intrinsically smooth, Fig. 6
 485 (Middle)(Left). The present estimation h^* is the optimal solution of a data-driven model combined with a physical-based
 486 model; no additional smoothing has been added, then the obtained value h^* is much less smooth than h_b . It is worth to
 487 notice that a non isotropic smoothing of the surface data (see the previous section) would have provided bed elevation
 488 patterns more correlated to the surface streamlines.

489 The difference between h^* and h_b is not correlated to the distance to the nearest measurement (flight track), on contrary to
 490 the empirically stated uncertainty in [14]. Indeed large corrections of h_b (up to 1500 m) are obtained close to flight tracks;

Table 2 Domain Ω_p =Ant1, information and results.

Domain Ω_p & mesh information				
Surface $ \Omega_p $	370809 km^2			
Mean ice thickness of h_b (Bedmap2)	2696.2 m			
# mesh vertices: in Ω_p / on flight tracks	57661 / 2152			
RU-SIA model output (with $\hat{\gamma}$)		Median	Mean	Max
$ H(h_b) - H^{obs} $ (before h -inversion)		7.0 m	10.3 m	53.0 m
$ H(h^*) - H^{obs} $ (after h -inversion)		2.6 m	3.7 m	44.6 m
Inferred RHS \dot{a}				
$ \dot{a}^* - \dot{a}_b $		0.7 cm/y	0.7 cm/y	3.6 cm/y
$ \dot{a}^* - \dot{a}_b / \dot{a}_b $		19.2 %	17.0 %	20.0 %
Inferred thickness h				
$ h^* - h_b $		275.2 m	356.8 m	1953.5 m
$ h^* - h_b / h_b $		10.4 %	13.3 %	65.7 %
Ice volume change in km^3 / in %		1.6 $10^4 km^3$ / 1.6 %		

Table 3 Domain Ω_p =Ant1. Comparison of the estimations if considering or not the flights tracks indicated in Fig.6 (Up)(L).

Inferred thickness difference	Median	Mean	Max
$ h^*(\Gamma_{tr}^{all}) - h^*(\Gamma_{tr}^{less}) $	151.3 m	196.9 m	1524.5 m
$ h^*(\Gamma_{tr}^{all}) - h^*(\Gamma_{tr}^{less}) / h^*(\Gamma_{tr}^{less}) $	5.6 %	6.6 %	80.0 %

close meaning at 1 – 2 minimal wave lengths of the RU-SIA model that is $\sim 20 - 40$ km, see e.g. in Fig. 6 (Down)(R) the areas around coordinates (2750,1950)(2900,1550)(3050,2050). At the opposite, h^* may remain very close to h_b in areas relatively far from any flight tracks.

Recall that the flight tracks are meshed as 1d lines; along these segments (≈ 2 km long), the inferred depth value can vary of ± 150 m around the measured values (inequality constraints imposed in the VDA processes). Therefore in the adjacent triangles (which are nearly equilateral with ≈ 2 km edges), the plotted mean depth values may be already much different than the measured ones. This large scale smoothing may explain the potential great differences between the two estimations even relatively close to the flight tracks.

The basic statistics presented in Tab. 2 show that after the VDA processes, the RU-SIA equation fits extremely well the measured surface elevation. The correction made on \dot{a} is relatively consequent, 17% in mean (it is the highest correction made among the 6 test cases Ant p). However it remains lower than the maximal authorised correction: $\pm 20\%$.

Finally the correction made on h_b is non negligible: 13.3% (356.8 m) in mean, with a 1.6% ($1.6 \cdot 10^4 km^3$) of volume change only (for 370809 km^2).

5.1.2 Ant1: if removing some flight tracks

In this paragraph, a new ice thickness estimation is computed. It differs from h^* since the flight tracks indicated in Fig.6 (Up)(L) are not considered anymore. The original complete set of flight tracks is denoted by Γ^{all} ; the partial one is denoted by Γ^{less} . In Γ^{less} case, the mesh of the entire area has been re-built (since in each mesh, the lineaic flight tracks are meshed differently). The inverse problem based on Γ^{less} is less constrained in particular the two VDA processes (Steps 1 and 3). Indeed the removed flight tracks are isolated, see Fig. 6(Up)(L); no constraint is imposed in the vicinity of these removed flight tracks anymore.

The statistical learning at Step 2) is unchanged, therefore values of $\bar{\gamma}$ are unchanged too. However the estimation $\hat{\gamma}$ defined by (11) is not the same since the Kriging step is changed. Indeed the latter is based on less flight tracks data. The difference between the two estimations ($\bar{\gamma}(\Gamma^{all}) - \bar{\gamma}(\Gamma^{less})$) is plotted in Fig. 7 (Up)(R). It can be noticed that $\hat{\gamma}$ is changed all over the domain and not particularly in the vicinity of the missing flight track. Indeed, the Kriging method (Step 2c) in Paragraph 2.3.2) aims at computing the minimal variance in norm 2 (least square) and not point-wise; hence the global change of $\hat{\gamma}$.

Next the inferred thickness h^* is different for two reasons: 1) values of γ are different; 2) the VDA process of Step 3) is not locally constrained at the missing flight tracks locations. The difference between the two thickness estimations is plotted in Fig. 7 (Down)(R). For a sake of readability, the legend in Fig. 7(Down)(R) has been bounded at ± 400 m (very few locations were greater than this bound). Basic statistics on the difference are presented in Tab. 3. Differences of 300 m correspond to $\approx 10 - 15\%$ of change. As expected, see (12), the variations of h are correlated to the variations of γ : compare Fig. 7 (Up)(R) to (Down)(R). Since the global change of γ , the change of h is not particularly important around the missing flight tracks.

Finally, it is worth to mention that the present inversion method is relatively global with local constraints ("in-situ" measurements along the flight tracks); it is not purely local inversions. In the present experiment, the obtained variations of h^* are roughly half of the ones obtained from h_b , see tables 2 and 3: difference of 6.6% in mean vs 13.3%, and 5.6% vs 10.4% for median values.

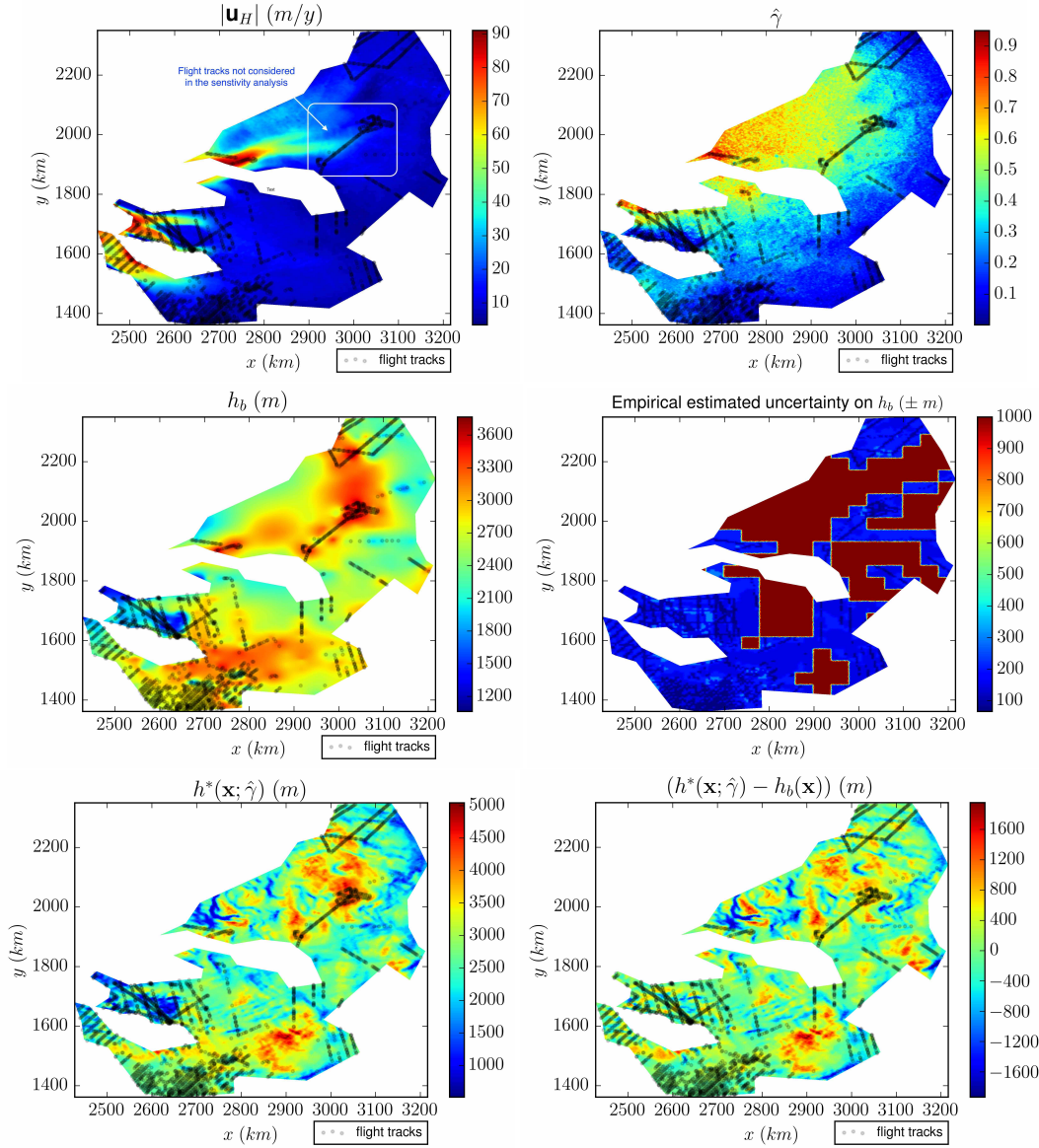


Fig. 6 Domain Ω_p =Ant1 (the plotted coordinates equal the Eastings-Northings plus (2800, 2800)km): (Up)(L) Surface velocity module $|u_H|$ and flight tracks (R) $\hat{\gamma}$ computed by NNRK, see (11). (Middle)(L) Thickness h_b from Bedmap2 [14] (R) Empirical uncertainty on h_b from [14]. (Down)(L) Inferred thickness with $\hat{\gamma}$: $h^*(\hat{\gamma})$ (R) Difference $(h^*(\hat{\gamma}) - h_b)$.

5.1.3 Ant1: with a different statistical learning method at Step 2)

It has been previously shown that uncertainties on γ generates uncertainties on h of same order of magnitude, see (12). In this paragraph, the influence of the statistical estimator considered at Step 2) is investigated. To do so, first we compare:

- $\bar{\gamma}$ obtained by ANN algorithm at Step 2a) to $\hat{\gamma}$ obtained by the complete NNRK algorithm (see (11));
- γ_{krig} obtained by an ordinary Kriging of values inferred along the flights tracks to $\hat{\gamma}$.

It may be a-priori guessed that a deep learning method (as the present NNRK algorithm) is more reliable than a Kriging inter/extrapolation of its values between the flight tracks.

Second, we compare the inferred thickness h^* obtained from $\hat{\gamma}$ (that is the original estimation plotted in Fig. 6 (Down)(L)) to the one obtained from γ_{krig} .

As expected, the difference between $\bar{\gamma}$ and $\hat{\gamma}$ (i.e. before and after the Kriging Step 2c) are localised in the vicinity of the flight tracks. In Ant1 case, these differences may be up to $\sim \pm 20\%$, see Fig. 8(Up)(R). More interestingly and as expected too, the differences between $\hat{\gamma}$ and γ_{krig} are not clearly correlated to the distance to the nearest flight track. The observed difference in Ant1 case may be consequent: $\sim \pm 40\%$, see Fig. 8(Middle)(R).

Next, like in the previous sensitivity test (and for the same reasons), the variations of h are correlated to the variations of γ : compare Fig. 8 (Middle)(R) to (Down)(R). Some statistics on the differences are presented in Tab. 4. Again the

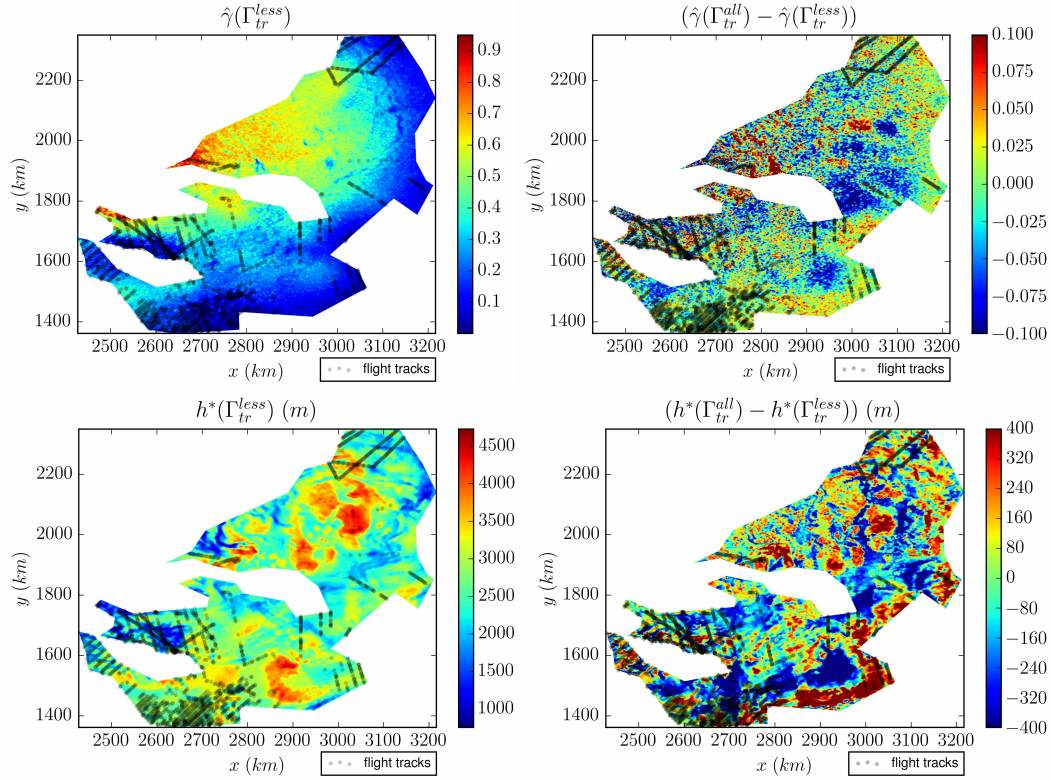


Fig. 7 Domain Ω_p =Ant1: comparison if not considering the flights tracks indicated in Fig.6 (Up)(L). (Up)(L) Field $\hat{\gamma}(\Gamma_{tr}^{less})$ (i.e. without the flights tracks indicated in Fig.6 (Up)(L)). (R) Difference $(\hat{\gamma}(\Gamma_{tr}^{all}) - \hat{\gamma}(\Gamma_{tr}^{less}))$. (Down)(L) Inferred thickness $h^*(\Gamma_{tr}^{less})$ (R) Difference between the two estimations: $(h^*(\Gamma_{tr}^{all}) - h^*(\Gamma_{tr}^{less}))$.

Table 4 Domain Ω_p =Ant1. Comparison the original thickness estimation (obtained using NNRK) to the one obtained using ordinary Kriging at Step2)

Inferred thickness difference	Median	Mean	Max
$ h^*(\hat{\gamma}) - h^*(\gamma_{krig}) $	145.5 m	183.5 m	1264.2 m
$ h^*(\hat{\gamma}) - h^*(\gamma_{krig}) / h^*(\hat{\gamma}) $	5.3%	7.2%	117.2%

544 obtained variations in h are roughly half of the ones obtained from h_b , see tables 2 and 4: difference of 7.2% in mean vs
 545 13.3%, and 5.3% vs 10.4% for median values.

546

547 5.2 Results for Ant3 area

548 Like Ant1, Ant3 area presents large uncovered areas during the airborne campaigns, corresponding to huge uncertainty
 549 on h_b , see Fig. 9.

550 5.2.1 The ice thickness estimation h^*

551 Like in Ant1 case, the only observed correlation is: γ is small if \mathbf{u}_H is small, see Fig. 9 (Up). Again the difference between
 552 h^* and h_b is uncorrelated to the distance to the nearest flight track (on contrary to the empirically established uncertainty
 553 in [14]). Large corrections of h_b (up to 700 m) are found close to flight tracks see e.g. the area around coordinates
 554 (5550,2800) in Fig. 6 (Down)(R); at the opposite, h^* may remain very close to h_b in areas relatively far from any flight
 555 tracks, see e.g. the area around coordinates (5250,2800).

556 The few statistics presented in Tab. 5 show that again after the VDA processes, the RU-SIA equation fits extremely
 557 well the measured surface elevation. The correction made on \hat{a} is much lower than the authorised maximal variation:
 558 11.2% in mean. In Ant3, the global correction made on h_b is relatively low: 6.6% in mean (3.5% median) with a 0.5%
 559 of volume change only. However in the most uncertain areas, the corrections made can be either low, see e.g. the areas
 560 around coordinates (5000,2650) (5200,3250), or important ($\pm \sim 700$ m), see e.g. the area around coordinates (5000,2650).

561

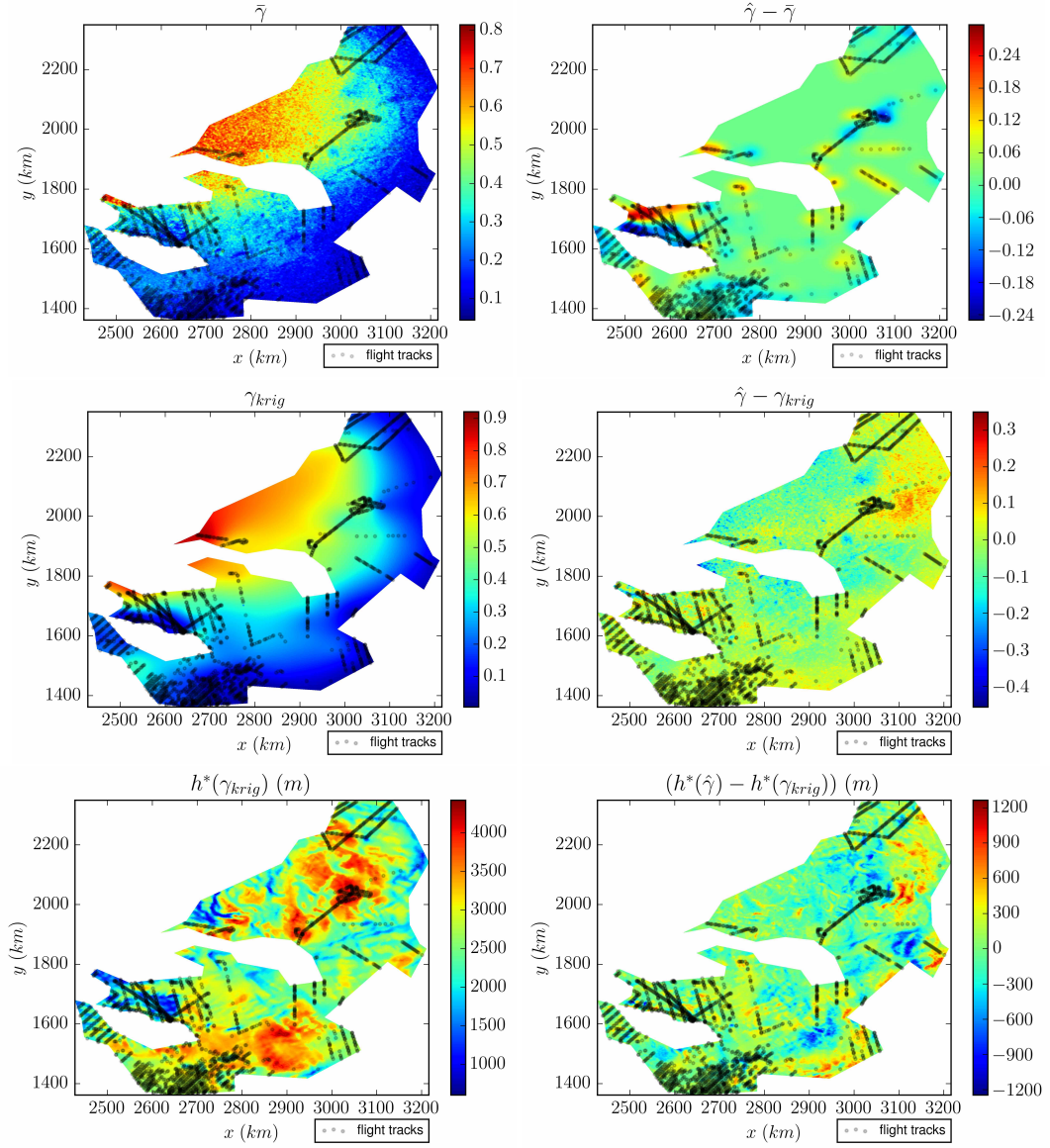


Fig. 8 Domain Ω_p =Ant1: comparison between different statistical learning methods at Step 2. (Up)(L) $\hat{\gamma}$ computed by ANN only. (R) Difference $(\hat{\gamma} - \bar{\gamma})$. (Middle)(L) γ_{krig} computed by ordinary Kriging. (R) Difference $(\hat{\gamma} - \gamma_{krig})$. (Down)(L) Inferred thickness $h^*(\gamma_{krig})$. (R) Difference $(h^*(\hat{\gamma}) - h^*(\gamma_{krig}))$.

Table 5 Domain Ω_p =Ant3, information and results.

Domain Ω_p & mesh information			
Surface $ \Omega_p $	250268 km ²		
Mean ice thickness of h_b (Bedmap2)	1822.8 m		
# mesh vertices: in Ω_p / on flight tracks	42881 / 2443		
RU-SIA model output (with $\hat{\gamma}$)	Median	Mean	Max
$ H(h_b) - H^{obs} $ (before h -inversion)	7.8 m	12.7 m	274.0 m
$ H(h^*) - H^{obs} $ (after h -inversion)	2.8 m	4.0 m	110.6 m
Inferred RHS \dot{a}			
$ \dot{a}^* - \dot{a}_b $	2.2 cm/y	2.5 cm/y	22.1 cm/y
$ \dot{a}^* - \dot{a}_b / \dot{a}_b $	11.2%	11.4%	20%
Inferred thickness h			
$ h^* - h_b $	70.0 m	124.5 m	862.2 m
$ h^* - h_b / h_b $	3.5%	6.6%	63.5%
Ice volume change in km ³ / in %	3.0 10 ³ km ³ / 0.5%		

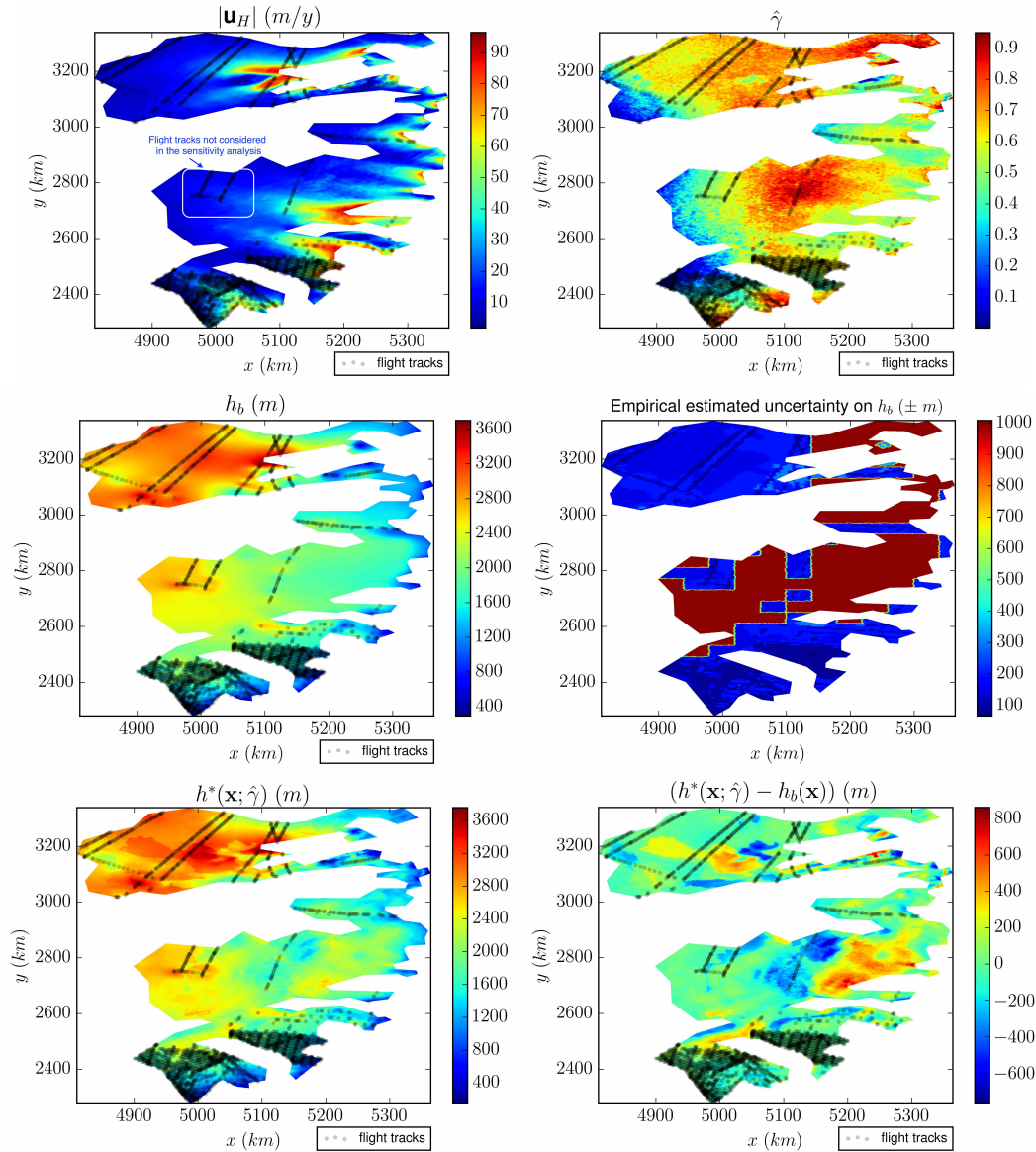


Fig. 9 Domain Ω_p =Ant3 (the plotted coordinates equal the Eastings-Northings plus (2800,2800)km): (Up)(L) Surface velocity module $|u_H|$ and flight tracks (R) $\hat{\gamma}$ computed by NNRK, see (11). (Middle)(L) Thickness h_b from Bedmap2 [14] (R) Empirical uncertainty on h_b from [14]. (Down)(L) Inferred thickness with $\hat{\gamma}$: $h^*(\hat{\gamma})$ (R) Difference $(h^*(\hat{\gamma}) - h_b)$.

Table 6 Domain Ω_p =Ant3. Comparison if considering or not the flights tracks indicated in Fig.9 (Up)(L).

Inferred thickness difference	Median	Mean	Max
$ h^*(\Gamma_{ir}^{all}) - h^*(\Gamma_{ir}^{less}) $	42.7m	56.3m	904.6m
$ h^*(\Gamma_{ir}^{all}) - h^*(\Gamma_{ir}^{less}) / h^*(\Gamma_{ir}^{less}) $	2.2%	2.8%	78.7%

5.2.2 Ant3: if removing some flight tracks

The ice thickness obtained if not considering the flight tracks indicated in Fig.9 (Up)(L) is compared to the original estimation h^* (the one plotted in Fig.9 (Down)(L)). For the same reason as in Ant1 case, both $\hat{\gamma}$ and h are changed all over the domain and in the vicinity of the missing flight track only. Largest changes are obtained in areas far to the missing tracks; also it may close to (assimilated) flight tracks, see e.g. the area around coordinates (5100,3200). The difference between the two estimations are plotted in Fig. 6 (Up)(R) and (Down)(R). (Again for a sake of readability, the legend in Fig. 10(Down)(R) has been bounded at ± 400 m; very few values being greater than this bound). Basic statistics on the difference are presented in Tab. 6. (A difference of 200m corresponds to $\approx 10\%$ of change). Again, the obtained variations on h are roughly half than the ones obtained from h_b , see tables 5 and 6: difference of 2.8% in mean vs 6.6%, and 2.2% vs 3.5% for median values.

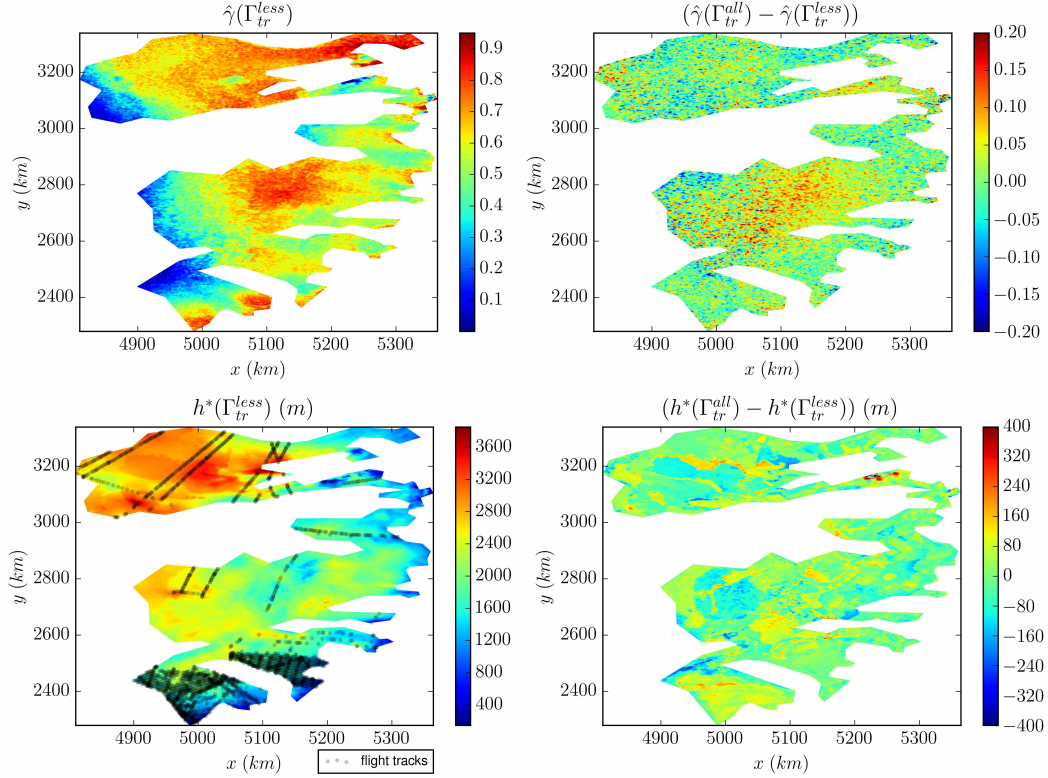


Fig. 10 Domain Ω_p =Ant3: comparison if not considering the flights tracks indicated in Fig.9 (Up)(L). (Up)(L) Field $\hat{\gamma}(\Gamma_{tr}^{less})$ (i.e. without the flights tracks indicated in Fig.9 (Up)(L)). (R) Difference $(\hat{\gamma}(\Gamma_{tr}^{all}) - \hat{\gamma}(\Gamma_{tr}^{less}))$. (Down)(L) Inferred thickness $h^*(\Gamma_{tr}^{less})$. (R) Difference between the two estimations: $(h^*(\Gamma_{tr}^{all}) - h^*(\Gamma_{tr}^{less}))$.

Table 7 Domain Ω_p =Ant3. Comparison of the original thickness estimation (obtained using NNRK) to the one obtained using ordinary Kriging at Step2).

Inferred thickness difference	Median	Mean	Max
$ h^*(\hat{\gamma}) - h^*(\gamma_{krig}) $	58.3 m	84.9 m	1418.7 m
$ h^*(\hat{\gamma}) - h^*(\gamma_{krig}) / h^*(\hat{\gamma}) $	2.9%	4.3%	124.5%

5.2.3 Ant3: with a different statistical learning method at Step 2)

Similarly to the Ant1 case, the difference between $\bar{\gamma}$ and $\hat{\gamma}$ (i.e. before and after the Kriging Step 2c) are mainly localised in the vicinity of the flight tracks with amplitudes up to $\sim \pm 0.20$, Fig. 11 (Up)(R). Again, the differences between $\hat{\gamma}$ and γ_{krig} are not in the vicinity of the tracks only; they may be anywhere. The observed difference is up to $\sim \pm 0.45$, see Fig. 11(Middle)(R). Statistics on the differences on the corresponding estimated thicknesses are presented in Tab. 7. The obtained differences on h are about one third lower than the ones obtained from h_b , see tables 2 and 7: difference of 4.3% in mean vs 6.6%, and 2.9% vs 3.5% for median values.

In summary, these empirical sensitivity analyses highlight the robustness and the reliability of the inversion method. If not considering some in-situ measurements (along some flight tracks), see figures 6 and 9 (Up)(L), the obtained differences on h are roughly half than the ones obtained between h^* and h_b .

If considering a simple Kriging method to estimate γ instead of the NNRK algorithm, the obtained differences on h in Ant1 case (resp. in Ant3 case) are roughly 1/2 (resp. 2/3) the ones obtained between h^* and h_b .

Therefore in all investigated cases, the obtained variations are lower than the ones obtained from h_b , see sections 5.1.1 and 5.2.1.

6 Conclusion

In this study, a method for inferring bedrock topography beneath glaciers at a wavelength $\sim 10\bar{h}$ (\bar{h} a characteristic thickness value) and with a resolution at $\sim \bar{h}$, is developed. The key elements of this inversion method are: 1) the Reduced Uncertainty flow model RU-SIA taking into account full physics (including non-uniform internal deformation due in particular to the temperature profile) and presenting a single dimensionless multi-physics parameter $\gamma(\mathbf{x})$; 2) two advanced

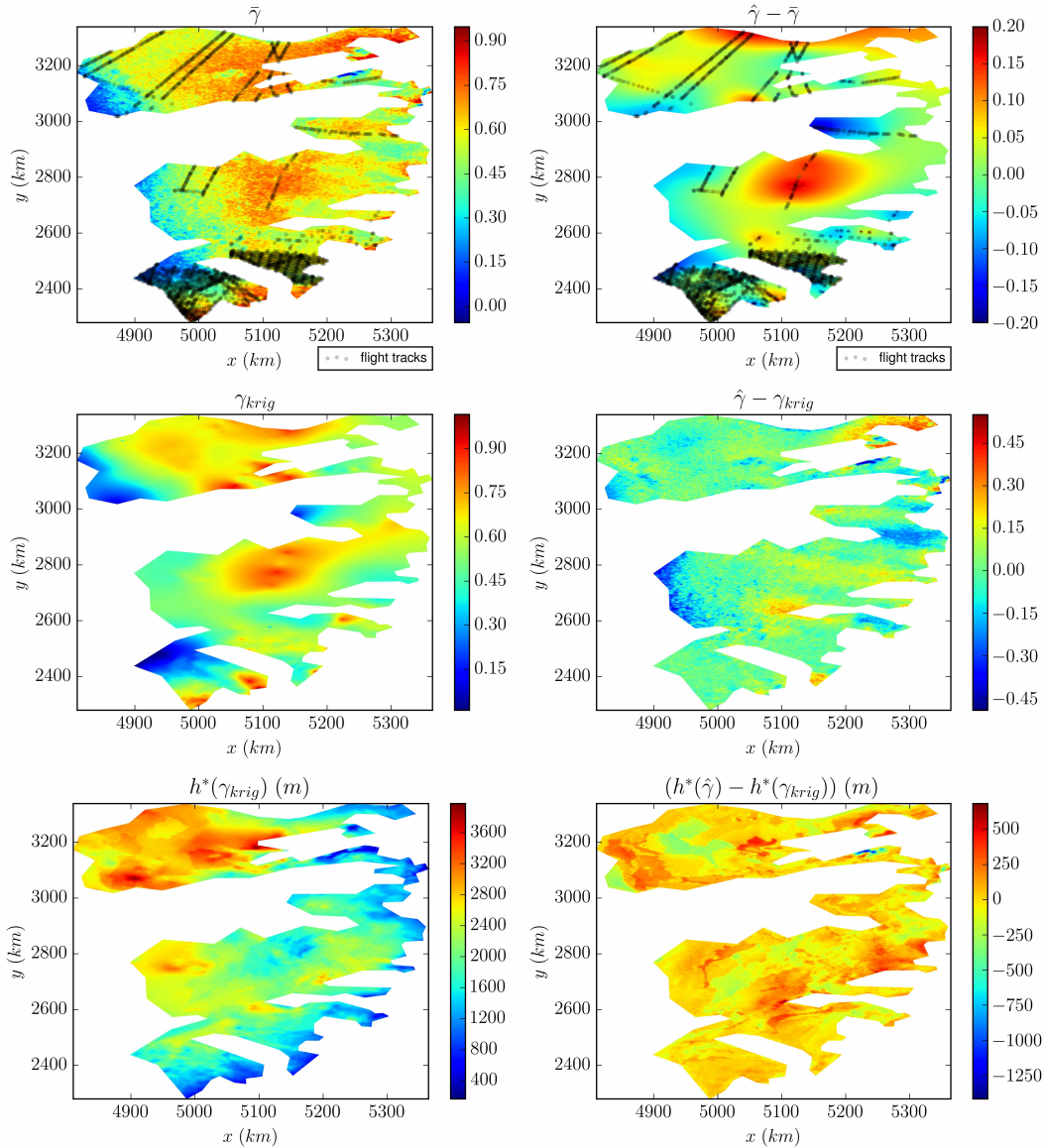


Fig. 11 Domain Ω_p =Ant3: comparison between different statistical learning methods at Step 2). (Up)(L) $\hat{\gamma}$ computed by ANN only. (R) Difference $(\hat{\gamma} - \tilde{\gamma})$. (Middle)(L) γ_{krig} computed by ordinary Kriging. (R) Difference $(\hat{\gamma} - \gamma_{krig})$. (Down)(L) Inferred thickness $h^*(\gamma_{krig})$. (R) Difference $(h^*(\hat{\gamma}) - h^*(\gamma_{krig}))$.

592 VDA processes; 3) an Artificial Neural Network (ANN) estimating $\gamma(\mathbf{x})$ from surface data and in situ measurements
 593 (acquired during airborne campaigns). An important feature of the method is to consider at each step robust and stable
 594 (or at least not trivially ill-posed) inverse problems. All steps of the algorithm have been thoroughly evaluated, partly
 595 in [39] and then fully in the present study, in particular the sensitivity of the estimation with respect to the presence
 596 or not of local in-situ measurements (airborne campaigns). The capabilities of ANN to infer the dimensionless multi-
 597 physics parameter γ from surface measurements alone are surprisingly excellent (at least for all datasets considered in
 598 EAIS). Note that it is easy to apply the only deep learning method (one of the steps in the present algorithm) to directly
 599 estimate the ice thickness h (instead of γ), by training the ANN from in-situ measurements acquired along the flight
 600 paths. We have performed such estimates. They turn out to be too dependent on the training data sets. Indeed, they are
 601 sensitive to the presence or absence of local in-situ measurements. In other words, these purely data-based estimates are
 602 insufficiently robust, unlike the γ estimates. Indeed, it seems more consistent to expect to be able to estimate an "intrinsic"
 603 dimensionless multi-physics parameter from the surface signature than from the ice depth. The present method allows
 604 to estimate the bed elevation between in-situ measurements, in poorly covered areas, and even in large uncovered areas,
 605 for example in EAIS where current estimates are based on gravimetry and therefore very uncertain. Indeed, the method
 606 developed is relatively global (with local constraints along the flight paths) and not purely local. We show here that the
 607 method is relatively robust. Estimates in the large uncovered areas remain as accurate as elsewhere. In the six large EAIS
 608 areas considered, the corrections made to Bedmap2 can be significant (up to 1000 m) even at 2 wavelengths (about 50

609 km) compared to "in-situ" measurements (flight paths). Outside the flight paths, the correction obtained for Bedmap2
 610 may or may not be significant, depending on the surface signature. The corrections obtained led to a total change in ice
 611 volume of 0.5 to 5.6%, depending on the area. In order to edit new bedrock maps from the present estimation method,
 612 additional research needs to be conducted, including adding additional prior information such as h regularization length
 613 scales. From a mathematical point of view, this can be easily done by adding additional regularization terms in the
 614 optimization formulations. This prior information should be introduced from geomorphological analysis and expertise. In
 615 our method the smoothing of the observation surface term $|\mathbf{u}_H|/\mathcal{S}$ can be defined as non-isotropic in order to distinguish
 616 the minimum wavelength of the streamline from that of the crossflow. The present bed estimation method can be applied
 617 to any ice sheet or ice cap, as long as the range of validity of the method is compatible with the observed flows, i.e.,
 618 strongly to moderately sheared flows, i.e., not almost purely sliding flows. On the other hand, we also show that these
 619 estimates connect continuously with those derived from mass conservation alone, and thus can complement the estimates
 620 established for areas of strongly sliding glaciers. Finally, the method can be easily extended to unsteady flows if the
 621 surface observations (elevation and velocity) are given in time and assuming that the initial condition is either unimportant
 622 on the time scale considered or approximately known.

623 *Acknowledgments & Authors' contributions*

624 J.M. has defined the research plan, has derived the equations and the computational methods and has written the manuscript.
 625 J.Z. has written the computational code, has performed all the numerical results (plus many exploring ones). Both have
 626 analyzed the numerical results.

627 J.Z. has been funded by a CNES TOSCA grant (oct. 2017-sept. 2018) during her post-doctoral stay at IMT - INSA
 628 Toulouse.

629 J.M. acknowledges M. Morlighem (Univ. California of Irvine) and H. Seroussi (JPL/Caltech-NASA) for numerous dis-
 630 cussions on datasets during his visits at UCI. The authors acknowledge reviewers for their kind comments which have
 631 helped to improve the manuscript presentation.

632 **References**

- 633 1. S. Arlot, A. Celisse, et al. A survey of cross-validation procedures for model selection. *Statistics surveys*, 4:40–79, 2010.
- 634 2. D. B. Bahr, W. T. Pfeffer, and G. Kaser. Glacier volume estimation as an ill-posed inversion. *Journal of Glaciology*, 60(223):922–934, 2014.
- 635 3. J. Bamber, J. Griggs, R. Hurkmans, J. Dowdeswell, S. Gogineni, I. Howat, J. Mougintot, J. Paden, S. Palmer, E. Rignot, et al. A new bed elevation
 636 dataset for greenland. *The Cryosphere*, 7(2):499–510, 2013.
- 637 4. JL Bamber, JL Gomez-Dans, and JA Griggs. A new 1 km digital elevation model of the antarctic derived from combined satellite radar and laser
 638 data—part I: Data and methods. *The Cryosphere*, 3(1):101–111, 2009.
- 639 5. Andrew F Bennett. *Inverse modeling of the ocean and atmosphere*. Cambridge University Press, 2005.
- 640 6. M. Boutounet, J. Monnier, and J.-P. Vila. Multi-regime shallow free surface laminar flow models for quasi-newtonian fluids. *European Journal of*
 641 *Mechanics-B/Fluids*, 55:182–206, 2016.
- 642 7. D. Brinkerhoff, A. Aschwanden, and M. Truffer. Bayesian inference of subglacial topography using mass conservation. *Frontiers in Earth Science*,
 643 4:8, 2016.
- 644 8. W.F. Budd and I.F. Allison. An empirical scheme for estimating the dynamics of unmeasured glaciers. In *Proceedings of the Moscow Symposium*
 645 *Snow and Ice*, pages 246–256, 1975.
- 646 9. G. Clarke, E. Berthier, C. Schoof, and A. Jarosch. Neural networks applied to estimating subglacial topography and glacier volume. *Journal of*
 647 *Climate*, 22(8):2146–2160, 2009.
- 648 10. L. Dalc n, R. Paz, and M. Storti. Mpi for python. *Journal of Parallel and Distributed Computing*, 65(9):1108–1115, 2005.
- 649 11. V Demyanov, M Kanevsky, S Chernov, E Savelieva, and V Timonin. Neural network residual kriging application for climatic data. *Journal of*
 650 *Geographic Information and Decision Analysis*, 2(2):215–232, 1998.
- 651 12. D. Farinotti, D. Brinkerhoff, D. Clarke, J. F rst, H. Frey, P. Gantayat, F. Gillet-Chaulet, C. Girard, M. Huss, P. Leclercq, et al. How accurate are
 652 estimates of glacier ice thickness ? results from itmix, the ice thickness models intercomparison experiment. *The Cryosphere*, 11(2):949–970,
 653 2017.
- 654 13. Christopher J Fogwill, Christian SM Turney, Katrin J Meissner, Nicholas R Golledge, Paul Spence, Jason L Roberts, Mathew H England, Richard T
 655 Jones, and Lionel Carter. Testing the sensitivity of the east antarctic ice sheet to southern ocean dynamics: past changes and future implications.
 656 *Journal of Quaternary Science*, 29(1):91–98, 2014.
- 657 14. P. Fretwell, H. Pritchard, D. Vaughan, J. Bamber, N. Barrand, R. Bell, C. Bianchi, R. Bingham, D. Blankenship, G. Casassa, et al. Bedmap2:
 658 improved ice bed, surface and thickness datasets for antarctica. *The Cryosphere*, 7(1), 2013.
- 659 15. C. Geuzaine and J.-F. Remacle. Gmsh: A 3-d finite element mesh generator with built-in pre-and post-processing facilities. *International journal*
 660 *for numerical methods in engineering*, 79(11):1309–1331, 2009.
- 661 16. X. Glorot, A. Bordes, and Y. Bengio. Deep sparse rectifier neural networks. In *Proceedings of the fourteenth international conference on artificial*
 662 *intelligence and statistics*, pages 315–323, 2011.
- 663 17. R. Greve and H. Blater. *Dynamics of Ice Sheets and Glaciers*. Advances in Geophysical and Environmental Mechanics and Mathematics. Springer-
 664 Verlag, 2009.
- 665 18. G. Gudmundsson. Transmission of basal variability to a glacier surface. *Journal of Geophysical Research: Solid Earth*, 108(B5), 2003.
- 666 19. G. Gudmundsson. Analytical solutions for the surface response to small amplitude perturbations in boundary data in the shallow-ice-stream
 667 approximation. *The Cryosphere*, 2(2):77–93, 2008.
- 668 20. S.A. Haben, A.S. Lawless, and N.K. Nichols. Conditioning of incremental variational data assimilation, with application to the met office system.
 669 *Tellus A: Dynamic Meteorology and Oceanography*, 63(4):782–792, 2011.
- 670 21. C. Heining and M. Sellier. Direct reconstruction of three-dimensional glacier bedrock and surface elevation from free surface velocity. *AIMS*
 671 *Geosciences*, 2:45–63, 2016.

- 672 22. V. Helm, A. Humbert, and H. Miller. Elevation and elevation change of greenland and antarctica derived from cryosat-2. *The Cryosphere*,
673 8(4):1539–1559, 2014.
- 674 23. R. Hindmarsh. A numerical comparison of approximations to the stokes equations in ice sheet and glacier modeling. *J. geophysical research*, 109,
675 2004.
- 676 24. R. Hindmarsh. Stress gradient damping of thermoviscous ice flow instabilities. *Journal of Geophysical Research: Solid Earth*, 111(B12), 2006.
- 677 25. M. Huss and D. Farinotti. A high-resolution bedrock map for the antarctic peninsula. *The Cryosphere*, 8(4):1261–1273, 2014.
- 678 26. M.F. Hutchinson. A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. *Journal of hydrology*,
679 106(3-4):211–232, 1989.
- 680 27. D. Jansen, H. Sandhäger, and W. Rack. Model experiments on large tabular iceberg evolution: ablation and strain thinning. *Journal of glaciology*,
681 51(174):363–372, 2005.
- 682 28. B. Kaltenbacher, A. Neubauer, and O. Scherzer. *Iterative regularization methods for nonlinear ill-posed problems*, volume 6. Walter de Gruyter,
683 2008.
- 684 29. B. Kamb and K. Echelmeyer. Stress-gradient coupling in glacier flow: I. longitudinal averaging of the influence of ice thickness and surface slope.
685 *Journal of Glaciology*, 32(111):267–284, 1986.
- 686 30. M. Kanevski, V. Timonin, and A. Pozdnukhov. *Machine learning for spatial environmental data: theory, applications, and software*. EPFL press,
687 2009.
- 688 31. D. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- 689 32. Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *nature*, 521(7553):436, 2015.
- 690 33. A. Lorenc. Optimal nonlinear objective analysis. *Quarterly Journal of the Royal Meteorological Society*, 114(479):205–240, 1988.
- 691 34. 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.
692 *The Cryosphere*, 8:721–741, 2014.
- 693 35. N. Martin and J. Monnier. Inverse rheometry and basal properties inference for pseudoplastic geophysical flows. *European Journal of Mechanics-*
694 *B/Fluids*, 50:110–126, 2015.
- 695 36. L. Michel, M. Picasso, D. Farinotti, A. Bauder, M. Funk, and H. Blatter. Estimating the ice thickness of mountain glaciers with an inverse approach
696 using surface topography and mass-balance. *Inverse Problems*, 29(3):035002, 2013.
- 697 37. L. Michel, M. Picasso, D. Farinotti, M. Funk, and H. Blatter. Estimating the ice thickness of shallow glaciers from surface topography and
698 mass-balance data with a shape optimization algorithm. *Computers & Geosciences*, 66:182–199, 2014.
- 699 38. J. Monnier and P.-E. des Bosc. Inference of the bottom properties in shallow ice approximation models. *Inverse Problems*, 33(11):115001, 2017.
- 700 39. J. Monnier and J. Zhu. Inference of the bottom topography in anisothermal mildly-sheared shallow ice flows. *Computer Methods in Applied*
701 *Mechanics and Engineering*, 348:954–977, 2019.
- 702 40. M. Morlighem, E. Rignot, J. Mouginot, H. Seroussi, and E. Larour. High-resolution ice-thickness mapping in south greenland. *Annals of Glaciol-*
703 *ogy*, 55(67):64–70, 2014.
- 704 41. M. Morlighem, E. Rignot, H. Seroussi, E. Larour, H. Ben Dhia, and D. Aubry. A mass conservation approach for mapping glacier ice thickness.
705 *Geophysical Research Letters*, 38(19), 2011.
- 706 42. M. Morlighem, C. Williams, E. Rignot, L. An, J.E. Arndt, J. Bamber, G. Catania, N. Chauché, J. Dowdeswell, B. Dorschel, et al. Bedmachine
707 v3: Complete bed topography and ocean bathymetry mapping of greenland from multibeam echo sounding combined with mass conservation.
708 *Geophysical research letters*, 44(21), 2017.
- 709 43. J. Mouginot, E. Rignot, B. Scheuchl, and R. Millan. Comprehensive annual ice sheet velocity mapping using landsat-8, sentinel-1, and radarsat-2
710 data. *Remote Sensing*, 9(4):364, 2017.
- 711 44. P.B. Price, O. Nagornov, R. Bay, D. Chirkin, Y. He, P. Miocinovic, A. Richards, K. Woschnagg, B. Koci, and V. Zagorodnov. Temperature profile
712 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–
713 7847, 2002.
- 714 45. U. Radok, D. Janssen, and W. Budd. Steady-state temperature profiles in ice sheets. *Bull. Int. Assoc. Scient. Hydrol*, 8(1):36, 1970.
- 715 46. L. Rasmussen. Bed topography and mass-balance distribution of columbia glacier, alaska, usa, determined from sequential aerial photography.
716 *Journal of Glaciology*, 34(117):208–216, 1988.
- 717 47. C. Schoof and C. Hindmarsh. Thin-film flows with wall slip: an asymptotic analysis of higher order glacier flow models. *Q. J. Mech. Appl. Math.*,
718 63(1):73–114, 2010.
- 719 48. M. Sellier. Inverse problems in free surface flows: a review. *Acta Mechanica*, 227(3):913–935, 2016.
- 720 49. H. Seroussi, M. Morlighem, E. Rignot, A. Khazendar, E. Larour, and J. Mouginot. Dependence of century-scale projections of the greenland ice
721 sheet on its thermal regime. *Int. Glacio. Soc.*, 59:1024–1034, 2013.
- 722 50. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting.
723 *The Journal of Machine Learning Research*, 15(1):1929–1958, 2014.
- 724 51. W. Van Pelt, J. Oerlemans, C. Reijmer, R. Pettersson, V. Pohjola, E. Isaksson, D. Divine, et al. An iterative inverse method to estimate basal
725 topography and initialize ice flow models. *The Cryosphere*, 7:987–1006, 2013.
- 726 52. J.M. 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.
727 Modelling the climate and surface mass balance of polar ice sheets using racmo2: Part 2: Antarctica (1979-2016). *The Cryosphere*, 12(4):1479–
728 1498, 2018.
- 729 53. G. Wahba. *Spline models for observational data*, volume 59. SIAM, 1990.
- 730 54. R. Williams, R. Hindmarsh, and R. Arthern. Calculating balance velocities with a membrane stress correction. *Journal of Glaciology*, 60(220):294–
731 304, 2014.
- 732 55. R. Winkelmann, M.A. Martin, M. Haseloff, T. Albrecht, E. Bueler, C. Khroulev, and A. Levermann. The potsdam parallel ice sheet model
733 (pism-pik) part 1: Model description. *The Cryosphere*, 5(3):715–726, 2011.

734 A Results for other EAIS areas

735 In this section, the estimation of ice thickness is performed in the four other areas indicated in Fig. 4: Ant2, Ant4, Ant5 and Ant6. Ant2 is located
736 upstream of Fisher and Merllor ice-streams (upstream Amery ice shelf). Ant4 is located upstream Totten ice-streams (Wilkes land and Terre Adélie),
737 one of the largest discharger of ice in EAIS. Ant5 is located upstream Ninnis and Mertz ice-streams in Terre Adélie and George V land. Ant6 is located
738 upstream Byrd ice-streams (east of Ross Ice shelf), see Fig. 4.

739 These areas have been relatively well covered during the airborne campaigns excepted the north-east part of Ant2, see Fig. 12. Distances between flight
740 tracks are relatively low, therefore the empirical uncertainty assigned to Bedmap2 estimations h_b is low too: $\approx \pm [100 - 250]$ m, see figures 12, 13, 14,
741 15 (Middle).

742 For each area, domain information and statistics on the numerical results are presented, see tables 8-11.

Table 8 Domain Ω_p =Ant2, information and results.

Domain Ω_p & mesh information			
Surface $ \Omega_p $	431860 km^2		
Mean ice thickness of h_b (Bedmap2)	2144.4 m		
# mesh vertices: in Ω_p / on flight tracks	65123/5194		
RU-SIA model output (with $\hat{\gamma}$)	Median	Mean	Max
$ H(h_b) - H^{obs} $ (before h -inversion)	8.7 m	17.3 m	149.9 m
$ H(h^*) - H^{obs} $ (after h -inversion)	3.9 m	5.1 m	49.2 m
Inferred RHS \dot{a}			
$ \dot{a}^* - \dot{a}_b $	0.6 cm/y	0.8 cm/y	4.0 cm/y
$ \dot{a}^* - \dot{a}_b / \dot{a}_b $	14.8 %	13.2 %	20 %
Inferred thickness h			
$ h^* - h_b $	171.1 m	302.4 m	2025.6 m
$ h^* - h_b / h_b $	8.3 %	14.2 %	80.0 %
Ice volume change in km^3 / in %	5.2 $10^4 km^3$ / 5.6 %		

Table 9 Domain Ω_p =Ant4, information and results.

Domain Ω_p & mesh information			
Surface $ \Omega_p $	439045 km^2		
Mean ice thickness of h_b (Bedmap2)	2745.4 m		
# mesh vertices: in Ω_p / on flight tracks	61219/4977		
RU-SIA model output (with $\hat{\gamma}$)	Median	Mean	Max
$ H(h_b) - H^{obs} $ (before h -inversion)	6.3 m	8.4 m	66.9 m
$ H(h^*) - H^{obs} $ (after h -inversion)	3.0 m	4.1 m	45.3 m
Inferred RHS \dot{a}			
$ \dot{a}^* - \dot{a}_b $	2.0 cm/y	2.7 cm/y	15.0 cm/y
$ \dot{a}^* - \dot{a}_b / \dot{a}_b $	8.7 %	10.1 %	20 %
Inferred thickness h			
$ h^* - h_b $	147.6 m	185.0 m	1241.6 m
$ h^* - h_b / h_b $	5.7 %	7.0 %	50.4 %
Ice volume change in km^3 / in %	1.5 $10^4 km^3$ / 1.5 %		

Table 10 Ant 5, information and results.

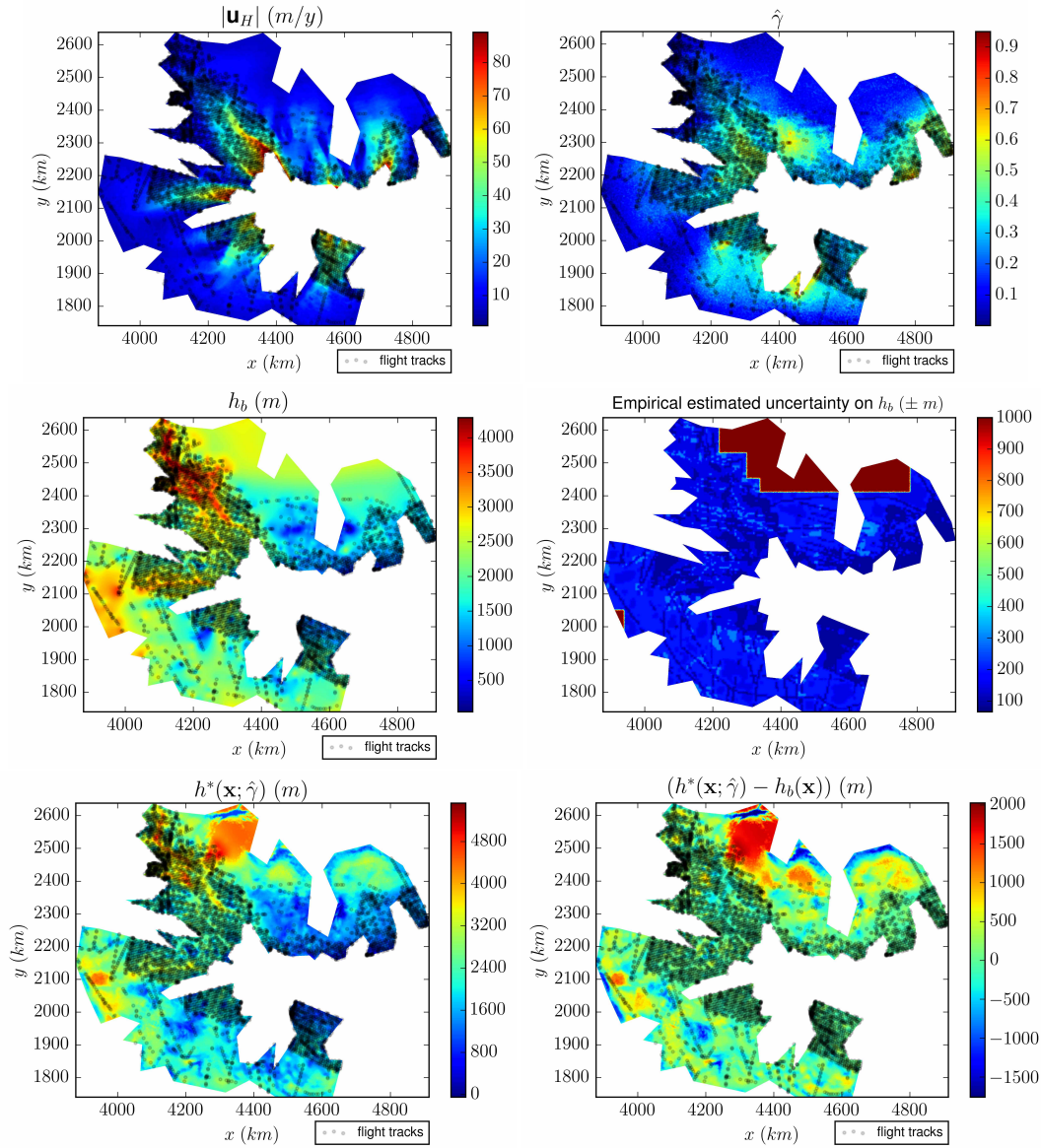
Domain Ω_p & mesh information			
Surface $ \Omega_p $	362019 km^2		
Mean ice thickness of h_b (Bedmap2)	2415.3 m		
# mesh vertices: in Ω_p / on flight tracks	41597/2351		
RU-SIA model output (with $\hat{\gamma}$)	Median	Mean	Max
$ H(h_b) - H^{obs} $ (before h -inversion)	18.3 m	23.9 m	173.3 m
$ H(h^*) - H^{obs} $ (after h -inversion)	3.9 m	5.1 m	43.1 m
Inferred RHS \dot{a}			
$ \dot{a}^* - \dot{a}_b $	3.9 cm/y	4.4 cm/y	14.5 cm/y
$ \dot{a}^* - \dot{a}_b / \dot{a}_b $	18.8 %	15.4 %	20 %
Inferred thickness h			
$ h^* - h_b $	275.1 m	373.1 m	1989.2 m
$ h^* - h_b / h_b $	12.1 %	15.9 %	68.8 %
Ice volume change in km^3 / in %	3.8 $10^4 km^3$ / 5.6 %		

743 As already noticed, RU-SIA model already fits well with the surface topography after the data-driven model (Step 2) only; that is with $\hat{\gamma}$ and h_b as
744 parameters in (1): see lines " $|H(h_b) - H_{obs}|$ (before h -inversion)" in the tables. Next RU-SIA model fits accurately the surface topography after the
745 re-calibration / estimation of h . Indeed, misfit values range within $\approx [4 - 5]$ m in mean; see lines " $|H(h_b) - H_{obs}|$ (after h -inversion)" in tables.
746 The corrections made on \dot{a} are $\sim [10 - 17]$ % in mean, that is within the a-priori uncertainty range indicated in [52].
747 The estimated thickness h^* is plotted in each case, see figures 12-15 (Down)(L); its difference with h_b is plotted (Down)(R). The corrections made
748 to h_b are non negligible: they are ranging within $[7.0 - 15.9]$ % in mean (corresponding to $[185 - 373]$ m , see the lines " $|h^* - h_b|$ " in tables). These
749 corrections lead to changes of total ice volume by $[1.5, 5.6]$ %. Again, the obtained difference with h_b is independent of the distance to the closest flight
750 track. Maximum values of correction to h_b can be locally high. In the uncovered north-east Ant2 area, corrections are up to ≈ 2000 m . At ≈ 50 km from
751 the nearest flight track (that is ~ 2 minimal wave lengths of the model) correction may reach 1000 m (even in area surrounded by well covered areas),
752 see e.g. figures 14 and 15 (Down)(R). Beyond ≈ 100 km from the nearest data, the correction of h_b may be significant or not, depending on the surface
753 signature, see e.g. figures 14 and 15 (Down)(R).

754

Table 11 Domain Ω_p =Ant6, information and results.

Domain Ω_p & mesh information			
Surface $ \Omega_p $	406388 km^2		
Mean ice thickness of h_b (Bedmap2)	2672.9m		
# mesh vertices: in Ω_p / on flight tracks	63981/3012		
RU-SIA model output (with $\hat{\gamma}$)			
$ H(h_b) - H^{obs} $ (before h -inversion)	Median	Mean	Max
	8.3m	11.0m	46.8m
$ H(h^*) - H^{obs} $ (after h -inversion)	2.7m	3.4m	21.3m
Inferred RHS \dot{a}			
$ \dot{a}^* - \dot{a}_b $	0.4 cm/y	0.5 cm/y	1.8 cm/y
$ \dot{a}^* - \dot{a}_b / \dot{a}_b $	15.8%	14.2%	20%
Inferred thickness h			
$ h^* - h_b $	218.6m	313.9m	1777.3m
$ h^* - h_b / h_b $	8.1%	11.6%	63.8%
Ice volume change in km^3 / in %	$2.6 \cdot 10^3 km^3$ / 1.7%		

**Fig. 12** Domain Ω_p =Ant2 (the plotted coordinates equal the Eastings-Northings plus (2800,2800)km): (Up)(L) Surface velocity module $|\mathbf{u}_H|$ and flight tracks. (R) $\hat{\gamma}$ computed by NNRK, see (11). (Middle)(L) Thickness h_b from Bedmap2 [14]. (R) Empirical uncertainty on h_b from [14]. (Down)(L) Inferred thickness with $\hat{\gamma}$: $h^*(\hat{\gamma})$. (R) Difference $(h^*(\hat{\gamma}) - h_b)$.

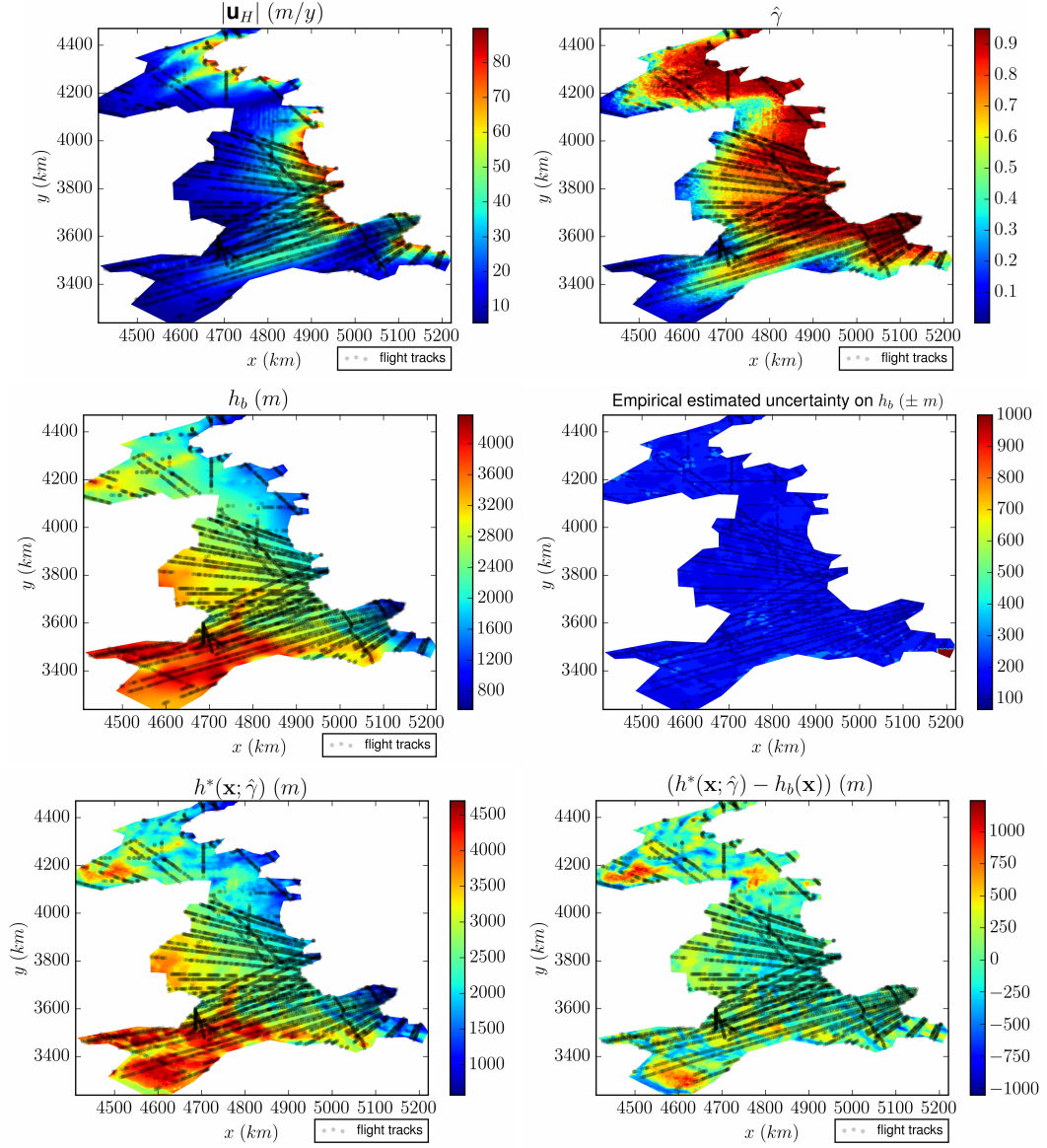


Fig. 13 Domain $\Omega_p = \text{Ant4}$ (the plotted coordinates equal the Eastings-Northings plus (2800, 2800)km): (Up)(L) Surface velocity module $|\mathbf{u}_H|$ and flight tracks. (R) $\hat{\gamma}$ computed by NNRK, see (11). (Middle)(L) Thickness h_b from Bedmap2 [14]. (R) Empirical uncertainty on h_b from [14]. (Down)(L) Inferred thickness with $\hat{\gamma}$: $h^*(\hat{\gamma})$. (R) Difference $(h^*(\hat{\gamma}) - h_b)$.

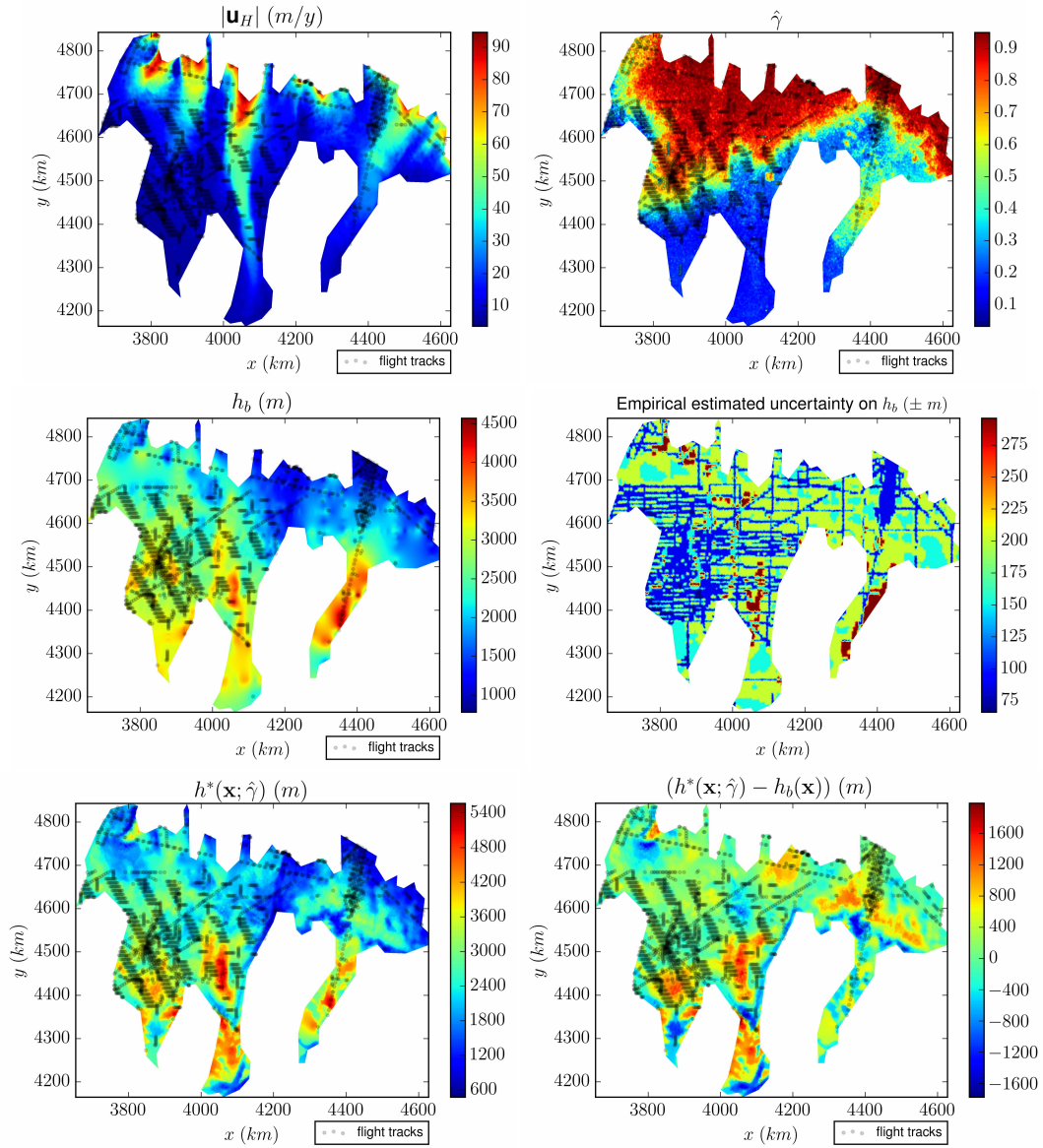


Fig. 14 Domain $\Omega_p=Ant5$ (the plotted coordinates equal the Eastings-Northings plus (2800,2800)km): (Up)(L) Surface velocity module $|\mathbf{u}_H|$ and flight tracks. (R) $\hat{\gamma}$ computed by NNRK, see (11). (Middle)(L) Thickness h_b from Bedmap2 [14]. (R) Empirical uncertainty on h_b from [14]. (Down)(L) Inferred thickness with $\hat{\gamma}$: $h^*(\hat{\gamma})$. (R) Difference $(h^*(\hat{\gamma}) - h_b)$.

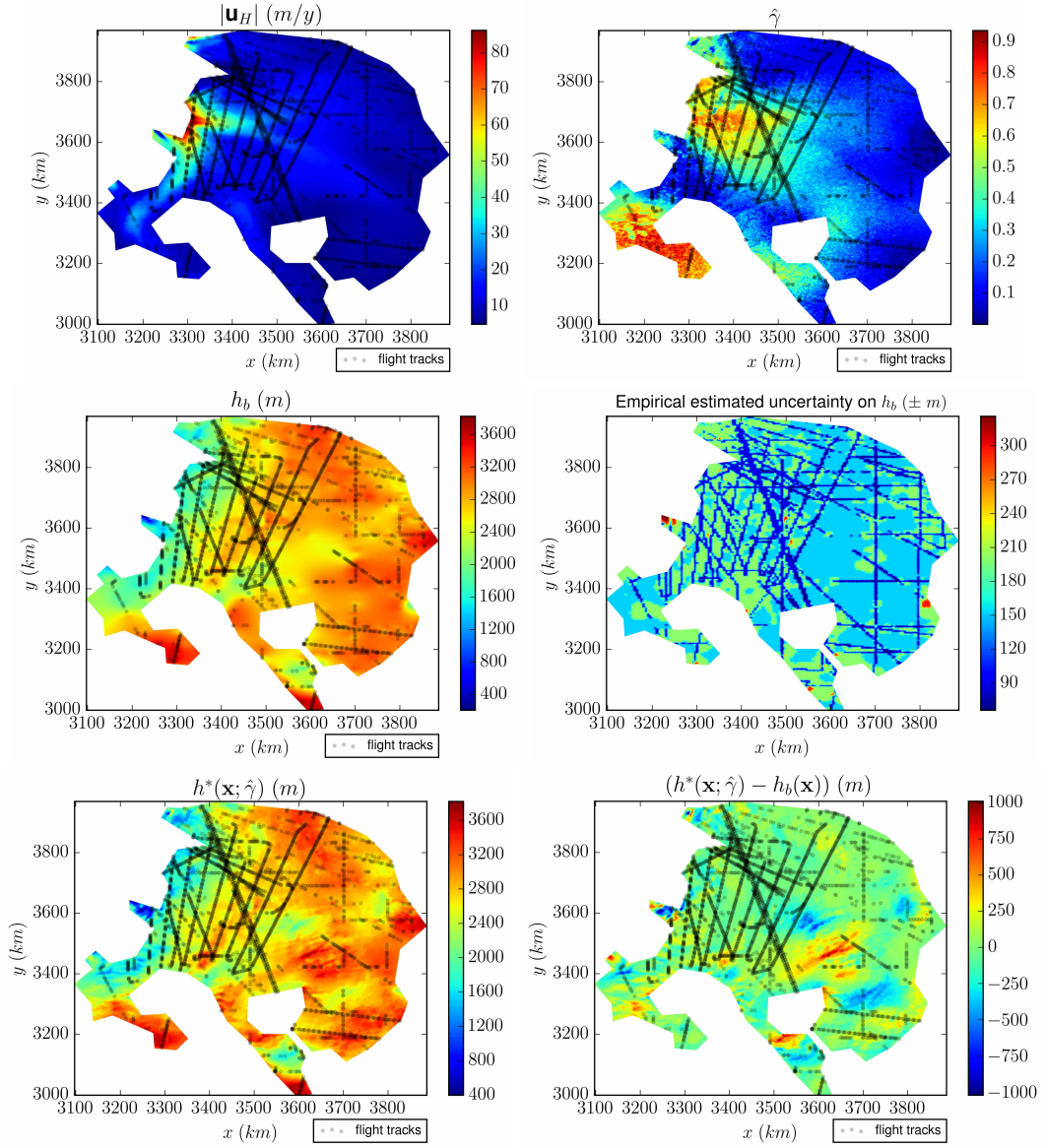


Fig. 15 Domain $\Omega_p = \text{Ant6}$ (the plotted coordinates equal the Eastings-Northings plus (2800,2800)km): (Up)(L) Surface velocity module $|\mathbf{u}_H|$ and flight tracks. (R) $\hat{\gamma}$ computed by NNRK, see (11). (Middle)(L) Thickness h_b from Bedmap2 [14]. (R) Empirical uncertainty on h_b from [14]. (Down)(L) Inferred thickness with $\hat{\gamma}$: $h^*(\hat{\gamma})$. (R) Difference $(h^*(\hat{\gamma}) - h_b)$.