Inference of the bed topography in poorly flew over ice-sheets areas from surface data and a reduced uncertainty flow model

Jerome Monnier, Jiamin Zhu

Inference of the bed topography in poorly flew over ice-sheets areas from surface data and a reduced uncertainty flow model

J. Monnier [a] and Jiamin Zhu[a]

[a] INSA & Institut de Mathématiques de Toulouse (IMT), France

ARTICLE HISTORY
Compiled November 19, 2018

ABSTRACT
A method to infer the bed elevation from glaciers surface measurements (elevation, velocity) and sparse in-situ thickness values is developed and assessed. This inversion method relies on: a statistical model (Deep Neural Network) based on the in-situ thickness measurements, the dedicated RU-SIA flow model (RU for Reduced Uncertainty) natively integrating the surface measurements (altimetry, InSAR) and advanced Variational Data Assimilation processes. The RU-SIA model takes into account basal slipperiness and non uniform vertical profiles (including thermal gradients) via an unique dimensionless parameter. The inversion method is robust; it may be applied to very poorly covered and uncovered areas during airborne campaigns as soon as flows are moderately sheared. Numerical inversions are performed for some large East Antarctica Ice Sheet areas presenting surface velocities ranging from ~ 5 to 80 m/y. Estimations are provided in uncovered areas during airborne campaigns hence presenting up to now highly uncertain bed elevation values. The estimations are valid for wave lengths greater than $\sim 10\bar{h}$ due to the considered shallow flow assumption, with a resolution at $\sim \bar{h}$ ($\bar{h}$ a characteristic thickness value). Detailed analysis and comparisons with the bed topography BedMap2 are presented.

KEYWORDS
Data assimilation; reduced flow model; surface data; bed inference; Antarctica.

1. Introduction

Bed topography elevation is a fundamental data to set up ice-sheets flow models; moreover if combined with the surface topography measurements (acquired by altimetry) it straightforwardly provides the ice volume. In Antarctica and Greenland, ice thickness measurements are available along airborne radio-echo sounding tracks see e.g. CReSIS RDS database [1]. The current databases are particularly consequent in fast ice costal stream areas; on the contrary they are generally very sparse inland and may be even nonexistent deep inland, see [1,2] and references therein. In other respects, numerous satellites have provided accurate measurements of ice sheets surfaces: altimeters provide surface elevation H at $\approx \pm 20$ cm for 1 km^2 pixels, see e.g. [1,3], radar interferometers (InSar) provide surface velocity $\mathbf{u}_H$ see e.g. [4]. Next the challenge is to fill up the bed elevation values between the thickness measurements;

CONTACT J. Monnier. Email: jerome.monnier@insa-toulouse.fr

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

this can be done by modelling and using all available data (i.e. dense and accurate surface data plus the sparse in-situ ones).
The first model one naturally mind to is an ordinary Kriging, that is a Gaussian process regression method providing the best linear unbiased prediction of the intermediate values. Indeed it is the most classical method used in geo-statistics, see e.g. [5]; it may be considered as a purely data-driven model. This statistical interpolation method has enabled to provide bed elevation estimations e.g. deep inland Greenland, see [6]. This technique leads to large uncertainties a dozen of km away of the measurements, see e.g. [2].
BedMap2 [2], the current reference bed topography elevation in Antarctica, has been built up from a compilation of consequent datasets. Inland between the measured ice thickness values (mainly acquired during airborne campaigns), bed elevation values are estimated by running the ArcGIS Topogrid routine (ESRI Ltd) based on the Anudem algorithm, see [7]. It is essentially a thin plate spline technique solved by smoothed finite differences, see [8]. Empirical uncertainties are provided, see [2] for more details. For cells located less than 20km from a measurement, [2] suggests that the extension error is an increasing function of the distance; beyond 20 km, the error would be uncorrelated with distance. For cells located more than 50 km from any measurement, the thickness estimation is based on gravity-field inversion, hence accompanied with very large uncertainty ($\pm \sim 1000\,m$ according to [2]).
To reduce uncertainties on the bed topography estimations, combining physical-based models with datasets is definitively a good direction. However more a physical model is complex, more it requires detailed information to be solved, and more difficult it is to invert. Thus the challenge is to employ a model sufficiently complex to be useful but simple enough to lead to well-posed inverse problem (i.e. not leading to severe equifinality issues) and stable (errors at input do not generate large drift of estimated values). Indeed in inverse modelling and data assimilation, equifinality issues are the common difficult pitfall to avoid, see e.g. [9] in different geophysical contexts. Moreover if the datasets are a-priori meaningful and large enough, the use of Deep Neural Networks is definitively a good direction too.
A comparison of inverse methods to estimate the ice thickness is presented in [10]. The comparison relies on numerous test cases representing a large spectrum of ice flow regimes (but no ice-sheet is considered); one of the constraint is that no prior thickness value is provided. The 15 inter-compared methods are classified by resolution type and not by domain of validity (e.g. in function of the flow regime). Numerical comparisons are presented; however no analysis on the equifinality issue(s) is proposed (well-posed vs ill-posed inverse formulations).
In fast ice streams, the flow may be considered as plug flows (pure sliding); then the inversion of the depth-integrated mass equation enables to fill up the gaps downstream (and upstream) the measurements, see e.g. [11,12]. The measurements locations (e.g. the flight tracks during airborne campaigns) have to be cross-lines and relatively dense. Indeed it is well known that this transport mass equation is intrinsically unstable to invert and it propagates errors, see e.g. [13] and references therein. To locally damp this feature, artificial diffusion regularising the equation may be introduced. The inversion of mass conservation proposed in [11,14], next combined with surface measurements by Variational Data Assimilation (VDA) in [15,16], has enabled to greatly improve the bed topography estimations under ice streams along flow lines in Greenland [12,16]. [17] applies a Bayesian framework to provide probability distributions of thickness assuming Gaussian covariance structures of input data (the algorithm being equivalent to ordinary Kriging if no data is available).

This Bayesian approach is particularly suitable for these inversions; however in this case, the estimations are based on mass conservation only.
For fully sheared flows, the isothermal SIA flow model (with no slip at bottom) has been inverted e.g. in [18,19], providing robust estimations but relevant in a restricted flow regime only. In moderately sheared flows, slipperiness at bottom has to be taken into account hence transforming the inverse problem to a severely ill-posed one. Indeed in this case, the flow models present various unknown parameters including the rate factor (internal deformation) and the basal slipperiness. [20] inverts the 1D depth-integrated SIA equation with slipperiness and a shape factor (modelling the 3D features of the flow) leading to an ill-posed inverse problem (the inversions are performed by imposing empirical constant values for the few unknown parameters). Based on the inversion of the complete hybrid SIA-SSA system PISM [21], [22] proposes an empirical iterative method to fit with the surface elevation by calibrating the bed elevation only. In [23] the SIA flow model with slipperiness at bottom is inverted by distinguishing different regimes, enabling to define well-posed inverse problems, hence providing stable and robust inversions. However the rate factor (depending in particular on the thermal field) is supposed to be constant; unfortunately this assumption is rarely satisfied.
Inland ice-sheets in mid-range surface velocity areas ($\approx [10-80]$ m/y in Antarctica), the flows are moderately sheared; they cannot reasonably be modelled using uniform models like plug flows or isothermal ones. In these areas, the measured surface features (elevation and velocity) are the signature of both the slipperiness at bottom *and* internal deformation; the latter depending on the constitutive ice behaviour *and* thermal gradients. Therefore inverting the surface data in moderately sheared flows is highly challenging.
The present study aims at solving this inverse problem: estimating the ice thickness in moderately sheared flows while taking into account the complete physics of the flow. To do so, a key ingredient is the Reduced Uncertainty SIA (RU-SIA) model derived in [24]. This flow model is dedicated to the present data inversions by natively integrating the surface measurements. It is a depth-integrated shallow flow model respecting a "well-balanced complexity" since 2D (and not 3D with mobile surface like full Stokes) but quite complete since mass and momentum equations are considered. Moreover the RU-SIA equation takes into account the non uniform internal properties (in particular due to varying vertical thermal profiles) while it presents an unique uncertain dimensionless parameter, see [23,24].
In other respects, inland moderately sheared flows have been poorly covered during the airborne campaigns (they are difficult and expansive to overflight). Therefore the chosen model needs to be stable and robust when inverted even in a lack of in-situ data. This mathematical feature is as important as the model consistency; this feature is all but trivial to obtain, see e.g. [13,23,25,26]. Moreover it would be highly valuable that its inversion capabilities would be as insensitive as possible to the measurements locations.
It is worth to point out that only an effective bed topography can be infered from the surface signature. Indeed the glaciers flows act as low-band filters: the bed variations are filtered by the flow with filtering features depending on the flow regime, see [27–30] for detailed analysis. To a flow regime and a flow model corresponds a minimal inferable wave length [28–30].
In the present study, a new inversion method to infer the bed elevation from surface measurements (elevation, velocity) and sparse in-situ thickness values is developed, assessed and applied to various inland East Antartica Ice Sheet (EAIS) areas. This

inversion method relies on the RU-SIA equation developed in [23,24], the advanced VDA process proposed in [24] and a Deep Neural Network (statistical learning). The latter aims at estimating the unique dimensionless parameter of the RU-SIA equation from the in-situ datasets (available along the flight tracks). The resulting inversion strategy is the same as the one assessed in detail in [24] excepted the statistical learning step which is an ordinary Kriging in [24] instead of a Neural Network Residual Kriging (NNRK) algorithm [31,32] here. The two global inversion algorithm versions are compared in two large EAIS areas.
It is worth mentioning that the present bed topography estimations may be applied to any ice-sheets areas (or large ice-caps) as soon as the model assumption is satisfied. The computed thickness estimations are valid at $\approx 10 \times \bar{h}$ wave length ($\bar{h}$ a characteristic thickness value). The present test areas have been selected since they respect the domain of validity of the method; moreover estimating more accurately the bed elevation in EAIS is not without interest. Indeed, global warming may threaten EAIS stability, particularly around some of the present selected areas, see e.g. [33].

The outline of the article is as follows. In Section 2, the Reduced Uncertainty (RU) equation developed in [23,24] is recalled; its domain of validity is highlighted; uncertainty range on its unique unknown dimensionless parameter is derived. Next, the inversion method is detailed. It is developed in three steps: following [24], Step 1) and Step 3) aim at inverting the RU-SIA equation by VDA (that is physical-based inversions). Step 2) is a deep-learning process (ANNK algorithm), that is a purely data-based inversion based on the in-situ data. In Section 3, the six considered EAIS areas (named Ant$p$, $p = 1, .., 6$) are presented; bases of the computations are highlighted; features of Bedmap2 estimations [2] are briefly recalled. Next each of the three computational steps of the inversion method is analysed in detail. In Section 4, the robustness of the inverse method is assessed, in particular its sensitivity with respect to the presence or not of additional flight tracks. These detailed analyses are presented for the two most poorly covered areas. A conclusion is proposed in Section 5. The computed thickness estimations for the four other areas (all being in EAIS) are presented in Appendix.

2. Method

In this section the method, algorithms developed to estimate the ice thickness h are detailed. The estimations are based on both physical model inversions and statistical learning. The physical-based inversions relies on the RU-SIA equation developed in [24] and a VDA process. The data driven inversion is based on a Deep Neural Network (DNN) aiming at estimating the single uncertain dimensionless multi-physics parameter γ of the RU-SIA equation. The global inversion algorithm is built up in three steps; each step combining optimally data and the model. Data employed in the two VDA processes are the surface ones (altimetry, InSAR and climatic term SMB) plus the in-situ thickness measurements. Data employed in DNN are in-situ ones (available along the flight tracks). The employed data are those available in BedMap2 database [2] plus [4] for surface velocities and [34] for the climatic term. The final output of the global inversion algorithm is the ice thickness h, hence the bed topography elevation b.

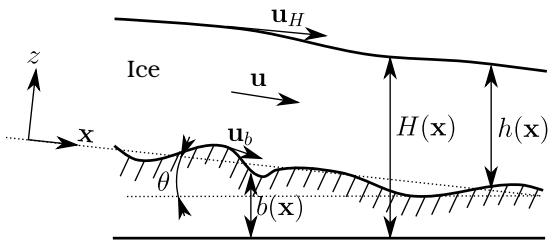

Figure 1. Schematic vertical view of the gravitational ice flow and notations

2.1. The RU-SIA equation

The RU-SIA equation is obtained by reformulating the depth-integrated SIA model with basal slipperiness, see e.g. [35] Chapter 5, with a non constant rate factor and by introducing the surface terms (elevation and velocity), see [23,24]. The resulting 2D depth-integrated model, see [24], is relevant for large scale on sheared flows with some slipperiness at bottom, with non constant vertical temperature profile and if the surface features are measured. The various multi-physics terms (constitutive law exponent, flow regime, temperature dependent term) can be gathered into the single parameter γ. Therefore γ contains all the physical parametrisation uncertainty. The RU-SIA model assumptions are the same as SIA (classical lubrication theory) with basal slipperiness that is: the flow is necessarily (moderately) sheared (normal stress components are negligible) and it is "shallow" (long wave assumption). 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; making this shallow flow models "large scale" models (see the discussion below for quantified minimal wave length).

The surface slope is denoted by $\mathcal{S} = |\nabla H|$; $\mathbf{u}_H$ is the surface velocity and $\mathbf{u}_b$ is the basal velocity.

The depth (ice thickness) is denoted by h, $h = (H - b)$ with H the ice surface elevation and b the bed elevation, see Fig. 1. θ is a potential mean slope value in the (x, y)-plane; $\theta = 0$ in the forthcoming test areas.

The depth-integrated flow model SIA model with slipperiness at bottom (see e.g. [35] Chapter 5) is derived in a non isothermal version in [23] providing the so-called xSIA (x for extended) equation. Moreover in [24] by taking advantage of the measured surface features (elevation and velocity norm), xSIA equation is re-formulated to obtain the RU-SIA equation (RU for Reduced Uncertainty). RU-SIA equation reads, [24]:

$$-div\left(\frac{|\mathbf{u}_H|}{\mathcal{S}}\,\gamma h\,\nabla H\right) = \dot{a} \tag{1}$$

where γ is the dimensionless multi-physics parameter of the model (single parameter if not considering the thickness h). The RHS $\dot{a}$ is the classical one defined by: $\dot{a} = (\partial_t h - a)$ with a the mass balance (accumulation and ablation). The parameter γ reads:

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

with R_s the slip ratio defined by: $R_s = 1 - \dfrac{|\mathbf{u}_b|}{|\mathbf{u}_H|}$. The parameter c_A is defined by:

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

with q the constitutive power-law exponent (q is classically set to 3) and $R_A = \dfrac{\bar{A}}{\underline{A}}$. $\bar{A}$ and $\underline{A}$ are depth-integrated quantities naturally appearing in the xSIA equation if the rate factor A depends on $(\mathbf{x}, z)$. Here A depends on $(\mathbf{x}, z)$ since depending on the non constant temperature vertical profile. The expressions of these parameters are, see [23]:

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

If the vertical profile of A is constant then: $\bar{A}(\mathbf{x}) = A(\mathbf{x}) = \underline{A}(\mathbf{x})\ \forall \mathbf{x}$. Therefore in the (unrealistic) isothermal case, A is a constant and: $\bar{A} = A = \underline{A}$. Moreover $R_A = 1 = c_A$ and for $q = 3$, γ simplifies as: $\gamma = \left(1 - \frac{1}{5}R_s\right)$.
Recall that $R_s = 1$ for no slip at bottom (fully sheared flow) and $R_s \approx 0.5$ for a moderately sheared flow.

In the flow model, $|\mathbf{u}_H|$ and $\mathcal{S}$ can be provided by InSAR and altimetry measurements respectively. Assuming that h (or equivalently b) is given, Eqn (1) in variable H contains γ as the single uncertain parameter despite arising from the integration of the full 3D Stokes free surface model with non constant constitutive law vertical profile.
Surface values H at boundaries of the geometrical domain (measured by altimetry) close this non linear diffusive equation in variable H.

On the RU-SIA equation domain of validity. Let us highlight the domain of validity of Eqn (1). Since it is an asymptotic depth-integrated model (long wave assumption, "shallow"), ε has to be $\lesssim 1/10$, see e.g. [36]. That means a "large scale", equivalently "long wave length", model valid for minimal wave length $L^* \approx 20$ km (resp. 30 km) if $H^* \approx 2$ km (resp. 3 km). Moreover since the basic scalings are those of sheared flows (with or without slipperiness at bottom), the model has a validity range in terms of flow regime too. The flow regime can be characterised by the slip ratio R_s. By construction, the SIA-like models (including xSIA and RU-SIA equations above) are valid for R_s ranging from ≈ 0.5 to 1, see [37–39] for analysis. This rough estimation in terms of R_s is numerically quantified in real world cases (including east Antarctica) in [40]. Indeed, a detailed assessment of the (isothermal, classical) SIA model is proposed in terms of minimal wave length vs surface velocity magnitude. (It can be assumed that the flow regime can be interpreted in terms of surface velocity norm too). This study [40] is based on the so-called MCL criteria (criteria proposed in [41] and defined as the length scale over which the terms of driving stress and drag are comparable). In particular it can be noticed in [40] 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; therefore showing the validity of the upper bound $\varepsilon \lesssim 1/10$. (On the contrary, [40] numerically confirms that faster ice streams are wrongly modelled by these SIA-like models since not including longitudinal stresses).

The six test areas Ant$p$ of the present study, see Fig. 4, have been selected upon their surface velocities; all of them ranging in $\approx [5-90]$ m/y. Since $H^* \approx 2-3$ km, this makes the RU-SIA equation accurate at wave length $L^* \approx 20-30$ km.

A link between the inversion of the mass equation and the inversion of RU-SIA equation. As already mentioned, [11,15,16] aim at infering h by inverting the mass equation $div\,(h\bar{\mathbf{u}}) = \dot{a}$ with $\bar{\mathbf{u}}$ the depth-averaged velocity. $\bar{\mathbf{u}}$ is related to $\mathbf{u}_H$ by: $\bar{\mathbf{u}} = \tilde{\alpha}\mathbf{u}_H$ with $\tilde{\alpha}$ empirically set; $\tilde{\alpha} \lesssim 1$ in fast flows (where this transport equation may be inverted in highly covered areas with cross-lines flight tracks). In the derivation of the RU-SIA equation, see [24], it is shown that: $\bar{\mathbf{u}} = -\frac{|\mathbf{u}_H|}{\mathcal{S}}\gamma\nabla H$. Therefore the empirical parameter $\tilde{\alpha}$ set in [11,15] is nothing else than the dimensionless multi-physics parameter γ: $\tilde{\alpha} = \gamma$ defined by (2). In plug like regimes (fast ice streams), $R_s \sim 0$ and the uncertainty on the internal deformation represented here by the parameters c_A and q is negligible.

2.2. Typical uncertainty on the multi-physics parameter γ

The a-priori estimations on the uncertain parameter γ, see (2), presented in [24] are recalled. Let us recall that the definition of γ contains multi-physics parameters: the constitutive law uncertainty (the exponent q), the vertical temperature profile (through the rate factor $A(z)$) and the flow regime (the slip ratio R_s). For a sake of simplicity, q is set to the widely employed value for glaciers flows, that is $q = 3$.
In the (unrealistic) isothermal case, $\gamma = (1 - 0.2\,R_s)$. Moreover since $R_s \in [\approx 0.5, 1]$ in the targeted areas, then the uncertainty on γ (fully due to the flow regime uncertainty in this case) would be relatively small, $\approx 10\%$ only.
In the real world, following the Arrhenius law and values, see e.g. [35, p.54], by considering typical ice-sheet vertical temperature profiles, [42,43], see also e.g. [44,45], the following vertical profile of $A(z)$ is considered, see Fig. 2 (Left):

$$A(z) = \begin{cases} A_a & \text{for } z \in [B, H] \\ \frac{A_a}{B-b}\left((1-k)z + kB - b\right) & \text{for } z \in [b, B] \end{cases} \tag{5}$$

with A_a and k given constants. Let us define: $B = (mh + b)$ with $m \in [0, 1]$. This particular vertical profile of A presents a boundary layer at bottom of thickness $(B - b) = mh$ varying linearly, Fig. 2 (Left). This schematic profile represents realistic ones present in east Antarctica, see e.g. [42,43]. The value of c_A vs m for different values of k is presented in Fig. 2 (Right). The case $m = 0$ corresponds to the isothermal case, $c_A = 1$. For relatively thin thermal boundary layer, c_A increases with mh; then c_A decreases again up to its minimal value $C_A^{min} < 1$ (minimal value obtained for the purely linear vertical profile, $m = 1$).

Let us consider typical temperature values in east Antartica: the bed at $0\,C^\circ$ and the surface at $-40\,C^\circ$. These values correspond to $A_a \approx 10^{-26}$ hence $k \approx 1000$, see Fig. 2 (right). (In Fig. 2, the value $k \approx 10$ would correspond to typical inland Greenland cases, see [24]). Moreover, assuming that the thickness boundary layer is: $m \in [0.1, 0.5]$; then: $c_A \in (3.11, 4.64)$, see Fig. 2. It follows from (2) that: $\gamma \approx [1 - (0.78 \pm 0.15)R_s]$.
The present basic uncertainty analysis based on typical values in the targeted areas

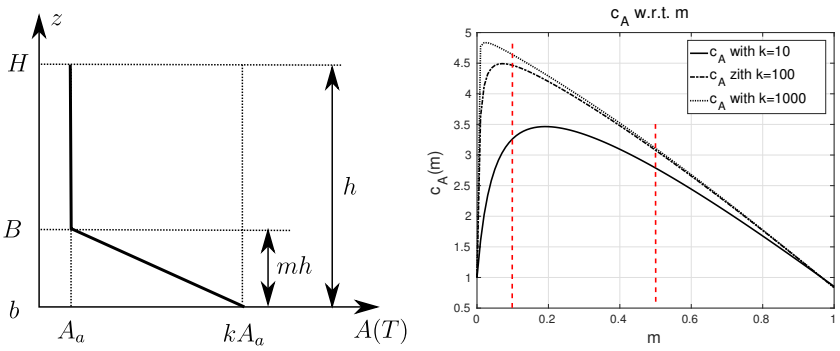

Figure 2. (Left) Typical vertical profile of rate factor $A(z)$, see (5). (Right) The parameter c_A vs m, see (3), with $k = 10, 100$ and 1000.

shows that the uncertainty on γ come similarly from the vertical thermal profile uncertainty (represented by the term $(c_A/(q+2))$) and the slip ratio R_s. Moreover it shows that the present vertical profiles, see (5) and Fig. 2 (Right), in the targeted regimes, would make vary γ within the interval $\approx]0, 0.7]$.
By relaxing the assumption of the imposed vertical profile (5), it seems reasonable to estimate the upper bound of γ by setting $c_A = 1$ and $R_s = 0.5$ which gives: $\gamma \in]0, \approx 0.9]$.
This multi-physics parameter γ is numerically infered following the inversion algorithm presented in next section. As it will be shown, if not imposing any bound on the infered values, the infered values will be obtained within the present a-priori estimated interval, excepted at very local locations where some pick values are reached, see next section.

2.3. The inversion method

The ice thickness h is inferred by a method developed in three steps. Step 1) and Step 3) are physical-based inversions: the RU-SIA equation (1) is inverted by VDA. (These inversions are the ones proposed and assessed in [24]). Step 2) is based on ANN algorithm; it is a purely data-driven inversion.

2.3.1. Sketch of the global inversion algorithm

The three steps of the inversion method are as follows.

`Step 1)` Estimation of the effective diffusivity $\eta = (\gamma h)$ in RU-SIA equation (1) by VDA. Given the surface measurements H and $|U_H|$, the effective diffusivity $\eta = (\gamma h)$ in (1) is infered by solving the following optimal control problem:

$$\min_{\mathbf{k}} \; g(\mathbf{k}) \quad \text{with } g(\mathbf{k}) = g_{obs}(\mathbf{k}) + \alpha g_{reg}(\mathbf{k}) \tag{6}$$

With: $\mathbf{k} \equiv \eta = (\gamma h)$,

$$g_{obs}(\cdot) = \frac{1}{2} \int_{\Omega} |H(\cdot)(\mathbf{x}) - H^{obs}(\mathbf{x})|^2 \chi_{tr}(\mathbf{x}) d\mathbf{x}, \tag{7}$$

χ_{tr} the restriction operator to the flight tracks, $g_{reg}(\cdot)$ a Tykhonov's regularization term, see e.g. [46]. In this step, it is defined as: $g_{reg}(\eta) = \frac{1}{2}\int_\Omega |\nabla\eta(\mathbf{x})|^2 d\mathbf{x}$. The weigh coefficient α is empirically set "at best". The surface elevation H_k satisfies the RU-SIA equation (1) (with Dirichlet boundary conditions). The gradient of the cost functional is computed by introducing the adjoint equation. The minimisation algorithm is a quasi-Newton method (L-BFGS algorithm from Python routine scipy.optimize.minimize). We refer to [24] for more details. The iterative minimization process is performed until convergence. Numerous numerical experiments demonstrate the convergence is very robust; see [24] for a thoroughly assessment of the method. In particular the optimal solution does not significantly depend on the smoothing length scale of the surface data (done here at $\approx$ 24 km , see next Section), nor on the first guess (being here h_b the BedMap2 value).
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) denoted by γ^*_{tr} are straightforwardly deduced: $\gamma^*_{tr} = \frac{\eta^*}{h_b}\chi_{tr}(\mathbf{x})$. These values are inputs of the next algorithm, Step 2).
Recall that all these values are valid (and accurate, see [24]) at the flow model scale, that means at $\approx$ 25 km minimal wave length; with values computed on $\approx$ 2 km grids.

`Step 2)` Extension of γ in the whole area by NNRK.
Given γ^*_{tr} (result of Step 1)), a NNRK algorithm, see e.g. [31,32], is developed to extend values of γ along the flights tracks to the whole area. The NNRK algorithm is a purely statistical learning algorithm; it is decomposed in few steps: an ANN estimator (deep learning) and an ordinary Kriging of the residuals. Details are presented in 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 infered simultaneously with the RHS $\dot{a}$ in (1) by an advanced formulation of VDA. The equations and the complete VDA algorithm are presented (and assessed) in detail in [24]. Let us recall them briefly.
Similarly to Step 1), the pair $(h, \dot{a})$ in (1) is infered by solving the optimal control problem (6) with g_{obs} defined by (7) but minimizing with respect to $\mathbf{k} = (h, \dot{a})$ (and not w.r.t. $\eta = \gamma h$ like in Step 1)). In (7) the surface elevation H_k satisfies (1) (plus Dirichlet boundary conditions).
In the present VDA formulation, the regularization term reads:

$$g_{reg}(h, \dot{a}) = \frac{1}{2}\|(h - h_b)\|_{C_h^{-1}} + \frac{1}{2}\|\dot{a} - \dot{a}_b\|_{C_a^{-1}} \tag{8}$$

with C_h^{-1} and C_a^{-1} covariance operators defining metrics. 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 [24] for details (and [47] for a method analysis in another physical context). Next a change of the control variable is made, see [24,47]. 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, see

[24].
In (6) the weight coefficient α is defined as a decreasing sequence following an iterative regularisation strategy, see [24] for details and [46] for a method analysis. This iterative regularisation strategy has demonstrated to improve the robustness and the convergence speed of the VDA algorithm too, [24,46].

The robustness of the present inversion algorithm is thoroughly assessed in [24], in particular the sensitivity of the present 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 $\approx$ 24 to 48 km; iv) the first guess (chosen here as the BedMap 2 value h_b).

Remark 1. It is shown in [24] that the explicit expression (2) of the multi-physics parameter γ enables to compute *a-posteriori* estimations of the (spatially distributed) slip ratio value R_s. This is a very interesting feature by making possible to confirm or not the good RU-SIA model consistency. In the test area considered in [24] (area included in the present Ant6, see Fig. 4), the a-posteriori analysis made in [24] remarkably confirms the good consistency of RU-SIA model.
Moreover based on some a-priori vertical thermal profile(s) e.g. the one defined by (5), RU-SIA equation (1) provides *a-posteriori* estimations of the effective thermal boundary layer thickness $(B-b)$, see Fig. 2 (Left). To our best knowledge such a-posteriori estimation of the vertical profiles is original; it may be interesting for various analyses. Moreover the considered a-priori vertical profiles could be adjust by constraining them with (the very few) in-situ measurements.

2.3.2. The Neural Network Residual Kriging (NNRK) algorithm (Step 2))

The NNRK algorithm is decomposed into the few following steps.

`Step 2a)` Considering all surface data (H, $\mathcal{S}$, $|\mathbf{u}_H|$, $\dot{a}_b$) along the flights tracks (in all areas Ant$p$) plus γ^*_{tr} computed at Step 1) (that is the "true" value), an estimator of γ is computed by an Artificial Neural Network (ANN) algorithm (deep learning), see e.g. [31,32]. This estimator is denoted by $\bar{\gamma}$.
The complete data sets along the flight tracks constitute the training data set; it is denoted by $\mathcal{D}$. $\mathcal{D}$ contains examples (I_i, O_i), $i = 1, \cdots, N_{ft}$, where $I_i = (H, \mathcal{S}, |\mathbf{u}_H|, \dot{a}_b)(\mathbf{x}_i)$ is the i-th input and $O_i = \gamma^*(\mathbf{x}_i)$ is the i-th expected output; N_{ft} is the number of examples in the training data set. The space coordinates $\mathbf{x}_i$, $i = 1, \cdots, N_{ft}$, are along the flight tracks.
The estimator $\bar{\gamma}$ is computed as the minimizer of the mean square misfit $\frac{1}{N_{ft}} \sum_{i=1}^{Nft} (O_i - \bar{\gamma}(I_i))^2$. This misfit (error function also called empirical risk) reads:

$$j(\mathcal{D}; \cdot) = \frac{1}{N_{ft}} \sum_{i=1}^{N_{ft}} [\gamma^*_{tr}(\mathbf{x}_i) - \bar{\gamma}(H, \mathcal{S}, |\mathbf{u}_H|, \dot{a}_b)(\mathbf{x}_i)]^2 \quad \text{for } \mathbf{x}_i \in \Gamma_{tr} \tag{9}$$

To solve this large dimensional optimisation problem (purely data-based), the very likely most efficient methods are ANNs with a few hidden layers (deep learning processes). Here, each of the 4 hidden layers contains 50 neurones, see

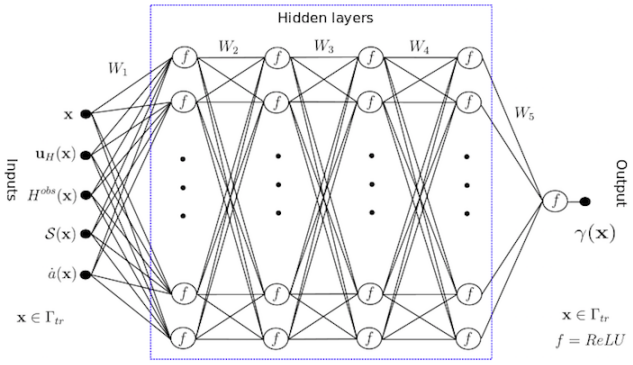

Figure 3. Artificial Neural Network (ANN) with four hidden layers: the first step of NNRK algorithm (Step 2a) of the global inversion algorithm)

Fig. 3. The most efficient activation function is chosen: the rectified linear unit (ReLU) function, see e.g. [48,49].The ANN is completely determined by the weight parameters $(W_1, \cdots, W_5)$, Fig. 3.
The training step consists to identify the optimal values of these parameters $W_j,\ j = 1, .., 5$. W_j are 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 training step is programmed using PyTorch, a deep learning framework in Python.
To minimize $j(\mathcal{D}; (W_1, \cdots, W_5))$ with respect to $\{W_j\}_j$, the classical Adam method [50], a first-order gradient-based stochastic optimization, is employed. The learning rate (the gradient descent step size) is classically adjusted during the optimization procedure. Also to avoid overfitting, the dropout method [51] is adopted. 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 reader may refer e.g. to [32] for more details and know-hows on these ANN and NNRK methods.

Let us point out that this NNRK algorithm has been trained by including h_b as an input parameter or not. The estimator taking h_b as an input parameter turned out to be very similar to the one without; it diminishes the cost value j very slightly only. This test strengthens our strategy to predict the dimensionless parameter γ of the flow model from the surface features only.

`Step 2b)` The K-fold cross-validation, see e.g. [52], is applied to assess the ANN model and analyse if the ANN can be generalized to independent data sets or not. Let us recall that K-fold cross-validation proceeds as follows, see e.g. [52]:

- Divide randomly the original training data set $\mathcal{D}$ into $K \in \mathbb{N}^+$ roughly equal parts;
- For each subset $\mathcal{D}_k$, $k = 1 \cdots K$, ANN trains from the other $(K-1)$ subsets $\mathcal{D}_i,\ i \neq k$. For $i = 1, \cdots, K$, denote $\mathcal{D}i_{test} = \mathcal{D}_i$ and $\mathcal{D}i_{train} = \cup_{j \neq i} \mathcal{D}_j$;
- Compute the error over each test set $j(\mathcal{D}i_{test})$, and choose the ANN providing the smallest total (training plus test) error, i.e. $j(\mathcal{D}i_{test}) + j(\mathcal{D}i_{train})$.

`Step 2c)` The residual value along the flights tracks is defined by: $\varepsilon_\gamma = (\gamma^*_{tr} - \bar{\gamma})$ with $\bar{\gamma}$

computed by ANN. An ordinary Kriging (with a spherical semi-variogram model) is performed to extend ε_γ all over the area; the obtained estimator is denoted by $\hat{\varepsilon}_\gamma$. By construction this residual satisfies: $\mathbb{E}(\varepsilon_\gamma) \approx 0$; also 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. [32] Chapter 3.
The final estimation in the whole area is denoted by $\hat{\gamma}$; it is obtained as the sum of the ANN estimation and the ordinary Kriging estimation of residuals :

$$\hat{\gamma}(\mathbf{x}) = \bar{\gamma}(\mathbf{x}) + \hat{\varepsilon}_\gamma(\mathbf{x}) \quad \text{for } \mathbf{x} \in \Omega_p \tag{10}$$

ANNs enable to find nonlinear trends between the data. Here the resulting estimator provides the dimensionless parameter γ (2) of RU-SIA flow model (1) in function of the surface data (altimetry, InSAR and $\dot{a}$) for each $\mathbf{x}$ in Ω_p. The inversion algorithm thoroughly analysed in [24] is the same as the present one excepted Step 2) which is an ordinary Kriging instead of NNRK. The difference between the resulting two estimators is analysed in next section. ANNs are very CPU-time and memory consuming, then the code has been written using the Python lib Mpi4py [53].

2.3.3. On the linked uncertainty between γ and h

In the global inversion algorithm described above, once the product $\eta = (\gamma h)$ accurately infered at Step 1) it remains to separate the effects, the surface signatures, of the two unknown fields γ and h. The parameter γ models the vertical profile by integrating the temperature vertical profile, the basal slipperiness and the constitutive law. Its computed value along the flight tracks at Step 1) is extended all over the domain by a data-driven model (statistical learning based on the airborne data sets). Next the thickness h is accurately computed / infered but based on the previous estimation of γ, see Eqn (1).
Let us denote by the superscript ^{t} the true (effective) value of the fields. If assuming that each computational inverse algorithm is accurate and robust (this can be assumed to be correct, see [24]) then it can be assumed that: $\gamma^t h^t \approx \gamma^* h^*$, where the * fields are the optimal values computed by the successive algorithms of Step 2) and Step 3). By defining $\delta h = (h^* - h^t)$ and $e_h = \delta h / h$ (and the equivalent quantities in γ if γ does not vanish), it follows that (at order 1):

$$e_h \approx -e_\gamma \text{ (where } \gamma \text{ does not tend to 0)} \tag{11}$$

In other words, Step 2) and Step 3) of the global inversion algorithm may propagate the error made on γ to h in the same order of magnitudes in % (but reversely). The computed fields at each step are analysed in next section.

In summary, the product (γh) is identified all over the domain; VDA techniques well controlled can provide accurate results. Next, the challenge is to separate the depth h from γ, γ the unique dimensionless (multi-physics) parameter of RU-SIA flow model. Then the following bet is made: given all surface features of the flow $(H, \mathcal{S}, |\mathbf{u}_H|, \dot{a})$, an ANN enables to provide an accurate estimation of γ all over the area. Deep learning algorithms such as the present one have surprisingly demonstrated to be extremely accurate in numerous applications including in physics; this is the case here: the assess-

ment made on the various $\mathcal{D}i_{test}$ sets demonstrate the high accuracy of the builded estimator. Finally given the estimation of γ, the advanced VDA process enables to provide reliable values of the ice thickness h.
Observe that it would be straightforward to use the same NNRK algorithm to directly estimate the thickness h all over the domain. However it seems definitively more consistent to estimate an unique dimensionless parameter of flow model from the surface features represented by this same model instead of an input data partially dependent only. Following this idea of purely data-driven estimations, [54] 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. Algorithm steps analysis

The inversion algorithm to estimate the bed elevation previously presented can be applied to any ice-sheets areas as soon as the RU-SIA model assumptions are satisfied. These assumptions are: a) a moderately sheared flow (in Antarctica this corresponds to surface velocities ranging approximatively in $[5, 100]$ m/y); b) if estimating the thickness at $\approx 10 \times \bar{h}$ wave length ($\bar{h}$ a characteristic thickness value). For a sake of detailed method analysis, here a few areas only are considered, see Fig. 4; all of them are in East Antarctica. The 6 considered test areas are denoted by Ant$p$, $p = 1, .., 6$, see Fig. 4. All of them are large upstream areas of major ice-sheds in East Antarctica; they have been selected since they respect the method domain of validity. They are relatively large (ranging from 250268 to 439045 km^2) and thick (the mean thickness value of h_b ranging from 1822 to 2745 m), see tables 2-A4 for more details. The exact coordinates characteristic points defining each area ($\approx 100 - 150$ points per area), see Fig. 4, are available on the open source computational software webpage[2]
It is worth to recall that estimating more accurately the bed elevation in these Ant$p$ areas is not without interest since global warming may threaten East Antarctica ice sheet stability as mentioned e.g. in [33].

3.1. Data and minimal wave length

In this paragraph, we describe the pre-treatment done on the surface data, in link with the long wave assumption of RU-SIA equation (1) (shallow flow model), and the numerical grids. Moreover we briefly recall how the BedMap2 ice thickness value h_b is obtained in [2] and how its uncertainty (as estimated in [2]) is generated.

3.1.0.1. Surface data smoothing & numerical grids.. The mean value of the ice thickness provided by BedMap2 [2] in the 6 Ant$p$ areas equals $\approx$ 2.7 km. As previously discussed, the RU-SIA equation is sufficiently accurate as soon as $\varepsilon = \frac{[H]}{[L]} \lesssim 0.1$. Therefore surface data $|\mathbf{u}_H|$ and H provided in databases need to be smoothed at the minimal wave length $L^* \approx 27$ km. To do so, a Gaussian with standard deviation $\sigma = 4$ km is applied to smooth the surface data (both to the

[2] J. Zhu. 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.*

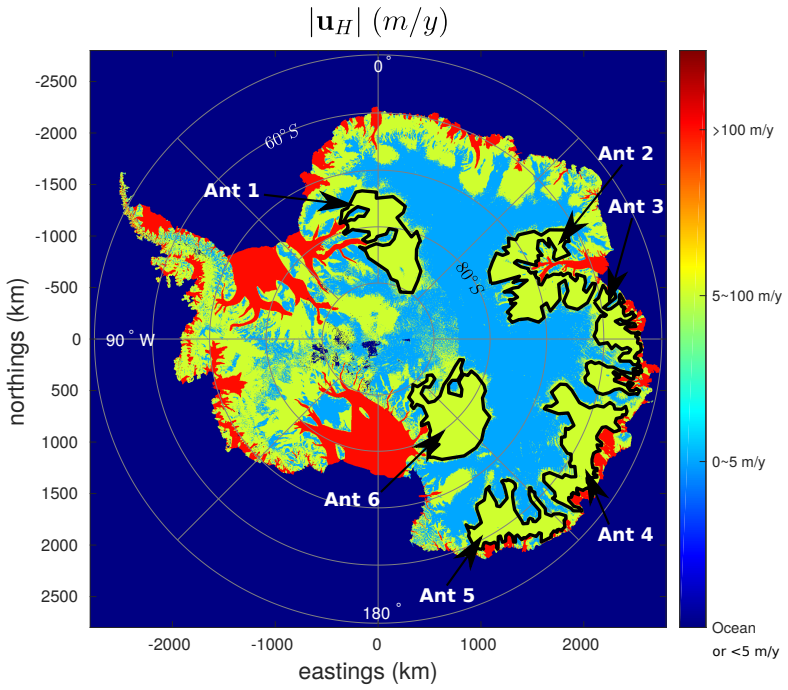

Figure 4. Location of the 6 test areas Ant$p$ (east Antarctica) with InSAR-based surface velocity values in m/y (from [4]).

elevation and the velocity norm). Therefore the smoothing effects are sensitive in disks of $\approx 2 \times 3\sigma = 24$ km of diameter.
Next the mesh is built up using Gmsh software [55] with a grid size $\delta x \approx 3$ km. Indeed $\delta x = 3$ km provides ≈ 10 points per minimal wave length L^*, therefore respecting the minimal number of points to properly approximate all fields. The lineic flight tracks are meshed with cells of $\delta x \approx 2$ km size. Along these flight tracks the imposed thickness measurements are interpolated values of the BedMap2 values [2].

3.1.0.2. On the reference bed topograghy Bedmap2.. Let us recall that for each VDA process of the present inversion algorithm, that is Step 1) and Step 3) (see paragraph 2.3.2), the first guesses are the Bedmap2 ice thickness value h_b. h_b and its a-priori uncertainty as derived in [2] are plotted in each test area Ant$p$, see figures 6, 9, A1, A2, A3 and A4 (Middle). The present thickness estimations h^* are compared to h_b, see the same figures but (Down)(R).
In BedMap2 database [2], the interpolation - extrapolation of airborne measurements are performed throughout the domain by the ArcGIS Topogrid routine (ESRI Ltd, ArcGIS 9); the latter is based on the ANUDEM algorithm [7]. This algorithm uses an iterative finite difference interpolation technique which is essentially a thin plate spline technique [8]. Next, empirical uncertainty values are stated as follows, see [2]. The thickness measurements are split into two datasets $(D1)$ and $(D2)$. Dataset $(D1)$ is used to build up an interpolation including at Dataset $(D2)$ location points; values of $(D2)$ being not used at this stage. Next, Dataset $(D2)$ is used to quantify the misfit with the "predicted" - interpolated values; and basic statistics on the results with dependence on the distance to data are deduced. For cells located between 5

and $20\,km$ from any data, [2] suggests that the interpolation error is an increasing function of distance from the closest data; beyond it would be not correlated. (Observe that this distance corresponds approximatively to the minimal wave length of the RU-SIA model). For cells that are more than 50 km from airborne measurements, the thickness estimation is based on gravity-field inversion (gravity-derived thickness); the proposed related uncertainty equals $\pm 1000\,m$. This is how the uncertainty values on Bedmap2 values h_b plotted in figures 6-A4 (Middle) (Right) are established.

3.2. Step 1: Estimation of η by VDA

The effective diffusivity η defined in (1), $\eta = (\gamma h)$, is estimated in each area Ant$p$ by VDA following Step 1), see 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; it has been thoroughly assessed in [24]. The stopping criteria is the stationarity of $\|\eta\|$. Here the RHS $\dot{a}$ provided by [34] is supposed to be exact. After convergence of the VDA process, given the ice thickness along the flights tracks, the optimal value γ^* is saved for Step 2), that is:

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

3.3. 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 Ant$p$, $p = 1, .., 6$ plus the values of γ^*_{tr} 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 of course completely independent of the 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; 1 example being constituted by 1 (input, output) pair of ANN.
It can be read on Fig. 5 (Left) that all ANN models trained from different training data sets $\mathcal{D}i_{train}$ have very close cost values $j(\mathcal{D}i_{train})$ defined by (9). Moreover, the cost values of each test sets $j(\mathcal{D}i_{test})$ are almost equal showing that all ANN models have very close prediction capability, all being excellent: after optimisation, $j(\mathcal{D}i_{test}) \approx 5\ 10^{-3}$ (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 very slightly smaller misfit value $(j(\mathcal{D}i_{train}) + j(\mathcal{D}i_{test}))$. The tests of prediction capabilities above demonstrate the robustness and accuracy of the trained ANN.
By performing the selected ANN ($\mathcal{D}2_{train}$ since being slightly more accurate), values of $\bar{\gamma}$ are computed in the whole domains. Next by performing the Kriging step (Step 2c) of NNRK algorithm, see paragraph 2.3.2), $\hat{\gamma}$ defined by (10) is obtained. Moreover following the a-priori estimation derived in Section 2.2, the upper bound 0.9 is imposed to the estimation.
For each test area, $\hat{\gamma}$ is plotted, see figures 6, 9, A1, A2, A3 and A4 (Up) (R). It is worth to notice that the imposed upper bound on $\hat{\gamma}$ (here 0.9) is active at very few locations only; moreover it is in great majority where the uncertainty on h_b is low, see figures 6, 9, A1, A2, A3 and A4 (Middle) (R).

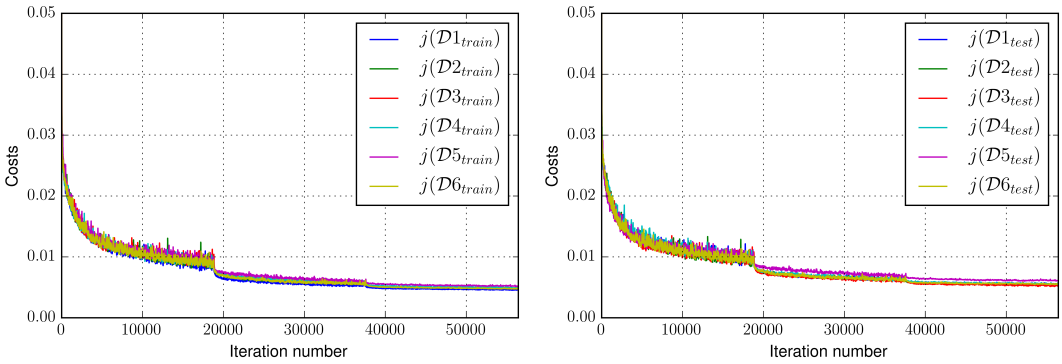

Figure 5. Misfit functional $j(\mathcal{D})$ defined by (9), vs minimization iterations for: (Left) different train sets $\mathcal{D}i_{train}$; (Right) different test sets $\mathcal{D}i_{tests}$. The learning rate (gradient descent step size) is decayed by 0.2 every 30 epochs (1 epoch = 1 forward pass and 1 backward pass of all the training examples), see e.g. [32] for details on this classical method.

Table 1. Difference between values of H the responses of RU-SIA equation (1) if changing the RHS value: $\dot{a}_b$ the Racmo2 value vs $\dot{a}^*$ the infered value at Step 3). Domain Ω_p=Ant1.

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

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

Once Step 2) is done, the dimensionless multi-physics parameter γ in RU-SIA equation (1) is given. Next the thickness h is infered 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: it has been thoroughly analysed in [24].

At this step, h is simultaneously infered with $\dot{a}$. Indeed this strategy enables to adjust the value of $\dot{a}$ within its uncertainty range, see [34]. Recall that this uncertainty is relatively low: $\sim \pm 20\%$. It can be noticed in tables 2, 5, A1, A2, A3 and A4 that the corrections made by the VDA stay in great majority within this a-priori uncertainty range. Indeed the infered values reach the authorised maximal variation at few locations only. In other words, this adjustment based on the physical-based model RU-SIA is consistent with the estimations proposed in [34] which is totally independent.

Of course, given the surface data, any change of $\dot{a}$ in RU-SIA equation (1) has an effect on the infered 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 infered 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 from Bedmap2 value h_b, see Tab. 2 in next section. This remark holds for all domains.

As already mentioned, the RU-SIA equation (1), the two VDA processes (i.e. Step 1) and Step 3)) and the present global inversion algorithm but with an ordinary Kriging at Step 2) have been assessed in detail in [24] on a test area included in Ant6. This thorough analysis focuses in particular on the robustness and the sensitivity of the algorithm(s) with respect to: i) the uncertainty on γ; ii) the proximity or not of

flight tracks; iii) the smoothing length scale of the surface data; iv) the first guess value (chosen here as BedMap 2 h_b).

3.5. On the RU-SIA model accuracy

For each test area Ant$p$, domain information and basic statistics on the results are presented, see tables 2, A1, 5, A2, A3, A4. Statistics on the computed surface elevation H as the output of RU-SIA model (1) are indicated. It can be noticed that the RU-SIA equation solved from Bedmap2 value h_b and the data-driven estimation $\hat{\gamma}$ (result of Step 2)) already fits with the measured surface elevation H^{obs}, see "Direct model validation" lines in the tables. This excellent accuracy based on the Bedmap2 bed elevation h_b (without any additional calibration of h) demonstrates the relevance and the validity of the physical-based RU-SIA model. Past Step 3) of the inversion algorithm (i.e. after the identification of h and $\dot{a}$ by VDA), of course the RU-SIA model fits even better with H^{obs}, see "$|H(h^*) - H^{obs}|$ (after h-inversion)" in the tables.

4. Results and sensitivity tests (Ant1 and Ant3 areas)

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

- `a)` some isolated (hence highly constraining) flight tracks are considered or not;
- `b)` the statistical - machine learning method of Step 2) is the NNRK algorithm described in paragraph 2.3.2, or an ordinary Kriging.

These comparisons aim at analysing the robustness and accuracy of the method.
Ant1 is a 370809 km^2 area north-east upstream of Bailey, Slessor and Recovery ice-streams; Ant3 is a 250268 km^2 area in Wilhelm and Queen Mary lands, upstream of Shackleton ice shelf and Davis sea.
Among the considered areas, Ant1 and Ant3 are those presenting the largest uncovered parts during airborne campaigns; they contain large areas where Bedmap2 estimation h_b is based on gravity field inversions, therefore presenting very large uncertainties.
For each test area, the domain information and basic statistics on the numerical results are presented in tables 2 and 5. The most relevant fields are plotted in figures 6 and 9: the surface velocity module $|\mathbf{u}_H|$ and the flight tracks locations at (Up)(L), the NNRK estimation $\hat{\gamma}$ defined by (10) at (Up) (R), the Bedmap2 value h_b at (Middle)(L), the uncertainty on h_b as presented in [2] at (Middle)(R), the present thickness estimation h^* at (Down)(L) and its difference with h_b at (Down)(R).

4.1. Results for Ant1 area

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

Table 2. Domain Ω_p=Ant1, information and results.

Domain Ω_p & mesh information			
Surface $\lvert\Omega_p\rvert$	$370809\,km^2$		
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
$\lvert H(h_b) - H^{obs}\rvert$ (before h-inversion)	$7.0\,m$	$10.3\,m$	$53.0\,m$
$\lvert H(h^*) - H^{obs}\rvert$ (after h-inversion)	$2.6\,m$	$3.7\,m$	$44.6\,m$
Infered RHS $\dot{a}$			
$\lvert\dot{a}^* - \dot{a}_b\rvert$	$0.7\,cm/y$	$0.7\,cm/y$	$3.6\,cm/y$
$\lvert\dot{a}^* - \dot{a}_b\rvert/\lvert\dot{a}_b\rvert$	19.2 %	17.0 %	20.0 %
Infered thickness h			
$\lvert h^* - h_b\rvert$	$275.2\,m$	$356.8\,m$	$1953.5\,m$
$\lvert h^* - h_b\rvert/\lvert h_b\rvert$	10.4%	13.3%	65.7%
Ice volume change in km^3 / in %	$1.6\ 10^4\,km^3$ / 1.6%		

4.1.1. Ant1: the ice thickness estimation h^*

Let us recall that $\hat{\gamma}$ is the NNRK estimated value of the multi-physics parameter γ defined by (2).No correlation is observed between h and γ; the only clearly observed correlation is : γ is small where $\lvert\mathbf{u}_H\rvert$ is small, Fig. 6 (Up). This observation is fully consistent with the a-priori analysis done in Section 2.2, see (2) and Fig. 2.
The estimated ice thickness h^* is much less smooth than the Bedmap2 estimation h_b, see Fig. 6. Recall that h_b values are thin plate spline based estimations (see [2] and paragraph 3.1), hence intrinsically smooth. The present estimation h^* is the optimal solution of a data-driven model combined with a physical-based model. It is remarkable that the difference between h^* and h_b is uncorrelated to the distance to the nearest flight tracks, on contrary to the empirically stated uncertainty in [2]. Indeed large corrections of h_b (up to 1500 m) are found close to flight tracks; close meaning at $1-2$ minimal wave lengths of RU-SIA model ($\sim 20-40$ 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) and along these segments (2 km long), the infered depth value is authorised to vary of ±150m around the measured values (inequality constraints imposed in the VDA processes). Therefore in the adjacent triangles (which are nearly equilateral with 2km sides), the plotted mean depth values may be already much different than the measured ones; hence the potential great differences between the two estimations even 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 with the surface data 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: ±20%.
Finally the correction made on h_b is noticeable: 13.3% (356.8 m) in mean, with a 1.6% ($1.6\ 10^4\,km^3$) of volume change only (for 370809 km^2).

4.1.2. 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}.

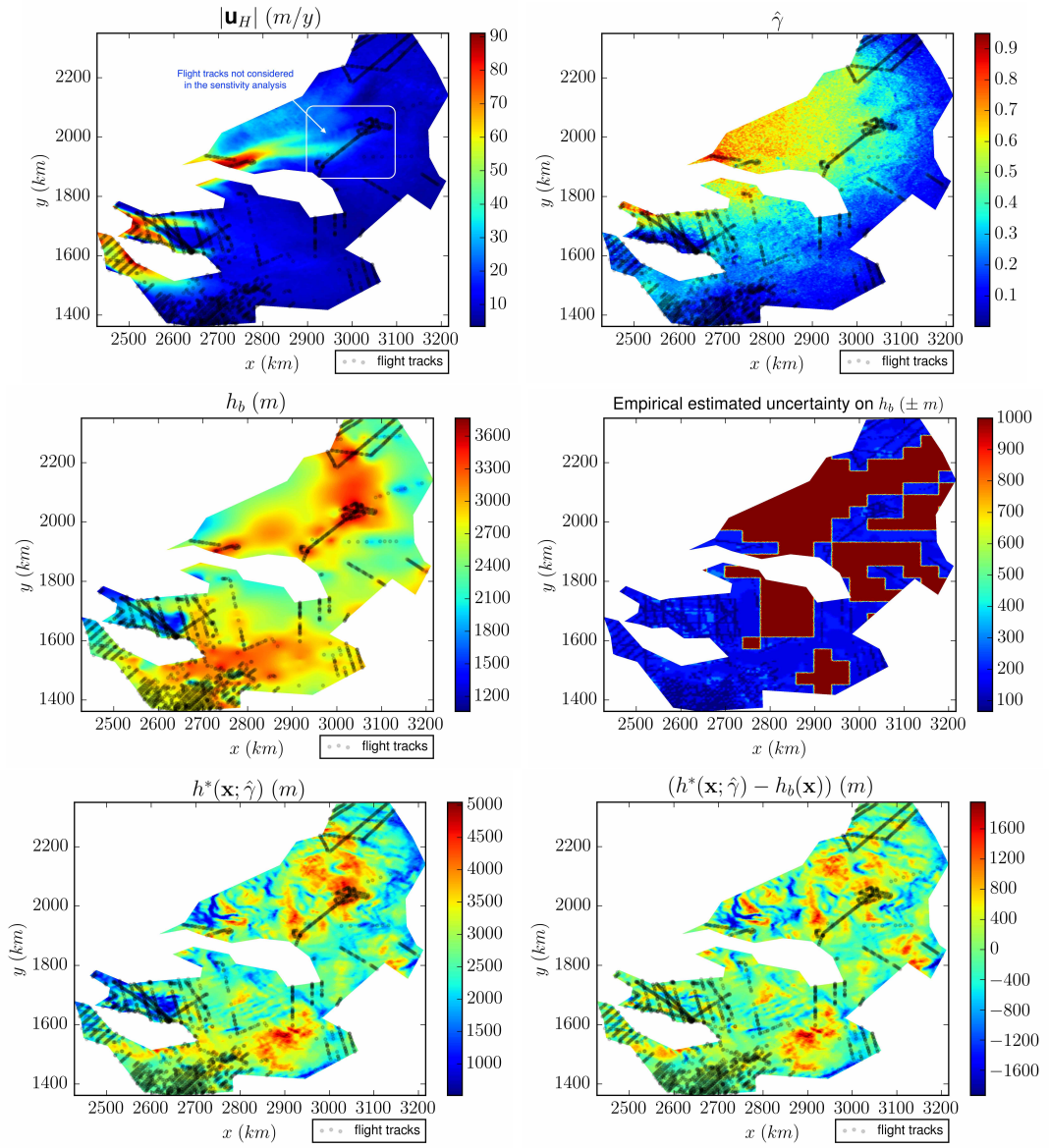

Figure 6. Domain Ω_p=Ant1: (Up)(L) Surface velocity module $|\mathbf{u}_H|$ and flight tracks (R) $\hat{\gamma}$ computed by NNRK, see (10). (Middle)(L) Thickness h_b from Bedmap2 [2] (R) Empirical uncertainty on h_b from [2]. (Down)(L) Infered thickness with $\hat{\gamma}$: $h^*(\hat{\gamma})$ (R) Difference $(h^*(\hat{\gamma}) - h_b)$.

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

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

In Γ^{less} case, the mesh of the entire area has to be re-builded since flight tracks are meshed. The inverse problem is less constrained. This is particularly true for the two VDA processes (steps 1 and 3) since the removed flight tracks are isolated, see Fig. 6(Up)(L), and no constraint is imposed anymore in the vicinity of these removed flight tracks.
The statistical learning at Step 2) is unchanged, therefore values of $\bar{\gamma}$ are unchanged too. However the estimation $\hat{\gamma}$ defined by (10) 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 $(\bar{\gamma}(\Gamma^{all}) - \hat{\gamma}(\Gamma^{less}))$ is plotted in Fig. 3 (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 discrepancies; hence the global change of $\hat{\gamma}$.
Next the infered 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. 3 (Down)(R). For sake of readability, the legend in Fig. 7(Down)(R) has been bounded at $\pm$400m (very few locations were greater than this bound). Basic statistics on the difference are presented in Tab. 3. Note that differences of 300 m correspond to $\approx 10-15\%$ of change. As expected, see e.g. (11), the variations of h are correlated to the variations of γ: compare Fig. 7 (Up)(R) to (Down)(R). And for the reason mentioned above (discussion on the change of γ values), the change of h is not particularly high around the missing flight tracks.
Finally, it is worth to mention that the present inversion method is relatively global with local constraints (the flight tracks); it not purely local inversions. In the present experiment, the obtained variations of h^* are roughly half than 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.

4.1.3. Ant1: with different statistical learning methods at Step 2)

It has been shown in Paragraph 2.3.3 that uncertainties on γ computed by the statistical learning method at Step 2) of the algorithm generates uncertainties on h potentially of same order of magnitude, see (11). Below the influence of the statistical estimator chosen at Step 2) is investigated. To do so, firstly we compare:

- `a)` $\bar{\gamma}$ obtained by ANN algorithm at Step 2a) to $\hat{\gamma}$ obtained by the complete NNRK algorithm, see (10);
- `b)` γ_{krig} obtained by an ordinary Kriging of the thickness measurements along the flights tracks to $\hat{\gamma}$ obtained by NNRK (see 10).

Secondly, we compare the infered 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

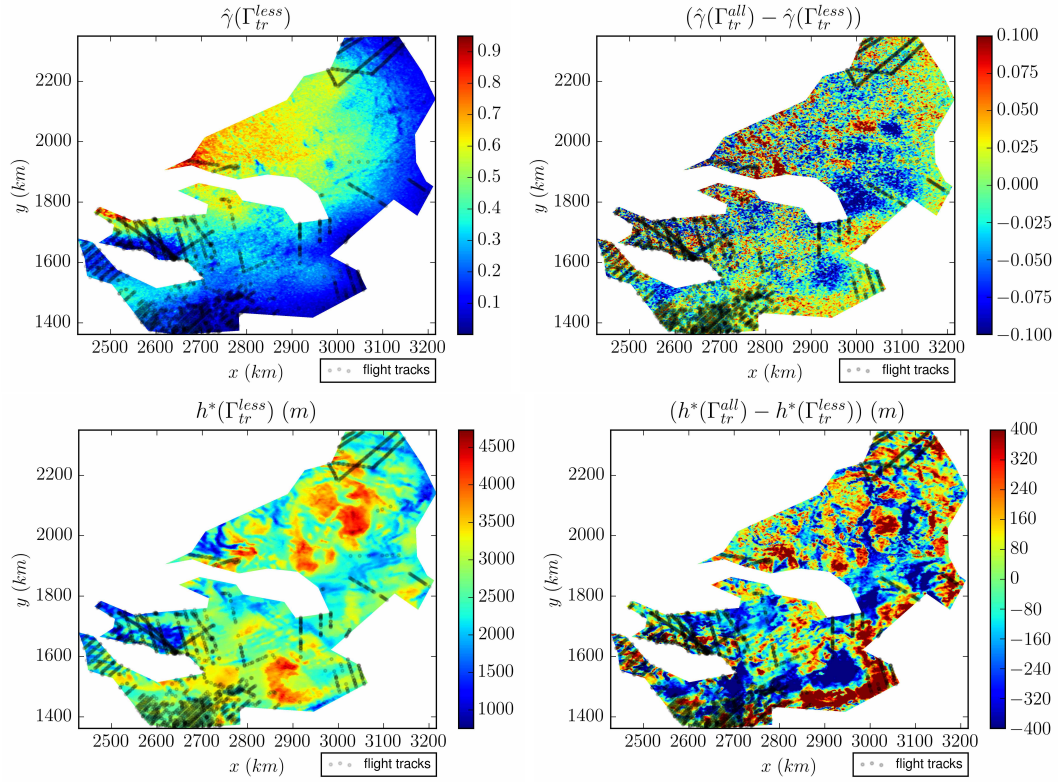

Figure 7. Domain Ω_p=Ant1: comparison if not considering the flights tracks indicated in Fig.6 (Up)(L). (Up)(L) Field $\hat{\gamma}(\Gamma_{tr}^{less})$ (i.e. without the flights tracks indicated in Fig.6 (Up)(L)). (R) Difference $(\hat{\gamma}(\Gamma_{tr}^{all}) - \hat{\gamma}(\Gamma_{tr}^{less}))$.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

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

2c), see paragraph 2.3.2) 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 from 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 obtained variations in h are roughly half than the ones obtained from h_b, see tables 2 and 4: difference of 7.2% in mean vs 13.3%, and 5.3% vs 10.4% for median values.

4.2. Results for Ant3 area

Similarly to Ant1, Ant3 presents large uncovered areas during the airborne campaigns, corresponding to huge uncertainty on h_b, Fig. 9.

4.2.1. Ant3: the ice thickness estimation h^*

Like in Ant1 case and as already noticed in [24] too, the only clearly observed correlation is: γ is small if $\mathbf{u}_H$ is small, Fig. 9 (Up). Again the difference between h^* and h_b is uncorrelated to the distance to the nearest flight track (on contrary to the empirically established uncertainty in [2]). Large corrections of h_b (up to 700 m) are found close to flight tracks see e.g. the area around coordinates (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 tracks, see e.g. the area around coordinates (5250,2800).
The few statistics presented in Tab. 5 show that again after the VDA processes, the RU-SIA equation fits extremely well with the surface elevation. The correction made on $\dot{a}$ is noticeably lower than the authorised maximal variation: 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% of volume change only. However in the most uncertain areas, the corrections made can be both low, see e.g. the areas around coordinates (5000,2650) (5200,3250), and important ($\pm \sim$ 700m), see e.g. the area around coordinates (5000,2650).

4.2.2. 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, $\hat{\gamma}$ next h are changed all over the domain and not particularly in the vicinity of the missing flight track. One of the largest changes are obtained in areas far form the missing tracks and close to considered ones, see e.g. the area around coordinates (5100,3200). The difference between the two estimations are

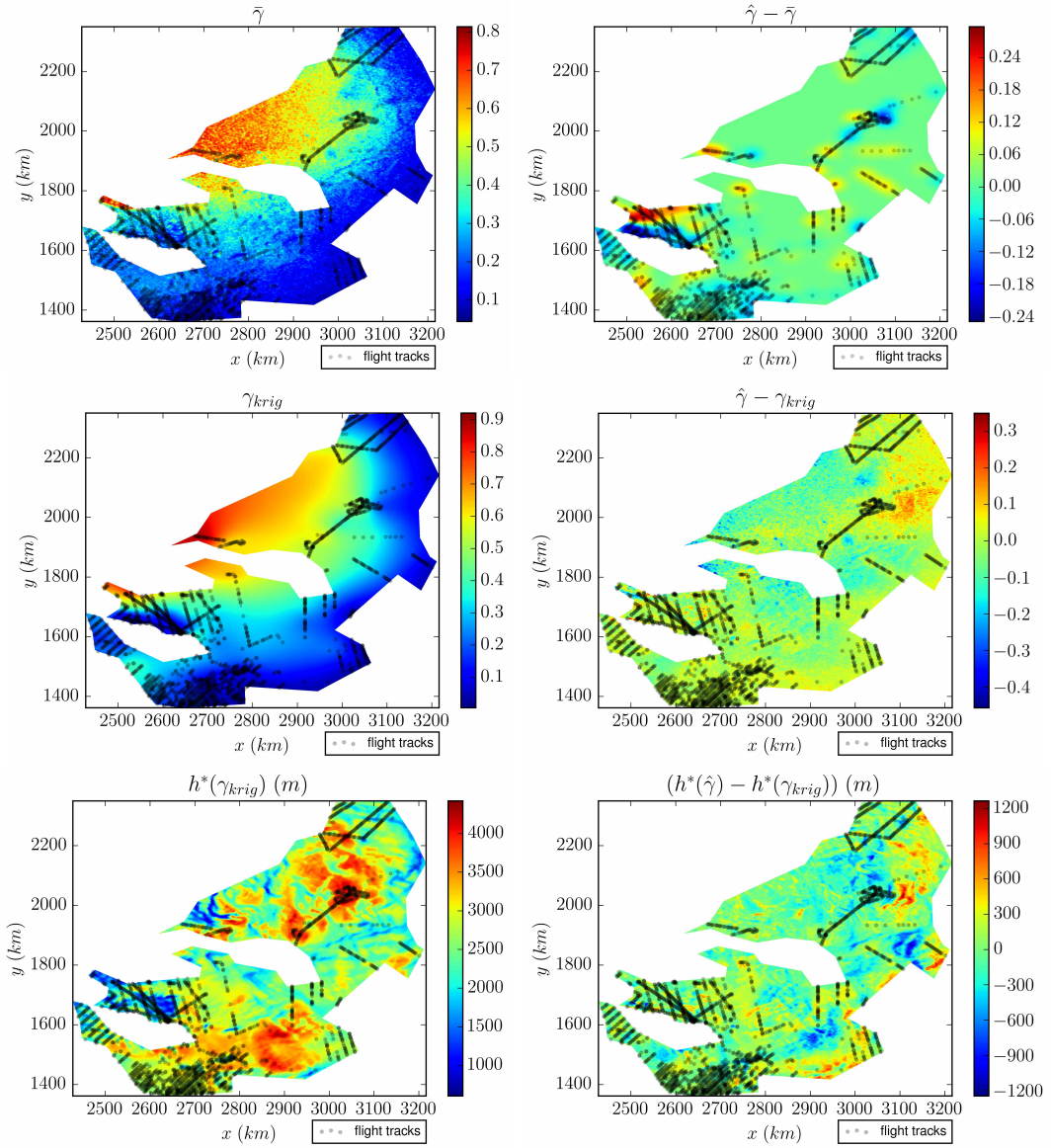

Figure 8. Domain Ω_p=Ant1: comparison between different statistical learning methods at Step 2). (Up)(L) $\bar{\gamma}$ computed by ANN only. (R) Difference $(\hat{\gamma} - \bar{\gamma})$. (Middle)(L) γ_{krig} computed by ordinary Kriging. (R) Difference $(\hat{\gamma} - \gamma_{krig})$. (Down)(L) Infered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

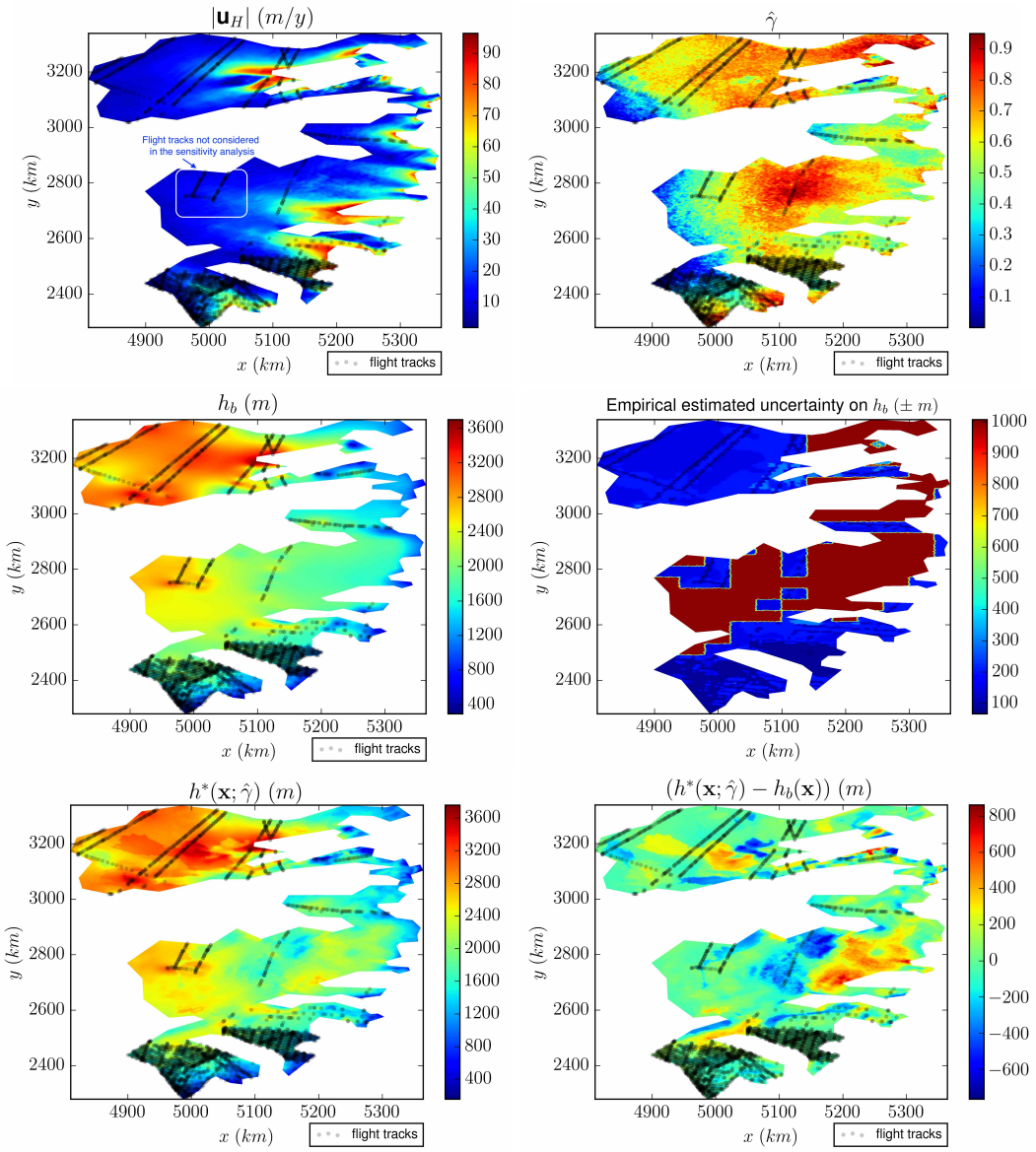

Figure 9. Domain Ω_p=Ant3: (Up)(L) Surface velocity module $|\mathbf{u}_H|$ and flight tracks (R) $\hat{\gamma}$ computed by NNRK, see (10). (Middle)(L) Thickness h_b from Bedmap2 [2] (R) Empirical uncertainty on h_b from [2]. (Down)(L) Infered thickness with $\hat{\gamma}$: $h^*(\hat{\gamma})$ (R) Difference $(h^*(\hat{\gamma}) - h_b)$.

Table 5. Domain Ω_p=Ant3, information and results.

Domain Ω_p & mesh information			
Surface $\lvert\Omega_p\rvert$		$250268\,km^2$	
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
$\lvert H(h_b) - H^{obs}\rvert$ (before h-inversion)	$7.8\,m$	$12.7\,m$	$274.0\,m$
$\lvert H(h^*) - H^{obs}\rvert$ (after h-inversion)	$2.8\,m$	$4.0\,m$	$110.6\,m$
Infered RHS $\dot{a}$			
$\lvert\dot{a}^* - \dot{a}_b\rvert$	$2.2\,cm/y$	$2.5\,cm/y$	$22.1\,cm/y$
$\lvert\dot{a}^* - \dot{a}_b\rvert/\lvert\dot{a}_b\rvert$	11.2 %	11.4 %	20 %
Infered thickness h			
$\lvert h^* - h_b\rvert$	$70.0\,m$	$124.5\,m$	$862.2\,m$
$\lvert h^* - h_b\rvert/\lvert h_b\rvert$	3.5%	6.6%	63.5%
Ice volume change in km^3 / in %		$3.0\ 10^3\,km^3$ / 0.5%	

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

Infered thickness difference	Median	Mean	Max
$\lvert h^*(\Gamma_{tr}^{all}) - h^*(\Gamma_{tr}^{less})\rvert$	$42.7\,m$	$56.3\,m$	$904.6\,m$
$\lvert h^*(\Gamma_{tr}^{all}) - h^*(\Gamma_{tr}^{less})\rvert/\lvert h^*(\Gamma_{tr}^{less})\rvert$	2.2%	2.8%	78.7%

plotted in Fig. 6 (Up)(R) and (Down)(R). (Again for sake of readability, the legend in Fig. 10(Down)(R) has been bounded at ±400m; very few values were greater than this bound). Basic statistics on the difference are presented in Tab. 6. (Note that e.g. a difference of 200m corresponds to $\sim 10\%$ of change). Again, the obtained variations on h are roughly half than the ones those 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.

4.2.3. Ant3: with different statistical learning methods at Step 2)

Similarly to Ant1 case, the difference between $\bar{\gamma}$ and $\hat{\gamma}$ (i.e. before and after the Kriging Step 2c)) are essentially localised in the vicinity of the flight tracks with amplitudes up to $\sim \pm 0.20$, Fig. 11 (Up)(R). Again too, the differences between $\hat{\gamma}$ and γ_{krig} are obviously not along the tracks but may be anywhere else; the observed difference may be up to $\sim \pm 0.45$, Fig. 11(Middle)(R). Statistics on the differences on the corresponding estimated thicknesses are presented in Tab. 7. The obtained variations in 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 conclusion, these comparisons based on one hand if removing some flight tracks, on the other hand if changing the statistical learning method, highlight the relative robustness and reliability of the proposed thickness estimations. If not considering a localised group of flight tracks, see figures 6 and 9 (Up)(L), the obtained variations on h are roughly half than the ones obtained from h_b. If considering a simple Kriging

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

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

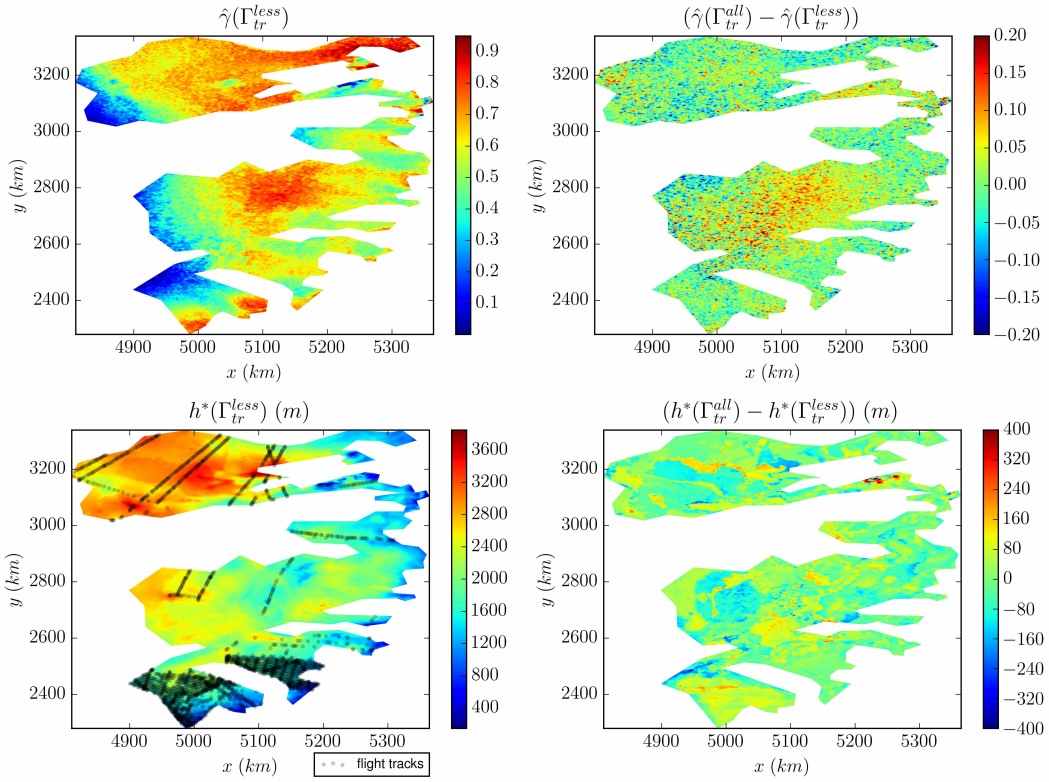

Figure 10. Domain Ω_p=Ant3: comparison if not considering the flights tracks indicated in Fig.9 (Up)(L). (Up)(L) Field $\hat{\gamma}(\Gamma_{tr}^{less})$ (i.e. without the flights tracks indicated in Fig.9 (Up)(L)). (R) Difference $(\hat{\gamma}(\Gamma_{tr}^{all}) - \hat{\gamma}(\Gamma_{tr}^{less}))$.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

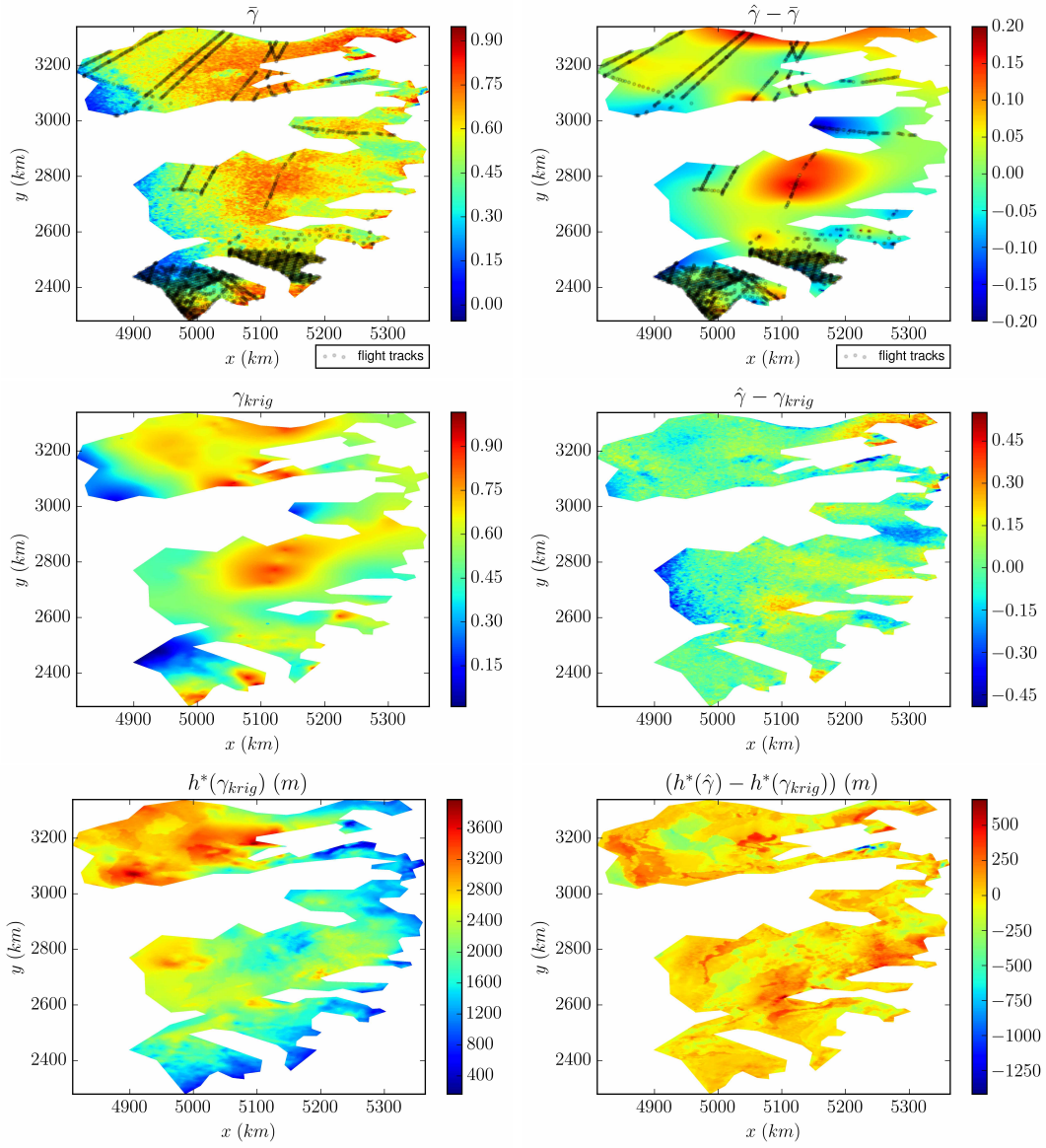

Figure 11. Domain Ω_p=Ant3: comparison between different statistical learning methods at Step 2). (Up)(L) $\bar{\gamma}$ computed by ANN only. (R) Difference $(\hat{\gamma} - \bar{\gamma})$. (Middle)(L) γ_{krig} computed by ordinary Kriging. (R) Difference $(\hat{\gamma} - \gamma_{krig})$. (Down)(L)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

method to estimate γ vs the NNRK algorithm, the obtained variations on h are roughly the half (lower than the ones obtained from h_b) in Ant1 case and two thirds in Ant3 case. Therefore in all cases the obtained variations are lower than the ones obtained from h_b. Moreover it may be guessed that the statistical learning method based on the measured data (the NNRK algorithm) should be more reliable than the "blind" Kriging extrapolation (which is not related to the datasets).

5. Conclusion

In this study, a method to infer the bedrock topography beneath glaciers at regional scale, at wave length $\sim 10\bar{h}$ (with a resolution at $\sim \bar{h}$, $\bar{h}$ a characteristic thickness value) is developed. The key ingredients of this inversion method are the following: a Reduced-Uncertainty flow model (RU-SIA) taking into account a complete physics (including depth-varying internal deformation) while presenting an unique dimensionless parameter γ; two advanced VDA processes; an Artificial Neural Network (ANN) estimating γ from the surface data and in-situ measurements (acquired during airborne campaigns). A guiding idea in the method development was to consider well-posed inverse problems (or at least not trivially ill-posed ones). All steps of the algorithm have been thoroughly assessed: the two VDA processes which have been proved to be robust, accurate and sufficiently insensitive to prior information (the first guess chosen here is BedMap2); the ANN prediction capabilities which turned out to be excellent on the considered datasets. It would be straightforward to apply the same ANN to directly estimate the ice thickness. However it is definitively more consistent to estimate an unique dimensionless parameter of the flow model from the surface features correctly represented by this same model, instead of a partially dependent parameter.
Numerous sensitivity analyses have been investigated, first in [24] and here in two large EAIS areas. In particular the estimation sensitivity with respect to: the presence or not of local in-situ measurements (flight tracks), different statistical learning methods (Kriging vs ANN), different grid resolutions and different scales of surface data smoothing.
The method enables to estimate the bed elevation between in-situ measurements, in poorly covered areas and even in large uncovered areas e.g. in EAIS where the current estimations are gravimetry-based (hence highly uncertain). The estimations in such uncovered areas may be as accurate as elsewhere since they are based on the surface signature (measured by altimetry and InSAR); it is shown that the inversion method is relatively global with local constraints (the flight tracks) and not purely local. In the considered six large EAIS areas, the corrections made to BedMap2 may be large (up to 1000 m) even close to flight tracks (e.g. at 2 wave lengths that is $\sim$ 50km). Away from in-situ measurements (flight tracks), the correction of BedMap2 may be significant or not, depending on the surface signature. These corrections led to a total ice volume change by $0.5 - 5.6\%$, depending on the area.
This bed estimation method may be applied to many other ice-sheets areas or ice-caps, as soon as respecting the method domain of validity, that is highly to moderately sheared flows (i.e. not slipping flows). The method may be extended to unsteady flows if the surface observations (elevation and velocity) are given in time and assuming that the initial condition is either not important (at the considered time scale) or approximatively known.

Acknowledgements

The first author acknowledges M. Morlighem and H. Seroussi from Univ. California of Irvine and JPL/Caltech-NASA for numerous and crucial discussions he had with them on real-world ice flows and datasets. Dr J. Zhu has been funded by a CNES TOSCA grant (oct. 2017-sept. 2018) during her post-doctoral stay at IMT - INSA Toulouse.

References

[1] Bamber J, Gomez-Dans J, Griggs J. A new 1 km digital elevation model of the antarctic derived from combined satellite radar and laser data–part 1: Data and methods. The Cryosphere. 2009;3(1):101–111.

[2] Fretwell P, Pritchard H, Vaughan D, et al. Bedmap2: improved ice bed, surface and thickness datasets for antarctica. The Cryosphere. 2013;7(1).

[3] Helm V, Humbert A, Miller H. Elevation and elevation change of greenland and antarctica derived from cryosat-2. The Cryosphere. 2014;8(4):1539–1559.

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

[5] Jean-Paul C, Pierre D. Geostatistics: modeling spatial uncertainty. John Wiley & Sons Inc, New York. 1999;:695.

[6] Bamber J, Griggs J, Hurkmans R, et al. A new bed elevation dataset for greenland. The Cryosphere. 2013;7(2):499–510.

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

[8] Wahba G. Spline models for observational data. Vol. 59. Siam; 1990.

[9] Bennett AF. Inverse modeling of the ocean and atmosphere. Cambridge University Press; 2005.

[10] Farinotti D, Brinkerhoff DJ, Clarke GK, et al. How accurate are estimates of glacier ice thickness? results from itmix, the ice thickness models intercomparison experiment. Cryosphere. 2017;11(2):949–970.

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

[12] Morlighem M, Rignot E, Mouginot J, et al. High-resolution ice-thickness mapping in south greenland. Annals of Glaciology. 2014;55(67):64–70.

[13] Michel L, Picasso M, Farinotti D, et al. Estimating the ice thickness of mountain glaciers with an inverse approach using surface topography and mass-balance. Inverse Problems. 2013;29(3):035002.

[14] Budd W, Allison I. An empirical scheme for estimating the dynamics of unmeasured glaciers. In: Proceedings of the Moscow Symposium ?Snow and Ice; 1975. p. 246–256.

[15] Morlighem M, Rignot E, Seroussi H, et al. A mass conservation approach for mapping glacier ice thickness. Geophysical Research Letters. 2011;38(19).

[16] Morlighem M, Williams CN, Rignot E, et al. Bedmachine v3: Complete bed topography and ocean bathymetry mapping of greenland from multibeam echo sounding combined with mass conservation. Geophysical research letters. 2017;44(21).

[17] Brinkerhoff D, Aschwanden A, Truffer M. Bayesian inference of subglacial topography using mass conservation. Frontiers in Earth Science. 2016;4:8.

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

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

[20] Huss M, Farinotti D. A high-resolution bedrock map for the antarctic peninsula. The Cryosphere. 2014;8(4):1261–1273.
[21] Winkelmann R, Martin M, Haseloff M, et al. The potsdam parallel ice sheet model (pismpik) part 1: Model description. The Cryosphere. 2011;5(3):715–726.
[22] Van Pelt W, Oerlemans J, Reijmer C, et al. An iterative inverse method to estimate basal topography and initialize ice flow models. The Cryosphere. 2013;7:987–1006.
[23] Monnier J, des Boscs P. Inference of the bottom properties in shallow ice approximation models. Inverse Problems. 2017;33(11):115001.
[24] Monnier J, J Z. Inference of the bed topography for non-isothermal moderately sheared ice flows. 2018;:In review.
[25] Bahr DB, Pfeffer WT, Kaser G. Glacier volume estimation as an ill-posed inversion. Journal of Glaciology. 2014;60(223):922–934.
[26] Sellier M. Inverse problems in free surface flows: a review. Acta Mechanica. 2016; 227(3):913–935.
[27] Gudmundsson G. Transmission of basal variability to a glacier surface. Journal of Geophysical Research: Solid Earth. 2003;108(B5).
[28] Gudmundsson G. Analytical solutions for the surface response to small amplitude perturbations in boundary data in the shallow-ice-stream approximation. The Cryosphere. 2008;2(2):77–93.
[29] Martin N, Monnier J. Adjoint accuracy for the full-stokes ice flow model: limits to the transmission of basal friction variability to the surface. The Cryosphere. 2014;8:721–741. Available from: http://www.the-cryosphere.net/8/721/2014.
[30] Martin N, Monnier J. Inverse rheometry and basal properties inference for pseudoplastic geophysical flows. European Journal of Mechanics-B/Fluids. 2015;50:110–126.
[31] Demyanov V, Kanevsky M, Chernov S, et al. Neural network residual kriging application for climatic data. Journal of Geographic Information and Decision Analysis. 1998; 2(2):215–232.
[32] Kanevski M, Timonin V, Pozdnukhov A. Machine learning for spatial environmental data: theory, applications, and software. EPFL press; 2009.
[33] Fogwill CJ, Turney CS, Meissner KJ, et al. Testing the sensitivity of the east antarctic ice sheet to southern ocean dynamics: past changes and future implications. Journal of Quaternary Science. 2014;29(1):91–98.
[34] van Wessem JM, Jan Van De Berg W, Noël BP, et al. Modelling the climate and surface mass balance of polar ice sheets using racmo2: Part 2: Antarctica (1979-2016). Cryosphere. 2018;12(4):1479–1498.
[35] Greve R, Blater H. Dynamics of ice sheets and glaciers. Springer-Verlag; 2009. Advances in Geophysical and Environmental Mechanics and Mathematics.
[36] Kamb B, Echelmeyer KA. Stress-gradient coupling in glacier flow: I. longitudinal averaging of the influence of ice thickness and surface slope. Journal of Glaciology. 1986; 32(111):267–284.
[37] Hindmarsh R. A numerical comparison of approximations to the stokes equations in ice sheet and glacier modeling. J geophysical research. 2004;109.
[38] Schoof C, Hindmarsh C. Thin-film flows with wall slip: an asymptotic analysis of higher order glacier flow models. Q Jl Mech Appl Math. 2010;63(1):73–114.
[39] Boutounet M, Monnier J, Vila JP. Multi-regime shallow free surface laminar flow models for quasi-newtonian fluids. European Journal of Mechanics-B/Fluids. 2016;55:182–206.
[40] Williams CR, Hindmarsh RC, Arthern RJ. Calculating balance velocities with a membrane stress correction. Journal of Glaciology. 2014;60(220):294–304.
[41] Hindmarsh RC. Stress gradient damping of thermoviscous ice flow instabilities. Journal of Geophysical Research: Solid Earth. 2006;111(B12).
[42] Price PB, Nagornov OV, Bay R, et al. Temperature profile for glacial ice at the south pole: Implications for life in a nearby subglacial lake. Proceedings of the National Academy of Sciences. 2002;99(12):7844–7847.
[43] Radok U, Jenssen D, Budd W. Steady-state temperature profiles in ice sheets. Bull Int

Assoc Scient Hydrol. 1970;8(1):36.
[44] Jansen D, Sandhäger H, Rack W. Model experiments on large tabular iceberg evolution: ablation and strain thinning. Journal of glaciology. 2005;51(174):363–372.
[45] Seroussi H, Morlighem M, Rignot E, et al. Dependence of century-scale projections of the greenland ice sheet on its thermal regime. Int Glacio Soc. 2013;59:1024–1034.
[46] Kaltenbacher B, Neubauer A, Scherzer O. Iterative regularization methods for nonlinear ill-posed problems. Vol. 6. Walter de Gruyter; 2008.
[47] Haben SA, Lawless AS, Nichols NK. Conditioning of incremental variational data assimilation, with application to the met office system. Tellus A: Dynamic Meteorology and Oceanography. 2011;63(4):782–792.
[48] Glorot X, Bordes A, Bengio Y. Deep sparse rectifier neural networks. In: Proceedings of the fourteenth international conference on artificial intelligence and statistics; 2011. p. 315–323.
[49] LeCun Y, Bengio Y, Hinton G. Deep learning. nature. 2015;521(7553):436.
[50] Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:14126980. 2014;.
[51] Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: a simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research. 2014;15(1):1929–1958.
[52] Arlot S, Celisse A, et al. A survey of cross-validation procedures for model selection. Statistics surveys. 2010;4:40–79.
[53] Dalcín L, Paz R, Storti M. Mpi for python. Journal of Parallel and Distributed Computing. 2005;65(9):1108–1115.
[54] Clarke GK, Berthier E, Schoof CG, et al. Neural networks applied to estimating subglacial topography and glacier volume. Journal of Climate. 2009;22(8):2146–2160.
[55] Geuzaine C, Remacle JF. Gmsh: A 3-d finite element mesh generator with built-in pre- and post-processing facilities. International journal for numerical methods in engineering. 2009;79(11):1309–1331.

Appendix A. Results for the four other EAIS areas

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 upstream of Fisher and Merllor ice-streams (upstream Amery ice shelf); Ant4 is located upstream Totten ice-streams (Wilkes land and Terre Adlie); Ant5 is located upstream Ninnis and Mertz ice-streams in Terre Adlie and George V land; Ant6 is located upstream Byrd ice-streams (east of Ross Ice shelf), see Fig. 4. These areas have been relatively well covered during the airborne campaigns excepted the north-east part of Ant2, see Fig. A1. Distances between flight tracks are relatively low, therefore the empirical uncertainty assigned to Bedmap2 estimations h_b is low too: $\approx \pm[100 - 250]$m, see figures A1, A2, A3, A4 (Middle).
For each area, domain information and statistics on the numerical results are presented, see tables A1, A2, A3, A4.
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 parameters in (1): see lines "($|H(h_b) - H_{obs}|$ (before h-inversion)" in the tables. Next RU-SIA model fits accurately with the surface topography after the 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.
The corrections made on $\dot{a}$ are $\sim [10 - 17]\%$ in mean, that is within the a-priori

Table A1. Domain Ω_p=Ant2, information and results.

Domain Ω_p & mesh information			
Surface $\lvert\Omega_p\rvert$		$431860\,km^2$	
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
$\lvert H(h_b) - H^{obs}\rvert$ (before h-inversion)	$8.7\,m$	$17.3\,m$	$149.9\,m$
$\lvert H(h^*) - H^{obs}\rvert$ (after h-inversion)	$3.9\,m$	$5.1\,m$	$49.2\,m$
Infered RHS $\dot{a}$			
$\lvert\dot{a}^* - \dot{a}_b\rvert$	$0.6\,cm/y$	$0.8\,cm/y$	$4.0\,cm/y$
$\lvert\dot{a}^* - \dot{a}_b\rvert/\lvert\dot{a}_b\rvert$	14.8 %	13.2 %	20 %
Infered thickness h			
$\lvert h^* - h_b\rvert$	$171.1\,m$	$302.4\,m$	$2025.6\,m$
$\lvert h^* - h_b\rvert/\lvert h_b\rvert$	8.3%	14.2%	80.0%
Ice volume change in km^3 / in %	$5.2\ 10^4\,km^3$ / 5.6%		

Table A2. Domain Ω_p=Ant4, information and results.

Domain Ω_p & mesh information			
Surface $\lvert\Omega_p\rvert$		$439045\,km^2$	
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
$\lvert H(h_b) - H^{obs}\rvert$ (before h-inversion)	$6.3\,m$	$8.4\,m$	$66.9\,m$
$\lvert H(h^*) - H^{obs}\rvert$ (after h-inversion)	$3.0\,m$	$4.1\,m$	$45.3\,m$
Infered RHS $\dot{a}$			
$\lvert\dot{a}^* - \dot{a}_b\rvert$	$2.0\,cm/y$	$2.7\,cm/y$	$15.0\,cm/y$
$\lvert\dot{a}^* - \dot{a}_b\rvert/\lvert\dot{a}_b\rvert$	8.7 %	10.1 %	20 %
Infered thickness h			
$\lvert h^* - h_b\rvert$	$147.6\,m$	$185.0\,m$	$1241.6\,m$
$\lvert h^* - h_b\rvert/\lvert h_b\rvert$	5.7%	7.0%	50.4%
Ice volume change in km^3 / in %	$1.5\ 10^4\,km^3$ / 1.5%		

uncertainty range indicated in [34].
The estimated thickness h^* is plotted in each case, see figures A1, A2, A3, A4 (Down)(L); its difference with h_b is plotted (Down)(R). The corrections made 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 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 track. Maximum values of correction to h_b can be locally high. In the uncovered north-east area in Ant2, these maximum values corrections reach $\approx$ 2000 m. At $\approx$ 50km from the nearest flight track (that is $\sim$ 2 minimal wave lengths of the model) correction may reach 1000m (even in area surrounded by well covered areas), see e.g. figures A3 and A4 (Down)(R). Beyond $\approx$ 100km from the nearest data, the correction of h_b may be significant or not, depending on the surface signature, see e.g. figures A3 and A4 (Down)(R).

Table A3. Ant 5, information and results.

Domain Ω_p & mesh information			
Surface $\lvert\Omega_p\rvert$		$362019\,km^2$	
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
$\lvert H(h_b) - H^{obs}\rvert$ (before h-inversion)	$18.3\,m$	$23.9\,m$	$173.3\,m$
$\lvert H(h^*) - H^{obs}\rvert$ (after h-inversion)	$3.9\,m$	$5.1\,m$	$43.1\,m$
Infered RHS $\dot{a}$			
$\lvert\dot{a}^* - \dot{a}_b\rvert$	$3.9\,cm/y$	$4.4\,cm/y$	$14.5\,cm/y$
$\lvert\dot{a}^* - \dot{a}_b\rvert/\lvert\dot{a}_b\rvert$	18.8 %	15.4 %	20 %
Infered thickness h			
$\lvert h^* - h_b\rvert$	$275.1\,m$	$373.1\,m$	$1989.2\,m$
$\lvert h^* - h_b\rvert/\lvert h_b\rvert$	12.1%	15.9%	68.8%
Ice volume change in km^3 / in %		$3.8\ 10^4\,km^3$ / 5.6%	

Table A4. Domain Ω_p=Ant6, information and results.

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

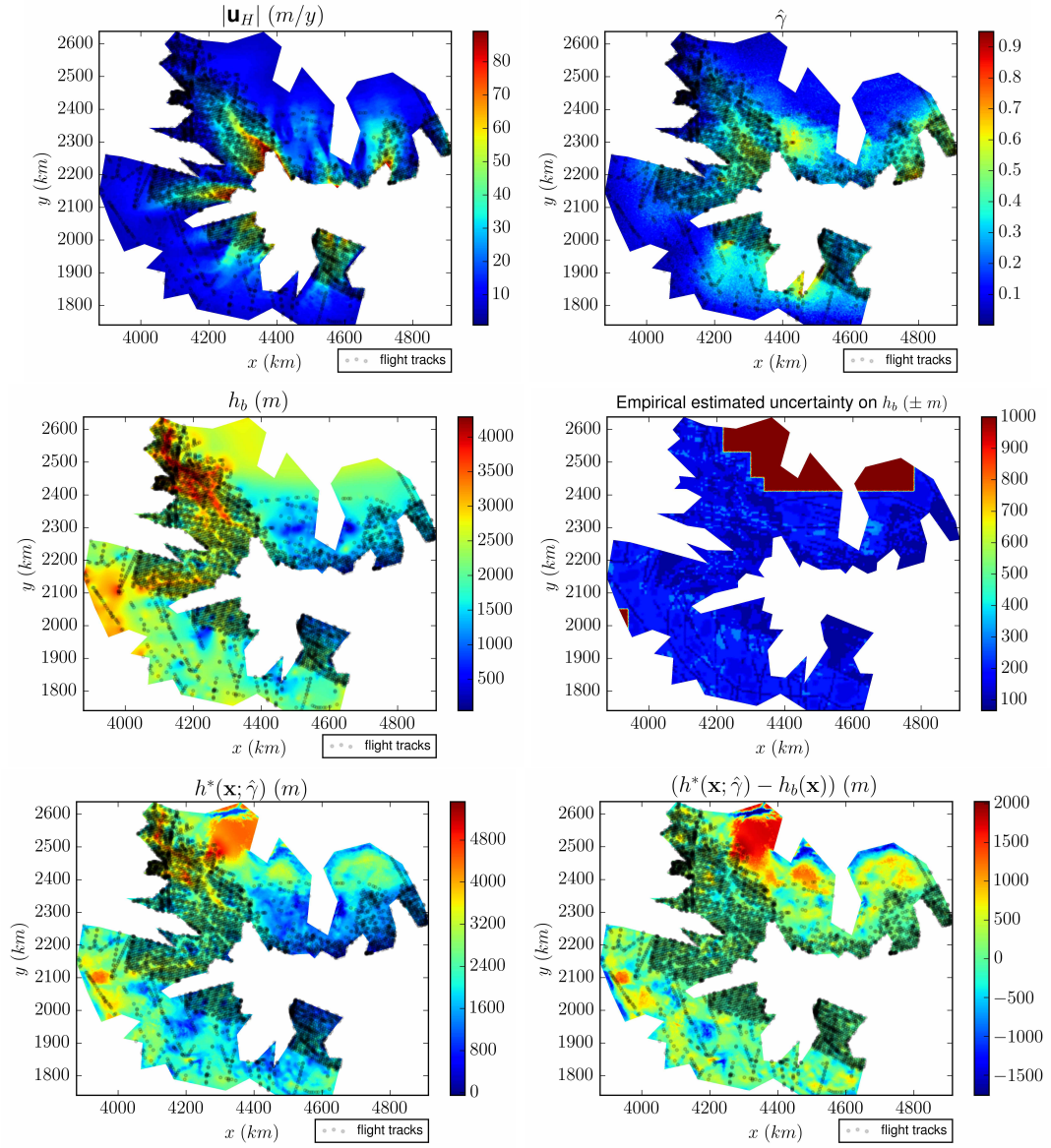

Figure A1. Domain Ω_p=Ant2: (Up)(L) Surface velocity module $|\mathbf{u}_H|$ and flight tracks. (R) $\hat{\gamma}$ computed by NNRK, see (10). (Middle)(L) Thickness h_b from Bedmap2 [2]. (R) Empirical uncertainty on h_b from [2]. (Down)(L) Infered thickness with $\hat{\gamma}$: $h^*(\hat{\gamma})$. (R) Difference $(h^*(\hat{\gamma}) - h_b)$.

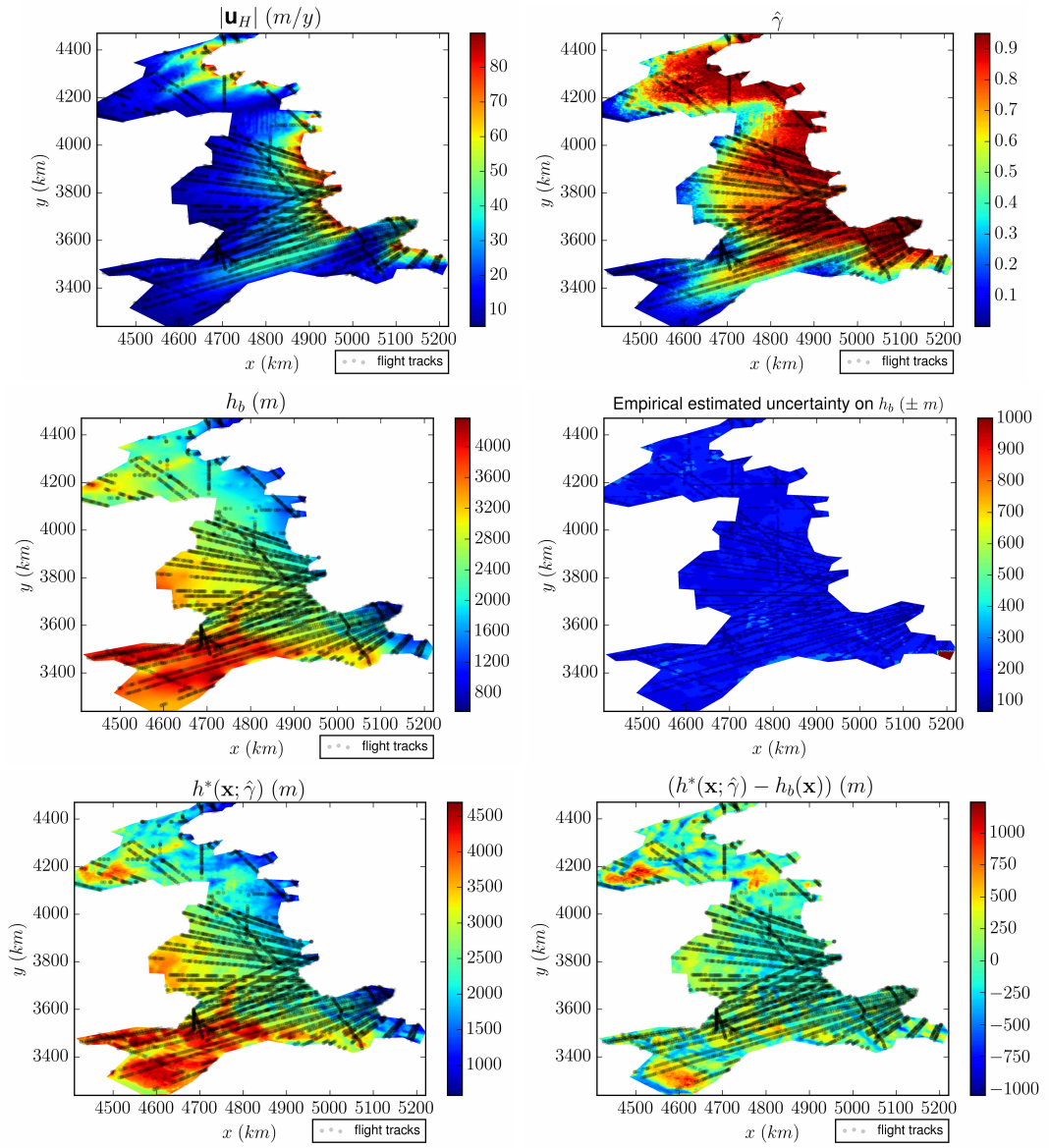

Figure A2. Domain Ω_p=Ant4: (Up)(L) Surface velocity module $|\mathbf{u}_H|$ and flight tracks. (R) $\hat{\gamma}$ computed by NNRK, see (10). (Middle)(L) Thickness h_b from Bedmap2 [2]. (R) Empirical uncertainty on h_b from [2]. (Down)(L) Infered thickness with $\hat{\gamma}$: $h^*(\hat{\gamma})$. (R) Difference $(h^*(\hat{\gamma}) - h_b)$.

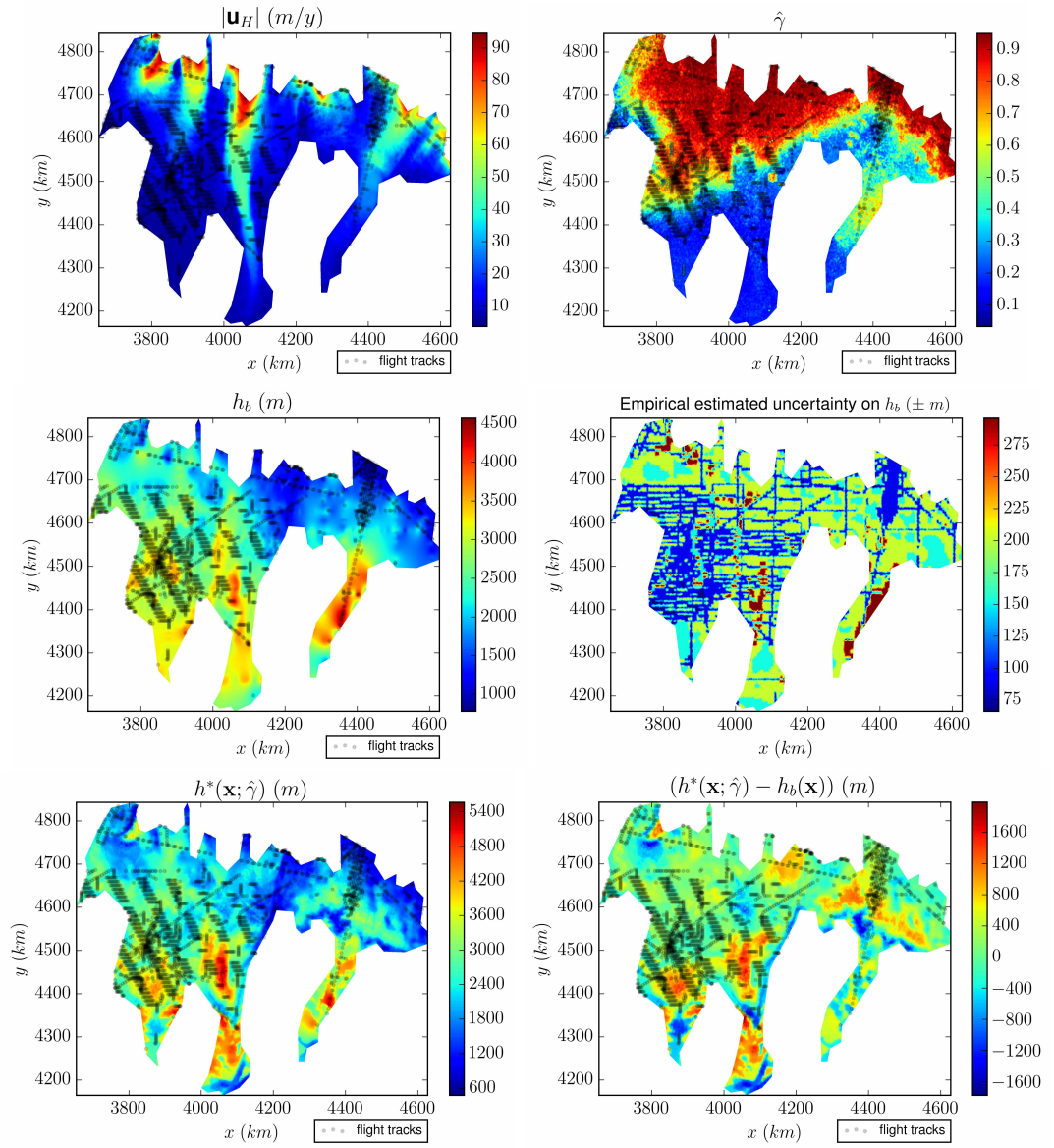

Figure A3. Ant 5: (Up)(L) Surface velocity module $|\mathbf{u}_H|$ and flight tracks. (R) $\hat{\gamma}$ computed by NNRK, see (10). (Middle)(L) Thickness h_b from Bedmap2 [2]. (R) Empirical uncertainty on h_b from [2]. (Down)(L) Infered thickness with $\hat{\gamma}$: $h^*(\hat{\gamma})$. (R) Difference $(h^*(\hat{\gamma}) - h_b)$.

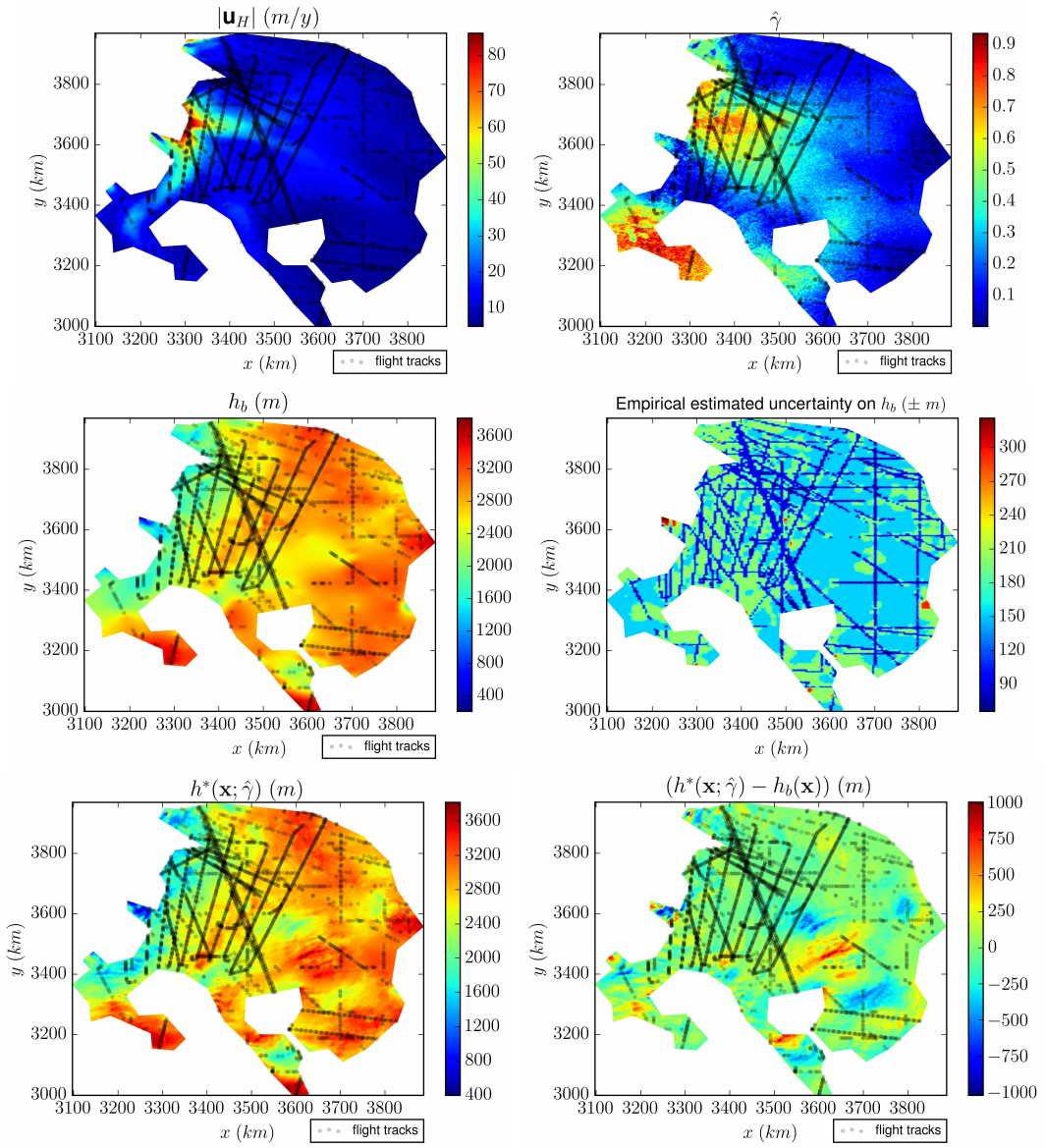

Figure A4. Domain Ω_p=Ant6: (Up)(L) Surface velocity module $|\mathbf{u}_H|$ and flight tracks. (R) $\hat{\gamma}$ computed by NNRK, see (10). (Middle)(L) Thickness h_b from Bedmap2 [2]. (R) Empirical uncertainty on h_b from [2]. (Down)(L) Infered thickness with $\hat{\gamma}$: $h^*(\hat{\gamma})$. (R) Difference $(h^*(\hat{\gamma}) - h_b)$.