Comput Geosci (2014) 18:969–988
DOI 10.1007/s10596-014-9439-6

ORIGINAL PAPER

Bedrock topography reconstruction of glaciers from surface topography and mass–balance data

Laurent Michel-Griesser · Marco Picasso · Daniel Farinotti · Martin Funk · Heinz Blatter

Received: 13 February 2014 / Accepted: 23 July 2014 / Published online: 23 August 2014

Abstract Three methods based on the three-dimensional shallow ice approximation of glacier flow are devised that infer a glacier's subglacial topography from the observation of its time-evolving surface and mass balance. The quasi-stationary inverse method relying on the apparent surface mass-balance description of the glacier's evolution is first exposed. Second, the transient inverse method that iteratively updates the bedrock topography with the surface topography discrepancy is formulated. Third, a shape optimization algorithm is presented. The aim of the paper is to collect these methods, analyze their differences, and identify what brings the sophistication of shape optimization for reconstructing subglacial topographies. The three methods are compared to the ice thickness estimation method (ITEM) on direct measurements on Gries glacier, Swiss Alps. The paper concludes with a detailed discussion on the sensitivity of the shape optimization method to the model parameters.

L. Michel-Griesser (✉) · M. Picasso
EPFL SB MATHICSE, Station 8, 1015 Lausanne, Switzerland
e-mail: laurent.griesser@gmail.com

D. Farinotti
German Research Centre for Geosciences (GFZ), Telegrafenberg, 14473 Potsdam, Germany

M. Funk
ETH Zurich, Laboratory of Hydraulics, Hydrology and Glaciology, Wolfgang-Pauli-Str. 27, 8093 Zurich, Switzerland

H. Blatter
ETH Zurich, Institute for Atmospheric and Climate Science, Universitätstrasse 16, 8092 Zurich, Switzerland

Keywords Glacier · Bedrock topography · Inverse problem · Tikhonov regularization · Transient inverse method · Quasi-stationary inverse method · Shape optimization algorithm

1 Introduction

Ice flow modelling requires accurate boundary conditions and model parameters. For instance, initializing a glacier model with an unsuitable approximation of subglacial topography may result in possibly significant errors in the predicted glacier's surface topography. Moreover, the glacier model's building blocks, namely ice rheology, surface mass-balance and sliding law, describe physics approximately with parameters that can hardly be determined accurately and are thus still the subject of current research [4, 5, 8, 24, 27, 35, 37, 42, 66].

Once the boundary conditions and the model parameters are known, the numerical simulation of glaciers can provide the desired forecasts. The most complete description of ice dynamics is provided by a non-linear Stokes approximation of ice flow coupled with an evolution equation for the glacier's surface that transports ice following the Stokes velocity [24, 40, 80]. Past reconstruction and future predictions of some Swiss glaciers was successfully performed with a three-dimensional model [41, 42] that usually relies on large computer resources, therefore making model inversions difficult. Hence, many lower-order models of ice flow have been developed in the past decades, among which the widespread shallow ice approximation (SIA) [23, 39, 58].

This paper focuses on the computation of subglacial topography from knowledge of surface topography evolution. Although this topography can actually be assessed

by geophysical means, for practical reasons, its measurement can only be performed at selected locations and for a limited sample of glaciers [12]. Therefore, several theoretical approaches to determine subglacial topography have been proposed. Assumptions of a parabolic cross section and perfectly plastic behaviour of glacier ice supply a rough estimation, where the shear stress at the glacier's base is supposed to be constant and close to the plastic yield stress of ice [63]. The latter estimation can be extended to account for the effect of side drag on the stress balance [48]. A similar approach, based on the parallel-sided slab approximation [23], empirically infers the glacier-specific plastic yield stress as a function of the glacier's elevation range and then relates it to the glacier's thickness and slope [28, 64]. An alternative method exists that relies on mass turnover and parallel-sided slabs [17]. Subglacial topography of a very large sample of glaciers is reconstructed with an even more simplified approach [36]. It is also estimated by means of inverse approaches relying on surface velocity measurements [15, 25, 52, 60, 68, 69, 73]. Most of these procedures are direct algorithms in the sense that they use simplified model equations from which subglacial topography can be deduced analytically. The transient and quasi-stationary inverse methods formulated in what follows are simple iterative methods that can be applied to any flow model.

However, the latter two inverse methods can be improved with shape optimization. On subglacial topography estimation by means of optimal control, the oldest contribution is an interpolation method that seeks the best solution fitting measured data, under the constraint that it also minimizes topography potential and curvature [38]. Subsequent work on optimal control algorithms that infer the ice thickness distribution of a glacier is rather sparse but literature has lately become more abundant on those that minimize the surface velocity misfit on steady geometries [52, 60]. Without direct consideration of surface velocity data, Clarke et al. [13] minimize the discrepancy between observed ice thickness and quasi-stationary shallow ice value. Optimal control is more commonly used to determine basal sliding law and rheology coefficients [4, 5, 21, 27, 44, 45, 47, 50, 51, 66, 75], where the misfit between computed and measured surface velocities is minimized with two and three-dimensional flow models of arbitrary order.

Many of the existing reconstruction methods rely on surface velocity measurements, which makes the problem of computing both basal topography and model parameters like rheology, surface mass-balance and sliding coefficients underdetermined. Because very little velocity information is available on Swiss Alps glaciers, the algorithms below compute the ice thickness distribution from observed surface topographies.

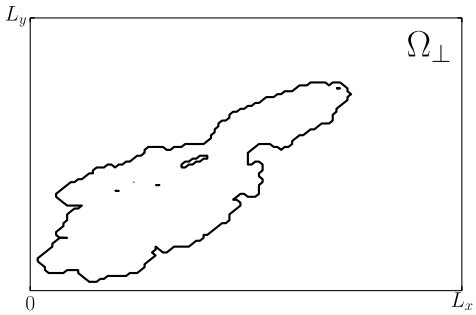
The three conceptually different algorithms presented in this paper are extensions to either three space dimensions or the optimal control framework of the transient and quasi-stationary inverse methods developed in [55, 56, 65]. Our previous contribution on the latter two methods introduced their two-dimensional formulation without consideration of sliding and did not compare them to the shape optimization algorithm. The main purpose of this paper is to compare the three methods theoretically and computationally. On synthetic data, the superiority of the shape optimization over the other inversion methods is demonstrated. The algorithms are also compared to the ice thickness estimation method (ITEM, [17]) for reconstructing the bedrock topography of Gries glacier, Swiss Alps, for which ice thickness measurements exist, thus allowing a comparison with real-world data. Despite its simple underlying ice flow model, the ITEM showed its efficiency and relevance in estimating basal topography of several glaciers [16, 17]. The advantages of the three methods over currently existing procedures are that they all incorporate sliding, time-dependence, and three-dimensional space in a physically consistent way with respect to the shallow ice model. In particular, they use neither interpolation nor any kind of surface topography filtering where the surface slope is too small.

The paper's outline is the following: first, the shallow ice equation is recalled. Next, the quasi-stationary inverse, transient inverse and shape optimization methods are formulated. Then, the three algorithms are compared in performance on synthetic data and the real-world geometry of Gries glacier, Swiss Alps. In the real-world case, the three methods are compared to the ITEM. Finally, the sensitivity of the shape optimization procedure to the model parameters is discussed and conclusions are drawn.

2 Forward problem

A three-dimensional glacier ice volume is considered whose time-dependent outline is contained in $\Omega_\perp = [0, L_x] \times [0, L_y]$ (see Fig. 1), from initial time t_i to final time t_f. Its bed and surface topographies are denoted by b and s, respectively, and its ice thickness by $\mathcal{H} = s - b$. Glacial isostasy is neglected, hence, the bedrock topography is time independent. In this contribution, the three-dimensional SIA of flow dynamics is considered, where the transient forward problem can be cast into: *given a bedrock b, an initial surface s_i and a surface mass-balance time series $\mathcal{B}$, find the ice thickness $\mathcal{H} : \Omega_\perp \times [t_i, t_f] \longrightarrow \mathbb{R}$ such that*

$$\begin{cases} \dfrac{\partial \mathcal{H}}{\partial t} = \nabla \cdot \Big(\mathcal{D} \nabla (b + \mathcal{H}) \Big) + \mathcal{B}, & \text{in } \Omega_\perp \times [t_i, t_f], \\ \mathcal{H} \geq 0, & \text{in } \Omega_\perp \times [t_i, t_f], \\ \mathcal{H} = 0, & \text{on } \partial\Omega_\perp \times [t_i, t_f], \\ \mathcal{H} = s_i - b, & \text{in } \Omega_\perp \times \{t_i\}, \end{cases} \tag{1}$$



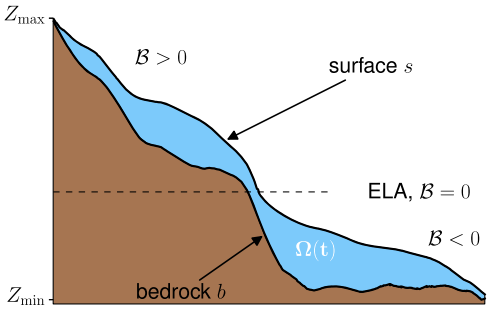


Fig. 1 *Left* Example ice extent in the (x, y)-plane of Gries glacier, Swiss Alps, that illustrates the notations introduced in the text. The map domain $\Omega_\perp$ consists of the whole rectangle $[0, L_x] \times [0, L_y]$. *Right* Profile of Gries glacier along a flow line. The bedrock topography b is the ice-lithosphere interface, while the surface topography is the ice—air interface. The ice domain is depicted in *blue*

where the total diffusivity $\mathcal{D}$ consists in two contributions, $\mathcal{D}_f$ and $\mathcal{D}_s$, due to flow dynamics and Weertman-type [76] sliding respectively, namely

$$\mathcal{D} = \mathcal{D}_f + \mathcal{D}_s, \tag{2}$$

where

$$\mathcal{D}_f = \Gamma \mathcal{H}^{n+2} \left\| \nabla(b+\mathcal{H}) \right\|^{n-1}, \qquad \mathcal{D}_s = \Gamma_s \mathcal{H}^{n+1} \left\| \nabla(b+\mathcal{H}) \right\|^{n-1}, \tag{3}$$

and the diffusion coefficients are defined by

$$\Gamma = 2\frac{A(\rho g)^n}{n+2}, \qquad \Gamma_s = \left(\frac{\rho g}{C}\right)^n, \tag{4}$$

where A is the rate factor, ρ is the ice density, g is the acceleration due to gravitation, $n \geq 1$ is the Glen's flow law exponent [20] and C is a position-dependent function [23, 39, 41].

The surface mass balance, in turn, is modelled with the following four-parameter expression:

$$\mathcal{B} = \frac{m_1 + m_2}{2}\left(b + \mathcal{H} - z_{\text{ELA}}\right) - \sqrt{\left(\frac{m_1 - m_2}{2}\left(b + \mathcal{H} - z_{\text{ELA}}\right)\right)^2 + \varepsilon}, \tag{5}$$

where z_{ELA} is the equilibrium line altitude (ELA), m_1 and m_2 are the mass–balance gradients below and above the ELA, respectively, and $\varepsilon > 0$ is a regularization parameter. Such a mass–balance parameterization describes snow accumulation above and snow ablation below the ELA, with rates m_2 and m_1, respectively, $m_2 \leq m_1$. The surface mass balance (5) is widely used with $\varepsilon = 0$.

From the computed thickness $\mathcal{H}$, the elevation s_f of the glacier's surface topography at final time is deduced from

$$s_f = b + \mathcal{H}\Big|_{t=t_f}. \tag{6}$$

An implicit finite-difference scheme, centred in space, that solves problem (1), is detailed in the Appendix.

3 Bedrock topography reconstruction methods

In this section, the inversion problem is first exposed. Then, the inversion methods are formulated, starting with the quasi-stationary inverse method [17, 55]. Next, the transient inverse method is formulated [55, 65], before the shape optimization algorithm is introduced. The computational details underlying the latter method are burdensome and reported by [56, 57]. A simplified analysis in two space dimensions without sliding is presented in the Appendix. Its formulation's complexity is its only inconvenience given that this method is, by far, the fastest and most accurate, as shown later.

3.1 Problem statement

Consider the transient problem corresponding to Eq. 1 in the case where the initial and final surface topographies and the surface mass–balance $\mathcal{B}$ are known and the bedrock topography is unknown. The initial surface topography is denoted by s_i and the observed final surface topography is s^{obs}. The aim is to find such a bedrock topography b that minimizes the discrepancy between the known final surface topography s^{obs} and the computed final surface topography $s_f = b + \mathcal{H}\big|_{t=t_f}$ with Eq. 1. The inversion problem has two unknowns: the surface geometry s, which is only known at times t_i and t_f and the bedrock topography b (or, equivalently, the ice thickness $\mathcal{H}$). The time-dependent inverse problem can be stated as follows: *given the glacier surface topographies s_i at time t_i and s^{obs} at time t_f as well as the surface mass-balance time series $\mathcal{B}$ in the time interval $[t_i, t_f]$, find the bedrock geometry $b : \Omega_\perp \to \mathbb{R}$ such that*

$$s^{\text{obs}} = s_f, \qquad \text{in} \qquad \Omega_\perp, \tag{7}$$

under the constraint that b and $\mathcal{H}$ satisfy the shallow ice Eq. 1. In other words, this means that a bedrock topography b is sought such that the glacier evolves from surface s_i to s^{obs} over the time interval $[t_i, t_f]$ under the influence of the surface mass–balance $\mathcal{B}$. In this paper, the glacier's outline is assumed as time dependent. In particular, it may be that part of the bedrock topography inside domain $\Omega_\perp$ is known (*e.g.* the red part of the bedrock topography in Fig. 2). Let $\Omega_\perp^* \subset \Omega_\perp$ be the domain where the bedrock topography is known (green part of the map domain in Fig. 2). In this domain, the bedrock topography is denoted by b^* so that $b = b^*$ is enforced in $\Omega_\perp^*$. Additionally, because the source term $\mathcal{B}$ can take negative values, the bedrock topography b must satisfy

$$b < \min\{s_i, s^{\text{obs}}\}, \qquad \text{in} \qquad \Omega_\perp \setminus \Omega_\perp^*, \tag{8}$$

to make any of the below methods applicable. Note the strict inequality. This leads to the definition of the space of admissible bedrock topographies [54, 57]

$$\mathcal{U}^{\text{ad}} = \Big\{ b \in \mathcal{C}^0(\Omega_\perp) : b = b^* \text{ in } \Omega_\perp^* \text{and } b \leq \min\{s_i, s^{\text{obs}}\} \text{in } \Omega_\perp \setminus \Omega_\perp^* \Big\}, \tag{9}$$

where $\mathcal{C}^0(\Omega_\perp)$ is the set of continuous functions on $\Omega_\perp$.

3.2 Quasi-stationary inverse method

The motivation behind the quasi-stationary inverse method is the straightforward automatization of procedure [17] in the shallow ice context. Basically, it consists in solving the stationary version of Eq. 1 with a modified source term that accounts for both the surface mass–balance and the measured surface topography evolution in time. In the quasi-stationary approximation, surface topography evolution is assumed uniform in time. In other words, this means that the ice thickness change rate $\frac{\partial \mathcal{H}}{\partial t}$ in Eq. 1 is approached with $\frac{s^{\text{obs}} - s_i}{t_f - t_i}$, so that the evolution $(1)_1$ is transformed into the stationary equation

$$\frac{s^{\text{obs}} - s_i}{t_f - t_i} = \nabla \cdot (\mathcal{D} \nabla s) + \mathcal{B}. \tag{10}$$

The so-called apparent surface mass-balance [17] is defined by

$$\tilde{\mathcal{B}} = \mathcal{B} - \frac{s^{\text{obs}} - s_i}{t_f - t_i}. \tag{11}$$

Then, the quasi-stationary formulation of the transient inverse problem in the SIA is: *given the glacier surface topographies s_i and s^{obs} as well as the surface mass–balance time series $\mathcal{B}$ in the time interval $[t_i, t_f]$, find the thickness $\mathcal{H} : \Omega_\perp \to \mathbb{R}$ such that*

$$\begin{cases} \nabla \cdot (\mathcal{D} \nabla s) = -\tilde{\mathcal{B}}, & \text{in } \Omega_\perp \setminus \Omega_\perp^* \\ \mathcal{H} > 0, & \text{in } \Omega_\perp \setminus \Omega_\perp^* \\ \mathcal{H} = 0, & \text{in } \Omega_\perp^*, \end{cases} \tag{12}$$

where s can be any linear combination of s_i and s^{obs} and $\tilde{\mathcal{B}}$ is given by Eq. 11. In this paper, $s = s^{\text{obs}}$ is chosen. When the thickness $\mathcal{H}$ is known, the bedrock topography can be deduced from the relation $b = s - \mathcal{H}$. However, in general, the divergence (12) cannot be solved analytically. An approximation of the solution can be found after transformation into the following regularized, pseudo-time-dependent formulation [46, 55]:

$$\begin{cases} -\varepsilon_{\text{QSIM}} \nabla \cdot (\tilde{\mathcal{D}} \nabla \mathcal{H}) + \dfrac{\partial \mathcal{H}}{\partial \theta} = \nabla \cdot (\mathcal{D} \nabla s) + \tilde{\mathcal{B}}, & \text{in } \Omega_\perp \setminus \Omega_\perp^* \\ \mathcal{H} > 0, & \text{in } \Omega_\perp \setminus \Omega_\perp^* \\ \mathcal{H} = 0, & \text{in } \Omega_\perp^*, \end{cases} \tag{13}$$

$\theta > 0$, where $\varepsilon_{\text{QSIM}} > 0$ is a regularization parameter ("QSIM" stands for quasi-stationary inverse method), θ is a virtual time, the surface topography s is fixed to $s = s^{\text{obs}}$ and

$$\tilde{\mathcal{D}} = \Big(\Gamma \mathcal{H} + \Gamma_s \Big) \mathcal{H}^{n+1} \Big\| \nabla \mathcal{H} \Big\|^{n-1}. \tag{14}$$

Virtual time θ in Eq. 13 is a purely numerical artifice introduced to solve problem Eq. 12. The stationary solution of Eq. 13 is a regularized approximation of the solution of problem Eq. 12. A numerical method that solves problem Eq. 13 is described in the Appendix. This method is easy to implement because it only requires slight modifications of

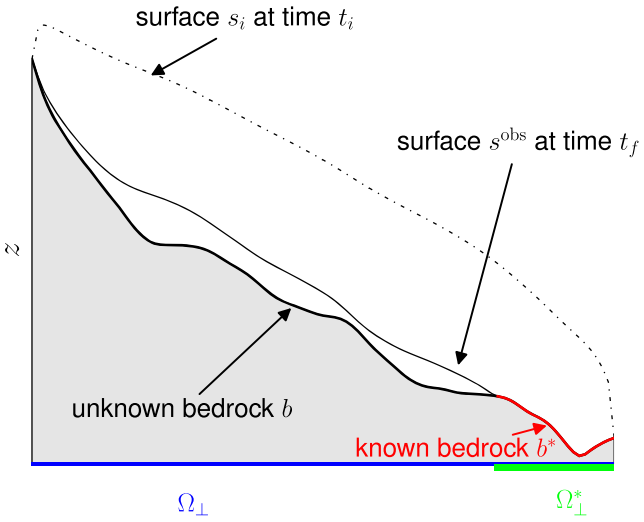


Fig. 2 Two-dimensional glacier profile illustrating the surface topographies s_i at time t_i and s^{obs} at time t_f. A profile of the glacier's map domain $\Omega_\perp$ is indicated in *blue* and the *subset* $\Omega_\perp^*$ of $\Omega_\perp$ where the bedrock topography is known and marked in *green*

the forward solver. However, it aims at finding a stationary geometry whose existence is not guaranteed a priori.

3.3 Transient inverse method

The most important drawback of the previous method is the linear approximation of the glacier's time evolution. The method tailored hereafter takes the glacier's transient evolution into account more accurately. Let $m > 0$ be an iteration index, b^m the bedrock at iteration m and b^0 some initial guess for b. In essence, the idea is to iteratively update the bedrock with the surface topography discrepancy at final time, which is equivalent to superimpose the current bedrock b^m and its updated version $s^{\text{obs}} - \mathcal{H}^m\big|_{t=t_f}$ [55, 65]:

$$\begin{aligned} b^{m+1} &= b^m + \beta\left(s^{\text{obs}} - s_f^m\right) = b^m + \beta\left(s^{\text{obs}} - b^m - \mathcal{H}^m\Big|_{t=t_f}\right) \\ &= (1-\beta)b^m + \beta\left(s^{\text{obs}} - \mathcal{H}^m\Big|_{t=t_f}\right), \end{aligned} \tag{15}$$

where $s_f^m = b^m + \mathcal{H}^m\big|_{t=t_f}$ is the approximated final surface topography at iteration m (see Fig. 3) and $0 < \beta \leq 1$ a relaxation parameter. In addition to this update, a smoothing of the computed ice thickness is necessary to provide convergence. This is done by including a Laplacian term in Eq. 15. Whatever the forward model is, the procedure remains unchanged: first, the forward simulation is performed in order to determine the final surface topography. Then, that topography is compared to the measured topography s^{obs} and the discrepancy is added to the bedrock. This approach is close to that exposed by [32] but, here, steady as well as transient problems can be solved. Therefore, the procedure is the following: let $\delta > 0$. Then,

1. **initialize** the bedrock topography with b^0
2. **while** $\|b^{m+1} - b^m\| > \delta$
   (a) **solve** the forward problem (1) with bedrock topography b^m and initial surface s_i, whose output is the ice thickness $\mathcal{H}^m\big|_{t=t_f}$ at final time;
   (b) **solve**

$$\begin{cases} H^m = \mathcal{H}^m\big|_{t=t_f} + \beta\left(b^m + \mathcal{H}^m\big|_{t=t_f} - s^{\text{obs}}\right) + \varepsilon_{\text{TIM}}\beta\triangle H^m, & \text{in } \Omega_\perp \setminus \Omega_\perp^*, \\ H^m = 0, & \text{in } \Omega_\perp^*, \end{cases} \tag{16}$$

   for H^m, where $0 < \beta \leq 1$ is a relaxation parameter and $\varepsilon_{\text{TIM}} \geq 0$ a regularization parameter;
   (c) **set**

$$b^{m+1} = s^{\text{obs}} - H^m. \tag{17}$$

This method aims at minimizing local surface topography discrepancy at time t_f. It is easy to implement because it relies on the iterative use of the forward model. Consequently, it has the further advantage of being as memory-demanding as the forward solver. However, the procedure is based on a local bedrock-to-surface perturbation transfer, which is not occurring in nature [67].

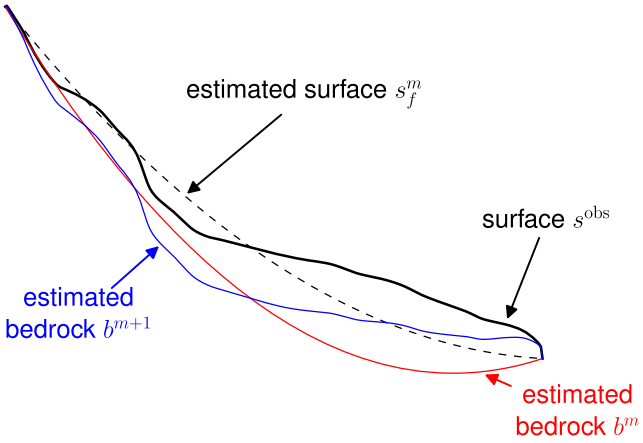


Fig. 3 Example glacier profile illustrating Eq. 15. The idea is to report the discrepancy $s^{\text{obs}} - s_f^m$ on the bedrock topography at each iteration. The bedrock is lifted where $s_f^m < s^{\text{obs}}$ and lowered otherwise

3.4 Shape optimization method

The last method consists in seeking the optimal bedrock elevation b such that the misfit between computed and observed surface topographies at final time t_f is as small as possible. This is an optimal control problem where the ice thickness $\mathcal{H} = s - b$ is termed "state variable" and the bedrock topography b "control variable". The misfit functional to be minimized is

$$\mathcal{J}(\mathcal{H}, b) = \frac{1}{2}\int_{\Omega_\perp} (s_f(\mathcal{H}, b) - s^{\text{obs}})^2 \mathrm{d}x + \varepsilon_{\text{SOA}}\mathcal{R}(b), \tag{18}$$

where the first term represents the misfit between computed and observed surface topographies, $\mathcal{R}$ is a Tikhonov regularization [74] and $\varepsilon_{\text{SOA}} \geq 0$ is a regularization parameter. Traditionally, the regularization's purpose is to minimize the topography's potential or even its curvature [38, 60], that is the bedrock topography's gradient or its Laplacian. However, [54] showed that such regularizations do not provide satisfactory estimations in practical applications of the shape optimization method introduced here. The more appropriate minimal perimeter constraint is advocated that turns $\mathcal{R}(b)$ into the area spanned by the bedrock topography b [57].

Minimizing misfit (18) is slightly different from minimizing discrepancy between computed and observed surface topographies, as addressed previously. Here, the cost

functional (18) is minimized, which corresponds to finding a bedrock topography that minimizes the overall surface topography misfit. In both the quasi-stationary and transient inverse methods, the surface topography discrepancy is minimized locally and everywhere in the computational domain.

The shallow ice Eq. 1 links control b to state variable $\mathcal{H}(b)$. The cost functional to be minimized can thus be expressed as

$$J(b) = \mathcal{J}(\mathcal{H}(b), b). \tag{19}$$

The shape optimization problem is cast into: *find*

$$\underset{b\in\mathcal{U}^{\mathrm{ad}}}{\operatorname{argmin}}\, J(b) \tag{20}$$

under the constraint that the ice thickness and the bedrock topography satisfy problem Eq. 1, where $\mathcal{U}^{\mathrm{ad}}$ is defined by Eq. 9. In order to minimize the cost functional (19), the solution of equation

$$\frac{\mathrm{d}J}{\mathrm{d}b}(b) = 0 \tag{21}$$

is sought in the space $\mathcal{U}^{\mathrm{ad}}$, which is realized by means of a quasi-Newton method [45, 62] coupled to a projection of b onto $\mathcal{U}^{\mathrm{ad}}$. Let m be an iteration index and b^m the bedrock topography at iteration m. The $m+1$ iterate is found by solving

$$b^{m+1} = b^m + \left(\frac{\mathrm{d}^2 J}{\mathrm{d}b^2}(b^m)\right)^{-1} \frac{\mathrm{d}J}{\mathrm{d}b}(b^m), \tag{22}$$

where the gradient of J is computed exactly with a primal–dual method while its Hessian is approximated with the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method [7, 18, 22, 72]. An intuition of the primal–dual method is given in the Appendix in the non-regularized, non-sliding, two-dimensional case. The full three-dimensional computation is reported by [57]. Basically, the method advocated here consists in augmenting cost functional (18) to a Lagrangian functional that incorporates the shallow ice constraints (1) in a "first discretize, then optimize" approach [33]. In the minimization process, forward problem (1) and a backward-in-time, diffusion–transport problem, the so-called primal and dual problems, respectively, are solved iteratively so that the cost functional's gradient can actually be evaluated. In summary, the shape optimization procedure is the following: let $\delta > 0$. Then,

1. **Set** glacier geometry with initial bedrock b^0 and surface s_i
2. **while** $|J(b^{m+1}) - J(b^m)| > \delta$
   (a) **solve** the primal problem (1)
   (b) **solve** the dual problem of Eq. 1
   (c) **compute** the gradient $\frac{\mathrm{d}J}{\mathrm{d}b}$ from the value of the state (solution to the primal problem) and the costate (solution to the dual problem) variables
   (d) **solve** Eq. 22 with the Bounded, Limited Memory, Variable Metric method (BLMVM) of the Toolkit for Advanced Optimization (TAO, [61]), where the Hessian is approximated with the BFGS method
   (e) **project** b^{m+1} onto space of admissible controls $\mathcal{U}^{\mathrm{ad}}$.

The procedure can be extended to concurrently compute some model parameters, like surface mass balance or rheology parameters [57], contrary to the quasi-stationary and transient inverse methods. This, however, goes beyond the scope of this contribution.

The shape optimization algorithm is a direct extension of the transient inverse method that takes into account the non-locality of the bedrock-to-surface perturbation transfers, which are known to be non-trivial [67]. To understand how the two methods are linked together, let us adopt a more straightforward way to update bedrock topography than the quasi-Newton method (22). The analysis that follows is exclusively aimed at building the theoretical relation between the two algorithms. In practice, bedrock topography reconstruction is performed with the quasi-Newton method (22). As in the transient inverse method, introduce a pseudo-time θ and consider the following pseudo-transport equation [3, 77]:

$$\frac{\partial b}{\partial \theta} + \boldsymbol{w}\cdot\nabla b = 0, \tag{23}$$

with pseudo-velocity

$$\boldsymbol{w} = \mathcal{V}\frac{\nabla b}{\|\nabla b\|^2}, \tag{24}$$

where $\mathcal{V}$ is defined, for $\varepsilon_{\mathrm{SOA}} = 0$, by weak expression [57]

$$\begin{aligned}\frac{\mathrm{d}J}{\mathrm{d}b}\hat{b} = \int_{\Omega_\perp} \mathcal{V}\hat{b}\mathrm{d} - s^{\mathrm{obs}} + \lambda\Big|_{t=t_i} b\Big]\mathrm{d}\boldsymbol{x}x = \int_{\Omega_\perp}\Big[\Big(s_f \\ - \int_{t_i}^{t_f}\int_{\Omega_\perp} \lambda\frac{\partial\mathcal{B}}{\partial b}\hat{b}\mathrm{d}\boldsymbol{x}\mathrm{d}t \\ - \int_{t_i}^{t_f}\int_{\Omega_\perp} \nabla\cdot\Big[(n-1)\mathcal{G}\nabla s\cdot\nabla\lambda\nabla s + \mathcal{D}\nabla\lambda\Big]\hat{b}\mathrm{d}\boldsymbol{x}\mathrm{d}t \\ - \int_{t_i}^{t_f}\int_{\Omega_\perp}\left(\frac{\rho g}{C}\right)^n \mathcal{H}^{n+1}\|\nabla s\|^{n-1}\nabla s\cdot\nabla\lambda\hat{b}\mathrm{d}\boldsymbol{x}\mathrm{d}t,\end{aligned} \tag{25}$$

for any continuous function $\hat{b}$, where λ is the dual variable of $\mathcal{H}$ and $\mathcal{G}$ is a function of $\mathcal{H}$ and b (see also Appendices). From transport Eq. 23, a variation δb of b obeys

$$\delta b = -\boldsymbol{w}\cdot\nabla b\delta\theta, \tag{26}$$

where $\delta\theta$ is a pseudo-time variation. The choice of pseudo-velocity (24) is made because, in this case,

$$\frac{\mathrm{d}J}{\mathrm{d}b}\delta b = -\int_{\Omega_\perp} \mathcal{V}^2 \delta\theta \mathrm{d}\boldsymbol{x} < 0, \tag{27}$$

that is, (23), ensures a decrease of cost functional J. Explicitly discretized in time, transport Eq. 23 becomes, for iteration index m,

$$\begin{aligned} b^{m+1} = b^m + \Delta\theta\left(s^{\mathrm{obs}} - s_f^m\right) \\ -\Delta\theta\Bigg[\lambda^m\Big|_{t=t_i} - \int_{t_i}^{t_f}\Big[\lambda^m\frac{\partial\mathcal{B}^m}{\partial b} \\ + \nabla\cdot\left((n-1)\mathcal{G}^m\nabla s^m\cdot\nabla\lambda^m\nabla s^m + \mathcal{D}^m\nabla\lambda^m\right) \\ + \left(\frac{\rho g}{C}\right)^n\left(\mathcal{H}^m\right)^{n+1}\|\nabla s^m\|^{n-1}\nabla s^m\cdot\nabla\lambda^m\Big]\mathrm{d}t\Bigg], \end{aligned} \tag{28}$$

which is precisely (15) augmented with additional terms that account for the way a bedrock topography perturbation is transferred up to the surface topography and the regularization. The quasi-Newton approach Eq. 22 is a sophisticated way of solving Eq. 28.

The quasi-Newton method (22) is much more difficult to implement than the previous transient and the quasi-stationary inverse methods. It requires solving of the dual problem: it is a backward-in-time diffusion-transport problem, which can be either computed by hand or generated automatically by means of automatic differentiation [31, 70]. Its solving relies on storage of the ice thickness $\mathcal{H}$ at every time step, which makes the method much more memory-demanding. Finally, contrary to the quasi-stationary and transient inverse methods, this sophisticated method needs external algorithms that make use of the computed gradient $\frac{\mathrm{d}J}{\mathrm{d}b}$ in order to calculate a reasonable descent direction as well as the BFGS approximation of the Hessian matrix $\frac{\mathrm{d}^2J}{\mathrm{d}b^2}$. All these algorithms can be used as a black box through the TAO library.

The three methods presented in this section depend on a regularization parameter ($\varepsilon_{\mathrm{QSIM}}$, $\varepsilon_{\mathrm{TIM}}$ and $\varepsilon_{\mathrm{SOA}}$) that sets a compromise between data under- and overfitting. In the transient inverse method and the shape optimization algorithm, the choice of this parameter is based on an L-curve criterion [11, 57, 65]. In the quasi-stationary inverse method, a small enough parameter that allows convergence is chosen.

To conclude, none of the proposed methods guarantees that the computed estimation corresponds to the global minimum of the cost functional. Genetic algorithms or particle swarm methods [34, 43] are known to be more efficient to get the global minimum at the expense of a much slower convergence.

4 Numerical results

The purpose of this section is to first validate the inversion methods presented in the previous section on synthetic data, compare their performance and then apply them to a measured real-world geometry. Both the synthetic and the real-world calculations are performed on the same glacier, namely Gries glacier, Swiss Alps, for which bedrock topography measurements are available [16]. Hereafter, the initial surface topography s_i is always the topography measured in 2003. In the synthetic validation, the input values of the model parameters are chosen arbitrarily and the final surface topography is generated by means of Eq. 1. The feasibility of the inversion of the shallow ice model is demonstrated in this case. A qualitative estimation of the error due to the purely mathematical inversion is shown, which is free of any other kind of errors, such as errors in the measurements or the physical model. In the real-world validation, the model parameters are set to their measured values on this particular glacier and the final surface topography is the surface topography measured in year 2007.

4.1 Synthetic validation

To validate the three algorithms, a set of bedrock topographies, surface topographies and surface mass–balance time series is assumed that satisfies Eq. 1 on a synthetic geometry inspired from Gries glacier, Swiss Alps (see the large map of Fig. 6). The time frame considered below is $t_f - t_i = 5$a (5 years) and sliding is included. In other words, the measured bedrock and initial surface geometries of the glacier [16] are used to construct the initial state of the reference

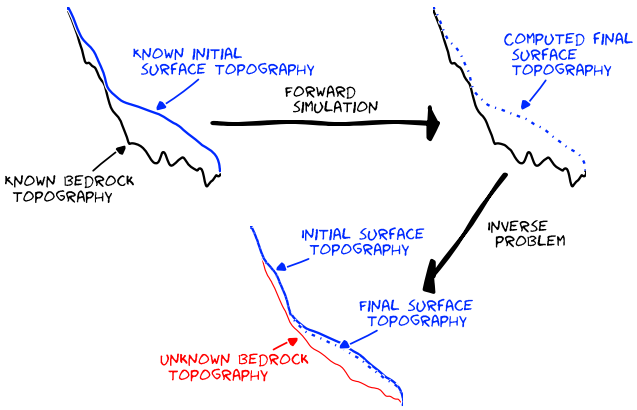


Fig. 4 Numerical validation process of the inversion methods: first, a forward simulation is performed over some time frame from a known bedrock and initial surface topography, whose outcome is a computed final surface topography (*top figures*). Then, the bedrock topography is assumed unknown and aimed at being reconstructed from the initial surface topography and the previously computed final surface topography. An initial guess is required for each method, depicted in *red* on the *bottom figure*

forward simulation so that the chosen geometry is a relevant representative of alpine glaciers. The target (or observed) surface topographies are generated synthetically by means of the shallow ice Eq. 1. The validation process is illustrated by Fig. 4: first, from the known bedrock and initial surface topographies, the simulation of the glacier's evolution with Eq. 1 provides a reference surface topography at the final time $t_f = 5$a. Then, the glacier's bedrock topography is assumed unknown and each algorithm is applied for its reconstruction.

For validations, the coefficient C in expression $(4)_2$ is set to $C = 1$ in the light-grey region of the map domain depicted by Fig. 6. The surface mass–balance parameters are arbitrarily set to $z_{\mathrm{ELA}} = 2,800$ m, $m_1 = m_2 = 0.01$ w.e.a^{-1} (water equivalent per year) and $\varepsilon = 0.1$ m^{2}a^{s-2}, while the rheology parameters are $A = 0.1$ bar^{-3}a^{-1} and $n = 3$.

The numerical experiments are performed on a Cartesian grid of space step $\Delta x = 50$ m. The time step is chosen in such a way that the numerical scheme is stable (see Appendix), namely $\Delta t = 10^{-2}$ a for both the transient inverse method and the shape optimization algorithm and $\Delta t = 10^{-6}$ a for the quasi-stationary inverse method. The regularization parameters of the quasi-stationary inverse method, the transient inverse method and the shape optimization algorithm are set to $\varepsilon_{\mathrm{QSIM}} = 0.1$, $\varepsilon_{\mathrm{TIM}} = 100$ and $\varepsilon_{\mathrm{SOA}} = 10^{-4}$, respectively. The regularization parameters are chosen in such a way that the methods converge within a reasonable time or satisfy an L-curve criterion [57, 65]. All the methods presented in this paper converge in the sense that the misfit between two successive bedrock topography iterates goes to zero when the number of iterations goes to infinity. Each algorithm is stopped when convergence is judged satisfactory, that is to say when the bedrock topography misfit is small enough. The transient inverse method's relaxation parameter is set to $\beta = 1$. Large values of β accelerate the inversion process at the expense of numerical stability. This parameter has an upper bound that should not be exceeded [65].

The fastest procedure is the shape optimization algorithm. The quasi-stationary inverse method converges with the highest difficulty. All these algorithms run within a few hours on a single processor. A systematic comparison of the time needed by the transient inverse method and the shape optimization algorithm is performed by [56] for the two-dimensional shallow ice model.

The results of the inversion of the synthetic data are shown in Figs. 5 and 6. For each method, Fig. 5 shows the

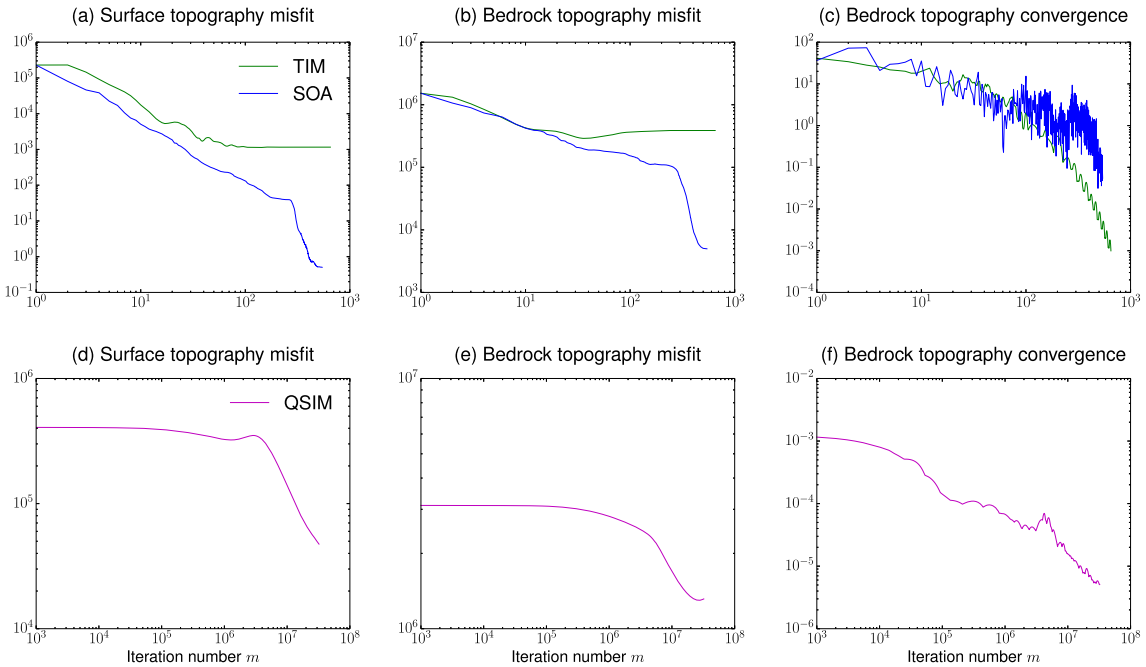


Fig. 5 Evolution of the surface and bedrock topography misfits (**a**, **b**, **d** and **e**) and convergence of the bedrock topography (**c** and **f**) as functions of the iteration number m when the transient inverse method (TIM, *green curve*), the shape optimization algorithm (SOA, *blue curve*) and the quasi-stationary inverse method (QSIM, *purple curve*) are applied to a synthetic geometry inspired from Gries glacier, Swiss Alps. Recall that the number of iterations is not comparable in the two rows. Plots **a** and **d** show the surface topography misfit (18) without regularization. Plots **b** and **e** display the bedrock topography misfit. In **c** and **f**, the maximal discrepancy between two successive bedrock topography estimations is plotted. The quasi-stationary inverse method is applied with $\varepsilon_{\mathrm{QSIM}} = 0.1$ and $\Delta t = 10^{-6}$a, the transient inverse method with $\beta = 1$, $\varepsilon_{\mathrm{TIM}} = 100$ and $\Delta t = 10^{-2}$a and the shape optimization algorithm with $\varepsilon_{\mathrm{SOA}} = 10^{-4}$ and $\Delta t = 10^{-2}$a. The grid resolution is $\Delta x = 50$m

differences in convergence and accuracy of the solution. It can be seen that the shape optimization algorithm lets the surface topography misfit (Eq. 18) decrease the most significantly. A minimal surface topography misfit is also achieved with the transient inverse method. However, it lies several orders of magnitude higher than that obtained with the shape optimization algorithm. The quasi-stationary inverse method provides another solution whose surface topography misfit is higher than that of the transient inverse method. The number of iterations in this case is much higher than in the two previous methods. In both the transient inverse method or the shape optimization algorithm, one iteration consists in essentially one simulation of the glacier's evolution over 5a followed by a smoothing or a dual step respectively. By contrast, one quasi-stationary iteration is the regularized simulation of the glacier's stationary evolution over one single time step Δt which makes this method the least demanding in terms of computational costs. The shape optimization algorithm is the most demanding because the dual problem must be solved for each time step. The transient inverse method has an intermediate computational cost since its smoothing procedure consists in solving a linear system only at the end of each simulation of the forward problem.

Figure 6 shows how accurate each inversion method is. Each subplot of Fig. 6 represents a profile across the glacier. Basically, all the estimations are satisfactory, the largest discrepancies being located at the glacier's tongue (profile (e)). The most accurate estimation is provided by the shape optimization algorithm that supplies a satisfactory ice thickness distribution everywhere in the domain, even near the glacier's tongue. The quasi-stationary inverse method performs the most poorly. In view of its simplicity, the results are, however, satisfactory. The obtained results confirm that the shape optimization algorithm is reliable and accurate. Various numerical experiments were performed

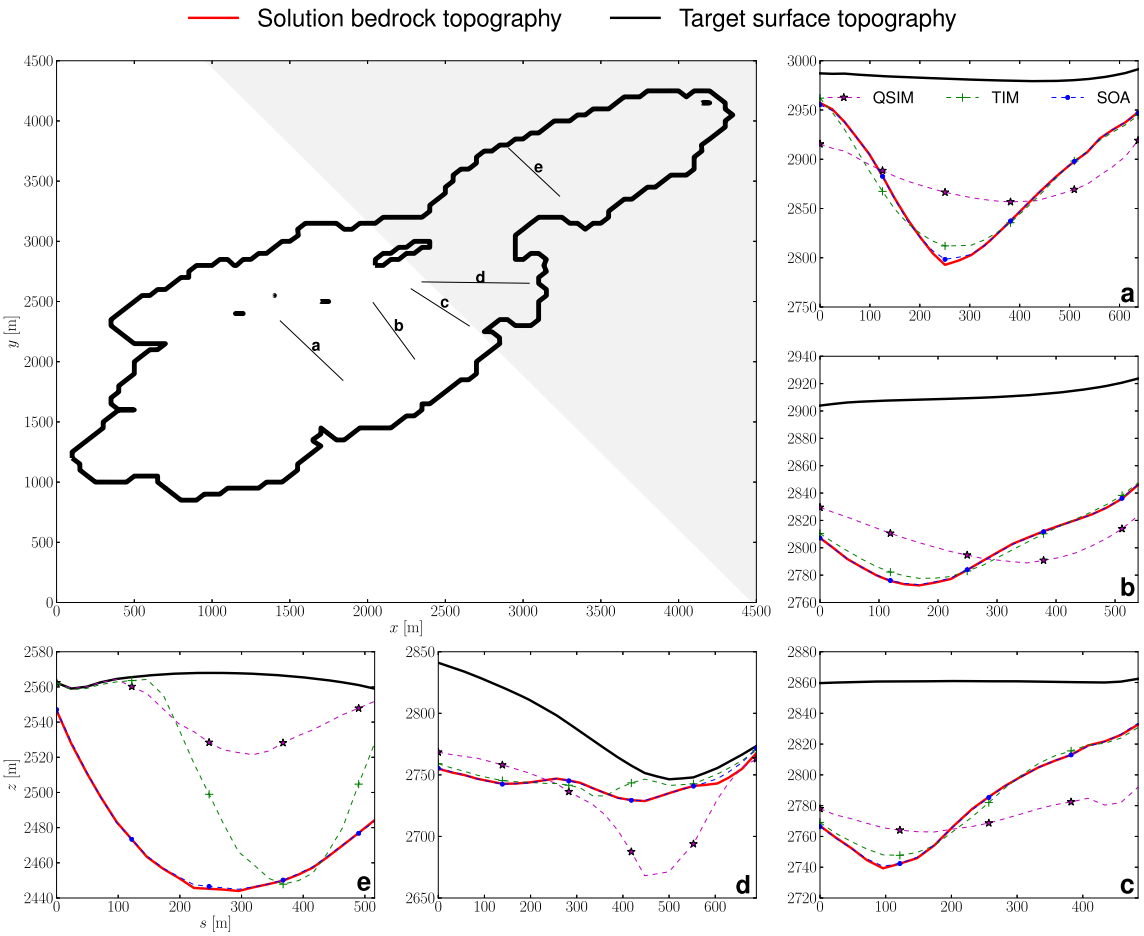


Fig. 6 Bedrock topography estimations along profile lines across a synthetic geometry inspired from Gries glacier, Swiss Alps. The large map delineates the glacier's outline at initial time and the sliding region (*light grey*), where the sliding coefficient is set to $C = 1$. On each subplot, the abscissa s and ordinate z represent the coordinate following the profile and the altitude of the depicted topographies, respectively. The depicted target surface topography (in *black*) is the surface obtained after simulation of the glacier's evolution with Eq. 1 from the solution bedrock topography, delineated by the *thick red curve*. The same colour code holds as in Fig 5. The best estimation is provided by the shape optimization algorithm

with different initial guesses, parameters and time frames [54]. Obviously, the larger the time interval between t_i and t_f, the larger the computational time. Moreover, with increasing time frame, the quasi-stationary inverse method performs more and more poorly and the shape optimization needs an initial guess that is closer and closer to the solution, essentially because it is based on the quasi-Newton method.

The main reason why the quasi-stationary inverse method performs the most poorly is the assumption that the ice thickness distribution evolves linearly in time. The apparent surface mass-balance description is too crude to make such a method accurate. In turn, the drawback of the transient inverse method is related to the way the bedrock-to-surface perturbation transfers occur [54, 67]. Due to their non-local behaviours, these transfers induce unwanted corrections to the bedrock topography that cannot be controlled efficiently and may impede convergence.

By contrast, the shape optimization method provides a bedrock topography estimation that is very close to the solution. This demonstrates the feasibility of an accurate bedrock topography reconstruction in the SIA. Because the synthetic inversions are performed with the exact physical model and input data, the defects of both the quasi-stationary and the transient inverse methods displayed at the glacier's tongue in Fig. 6e are necessarily due to the numerical method applied.

4.2 Real-world validation

The three methods are now applied to the measured surface topographies of Gries glacier in years 2003 and 2007. The surface mass–balance parameters were computed by [35] and correspond to the approximated annual averages of the mass-balance gradient and the ELA. Their values are $z_{\mathrm{ELA}} = 2,937$ m, $m_1 = m_2 = 0.0088$ w.e.a^{-1}. The surface mass-balance regularization parameter is set to $\varepsilon = 0.1$ m^{2}a^{-2}. No sliding is assumed (i.e. $C = \infty$ in Eq. 4) and the rheology parameters are set to $n = 3$ and $A = 0.076$ bar^{-3}a^{-1}. The regularization and relaxation parameters are set to $\varepsilon_{\mathrm{QSIM}} = 0.1$, $\varepsilon_{\mathrm{TIM}} = 100$, $\beta = 0.25$ and $\varepsilon_{\mathrm{SOA}} = 0.03$. None of the inversions is forced with the available bedrock topography measurements. The purpose is to compare the output of the three methods to actual measurements without forcing them with more data than the surface topography and mass–balance measurements.

The various bedrock topography estimations along some of the profiles represented in Fig. 6 are depicted by Fig. 7. The figure additionally displays the corresponding result of the ice thickness estimation method (ITEM) of [16]. The bedrock topography estimation is the most accurate on the top of the glacier (Fig. 7a) in each case. Near the glacier's tongue (Fig. 7e), most of the algorithms underestimate the thickness. In the middle of the glacier (Fig. 7b, d), it has

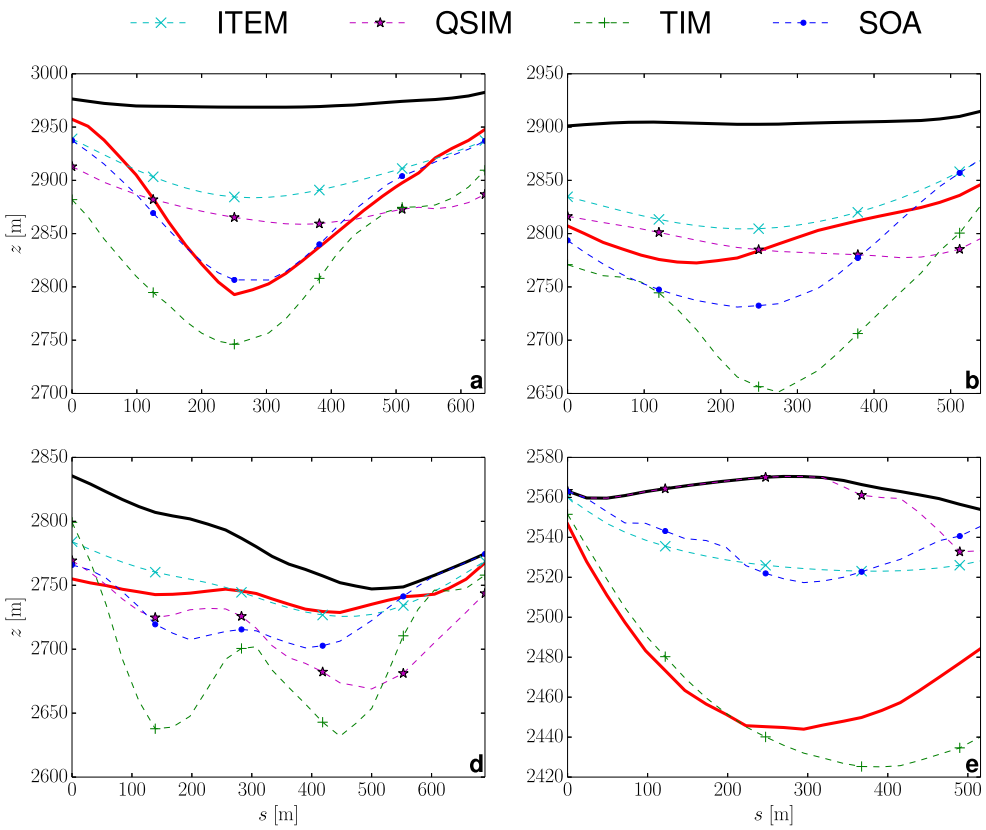


Fig. 7 Estimation of the bedrock topography from measurements collected on Gries glacier, Swiss Alps, with the ice thickness estimation method (*ITEM, cyan*), the quasi-stationary inverse method (*QSIM, magenta*), the transient inverse method (*TIM, green*) and the shape optimization algorithm (*SOA, blue*). The *black and red thick curves* delineate the observed final surface and the bedrock topographies, respectively. The abscissa s and ordinate z represent the coordinate following the profile and the altitude of the depicted topographies respectively. Bedrock topography measurements and sliding are not taken into account in the computations. The result of the ITEM is comparable to that of the QSIM

the tendency to be rather overestimated. Despite its performance on synthetic data, the shape optimization method provides an estimation that is particularly far from the measured topography near the glacier's tongue. Because the numerical efficiency of the method was demonstrated in the previous section, such a result can only be explained by the choice of too simple ice flow and surface mass–balance models and by the measurement errors contained in the surface topographies s_i and s_f.

4.3 Discussion

This section is devoted to a short discussion on the sensitivity of the shape optimization algorithm to the aforementioned quantities that are prone to errors when applied to real-world data. The sensitivity of the other methods is similar and was already discussed in the two-dimensional case for the surface mass balance and the surface topography [55].

In view of the investigation of the bedrock-to-surface perturbation transfers in [67], it is clear that the bedrock topography estimation is very sensitive to defects in the surface topography. Indeed, small perturbations of said topography are responsible for large oscillations in the subglacial topography estimation [55–57], which can be smoothed by means of a suitable regularization. However, the larger the regularization's effects, the less physically consistent the resulting estimation. In the analysis below, the regularization parameter is set to $\varepsilon_{\mathrm{SOA}} = 0.03$ in all cases.

Surface mass–balance data are available only for a limited number of glaciers over a limited time frame. Hence, calibration of parameters m_1, m_2, and z_{ELA} in [35] is subject to errors. Figures 8 and 9 show how the bedrock topography estimation is perturbed by the variation of one of these parameters while the other ones are kept fixed. As before, both the melting and accumulation rates are assumed to hold the same value. According to Fig. 8, an underestimated melting rate m_1 is responsible for an underestimation of the glacier's ice thickness everywhere in the domain. Larger values of m_1, that is to say in the neighbourhood of $m_1 = 0.01$ w.e.a^{-1}, yield more reasonable bedrock topographies. If m_1 is overestimated, then the ice thickness is essentially overestimated everywhere in the domain, the glacier's tongue being the most suffering region (Figs. 8 and 9e). This result is not surprising, given the fact that the optimization method is based on the comparison between a measured and a computed topography at final time. Large melting rates are responsible for fast melting at the glacier's tongue. Because the method's purpose is to let the computed surface topography match the measured one, a larger ice thickness is provided in this case than with small melting rates. The sensitivity of the bedrock topography estimation to the melting rate is high. By contrast, the effects of a misestimated equilibrium line altitude are smaller, as depicted by Fig. 9.

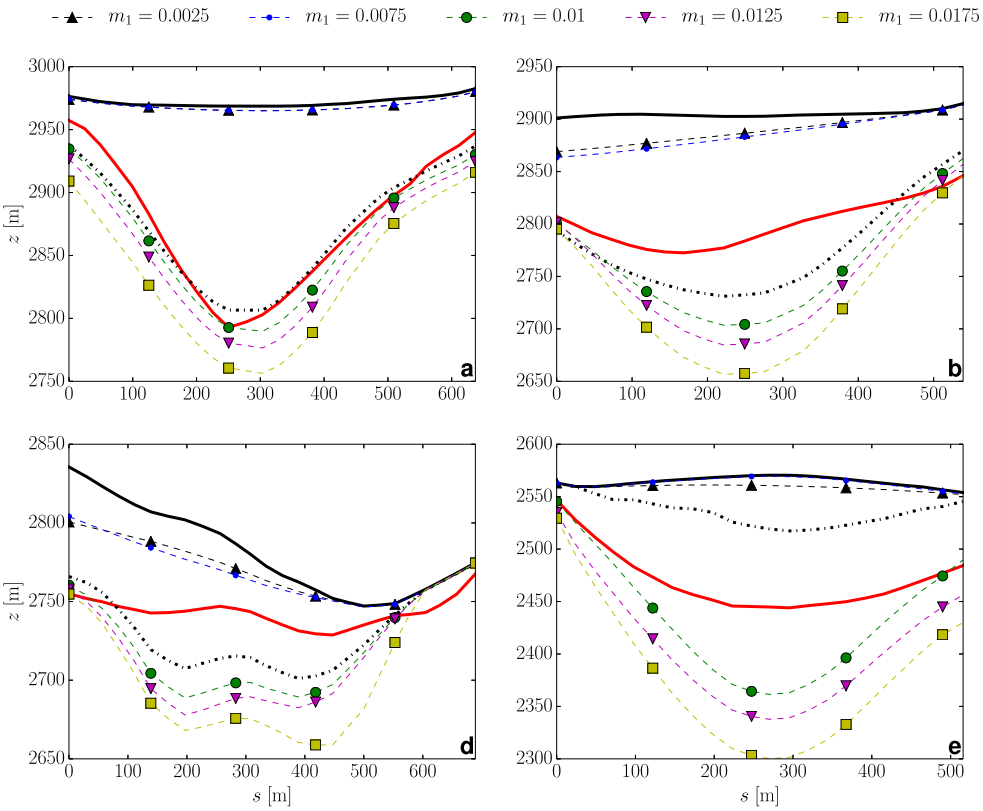


Fig. 8 Subglacial topography estimation with the shape optimization algorithm for various values of the melting rate m_1 (in w.e.a^{-1}), $m_2 = m_1$, $z_{\mathrm{ELA}} = 2,937$ m, rate factor $A = 0.076$ bar^{-3} a^{-1} and no sliding. The regularization parameter is set to $\varepsilon_{\mathrm{SOA}} = 0.03$. The *thick black, red and dash-dotted black curves* delineate the measured surface topography in 2007, the measured bedrock topography and the bedrock topography estimation obtained in Fig. 7, respectively. The *thin coloured curves* delineate the subglacial topography estimation with the shape optimization algorithm for various values of m_1. The abscissa s and ordinate z represent the coordinate following the profile and the altitude of the depicted topographies, respectively. The profile letters **a** to **e** refer to those shown on the large map of Fig. 6

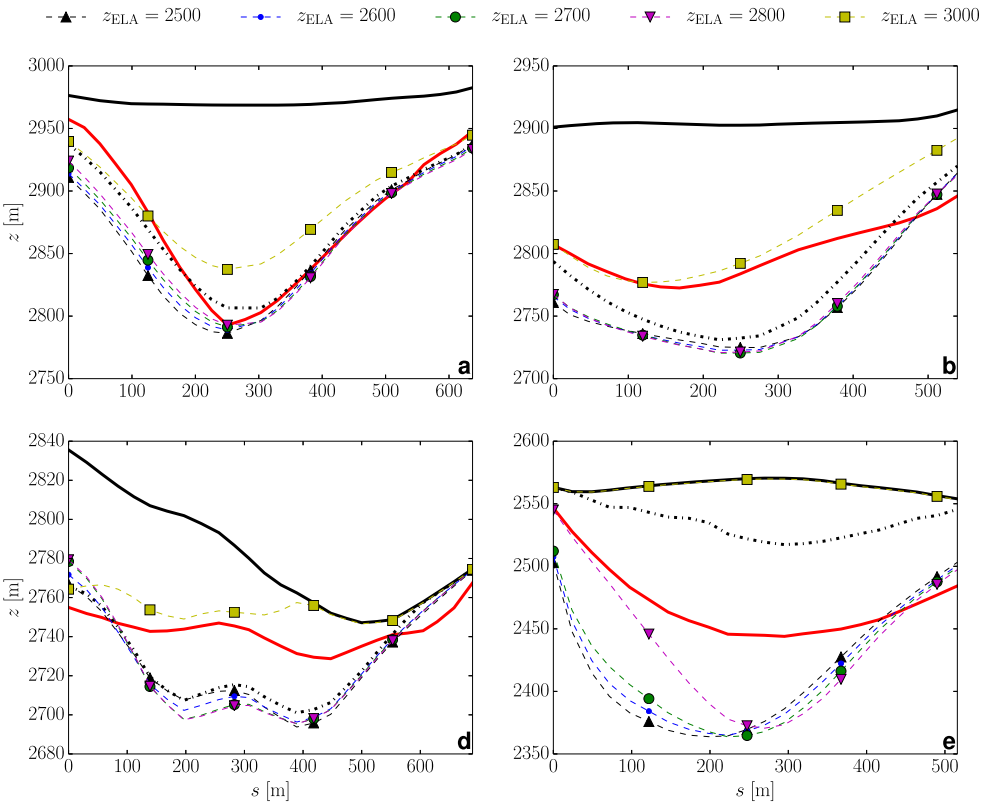


Fig. 9 Subglacial topography estimation with the shape optimization algorithm for various values of the equilibrium line altitude z_{ELA} (in meters), $m_1 = m_2 = 0.0088$ w.e.a^{-1}, rate factor $A = 0.076$ bar^{-3} a^{-1} and no sliding. The regularization parameter is set to $\varepsilon_{SOA} = 0.03$. The *thick black, red and dash-dotted black curves* delineate the measured surface topography in 2007, the measured bedrock topography and the bedrock topography estimation obtained in Fig. 7, respectively. The *thin coloured curves* delineate the subglacial topography estimation with the shape optimization algorithm for various values of z_{ELA}. The abscissa s and ordinate z represent the coordinate following the profile and the altitude of the depicted topographies, respectively. The profile letters **a** to **e** refer to those shown on the large map of Fig. 6

Again, the tongue is the most sensitive region of the glacier, where fluctuations in ice thickness are very large depending on the input z_{ELA}. With increasing z_{ELA}, the glacier's tongue is more likely to be thinner because melting is more important. Therefore, the whole glacier must be globally thinner. Conversely, a low equilibrium line altitude favours more accumulation, a consequence of which is an essentially thicker glacier. The effects of the surface mass-balance parameters are basically linear in these parameters.

The influence of rate factor A on the provided estimation is reported by Fig. 10. The smaller the rate factor, the higher the ice viscosity and, therefore, the lower the ice velocity. In this case, the whole glacier dynamics are slowed. It is difficult to explain clearly the effects depicted by Fig. 10 because they are strongly non-linear functions of the rate factor. The most significant effects of A are observed near the glacier's tongue (profile **e**). In the middle and at the top of the glacier, the effects are not significant.

Lastly, the effects of a finite sliding coefficient are shown on Fig. 11. In the real-world validation above, no sliding was assumed, which is unlikely in nature. As with an increasing rate factor, a decreasing sliding coefficient increases the ice velocity. While the effects of the rate factor are localized near the glacier's tongue, the sliding coefficient significantly changes the ice thickness everywhere in the domain, the largest differences occurring near the tongue.

The aforementioned parameters can be computed concurrently with the bedrock topography when surface ice velocities are available. However, it is difficult to estimate some of them, like namely the ELA and the shape of the sliding domain, because of the problem's ill-posedness [57]. Moreover, the available surface ice velocity data are usually insufficient to infer precise sliding and surface mass–balance distributions.

In addition to the incertitudes on the model parameters, approximation errors due to the shallow ice model (1) of flow are responsible for the misestimation of the subglacial topography. The Stokes approximation [23, 24, 41, 80] is better suited to the description of a mountain glacier's ice flow. However, the optimal control of a Stokes ice flow addresses several additional difficulties. The main issues are due to mesh deformation. There are basically two ways to tackle the problem: the mesh can be either deformed or fixed during the whole optimization process. Since topology changes often occur in glaciers, an extension of the so-called variational volume-of-fluid or level-set optimization [1, 3, 14, 79] is advocated for this purpose, where the mesh is never deformed. This, however, goes beyond the scope of this contribution.

Finally, it must be remembered that the above estimations can obviously be improved significantly when the available bedrock topography measurements are taken into account. At least two fundamentally different ways to proceed can be

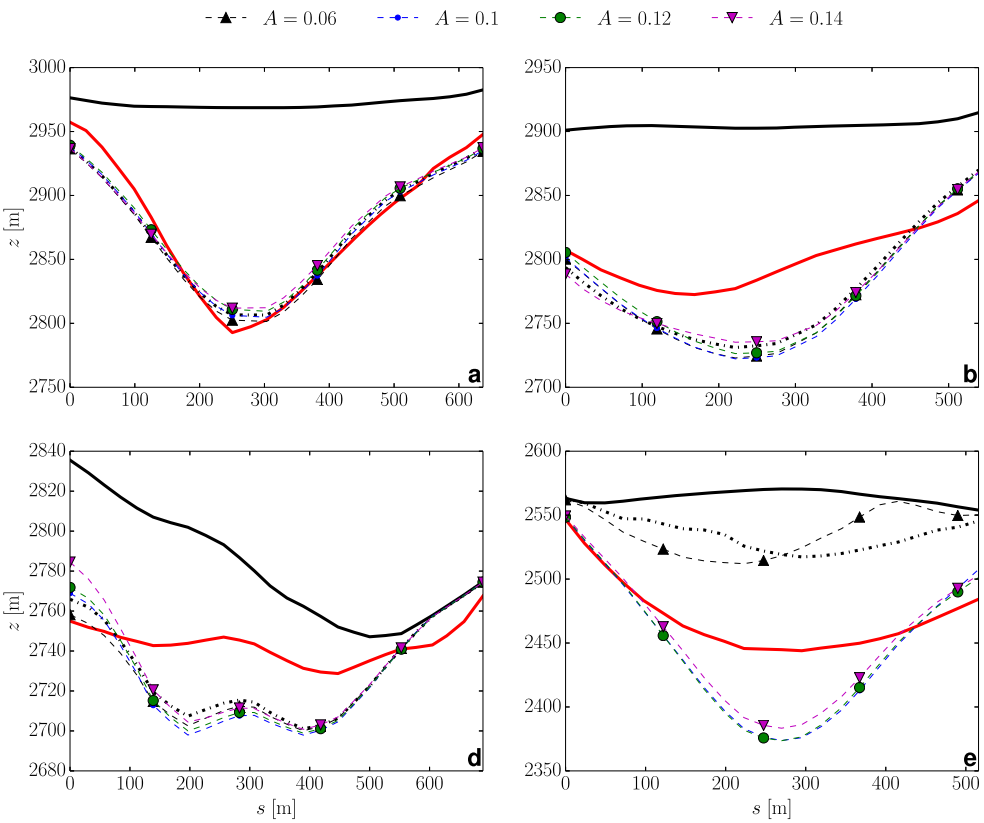


Fig. 10 Subglacial topography estimation with the shape optimization algorithm for various values of the rate factor A (in $\text{bar}^{-3}\text{a}^{-1}$), surface mass–balance parameters $m_1 = m_2 = 0.0088$ w.e.a^{-1}, $z_{\text{ELA}} = 2,937$ m and no sliding. The regularization parameter is set to $\varepsilon_{\text{SOA}} = 0.03$. The *thick black, red and dash-dotted black curves* delineate the measured surface topography in 2007, the measured bedrock topography and the bedrock topography estimation obtained in Fig. 7, respectively. The *thin coloured curves* delineate the subglacial topography estimation with the shape optimization algorithm for various values of A. The abscissa s and ordinate z represent the coordinate following the profile and the altitude of the depicted topographies, respectively. The profile letters **a** to **e** refer to those shown on the large map of Fig. 6

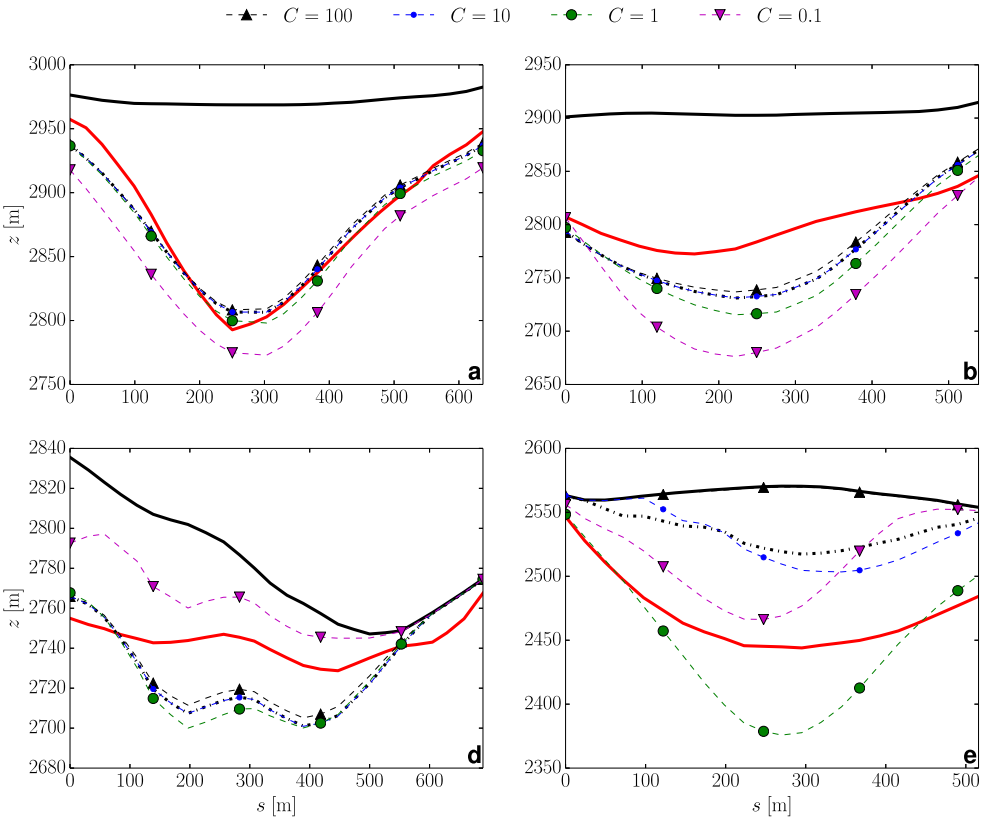


Fig. 11 Subglacial topography estimation with the shape optimization algorithm for various values of the sliding coefficient C (in bar $\text{a}^{\frac{1}{3}}$ $\text{m}^{-\frac{1}{3}}$), surface mass-balance parameters $m_1 = m_2 = 0.0088$ w.e.a^{-1}, $z_{\text{ELA}} = 2,937$ m and rate factor $A = 0.076$ $\text{bar}^{-3}\text{a}^{-1}$. The regularization parameter is set to $\varepsilon_{\text{SOA}} = 0.03$. The *thick black, red and dash-dotted black curves* delineate the measured surface topography in 2007, the measured bedrock topography and the bedrock topography estimation obtained in Fig. 7, respectively. The *thin coloured curves* delineate the subglacial topography estimation with the shape optimization algorithm for various values of A. The abscissa s and ordinate z represent the coordinate following the profile and the altitude of the depicted topographies, respectively. The profile letters **a** to **e** refer to those shown on the large map of Fig. 6

applied. First, in the shape optimization algorithm, the cost functional (18) can be augmented with the misfit between the measured and the computed bedrock topographies. Similarly, in the transient and quasi-stationary inverse methods, the ice thickness can be forced to take fixed values where bedrock topography measurements are available. In this case, each method supplies an estimation that is as close as desired to the measurements at the expense of worse convergence because the system is more constrained. Second, each method can be wrapped by yet another minimization method, for example, a particle swarm method [43], whose aim is to minimize the bedrock topography misfit when it is the output of one of the inversion methods presented above and for which the control variables are the model parameters. Nevertheless, a detailed description of such a meta-optimization algorithm [53] goes beyond the scope of this contribution.

5 Conclusion

This paper focuses on the analytical and numerical comparison of three conceptually different methods aiming at the estimation of a glacier's ice thickness distribution. Each approach is based on the three-dimensional, time-dependent, shallow ice approximation of glacier flow. The quasi-stationary inverse method handles the inverse problem as a stationary transport problem with the approximation of a uniform-in-time surface topography evolution. The transient inverse method, in turn, relies on the assumption of a local bedrock-to-surface topography response, which is only partially true, even in the SIA. The glacier's final surface topography is computed iteratively and compared to the observed surface topography. Their discrepancy is then reported onto bedrock topography until convergence is reached. These two procedures demand very little work once a forward model solver is available and only require small computer resources. However, they converge slowly. Based on the optimal control of the surface topography misfit, the more sophisticated shape optimization algorithm converges better and can compute model parameters concurrently with bedrock topography if surface ice velocity observations are available [57] at the expense of being both more difficult to implement and more memory-demanding. A mathematical relationship between this algorithm and the transient inverse method is established that describes how bedrock-to-surface perturbation transfers are taken into account to improve convergence.

The presented methods are validated on a synthetic geometry generated by the forward solver. Convergence, number of iterations, bedrock topography misfit and surface topography misfit are depicted for each method for the sake of comparison. The shape optimization algorithm converges within the smallest number of iterations. Given that each iteration of the method runs one single forward simulation most of the time (in addition to solving its corresponding linear dual equation), this makes this algorithm the fastest of the set. Moreover, it inverts the shallow ice equation with the highest accuracy. It is indeed responsible for an almost perfect agreement between the synthetic and the reconstructed bedrock topographies while the estimations obtained by means of the other methods suffer from significant defects, especially near the glacier's tongue. This shows that the shallow ice equation can be efficiently inverted with accuracy without any other artifice than the Tikhonov regularization.

Each method provides a satisfactory bedrock topography estimation when applied to real-world input measurements observed on Gries glacier, Swiss Alps, for which bedrock topography profiles are available [17]. The algorithms are compared in performance to the ITEM. The quasi-stationary inverse method is the closest estimation to the ITEM. The shape optimization algorithm and the transient inverse method provide improved bedrock topography estimations, especially at locations where the ice thickness is large. The deviations of obtained estimations from measured bedrock topography are much larger than in the synthetic cases. This phenomenon is explained by the possible unreliability of the input model parameters (that is the surface mass–balance, sliding and rheology parameters), as shown in various sensitivity experiments presented in the previous section. The measurement errors in surface topography are also responsible for defects in subglacial topography estimation. The best trade-off between data under- and overfitting is set by the value of the regularization parameter for each method, which is chosen by means of an L-curve criterion [57, 65]. Finally, the model order of the ice flow approximation plays a role in the bedrock topography reconstruction. However, the inversion of higher-order models is much harder and more resource-consuming. In order to keep the procedures as fast as possible, small enhancements of the SIA may be incorporated to make the presented algorithms more suitable for mountain glaciers [59].

The results presented in this contribution show that shape optimization of the ice flow model supplies the most reliable, versatile and physical subglacial topography estimation, which is sufficient motivation for developing its extensions to higher order flow models.

Acknowledgments The authors thank the anonymous reviewers for their relevant comments and Gil Michel for his useful advice on English writing.

Appendix

A Forward numerical scheme

An implicit, centred-in-space, finite-difference scheme is advocated to solve Eq. 1. Let $\Delta x > 0$ be some space step and $(\xi_j, \zeta_k) = (j,k)\Delta x, 0 \le j \le N_x+1, 0 \le k \le N_y+1$, vertices of the so-called "direct grid" $\tilde{\Omega}_\perp$ of $\Omega_\perp$. In particular, $L_x = (N_x+1)\Delta x$ and $L_y = (N_y+1)\Delta x$. Moreover, let $\Delta t > 0$ be some time step and $t^\ell = t_i + \ell\Delta t, 0 \le \ell \le M+1$, such that $t^0 = t_i$ and $t^{M+1} = t_f$.

Let us denote the approximation of the bedrock topography $b(\xi_j, \zeta_k)$, ice thickness $\mathcal{H}(\xi_j, \zeta_k, t^\ell)$, surface topography $s(\xi_j, \zeta_k, t^\ell)$ and surface mass-balance $\mathcal{B}(\xi_j, \zeta_k, t^\ell)$ at location (ξ_j, ζ_k) and time t^ℓ with $b_{jk}, \mathcal{H}^\ell_{jk}, s^\ell_{jk} = b_{jk}+\mathcal{H}^\ell_{jk}$, and $\mathcal{B}^\ell_{jk}$. Denote by $\boldsymbol{b}, \boldsymbol{\mathcal{H}}^\ell, \boldsymbol{s}^\ell$ and $\boldsymbol{\mathcal{B}}^\ell$ as the corresponding vectors. Moreover, let us write $\boldsymbol{s}_i$ and $\boldsymbol{s}_f$ as the vectors of components $s_i(\xi_j, \zeta_k)$ and $(s_f)_{jk} = b_{jk} + \mathcal{H}^{M+1}_{jk}$, that are the glacier's surface at initial and final times, respectively. Recall that the initial surface topography s_i is known while s_f is computed with the numerical scheme. The notations introduced here remain valid throughout all appendices.

The solving process is based on the equation's time-splitting. The finite-difference scheme, centred in space, discretizing (1), consists in initializing the ice thickness with

$$\mathcal{H}^0_{jk} = s_i(\xi_j, \zeta_k) - b_{jk} \tag{29}$$

for $1 \le j \le N_x$, $1 \le k \le N_y$; then, for $0 \le \ell \le M$, first solve the advection-diffusion scheme [2, 9, 10, 49, 71]

$$\frac{\overline{\mathcal{H}}^\ell_{jk} - \mathcal{H}^\ell_{jk}}{\Delta t} = \mathcal{T}^\ell_{jk}, \tag{30}$$

where $\mathcal{T}^\ell_{jk}$ is the discretization

$$\begin{aligned}\mathcal{T}^\ell_{jk} = {} & \frac{1}{2\Delta x}\left[\mathcal{D}^\ell_{j+\frac{1}{2},k}\frac{s^\ell_{j+1,k}-s^\ell_{jk}}{\Delta x} - \mathcal{D}^\ell_{j-\frac{1}{2},k}\frac{s^\ell_{jk}-s^\ell_{j-1,k}}{\Delta x}\right. \\ & \left. + \mathcal{D}^\ell_{j,k+\frac{1}{2}}\frac{s^\ell_{j,k+1}-s^\ell_{jk}}{\Delta x} - \mathcal{D}^\ell_{j,k-\frac{1}{2}}\frac{s^\ell_{jk}-s^\ell_{j,k-1}}{\Delta x}\right] \\ & + \frac{1}{2\Delta x}\left[\overline{\mathcal{D}}^\ell_{j+\frac{1}{2},k}\frac{\overline{s}^\ell_{j+1,k}-\overline{s}^\ell_{jk}}{\Delta x} - \overline{\mathcal{D}}^\ell_{j-\frac{1}{2},k}\frac{\overline{s}^\ell_{jk}-\overline{s}^\ell_{j-1,k}}{\Delta x}\right. \\ & \left. + \overline{\mathcal{D}}^\ell_{j,k+\frac{1}{2}}\frac{\overline{s}^\ell_{j,k+1}-\overline{s}^\ell_{jk}}{\Delta x} - \overline{\mathcal{D}}^\ell_{j,k-\frac{1}{2}}\frac{\overline{s}^\ell_{jk}-\overline{s}^\ell_{j,k-1}}{\Delta x}\right]\end{aligned} \tag{31}$$

of $\nabla\cdot(\mathcal{D}\nabla s)$, $\overline{s}^\ell_{jk} = b_{jk} + \overline{\mathcal{H}}^\ell_{jk}$ and the diffusivity's discretization is given later below; second, solve the mass-balance equation

$$\frac{\overline{\overline{\mathcal{H}}}^\ell_{jk} - \overline{\mathcal{H}}^\ell_{jk}}{\Delta t} = \vartheta(\overline{\mathcal{H}}_{jk})\overline{\overline{\mathcal{B}}}^\ell_{jk}, \tag{32}$$

where $\overline{\overline{\mathcal{B}}}^\ell_{jk}$ is the surface mass-balance (5) evaluated with $\overline{\overline{s}}^\ell_{jk} = b_{jk} + \overline{\overline{\mathcal{H}}}^\ell_{jk}$ and ϑ is the Heaviside function

$$\vartheta(x) = \begin{cases} 1, & \text{if } x > 0, \\ 0, & \text{otherwise;} \end{cases} \tag{33}$$

finally, project the thickness such that it is positive:

$$\mathcal{H}^{\ell+1}_{jk} = \left[\overline{\overline{\mathcal{H}}}^\ell_{jk}\right]^+ = \vartheta(\overline{\overline{\mathcal{H}}}^\ell_{jk})\overline{\overline{\mathcal{H}}}^\ell_{jk}. \tag{34}$$

In Eq. 31, $\mathcal{D}^\ell_{j-\frac{1}{2},k-\frac{1}{2}}$ is the discretization of the diffusivity (2)-(3) at position $(j-\frac{1}{2}, k-\frac{1}{2})\Delta x$ and time t^ℓ defined by

$$\begin{aligned}\mathcal{D}^\ell_{j-\frac{1}{2},k-\frac{1}{2}} &= \left(\Gamma\mathcal{H}^\ell_{j-\frac{1}{2},k-\frac{1}{2}} + (\Gamma_s)_{j-\frac{1}{2},k-\frac{1}{2}}\right)\mathcal{K}^{n+1,n-1}_{j-\frac{1}{2},k-\frac{1}{2}}(\ell) \\ \mathcal{D}^\ell_{j+\frac{1}{2},k} &= \frac{1}{2}\left(\mathcal{D}^\ell_{j+\frac{1}{2},k-\frac{1}{2}} + \mathcal{D}^\ell_{j+\frac{1}{2},k+\frac{1}{2}}\right) \\ \mathcal{D}^\ell_{j,k+\frac{1}{2}} &= \frac{1}{2}\left(\mathcal{D}^\ell_{j-\frac{1}{2},k+\frac{1}{2}} + \mathcal{D}^\ell_{j+\frac{1}{2},k+\frac{1}{2}}\right),\end{aligned}$$

$\overline{\mathcal{D}}$ is the diffusivity $\mathcal{D}$ evaluated with $\overline{\mathcal{H}}$ and

$$\begin{aligned}\mathcal{K}^{p,q}_{j-\frac{1}{2},k-\frac{1}{2}}(\ell) &= \left(\mathcal{H}^\ell_{j-\frac{1}{2},k-\frac{1}{2}}\right)^p\left(\alpha^\ell_{j-\frac{1}{2},k-\frac{1}{2}}\right)^q, \\ \mathcal{H}^\ell_{j-\frac{1}{2},k-\frac{1}{2}} &= \frac{\mathcal{H}^\ell_{j-1,k-1} + \mathcal{H}^\ell_{j,k-1} + \mathcal{H}^\ell_{j-1,k} + \mathcal{H}^\ell_{j,k}}{4}, \\ \alpha^\ell_{j-\frac{1}{2},k-\frac{1}{2}} &= \left[\left(\frac{s^\ell_{jk} - s^\ell_{j-1,k} + s^\ell_{j,k-1} - s^\ell_{j-1,k-1}}{2\Delta x}\right)^2\right. \\ &\quad \left. + \left(\frac{s^\ell_{jk} - s^\ell_{j,k-1} + s^\ell_{j-1,k} - s^\ell_{j-1,k-1}}{2\Delta x}\right)^2\right]^{\frac{1}{2}}.\end{aligned}$$

The half-integer indices refer to the mid-points $\left(j-\frac{1}{2}, k-\frac{1}{2}\right)\Delta x$ that are artificially introduced for the centred evaluation of the involved gradients.

Theoretically, the numerical scheme (29)–(34) is likely to suffer from convergence issues because it is solved with a Newton method at each time iteration and, hence, a good initial guess must be provided each time. However, in all the computations performed throughout this paper, no such problem arises. Furthermore, although it is second order in both time and space, the approximation of the glacier's boundaries is first order in space, hence, the scheme is first order in space overall. This can be improved by using either finite-elements or sophisticated, finite-difference techniques that take the boundaries into account more appropriately. A large precision on the SIA is not useful for this article's purposes considering that it is already a rough approximation of a Stokes ice flow. This scheme was validated by [57] on Halfar's glacier [29, 30].

B Quasi-stationary Inverse Method

The notations and definitions are the same as in the previous section. For the sake of simplicity and small computational times, the following finite-difference scheme, centred in space, explicit in time, is advocated. It is deduced from the semi-implicit discretization of Eq. 1 [2, 54, 55] where the surface topography is fixed and the ice thickness is sought: initialize the ice thickness with

$$\mathcal{H}^0_{jk} = \mathcal{H}_0(\xi_j, \zeta_k),$$

for $1 \le j \le N_x$, $1 \le k \le N_y$, where $\mathcal{H}_0$ is an initial guess for the ice thickness at initial time; then, for $\ell \ge 0$, first solve the regularized advection scheme

$$\begin{aligned}\frac{\overline{\mathcal{H}}^\ell_{jk}-\mathcal{H}^\ell_{jk}}{\Delta t} = \frac{1}{\Delta x}\Bigg[&\mathcal{D}^\ell_{j+\frac12,k}\frac{s_{j+1,k}-s_{jk}}{\Delta x} - \mathcal{D}^\ell_{j-\frac12,k}\frac{s_{jk}-s_{j-1,k}}{\Delta x}\\ &+\mathcal{D}^\ell_{j,k+\frac12}\frac{s_{j,k+1}-s_{jk}}{\Delta x} - \mathcal{D}^\ell_{j,k-\frac12}\frac{s_{jk}-s_{j,k-1}}{\Delta x}\Bigg]\\ +\frac{\varepsilon_{\mathrm{QSIM}}}{\Delta x}\Bigg[&\tilde{\mathcal{D}}^\ell_{j+\frac12,k}\frac{\overline{\mathcal{H}}^\ell_{j+1,k}-\overline{\mathcal{H}}^\ell_{jk}}{\Delta x} - \tilde{\mathcal{D}}^\ell_{j-\frac12,k}\frac{\overline{\mathcal{H}}^\ell_{jk}-\overline{\mathcal{H}}^\ell_{j-1,k}}{\Delta x}\\ &+\tilde{\mathcal{D}}^\ell_{j,k+\frac12}\frac{\overline{\mathcal{H}}^\ell_{j,k+1}-\overline{\mathcal{H}}^\ell_{jk}}{\Delta x} - \tilde{\mathcal{D}}^\ell_{j,k-\frac12}\frac{\overline{\mathcal{H}}^\ell_{jk}-\overline{\mathcal{H}}^\ell_{j,k-1}}{\Delta x}\Bigg]\end{aligned} \tag{35}$$

where

$$\begin{aligned}\tilde{\mathcal{D}}^\ell_{j-\frac12,k-\frac12} = &\left(\Gamma\mathcal{H}^\ell_{j-\frac12,k-\frac12} + (\Gamma_s)_{j-\frac12,k-\frac12}\right)\\ &\times\left(\mathcal{H}^\ell_{j-\frac12,k-\frac12}\right)^{n+1}\left(\tilde{\alpha}^\ell_{j-\frac12,k-\frac12}\right)^{n-1},\end{aligned}$$

with

$$\begin{aligned}\tilde{\alpha}^\ell_{j-\frac12,k-\frac12} = \Bigg[&\left(\frac{\mathcal{H}^\ell_{jk}-\mathcal{H}^\ell_{j-1,k}+\mathcal{H}^\ell_{j,k-1}-\mathcal{H}^\ell_{j-1,k-1}}{2\Delta x}\right)^2\\ &+\left(\frac{\mathcal{H}^\ell_{jk}-\mathcal{H}^\ell_{j,k-1}+\mathcal{H}^\ell_{j-1,k}-\mathcal{H}^\ell_{j-1,k-1}}{2\Delta x}\right)^2\Bigg]^{\frac12},\end{aligned}$$

is the discrete counterpart of Eq. 14; second, solve the mass–balance equation

$$\frac{\overline{\overline{\mathcal{H}}}^\ell_{jk}-\overline{\mathcal{H}}^\ell_{jk}}{\Delta t} = \tilde{\mathcal{B}}^\ell_{jk},$$

where $\tilde{\mathcal{B}}^\ell_{jk}$ is the apparent surface mass balance (11) evaluated with s_{jk}; finally, project the thickness such that it is positive:

$$\mathcal{H}^{\ell+1}_{jk} = \left[\overline{\overline{\mathcal{H}}}^\ell_{jk}\right]^+,$$

where the notation $[\cdot]^+$ is defined by Eq. 34. Equations 35, (B) and (B) are solved for every couple of indices (j,k) lying in such grid positions that the ice thickness is known to be larger than zero. Once the forward scheme (29)–(34) has been implemented, Eqs. 35 and (B) are slight modifications of their direct counterparts (31) and (32). Naturally, other numerical schemes can be considered. In this paper, this explicit version is used in order to keep the method as simple as possible. The outcome of this algorithm is a converged ice thickness distribution from which the bedrock topography can obviously be deduced.

C Shape optimization method

The method is presented with continuous equations in the core of this paper, for the sake of simplicity. However, the minimization is performed on the discrete counterparts of all the introduced quantities. The purpose of this section is to formulate the basic ideas underlying the method. The computations are essentially the same in two- as in three-space dimensions. They are much more cumbersome in three dimensions. Hence, the two-dimensional version of scheme (29)–(34) is considered here without sliding. The calculations in the three-dimensional case are reported by [57]. The notations are the same as before. In two space dimensions, the forward scheme is the following: initialize the ice thickness with

$$\mathcal{H}^0_j = s_i(\xi_j) - b_j, \tag{36}$$

for $1 \le j \le N_x$; then, for $0 \le \ell \le M$, first solve the advection-diffusion scheme

$$\begin{aligned}\frac{\overline{\mathcal{H}}^\ell_j-\mathcal{H}^\ell_j}{\Delta t} = \mathcal{T}^\ell_j = \frac{1}{2\Delta x}\Bigg[&\mathcal{D}^\ell_{j+\frac12}\frac{s^\ell_{j+1}-s^\ell_j}{\Delta x} - \mathcal{D}^\ell_{j-\frac12}\frac{s^\ell_j-s^\ell_{j-1}}{\Delta x}\\ &+\overline{\mathcal{D}}^\ell_{j+\frac12}\frac{\overline{s}^\ell_{j+1}-\overline{s}^\ell_j}{\Delta x} - \overline{\mathcal{D}}^\ell_{j-\frac12}\frac{\overline{s}^\ell_j-\overline{s}^\ell_{j-1}}{\Delta x}\Bigg],\end{aligned} \tag{37}$$

where the diffusivity's discretization is

$$\mathcal{D}^\ell_{j-\frac12} = \Gamma\left(\frac{\mathcal{H}^\ell_j+\mathcal{H}^\ell_{j-1}}{2}\right)^{n+2}\left|\frac{s^\ell_j-s^\ell_{j-1}}{\Delta x}\right|^{n-1}, \tag{38}$$

and $\overline{\mathcal{D}}$ is the diffusivity $\mathcal{D}$ evaluated with $\overline{\mathcal{H}}$; second, solve the mass–balance equation

$$\frac{\overline{\overline{\mathcal{H}}}^\ell_j-\overline{\mathcal{H}}^\ell_j}{\Delta t} = \vartheta(\overline{\mathcal{H}}_j)\overline{\overline{\mathcal{B}}}^\ell_j, \tag{39}$$

where $\overline{\overline{\mathcal{B}}}^\ell_j$ is the surface mass balance (5) evaluated with $\overline{\overline{s}}^\ell_j = b_j + \overline{\overline{\mathcal{H}}}^\ell_j$ and ϑ is the Heaviside function (33); finally, project the thickness such that it is positive:

$$\mathcal{H}^{\ell+1}_j = \left[\overline{\overline{\mathcal{H}}}^\ell_j\right]^+ = \vartheta(\overline{\overline{\mathcal{H}}}^\ell_j)\overline{\overline{\mathcal{H}}}^\ell_j. \tag{40}$$

The shape optimization method consists in determining the optimal elevations b_j, $1 \leq j \leq N_x$, of the bedrock topography such that the misfit between the computed and the observed surface topographies at final time t_f is as small as possible. This is an optimal control problem where the ice thickness $\boldsymbol{\mathcal{H}} = \boldsymbol{s} - \boldsymbol{b}$ is termed "state variable" and the bedrock topography $\boldsymbol{b}$ "control variable." For the sake of simplicity, let us consider the unregularized misfit functional

$$\mathcal{J}_h(\boldsymbol{\mathcal{H}}, \boldsymbol{b}) = \frac{1}{2} \left\| \boldsymbol{s}_f(\boldsymbol{\mathcal{H}}, \boldsymbol{b}) - \boldsymbol{s}^{\mathbf{obs}} \right\|^2, \tag{41}$$

where $\| \cdot \|$ stands for the usual Euclidean norm, $\boldsymbol{s}_f = \boldsymbol{b} + \boldsymbol{\mathcal{H}}^{M+1}$ is the computed surface at time $t_f = t^{M+1}$, and $\boldsymbol{s}^{\mathbf{obs}}$ the vector of components $s^{\mathrm{obs}}(\xi_j)$. This represents the misfit between the computed and the observed surface topographies. Now, the numerical scheme (36)–(40) links the control $\boldsymbol{b}$ to the state variable $\boldsymbol{\mathcal{H}}(\boldsymbol{b})$. The cost functional to be minimized can thus be expressed as

$$J_h(\boldsymbol{b}) = \mathcal{J}_h(\boldsymbol{\mathcal{H}}(\boldsymbol{b}), \boldsymbol{b}). \tag{42}$$

The shape optimization problem is cast into: *find*

$$\underset{\boldsymbol{b} \in \mathcal{U}_h^{\mathrm{ad}}}{\operatorname{argmin}} J_h(\boldsymbol{b}) \tag{43}$$

under the constraint that the ice thickness and the bedrock topography satisfy scheme (36)–(40), where $\mathcal{U}_h^{\mathrm{ad}}$ is the discrete counterpart of Eq. 9. In order to minimize the cost functional (42), the solution of the equation

$$\frac{\mathrm{d}J_h}{\mathrm{d}\boldsymbol{b}}(\boldsymbol{b}) = 0 \tag{44}$$

is sought in the space $\mathcal{U}_h^{\mathrm{ad}}$, which is realized by means of a quasi-Newton method while projecting the control variable onto $\mathcal{U}_h^{\mathrm{ad}}$. Let m be an iteration index and $\boldsymbol{b}^m$ the bedrock topography at iteration m. The $m+1$ iterate is found by solving

$$\boldsymbol{b}^{m+1} = \boldsymbol{b}^m + \left(\frac{\mathrm{d}^2 J_h}{\mathrm{d}\boldsymbol{b}^2}(\boldsymbol{b}^m) \right)^{-1} \frac{\mathrm{d}J_h}{\mathrm{d}\boldsymbol{b}}(\boldsymbol{b}^m), \tag{45}$$

where the gradient of J_h is computed exactly with a primal–dual method and the Hessian of J_h is approximated from the value of J_h and its derivative with the BFGS method. The primal–dual method aims at computing the cost functional's gradient and basically consists in finding the stationary points of the Lagrangian functional

$$\begin{aligned}
\mathcal{L}_h(\boldsymbol{\mathcal{H}}, \overline{\boldsymbol{\mathcal{H}}}, \overline{\overline{\boldsymbol{\mathcal{H}}}}, \boldsymbol{\lambda}, \overline{\boldsymbol{\lambda}}, \overline{\overline{\boldsymbol{\lambda}}}, \boldsymbol{b}) &= \mathcal{J}_h(\boldsymbol{\mathcal{H}}, \boldsymbol{b}) + \left[\boldsymbol{\mathcal{H}}^0 - \boldsymbol{s}_i + \boldsymbol{b} \right]^T \boldsymbol{\lambda}^0 \\
&+ \sum_{\ell=0}^{M} \Delta t \left[\frac{\overline{\boldsymbol{\mathcal{H}}}^\ell - \boldsymbol{\mathcal{H}}^\ell}{\Delta t} - \boldsymbol{\mathcal{T}}^\ell \right]^T \overline{\boldsymbol{\lambda}}^\ell \\
&+ \sum_{\ell=0}^{M} \Delta t \left[\frac{\overline{\overline{\boldsymbol{\mathcal{H}}}}^\ell - \overline{\boldsymbol{\mathcal{H}}}^\ell}{\Delta t} - \vartheta(\overline{\boldsymbol{\mathcal{H}}}^\ell) \overline{\overline{\boldsymbol{\mathcal{B}}}}^\ell \right]^T \overline{\overline{\boldsymbol{\lambda}}}^\ell \\
&+ \sum_{\ell=0}^{M} \left[\boldsymbol{\mathcal{H}}^{\ell+1} - \left[\overline{\overline{\boldsymbol{\mathcal{H}}}}^\ell \right]^+ \right]^T \boldsymbol{\lambda}^{\ell+1},
\end{aligned} \tag{46}$$

where $\overline{\overline{\boldsymbol{\mathcal{B}}}}^\ell$ is the surface mass balance evaluated in $\overline{\overline{\boldsymbol{\mathcal{H}}}}^\ell$, $\boldsymbol{\lambda}^\ell$, $\overline{\boldsymbol{\lambda}}^\ell$ and $\overline{\overline{\boldsymbol{\lambda}}}^\ell$ are the costate variables of $\boldsymbol{\mathcal{H}}^\ell$, $\overline{\boldsymbol{\mathcal{H}}}^\ell$ and $\overline{\overline{\boldsymbol{\mathcal{H}}}}^\ell$, that is the Lagrange multipliers of constraints (36), (37), (39) and (40). Note that the inequality constraint $\mathcal{H}_j^\ell \geq 0$ is transformed into an equality constraint and, as a consequence, is included in the Lagrangian. The solutions of the minimization problem are then among the stationary points of $\mathcal{L}_h$ that satisfy optimality condition [6, 19, 26, 78]

$$\nabla \mathcal{L}_h(\boldsymbol{\mathcal{H}}, \overline{\boldsymbol{\mathcal{H}}}, \overline{\overline{\boldsymbol{\mathcal{H}}}}, \boldsymbol{\lambda}, \overline{\boldsymbol{\lambda}}, \overline{\overline{\boldsymbol{\lambda}}}, \boldsymbol{b}) = 0. \tag{47}$$

From this condition, the dual problem of Eqs. 36 – 40 and, therefore, the gradient of J_h, with respect to the bedrock topography, can be deduced.

Stationarity of $\mathcal{L}_h$, with respect to the costate variables $\boldsymbol{\lambda}$, $\overline{\boldsymbol{\lambda}}$ and $\overline{\overline{\boldsymbol{\lambda}}}$ yields constraint Eqs. 36–40. Stationarity of $\mathcal{L}_h$ with respect to the state variables $\boldsymbol{\mathcal{H}}$, $\overline{\boldsymbol{\mathcal{H}}}$ and $\overline{\overline{\boldsymbol{\mathcal{H}}}}$ corresponds to the dual problem, which can be cast into the following formulation: finalize the dual variable with

$$\lambda_j^{M+1} = s^{\mathrm{obs}}(\xi_j) - (s_f)_j. \tag{48}$$

Then, for $\ell = M, M-1, \ldots, 1, 0$, first solve

$$\overline{\overline{\lambda}}_j^\ell = \frac{\vartheta(\overline{\overline{\mathcal{H}}}_j^\ell) \lambda_j^{\ell+1}}{1 - \Delta t \vartheta(\overline{\mathcal{H}}_j^\ell) \frac{\partial \overline{\overline{\mathcal{B}}}_j^\ell}{\partial \mathcal{H}_j}}, \tag{49}$$

where ϑ is the Heaviside function (33); second, solve the discretized linear, backward advection–diffusion equation

$$\begin{aligned}
-\frac{\overline{\overline{\lambda}}_j^\ell - \overline{\lambda}_j^\ell}{\Delta t} &= \frac{n}{2\Delta x} \left(\overline{\mathcal{D}}_{j+\frac{1}{2}}^\ell \frac{\overline{\lambda}_{j+1}^\ell - \overline{\lambda}_j^\ell}{\Delta x} - \overline{\mathcal{D}}_{j-\frac{1}{2}}^\ell \frac{\overline{\lambda}_j^\ell - \overline{\lambda}_{j-1}^\ell}{\Delta x} \right) \\
&- \frac{n+2}{4} \left(\overline{\mathcal{F}}_{j+\frac{1}{2}}^\ell \frac{\overline{s}_{j+1}^\ell - \overline{s}_j^\ell}{\Delta x} \frac{\overline{\lambda}_{j+1}^\ell - \overline{\lambda}_j^\ell}{\Delta x} \right. \\
&\left. + \overline{\mathcal{F}}_{j-\frac{1}{2}}^\ell \frac{\overline{s}_j^\ell - \overline{s}_{j-1}^\ell}{\Delta x} \frac{\overline{\lambda}_j^\ell - \overline{\lambda}_{j-1}^\ell}{\Delta x} \right);
\end{aligned} \tag{50}$$

finally, solve

$$-\frac{\overline{\lambda}_j^\ell - \lambda_j^\ell}{\Delta t} = \frac{n}{2\Delta x}\left(\mathcal{D}^\ell_{j+\frac{1}{2}}\frac{\overline{\lambda}^\ell_{j+1}-\overline{\lambda}^\ell_j}{\Delta x} - \mathcal{D}^\ell_{j-\frac{1}{2}}\frac{\overline{\lambda}^\ell_j - \overline{\lambda}^\ell_{j-1}}{\Delta x}\right)$$
$$-\frac{n+2}{4}$$
$$\left(\mathcal{F}^\ell_{j+\frac{1}{2}}\frac{s^\ell_{j+1}-s^\ell_j}{\Delta x}\frac{\overline{\lambda}^\ell_{j+1}-\overline{\lambda}^\ell_j}{\Delta x} + \mathcal{F}^\ell_{j-\frac{1}{2}}\frac{s^\ell_j - s^\ell_{j-1}}{\Delta x}\frac{\overline{\lambda}^\ell_j - \overline{\lambda}^\ell_{j-1}}{\Delta x}\right), \quad (51)$$

where

$$\mathcal{F}^\ell_{j-\frac{1}{2}} = \Gamma\left(\frac{\mathcal{H}^\ell_j + \mathcal{H}^\ell_{j-1}}{2}\right)^{n+1}\left|\frac{s^\ell_j - s^\ell_{j-1}}{\Delta x}\right|^{n-1}$$

and accordingly for $\overline{\mathcal{F}}^\ell_{j-\frac{1}{2}}$. Equations 50 and 51 constitute the discretization of the continuous, backward diffusion-transport equation

$$-\frac{\partial\lambda}{\partial t} = n\nabla\cdot\left(\mathcal{D}\nabla\lambda\right) - (n+2)\mathcal{F}\frac{\partial s}{\partial x}\frac{\partial\lambda}{\partial x},$$

in $\Omega_\perp \times [t_i, t_f]$, where

$$\mathcal{F} = \Gamma\mathcal{H}^{n+1}\left|\frac{\partial s}{\partial x}\right|^{n-1}.$$

Finally, stationarity of $\mathcal{L}_h$ with respect to control $\boldsymbol{b}$ gives

$$\frac{\partial\mathcal{L}_h}{\partial b_j} = \lambda^0_j + \left((s_f)_j - s^{\mathrm{obs}}(\xi_j)\right) - \Delta t\sum_{\ell=0}^{M}\vartheta(\overline{\mathcal{H}}^\ell_j)\frac{\partial\overline{\overline{\mathcal{B}}}^\ell_j}{\partial b_j}\overline{\overline{\lambda}}^\ell_j$$
$$-\frac{n\Delta t}{2\Delta x}\sum_{\ell=0}^{M}\left[\left(\mathcal{D}^\ell_{j+\frac{1}{2}} + \overline{\mathcal{D}}^\ell_{j+\frac{1}{2}}\right)\frac{\overline{\lambda}^\ell_{j+1}-\overline{\lambda}^\ell_j}{\Delta x}\right.$$
$$\left.-\left(\mathcal{D}^\ell_{j-\frac{1}{2}} + \overline{\mathcal{D}}^\ell_{j-\frac{1}{2}}\right)\frac{\overline{\lambda}^\ell_j - \overline{\lambda}^\ell_{j-1}}{\Delta x}\right],$$

which is the discrete counterpart of the continuous expression

$$\frac{\partial\mathcal{L}}{\partial b}\hat{b} = \int_{\Omega_\perp}\left[\left(s_f - s^{\mathrm{obs}} + \lambda\Big|_{t=t_i}\right)\hat{b}\right]\mathrm{d}\boldsymbol{x}$$
$$-\int_{t_i}^{t_f}\int_{\Omega_\perp}\lambda\frac{\partial\mathcal{B}}{\partial b}\hat{b}\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t - n\int_{t_i}^{t_f}\int_{\Omega_\perp}\nabla\cdot\left(\mathcal{D}\nabla\lambda\right)\hat{b}\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t$$

for any continuous function $\hat{b}$.

When the state and costate variables satisfy the constraint Eqs. 36–40 and the dual problem (48)–(51), the Lagrangian (46) is precisely the cost functional (42), namely

$$\mathcal{L}_h(\boldsymbol{\mathcal{H}}(\boldsymbol{b}), \boldsymbol{\lambda}(\boldsymbol{b}), \boldsymbol{b}) = \mathcal{J}_h(\boldsymbol{\mathcal{H}}(\boldsymbol{b}), \boldsymbol{b}) = J_h(\boldsymbol{b}). \quad (52)$$

The dual problem links the control $\boldsymbol{b}$ to the costate variable $\boldsymbol{\lambda}(\boldsymbol{b})$. Because of Eqs. 36–40 and 48–51, the derivative of Eq. 52 with respect to $\boldsymbol{b}$ is, by the chain rule,

$$\frac{\mathrm{d}J_h}{\mathrm{d}b_j}(\boldsymbol{b}) = \frac{\partial\mathcal{L}_h}{\partial b_j}(\boldsymbol{\mathcal{H}}(\boldsymbol{b}), \boldsymbol{\lambda}(\boldsymbol{b}), \boldsymbol{b}).$$

The computations in the three-dimensional case are performed in the same way. The arising expressions are, however, much more complicated.

References

1. Abe, K., Koro, K.: A topology optimization approach using VOF method. Struct. Multidiscip. Optim. **31**, 470–479 (2006). doi:10.1007/s00158-005-0582-5
2. Adalgeirsdottir, G.: Flow dynamics of Vatnajökull ice cap, Iceland. Mitteilung 181, Versuchsanstalt für Wasserbau, Hydrologie und Glaziologie der ETH Zürich (2003)
3. Allaire, G., Jouve, F., Toader, A.M.: Structural optimization using sensitivity analysis and a level-set method. J. Comput. Phys. **194**, 363–393 (2004). doi:10.1016/j.jcp.2003.09.032
4. Arthern, R.J., Gudmundsson, G.H.: Initialization of ice-sheet forecasts viewed as an inverse Robin problem. J. Glaciol. **56**, 527–533 (2010)
5. Avdonin, S., Kozlov, V., Maxwell, D., Truffer, M.: Iterative methods for solving a nonlinear boundary inverse problem in glaciology. Journal of inverse and ill-posed problems **17**, 239–258 (2009). doi:10.1515/JIIP.2009.018
6. Becker, R., Kapp, H., Rannacher, R.: Adaptive finite element methods for optimal control of partial differential equations: basic concept. SIAM J. Control. Optim. **39**(1), 113–132 (2000). doi:10.1137/S0363012999351097
7. Broyden, C.G.: The convergence of a class of double-rank minimization algorithms 1. general considerations. J. Appl. Math. **6**(1), 76–90 (1970)
8. Budd, W.F., Warner, R.C., Jacka, T.H., Li, J., Treverrow, A.: Ice flow relations for stress and strain-rate components from combined shear and compression laboratory experiments. J. Glaciol. **59**(214), 374–392 (2013). doi:10.3189/2013JoG12J106
9. Bueler, E.: Numerical approximation of a two–dimensional thermomechanical model for ice flow. Department of mathematical sciences, technical report 02-02, University of Alaska, Fairbanks (2002)
10. Bueler, E., Lingle, C.S., Kallen-Brown, J.A., Covey, D.N., Bowman, L.N.: Exact solutions and verification of numerical models for isothermal ice sheets. J. Glaciol. **51**(173), 291–306 (2005). doi:10.3189/172756505781829449
11. Calvetti, D., Morigi, S., Reichel, L., Sgallari, F.: Tikhonov regularization and the L-curve for large discrete ill-posed problems. J. Comput. Appl. Math. **123**(12), 423–446 (2000). doi:10.1016/S0377-0427(00)00414-3. Numerical Analysis 2000. Vol. III: Linear Algebra
12. Christensen, E.L., Reeh, N., Forsberg, R., Jorgensen, J.H., Skou, N., Woelders, K.: A low-cost glacier-mapping system. J. Glaciol. **46**(154), 531–537 (2000). doi:10.3189/172756500781833142
13. Clarke, G.K.C., Anslow, F.S., Jarosch, A.H., Radic, V., Menounos, B., Bolch, T., Berthier, E.: Ice volume and subglacial topography for western Canadian glaciers from mass balance fields thinning rates, and a bed stress model. J. Clim. **26**, 4282–4303 (2013). doi:10.1175/JCLI-D-12-00513.1
14. Deng, Y., Zhang, P., Liu, Y., Wu, Y., Liu, Z.: Optimization of unsteady incompressible Navier Stokes flows using variational level set method. Int. J. Numer. Methods Fluids **71**(12), 1475–1493 (2013). doi:10.1002/fld.3721
15. Farinotti, D., Corr, H., Gudmundsson, G.H.: The ice thickness distribution of Flask Glacier, Antarctic Peninsula, determined by combining radio-echo soundings, surface velocity data, and flow modelling. Ann. Glaciol. **54**(63), 18–24 (2012). doi:10.3189/2013AoG63A603

16. Farinotti, D., Huss, M., Bauder, A., Funk, M.: An estimate of the glacier ice volume in the Swiss Alps. Glob. Planet. Chang. **68**(3), 225–231 (2009)
17. Farinotti, D., Huss, M., Bauder, A., Funk, M., Truffer, M.: A method to estimate the ice volume and ice-thickness distribution of alpine glaciers. J. Glaciol. **55**(191), 422–430 (2009)
18. Fletcher, R.: A new approach to variable metric algorithms. The Computer Journal **13**(3), 317–322 (1970)
19. Gallaire, F.: Adjoint methods for flow control. European Research Community on Flow. Turbulence and Combustion Bulletin **73**, 27–32 (2007)
20. Glen, J.W.: The flow law of ice. IUGG/IAHS Symposium of Chamonix IAHS Publication **47**, 171–183 (1958)
21. Goldberg, D.N., Sergienko, O.V.: Data assimilation using a hybrid ice flow model. Cryosphere **5**(2), 315–327 (2011). doi:10.5194/tc-5-315-2011
22. Goldfarb, D.: A family of variable-metric methods derived by variational means. Math. Comput. **24**, 23–26 (1970)
23. Greve, R., Blatter, H.: Dynamics of ice sheets and glaciers, 1st edn. Springer (2009)
24. Gudmundsson, G.H.: A three-dimensional numerical model of the confluence area of Unteraargletscher, Bernese Alps, Switzerland. J. Glaciol. **45**(150), 219–230 (1999). doi:10.3189/002214399793377086
25. Gudmundsson, G.H., Thorsteinsson, T., Raymond, C.: Inferring bed topography and stickiness from surface data on ice streams. American Geophysical Union Fall Meeting Abstracts. Abstract number IP21A-0678 (2001)
26. Guégan, A., Schmid, P.J., Huerre, P.: Optimal energy growth and optimal control in swept Hiemenz flow. J. Fluid Mech. **566**, 11–45 (2006). doi:10.1017/S0022112006001303
27. Habermann, M., Maxwell, D., Truffer, M.: Reconstruction of basal properties in ice sheets using iterative inverse methods. J. Glaciol. **58**(210), 795–807 (2012). doi:10.3189/2012JoG11J168
28. Haeberli, W., Hoelzle, M.: Application of inventory data for estimating characteristics of and regional climate-change effects on mountain glaciers: a pilot study with the European Alps. Ann. Glaciol. **21**, 206–212 (1995)
29. Halfar, P.: On the dynamics of the ice sheets. Journal of Geophysical Research: Oceans **86**, 11,065–11,072 (1981). doi:10.1029/JC086iC11p11065
30. Halfar, P.: On the dynamics of the ice sheets 2. Journal of Geophysical Research: Oceans **88**, 6043–6051 (1983). doi:10.1029/JC088iC10p06043
31. Heimbach, P., Bugnion, V.: Greenland ice-sheet volume sensitivity to basal, surface and initial conditions derived from an adjoint model. Ann. Glaciol. **50**(52), 67–80 (2009). doi:10.3189/172756409789624256
32. Heining, C.: Velocity field reconstruction in gravity-driven flow over unknown topography. Phys. Fluids **23**(3), 032,101 (2011). doi:10.1063/1.3559144
33. Hinze, M., Pinnau, R., Ulbrich, M., Ulbrich, S.: Optimization with PDE Constraints. Springer (2009)
34. Hofer, T.: Numerical simulation and optimization of the alumina distribution in an aluminium electrolysis pot. doi:10.5075/epfl-thesis-5023. EPFL PhD thesis no. 5023 (2011)
35. Huss, M., Bauder, A., Funk, M., Hock, R.: Determination of the seasonal mass balance of four Alpine glaciers since 1865. J. Geophys. Res. **113**, F01,015 (2008). doi:10.1029/2007JF000803
36. Huss, M., Farinotti, D.: Distributed ice thickness and volume of all glaciers around the globe. J. Geophys. Res-Earth. (2012). doi:10.1029/2012JF002523
37. Huss, M., Hock, R., Bauder, A., Funk, M.: 100-year mass changes in the swiss alps linked to the atlantic multidecadal oscillation. Geophys. Res. Lett. **37** (2010). doi:10.1029/2010GL042616
38. Hutchinson, M.F.: A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. J. Hydrol. **106**, 211–232 (1989)
39. Hutter, K.: Theoretical Glaciology. Reidel (1983)
40. Jouvet, G.: Modélisation, analyse mathématique et simulation numérique de la dynamique des glaciers. doi: 10.5075/epfl-thesis-4677. EPFL PhD thesis no. 4677 (2010)
41. Jouvet, G., Huss, M., Blatter, H., Picasso, M., Rappaz, J.: Numerical simulation of Rhonegletscher from 1874 to 2100. J. Comp. Phys. **228**, 6426–6439 (2009)
42. Jouvet, G., Huss, M., Funk, M., Blatter, H.: Modelling the retreat of Grosser Aletschgletscher, Switzerland, in a changing climate. J. Glaciol. **57**(206), 1033–1045 (2011). doi:10.3189/002214311798843359
43. Kennedy, J., Eberhart, R.: Particle swarm optimization. IEEE International Conference on Neural Networks **4**, 1942–1948 (1995). doi:10.1109/ICNN.1995.488968
44. Khazendar, A., Rignot, E., Larour, E.: Larsen B Ice Shelf rheology preceding its disintegration inferred by a control method. Geophys. Res. Lett. **34** (2007). doi:10.1029/2007GL030980
45. Kirner, P.: Modélisation mathématique et simulation numérique des phénomènes dynamiques et thermiques apparaissant dans un glacier. doi:10.5075/epfl-thesis-3840. EPFL PhD thesis no. 3840 (2007)
46. Kuzmin, D.: A guide to numerical methods for transport equations. Free CFD Books. University of Nürnberg (2010)
47. Larour, E., Rignot, E., Joughin, I., Aubry, D.: Rheology of the Ronne Ice Shelf, Antarctica, inferred from satellite radar interferometry data using an inverse control method. Geophys. Res. Lett. **32**(5) (2005). doi:10.1029/2004GL021693
48. Li, H.L., Ng, F., Li, Z.Q., Qin, D.H., Cheng, G.D.: An extended "perfect-plasticity" method for estimating ice thickness along the flow line of mountain glaciers. Geophys. Res. Lett. **117**, F01020 (2012). doi:10.1029/2011JF002104
49. Mahaffy, M.W.: 3-dimensional numerical model of ice sheets - tests on barnes ice cap. J. Geophys. Res. - Oceans Atmos. **81**(6), 1059–1066 (1976). doi:10.1029/JC081i006p01059
50. Martin, N.: Modélisations directes et inverses d'écoulements géophysiques par méthodes variationnelles - application à la glaciologie. Ph.D. thesis, Université de Toulouse (2013)
51. Maxwell, D., Truffer, M., Avdonin, S., Stuefer, M.: An iterative scheme for determining glacier velocities and stresses. J. Glaciol. **54**, 888–898 (2008)
52. McNabb, R.W., Hock, R., O'Neel, S., Rasmussen, L.A., Ahn, Y., Braun, M., Conway, H.B., Herreid, S., Joughin, I.R., Pfeffer, W.T., Smith, B.E., Truffer, M.: Using surface velocities to calculate ice thickness and bed topography: a case study at Columbia Glacier, Alaska, USA. J. Glaciol. **58**(212), 1151–1164 (2012). doi:10.3189/2012JoG11J249
53. Mercer, R.E., Sampson, J.R.: Adaptive search using a reproductive meta-plan. Kybernetes **7**(3), 215–228 (1978). doi:10.1108/eb005486
54. Michel, L.: Estimating the ice thickness of mountain glaciers from surface topography and mass-balance data. doi: 10.5075/epfl-thesis-5940. EPFL PhD thesis no. 5940 (2013)
55. Michel, L., Picasso, M., Farinotti, D., Bauder, A., Funk, M., Blatter, H.: Estimating the ice thickness of mountain glaciers with an inverse approach using surface topography and mass-balance. Inverse Problems **29**(3), 035,002 (2013). doi:10.1088/0266-5611/29/3/035002
56. Michel, L., Picasso, M., Farinotti, D., Bauder, A., Funk, M., Blatter, H.: Estimating the ice thickness of mountain glaciers with a shape optimization algorithm using surface topography and mass-balance. J. Inverse Ill-Posed Probl. (2014). doi:10.1515/jip-2013-0016. (in press)

57. Michel, L., Picasso, M., Farinotti, D., Funk, M., Blatter, H.: Estimating the ice thickness of shallow glaciers from surface topography and mass-balance data with a shape optimization algorithm. Comput. Geosci. **66**, 182–199 (2014)
58. Morland, L.W.: Thermomechanical balances of ice sheet flows. Geophysical & Astrophysical Fluid Dynamics **29**(1-4), 237–266 (1984). doi:10.1080/03091928408248191
59. Morland, L.W., Johnson, I.R.: Effects of bed inclination and topography on steady isothermal ice sheets. J. Glaciol. **28**(98), 71–90 (1982)
60. Morlighem, M., Rignot, E., Seroussi, H., Larour, E., Dhia, H.B., Aubry, D.: A mass conservation approach for mapping glacier ice thickness. Geophys. Res. Lett. **38**, L19,503 (2011). doi:10.1029/2011GL048659
61. Munson, T., Sarich, J., Wild, S., Benson, S., McInnes, L.C.: Tao 2.0 users manual. Tech. rep.rep., Mathematics and Computer Science Division, Argonne National Laboratory (2012). Http://www.mcs.anl.gov/tao
62. Nocedal, J., Wright, S.: Numerical optimization, 2nd edn. Springer Series in Operations Research and Financial Engineering. Springer (2006)
63. Nye, J.F.: The mechanics of glacier flow. J. Glaciol. **2**(12), 82–93 (1952)
64. Paul, F., Linsbauer, A.: Modeling of glacier bed topography from glacier outlines, central branch lines, and a DEM. Int. J. Geogr. Inf. Sci. **26**(7), 1173–1190 (2012). doi:10.1080/13658816.2011.627859
65. van Pelt, W.J.J., Oerlemans, J., Reijmer, C.H., Pettersson, R., Pohjola, V.A., Isaksson, E., Divine, D.: An iterative inverse method to estimate basal topography and initialize ice flow models. Cryosphere **7**(3), 987–1006 (2013). doi:10.5194/tc-7-987-2013
66. Petra, N., Zhu, H., Stadler, G., Hughes, T.J.R., Ghattas, O.: An inexact Gauss Newton method for inversion of basal sliding and rheology parameters in a nonlinear Stokes ice sheet model. J. Glaciol. **58**(211), 889–903 (2012). doi:10.3189/2012JoG11J182
67. Raymond, M.J., Gudmundsson, G.H.: On the relationship between surface and basal properties on glaciers, ice sheets, and ice streams. J. Geophys. Res. - Solid Earth **110**, B08,411 (2005). doi:10.1029/2005JB003681
68. Raymond, M.J., Gudmundsson, G.H.: Estimating basal properties of glaciers from surface measurements: a non-linear Bayesian inversion approach. Cryosphere Discuss. **3**(1), 181–222 (2009). doi:10.5194/tcd-3-181-2009
69. Raymond-Pralong, M., Gudmundsson, G.H.: Bayesian estimation of basal conditions on Rutford Ice Stream, West Antarctica, from surface data. J. Glaciol. **57**(202), 315–324 (2011)
70. Roth, R., Ulbrich, S.: A discrete adjoint approach for the optimization of unsteady turbulent flows. Flow, Turbulence and Combustion **90**(4), 763–783 (2013). doi:10.1007/s10494-012-9439-3
71. Schafer, M., Meur, E.L.: Improvement of a 2-D SIA ice-flow model: application to Glacier de Saint-Sorlin, France. J. Glaciol. **53**(183), 713–722 (2007). doi:10.3189/002214307784409234
72. Shanno, D.F.: Conditioning of quasi-Newton methods for function minimization. Math. Comput. **24**, 647–656 (1970)
73. Thorsteinsson, T., Raymond, C.F., Gudmundsson, G.H., Bindschadler, R.A., Vornberger, P., Joughin, I.: Bed topography and lubrication inferred from surface measurements on fast-flowing ice streams. J. Glaciol. **49**(167), 481–490 (2003). doi:10.3189/172756503781830502
74. Tikhonov, A.: Solution of incorrectly formulated problems and the regularization method. Soviet Math. Doklady **4**, 1035–1038 (1963)
75. Vieli, A., Payne, A.J.: Application of control methods for modelling the flow of Pine Island Glacier, West Antarctica. Ann. Glaciol. **36**(1), 197–204 (2003). doi:10.3189/172756403781816338
76. Weertman, J.: On the sliding of glaciers. J. Glaciol. **3**(21), 33–38 (1957)
77. Yamada, T., Izui, K., Nishiwaki, S., Takezawa, A.: A topology optimization method based on the level set method incorporating a fictitious interface energy. Comput. Methods Appl. Mech. Eng. **199**(4548), 2876–2891 (2010). doi:10.1016/j.cma.2010.05.013
78. Yamaleev, N., Diskin, B., Nielsen, E.: Adjoint-based methodology for time-dependent optimization. 12th The American Institute of Aeronautics and Astronautics/International Society for Structural and Multidisciplinary Optimization (AIAA/ISSMO) Multidisciplinary Analysis and Optimization Conference **MAO-10** (2008)
79. Zhou, S., Li, Q.: A variational level set method for the topology optimization of steady-state NavierStokes flow. J. Comput. Phys. **227**(24), 10,178–10,195 (2008). doi:10.1016/j.jcp.2008.08.022. http://www.sciencedirect.com/science/article/pii/S002199910800449X
80. Zwinger, T., Greve, R., Gagliardini, O., Shiraiwa, T., Lyly, M.: A full Stokes-flow thermo-mechanical model for firn and ice applied to the Gorshkov crater glacier, Kamchatka. Ann. Glaciol. **45**, 29–37 (2007)