Supplement of The Cryosphere, 11, 949–970, 2017
http://www.the-cryosphere.net/11/949/2017/
doi:10.5194/tc-11-949-2017-supplement

The Cryosphere Open Access EGU

Supplement of

How accurate are estimates of glacier ice thickness? Results from ITMIX, the Ice Thickness Models Intercomparison eXperiment

Daniel Farinotti et al.

Correspondence to: Daniel Farinotti (daniel.farinotti@ethz.ch)

S 1 Supplementary Model Descriptions

NOTE: The model descriptions are presented in alphabetical order. The following abreviations are used: SMB = Surface Mass Balance, DEM= Digital Elevation Model, ∂h/∂t = rate of ice thikckness change.

S 1.1 "Brinkerhoff" – Brinkerhoff et al. (2016)

The method of Brinkerhoff et al. (2016) poses the problem of finding bedrock elevations in the context of Bayesian inference. The model uses the prior hypotheses that bed elevations and the ice flux divergence (SMB-∂h/∂t) can be modelled as Gaussian random fields with assumed covariance but unknown mean. The choice of covariance function enforces strong prior information about smoothness. Depth averaged velocities are found by solving the continuity equation. The likelihood assumes that velocities and the ice flux divergence are normally distributed with known covariance around the supplied data. With prior and likelihood in hand, the model uses the Metropolis-Hastings algorithm (Hastings, 1970) to generate samples from the posterior distribution of bed elevations.

For ITMIX, only the maximum likelihood solution is reported (not the full posterior distribution). Because of the Gaussian distribution of observations and priors, the choice of observation and prior covariance does not affect the maximum likelihood solution. The model was applied only to the synthetic cases, as velocity fields provided for the real cases had either insufficient spatial coverage or non-physical behaviour incompatible with the assumed model physics.

S 1.2 "Brinkerhoff-v2" – Brinkerhoff (unpublished)

The general idea behind "Brinkerhoff-v2" is to find a bedrock topography that – when forcing a given ice flow model with a prescribed SMB field – gives rise to the observed surface geometry Mathematically, the following inverse problem is solved: Find the bedrock topography $B(x, y)$ that minimizes the functional

$$I = \int_{\bar{\Omega}} \frac{(\hat{S} - \hat{S}_{\text{obs}})^2}{2} + \gamma \nabla B \cdot \nabla B \, \mathrm{d}\bar{\Omega} + \lambda \int_{\Gamma} (S_{\text{obs}} - B) \, \mathrm{d}\Gamma, \tag{1}$$

where $\bar{\Omega}$ and Γ are the glacier extent and boundary, respectively, $\hat{S}$ and $\hat{S}_{\text{obs}}$ are the modelled and observed surface elevations smoothed over a length scale l, γ is a smoothness parameter, and λ is a Lagrange multiplier. The first term in I quantifies the misfit between modelled and observed surface elevations. The second term is a Tikhonov regularization term which imposes smoothness on the computed bed elevation. The third term – which is not applied at ice divides with non-zero thickness – imposes the bed elevation to be equal to the surface elevation at the observed ice boundaries.

$S = B + H$ is the steady state solution to the continuity equation

$$\nabla \cdot \bar{\mathbf{u}} H = \dot{b} - \left. \frac{\partial S}{\partial t} \right|_{\text{obs}}, \tag{2}$$

subject to the inequality constraint $S \geq B$. The latter ensures that ice thickness H is positive. We treat $\frac{\partial S}{\partial t}|_{\text{obs}}$ as an observed quantity rather than as a true time derivative, leading to the concept of "effective mass balance" $\dot{b}_{\text{eff}} = \dot{b} - \frac{\partial S}{\partial t}|_{\text{obs}}$. In the continuity equation, $\bar{\mathbf{u}}$ is the depth averaged velocity, and $\dot{b}$ the climatic mass balance. We compute the velocity field using the Blatter-Pattyn approximation to the Stokes' equations (Pattyn, 2003)

$$\frac{\partial}{\partial x}\left[4\eta \frac{\partial u}{\partial x} + 2\eta \frac{\partial v}{\partial y}\right] + \frac{\partial}{\partial y}\left[\eta \frac{\partial u}{\partial y} + \eta \frac{\partial v}{\partial x}\right] + \frac{\partial}{\partial z}\eta \frac{\partial u}{\partial z} = \rho g \frac{\partial S}{\partial x} \tag{3}$$

$$\frac{\partial}{\partial x}\left[\eta \frac{\partial u}{\partial y} + \eta \frac{\partial v}{\partial x}\right] + \frac{\partial}{\partial y}\left[2\eta \frac{\partial u}{\partial x} + 4\eta \frac{\partial v}{\partial y}\right] + \frac{\partial}{\partial z}\eta \frac{\partial v}{\partial z} = \rho g \frac{\partial S}{\partial y} \tag{4}$$

subject to a stress free upper surface and a basal shear stress given by

$$\tau_{b,i} = -\beta^2 \mathbf{u}_i, \tag{5}$$

where β^2 is a traction parameter.
Because we solve these equations over the observed glacier extent (and do not allow for boundary migration), we must enforce the condition that velocity be tangential to the boundaries

$$\mathbf{u} \cdot \mathbf{n} = 0, \tag{6}$$

where $\mathbf{n}$ is the surface normal. We allow tangential velocity to be free. Such a condition also amounts to a zero-flux boundary for the mass conservation equation, which may usually only be specified at one point on a particle trajectory. Here we apply it over the entire domain boundary, which imposes an additional constraint on the effective mass balance field. Integrating the continuity equation over the entire domain and integrating by parts we find that

$$\int_\Gamma \bar{\mathbf{u}} H \cdot \mathbf{n} \, \mathrm{d}\Gamma = \int_{\bar{\Omega}} \dot{b}_{\text{eff}} \, \mathrm{d}\bar{\Omega}. \tag{7}$$

The no-penetration boundary condition makes the left side zero, implying that the effective mass balance must also integrate to zero for the continuity equation to have a valid solution. We enforce this constraint by applying an additive constant to the observations of effective mass balance. In the absence of additional considerations, this also implies that the solution is permanently sensitive to initial conditions, in that the mass of the system remains constant. However, since thickness is constrained to be positive, this constraint effectively acts as an internal source term. As such we initialize our solution procedure from a zero ice thickness state, yielding as a steady-state solution the lowest-volume ice thickness distribution consistent with continuity and the Blatter-Pattyn equations.
We discretize the continuity and Blatter-Pattyn equations using linear finite elements (Zienkiewicz and Taylor, 2000). The continuity equation is hyperbolic, so we stabilize it using a streamline upwind Petrov-Galerkin method (Brooks and Hughes, 1982). We simultaneously solve both equations with the variational inequality solver SNES VI (Balay et al., 2016) that uses Newton's method for dealing with the non-linear viscosity and coupling terms and an active set method for dealing with the bounds. We use pseudo-transient continuation to find the steady state solution to the coupled system (Kelley and Keyes, 1998).
With a mean of computing S, we use a heuristic fixed point method to find the approximate minimum of I. At each iteration we compute the steady surface elevation, and then use Gaussian smoothing over the length scale l to find $\hat{S}$. We set l to be approximately one ice thickness, though this is necessarily a rough estimate since we do not know the thickness *a priori*. We then use the difference between this and the observed surface elevations smoothed over the same length scale to compute a new bedrock elevation according to:

$$\frac{B - B_0}{\Delta\tau} = -\frac{\hat{S} - \hat{S}_{\text{obs}}}{\Delta\tau} + \gamma \nabla \cdot \nabla B, \tag{8}$$

where $\Delta\tau$ is a fictitious time scale of relaxation, and the solution is subject to the Dirichlet boundary condition $B = S$ on all boundaries where ice is known to have zero thickness. We alternate between using this equation to update bedrock elevations and producing new steady state surface elevations. We progressively reduce the regularization parameter γ until the average surface elevation error falls below an arbitrary threshold, which we set to 2m for this paper.
As an optional final step, if surface velocities are available, we invert for basal traction using what have become standard techniques in ice sheet modelling known as adjoint methods (e.g. MacAyeal, 1993; Brinkerhoff and Johnson, 2013). We perform this procedure as an outer iteration relative that described above. Experimentation has shown that only a few iterations are usually required for the basal traction field to converge. In the absence of velocity observations,

β^2 is tuned such that approximately half of the resulting surface velocity field is due to sliding. A uniform flow rate factor of $A = 3.17 \times 10^{-24}\,\mathrm{Pa}^{-3}\,\mathrm{s}^{-1}$ was used. The model was applied to all cases that were topologically contiguous and for which SMB was available.

S 1.3 "Farinotti" (ITEM) – Farinotti et al. (2009b)

The method of Farinotti et al. (2009b) (also referred to as "ITEM", e.g. Farinotti et al., 2009a; Gabbi et al., 2012), is based on mass conservation and principles of ice flow dynamics. Basically, the approach estimates the ice volume flux across profiles located along manually prescribed ice flow lines, and converts it into ice thickness by using Glen's flow law (Glen, 1955). For any point along a given flowline, the ice volume flux is approximated by integrating the ice flux divergence (difference between SMB and $\partial \mathrm{h}/\partial \mathrm{t}$, sometimes referred to as "apparent mass balance") upstream of that point, whereas only the area within manually prescribed "ice flow catchments" is considered. The ice flux divergence is assumed to have a linear dependence on elevation, and two separate gradients are prescribed for the glacier ablation and accumulation zone. The ice thickness obtained along the individual flow lines is then interpolated across the glacier, and the local surface slope is used to modulate the resulting local ice thickness.
With the exception of the correction factor C (cf. Equation 7 in Farinotti et al., 2009b), the same model parameters as in Farinotti et al. (2009a) (see their Tab. 2) were used for ITMIX. C was set to 0.65, i.e. a value about 15 % higher than in the original publications. This is because Gabbi et al. (2012) suggested a systematic overestimation in the derived ice thickness distributions, and because a higher value of C translates into a lower ice thickness. Although the approach was originally designed for mountain glaciers only, it was applied to all 21 test cases, including ice caps.

S 1.4 "Fuerst" – Fürst et al., unpublished

The approach by Fürst et al. (unpublished at the time ITMIX was conducted; now presented in The Cryosphere Discussions (Fürst et al., 2017)) is a minimization approach based on mass conservation. It maps glacier ice thickness H following an idea similar to the one by Morlighem et al. (2011) (cf. Sec. S 1.13). Required input comprises distributed fields of (a) surface mass balance $\dot{\mathrm{b}}$, (b) rate of ice thickness change $\partial \mathrm{H}/\partial \mathrm{t}$, and (c) surface ice flow velocity $\mathbf{u} = (u_1, u_2)$. To solve for mass conservation, the Elmer/Ice model (Gillet-Chaulet et al., 2012; Gagliardini et al., 2013) is used. Discretization is performed with the stabilised streamline upwind Petrov-Galerkin (SUPG) scheme (Brooks and Hughes, 1982). Along the domain margin, the ice thickness is set to zero as a Dirichlet boundary condition.
Inconsistencies in the input fields transmit to the thickness solution H, which might therefore show strong variations and negative values. To reduce such deficiencies, an optimisation is pursued which iteratively updates the control variables $\dot{\mathrm{b}}$ and $\mathbf{u}$. For this purpose, the following cost function J is minimized:

$$
\begin{aligned}
J = \quad & \lambda_{\mathrm{pos}} \cdot \int_\Omega H^2 \int_{-\infty}^{H} \delta(s) ds \, d\Omega + \\
+ \quad & \lambda_{\mathrm{regH}} \cdot \int_\Omega \left(\frac{\partial H}{\partial x}\right)^2 + \left(\frac{\partial H}{\partial y}\right)^2 d\Omega + \\
+ \quad & \lambda_{\dot{\mathrm{b}}} \cdot \int_\Omega \left(\dot{\mathrm{b}} - \dot{\mathrm{b}}^{\mathrm{data}}\right)^2 d\Omega + \\
+ \quad & \lambda_{\mathrm{U}} \cdot \sum_{\mathrm{i}=1}^{2} \int_\Omega \left(u_{\mathrm{i}} - u_{\mathrm{i}}^{\mathrm{data}}\right)^2 d\Omega + \\
+ \quad & \lambda_{\mathrm{regU}} \cdot \sum_{\mathrm{i}=1}^{2} \int_\Omega \left(\frac{\partial u_{\mathrm{i}}}{\partial x}\right)^2 + \left(\frac{\partial u_{\mathrm{i}}}{\partial y}\right)^2 d\Omega
\end{aligned}
\tag{9}
$$

Here, Ω is the ice-covered domain and λ_i ($i = [\text{pos, regH, } \dot{b}\text{, U, regU}]$) are weighting parameters. Superscripts "data" denote the initially provided input fields (measurements). In this way, J penalises negative ice thicknesses (first term in Eq. 9), strong variations in the ice thickness (second term), the mismatch between updated and initial fields of the control variables (third and fourth term), as well as strong spatial variations in the velocity components (last term). The multi-variate optimisation requires derivatives of J with respect to all control variables. These are calculated using the adjoint system. The iterative optimisation preferentially modifies $\dot{b}$ because magnitudes differ between the control variables. To align relative change values, a scaling factor of 10^{-1} was introduced for the velocity derivatives. For the optimisation of J, we apply the "M1QN3" module (Gilbert and Lemaréchal, 1989) that can solve large-scale unconstrained minimisation problems.
For the synthetic cases in ITMIX, good performance was achieved using $\lambda_{\text{pos}} = 1$, $\lambda_{\text{regH}} = 3{\cdot}10^{-5}$, $\lambda_{\dot{b}} = 8 \cdot 10^{-4}$, $\lambda_{\text{u}} = 8 \cdot 10^{-1}$ and $\lambda_{\text{regU}} = 8 \cdot 10^{-2}$. Two real-world tests cases with sufficient input data (Unteraar and Austfonna) were considered as well. On Unteraar Glacier, no velocity information was available in the accumulation area, which required a reduction of λ_{u} in this area by a factor $5 \cdot 10^2$. On Austfonna, the provided velocity vectors produced unrealistic streamlines for an ice cap. Therefore, flow directions were prescribed from surface slopes. Additionally, surface elevation changes were manually set to zero as up-slope ice flow would have resulted otherwise.

S 1.5 "Gantayat" – Gantayat et al. (2014)

The approach by Gantayat et al. (2014) is based on the shallow ice approximation (e.g. Cuffey and Paterson, 2010) and Glen's flow law (Glen, 1955). It solves the equation

$$u_{\text{s}} = u_{\text{b}} + \frac{2A}{n+1} \left(f \rho g H \sin \alpha\right)^n \tag{10}$$

for ice thickness H, where u_{s} and u_{b} are surface and basal velocities respectively, A is the flow rate factor, $n = 3$ the creep exponent, f a shape factor, $\rho = 900\,\text{kg}\,\text{m}^{-3}$ the ice density, $g = 9.81\,\text{m}\,\text{s}^{-2}$ the gravitational acceleration, and α the surface slope. The equation is solved within 100 m elevation bands and the result is smoothed with a kernel of 3×3 grid cells to obtain the final ice thickness distribution.
For ITMIX, $A = 3.2 \times 10^{-24}\,\text{Pa}^{-3}\,\text{s}^{-1}$ and $f = 0.8$ where chosen for all test cases; u_{s} was obtained from the provided surface velocity fields; and $u_{\text{b}} = 0.25\,u_{\text{s}}$ was assumed (Gantayat et al., 2014). Because of the data requirements, only 7 test cases were considered.

S 1.6 "Gantayat-v2" – modified Gantayat et al. (2014)

"Gantayat-v2" is a modified version of the approach by Gantayat et al. (2014). Instead of solving Equation 10 for elevation bands, the same equation is first solved along discrete points of manually digitized branchlines (e.g. Linsbauer et al., 2012), and the resulting ice thickness is spatially interpolated by assuming zero ice thickness at the glacier margin. For the interpolation, the ANUDEM-algorithm (Hutchinson, 1989) is used.
For ITMIX, individual branchlines were generated requiring (a) a lateral spacing between adjacent lines of about ∼200 m, (b) a minimal distance of ∼100 m from the glacier margin, and (c) that branchlines from individual glacier tributaries gradually merge with the branchlines of the main stream. The same 7 test cases as for "Gantayat" were considered, and the same parameters used.

S 1.7 "GCbedstress" – Clarke et al. (2013)

The bed-stress method of Clarke et al. (2013) shares many conceptual features with Farinotti et al. (2009b) but differs in its implementation. Glacier flowsheds are hand-delineated and

then transversely dissected by ladder-like "rungs" that represent flux gates oriented roughly perpendicular to the local ice flow direction. Ice discharge through these flux gates is calculated by integrating the apparent balance (SMB$-\partial$h/∂t) over the upstream area associated with each gate. Discharge values associated with each rung are then applied to intervening cells by interpolation in a process that is equivalent to inserting rungs. The average ice flux per unit width of channel is found by dividing ice discharge by the channel width for each cell. This width is taken to be the sum of the distance from the nearest channel boundaries to the left and right of the downflow direction. "Raw" ice thickness estimates are obtained from the ice flux using Glen's law (flow rate factor $A = 2.4 \times 10^{-24}\,\mathrm{Pa}^{-3}\,\mathrm{s}^{-1}$), together with the estimated surface slope for each cell and the inclined slab flow assumption. The raw estimates are then smoothed by minimizing a cost function that negotiates a tradeoff between accepting the raw estimates or maximizing the smoothness of the solution. Zero ice thickness is used as a boundary condition at ice-free margins.

When SMB fields were not provided within ITMIX, these were constructed assuming a linear variation with elevation above and below the estimated equilibrium line altitude. When ∂h/∂t fields were lacking these were assumed to vanish, or take a constant value, or vary linearly with elevation, depending on available information. No sliding was assumed, which leads to systematically higher ice thickness estimates than in the case of sliding. The tradeoff between raw estimates and smoothness was found to depend on the grid spacing in a manner unforeseen by Clarke et al. (2013) so that the assigned parameter χ (see their Eq. 6) differs among the considered test cases.

S 1.8 "GCneuralnet" – Clarke et al. (2009)

The artificial neural net (ANN) method of Clarke et al. (2009) is based on the assumption that presently glacierized areas denuded of their ice cover would resemble nearby ice-free landscapes. This assumption relies on the geomorphic premise that *"landscape signatures of glaciation are the expression of regional influences such as geology, climate, and the intensity of past glaciations"* (Clarke et al., 2009). The ice thickness is estimated considering the minimum range distance from an on-glacier site to enclosing valley walls in eight compass directions (45° aperture) using an azimuthal stencil that has two elevation layers. The maximum search range for the stencil is limited by the map dimensions and other considerations. To train the ANN the stencil is centred on an ice-free cell and range distances to valley walls are measured for each sector and layer of the stencil as the thickness of ice cover is increased.

The surface DEMs provided within ITMIX are unsuitable for direct application of the method. This is because the DEMs either lack elevation data beyond the glacier margins, or because the tight framing may cause the stencil to probe the frame edge. The provided map domains where therefore artificially expanded by reflecting elevation and ice mask data at the frame boundaries. This expedient falsifies the topography and ice cover beyond the frame boundaries. It can only be justified because it allows stencil calculations to proceed and the ANN method to be included in ITMIX. Note, moreover, that the geomorphic premise makes the ice caps (Academy, Austfonna and Devon) and crater mountain glaciers (Elbrus and Mocho) considered within ITMIX unsuitable candidates for the ANN method. This is because little about the nature of their subglacial topography can be inferred from the geometric character of the surrounding ice-free terrain.

S 1.9 "Huss" (HF-model) – Huss and Farinotti (2012)

The method of Huss and Farinotti (2012) further develops the approach by Farinotti et al. (2009b). It avoids the digitization of glacier flowlines, includes additional physics (e.g. basal sliding, longitudinal variations in the valley shape factor, influence of ice temperature and climatic regime), and is applicable at the global scale. Glacier hypsometry and surface character-

istics (mean slope and width) are evaluated for 10 m elevation bands, and all calculations are performed using this simplified 2D shape. Apparent mass balance gradients for the ablation and accumulation area (see Farinotti et al., 2009b) are estimated based on the continentality of the glacier, which is derived from local equilibrium line altitudes. Ice volume fluxes along the glacier are converted into ice thickness using an integrated form of Glen's flow law. The variations in the valley shape factor and the basal shear stress in the longitudinal glacier profile are taken into account. Simple parametrisations describe both the temperature-dependence of the flow rate factor and the variability in basal sliding. Calculated mean elevation-band thickness is extrapolated to each cell of a regular grid considering local surface slope, and the distance from the glacier margin. For marine-terminating glaciers, a fixed ice volume flux is prescribed at the glacier terminus.

For ITMIX, all model parameters were set to the values used in Huss and Farinotti (2012). For the synthetic cases, geographical position and continentality defining the spatial variation of the model parameters, were not provided within ITMIX. These variables were thus estimated from the supplied surface mass balance distribution.

S 1.10 "Linsbauer" (GlabTop) – Linsbauer et al. (2009, 2012)

The *Glacier bed Topography (GlabTop)* method (labelled "Linsbauer" throughout the manuscript) by Linsbauer et al. (2009, 2012) uses an empirical relation between average basal shear stress τ and glacier elevation range (Haeberli and Hoelzle, 1995) for calculating the ice thickness at individual points along manually digitized glacier branchlines. From τ and the zonal surface slope α (computed within 50 m elevation bins along the branchlines) the ice thickness h is calculated as $h = \tau/(f\rho g \sin\alpha)$, where f is a shape factor, $\rho = 900\,\mathrm{kg\,m^{-3}}$ the ice density, and $g = 9.81\,\mathrm{m\,s^{-2}}$ gravitational acceleration. The dependence on α implies thin ice where the glacier surface is steep and thick ice where it is flat. A distributed ice thickness is obtained by interpolating the estimated point-values within the glacier outlines.

For ITMIX, branchlines covering all glacier branches and tributaries were digitized manually. A maximal value of $\tau = 150\,\mathrm{kPa}$ was assumed and $f = 0.8$ was set. The remaining parameters, as well as the interpolation algorithms used for modelling, are the same as in Linsbauer et al. (2012). Note that GlabTop was designed for alpine glaciers and has not been applied for ice caps so far. Within ITMIX, however, it was applied to all 21 test cases, including ice caps.

S 1.11 "Machguth" (GlabTop2) – Frey et al. (2014)

GlabTop2, labelled with "Machguth" and fully described in Frey et al. (2014), is based on the same concept as the "Linsbauer" model (see previous section). In particular, local ice thickness is calculated from an estimate of the basal shear stress and the surface slope. The laborious process of manually drawing branchlines, however, is rendered obsolete by computing the surface slope from the average slope of all grid cells within a predefined elevation buffer. The method is entirely grid-based and first calculates the ice thickness at a set of randomly selected grid cells. In a second step, this thickness is interpolated to the entire glacier area. To achieve realistic glacier cross-sections, the interpolation scheme assigns a minimum, non-zero ice thickness to all grid cells directly adjacent to the glacier margin.

For ITMIX, the identical settings as in Frey et al. (2014) were used. A maximal ice thickness of 1000 m, however, was introduced to avoid excessive glacier thickness in very flat areas of ice caps. To avoid influence of small-scale surface undulations on modelled ice thickness, all provided DEMs were down-sampled to 75 m cell size for the calculations, and then re-interpolated to the original resolution. The method is fully automated and fast, but includes a non-physical, tunable parameter controlling the random point sampling. The model has been shown to be well-suited for mountain glaciers (Frey et al., 2014) but was not applied to ice caps so far. All test cases besides Austfonna were considered.

S 1.12 "Maussion" (OGGM) – Maussion et al. (unpublished)

The *Open Global Glacier Model (OGGM)* by Maussion et al. (unpublished) was primarily developed for the dynamical modelling of mountain glaciers and the estimation of their total volume. Estimation of a distributed subglacial topography is implemented through a procedure that extends and fully automatizes the method of Farinotti et al. (2009b). Automatisation is achieved by generating multiple flowlines according to Kienholz et al. (2014) (Fig. S1a).

At the junction between a tributary and the main glacier stem, the tributary's catchment area is computed using a flow-routing approach. The geometrical width of each cross section along the flowline can then be computed from the intersection of a normal vector with either (a) the glacier outline or (b) the outline of the tributary's catchment. While these geometrical widths are visually accurate (cf. Fig. S1b), their cumulative area often does not match the total area of the glacier. The area-altitude distribution of the so-generated "flowline glacier" should, however, match the actual one, and the geometrical widths are thus corrected with an altitudinal, multiplicative factor. This ensures the total area of the glacier to be reproduced (Fig. S1b and S1c).

After these geometrical operations, the mass balance for each point on the flowline is derived from monthly temperature and precipitation data (extracted from the CRU dataset; Harris et al., 2014) and the temperature index model by Marzeion et al. (2012). In a nutshell, the calibration procedure searches for a 31-years climate period in the past for which the glacier would have been in equilibrium while keeping its modern-time geometry. This "equilibrium mass-balance" is then assumed to be equal to the "apparent mass-balance" as defined by Farinotti et al. (2009b). The major implication for the inversion in comparison to Farinotti et al. (2009b) is that the OGGM mass-balance profiles are not parametrized but actually depend on the local climate. This dependence is mostly related to the amounts of solid precipitation.

The local ice volume flux q is then calculated by integrating the resulting mass-balance field upstream of each grid point. Mass conservation is given by

$$q = uS \tag{11}$$

where the ice velocity u and the glacier cross-section area S are given by:

$$u = \frac{2A}{n+2} h (\rho g h \alpha)^n \qquad \text{and} \qquad S = \frac{2}{3} w h, \tag{12}$$

where A is the flow rate factor, n Glen's flow law exponent, ρ the ice density, g the gravitational acceleration, α the surface slope, and w the cross-section width. Combination of the above equations allows for the ice thickness h to be computed. Note that in this step, OGGM assumes no sliding and a parabolic shape of the glacier bedrock. The flowline's thickness is finally interpolated to the actual glacier geometry. OGGM uses a 2D cubic interpolation weighted with the inverse sinus of the local slope. When doing so, the previously computed total volume is conserved. This procedure is rather coarse and might be further developed in the future.

Glen's flow rate factor A is a free calibration parameter in OGGM. For ITMIX, A was calibrated with the "observed" average glacier thicknesses reported in GlaThiDa v1 (Gärtner-Roer et al., 2014). During calibration, GlaThiDa entries referring to any of the ITMIX test cases where omitted, and the resulting value of A applied to all test cases. We found an optimised value for A of 7.73×10^{-24} $\mathrm{Pa^{-3}\,s^{-1}}$, which is about three times larger that the typical value suggested by Cuffey and Paterson (2010). This is consistent with the neglection of basal sliding, which has to be compensated by less stiff ice. The result, however, should not be over-interpreted, as the calibration also compensates for a series of additional uncertainties in the inversion.

For ITMIX, the ice caps were handled as the sum of individual glacier basins as provided by the Randolph Glacier Inventory v5.0 (Arendt et al., 2015). Further information on OGGM, including its code, are available at `www.oggm.org`. More information about the ITMIX-specific calibration

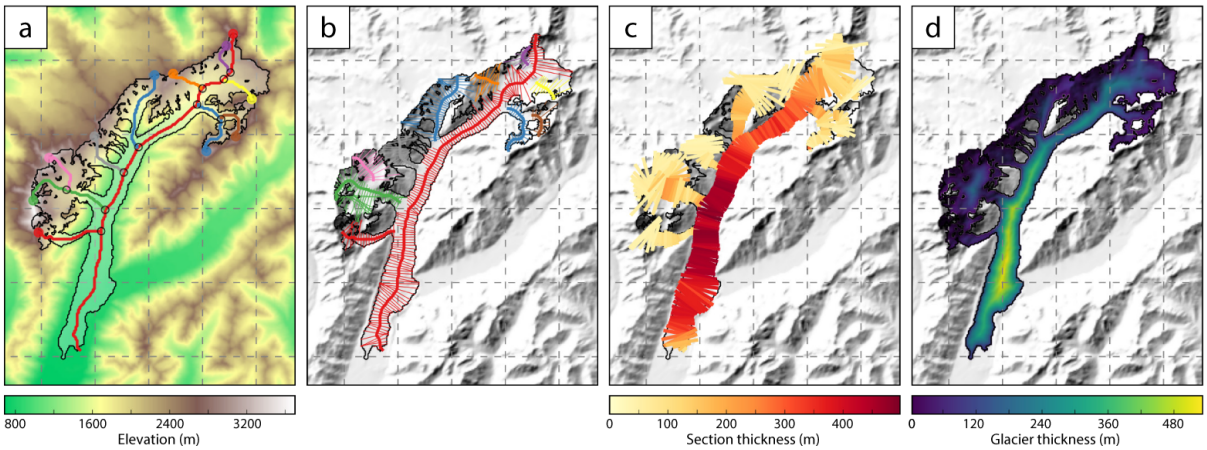

Figure S1: Example of the OGGM inversion workflow applied to the Tasman Glacier, New Zealand: (a) definition of a local map and computation of the flowlines; (b) computation of the glacier geometrical widths; (c) inversion of the glacier thickness on the width-corrected flowlines; (d) spatial interpolation of the ice thickness. Figure from `https://github.com/OGGM/oggm`, Copyright: OGGM authors.

procedure can be found at `http://fabienmaussion.info/2016/06/18/itmix-experiment-phase1/`. ITMIX results were computed by using OGGM version v0.1.1 (Maussion et al., 2017).

S 1.13 "Morlighem" – Morlighem et al. (2011)

The method of Morlighem et al. (2011) was originally designed to fill gaps between ground-penetrating radar measurements over the Greenland and Antarctic ice sheets. Based on mass conservation, it computes the ice thickness by requiring the ice flux divergence to be balanced by the rate of thickness change and the net surface and basal mass balances. The strength of the method lies in the capability of including direct ice thickness measurements (e.g. Morlighem et al., 2014), which is achieved by optimizing the ice flux divergence and depth-averaged velocities to minimize the misfit between observed and modelled thickness.

For ITMIX, the optimization sequence was not used, as no ice thickness measurements were provided. The algorithm thus relied solely on mass conservation. For the test cases including surface velocities, these were assumed to be equal to the depth averaged velocities. In the other cases, the shallow ice approximation was used together with an assumption of no-sliding to convert the computed ice mass flux into ice thickness (flow rate factor $A = 9.3 \times 10^{-25}\,\mathrm{Pa}^{-3}\,\mathrm{s}^{-1}$). The method was only applied to test cases providing SMB.

S 1.14 "RAAJgantayat" – Re-implemented from Gantayat et al. (2014)

"RAAJgantayat" is a set of solutions derived with an independent re-implementation of the approach by Gantayat et al. (2014) (see Sec. S 1.5 for a description). This particular version was applied to four valley glaciers (Tasman, Unteraar, Brewster, North Glacier) and one synthetic case (Synthetic1).

For Brewster and North Glacier, the required surface flow velocity fields were obtained by interpolating the provided point velocities with a standard inverse-distance weighting technique. For the other cases, the available velocity field were used directly. To avoid unrealistically large ice thicknesses, distributed surface slope values derived from the provided DEMs were filtered so that values below 2° were eliminated. For all test cases, the shape factor (flow rate factor) was set to $f = 0.75$ ($A = 3.2 \times 10^{-24}\,\mathrm{Pa}^{-3}\,\mathrm{s}^{-1}$). All other parameters (including ice density, creep exponent, gravitational acceleration) were set to the same values as in Gantayat et al. (2014).

S 1.15 "RAAJglabtop2" – Re-implemented from Linsbauer et al. (2009)

Similarly as above, "RAAJglabtop2" is a set of solutions stemming from an independent re-implementation of an existing model, i.e. the model GlabTop2 by Frey et al. (2014) (see Sec. S 1.11). Results were generated for three real-world geometries (Tasman, North Glacier, Unteraar) and one synthetic test case (Synthetic2).

For the simulations, a shape factor of $f = 0.74$ was used. The average basal shear stress τ, derived from the glacier's elevation range, was set to 130, 150, 150, and 97 kPa for the test cases North Glacier, Tasman, Unteraar, and Synthetic2, respectively. All other parameters were set to the values given in Frey et al. (2014).

S 1.16 "Rabatel" – Rabatel et al. (unpublished)

The "Rabatel" method uses glacier surface flow velocities and SMB to quantify the ice flux for individual cross sections perpendicular to the central flow line of the glacier. The ice thickness is then quantified along these cross sections and a spatial distribution obtained through interpolation between various cross sections. The approach can be divided into four main steps:

1. Definition of the cross sections: The central flow line is delineated from the uppermost elevation of the glacier to the glacier snout, perpendicular to the contour lines. Cross sections are defined perpendicular to the central flow line, and chosen to cover the glacier homogeneously. In the ablation zone, the limits of the cross sections are the glacier edges. Where tributaries are present, or in the accumulation zone, the cross sections are delimited using a threshold on the surface flow velocity. The threshold is a parameter that is either fixed across the glacier, or adjusted for individual cross sections. When thickness measurements are available, the threshold can be adjusted to accommodate them.

2. Quantification of ice flux and average thickness for individual cross sections: As in Farinotti et al. (2009b), the gridded glacier surface topography and an elevation gradient $\partial\widetilde{b}/\partial z$ are used to calculate the "apparent mass-balance" $\widetilde{b}_i$ for each grid cell i. The ice flux Q_j of each cross-section j is then computed from the sum of $\widetilde{b}_i$ of all n grid cells that contribute to the flux at that location, i.e.

$$Q_j = \sum_n \widetilde{b}_i. \tag{13}$$

On the other hand, Q_j can be computed from the average velocity of the cross section $\overline{v}_j$, and its surface area S_j:

$$Q_j = \overline{v}_j S_j. \tag{14}$$

According to Cuffey and Paterson (2010), the depth-averaged ice flow velocity typically corresponds to 80 % of the flow velocity at the glacier surface. Combining equations (13) and (14), and knowing the distribution of $\widetilde{b}$, the surface area of each cross section can be computed from the average surface flow velocity $\overline{v}_{\mathrm{surf}}$ of the cross section:

$$S_j = \frac{\sum_n \widetilde{b}_i}{0.8\ \overline{v}_{\mathrm{surf}}}. \tag{15}$$

Since the length l_j of the cross section is known, the average ice thickness for that cross section can be computed as $\overline{h}_j = S_j/l_j$.

3. Ice thickness distribution along the cross section: The average ice thickness is "distributed" along the cross section by using the surface flow velocity pattern. For that, it is assumed that the ice thickness h_i for each grid cell along the cross section follows the distribution of the surface flow velocity, i.e.

$$h_i = v_i \overline{h}_j / \overline{v}_j. \tag{16}$$

4. Extrapolation at the glacier scale: The final step consists in extrapolating the thickness calculated for individual cross sections at the glacier scale. For ITMIX, in which the method was applied to the test case "Synthetic1" only, this was done by using a Kriging method.

S 1.17 "VanPeltLeclercq" – Adapted from van Pelt et al. (2013)

The "VanPeltLeclercq" approach iteratively derives a distributed glacier bed topography by minimizing the mismatch between modelled and observed glacier surface elevations (e.g. Leclercq et al., 2012; Michel et al., 2013; van Pelt et al., 2013). Following van Pelt et al. (2013), repeated time-dependent model runs are performed with a model for ice dynamics. In contrast to van Pelt et al. (2013), who used the Parallel Ice Sheet Model (PISM, Bueler and Brown, 2009) to reconstruct basal topography in both synthetic and real cases, the dynamical model "SIADYN" is used. SIADYN is based on the vertically integrated shallow ice approximation (e.g. Hutter, 1983), including Weertman sliding (Huybrechts, 1991), and is part of the ICEDYN package (Sect. 3.3 in Reerink et al., 2010). Model runs are stopped at the time the provided surface DEMs refer to. After every run, the misfit between modelled and observed surface elevation is computed. The bed topography is then adjusted by a fraction of this misfit, thus resulting in a new bed topography for a next model iteration. In test-cases providing velocity data, the iterative procedure is stopped when a minimum velocity misfit is achieved. In absence of such data, the procedure is terminated when the average surface elevation misfit drops below 5 m.
Only test cases providing SMB information were considered, as the forward model requires SMB input. Since ITMIX does not provide SMB time series but just one SMB field, the model was run with constant forcing until equilibrium. This may result in a mismatch between modelled and observed glacier extent at the time the surface topography was measured. To avoid this, SMB fields of the individual test cases were adjusted with a constant offset. This SMB offset was optimized for each glacier individually. Higher accuracy in the derived bedrock elevations can potentially be obtained when time-dependent SMB data are available. When velocity observations were available, these were additionally used to tune the model parameters affecting basal sliding and deformational flow.

S 2 Supplementary Figures and Tables

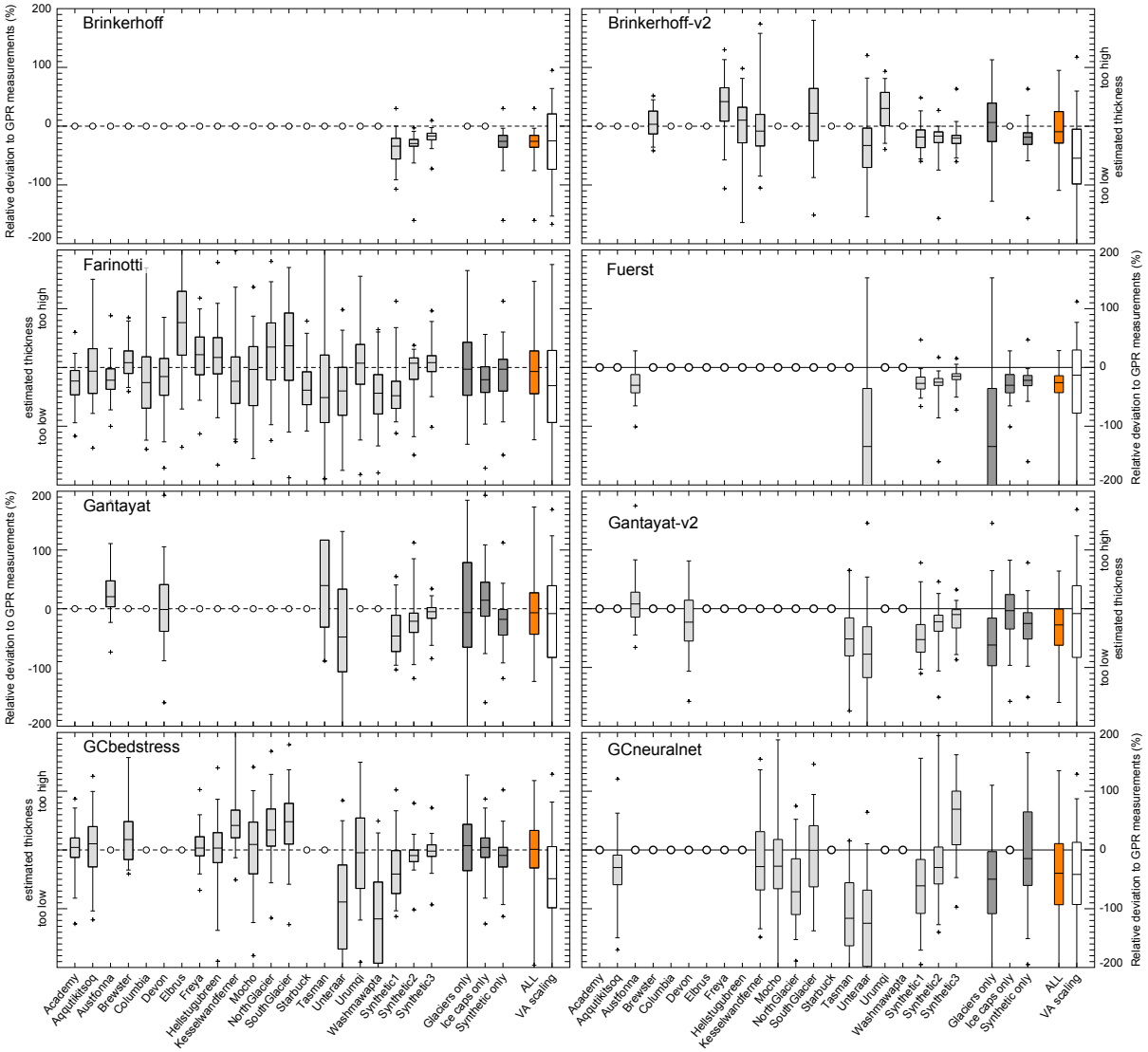

Figure S2: Differences between estimated and measured ice thickness for all test cases. Models are shown in alphabetical order (models not displayed in the figure are found in Fig. S2 on the next page). Boxplots show minimum and maximum values (crosses), the 95 % confidence interval (whiskers), the interquartile range (box) and the median (lines within box). [Continued on next page]

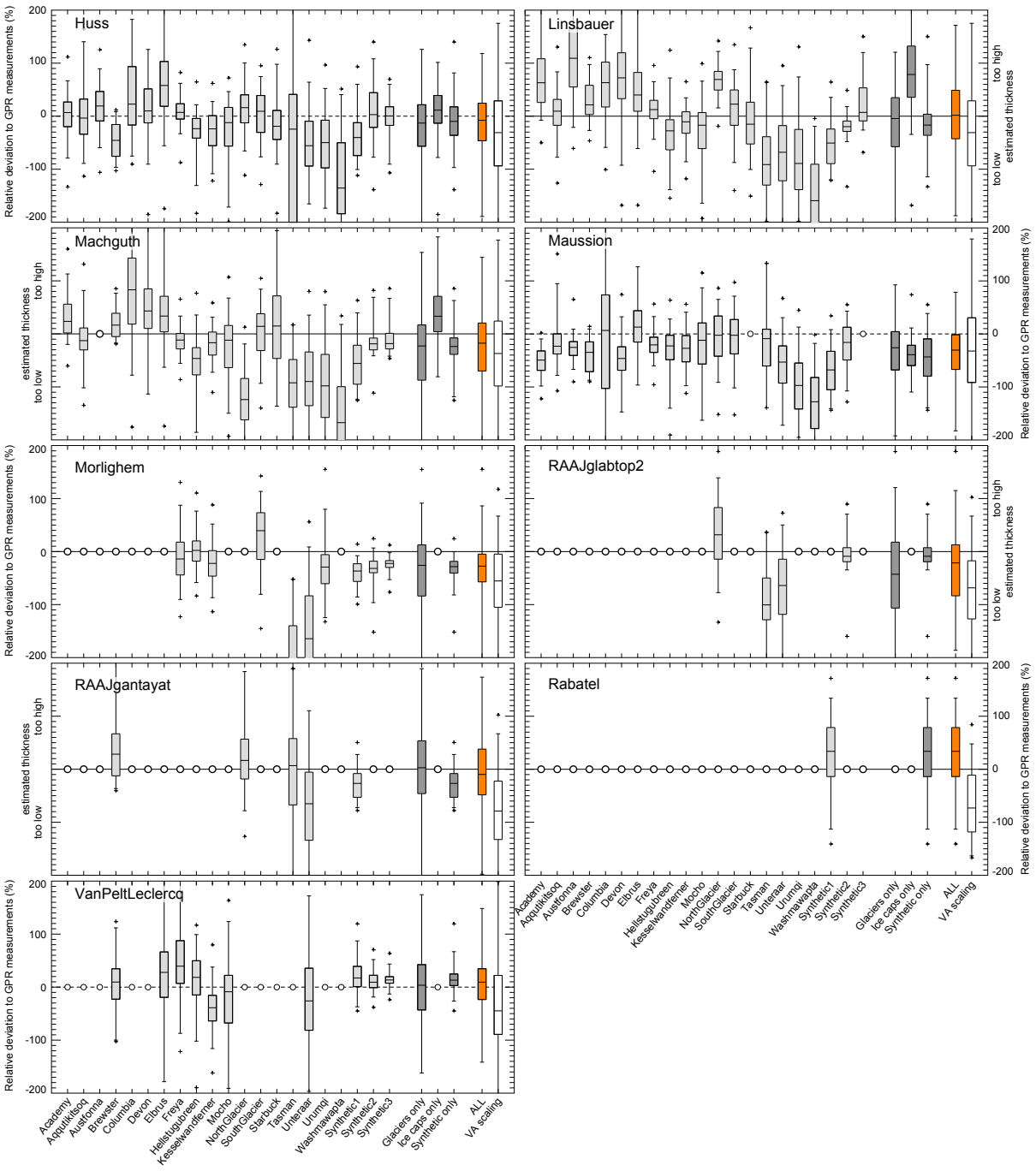

Figure S2: [Continued from previous page]

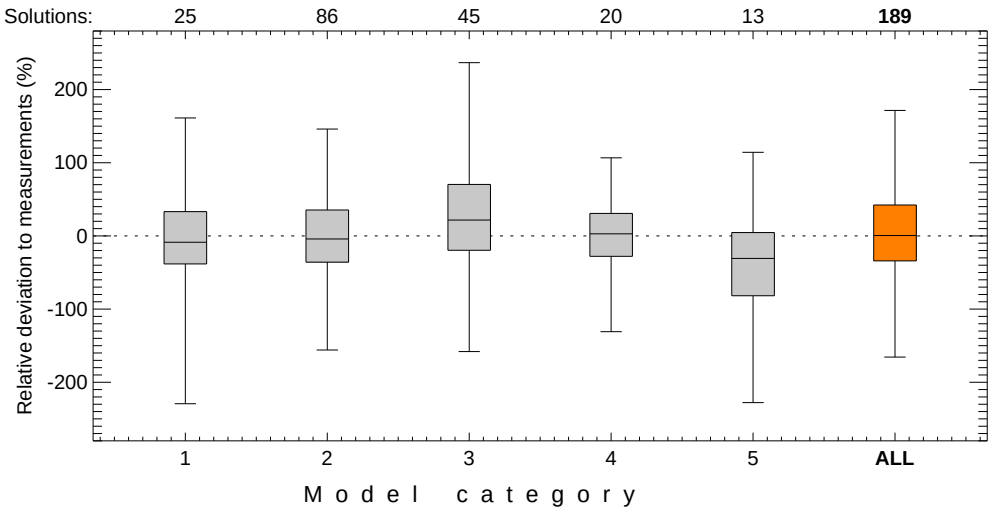

Figure S3: Differences between estimated and measured ice thickness, pooled for the individual model categories defined in Section 4 of the main text. The total number of solutions provided by models of a given category is given. Categories include (1) minimization approaches, (2) mass conserving approaches, (3) shear-stress based approaches, (4) velocity-based approaches, and (5) other approaches. Boxplots show, the 95 % confidence interval (whiskers), the interquartile range (box) and the median (lines within box).

Table S1: Years the input data used during the experiment are referring to. OL: Glacier outline, DEM: digital elevation model of the glacier surface, SMB: surface mass balance, Vel.: ice flow velocity at the surface, ∂h/∂t: rate of ice thickness change, H: ice thickness measurements. References for data sources are given in Table 1 of the main article.

Test case	OL	DEM	SMB	Vel.	∂h/∂t	H
Academy	1997	1997	-	-	-	1997
Aqqutikitsoq	2014	2014	-	-	-	2014
Austfonna	2008	2007	2004-2013	1995-1996	2003-2009	1983
Brewster	1997	1997	2004-2008	2004-2008	-	1997
Columbia	2007	2007	-	-	-	2010
Devon	1999	2000	-	2007-2008	-	2000
Elbrus	1997	1997	1984-2010	-	1957-1997	2005-2007,2013,2014
Freya	2013	2013	2008-2014	-	-	2013
Hellstugubreen	2003	2009	1997-2010	2011-2013	1997-2009	2011
Kesselwandferner	1997	1997	1988-1997	-	-	1995
Mocho	2000	2013	2006-2014	-	-	2013
North Glacier	2007	2007	-	2006-2012	-	2008,2009,2011
South Glacier	2007	2007	2007-2012	2006-2014	-	2008,2009,2011
Starbuck	2003	2003	-	-	-	2013
Tasman	1986	1986	2000-2010	2000-2011	-	1971-1973
Unteraar	2003	2003	1997-2003	1997	1997-2003	1997,1998,2000
Urumqi	2012	2012	2011-2014	-	-	2012
Washmawapta	2007	2007	-	-	-	2006

Table S2: Overview of the values used for the rate factor A in Glen's flow law (Glen, 1955). "constant/variable" indicates whether the particular model applies a constant rate factor to all test cases, or whether A is varied on a case-by-case basis. "tuned/prescribed" indicates whether the rate factor is used as a tuning parameter or is prescribed *a priori*. Models that do not include a rate factor are labelled with "-".

Model	A ($\mathrm{Pa}^{-3}\,\mathrm{s}^{-1}$)	constant/variable	tuned/prescribed
Brinkerhoff	3.17×10^{-24}	constant	prescribed
Brinkerhoff-v2	3.17×10^{-24}	constant	prescribed
Farinotti	2.40×10^{-24}	constant	prescribed
Fuerst	-	-	-
Gantayat	3.20×10^{-24}	constant	prescribed
Gantayat-v2	3.20×10^{-24}	constant	prescribed
GCbedstress	2.40×10^{-24}	constant	prescribed
GCneuralnet	-	-	-
Huss	$1.07 - 2.38 \times 10^{-24}$	variable[(1)]	prescribed
Linsbauer	-	-	-
Machguth	-	-	-
Maussion	7.73×10^{-24}	constant	tuned[(2)]
Morlighem	0.93×10^{-24}	constant	prescribed
RAAJgantayat	3.20×10^{-24}	constant	prescribed
RAAJglabtop2	-	-	-
Rabatel	-	-	-
VanPeltLeclercq	$0.48 - 1.32 \times 10^{-24}$	variable	prescribed

[(1)] A is a function of ice temperature. For details, refer to Huss and Farinotti (2012).
[(2)] Tuned to measurements within GlaThiData (WGMS, 2016). See Sec. S 1.12 for details.

Table S3: Case-by-case model performance. For each test case and every participating model, the average (avg), median (med), interquartile range (IQR), and 95 % confidence interval (95 %) of the deviations from ice thickness measurements are given (indicators). Values are percental deviations from the mean ice thickness. The numbers in parenthesis are the rank of every indicator (first three ranks are highlighted). The average rank of the four indicators is given in the last column (AVG).

Test case / Model	avg	med	IQR	95%	AVG
Academy					
1 GCbedstress	3 (**2**)	4 (**1**)	±17 (**1**)	±77 (5)	2.2
2 Farinotti	−27 (**3**)	−23 (**3**)	±21 (**3**)	±59 (**2**)	2.8
3 Huss	2 (**1**)	7 (**2**)	±23 (4)	±73 (4)	2.8
4 Maussion	−51 (5)	−49 (5)	±18 (**2**)	±44 (**1**)	3.2
5 Machguth	32 (4)	24 (4)	±27 (5)	±66 (**3**)	4.0
6 Linsbauer	77 (6)	63 (6)	±41 (6)	±128 (6)	6.0
Aqqutikitsoq					
1 Linsbauer	7 (4)	9 (**3**)	±24 (**3**)	±81 (**1**)	2.8
2 Huss	2 (**2**)	−4 (**1**)	±33 (5)	±115 (7)	3.8
3 Machguth	−11 (5)	−13 (5)	±21 (**2**)	±92 (**3**)	3.8
4 Maussion	−16 (6)	−23 (6)	±19 (**1**)	±86 (**2**)	3.8
5 Farinotti	1 (**1**)	−6 (**2**)	±38 (7)	±114 (6)	4.0
6 GCbedstress	6 (**3**)	11 (4)	±34 (6)	±102 (4)	4.2
7 GCneuralnet	−34 (7)	−30 (7)	±25 (4)	±106 (5)	5.8
Austfonna					
1 Gantayat-v2	9 (**1**)	8 (**1**)	±21 (4)	±64 (4)	2.5
2 Maussion	−27 (5)	−26 (5)	±13 (**1**)	±38 (**1**)	3.0
3 Farinotti	−20 (**3**)	−22 (4)	±17 (**3**)	±52 (**3**)	3.2
4 Fuerst	−26 (4)	−30 (6)	±16 (**2**)	±47 (**2**)	3.5
5 Huss	18 (**2**)	19 (**2**)	±28 (6)	±74 (6)	4.0
6 Gantayat	28 (6)	21 (**3**)	±22 (5)	±67 (5)	4.8
7 Linsbauer	118 (7)	109 (7)	±55 (7)	±166 (7)	7.0
Brewster					
1 Brinkerhoff-v2	5 (**2**)	3 (**1**)	±19 (**2**)	±40 (**1**)	1.5
2 Farinotti	13 (**3**)	8 (**2**)	±19 (**1**)	±57 (5)	2.8
3 Machguth	20 (4)	17 (4)	±22 (**3**)	±47 (**2**)	3.2
4 VanPeltLeclercq	4 (**1**)	10 (**3**)	±29 (6)	±106 (8)	4.5
5 Linsbauer	29 (6)	21 (6)	±28 (4)	±62 (6)	5.5
6 Maussion	−40 (7)	−35 (8)	±28 (5)	±48 (**3**)	5.8
7 GCbedstress	27 (5)	18 (5)	±32 (8)	±96 (7)	6.2
8 Huss	−46 (9)	−46 (9)	±30 (7)	±53 (4)	7.2
9 RAAJgantayat	41 (8)	28 (7)	±40 (9)	±121 (9)	8.2
Columbia					
1 Farinotti	−14 (**1**)	−26 (**3**)	±44 (**2**)	±147 (**3**)	2.2
2 Huss	38 (**3**)	23 (**2**)	±55 (**3**)	±129 (**2**)	2.5
3 Linsbauer	58 (4)	63 (4)	±42 (**1**)	±106 (**1**)	2.5
4 Maussion	−18 (**2**)	6 (**1**)	±88 (5)	±238 (5)	3.2
5 Machguth	87 (5)	83 (5)	±62 (4)	±184 (4)	4.5
Devon					
1 Farinotti	−15 (**2**)	−16 (**3**)	±31 (**2**)	±106 (4)	2.8
2 Gantayat	3 (**1**)	−1 (**1**)	±40 (6)	±97 (**3**)	2.8
3 Huss	17 (**3**)	10 (**2**)	±32 (**3**)	±108 (5)	3.2
4 Gantayat-v2	−20 (4)	−23 (4)	±35 (4)	±94 (**2**)	3.5
5 Maussion	−49 (6)	−46 (6)	±22 (**1**)	±90 (**1**)	3.5
6 Machguth	47 (5)	43 (5)	±37 (5)	±166 (6)	5.2
7 Linsbauer	78 (7)	72 (7)	±43 (7)	±177 (7)	7.0

Continued on next page.

Table S3: Continued from previous page.

Test case / Model		avg		med		IQR		95%		AVG
Elbrus										
1	Maussion	15	(**1**)	13	(**1**)	±30	(**1**)	±112	(**1**)	1.0
2	Machguth	49	(**3**)	34	(**3**)	±34	(**2**)	±167	(4)	3.0
3	Linsbauer	54	(4)	40	(4)	±37	(**3**)	±166	(**3**)	3.5
4	VanPeltLeclercq	23	(**2**)	28	(**2**)	±43	(5)	±194	(6)	3.8
5	Huss	66	(5)	58	(5)	±42	(4)	±142	(**2**)	4.0
6	Farinotti	79	(6)	76	(6)	±54	(6)	±171	(5)	5.8
Freya										
1	GCbedstress	6	(**1**)	3	(**1**)	±16	(4)	±50	(4)	2.5
2	Huss	10	(**3**)	7	(**2**)	±14	(**3**)	±48	(**3**)	2.8
3	Machguth	−13	(5)	−12	(**3**)	±14	(**2**)	±45	(**1**)	2.8
4	Maussion	−20	(6)	−21	(6)	±14	(**1**)	±46	(**2**)	3.8
5	Linsbauer	12	(4)	12	(4)	±17	(5)	±54	(5)	4.5
6	Morlighem	−10	(**2**)	−14	(5)	±31	(7)	±89	(8)	5.5
7	Farinotti	22	(7)	22	(7)	±32	(8)	±78	(6)	7.0
8	Brinkerhoff-v2	37	(8)	42	(9)	±28	(6)	±85	(7)	7.5
9	VanPeltLeclercq	57	(9)	39	(8)	±40	(9)	±168	(9)	8.8
Hellstugubreen										
1	Morlighem	2	(**2**)	2	(**1**)	±19	(**2**)	±67	(**1**)	1.5
2	GCbedstress	−2	(**1**)	3	(**2**)	±25	(4)	±112	(8)	3.8
3	Huss	−27	(6)	−24	(7)	±19	(**1**)	±76	(**2**)	4.0
4	Maussion	−29	(7)	−23	(6)	±23	(**3**)	±84	(**3**)	4.8
5	Farinotti	19	(5)	17	(4)	±31	(8)	±97	(4)	5.2
6	Brinkerhoff-v2	−3	(**3**)	10	(**3**)	±30	(7)	±123	(9)	5.5
7	VanPeltLeclercq	14	(4)	18	(5)	±32	(9)	±101	(5)	5.8
8	Linsbauer	−33	(8)	−28	(8)	±28	(6)	±105	(6)	7.0
9	Machguth	−54	(9)	−46	(9)	±26	(5)	±111	(7)	7.5
Kesselwandferner										
1	Linsbauer	−14	(**2**)	−11	(**2**)	±20	(**1**)	±60	(**2**)	1.8
2	Machguth	−17	(4)	−17	(**3**)	±22	(**2**)	±52	(**1**)	2.5
3	Brinkerhoff-v2	−4	(**1**)	−8	(**1**)	±27	(7)	±122	(8)	4.2
4	Morlighem	−20	(6)	−22	(4)	±24	(4)	±69	(5)	4.8
5	Maussion	−28	(7)	−27	(7)	±25	(6)	±69	(4)	6.0
6	Huss	−30	(8)	−24	(6)	±28	(8)	±68	(**3**)	6.2
7	Farinotti	−20	(5)	−24	(5)	±40	(9)	±129	(9)	7.0
8	VanPeltLeclercq	−41	(9)	−39	(9)	±24	(5)	±77	(6)	7.2
9	GCbedstress	52	(10)	42	(10)	±24	(**3**)	±111	(7)	7.5
10	GCneuralnet	−15	(**3**)	−28	(8)	±50	(10)	±135	(10)	7.8
Mocho										
1	GCbedstress	1	(**1**)	9	(**3**)	±44	(6)	±112	(**3**)	3.2
2	Machguth	−27	(6)	−12	(4)	±40	(4)	±109	(**1**)	3.8
3	Farinotti	−17	(**2**)	−3	(**1**)	±50	(8)	±121	(5)	4.0
4	Huss	−28	(7)	−12	(6)	±36	(**2**)	±109	(**2**)	4.2
5	Maussion	−23	(4)	−12	(5)	±39	(**3**)	±125	(6)	4.5
6	VanPeltLeclercq	−21	(**3**)	−9	(**2**)	±45	(7)	±157	(7)	4.8
7	Linsbauer	−33	(8)	−17	(7)	±34	(**1**)	±115	(4)	5.0
8	GCneuralnet	−23	(5)	−28	(8)	±42	(5)	±216	(8)	6.5

Continued on next page.

Table S3: Continued from previous page.

Test case / Model	avg		med		IQR		95%		AVG
North Glacier									
1 Huss	14	(**2**)	16	(**2**)	±26	(**2**)	±83	(**3**)	2.2
2 Maussion	−6	(**1**)	−3	(**1**)	±38	(5)	±78	(**2**)	2.2
3 Linsbauer	68	(8)	69	(7)	±18	(**1**)	±53	(**1**)	4.2
4 GCbedstress	37	(6)	34	(5)	±32	(**3**)	±92	(4)	4.5
5 RAAJgantayat	24	(**3**)	17	(**3**)	±38	(4)	±131	(9)	4.8
6 RAAJglabtop2	34	(5)	32	(4)	±49	(9)	±108	(7)	6.2
7 Farinotti	27	(4)	35	(6)	±49	(8)	±122	(8)	6.5
8 GCneuralnet	−60	(7)	−71	(8)	±47	(7)	±102	(6)	7.0
9 Machguth	−122	(9)	−124	(9)	±39	(6)	±92	(5)	7.2
South Glacier									
1 Maussion	−6	(**3**)	−3	(**2**)	±32	(**1**)	±87	(**2**)	2.0
2 Huss	4	(**1**)	9	(**3**)	±35	(5)	±76	(**1**)	2.5
3 Machguth	5	(**2**)	14	(4)	±35	(4)	±89	(4)	3.5
4 Linsbauer	14	(5)	23	(6)	±33	(**2**)	±87	(**3**)	4.0
5 GCneuralnet	−10	(4)	−1	(**1**)	±52	(8)	±116	(7)	5.0
6 Brinkerhoff-v2	24	(6)	22	(5)	±45	(7)	±134	(8)	6.5
7 Morlighem	29	(7)	39	(8)	±44	(6)	±97	(5)	6.5
8 GCbedstress	45	(9)	48	(9)	±35	(**3**)	±97	(6)	6.8
9 Farinotti	35	(8)	37	(7)	±57	(9)	±140	(9)	8.2
Starbuck									
1 Huss	−14	(**2**)	−19	(**3**)	±27	(**1**)	±94	(**2**)	2.0
2 Linsbauer	−7	(**1**)	−15	(**2**)	±39	(**3**)	±114	(**3**)	2.2
3 Farinotti	−34	(4)	−39	(4)	±28	(**2**)	±83	(**1**)	2.8
4 Machguth	15	(**3**)	15	(**1**)	±59	(4)	±166	(4)	3.0
Tasman									
1 Maussion	−19	(**2**)	−9	(**2**)	±35	(**2**)	±136	(6)	3.0
2 Gantayat-v2	−45	(4)	−51	(6)	±32	(**1**)	±120	(**3**)	3.5
3 RAAJgantayat	−10	(**1**)	7	(**1**)	±63	(8)	±225	(9)	4.8
4 Machguth	−94	(9)	−92	(8)	±45	(4)	±117	(**2**)	5.8
5 Linsbauer	−84	(7)	−91	(7)	±46	(5)	±132	(5)	6.0
6 RAAJglabtop2	−91	(8)	−100	(9)	±39	(**3**)	±120	(4)	6.0
7 Farinotti	−19	(**3**)	−51	(5)	±57	(7)	±225	(10)	6.2
8 Gantayat	52	(5)	39	(4)	±74	(9)	±169	(7)	6.2
9 GCneuralnet	−106	(10)	−116	(10)	±54	(6)	±111	(**1**)	6.8
10 Huss	−73	(6)	−24	(**3**)	±125	(11)	±271	(11)	7.8
11 Morlighem	−237	(11)	−214	(11)	±102	(10)	±196	(8)	10.0
Unteraar									
1 Brinkerhoff-v2	−39	(**2**)	−33	(**2**)	±33	(**1**)	±118	(**3**)	2.0
2 Farinotti	−43	(4)	−40	(**3**)	±41	(**3**)	±119	(4)	3.5
3 Maussion	−58	(6)	−53	(5)	±35	(**2**)	±101	(**1**)	3.5
4 Huss	−52	(5)	−56	(6)	±42	(4)	±115	(**2**)	4.2
5 VanPeltLeclercq	−24	(**1**)	−26	(**1**)	±59	(9)	±184	(14)	6.2
6 RAAJglabtop2	−69	(8)	−64	(7)	±52	(8)	±142	(8)	7.8
7 Linsbauer	−70	(9)	−68	(9)	±52	(7)	±135	(7)	8.0
8 Gantayat	−43	(**3**)	−48	(4)	±70	(13)	±184	(13)	8.2
9 Gantayat-v2	−78	(10)	−77	(10)	±43	(5)	±145	(10)	8.8
10 Machguth	−88	(11)	−90	(12)	±51	(6)	±131	(6)	8.8
11 RAAJgantayat	−67	(7)	−65	(8)	±64	(10)	±171	(12)	9.2
12 GCbedstress	−94	(12)	−89	(11)	±72	(14)	±144	(9)	11.5
13 Morlighem	−146	(15)	−164	(15)	±66	(12)	±129	(5)	11.8
14 GCneuralnet	−137	(14)	−125	(13)	±65	(11)	±171	(11)	12.2
15 Fuerst	−113	(13)	−135	(14)	±86	(15)	±228	(15)	14.2

Continued on next page.

Table S3: Continued from previous page.

Test case / Model		avg		med		IQR		95%		AVG
Urumqi										
1	Brinkerhoff-v2	29	(**3**)	30	(4)	±28	(**2**)	±55	(**1**)	2.5
2	Morlighem	−31	(4)	−29	(**3**)	±27	(**1**)	±102	(**2**)	2.5
3	Farinotti	7	(**2**)	7	(**2**)	±34	(**3**)	±139	(8)	3.8
4	GCbedstress	−1	(**1**)	−5	(**1**)	±60	(8)	±134	(6)	4.0
5	Huss	−54	(5)	−50	(5)	±45	(5)	±113	(4)	4.8
6	Maussion	−97	(8)	−97	(7)	±43	(4)	±104	(**3**)	5.5
7	Linsbauer	−78	(6)	−89	(6)	±57	(6)	±136	(7)	6.2
8	Machguth	−92	(7)	−98	(8)	±60	(7)	±129	(5)	6.8
Washmawapta										
1	Farinotti	−44	(**1**)	−44	(**1**)	±33	(**1**)	±97	(**1**)	1.0
2	GCbedstress	−121	(**2**)	−117	(**2**)	±69	(4)	±141	(**3**)	2.8
3	Maussion	−128	(4)	−128	(**3**)	±48	(**2**)	±106	(**2**)	2.8
4	Huss	−122	(**3**)	−136	(4)	±67	(**3**)	±155	(5)	3.8
5	Linsbauer	−165	(5)	−159	(5)	±70	(5)	±154	(4)	4.8
6	Machguth	−172	(6)	−167	(6)	±74	(6)	±177	(6)	6.0
Synthetic1										
1	Brinkerhoff-v2	−20	(**2**)	−19	(**2**)	±15	(**2**)	±41	(**3**)	2.2
2	Fuerst	−27	(**3**)	−27	(4)	±10	(**1**)	±25	(**1**)	2.2
3	VanPeltLeclercq	19	(**1**)	17	(**1**)	±19	(5)	±62	(6)	3.2
4	RAAJgantayat	−28	(4)	−27	(**3**)	±22	(6)	±50	(5)	4.5
5	Brinkerhoff	−38	(7)	−34	(6)	±18	(4)	±46	(4)	5.2
6	Morlighem	−39	(9)	−37	(7)	±17	(**3**)	±39	(**2**)	5.2
7	Gantayat	−39	(10)	−46	(10)	±31	(9)	±68	(7)	9.0
8	Huss	−38	(8)	−40	(8)	±31	(10)	±80	(12)	9.5
9	Farinotti	−41	(11)	−48	(11)	±23	(7)	±80	(11)	10.0
10	Gantayat-v2	−45	(12)	−53	(13)	±23	(8)	±74	(8)	10.2
11	Rabatel	29	(5)	34	(5)	±46	(16)	±123	(15)	10.2
12	GCbedstress	−34	(6)	−41	(9)	±37	(13)	±85	(14)	10.5
13	Linsbauer	−54	(14)	−51	(12)	±33	(11)	±78	(10)	11.8
14	Maussion	−69	(16)	−68	(16)	±36	(12)	±74	(9)	13.2
15	Machguth	−57	(15)	−56	(14)	±37	(14)	±82	(13)	14.0
16	GCneuralnet	−50	(13)	−61	(15)	±46	(15)	±163	(16)	14.8
Synthetic2										
1	GCbedstress	−9	(4)	−10	(5)	±10	(5)	±30	(**2**)	4.0
2	RAAJglabtop2	−2	(**1**)	−9	(**3**)	±13	(9)	±53	(8)	5.2
3	VanPeltLeclercq	11	(6)	9	(4)	±12	(8)	±35	(4)	5.5
4	Linsbauer	−18	(8)	−20	(9)	±10	(4)	±33	(**3**)	6.0
5	Brinkerhoff-v2	−22	(11)	−17	(7)	±9	(**3**)	±37	(5)	6.5
6	Farinotti	−5	(**2**)	7	(**2**)	±18	(12)	±74	(11)	6.8
7	Brinkerhoff	−30	(14)	−30	(13)	±6	(**1**)	±27	(**1**)	7.2
8	Machguth	−13	(7)	−19	(8)	±11	(7)	±55	(9)	7.8
9	Fuerst	−28	(13)	−25	(12)	±6	(**2**)	±40	(6)	8.2
10	Huss	11	(5)	3	(**1**)	±33	(15)	±93	(14)	8.8
11	Maussion	−21	(10)	−16	(6)	±31	(13)	±75	(12)	10.2
12	Gantayat	−19	(9)	−21	(10)	±16	(11)	±90	(13)	10.8
13	Gantayat-v2	−26	(12)	−22	(11)	±14	(10)	±66	(10)	10.8
14	Morlighem	−32	(15)	−32	(15)	±11	(6)	±52	(7)	10.8
15	GCneuralnet	−7	(**3**)	−30	(14)	±31	(14)	±161	(15)	11.5

Continued on next page.

Table S3: Continued from previous page.

Test case / Model		avg		med		IQR		95%		AVG
Synthetic3										
1	GCbedstress	−3	(**2**)	−2	(**2**)	±10	(7)	±34	(6)	4.2
2	Fuerst	−17	(7)	−15	(8)	±5	(**1**)	±28	(**3**)	4.8
3	Brinkerhoff	−18	(8)	−17	(9)	±6	(**2**)	±18	(**1**)	5.0
4	VanPeltLeclercq	14	(6)	14	(7)	±6	(**3**)	±28	(4)	5.0
5	Gantayat	−9	(5)	−5	(**3**)	±9	(6)	±42	(7)	5.2
6	Huss	−1	(**1**)	0	(**1**)	±18	(11)	±75	(12)	6.2
7	Farinotti	7	(**3**)	8	(5)	±14	(8)	±64	(10)	6.5
8	Morlighem	−24	(12)	−22	(12)	±6	(4)	±25	(**2**)	7.5
9	Brinkerhoff-v2	−23	(10)	−20	(11)	±7	(5)	±31	(5)	7.8
10	Machguth	−8	(4)	−18	(10)	±15	(9)	±54	(9)	8.0
11	Gantayat-v2	−18	(9)	−10	(6)	±16	(10)	±46	(8)	8.2
12	Linsbauer	23	(11)	7	(4)	±31	(12)	±73	(11)	9.5
13	GCneuralnet	58	(13)	69	(13)	±46	(13)	±105	(13)	13.0

Supplementary References

Arendt, A., , and 87 others (2015). *Randolph Glacier Inventory [5.0]: A Dataset of Global Glacier Outlines, Version 5.0*. Global Land Ice Measurements from Space (GLIMS), Boulder, Colorado, USA. Digital Media.

Balay, S., Abhyankar, S., Adams, M. F., Brown, J., Brune, P., Buschelman, K., Dalcin, L., Eijkhout, V., Gropp, W. D., Kaushik, D., Knepley, M. G., McInnes, L. C., Rupp, K., Smith, B. F., Zampini, S., Zhang, H., and Zhang, H. (2016). PETSc Web page.

Brinkerhoff, D. J., Aschwanden, A., and Truffer, M. (2016). Bayesian inference of subglacial topography using mass conservation. *Frontiers in Earth Science*, 4(8):1–15. doi: 10.3389/feart.2016.00008.

Brinkerhoff, D. J. and Johnson, J. V. (2013). Data assimilation and prognostic whole ice sheet modelling with the variationally derived, higher order, open source, and fully parallel ice sheet model VarGlaS. *The Cryosphere*, 7(4):1161–1184. doi: 10.5194/tc-7-1161-2013.

Brooks, A. N. and Hughes, T. J. (1982). Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. *Computer methods in applied mechanics and engineering*, 32(1-2):199–259. doi: 10.1016/0045-7825(82)90071-8.

Bueler, E. and Brown, J. (2009). Shallow shelf approximation as a "sliding law" in a thermomechanically coupled ice sheet model. *Journal of Geophysical Research: Earth Surface*, 114(F3):F03008. doi: 10.1029/2008JF001179.

Clarke, G. K. C., Anslow, F. S., Jarosch, A. H., Menounos, B., Bolch, T., and Berthier, E. (2013). Ice volume and subglacial topography for western Canadian glaciers from mass balance fields, thinning rates, and a bed stress model. *Journal of Climate*, 26(12):4282–430. doi: 10.1175/JCLI-D-12-00513.1.

Clarke, G. K. C., Berthier, E., Schoof, C. G., and Jarosch, A. H. (2009). Neural networks applied to estimating subglacial topography and glacier volume. *Journal of Climate*, 22:2146–2160. doi: 10.1175/2008JCLI2572.1.

Cuffey, K. and Paterson, W. (2010). *The Physics of Glaciers*. Elsevier, Oxford, U.K., fourth edition.

Farinotti, D., Huss, M., Bauder, A., and Funk, M. (2009a). An estimate of the glacier ice volume in the Swiss Alps. *Global and Planetary Change*, 68:225–231. doi: 10.1016/j.gloplacha.2009.05.004.

Farinotti, D., Huss, M., Bauder, A., Funk, M., and Truffer, M. (2009b). A method to estimate ice volume and ice thickness distribution of alpine glaciers. *Journal of Glaciology*, 55(191):422–430. doi: 10.3189/002214309788816759.

Frey, H., Machguth, H., Huss, M., Huggel, C., Bajracharya, S., Bolch, T., Kulkarni, A., Linsbauer, A., Salzmann, N., and Stoffel, M. (2014). Estimating the volume of glaciers in the Himalayan–Karakoram region using different methods. *The Cryosphere*, 8(6):2313–2333. doi: 10.5194/tc-8-2313-2014.

Fürst, J. J., Gillet-Chaulet, F., Benham, T. J., Dowdeswell, J. A., Grabiec, M., Navarro, F., Pettersson, R., Moholdt, G., Nuth, C., Sass, B., Aas, K., Fettweis, X., Lang, C., Seehaus, T., and Braun, M. (2017). Application of a two-step approach for mapping ice thickness to various glacier types on Svalbard. *The Cryosphere Discussions*, 2017:1–43. doi: 10.5194/tc-2017-30.

Gabbi, J., Farinotti, D., Bauder, A., and Maurer, H. (2012). Ice volume distribution and implications on runoff projections in a glacierized catchment. *Hydrology and Earth System Sciences*, 16:4543–4556. doi: 10.5194/hess-16-4543-2012.

Gagliardini, O., Zwinger, T., Gillet-Chaulet, F., Durand, G., Favier, L., de Fleurian, B., Greve, R., Malinen, M., Martin, C., Raback, P., Ruokolainen, J., Sacchettini, M., Schafer, M., Seddik, H., and Thies, J. (2013). Capabilities and performance of Elmer/Ice, a new-generation ice sheet model. *Geoscientific Model Development*, 6(4):1299–1318. doi: 10.5194/gmd-6-1299-2013.

Gantayat, P., Kulkarni, A., and Srinivasan, J. (2014). Estimation of ice thickness using surface velocities and slope: case study at Gangotri Glacier, India. *Journal of Glaciology*, 60(220):277–282. doi: 10.3189/2014JoG13J078.

Gärtner-Roer, I., Naegeli, K., Huss, M., Knecht, T., Machguth, H., and Zemp, M. (2014). A database of worldwide glacier thickness observations. *Global and Planetary Change*, 122:330–344. doi: 10.1016/j.gloplacha.2014.09.003.

Gilbert, J. C. and Lemaréchal, C. (1989). Some numerical experiments with variable-storage quasi-Newton algorithms. *Mathematical Programming*, 45:407–435. doi: 10.1007/BF01589113.

Gillet-Chaulet, F., Gagliardini, O., Seddik, H., Nodet, M., Durand, G., Ritz, C., Zwinger, T., Greve, R., and Vaughan, D. G. (2012). Greenland ice sheet contribution to sea-level rise from a new-generation ice-sheet model. *The Cryosphere*, 6(6):1561–1576. doi: 10.5194/tc-6-1561-2012.

Glen, J. W. (1955). The creep of polycrystalline ice. *Proceedings of the Royal Society of London, Series A*, 228(1175):519–538.

Haeberli, W. and Hoelzle, M. (1995). Application of inventory data for estimating characteristics of and regional climate-change effects on mountain glaciers: a pilot study with the European Alps. *Annals of Glaciology*, 21:206–212. doi: 10.3198/1995AoG21-1-206-212.

Harris, I., Jones, P. D., Osborn, T. J., and Lister, D. H. (2014). Updated high-resolution grids of monthly climatic observations the CRU TS3.10 Dataset. *International Journal of Climatology*, 34(3):623–642. doi 10.1002/joc.3711.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika*, 57:97109. doi: 10.1093/biomet/57.1.97.

Huss, M. and Farinotti, D. (2012). Distributed ice thickness and volume of all glaciers around the globe. *Journal of Geophysical Research*, 117:F04010. doi: 10.1029/2012JF002523.

Hutchinson, M. F. (1989). A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. *Journal of Hydrology*, 106(3):211–232. doi: 10.1016/0022-1694(89)90073-5.

Hutter, K. (1983). *Theoretical glaciology; material science of ice and the mechanics of glaciers and ice sheets*. D. Reidel Publishing Company, Tokyo, Terra Scientific Publishing Company.

Huybrechts, P. (1991). *The Antartic ice sheet and enviremental change: a three-dimensional modelling study*. PhD thesis, Vrije Universiteit of Brussel. doi: 10013/epic.12054.d001.

Kelley, C. T. and Keyes, D. E. (1998). Convergence analysis of pseudo-transient continuation. *SIAM Journal on Numerical Analysis*, 35(2):508–523. doi: 10.1137/S0036142996304796.

Kienholz, C., Rich, J. L., Arendt, A. A., and Hock, R. (2014). A new method for deriving glacier centerlines applied to glaciers in Alaska and northwest Canada. *The Cryosphere*, 8(2):503–519. doi: 10.5194/tc-8-503-2014.

Leclercq, P. W., Pitte, P., Giesen, R. H., Masiokas, M. H., and Oerlemans, J. (2012). Modelling and climatic interpretation of the length fluctuations of Glaciar Frias (north Patagonian Andes, Argentina) 1639-2009 AD. *Climate of the Past*, 8(5):1385–1402. doi: 10.5194/cp-8-1385-2012.

Linsbauer, A., Paul, F., and Haeberli, W. (2012). Modeling glacier thickness distribution and bed topography over entire mountain ranges with GlabTop: Application of a fast and robust approach. *Journal of Geophysical Research*, 117:F03007. doi: 10.1029/2011JF002313.

Linsbauer, A., Paul, F., Hoelzle, M., Frey, H., and Haeberli, W. (2009). The Swiss Alps without glaciers - A GIS-based modelling approach for reconstruction of glacier beds. In *Proceedings of Geomorphometry 2009*, pages 243–247, Zurich. Available from `www.geomorphometry.org/Linsbauer2009`.

MacAyeal, D. R. (1993). A tutorial on the use of control methods in ice-sheet modeling. *Journal of Glaciology*, 39(131):91–98. doi: 10.3198/1993JoG39-131-91-98.

Marzeion, B., Jarosch, A., and Hofer, M. (2012). Past and future sea-level change from the surface mass balance of glaciers. *The Cryosphere*, 6:1295–1322. doi:10.5194/tc-6-1295-2012.

Maussion, F., Rothenpieler, T., Marzeion, B., Landmann, J., Oesterle, F., Jarosch, A., Recinos, B., and Vlug, A. (2017). Oggm/oggm: v0.1.1. doi: 10.5281/zenodo.292630.

Michel, L., Picasso, M., Farinotti, D., Funk, M., and Blatter, H. (2013). Estimating the ice thickness of mountain glaciers with an inverse approach using surface topography and mass-balance. *Inverse Problems*, 29(3):035002. doi: 10.1088/0266-5611/29/3/035002.

Morlighem, M., Rignot, E., Mouginot, J., Seroussi, H., and Larour, E. (2014). Deeply incised submarine glacial valleys beneath the Greenland Ice Sheet. *Nature Geoscience*, 7(6):418–422. doi: 10.1038/ngeo2167.

Morlighem, M., Rignot, E., Seroussi, H., Larour, E., Dhia, H. B., and Aubry, D. (2011). A mass conservation approach for mapping glacier ice thickness. *Geophysical Research Letters*, 38:L19503. doi: 10.1029/2011GL048659.

Pattyn, F. (2003). A new three-dimensional higher-order thermomechanical ice sheet model: Basic sensitivity, ice stream development, and ice flow across subglacial lakes. *Journal of Geophysical Research: Solid Earth*, 108(B8):2382. doi: 10.1029/2002JB002329.

Reerink, T. J., Kliphuis, M. A., and van de Wal, R. S. W. (2010). Mapping technique of climate fields between GCM's and ice models. *Geoscientific Model Development*, 3(1):13–41. doi: 10.5194/gmd-3-13-2010.

van Pelt, W. J. J., Oerlemans, J., Reijmer, C. H., Pettersson, R., Pohjola, V. A., Isaksson, E., and Divine, D. (2013). An iterative inverse method to estimate basal topography and initialize ice flow models. *The Cryosphere*, 7:987–1006. doi: 10.5194/tc-7-987-2013.

WGMS (2016). Glacier thickness database 2.0. I. Gärtner-Roer, L. M. Andreassen, E. Bjerre, D. Farinotti, A. Fischer, M. Fischer, K. Helfricht, M. Huss, T. Knecht, S. Kutuzov, J. Landmann,I. Lavrentiev, H. Li, Z. Li, H. Machguth, K. Naegeli, F. Navarro, A. Rabatel, P. Stentoft, M. Zemp (eds.), World Glacier Monitoring Service, Zurich, Switzerland. doi: 10.5904/wgms-glathida-2016-07.

Zienkiewicz, O. C. and Taylor, R. L. (2000). *The finite element method: The basis*. Referex Engineering. Butterworth-Heinemann.