A framework for time-dependent Ice Sheet Uncertainty Quantification, applied to three West Antarctic ice streams

Beatriz Recinos[1], Daniel Goldberg[1], James R. Maddison[2], and Joe Todd[1]

[1]School of GeoSciences, The University of Edinburgh, UK.
[2]School of Mathematics and Maxwell Institute for Mathematical Sciences, The University of Edinburgh, UK.

Correspondence: B. Recinos (beatriz.recinos@ed.ac.uk)

Abstract. Ice sheet models are the main tool to generate forecasts of ice sheet mass loss; a significant contributor to sea-level rise, thus knowing the likelihood of such projections is of critical societal importance. However, to capture the complete range of possible projections of mass loss, ice sheet models need efficient methods to quantify the forecast uncertainty. Uncertainties originate from the model structure, from the climate and ocean forcing used to run the model and from model calibration. Here we quantify the latter, applying an error propagation framework to a realistic setting in West Antarctica. As in many other ice-sheet modelling studies we use a control method to calibrate grid-scale flow parameters (parameters describing the basal drag and ice stiffness) with remotely-sensed observations. Yet our framework augments the control method with a Hessian-based Bayesian approach that estimates the posterior covariance of the inverted parameters. This enables us to quantify the impact of the calibration uncertainty on forecasts of sea-level rise contribution or volume above flotation (VAF), due to the choice of different regularisation strengths (prior strengths), sliding laws and velocity inputs. We find that by choosing different satellite ice velocity products our model leads to different estimates of VAF after 40 years. We use this difference in model output to quantify the variance that projections of VAF are expected to have after 40 years and identify prior strengths that can reproduce that variability. We demonstrate that if we use prior strengths suggested by L-curve analysis, as is typically done in ice-sheet calibration studies, our uncertainty quantification is not able to reproduce that same variability. The regularisation suggested by the L-curves is too strong and thus propagating the observational error through to VAF uncertainties under this choice of prior leads to errors that are smaller than those suggested by our 2-member "sample" of observed velocity fields. Additionally, our experiments suggest that large amounts of data points may be redundant, with implications for the error propagation of VAF.

1 Introduction

Ice sheet models are important tools not only for generating knowledge, but also for operational forecasts. In this way, they are analogous to weather models and oceanographic models and have emerged as the de facto standard for generating projections of ice sheet contribution to sea-level rise. However, quantifying the uncertainty in forecasts produced by these models remain one of the most challenging goals of scientific inquiry (Aschwanden et al., 2021). Here, we seek to characterise the uncertainty in model projections of marine ice sheet loss which arises from calibration with data.

The paradigm of ice-sheet projection is the calibration of the model parameters with observations (via control methods e.g.
Macayeal, 1992) followed by running of the calibrated model forward in time forced by future ocean and climate scenarios.
The process is uncertain due to (i) model/structural uncertainty (i.e. uncertainty in the formulation of the model and its ability to
represent the physics of the system), (ii) uncertainty in external forcing (e.g. ocean melting of ice shelves), and (iii) calibration
uncertainty (i.e. the uncertainty in calibrated parameters, sometimes referred to as parametric uncertainty). In this study we
use control methods and a Bayesian inference approach to characterise (iii). The Bayesian framework computes posterior
information given the assumed model and external forcing. We do not attempt to quantify (i) and (ii) but we discuss how these
uncertainties can be quantified and incorporated into our error propagation framework.

The use of control methods ("inverse methods") in ice-sheet modelling dates back to Macayeal (1992). Since then, their
use in estimating basal and internal conditions (hidden properties) of glaciers and ice sheets from measured surface velocities
has become widespread (e.g. Sergienko et al., 2008; Morlighem et al., 2010; Cornford et al., 2015; Hill et al., 2021, to name
a few). This is mostly due to the ability of these methods to perform large-scale inversions via the minimisation of a cost
function, thus allowing a better representation of basal and rheological conditions to which the ice flow is sensitive (Barnes
et al., 2021). However, these data assimilation techniques are not well posed (Petra et al., 2014) and a unique solution is never
guaranteed, regardless of the control method used (Barnes et al., 2021). Control methods have regularisation terms which need
to be chosen in order to impose smoothness on the inverted parameters (Koziol et al., 2021). In many studies, the strength
of the regularisation is determined heuristically through L-curve analysis (Gillet-Chaulet et al., 2012; Barnes et al., 2021).
Additionally, control methods do not provide calibration uncertainty. They can be interpreted as methods that return only the
mode of a posterior probability density function (PDF) of the inverted model parameters, which does not fully characterise
calibration uncertainty (Koziol et al., 2021) nor does it propagate the observational uncertainty onto projections of sea-level
rise.

Previous works attempt to quantify uncertainty by considering the forcing uncertainty (Tsai et al., 2017; Robel et al., 2019;
Levermann et al., 2020) or structural uncertainty (Hill et al., 2021). Others consider calibration uncertainty (Isaac et al., 2015;
DeConto and Pollard, 2016; Brinkerhoff et al., 2021; Brinkerhoff, 2022) but use low-dimensional parameter sets to describe
the ice rheology and basal friction. Here we carry out the first assessment of calibration uncertainty using a time-dependent
marine ice-sheet model in which the calibration of the ice dynamic parameters scale with the dimension of the numerical grid
(see Fig. 1).

The uncertainty associated with ice-sheet model calibration in this sense can be quantified through Bayesian inference
(Tarantola, 2005; Stuart, 2010), in which prior knowledge is "updated" with observational evidence (Koziol et al., 2021).
The solution of the Bayesian inference problem in our framework hence takes the form of a very high-dimensional posterior
PDF. The calculation of covariances from integration using this PDF, due to the complexity of the Stokes partial differential
equations and our domain, is not possible via state-of-the-art Markov chain Monte Carlo (MCMC) methods and prohibited
by computational expense (Isaac et al., 2015; Koziol et al., 2021). However, it can be shown that under certain assumptions,
the posterior covariance of the inverted parameters can be characterised by the inverse of the Hessian (the matrix of second
derivatives) of the cost function with respect to the inverted parameters (Thacker, 1989; Kalmikov and Heimbach, 2014; Petra

et al., 2014; Isaac et al., 2015). Our framework augments the control method by using this Hessian-based Bayesian approach

60   that not only inverts for the ice dynamic parameters such that model velocities match observations, but characterises the posterior covariance of each inverted parameter (also referred to as control parameters in this study).

We perform a joint inversion for the basal sliding coefficient and the rheological parameter for describing ice stiffness. Beginning with a cost function definition which allows for velocity data to be imposed at arbitrary locations (i.e. a point-cloud), we generate a low-rank update approximation to the posterior covariance of the control parameters via the use of the

65   Hessian of the cost function, and find the sensitivities of a time-evolving Quantity of Interest (QoI) to the control parameters. We then project the covariance on to the resulting linear sensitivity to estimate the growth of the QoI uncertainty over time; here our QoI is the sea-level rise contribution or volume above flotation (VAF).

We apply for the first time this error propagation framework to a realistic setting (three ice streams in West Antarctica) and present several model experiments that explore the impact on the uncertainty in forecasts of VAF due to the choice of different

70   strengths of priors, sliding laws and velocity inputs. We find that significant differences in satellite ice velocity products (particularly at the ice margins) can lead to different projected estimates of sea-level rise contributions or VAF trajectories. We also find that the choice of regularisation strength or regularisation parameters, suggested by L-curve analysis – a common means of estimating such parameters – may lead to an overly informative prior and hence underestimate the variability in such projections.

75   We investigate the effect that data density (density of observed velocity data points) has on the resulting inference. This diagnostic suggests that large amount of data points may be redundant when inverting for the control parameters, with potential implications for observational velocity error models and how they inform the uncertainty in our projections. Additionally, we test our inversion results against the numerical framework of a different ice sheet model (i.e. STREAMICE module of MITgcm Goldberg and Heimbach, 2013), in order to qualitatively inspect model structural uncertainty and forcing uncertainty.

80   2   Methods

The mathematical framework of `FEniCS_ice` is explained in detail in Koziol et al. (2021). In this section we summarise the model physics, the data assimilation techniques used for the calibration of two key ice dynamic parameters, and explain how we quantify calibration uncertainty in projections of sea-level rise contributions or volume above flotation (VAF).

2.1   Physics

85   `FEniCS_ice` solves the Stokes equation by implementing the well-known Shallow Shelf Approximation (SSA; MacAyeal, 1989; Schoof, 2006; Shapero et al., 2021; Hill et al., 2021). The ice velocity $\mathbf{u}$ is vertically integrated and has two components: internal deformation and basal sliding (see Sect. 3 from Koziol et al., 2021, for details). The model uses data assimilation methods to optimise these velocity components based on observations by estimating two "hidden" properties of the ice; i) the basal friction coefficient (α) in the sliding law, and ii) the rheological parameter for describing ice stiffness (β) in Glen's flow

90   law (both properties are referred to as control parameters in this study). In this section we define the control parameters and the time-dependent SSA whereas the details of the inverse methodology are explained in Sect. 2.4.

2.1.1   Ice rheology and basal sliding

We define the ice viscosity ν, which depends on the strain-rate tensor ε_e as

$$\nu = \frac{1}{2} B \varepsilon_e^{\frac{1-n}{2n}}.$$

95   B is generally referred to as the "stiffness" of the ice and is thought to depend on ice temperature (Pattyn, 2010). Here we define the control parameter β as the square root of that stiffness where $\beta = \sqrt{B} = \sqrt{A^{-1/n}}$. A in this definition is the rate factor commonly known as the ice creep parameter in Glen's ice flow law (Glen and Perutz, 1955) and n is the exponent of Glen's flow law with the widely accepted value of 3 (Cuffey and Paterson, 2010).

Basal sliding is considered the dominant component of surface velocities in fast-flowing ice streams (Hill et al., 2021),
100   making the time-dependent part of the ice sheet model sensitive to the choice of sliding law (Brondex et al., 2019; Barnes and Gudmundsson, 2022) thus, we consider two different sliding laws. The first is the Weertman–Budd sliding law (Weertman, 1957; Budd et al., 1979; Budd and Jenssen, 1987) defined here as

$$\tau_{\mathbf{b}} = \alpha^2 N^{1/3} u^{-2/3} \mathbf{u} \tag{1}$$

where $\tau_{\mathbf{b}}$ is basal stress, α is the scalar, spatially varying sliding coefficient, u is ice speed, and N is the effective pressure.
105   Here N is defined as

$$N = \rho_i g H + \min(0, R) \rho_w g \tag{2}$$

where ρ_i and ρ_w are ice and ocean densities, g is the magnitude of the gravitational acceleration, H is the ice thickness and R is the bed elevation (Koziol et al., 2021). Furthermore basal stress is nonzero only where ice is grounded, i.e. where $\rho_i g H + \rho_w R > 0$. The second sliding law considered is often referred to as the Cornford sliding law (Asay-Davis et al., 2016;
110   Cornford et al., 2020) and is defined as

$$\tau_{\mathbf{b}} = \frac{\mu \alpha^2 N u^{\frac{1-m}{m}}}{[\alpha^{2m} u + (\mu N)^m]^{1/m}} \mathbf{u} \tag{3}$$

where $\mu = \frac{1}{2}$ and $m = 3$. A key property of both sliding laws is that as the grounding line is approached, effective pressure becomes small, leading to a smoother transition across the grounding line in terms of the basal drag from floating to grounded conditions.

115   While additional sliding laws have been proposed and are now implemented within a number of existing ice flow models (Hill et al., 2021), in this study we use the Weertman–Budd sliding law for most of our experiments, as it is one of the most commonly used. However, we trial our error model framework using the Cornford sliding law (Asay-Davis et al., 2016) and compare the results of both sliding laws in Sect. 5.3.

2.1.2 Time-dependent ice sheet model

120 The resulting calibrated fields of α and β (see Sect. 2.4 for details regarding the parameters calibration) are then input into our forward-in-time simulations where the continuity equation is solved:

$$H_t + \nabla \cdot (H\mathbf{u}) = b. \tag{4}$$

b represents localised changes in mass at the surface and/or the base of the ice sheet, i.e. accumulation due to snow-fall or basal melting of the ice-shelf by the ocean. We assume a constant surface mass balance field (i.e. surface mass balance of 0.38

125 mm of sea level equivalent based in Arthern et al., 2006) and implement a simple depth-dependent parameterisation of ocean melt rate m, which gives the melt rate as a function of ice-shelf draft only. Such parameterisations have been used previously to examine the response of marine ice sheets to ice-shelf melting (e.g., Favier et al., 2014; Seroussi et al., 2017; Lilien et al., 2019; Robel et al., 2019). The form we use is

$$m(z_b) = \frac{M_{max}}{2}\left(1 + \tanh\left[2\left(\frac{z_b - z_{th}}{z_{th}}\right)\right]\right) \tag{5}$$

130 where z_b is ice-shelf depth, M_{max} is the maximum melt rate and z_{th} represents the depth of the ocean thermocline. m is nonzero only where ice-shelf thickness H is below floatation, and is also set to zero where thickness H is below $10\,\mathrm{m}$.

We use such a parameterisation because our aim is to study glaciers which are strongly forced by modified Circumpolar Deep Water (CDW), which is present on certain parts of the Antarctic continental shelf as a warm deep layer overlain by cold surface-modified waters (e.g., Jacobs et al., 2011; Dutrieux et al., 2014; Jenkins, 2016; Jenkins et al., 2018). The form of (5) is

135 chosen because the melt profile transitions from low melt rates above the thermocline depth z_{th} to strong melting at depth, and saturates at M_{max} rather than growing without bound. Defining the parameterisation in this way rather than a piecewise-linear function helps maintain differentiability which aids the later application of algorithmic differentiation (Section 2.5). We discuss our particular choice of M_{max} and z_{th} below in Sect. 3.

The continuity equation is solved with the purpose of finding the loss of ice volume above flotation (VAF), the volume of

140 ice that can contribute to sea level at a certain time T (e.g. $T = 40$ years) which is defined as

$$Q_T^{VAF} = \int\limits_\Omega (H - H_f)^+ dA. \tag{6}$$

where H_f is the *floatation thickness* defined by $-R\left(\frac{\rho_w}{\rho_i}\right)$, and Ω is the computational domain (see Sect. 2.2 and Sect. 3 for details). Note that we have simplified the ice sheet surface mass balance and the basal melting of the ice shelf thus calculations of future VAF estimates presented here do not constitute realistic projections. However, equations (4) and (6) are convenient to

145 calculate such projections and sufficiently nontrivial and nonlinear that the effect of uncertainty arising from the calibration of α and β with observations can be seen.

2.2 Discretisation

We solve the Shallow-Shelf Approximation (SSA) momentum balance as well as (4) using the FEniCS finite-element software library (Alnæs et al., 2015). We discretise velocity $\mathbf{u}$, bathymetry R, and drag and stiffness parameters α and β using first-order

150 continuous Lagrange elements on a triangular finite element mesh. Thickness is defined to be constant within elements (a DG(0) discretisation). The only non-standard aspect of the formulation is in the weak definition of the driving stress $\rho g H \nabla z_{surface}$, which is written as $\mathcal{F} + \mathcal{W} \nabla R$ (see Sect. 3 in Koziol et al., 2021, for $\mathcal{F}$ and $\mathcal{W}$ definitions and the SSA formulation). This formulation is equivalent to the more standard form of driving stress when H is discretised using continuous finite elements, but with this form H can be discretised using zero-order discrete-galerkin (DG(0)) elements as well. The continuity equation (4)

155 is solved using a simple first-order upwind scheme, which is found to be more stable when using a DG(0) thickness function. Details of the mesh generation are explained in Sect. 3 when we discuss the study area.

2.3 Notation

To facilitate readability of this and subsequent sections we adopt formatting conventions for different mathematical objects. Coefficient vectors corresponding to finite-element functions appear as $\bar{c}$; other vectors and vector-valued functions as $\breve{d} \in \mathbb{R}^q$;

160 and matrices as $\mathbf{E}$.

2.4 Cost function J^c

To calibrate the basal sliding coefficient (α) and the rheological parameter for describing ice stiffness (β) we apply data assimilation techniques usually used in glaciology (Morlighem et al., 2010; Joughin et al., 2010; Cornford et al., 2015), where the aim is to find the parameter sets which gives the best fit to ice velocity observations. Our approach augments such data

165 assimilation techniques by using a Hessian-based Bayesian approach to characterise uncertainty of α and β. In Sect. 2.5 we describe how we propagate the errors that result from this calibration into projections of VAF. Here we describe how we invert for the control parameters via the minimisation of a scalar cost function which takes the general form

$$J^c = J^c_{mis} + J^c_{reg}. \tag{7}$$

J^c_{mis}, the misfit cost, is half the square-integral of the misfit between the surface velocity of the ice model and remotely-

170 sensed observations, normalised by the observational standard deviation. These terms are discretised to implement the control method (as described in Sect. 2.2). The misfit cost is

$$J^c_{mis} = \frac{1}{2} \| \breve{u}_{obs} - \breve{u} \|^2_{\mathbf{\Gamma}^{-1}_{\mathbf{obs}}} \tag{8}$$

Here $\breve{u}_{obs}$ is the observed velocity given as cloud point data (location and velocity value) and $\breve{u}$ is the velocity estimated via the SSA approximation, interpolated at $\breve{u}_{obs}$ coordinates. As error covariance is not given, $\mathbf{\Gamma}_{\mathbf{obs}}$ is a diagonal matrix containing

175 the standard deviation (STD) of the observations – note that this neglects observational covariance (see Sect. 3.1.3 and Sect. 6.4 where we discuss observational error covariance).

J^c_{reg}, the regularisation cost, is imposed to prevent instabilities, and is typically chosen as a Tikhonov operator which penalises the square-integral of the gradient of the parameter field (e.g., Morlighem et al., 2010; Cornford et al., 2015). It is defined as

180 $$J^c_{reg} = \frac{1}{2} \| \bar{c} - \bar{c}_0 \|^2_{\mathbf{\Gamma}^{-1}_{\mathbf{prior}}}, \tag{9}$$

where $\overline{c}$ is our hidden field, which depends on both control parameters $\overline{c} = (\alpha, \beta)$. $\overline{c}_0$ is the prior mean and the symmetric positive definite $\mathbf{\Gamma_{prior}}$ is the prior covariance of the control parameters, the construction of which is described below.

2.4.1 Prior distribution of parameters

The prior covariance between α and β is set to zero and, following from Koziol et al. (2021), we define a prior covariance for each of α and β

$$\mathbf{\Gamma_{prior}} = \mathbf{L}^{-1}\mathbf{M}\mathbf{L}^{-1}, \tag{10}$$

where $\mathbf{M}$ is the finite element *mass matrix*, and $\mathbf{L}$ is the matrix arising from the finite-element discretisation of the differential operator

$$\mathcal{L}(\cdot) \equiv \gamma \nabla^2(\cdot) - \delta(\cdot). \tag{11}$$

α_0, the prior mean of α, is zero. The prior mean of β is given by

$$\beta_0 = \left(I - \frac{\gamma}{\delta}\nabla^2 \right)^{-1} \beta_{bgd}, \tag{12}$$

where β_{bgd} is the initial guess, described in Section 3. Our form differs from the square-gradient regularisation sometimes used in control methods (e.g., Morlighem et al., 2010; Cornford et al., 2015), but avoids infinite variance as the mesh is refined (Bui-Thanh et al., 2013).

An important aspect of implementing a prior is to choose the strength with which it is imposed. Here γ and δ in (11) determine the strength of our prior. If the prior is strong (e.g. γ is very small) it will force the posterior to match the prior; dragging the posterior PDF away from the control parameter values which the data suggests are more likely (see Sect. 2.5 for the definition of the posterior PDF of the control parameters). A weak prior imposes only weak restrictions on the plausible range of parameters – potentially beyond the physically plausible range expected a priori. A weak prior can also present practical difficulties, as the posterior maximiser can contain undesirable and nonphysical features that lead to non-stable solutions of the time-dependent SSA. As γ and δ are nonintuitive, we make use of the following expressions for a characteristic pointwise variance $\sigma^2_{c(0)}$ and auto-covariance length scale l_0 of each control parameter (see Sect. 2.2 in Lindgren et al., 2011, for details):

$$\sigma^2_{c(0)} = \frac{1}{(4\pi\gamma\delta)} \tag{13}$$

$$l_{c(0)} = \sqrt{\frac{\gamma}{\delta}} \tag{14}$$

In in Sect. 5.1, we show the impact on the VAF projection uncertainty due to the choice of different prior strengths.

2.5 Error propagation framework

Finally, our goal is to find the **posterior probability density function** (PDF) of the control parameters ($\overline{c}$) given the observational data ($\breve{u}_{obs}$); $p(\overline{c}|\breve{u}_{obs})$, and propagate forward the associated uncertainty in time-dependent VAF (denoted here as just

Q_T for short). The error propagation framework used here follows from Isaac et al. (2015) and similar studies (Bui-Thanh

210   et al., 2013; Petra et al., 2014), and has been described in detail by Koziol et al. (2021) – here the key elements are summarised.

The cost function (7) can be interpreted in a Bayesian sense, with J^c_{reg} being the negative logarithm of the prior density, and J^c_{mis} being the negative logarithm of the likelihood density (each up to a normalisation term). J^c is then the negative logarithm of the posterior PDF of the control parameters conditioned on the data. The mode of the posterior PDF (i.e. the maximum a posteriori, or MAP, estimate, assumed here to be unique) is given by the value of the parameters at which J^c

215   obtains its minimum. In the case of a linear model, the posterior inverse covariance, denoted $\mathbf{\Gamma}^{-1}_{\mathbf{post}}$, is given by the Hessian matrix (here referred to as the 'Hessian') of J^c evaluated at the MAP point. In the general case the Hessian defines a Gaussian approximation for the posterior PDF (as the second order approximation for its negative logarithm at the MAP point) and thus defines an approximation for the inverse posterior covariance.

If we have estimates of VAF at a given time (6), which depend linearly on the control parameters, and if the posterior ($\mathbf{\Gamma}_{\mathbf{post}}$)

220   is Gaussian, then the posterior variance of VAF at a time T, is given by

$$\sigma^2(Q_T) = \left(\frac{\partial Q_T}{\partial \overline{c}}\right)^T \mathbf{\Gamma}_{post} \left(\frac{\partial Q_T}{\partial \overline{c}}\right) \tag{15}$$

with derivatives evaluated at the minimiser of J^c. In the case that $\mathbf{\Gamma}_{\mathbf{post}}$ is not Gaussian, or estimates of VAF depend nonlinearly on the control parameters, (15) yields to an approximation of that posterior variance $\sigma^2(Q_T)$. We discuss in Sect. 6 the limitations of these assumptions.

225   We use the time-dependent adjoint capabilities of `FEniCS_ice` to find the sensitivities of VAF to the control parameters ($\frac{\partial Q_T}{\partial \overline{c}}$), for discrete values of T over 40 years. The inverse Hessian $\mathbf{\Gamma}_{\mathbf{post}}$ is itself approximated using a low-rank update to the prior covariance,

$$\mathbf{\Gamma}_{post} \approx \mathbf{\Gamma}_{prior} - \mathbf{C}_r \mathbf{\Lambda}_r (\mathbf{I}_r + \mathbf{\Lambda}_r)^{-1} \mathbf{C}^T_r, \tag{16}$$

where $\mathbf{\Lambda}_r$ and $\mathbf{C}_r$, respectively, represent the leading r eigenvalues and eigenvectors of the *prior-preconditioned misfit Hessian*:

230

$$\tilde{\mathbf{H}}_{mis} = \mathbf{\Gamma}_{prior} \left(\frac{\partial^2 J^c_{mis}}{\partial \overline{c}^2}\right). \tag{17}$$

Notably, this decomposition has the quality that the leading eigenvectors (those with the largest eigenvalues) are those *most informed* by the data. The leading eigenvectors define the components of the control parameters for which the observations change the estimated posterior uncertainty, relative to the prior uncertainty, by the largest factor. Thus the retained eigenvectors

235   of the Hessian, inform in which space of our mesh the model inversion gained the most information from the observations and the prior (see Fig. 9 and Sect. 5.1 for details).

Computationally the key ingredients to compute $\sigma^2(Q_T)$ are the ability to find a minimiser of J^c, the ability to compute the derivatives of VAF with respect to the control parameters ($\frac{\partial Q_T}{\partial \overline{c}}$), and the ability to compute Hessian information. The minimisation of J^c can be accelerated using gradient-based methods, if J^c can itself be differentiated with respect to $\overline{c}$. Here the

240   required first and second derivative information is obtained using `tlm_adjoint` (Maddison et al., 2019), with L-BFGS (Zhu

et al., 1997; Morales and Nocedal, 2011) used to perform the minimisation of J^c and SLEPc (Hernandez et al., 2005, 2007) used to calculate the eigendecomposition. Important points to make are that the eigenproblem requires only the *action* of the misfit Hessian, which would be computationally infeasible to form in full. Additionally, the Hessian takes account of the full nonlinearity of the ice-sheet model, in contrast with the Gauss-Newton approximation to the Hessian (Shapero et al., 2021). In Koziol et al. (2021) a comparison was made between the two in the context of an idealised problem, and results were minimal. For more details on the error propagation framework, see Koziol et al. (2021).

For all the experiments presented in this study, we calculate up to 10^4 (out of 10^5) eigenvalues and eigenvectors to ensure the convergence of $\sigma^2(Q_T)$ against the number of eigenvalues (see results in Sect. 5). The uncertainty of VAF at discrete times $\sigma^2(Q_T)$ is then found using (15), which can then be linearly interpolated to find a "trajectory" of uncertainty.

3   Study area, model domain and data sources

Our study area, shown in Fig. 1, covers part of the Amundsen Sea Embayment (ASE) in West Antarctica and includes three ice streams: Pope, Smith and Kohler Glaciers (PSK), as well as, the Dotson and Crosson ice shelves. PSK glaciers have exhibited some of the highest retreat rates in Antarctica throughout the satellite observing record, with their grounding lines receding over 30 km in recent decades (Scheuchl et al., 2016; Goldberg and Holland, 2022). Their catchment can potentially contribute up to 6 cm to the global mean sea level (Morlighem et al., 2020), double the global mean sea level contribution of the inventory of Earth's mountain glaciers (when excluding the Antarctic and Greenland periphery, Hock et al., 2023). A complete collapse of the ice shelves in this area would likely lead to accelerated mass loss from adjacent ice streams, including Thwaites Glacier (Goldberg and Holland, 2022). Previous modelling studies have shown that past and future retreat of these glaciers is strongly tied to ocean-forced melting, but that the method of calibration may affect projected rates of ice loss as well (Lilien et al., 2019; Goldberg and Holland, 2022). As such, and due to the vast quantity of data available for this region, we choose this area to test our model error framework in a realistic setting.

The domain is set up by generating an unstructured finite element mesh using time-averaged strain rates computed from satellite velocity observations (MEaSUREs v1.0 1996 - 2012, Rignot et al., 2014). Additionally, BedMachine Antarctica v2.0 (Morlighem et al., 2020) is used to provide geometry field information and the raster mask from which we define our boundary conditions; ice/ocean (calving) and ice/ice (edge of domain) boundaries in Fig. 1. The mesh generation occurs in two phases, first by generating an initial uniform-resolution mesh of 1000 elements with the mesh generator Gmsh (v.4.8.4 Geuzaine and Remacle, 2009) and second by refining that mesh with the calculated strain metric in the MMG software (v5.5.2 Dobrzynski, 2012). This generates a finer triangular mesh in the areas of the domain where high resolution is needed (e.g. close the calving front and in areas where velocities are higher in Fig. 2). The mesh resolution is highly heterogeneous and depends on the observed strain rates, with a minimum resolution of approx. 200 m. BedMachine Antarctica v2.0 (Morlighem et al., 2020) is also used to define the model's bed, ice thickness and surface elevation fields. The initial guess for β, β_{bgd}, is generated from the temperature dataset of Pattyn (2010). Based on coupled ice sheet-ocean modelling for the region (Goldberg and Holland, 2022), spatially uniform melt parameters of $M_{max} = 30$ m yr^{-1} and $z_{th} = 600$ m were chosen.

3.1 Velocity input data sources

275 In the last two decades, ice velocity mapping at the continental scale (Rignot et al., 2011; Gardner et al., 2019, 2018) has allowed major advances in the study of polar regions by providing complete observations of the complex flow pattern of ice sheets and glaciers (Mouginot et al., 2017). Much emphasis has been put on the fast processing of large data volumes and products with complete spatial coverage. However the metadata of such measurements is often highly simplified regarding the measurements precision and uncertainty (Altena et al., 2022). Moreover, the methods used to estimate errors in the observed

280 velocities tend to often produce errors that are unrealistically small (see Fig. 2 or Gardner et al., 2019). A quantification of the error estimation or dispersion (standard deviation) for each individual velocity measurement can be important for the inversion of unknown ice dynamic parameters (e.g. the basal friction coefficient α). Errors in the velocity data can propagate into derived results in a complex way, making model outcomes very sensitive to velocity noise and outliers (Altena et al., 2022). Therefore, we use two satellite velocity products to carry out inversion experiments and calibration uncertainty propagation; MEaSUREs

285 InSAR-Based Antarctica Ice Velocity Map (MEaSUREs v2.0. Rignot et al., 2017; Mouginot et al., 2017) and ITS_LIVE surface velocities (Gardner et al., 2019, 2018). To avoid large data gaps in the observations we focus on data acquired between 2013 and 2014 (see Fig 2). MEaSUREs provides surface velocities from July 2013 to July 2014 and ITS_LIVE from January to December 2014, thus we investigated the effect of the 6 month offset between both products, which turned out to be negligible (see Fig. A1 of Appendix A). In this section, we describe the acquisition sensors and standard deviation (STD) of each data-set,

290 as this is relevant to understanding the differences between each product, our experimental design and our results (see Sect. 4.2 and 5.2).

3.1.1 MEaSUREs v2.0

The grid spacing of this data set is 450 m. According to the product metadata (Rignot et al., 2017; Mouginot et al., 2017), the 2013–2014 year is a result of the data gathered by several instruments: RADARSAT-2 (CSA, 2012-2016), Sentinel-1

295 (Copernicus/ESA/EU, 2014-2016) and Landsat-8 (2013-2016). Landsat-8 is an optical sensor and it has mapped most of the ice sheet interior and the Antarctic coast, whereas RADARSAT-2 and Sentinel-1, are C-band synthetic aperture radar (SAR) instruments and have mostly captured velocities in the coast. Mouginot et al. (2017) notice that along the Antarctic coast, large differences ($\geq$50 m.yr^{-1}) between Landsat-8 and SAR based velocities are found, which can be due to stronger weather, ionospheric noise, and ongoing velocity changes.

3.1.2 ITS_LIVE

300

The grid spacing of this data set is 240 m. Surface velocities are derived only from optical sensor imagery (Landsat 4, 5, 7, and 8) using the auto-RIFT feature tracking processing chain described in Gardner et al. (2018). Data scarcity and/or low radiometric quality are significant limiting factors for many regions in the earlier product years. However annual coverage is nearly complete for the years following the Landsat 8 launch in 2013 (Gardner et al., 2019).

3.1.3 Observational Error model

The construction of $\mathbf{\Gamma}_{obs}$ based on reported errors deserves attention. Neither velocity product reports information on spatial error covariance, so $\mathbf{\Gamma}_{obs}$ is diagonal for both products. We interpret the likelihood PDF, $p(\breve{u}|\bar{c})$, as the density associated with the likelihood for a *single outcome* of an observation, as opposed to the distribution of the *average* outcome over an ensemble of observations. Essentially, we consider the standard deviation of observations, as opposed to the standard error of a sample mean.

The MEaSUREs product reports both error and standard deviation (STD), and we use the latter to construct $\mathbf{\Gamma}_{obs}$. The ITS_LIVE product does not report standard deviation, but gives the number (count) of measurements for each data point, and expresses error variance as an inverse weighted sum of individual measurement variances (Gardner et al., 2019). We therefore express standard deviation of each velocity component as

$$\text{STD}_{ITS} = count^{\frac{1}{2}} \times err_{ITS}. \tag{18}$$

Note this formula assumes uniform variance over all measurements contributing to a data point, which is not likely to be true.

In Koziol et al. (2021) it is shown for an idealised problem that the diagonality of $\mathbf{\Gamma}_{obs}$ leads to ever-decreasing posterior uncertainty as data density is increased. This is only an issue if errors correlate over the scale of separation of data points, but assessing error covariance is beyond the scope of this study. Still, this deficiency guides our investigation of the impacts of data density, described below in Section 4.

4 Experimental design and rationale

All inversion methods contain regularisation parameters which must be chosen (Barnes et al., 2021) and L-curve analysis (e.g. Fürst et al., 2015; Jay-Allemand et al., 2011; Gillet-Chaulet et al., 2012; Barnes et al., 2021) is a commonly used technique to make an informative guess regarding the value of these parameters – although there are alternative approaches (see Sect. 2.4.1 or Waddington et al., 2007; Habermann et al., 2013). Another common aspect of inversions in ice sheet modelling, due to data availability, is to use only one type of remotely-sensed ice velocity product for the calibration of the control parameters. In this section we study how these ongoing practices can impact the forecast of VAF and its uncertainty (see Sect. 4.1 and 4.2). Additionally, we assess the effect that data density (i.e. decreasing the number of observations) has on the inference (see Sect. 4.3). The experiments described in this section lay the groundwork for the model configurations used in Sect. 5.

4.1 L-curve analysis on the control parameters

L-curves are used to visualise the trade-off between the magnitude of the regularisation term (how much the control parameters should vary) and the quality of the fit (how well can we reproduce observations) and are generally created by plotting the regularisation terms against the misfit. However, there are several ways to construct an L-curve as it depends on the cost function definition. For consistency with control-method applications in the ice-sheet modelling literature (e.g. Barnes et al., 2021), we generate L-curves by varying the smoothing parameters γ and δ, rather than the variance and length scale arising

from a physical interpretation of the prior (Sect. 2.4.1). To shorten this analysis we show only L-curves for γ in Fig. 3. The L-curves presented in figures 3e and f are created by using ITS_LIVE surface velocities to find the misfit J^c_{mis} (7) and by plotting the regularisation terms against the cost function value J^c, as the regularisation terms γ_α and γ_β (11) vary over several orders of magnitude (10^4 to 10^{-4}).

340   In order to understand the effect that the strength of the regularisation (or prior strength) has on the control parameters, we show α and β spatial distributions computed using the extreme values of the L-curves (see figures 3a to d). If the prior strength is strong (i.e. a large γ_α) the inverted parameter field (in this case the sliding coefficient) is relatively smooth (see Fig. 3b) and J^c generally small. The L-curve for γ_α in Fig. 3e suggests a $\gamma_\alpha = 100.0$ as a reasonable trade-off between the cost function value and the regularization term. For γ_β this value is one order of magnitude smaller ($\gamma_\beta = 10.0$, see Fig. 3f). For δ_γ and δ_β

345   the L-curve analysis (not shown) suggests a value of 1×10^{-5}.[1] We used those values to conduct the rest of the experiments presented in Sect. 4.2 and 4.3.

4.2   Model output computed with different ice velocity observations

We use the regularisation parameters found in the previous section and run all stages of the error model framework (all methods in Sect. 2) twice using different satellite velocity products for each run; MEaSUREs and ITS_LIVE. We compare the observed

350   ice velocity from both products in Fig. 4b and find significant differences (≥ 100 m.yr^{-1}), especially at the ice margins. The assimilated states of FEniCS_ice reproduce these differences as shown in Fig. 4a, were we show modelled velocity differences between the two runs. Consequently; the inverted parameters from both runs are also different (see figures 4c and d) and thus are the projections of VAF (see Fig 5a). Differences in the output from both inversions are particularly large at the ice margins, and in the case of the ice stiffness parameter β, the largest differences are found at the Crosson and Dotson ice

355   shelves (see Fig 4d).

    Fig. 5b shows estimated posterior uncertainty of VAF loss after 40 years from our Hessian-based framework in our L-curve informed workflows. The uncertainty estimates are on the order of 10^{10} m^3, or 10 km^3. Meanwhile, the difference in VAF loss between the MEaSUREs- and ITS_LIVE-based trajectories is $\sim O(10^{11}$ m$^3)$, i.e. an order of magnitude larger. These results are seemingly at odds. In other words, our error propagation framework suggests a forecast uncertainty that is one

360   order of magnitude smaller than the variability in VAF found when using two different satellite velocity products. This leads to a contradiction, as it suggests that our observed variability in VAF loss is extremely unlikely. This contradiction could arise from one or more of the following: i) the regularisation suggested by the L-curve analysis is too strong, i.e. the prior is overly informative; ii) the observational error covariance matrix used for ITS_LIVE does not capture the true variability of the velocity field shown in Fig. 4b (see Sect. 3.1.3); and/or iii) there are too many data points informing our cost function. We address point

365   (iii) in the next section whereas points (i) and (ii) are addressed in sections 5.2 and 5.1.

---

[1]For each regularisation parameter units please refer to Table 1.

4.3 Effect of observational data subsampling

In this section we develop a metric to evaluate the quality of the model's inversions if we decrease the number of observations. In other words, we study the effect that different data densities have on the cost function performance. Similar techniques have been used to gather information on cross-correlations between two different sets of observations (e.g. covariance between winds speed observations in Desroziers et al., 2005), here we apply a similar diagnostic test to study the covariance between adjacent velocity observations from the same product. Additionally, Koziol et al. (2021) show that data density affects the posterior covariance of the control parameters for an idealised experiment, thus we use the results of this metric to test in Sect. 5.2, how data density affects the posterior uncertainty of VAF. The test has the following steps:

1. Produce several training data-sets of observed ice velocity by retaining different percentages of data points from a given set of observations (e.g. ITS_LIVE).

2. Use those training sets to compute inversions of α and β using our L-curve informed prior configuration.

3. Use the α and β results from step 2 to evaluate the cost function on velocity points that were not used to compute the inversions (i.e. using a validation data-set; observations from a different product such as MEaSUREs).

Based on both products metadata (see Sect. 3.1), we consider MEaSUREs and ITS_LIVE velocities to be two independent realisations of the state of the ice sheet at a given time and location. Hence, we use ITS_LIVE velocities for training and MEaSUREs for validation.

To construct the training sets, we divide the domain into cells of different sizes (different grid-spacing) and systematically drop observations by iterating over the x and y directions of the ITS_LIVE grid. We select sub-samples of the data by retaining corner and center observations from each cell per iteration – i.e. upper and middle cell-points. An example of a training data set is shown in Fig. 6a, where we retain only 1.6% of the velocity observations. To construct the validation set, we downscale MEaSUREs to the ITS_LIVE resolution and drop problematic data points (see figures 6d to f), i.e. locations where MEaSUREs and ITS_LIVE present velocity differences higher than $50 \ \mathrm{m.yr}^{-1}$ (as shown in Fig. 4b).

Results of the test (see Fig. 7) demonstrate that our framework provides robust inversions for the drag and stiffness parameters α and β. Additionally, these results reveal that the value of the cost function J^c does not change significantly even when we retain only 1.6% of the data, which suggests that a large amount of data points may be redundant when inverting for the control parameters, with potential implications for the error propagation of VAF (see the results from Sect. 5.2). However, observations are needed to inform the model in critical areas of the domain (i.e. at the grounding line or calving fronts) and if we drop observations in a random manner the performance of the cost function may decrease.

5 Results

Results from Sect. 4.2 show that if we use the prior strength suggested by the L-curves and the original ITS_LIVE velocity and standard deviation (STD), our framework underestimates the posterior uncertainty of VAF loss after 40 years by one order

of magnitude (see Fig. 5b). More precisely, we calculate a posterior uncertainty which suggests that the difference in projected VAF (estimated by using two velocity products that nominally observe the same physical properties to calibrate our model), is extremely unlikely (see Fig. 5a). To achieve a posterior uncertainty that reflects the same order of magnitude – i.e. $\sim \mathrm{O}(10^{11}$

400    $\mathrm{m}^3)$, we carry out two more experiments aiming to understand how the uncertainty in forecast of VAF is affected by the use of different strengths of prior (Sect. 5.1) and different versions of the ITS_LIVE data (Sect. 5.2). Additionally, we study the impact of using different sliding laws on the posterior uncertainty of VAF (Sect. 5.3).

5.1   Impact of using different prior strengths on the posterior uncertainty of VAF

We keep the same velocity input for all model configurations trialed in this section (i.e. retaining only 1.6% of the ITS_LIVE

405   data and adjusting the observations STD, as explained in Sect. 5.2), but vary the strengths of the prior. We experiment with the variance $\sigma^2_{c(0)}$ and auto-covariance length scale $l_{c(0)}$ of each control parameter instead of using priors suggested by L-curve analysis, as these definitions (see Sect. 2.4.1) have a more physical meaning. We calculate prior strengths using (13) and by making an informed guess on $\sigma^2_{c(0)}$ and $l_{c(0)}$ based on existing prior knowledge and physical concepts that define each control parameter.

410   From the literature (Pattyn, 2010; Khazendar et al., 2011; Still et al., 2022) we know that the spatial pattern of the ice stiffness is not uncorrelated. The advection of colder tributary glacier ice onto the ice shelf is well represented by the vast expanses of stiffer ice originating at the grounding line and extending downstream for tens of kilometers (see panels c and d of Fig. 3), whereas observed deformation patterns at the shear margins (Khazendar et al., 2011; Still et al., 2022) suggest the presence of weaker deformable ice, where the prominent formation of crevasses occur. Thus we keep an auto-covariance length scale $l_{\beta(0)}$

415   for β equal to 1 km in all prior configurations, as crevasses can be present within that length scale. In future studies this length scale could be set by conducting a detailed spatial statistical analysis of crevasse maps derived from remote sensing, which is beyond the scope of this study. For $\sigma^2_{\beta(0)}$ we trial variance values computed from the STD of the ice stiffness (B) initial guess (see details in Sect. 3).

For the sliding parameter we define auto-covariance length scales $l_{\alpha(0)}$, that are slightly larger than those assumed for β

420   (i.e. 2 to 3 km); as observations from airborne radar over the ice sheet (De Rydt et al., 2013) verify that for fast-flowing ice streams, the surface topography carries important information about the bed with wavelengths between 1 and 20 times the mean ice thickness (≥ 1 km) thus controlling basal sliding at similar scales. Additionally, model experiments described in Gudmundsson (2008) show that the SSA overestimates the effects of bed slipperiness perturbations on the surface profile for wavelengths less than about 5 to 10 times the mean ice thicknesses, the exact number depending on values of surface slope

425   and slip ratio. Variance values are less intuitive, thus we trial $\sigma^2_{\alpha(0)}$ values over several orders of magnitude in order to vary the prior strength imposed on α.

The resultant prior strengths are shown in Table 1 and the estimated posterior uncertainty of VAF for each prior configuration is shown in Fig.8a (solid lines), both are ordered from weak to strong. We find that a strong prior on the sliding parameter (such as the one suggested by the L-curve analysis), suppresses the error propagation from the satellite data onto projections of VAF.

430   Most of our prior experiments focus on α however, we also trial our error propagation framework changing the variance of β

(see Table 1). Compared to prior experiments run on α, changing β priors show little influence in the posterior uncertainty of VAF. We also test weaker priors for both parameters, but these experiments lead to non-stable solutions of the time-dependent SSA as the parameters present undesirable and nonphysical features (not shown).

Our low-rank approximation to the Hessian (16) makes the assumption that if additional eigenvectors were retained (i.e. if r were larger) the estimated posterior uncertainty $\sigma(Q_T)$ would not change considerably. To test this we examine the marginal change in $\sigma(Q_T)$ for each r (see Fig. 8c), which exhibits an approximately exponential decay with r. To estimate the effect of the low-rank approximation we assume that the decay rate holds up to $r = N$, where N is the full problem size, and that all neglected terms in (16) make a negative contribution to $\sigma(Q_T)$ – i.e. we estimate the "worst case" where every extra eigenvector/eigenvalue calculated decreases the uncertainty (See Appendix B for details of this estimate). The resulting estimated STD's of VAF for an infinite number of eigenvalues σ_{full}^{est} are shown in the captions of Fig. 8c, and indicate that for all prior strengths, even in the worst case, posterior uncertainties decrease by a small proportion and more importantly are not as small as those values seen in our L-curve investigations. We perform this same check in all remaining experiments and observe similar results (see Fig 8d and Fig. 10f).

Finally, the retained eigenvectors from the Hessian can be interpreted as those modes in the parameter space that change the approximated posterior uncertainty relative to the prior uncertainty. We show in Fig. 9 the leading eigenvectors (see panels a to f) for the prior configuration highlighted in Table 1. The dominant eigenvectors of the ice stiffness parameter field β (panels b, d, and f), suggest that we gained more information regarding this parameter at the grounding line. The eigenvector corresponding to the smallest eigenvalue (Fig. 9h), is increasingly more oscillatory (and thus informs at a smaller length scales in the parameter space) and is increasingly relatively less informed by the velocity observations; such patterns are also present in similar studies (e.g. Isaac et al., 2015).

5.2 Impact of using different ice velocity observations on the posterior uncertainty of VAF

Contrary to the previous section, here we keep the same prior strength for all the experiments but modify the velocity input. We modify the original ITS_LIVE data by decreasing the amount of observations and by adjusting the STD of each velocity component to match the following condition:

$$vx_{std} \rightarrow \max(vx_{std}, \quad \text{abs}(vx_I - vx_M)), \tag{19}$$

$$vy_{std} \rightarrow \max(vy_{std}, \quad \text{abs}(vy_I - vy_M)). \tag{20}$$

where vx_{std} and vy_{std} are standard deviations of velocity components, and I and M subscripts refer to ITS_LIVE and MEaSUREs, respectively. In other words, where the original uncertainty of the data is small, we replace those coordinates STD with the absolute difference between MEaSUREs and ITS_LIVE velocities at that same location. Figures 6b and c show this error adjustment for each velocity component. We generate three versions of the ITS_LIVE data by; i) retaining all data points but adjusting the STD, ii) retaining only 1.6% of the data (inline with the results from the observational data subsampling test) and adjusting the STD and iii) retaining only 1.6% of the data but keeping the original STD.

We run our error propagation framework using these three data sets (and the weak prior configuration highlighted in Table 1) and compute VAF posterior uncertainties as shown in Fig. 8b (see solid lines); uncertainties in this figure are plotted from high to low (from dark to light blue) and represent a 95% confidence interval. The effect of retaining only 1.6% of the observational data leads to a slight decrease in posterior VAF uncertainty – which is counterintuitive as one would expect fewer observations to give a larger posterior calibration uncertainty. However, the VAF uncertainty (15) is derived from both calibration uncertainty and VAF sensitivity. The latter differs between the experiments, as can be seen from the projection of prior uncertainty on to VAF sensitivity (blue dashed lines). The adjustment of observational STD increases VAF uncertainty at approximately 5 to 15 years, but has less impact after. Importantly however, the overall impact on the posterior uncertainty of VAF loss after 40 years is small relative to the effect of changing the prior strength (see solid line differences between panels a and b of Fig. 8).

5.3   Impact of using different sliding laws on the posterior uncertainty of VAF

In previous configurations we use the Weertman-Budd sliding law, but due to the sensitivity of the time-dependent ice sheet model to the choice of sliding law (Brondex et al., 2019; Kazmierczak et al., 2022; Barnes and Gudmundsson, 2022) we trial our error propagation framework using the Cornford law and compare qualitative differences from both laws in Fig 10. We use the same velocity constraints (i.e. same as in Sect. 5.2) and consider only a single prior distribution (the highlighted parameters in Table 1) but modifying $\sigma_{\alpha(0)}^2$ as basal stress does not scale with effective stress in the interior.

Using the inverse of the low rank update approximation for the cost function Hessian (Γ_{post}) we can estimate the posterior standard deviation (STD) of α and β. We divide the mesh into 'patches' of approximately 1 km in diameter, and for each patch we compute the mean of each control parameter. We treat this mean as a new quantity of interest (QoI) and compute its STD via the same framework as projections of VAF loss (see Sect. 2.5). Essentially, we visualise in panels a to c of Fig. 10 the posterior of a 'local average' of α and β.

For both sliding laws the sliding parameter α is more uncertain close to the grounding line and at the Bear Peninsula (see Fig. 10 panels a and b) where uncertainties from the ITS_LIVE product are higher (see Fig 2 panels b and c). The large uncertainty just at the grounding line in the Cornford results is due to the insensitivity of basal stress to α when the ice is near floatation (see sensitivity analysis below). For the ice stiffness parameter β the most uncertain areas of our domain are the grounding lines of the PSK glaciers and the Crosson ice shelf (see panel c of Fig. 10) – these are the areas where the speed from the two satellite products show significant differences (see Fig. 4b).

For both sliding laws, VAF uncertainty reaches a similar order of magnitude $\sim O(10^{11})$ m^3 at year 40 (Fig 10e). However, a quantitative comparison is somewhat misleading, as the impact of prior strength is is not investigated for the Cornford law. There are qualitative differences however: the posterior uncertainty of VAF for each sliding law saturates at a different rate, with the posterior uncertainty of the Cornford law configuration growing at a faster rate after year 10. We compare sensitivity maps of the model's VAF estimates to the basal friction coefficient α^2 at year 10 and 40, normalised to year 40 sensitivities for the respective experiment (see Fig. 11). VAF sensitivities at year 10 computed using the Weertman-Budd law have a higher sensitivity to the basal friction coefficient relative to those computed using the Cornford law – particularly at the grounding line of Kohler Glacier (see panels a and c of Fig. 11). Additionally, Fig. 11 shows that at year 40 both sliding laws have similar

sensitivities. In this section we only show sensitivity maps for α^2, sensitivities to the ice stiffness are shown in Fig. A2 of Appendix A.

6   Discussion

500   The efficiency of our error propagation framework allows us to explore how different prior strengths, velocity inputs and sliding laws affect the uncertainty of VAF projections. We find that by choosing different satellite ice velocity products (that nominally observed the same physical properties to calibrate `FEniCS_ice`) our model leads to different estimates of VAF loss after 40 years (see Sect. 4.2). This effect may be less important for ice streams strongly coupled to ocean forcing (Lilien et al., 2019; Goldberg and Holland, 2022), but could be more influential for unstable margins (Joughin et al., 2014). We use such difference
505   to quantify (in a robust way) the variance that projections of VAF are expected to have after 40 years and identify prior strengths that can reproduce this expected variability (see Sect. 5.1). We demonstrate that if we use prior strengths suggested by the L-curve analysis, as is typically done in ice-sheet calibration studies, our uncertainty quantification is not able to reproduce this level of variability. The regularisation suggested by the L-curves is too strong and thus suppresses the error propagation from the satellite data into the QoI, resulting in VAF projections with quantified uncertainties that are smaller than those suggested
510   by our 2-member "sample" of observed velocity fields. Additionally, our analysis suggests that the error given by the velocity data cannot fully explain the variability in ice velocities observed in Fig. 4b and that large amounts of data points may be redundant, with implications for the error propagation of the QoI.

Our ice sheet flow model described in Sect. 2 can be thought of as a (nonlinear) *mapping* from a set of input fields, which might be unobservable or poorly known (α and β fields), to a set of output fields, which might correspond to observable
515   quantities (e.g. satellite surface velocity observations). In `FEniCS_ice`, the parameter-to-observable map $\breve{f}$ is a composition of two functions: the solution of the SSA equations (see Sect. 3 of Koziol et al., 2021, for details) and the misfit term J^c_{mis} (8).

Our error propagation framework considers the ice sheet inverse problem as a linearised inverse problem; by linearisation we mean that $\breve{f}$ is linearised about the MAP point. Thus the framework relies on a number of key assumptions related to this and other issues:

520   1. (i) The observational errors and prior distributions are defined as Gaussian distributions, (ii) the parameter-to-observable map $\breve{f}$ is linear (or close to linear), and (iii) the Quantity of interest (i.e. VAF) at a given time depends linearly (or nearly linearly) on the control parameters – in other words, the parameter-to-QoI map is close to linear.

2. The difference between velocities predicted by the model and the observations is due only to measurement errors (we assume zero model error see Sect. 2.4 – or more precisely consider conditional posterior information given the model).

525   3. The observational error covariance matrix is diagonal, i.e. errors in observations do not correlate spatially.

4. The posterior covariance of the control parameters Γ_{post} is fully sampled with the number of eigenvectors and eigenvalues that we retain from the Hessian.

17

Note (1).(i,ii) above implies a Gaussian posterior. We already test (4) in Sect. 5.1 and address (1)-(3) in the following subsections.

6.1 Linear dependence of parameter-to-observable and Quantity of Interest maps with respect to the control parameters.

`FEniCS_ice` computes a second-order approximation to the posterior covariance of the control parameters Γ_{post} (via the eigendecomposition of the cost function Hessian evaluated at the MAP point, see 2.5) and propagates forward the associated calibration uncertainty in time-dependent estimates of VAF loss (our QoI). The posterior PDF of $\bar{c}$ is not guaranteed to be Gaussian due to the nonlinearity of the Stokes equations that describe $\breve{f}$. Furthermore, the propagation step (Eq. 15) is based on a linear transformation of a Gaussian random variable, and assumes that $Q_T(\bar{c})$, the parameter-to-QoI map, is well-described by linear sensitivities.

Petra et al. (2014) test the Gaussianity of the parameter-to-observable map by sampling from the posterior PDF of the hidden field $\bar{c}$ via two different Markov chain Monte Carlo (MCMC) sampling methods (Tierney, 1994); the Newton MCMC method and a new stochastic Newton method with a MAP-based Hessian. They solve a two-dimensional flow-line ice sheet inverse problem with a moderate number of parameters (~ 100) and conclude that the most non-Gaussian behavior for $\breve{f}$ occurs in the directions of the domain with the largest variance in the observations, thus $\breve{f}$ deviates from a linear approximation in those directions (Petra et al., 2014; Isaac et al., 2015). Nevertheless, in the directions where the variance is large, the posterior covariance of the control parameters will be significantly influenced by the prior – which is also defined as Gaussian thus it is expected that the Gaussianity assumption will hold and $\breve{f}$ will be weakly linear in such directions. Therefore, a Hessian-based approximation (such as the one describe in Sect. 2.5) to the posterior covariance of the parameters may be appropriate despite the nonlinearity of $\breve{f}$ (Petra et al., 2014).

Koziol et al. (2021) test the linearity of the `FEniCS_ice` parameter-to-QoI map for an idealised ice sheet flow problem (Pattyn et al., 2008) through a simple Monte Carlo sampling of the posterior PDF of $\bar{c}$. The study finds strong agreement with the linearly propagated posterior covariance when there is a moderately strong prior, but slightly poorer agreement with a weak prior.

Unfortunately, due to the size of our parameter space, testing the Gaussianity of the posterior PDF of $\bar{c}$ is beyond the scope of our study. Similarly, sampling the posterior PDF to validate the propagation of calibration uncertainty to the QoI as in Koziol et al. (2021) would be intractable for our more realistic setting. Instead, we develop a simple test to check the linearity of the parameter-to-QoI map and how this linearity is affected when we impose different strengths of prior. We test the linearity of the parameter-to-QoI mapping by using data from Sect. 4.2 to compute the following dot products:

$$\frac{\partial Q_T^I}{\partial \alpha_I} \cdot (\alpha_I - \alpha_M), \tag{21}$$

$$\frac{\partial Q_T^I}{\partial \beta_I} \cdot (\beta_I - \beta_M). \tag{22}$$

Here I and M indicate model output computed by either using ITS_LIVE or MEaSUREs velocities. We visualise the
560 linearity of the VAF operator by plotting each dot product together with the absolute difference between VAF trajectories computed using ITS_LIVE and MEaSUREs. Additionally, we repeat this test for a stronger prior configuration by imposing a strong regularisation on β (stronger than the one suggested by the L-curve analysis, $\gamma_\beta = 100.0$).

Results from both tests are shown in Fig 12 and verify that VAF estimates over time are highly dependent on α and that the linearity of the parameter-to-QoI map depends on the strength of the regularisation (as in Fig. 12b and in Koziol et al., 2021).
565 The main objective of this study is to propagate calibration uncertainty into projections of VAF loss. We find that in order to do so, we must impose a weaker prior on the control parameters than widely-used methods (i.e. L-curve analysis) would suggest. But as shown above, in doing so we might need to compromise on the linearity of the parameter-to-QoI map. Moreover, as shown in Koziol et al. (2021), a weaker prior means a weaker spectral decay of the prior-preconditioned Hessian spectrum, requiring to retain more of its eigenvectors (see also Sect. 5.1).

570 In other words, to avoid the prior probability from overwhelming the likelihood in our Bayesian inversion, we are required to examine a regime where we compromise the linearity of the time-dependent model in certain areas of the domain. Still, we expect that the framework can provide an "order-of-magnitude" estimate of the contribution of calibration uncertainty to QoI uncertainty. Although not previously applied to a problem as large as the present study, Stochastic Newton MCMC (Martin et al., 2012; Petra et al., 2014), which does not rely on a Gaussian assumption, may provide a more robust estimate
575 in such regimes. Importantly, to be tractable this method requires a reasonable estimate of the posterior density (the "proposal density") – and such an estimate can be provided using the low-rank Hessian approximation generated within our framework. Thus Stochastic Newton MCMC may be a viable approach for non-Gaussian uncertainty quantification in future studies.

6.2 Qualitative inspection of the model's structural and forcing uncertainty

We only quantify calibration (parametric) uncertainty in projections of marine ice sheet loss. We do not quantify structural
580 or model uncertainty, i.e. errors that arise from the discretisation of the inverse problem (Barnes et al., 2021) or from the formulation of the model and its ability to represent the physics of the system (Hill et al., 2021). In Sect. 5.3 we trial our error propagation framework with different sliding laws and examine the implications for projections of VAF loss (see Fig. 10); though quantifying the likelihood of various sliding-law formulations is beyond the scope of our study.

In this section we look at uncertainty due to the use of different physics and discretisation to solve the ice-sheet momentum
585 balance. We do this by using a second ice sheet model: the STREAMICE module of MITgcm (Goldberg and Heimbach, 2013), which solves a depth-integrated balance that accounts for vertical shearing (absent from the shallow ice-stream approximation; Goldberg, 2011). STREAMICE solves the momentum balance on a regular rectangular grid, a distinct discretisation from `FEniCS_ice`. With a uniform 500 m grid, we simulate with STREAMICE an instantaneous velocity field (without time evolution), using the inverted parameter fields of `FEniCS_ice` (interpolated to the STREAMICE grid) and the same geometry
590 and boundary conditions described in Sect. 3. The particular fields of α and β are from our L-curve analysis (Sect. 4.1). We compare both models' surface velocities and find differences on the order of 100-200 myr^{-1}, particularly in the fastest-flowing ice areas and on the ice shelves (see Fig. A3). `FEniCS_ice` and STREAMICE have different approximations to Stokes flow,

employ different treatments of the grounding line in their equations, and have very different resolution, which may lead to this disagreement. Barnes et al. (2021) find similar results, when two other adaptive mesh, finite-element SSA models are

595   compared to STREAMICE through the same diagnostic experiment. The authors show that these diagnostic calculations are not indicative of the performance of the models in time-dependent simulations (see Fig. 6 of Barnes et al., 2021, where all models reach similar projections of VAF).

We emphasise this comparison cannot quantify structural uncertainty, but can inform us (qualitatively) of the effects of implementing different discretisations and grounding line formulations in the model numerics.

6.3   Relevance of calibration uncertainty versus structural and forcing uncertainty

600

As previously mentioned, structural uncertainty is neglected in our study and we use a very simple ocean forcing parameterisation, for which uncertainties are not considered. We make clear that our aim is to quantify calibration uncertainty alone; however, it is only worth doing so if the contribution of calibration uncertainty to forecast uncertainty is non-negligible, and/or the framework represents nontrivial steps toward incorporating these other sources of uncertainty. Regarding the former, the

605   existing literature provides some clues as to whether calibration uncertainty is important. Goldberg and Holland (2022) carry out coupled ice sheet-ocean modelling experiments for the PSK glaciers region, and show that the type of calibration of ice model parameters (i.e. whether fit to observed thinning is accounted for) strongly determines ice loss over 20-30 years; beyond this point, ice loss depends on far-field ocean conditions. For other catchments, this "crossover time" could be shorter, or longer – meaning that uncertainty in calibration could inform projection uncertainty on the multidecadal scale before it is overtaken

610   by climate uncertainty. The short-term persistence of calibration errors is echoed in other types of cryospheric modelling: Aschwanden and Brinkerhoff (2022) showed that the introduction of satellite-based information strongly reduced uncertainty in short-term projections of Greenland ice loss, but that this relative information gain was greatly reduced by 2100, particularly under strong climate forcing scenarios. Still, calibration uncertainties should not be dismissed even if they are overwhelmed by climate forcing on long time scales: there are strong reasons why short-term (multidecadal) projections of ice loss are key

615   for planning and mitigation (Bassis, 2022).

Moreover, our framework of estimating calibration uncertainty can easily be expanded to account for forcing uncertainty. Provided that forcing uncertainty is independent of parameter uncertainty, the contribution of forcing to projection uncertainty is additive, and can be found using an expression similar to Eq. 15. Importantly, such a calculation is independent from the estimation of posterior parameter uncertainty through eigendecomposition of the Hessian – which is by far the most costly

620   component. This is not true of model uncertainty: our likelihood PDF $p(\breve{u}|\bar{c})$, which gives the probability of observable velocity conditioned on parameters and the model, and hence neglects model uncertainty. A potential way to incorporate model uncertainty – once it is quantified – is to adjust the observational error covariance used in the likelihood. A similar approach has been used in the Bayesian Error Approximation method of Babaniyi et al. (2021).

6.4 Accuracy of observational error model

625   We draw the tentative conclusion that, for our study area, the choice of prior distribution informed by L-curve analysis is overly informative and underestimates calibration uncertainty. This is based on the fact that, with such a prior, the posterior VAF is an order of magnitude smaller than the variability in VAF between two widely used velocity products as constraining data. Essentially, the two products are treated as a 2-member sample from a distribution describing the true surface velocities. While this is a very small sample size, our assessment makes the assumption that (i) the posterior VAF distribution is Gaussian (which

630 is explored above) and (ii) the two members are likely sampling outcomes under our observational error model – and therefore that the variation of $\sim O(10^{11} \text{ m}^3)$ is not a statistical outlier – thus our Hessian-based assessment must be too small.

    A further assumption in our assessment is that our observational error model is accurate. As described in Sect. 3.1.3, we use reported errors and standard deviations as diagonal terms in $\mathbf{\Gamma_{obs}}$ and assume zero spatial error covariance. Error magnitudes may be underestimated – although we somewhat account for this by adjusting observational errors based on differences between

635 the products (Sect. 5.2). Additionally, not accounting for spatial error correlation could underestimate calibration uncertainty, as shown in the idealised experiments of Koziol et al. (2021). It is possible that improved assessments of spatial observational error covariance may be needed to accurately quantify calibration uncertainty when calibrating ice-sheet models with satellite-based data. Such approaches have been used in weather data assimilation Tabeart et al. (2020).

7 Conclusions

640 This study set out to apply `FEniCS_ice` error propagation framework to a realistic setting in West Antarctica (which includes three ice streams: Pope, Smith and Kohler Glaciers) and infer from satellite velocity observations two important unknown parameters in ice dynamics and its uncertainties; the basal sliding friction coefficient (α) and the rheological parameter for describing ice stiffness (β). As in many other ice-sheet modelling studies we use a control method to calibrate grid-scale flow parameters. However, our framework augments the control method with a Hessian-based Bayesian inference approach, which

645 characterises the posterior covariance of the inverted parameters. We project calibration uncertainty forward in time and onto projections of volume above flotation (VAF).

    We find that by choosing different satellite ice velocity products (that nominally observed the same physical properties to calibrate `FEniCS_ice`) our model leads to different estimates of VAF after 40 years or to different projections of sea-level rise contribution. We use this difference in model output as an order-of-magnitude estimate of the variance that projections of

650 VAF should have after 40 years and identify prior strengths that can reproduce that variability. We demonstrate that if we use prior strengths suggested by L-curves, as is typically done in ice-sheet calibration studies, our uncertainty quantification is not able to reproduce that same variability. The regularisation suggested by the L-curves is too strong and thus suppresses the error propagation from the satellite data into VAF projections with quantified uncertainties that are smaller than those suggested by our 2-member "sample" of observed velocity fields. Additionally, our experiments suggest that large amounts of data points

655 may be redundant, with implications for the error propagation of VAF.

We inspect qualitatively one aspect of structural uncertainty by trialing our error propagation framework with two different sliding laws (Weertman-Budd and Cornford laws). The posterior uncertainty of VAF evolves differently for the two parameterisations, with the Weertman-Budd uncertainty saturating relatively quickly while that of Cornford steadily increases. This may be due to differing patterns of sensitivity of VAF to the sliding parameters, particularly near the grounding line.

660 Finally, our framework alone does not fully quantify sea-level rise forecast uncertainty, but represents an important step. Further improvements to our method could be to (i) quantify calibration uncertainty through Stochastic Newton Markov chain Monte Carlo (MCMC) using our Hessian eigendecomposition as the proposal density, (ii) take into account model error in the likelihood probability density function and (iii) take into account forcing uncertainty in the error propagation framework.

Code availability. The version of `tlm_adjoint` used in this manuscript is available in a permanent DOI repository (https://zenodo.
665 org/badge/latestdoi/452296839). The `FEniCS_ice` software together with the application of the code to a real domain are coded in the Python language and licensed under the GPL-3.0 license. The latest version of the `FEniCS_ice` code is available on Github (https://github.com/EdiGlacUQ/fenics_ice), the documentation website of the model is under construction but a user guide is provided (https://github.com/EdiGlacUQ/fenics_ice/tree/main/user_guide). The code used to generate all figures and analyses of this study is available in a permanent DOI repository (https://doi.org/10.5281/zenodo.7615259) and on Github (https://github.com/bearecinos/smith_glacier) as well as
670 the `FEniCS_ice` version used for this study (https://zenodo.org/badge/latestdoi/101511241). We have constructed a documentation website for the application of the model to the PSK Glaciers domain (https://github.com/bearecinos/smith_glacier/wiki) where we explain in detail the installation of the code, the preparation of input data and how to run and visualise the experiments presented in this study.

Data availability. The output data of the model is available in the following permanent DOI repository https://doi.org/10.5281/zenodo.7612243. Information about how to read and plot the data can be found in the Smith repository wiki, see: https://github.com/bearecinos/
675 smith_glacier/wiki.

Appendix A: Extra figures

A1 ITS_LIVE 6 month offset speed change

To study the 6 month offset between ITS_LIVE and MEaSUREs velocities – i.e. from July 2014 to December 2014, we subtract from the 2018 ITS_LIVE velocity mosaic, data acquired in 2014 and divide this by 8 in order get the monthly changes (see
680 Fig. A1a). The effect of the 6 month offset between both products is negligible compared to the difference observed in Fig. 4b and to the speed ratio shown in Fig. A1b. However, there are significant differences (over a small area) at the calving front of the Crosson Ice shelf (see Fig. A1a).

The Cryosphere
Discussions

A2   Sensitivity of VAF to the ice stiffness (B)

We compare sensitivity maps of the model's VAF estimates to the ice stiffness B (or β^2) at year 10 and 40 (see Fig. A2). VAF

685   projections are more sensitive to the ice stiffness at the grounding line of the PSK Glaciers and at the Crosson Ice shelf. In future studies, these sensitivity maps and the ice stiffness spatial distribution could be correlated to detailed spatial maps of crevasses in the area.

Appendix B: Convergence of the estimated posterior uncertainty $\sigma(Q_T)$ with the number of eigenvectors

For each successive eigenvalue-eigenvector pair (λ_r, C_r) we construct the low-rank approximation to the posterior covariance

690   using Eq. 16, and find the associated approximation to $\sigma(Q_T)$ by projecting the estimated covariance on to the QoI (Eq. 15). We refer to this iterate here as σ_r – the posterior QoI uncertainty using the leading r eigenvectors – and to the difference $\sigma_r - \sigma_{r-1}$ as $\Delta\sigma_r$.

We observe that, for sufficiently large r, the absolute change with r can be represented reasonably well by an exponential decay i.e.

695    $$|\Delta\sigma_r| = d_0 b^r \tag{B1}$$

for some b smaller than 1 (Fig. 8(c,d)). Assuming this to hold as r becomes large, we can estimate a lower bound for $\sigma(Q_T) = \sigma_N$ (where N is the parameter dimension) with a geometric sum. Specifically, we find the d_0 and b that best fit $|\Delta\sigma_r|$ for $r_{th} \leq r \leq r_M$, where r_{th} is inferred from the decay of $|\Delta\sigma_r|$ and r_N is the number of eigenpairs retained (in our case 10^4). The relationship given by Eq. B1 then implies for $M > N$

700    $$\sigma_M = \sigma_N + \sum_{r=N+1}^{M-1} \Delta\sigma_r$$
$$\leq \sigma_N - d_0 b^N \left(\frac{1 - b^{M-N}}{1 - b} \right)$$
$$< \sigma_N - \frac{d_0 b^N}{1 - b}. \tag{B2}$$

In Sect. 5 we use this result (with $r_{th} = 3000$) to estimate a lower bound for the posterior uncertainty of Q_T without low-rank approximation. The calculation is done at the final time i.e. for $T = 40$ a only. We emphasise that this calculation is purely

705   heuristic, and we are unaware of a theoretical lower bound for σ_M. Due to the tendency of the shallow-shelf approximation to filter high spatial frequencies in basal parameters (Gudmundsson, 2008), it is unlikely that $|\Delta\sigma_r|$ will decay more slowly than predicted by Eq. B1, and it may even decay more quickly. However due to the large size of the parameter space (10^5) it is not tractable to find the full spectrum, so the estimate is not testable for this problem.

Finally, other studies use eigenvalue magnitude as a criterion for truncating the spectrum (e.g., Isaac et al., 2015). More

710   specifically, eigenvectors are retained up to an index r such that $\frac{\lambda_r}{\lambda_r+1} \ll 1$. We note this constraint alone does not ensure

that the contribution to QoI uncertainty arising from the truncated part of the spectrum is negligible, even if the marginal contribution associated with each individual eigenpair is small.

References

725   Alnæs, M., Blechta, J., Hake, J., Johansson, A., Kehlet, B., Logg, A., Richardson, C., Ring, J., Rognes, M. E., and Wells, G. N.: The FEniCS Project Version 1.5, Archive of Numerical Software, 3, 9–23, 2015.

Altena, B., Kääb, A., and Wouters, B.: Correlation dispersion as a measure to better estimate uncertainty in remotely sensed glacier displacements, The Cryosphere, 16, 2285–2300, https://doi.org/10.5194/tc-16-2285-2022, 2022.

Arthern, R. J., Winebrenner, D. P., and Vaughan, D. G.: Antarctic snow accumulation mapped using polarization of 4.3-cm wavelength
730   microwave emission, Journal of Geophysical Research: Atmospheres, 111, https://doi.org/https://doi.org/10.1029/2004JD005667, 2006.

Asay-Davis, X. S., Cornford, S. L., Durand, G., Galton-Fenzi, B. K., Gladstone, R. M., Gudmundsson, G. H., Hattermann, T., Holland, D. M., Holland, D., Holland, P. R., Martin, D. F., Mathiot, P., Pattyn, F., and Seroussi, H.: Experimental design for three interrelated marine ice sheet and ocean model intercomparison projects: MISMIP v. 3 (MISMIP+), ISOMIP v. 2 (ISOMIP+) and MISOMIP v. 1 (MISOMIP1), Geoscientific Model Development, 9, 2471–2497, https://doi.org/10.5194/gmd-9-2471-2016, 2016.

735   Aschwanden, A. and Brinkerhoff, D.: Calibrated Mass Loss Predictions for the Greenland Ice Sheet, Geophysical Research Letters, 49, e2022GL099 058, https://doi.org/10.1029/2022GL099058, 2022.

Aschwanden, A., Bartholomaus, T. C., Brinkerhoff, D. J., and Truffer, M.: Brief communication: A roadmap towards credible projections of ice sheet contribution to sea level, The Cryosphere, 15, 5705–5715, https://doi.org/10.5194/tc-15-5705-2021, 2021.

Babaniyi, O., Nicholson, R., Villa, U., and Petra, N.: Inferring the basal sliding coefficient field for the Stokes ice sheet model under
740   rheological uncertainty, The Cryosphere, 15, 1731–1750, https://doi.org/10.5194/tc-15-1731-2021, 2021.

Barnes, J. M. and Gudmundsson, G. H.: The predictive power of ice sheet models and the regional sensitivity of ice loss to basal sliding parameterisations: a case study of Pine Island and Thwaites glaciers, West Antarctica, The Cryosphere, 16, 4291–4304, https://doi.org/10.5194/tc-16-4291-2022, 2022.

Barnes, J. M., Dias dos Santos, T., Goldberg, D., Gudmundsson, G. H., Morlighem, M., and De Rydt, J.: The transferability of adjoint
745   inversion products between different ice flow models, The Cryosphere, 15, 1975–2000, https://doi.org/10.5194/tc-15-1975-2021, 2021.

Bassis, J.: Quit Worrying About Uncertainty in Sea Level Projections, Eos, Transactions American Geophysical Union, 102, https://doi.org/https://doi.org/10.1029/2021EO210632, 2022.

Brinkerhoff, D., Aschwanden, A., and Fahnestock, M.: Constraining subglacial processes from surface velocity observations using surrogate-based Bayesian inference, Journal of Glaciology, 67, 385–403, https://doi.org/10.1017/jog.2020.112, 2021.

750   Brinkerhoff, D. J.: Variational inference at glacier scale, Journal of Computational Physics, 459, 111 095, https://doi.org/10.1016/j.jcp.2022.111095, 2022.

Brondex, J., Gillet-Chaulet, F., and Gagliardini, O.: Sensitivity of centennial mass loss projections of the Amundsen basin to the friction law, The Cryosphere, 13, 177–195, https://doi.org/10.5194/tc-13-177-2019, 2019.

Budd, W. F. and Jenssen, D.: Numerical Modelling of the Large-Scale Basal Water Flux under the West Antarctic Ice Sheet, in: Dynamics of
755   the West Antarctic Ice Sheet, edited by Van der Veen, C. J. and Oerlemans, J., pp. 293–320, Springer Netherlands, Dordrecht, 1987.

Budd, W. F., Keage, P. L., and Blundy, N. A.: Empirical Studies of Ice Sliding, Journal of Glaciology, 23, 157–170, https://doi.org/10.3189/S0022143000029804, 1979.

Bui-Thanh, T., Ghattas, O., Martin, J., and Stadler, G.: A computational framework for infinite-dimensional Bayesian inverse problems. Part I: The linearized case, with application to global seismic inversion, arXiv e-prints, arXiv:1308.1313, 2013.

760 Cornford, S. L., Martin, D. F., Payne, A. J., Ng, E. G., Le Brocq, A. M., Gladstone, R. M., Edwards, T. L., Shannon, S. R., Agosta, C., van den Broeke, M. R., Hellmer, H. H., Krinner, G., Ligtenberg, S. R. M., Timmermann, R., and Vaughan, D. G.: Century-scale simulations of the response of the West Antarctic Ice Sheet to a warming climate, The Cryosphere, 9, 1579–1600, https://doi.org/10.5194/tc-9-1579-2015, 2015.

Cornford, S. L., Seroussi, H., Asay-Davis, X. S., Gudmundsson, G. H., Arthern, R., Borstad, C., Christmann, J., Dias dos Santos, T., Feld-
765 mann, J., Goldberg, D., Hoffman, M. J., Humbert, A., Kleiner, T., Leguy, G., Lipscomb, W. H., Merino, N., Durand, G., Morlighem, M., Pollard, D., Rückamp, M., Williams, C. R., and Yu, H.: Results of the third Marine Ice Sheet Model Intercomparison Project (MISMIP+), The Cryosphere, 14, 2283–2301, https://doi.org/10.5194/tc-14-2283-2020, 2020.

Cuffey, K. and Paterson, W.: The Physics of Glaciers, 4th Edition, Academic Press, 2010.

De Rydt, J., Gudmundsson, G. H., Corr, H. F. J., and Christoffersen, P.: Surface undulations of Antarctic ice streams tightly controlled by
770 bedrock topography, The Cryosphere, 7, 407–417, https://doi.org/10.5194/tc-7-407-2013, 2013.

DeConto, R. M. and Pollard, D.: Contribution of Antarctica to past and future sea-level rise, Nature, 531, 591–597, https://doi.org/10.1038/nature17145, 2016.

Desroziers, G., Berre, L., Chapnik, B., and Poli, P.: Diagnosis of observation, background and analysis-error statistics in observation space, Quarterly Journal of the Royal Meteorological Society, 131, 3385–3396, https://doi.org/https://doi.org/10.1256/qj.05.108, 2005.

775 Dobrzynski, C.: MMG3D: User Guide, Technical Report RT-0422, INRIA, https://hal.inria.fr/hal-00681813, 2012.

Dutrieux, P., De Rydt, J., Jenkins, A., Holland, P., Ha, H., Lee, S., Steig, E., Ding, Q., Abrahamsen, E., and Schröder, M.: Strong Sensitivity of Pine Island Ice-Shelf Melting to Climatic Variability, Science, 343, 174–178, https://doi.org/10.1126/science.1244341, 2014.

Favier, L., Durand, G., Cornford, S. L., Gudmundsson, G. H., Gagliardini, O., Gillet-Chaulet, F., Zwinger, T., Payne, A., and Brocq, A. M. L.: Retreat of Pine Island Glacier controlled by marine ice-sheet instability, Nature Climate Change, 4, 117–121,
780 https://doi.org/10.1038/nclimate2094, 2014.

Fürst, J. J., Durand, G., Gillet-Chaulet, F., Merino, N., Tavard, L., Mouginot, J., Gourmelen, N., and Gagliardini, O.: Assimilation of Antarctic velocity observations provides evidence for uncharted pinning points, The Cryosphere, 9, 1427–1443, https://doi.org/10.5194/tc-9-1427-2015, 2015.

Gardner, A. S., Moholdt, G., Scambos, T., Fahnstock, M., Ligtenberg, S., van den Broeke, M., and Nilsson, J.: Increased West Antarctic and
785 unchanged East Antarctic ice discharge over the last 7 years, The Cryosphere, 12, 521–547, https://doi.org/10.5194/tc-12-521-2018, 2018.

Gardner, A. S., Fahnestock, M. A., and Scambos, T. A.: ITS_LIVE Regional Glacier and Ice Sheet Surface Velocities. Data archived at National Snow and Ice Data Center, https://doi.org/10.5067/6II6VW8LLWJ7, accessed: February 13, 2023., 2019.

Geuzaine, C. and Remacle, J.-F.: Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities, International Journal for Numerical Methods in Engineering, 79, 1309–1331, https://doi.org/https://doi.org/10.1002/nme.2579, 2009.

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

Glen, J. W. and Perutz, M. F.: The creep of polycrystalline ice, Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, 228, 519–538, https://doi.org/10.1098/rspa.1955.0066, 1955.

795 Goldberg, D. N.: A variationally derived, depth-integrated approximation to a higher-order glaciological flow model, Journal of Glaciology, 57, 157–170, https://doi.org/10.3189/002214311795306763, 2011.

Goldberg, D. N. and Heimbach, P.: Parameter and state estimation with a time-dependent adjoint marine ice sheet model, The Cryosphere, 7, 1659–1678, https://doi.org/10.5194/tc-7-1659-2013, 2013.

Goldberg, D. N. and Holland, P. R.: The Relative Impacts of Initialization and Climate Forcing in Coupled Ice Sheet-Ocean Modeling: Application to Pope, Smith, and Kohler Glaciers, Journal of Geophysical Research: Earth Surface, 127, e2021JF006 570, https://doi.org/https://doi.org/10.1029/2021JF006570, e2021JF006570 2021JF006570, 2022.

Gudmundsson, G. H.: Analytical solutions for the surface response to small amplitude perturbations in boundary data in the shallow-ice-stream approximation, The Cryosphere, 2, 77–93, https://doi.org/10.5194/tc-2-77-2008, 2008.

Habermann, M., Truffer, M., and Maxwell, D.: Changing basal conditions during the speed-up of Jakobshavn Isbræ, Greenland, The Cryosphere, 7, 1679–1692, https://doi.org/10.5194/tc-7-1679-2013, 2013.

Hernandez, V., Roman, J. E., and Vidal, V.: SLEPc: A Scalable and Flexible Toolkit for the Solution of Eigenvalue Problems, ACM Trans. Math. Softw., 31, 351–362, https://doi.org/10.1145/1089014.1089019, 2005.

Hernandez, V., Roman, J. E., Vidal, V., and A., T.: Krylov-Schur Methods in SLEPc, Tech. rep., Universidad Politecnica de Valencia, 2007.

Hill, E. A., Rosier, S. H. R., Gudmundsson, G. H., and Collins, M.: Quantifying the potential future contribution to global mean sea level from the Filchner–Ronne basin, Antarctica, The Cryosphere, 15, 4675–4702, https://doi.org/10.5194/tc-15-4675-2021, 2021.

Hock, R., Maussion, F., Marzeion, B., and Nowicki, S.: What is the global glacier ice volume outside the ice sheets?, Journal of Glaciology, p. 1–7, https://doi.org/10.1017/jog.2023.1, 2023.

Isaac, T., Petra, N., Stadler, G., and Ghattas, O.: Scalable and efficient algorithms for the propagation of uncertainty from data through inference to prediction for large-scale problems, with application to flow of the Antarctic ice sheet, Journal of Computational Physics, 296, 348–368, https://doi.org/https://doi.org/10.1016/j.jcp.2015.04.047, 2015.

Jacobs, S. S., Jenkins, A., Giulivi, C. F., and Dutrieux, P.: Stronger ocean circulation and increased melting under Pine Island Glacier ice shelf, Nature Geoscience, 4, 519–523, https://doi.org/10.1038/ngeo1188, 2011.

Jay-Allemand, M., Gillet-Chaulet, F., Gagliardini, O., and Nodet, M.: Investigating changes in basal conditions of Variegated Glacier prior to and during its 1982–1983 surge, The Cryosphere, 5, 659–672, https://doi.org/10.5194/tc-5-659-2011, 2011.

Jenkins, A.: A Simple Model of the Ice Shelf-Ocean Boundary Layer and Current, Journal of Physical Oceanography, 46, 1785–1803, https://doi.org/10.1175/JPO-D-15-0194.1, 2016.

Jenkins, A., Shoosmith, D., Dutrieux, P., Jacobs, S., Kim, T. W., Le, S. H., Ha, H. K., and Stammerjohn, S.: West Antarctic Ice Sheet retreat in the Amundsen Sea driven by decadal oceanic variability, Nat. Geoscience, 11, 733–738, https://doi.org/https://doi.org/10.1038/s41561-018-0207-4DO, 2018.

Joughin, I., Smith, B., and Holland, D. M.: Sensitivity of 21st Century Sea Level to Ocean-Induced Thinning of Pine Island Glacier, Antarctica, Geophys. Res. Lett., 37, L20 502, https://doi.org/10.1029/2010GL044819, 2010.

Joughin, I., Smith, B. E., and Medley, B.: Marine Ice Sheet Collapse Potentially Under Way for the Thwaites Glacier Basin, West Antarctica, Science, 344, 735–738, https://doi.org/10.1126/science.1249055, 2014.

Kalmikov, A. G. and Heimbach, P.: A Hessian-Based Method for Uncertainty Quantification in Global Ocean State Estimation, SIAM Journal on Scientific Computing, 36, S267–S295, https://doi.org/10.1137/130925311, 2014.

Kazmierczak, E., Sun, S., Coulon, V., and Pattyn, F.: Subglacial hydrology modulates basal sliding response of the Antarctic ice sheet to climate forcing, The Cryosphere, 16, 4537–4552, https://doi.org/10.5194/tc-16-4537-2022, 2022.

Khazendar, A., Rignot, E., and Larour, E.: Acceleration and spatial rheology of Larsen C Ice Shelf, Antarctic Peninsula, Geophysical Research Letters, 38, https://doi.org/https://doi.org/10.1029/2011GL046775, 2011.

835 Koziol, C. P., Todd, J. A., Goldberg, D. N., and Maddison, J. R.: fenics_ice 1.0: a framework for quantifying initialization uncertainty for time-dependent ice sheet models, Geoscientific Model Development, 14, 5843–5861, https://doi.org/10.5194/gmd-14-5843-2021, 2021.

Lannelongue, L., Grealey, J., and Inouye, M.: Green Algorithms: Quantifying the Carbon Footprint of Computation, Advanced Science, 8, 2100 707, https://doi.org/https://doi.org/10.1002/advs.202100707, 2021.

Levermann, A., Winkelmann, R., Albrecht, T., Goelzer, H., Golledge, N. R., Greve, R., Huybrechts, P., Jordan, J., Leguy, G., Martin, D.,
840 Morlighem, M., Pattyn, F., Pollard, D., Quiquet, A., Rodehacke, C., Seroussi, H., Sutter, J., Zhang, T., Van Breedam, J., Calov, R., DeConto, R., Dumas, C., Garbe, J., Gudmundsson, G. H., Hoffman, M. J., Humbert, A., Kleiner, T., Lipscomb, W. H., Meinshausen, M., Ng, E., Nowicki, S. M. J., Perego, M., Price, S. F., Saito, F., Schlegel, N.-J., Sun, S., and van de Wal, R. S. W.: Projecting Antarctica's contribution to future sea level rise from basal ice shelf melt using linear response functions of 16 ice sheet models (LARMIP-2), Earth System Dynamics, 11, 35–76, https://doi.org/10.5194/esd-11-35-2020, 2020.

845 Lilien, D. A., Joughin, I., Smith, B., and Gourmelen, N.: Melt at grounding line controls observed and future retreat of Smith, Pope, and Kohler glaciers, The Cryosphere, 13, 2817–2834, https://doi.org/https://doi.org/10.5194/tc-13-2817-2019, 2019.

Lindgren, F., Rue, H., and Lindström, J.: An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 73, 423–498, https://doi.org/https://doi.org/10.1111/j.1467-9868.2011.00777.x, 2011.

850 MacAyeal, D. R.: Large-scale ice flow over a viscous basal sediment: Theory and application to Ice Stream B, Antarctica, Journal of Geophysical Research, 94, 4071–4087, 1989.

Macayeal, D. R.: The Basal Stress Distribution of Ice Stream E, Antarctica, Inferred by Control Methods, JOURNAL OF GEOPHYSICAL RESEARCH, 97, 595–603, 1992.

Maddison, J. R., Goldberg, D. N., and Goddard, B. D.: Automated Calculation of Higher Order Partial Differential Equation Constrained
855 Derivative Information, SIAM Journal on Scientific Computing, 41, C417–C445, https://doi.org/10.1137/18M1209465, 2019.

Martin, J., Wilcox, L. C., Burstedde, C., and Ghattas, O.: A Stochastic Newton MCMC Method for Large-Scale Statistical Inverse Problems with Application to Seismic Inversion, SIAM Journal on Scientific Computing, 34, A1460–A1487, https://doi.org/10.1137/110845598, 2012.

Morales, J. L. and Nocedal, J.: Remark on "Algorithm 778: L-BFGS-B: Fortran Subroutines for Large-Scale Bound Constrained Optimiza-
860 tion", ACM Trans. Math. Softw., 38, https://doi.org/10.1145/2049662.2049669, 2011.

Morlighem, M., Rignot, E., Seroussi, G., Larour, E., Ben Dhia, H., and Aubry, D.: Spatial patterns of basal drag inferred using control methods from a full-Stokes and simpler models for Pine Island Glacier, West Antarctica, Geophys. Res. Lett., 37, L14 502, https://doi.org/10.1029/2010GL043853, 2010.

Morlighem, M., Rignot, E., Binder, T., Blankenship, D., Drews, R., Eagles, G., Eisen, O., Ferraccioli, F., Forsberg, R., Fretwell, P., Goel,
865 V., Greenbaum, J. S., Gudmundsson, H., Guo, J., Helm, V., Hofstede, C., Howat, I., Humbert, A., Jokat, W., Karlsson, N. B., Lee, W. S., Matsuoka, K., Millan, R., Mouginot, J., Paden, J., Pattyn, F., Roberts, J., Rosier, S., Ruppel, A., Seroussi, H., Smith, E. C., Steinhage, D., Sun, B., Broeke, M. R. v. d., Ommen, T. D. v., Wessem, M. v., and Young, D. A.: Deep glacial troughs and stabilizing ridges unveiled beneath the margins of the Antarctic ice sheet, Nature Geoscience, 13, 132–137, https://doi.org/10.1038/s41561-019-0510-8, 2020.

Mouginot, J., Rignot, E., Scheuchl, B., and Millan, R.: Comprehensive Annual Ice Sheet Velocity Mapping Using Landsat-8, Sentinel-1, and
870 RADARSAT-2 Data, Remote Sensing, 9, https://doi.org/10.3390/rs9040364, 2017.

Pattyn, F.: Antarctic subglacial conditions inferred from a hybrid ice sheet/ice stream model, Earth and Planetary Science Letters, 295, 451–461, https://doi.org/https://doi.org/10.1016/j.epsl.2010.04.025, 2010.

Pattyn, F., Perichon, L., Aschwanden, A., Breuer, B., de Smedt, B., Gagliardini, O., Gudmundsson, G. H., Hindmarsh, R. C. A., Hubbard, A., Johnson, J. V., Kleiner, T., Konovalov, Y., Martin, C., Payne, A. J., Pollard, D., Price, S., Rückamp, M., Saito, F., Souček, O., Sugiyama, S., and Zwinger, T.: Benchmark experiments for higher-order and full-Stokes ice sheet models (ISMIP HOM), The Cryosphere, Volume 2, Issue 2, 2008, pp.95-108, 2, 95–108, 2008.

Petra, N., Martin, J., Stadler, G., and Ghattas, O.: A Computational Framework for Infinite-Dimensional Bayesian Inverse Problems, Part II: Stochastic Newton MCMC with Application to Ice Sheet Flow Inverse Problems, SIAM Journal on Scientific Computing, 36, A1525–A1555, https://doi.org/10.1137/130934805, 2014.

Rignot, E., Mouginot, J., and Scheuchl, B.: Ice Flow of the Antarctic Ice Sheet, Science, 333, 1427–1430, https://doi.org/10.1126/science.1208336, 2011.

Rignot, E., Mouginot, J., and Scheuchl, B.: MEaSUREs InSAR-Based Ice Velocity of the Amundsen Sea Embayment, Antarctica, Version 1, https://doi.org/10.5067/MEASURES/CRYOSPHERE/nsidc-0545.001, 2014.

Rignot, E., Mouginot, J., and Scheuchl, B.: MEaSUREs InSAR-Based Antarctica Ice Velocity Map, Version 2, https://doi.org/10.5067/D7GK8F5J8M8R, 2017.

Robel, A. A., Seroussi, H., and Roe, G. H.: Marine ice sheet instability amplifies and skews uncertainty in projections of future sea-level rise, Proceedings of the National Academy of Sciences, 116, 14 887–14 892, https://doi.org/10.1073/pnas.1904822116, 2019.

Scheuchl, B., Mouginot, J., Rignot, E., Morlighem, M., and Khazendar, A.: Grounding line retreat of Pope, Smith, and Kohler Glaciers, West Antarctica, measured with Sentinel-1a radar interferometry data, Geophysical Research Letters, 43, 8572–8579, https://doi.org/https://doi.org/10.1002/2016GL069287, 2016.

Schoof, C.: A variational approach to ice stream flow, Journal of Fluid Mechanics, 556, 227–251, https://doi.org/10.1017/S0022112006009591, 2006.

Sergienko, O. V., MacAyeal, D. R., and Thom, J. E.: Reconstruction of snow/firn thermal diffusivities from observed temperature variation: application to iceberg C16, Ross Sea, Antarctica, 2004–07, Annals of Glaciology, 49, 91–95, https://doi.org/10.3189/172756408787814906, 2008.

Seroussi, H., Nakayama, Y., Larour, E., Menemenlis, D., Morlighem, M., Rignot, E., and Khazendar, A.: Continued retreat of Thwaites Glacier, West Antarctica, controlled by bed topography and ocean circulation, Geophysical Research Letters, pp. 6191—-6199, https://doi.org/10.1002/2017GL072910, 2017.

Shapero, D. R., Badgeley, J. A., Hoffman, A. O., and Joughin, I. R.: icepack: a new glacier flow modeling package in Python, version 1.0, Geoscientific Model Development, 14, 4593–4616, https://doi.org/10.5194/gmd-14-4593-2021, 2021.

Still, H., Hulbe, C., Forbes, M., Prior, D. J., Bowman, M. H., Boucinhas, B., Craw, L., Kim, D., Lutz, F., Mulvaney, R., and Thomas, R. E.: Tidal Modulation of a Lateral Shear Margin: Priestley Glacier, Antarctica, Frontiers in Earth Science, 10, https://doi.org/10.3389/feart.2022.828313, 2022.

Stuart, A. M.: Inverse problems: A Bayesian perspective, Acta Numerica, 19, 451–559, https://doi.org/10.1017/S0962492910000061, 2010.

Tabeart, J. M., Dance, S. L., Lawless, A. S., Migliorini, S., Nichols, N. K., Smith, F., and Waller, J. A.: The impact of using reconditioned correlated observation-error covariance matrices in the Met Office 1D-Var system, Quarterly Journal of the Royal Meteorological Society, 146, 1372–1390, https://doi.org/https://doi.org/10.1002/qj.3741, 2020.

Tarantola, A.: Inverse Problem Theory and Methods for Model Parameter Estimation, Society for Industrial and Applied Mathematics, https://doi.org/10.1137/1.9780898717921, 2005.

910   Thacker, W. C.: The role of the Hessian matrix in fitting models to measurements, J. Geophys. Res., 94, 6177–6196, https://doi.org/10.1029/JC094iC05p06177, 1989.

Tierney, L.: Markov Chains for Exploring Posterior Distributions, The Annals of Statistics, 22, 1701 – 1728, https://doi.org/10.1214/aos/1176325750, 1994.

Tsai, C.-Y., Forest, C. E., and Pollard, D.: Assessing the contribution of internal climate variability to anthropogenic changes in ice sheet
915   volume, Geophysical Research Letters, 44, 6261–6268, https://doi.org/https://doi.org/10.1002/2017GL073443, 2017.

Waddington, E. D., Neumann, T. A., Koutnik, M. R., Marshall, H.-P., and Morse, D. L.: Inference of accumulation-rate patterns from deep layers in glaciers and ice sheets, Journal of Glaciology, 53, 694–712, https://doi.org/10.3189/002214307784409351, 2007.

Weertman, J.: On the Sliding of Glaciers, Journal of Glaciology, 3, 33–38, https://doi.org/10.3189/S0022143000024709, 1957.

Zhu, C., Byrd, R. H., Lu, P., and Nocedal, J.: Algorithm 778: L-BFGS-B: Fortran Subroutines for Large-Scale Bound-Constrained Optimiza-
920   tion, ACM Trans. Math. Softw., 23, https://doi.org/10.1145/279232.279236, 1997.

Table 1. Prior strength configurations used in Sect. 5.1, based on the pointwise standard deviation $\sigma_{c(0)}$ and auto-covariance length scale $l_{c(0)}$ of each control parameter, ordered from weak to strong.

Prior configurations	$\sigma_{\alpha(0)}$	$l_{\alpha(0)}$	γ_α	δ_α	$\sigma_{\beta(0)}$	$l_{\beta(0)}$	γ_β	δ_β
weak	1000	3000	0.85	9.40×10^{-8}	30	1000	9.4	9.40×10^{-6}
	510	2000	1.1	2.80×10^{-7}	30	1000	9.4	9.40×10^{-6}
$\downarrow$	**500**	**3000**	**1.7**	**1.90×10^{-7}**	**30**	**1000**	**9.4**	**9.40×10^{-6}**
	500	3000	1.7	1.90×10^{-7}	60	1000	4.7	4.70×10^{-6}
strong	150	3000	5.6	6.20×10^{-7}	30	1000	9.4	9.40×10^{-6}

The configuration in bold is also used in the experiments of Sect. 5.2 and 5.3. The units of $\sigma_{\alpha(0)}$ are $\mathrm{m}^{-1/6}\,\mathrm{yr}^{1/6}\,\mathrm{Pa}^{1/2}$ and $\sigma_{\beta(0)}$ are $\mathrm{Pa}^{1/2}\,\mathrm{yr}^{1/6}$. The unit of the auto-covariance length scale $l_{c(0)}$ is reported in m. Following (13), the units of γ_α are $\mathrm{m}^{7/6}\,\mathrm{yr}^{-1/6}\,\mathrm{Pa}^{-1/2}$, γ_β are $\mathrm{m}\,\mathrm{Pa}^{-1/2}\,\mathrm{yr}^{-1/4}$, δ_α are $\mathrm{m}^{-5/6}\,\mathrm{yr}^{-1/6}\,\mathrm{Pa}^{-1/2}$ and δ_α are $\mathrm{m}^{-1}\,\mathrm{Pa}^{-1/2}\,\mathrm{yr}^{-1/4}$.
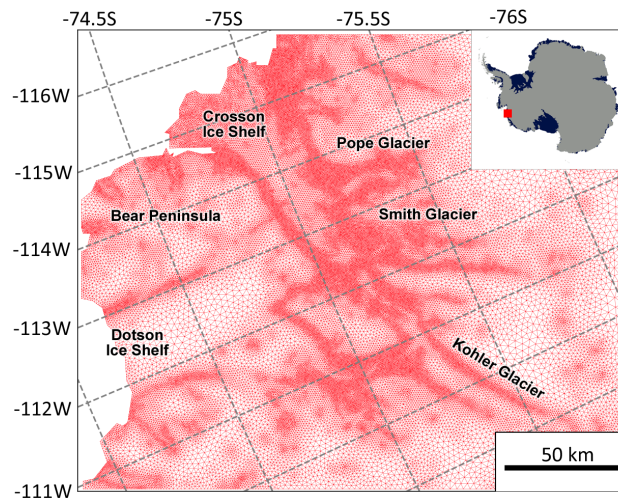
Figure 1. Variable resolution mesh of the ice streams region. The resolution depends on observed strain rates derived by using satellite velocity data (MEaSUREs v1.0 1996 - 2012, Rignot et al., 2014) and BedMachine Antarctica v2.0 (Morlighem et al., 2020). The boundaries to the East and South are entirely ice-ice boundaries, whereas the North and West features calving fronts where ice meets the ocean.
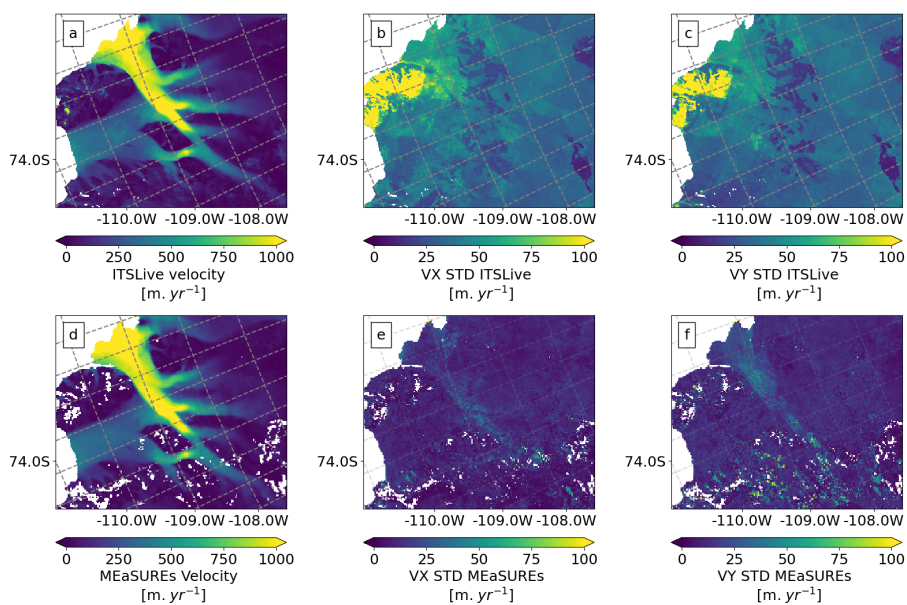
Figure 2. Observational input data. Satellite surface velocity observations (vector magnitude) and standard deviation (STD) of the velocity components (vx and vy) from ITS_LIVE (**a, b, and c.** Gardner et al., 2019, 2018) acquired from January to December 2014, and from MEaSUREs v2.0 (**d, e, and f.** Rignot et al., 2017) acquired from July 2013 to July 2014. For details on observational error model (STD estimates) see Sect. 3.1.3.
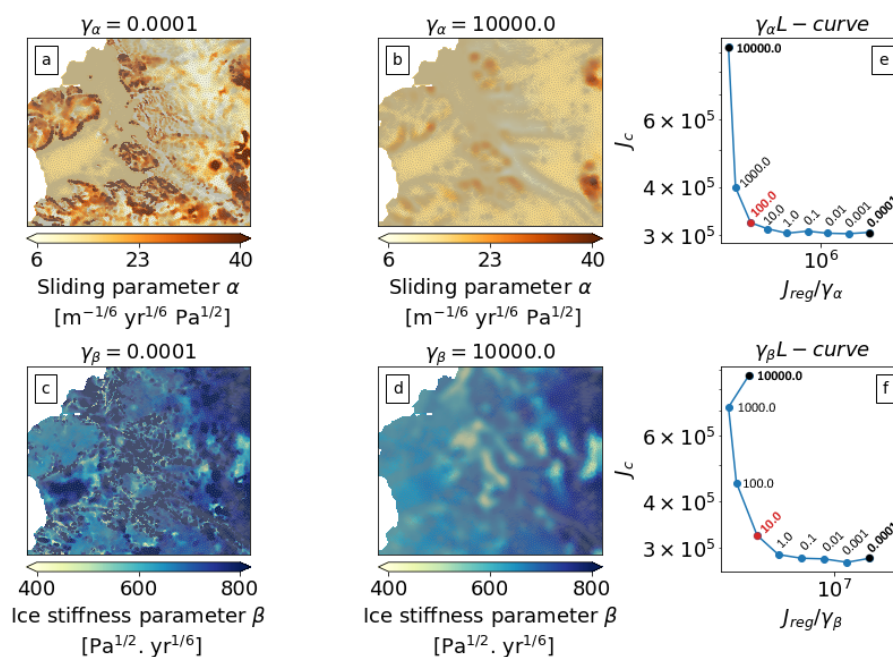
Figure 3. *L*-curve analysis output. **a and b:** Sliding parameter (α) computed using extreme γ_α values (bold values in panel e). **c and d:** Ice stiffness parameter (β) computed using extreme γ_β values (bold values in panel g). **e and g:** *L*-curve analysis for γ_α and γ_β, the optimal values suggested by the *L*-curves are highlighted in red.
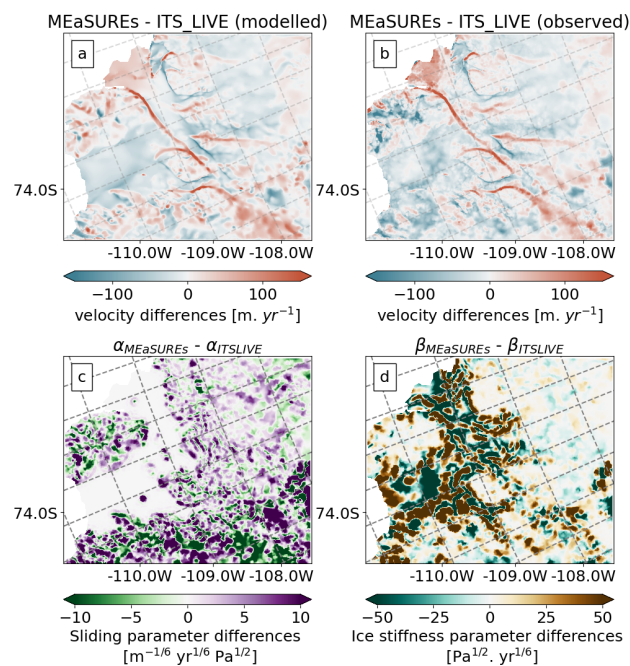
Figure 4. Inversion output differences after using two different satellite velocity products (MEaSUREs and ITSlive) to calibrate the ice dynamic parameters. **a:** Modelled velocity differences **b:** Observed velocity differences. **c:** Sliding parameter (α) differences. **d:** Ice stiffness parameter differences (β).
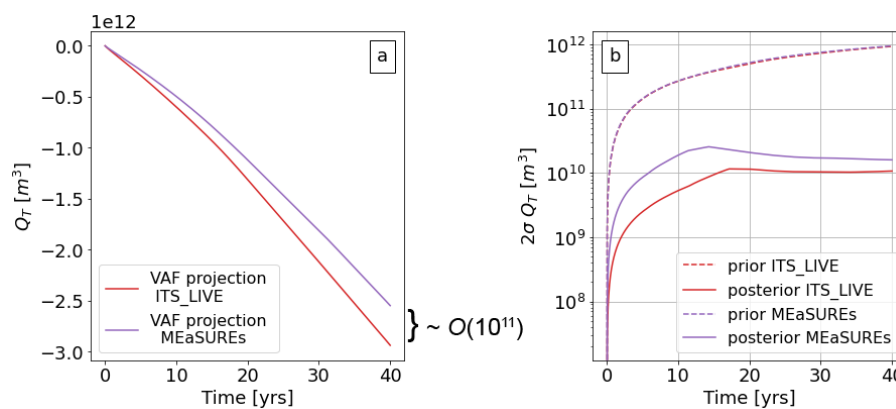
Figure 5. a VAF trajectories using different velocity products and the regularisation terms suggested by the L-curves ($\gamma_\alpha = 100$, $\gamma_\beta = 10$).
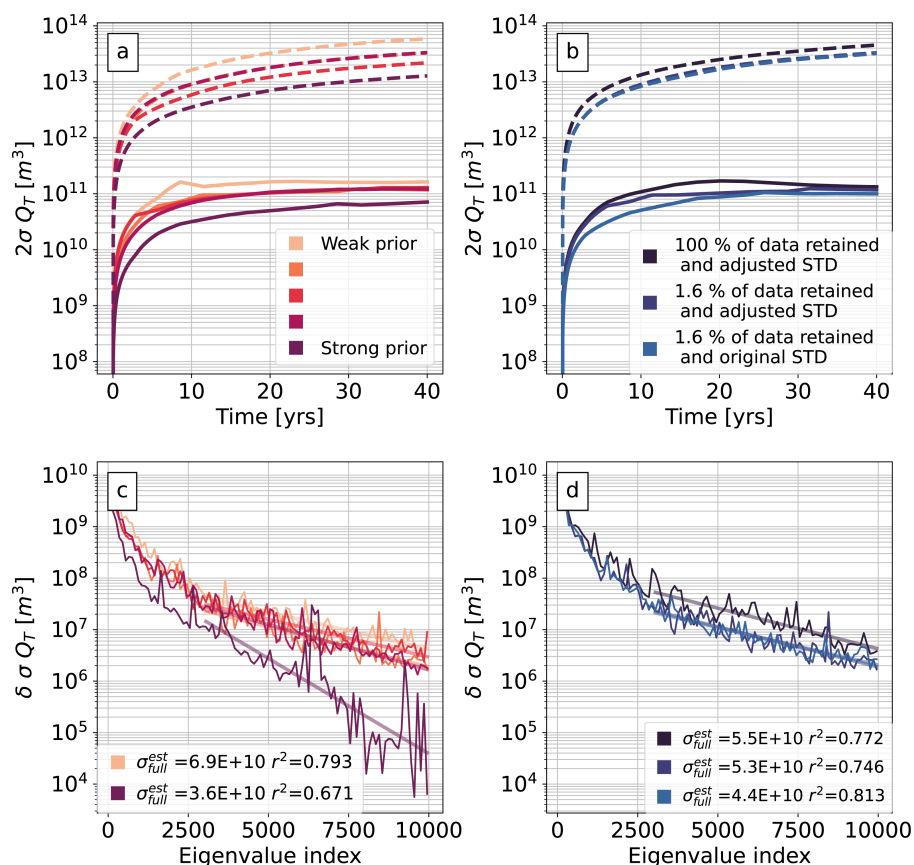
b: Hessian-based prior (dash lines) and posterior (solid line) uncertainties of VAF over time ($2\sigma Q_T$, represents 95% confidence interval).
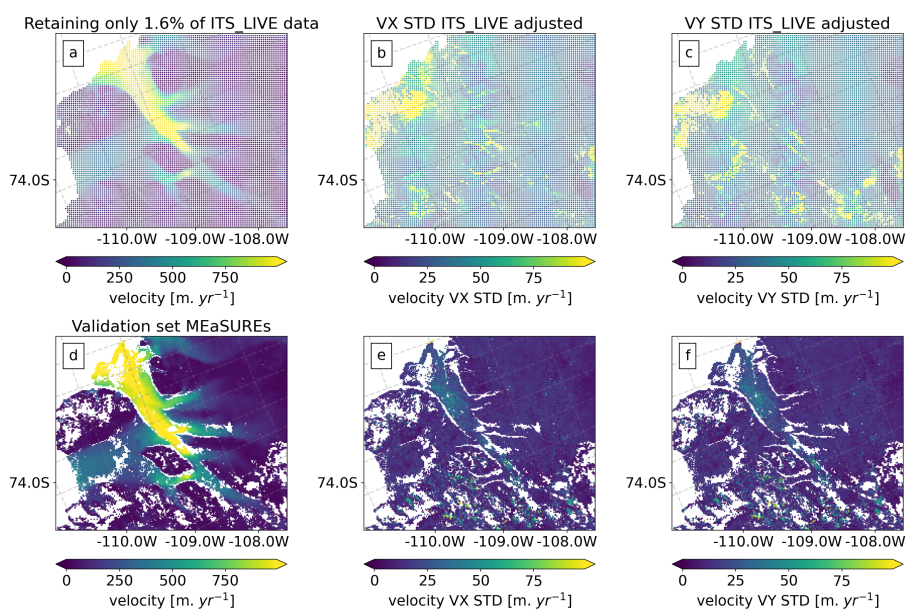
Figure 6. Overview of the input data used for the observational data subsampling test and for the experiments presented in Sect. 5. **a:** example of a training data set from ITS_LIVE where only 1.6% of the data points are retained, **b and c:** ITS_LIVE uncertainty in the x and y direction with the same data density as in (a) and with the STD of each component adjusted (see Sect. 5.2 for details). **d to f:** MEaSUREs validation data set used for the test.
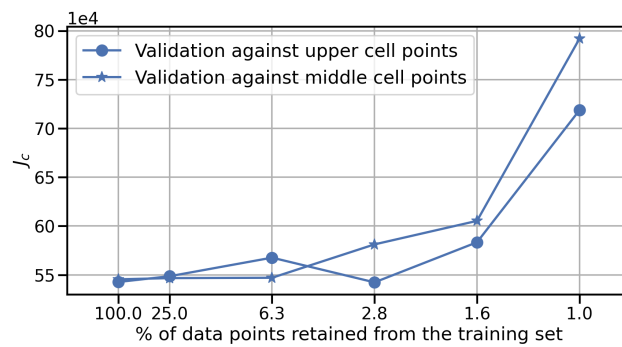
Figure 7. Observational data subsampling results. J^c performance if retaining a different number of observations for each training set.

Figure 8. Hessian-based prior (dash lines) and posterior (solid line) uncertainties of VAF over time ($2\sigma Q_T$, represents 95% confidence interval) computed using; **a:** different strengths of prior and **b:** different versions of the ITS_LIVE data (i.e. different data density and STD, see details in Sect.5.2). **c and d**: Rate of change for the posterior uncertainty of VAF ($\delta\sigma Q_T$) against the number of eigenvalues calculated; statistics are shown in the lower corners, i.e. σ_{full}^{est} the estimated STD of VAF for an infinite number of eigenvalues and the decreasing trend coefficient of determination (r^2).
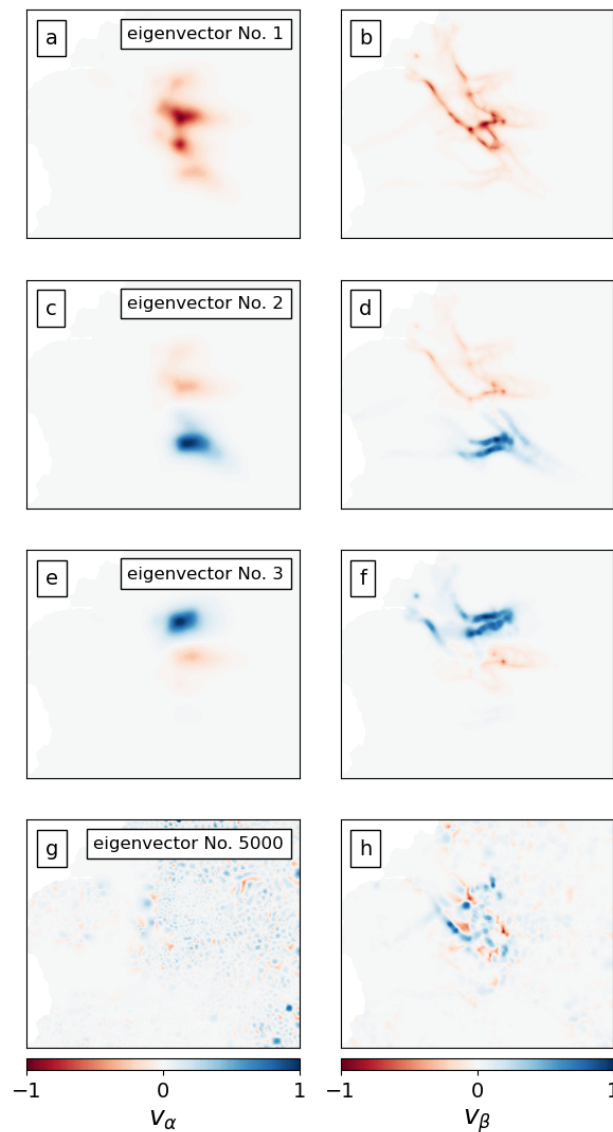
Figure 9. Eigenvectors (v) of the Hessian. Each eigenvector has an α component (right column) and a β component (second column). Each component is scaled to have a maximum magnitude of one and the scaling factor is shown in each vector legend. Ordered from large to small (from top to bottom), these v correspond to the 1st, 2nd, 3rd, and 5000th eigenvalues.
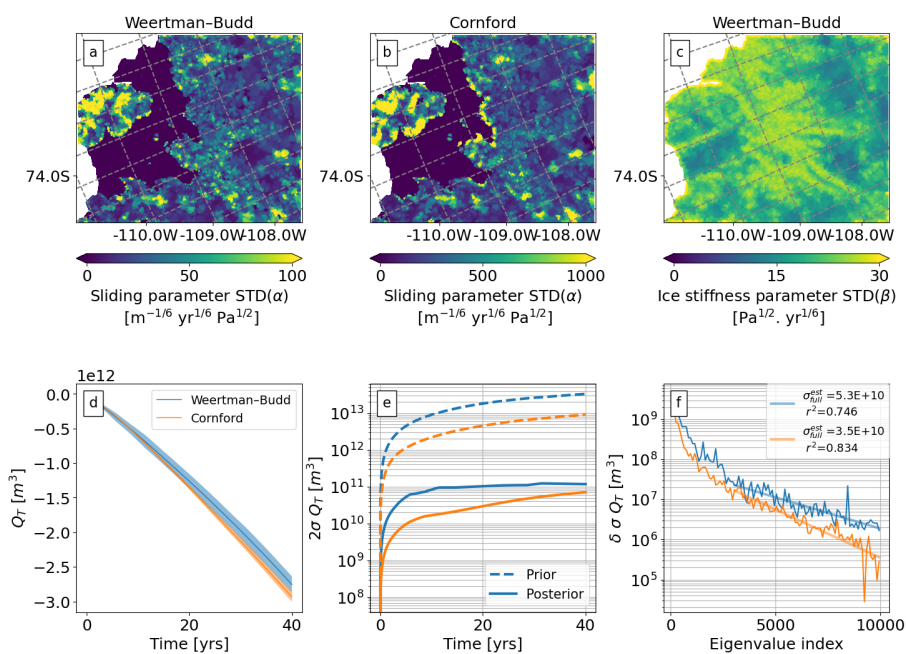
Figure 10. Model output when using two different sliding laws. Point-wise STD of the sliding parameter α for the **(a)** Weertman-Budd and **(b)** Cornford law. **c:** Point-wise STD of the ice stiffness parameter β (independent of the sliding law). **d:** VAF trajectories using the different sliding laws and the highlighted weak prior from Table 1. **e:** Hessian-based prior (dash lines) and posterior (solid lines) uncertainties of VAF over time ($2\sigma Q_T$, represents 95% confidence interval). **f:** Rate of change for the posterior uncertainty of VAF ($\delta\sigma Q_T$) against the number of eigenvalues calculated; statistics are shown in the lower corners, i.e. σ_{full}^{est} the estimated STD of VAF for an infinite number of eigenvalues and the decreasing trend coefficient of determination (r^2).
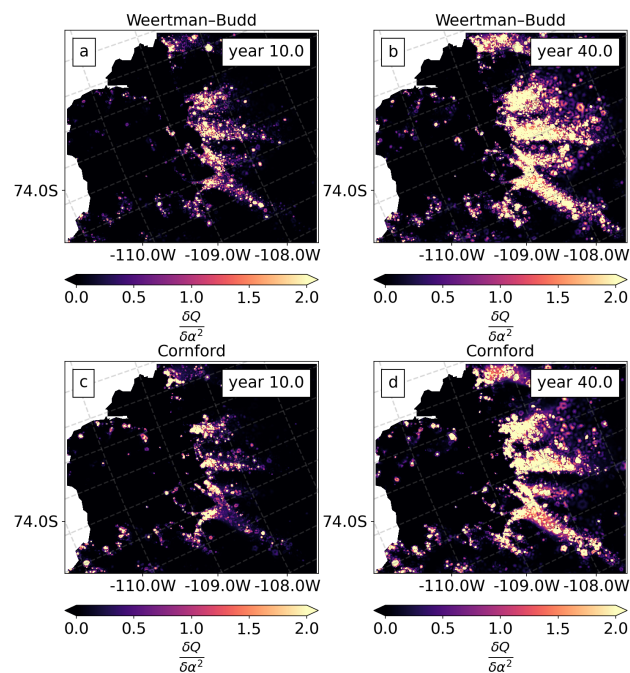
Figure 11. Sensitivity maps of the model's volume above flotation to the basal friction coefficient α^2; for year 10 and year 40, when using two different sliding laws. Units of α^2 are m$^{-1/3}$ yr$^{1/3}$ Pa. **a and b:** Weertman-Budd and **c and d:** Cornford law. These visualise the node-wise sensitivity given the choice of mesh and finite element discretisation.
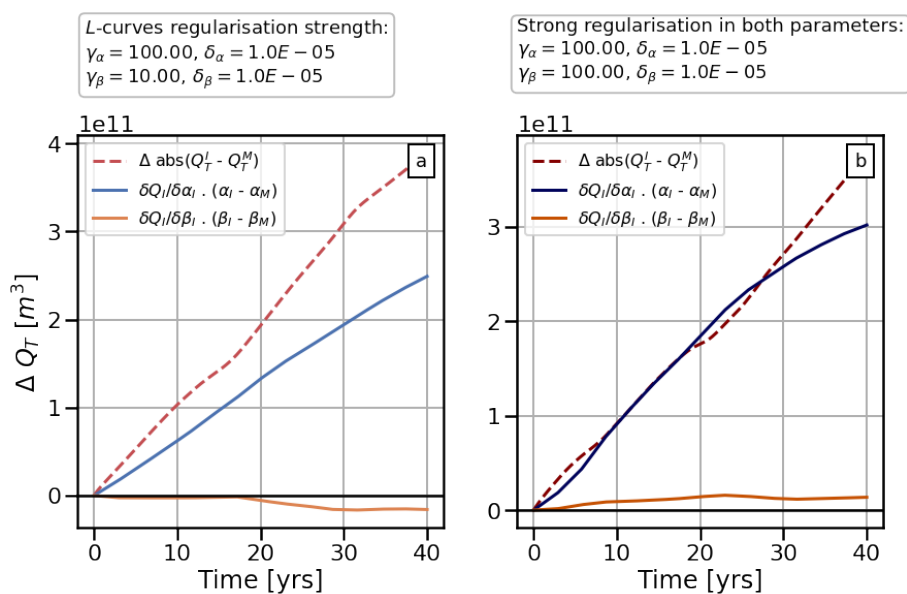
Figure 12. Linearity test results. Absolute difference of VAF trajectories estimated using different satellite velocity products (red dotted lines) and dot products results (solid lines) from (21) and (22) plotted in blue and yellow respectively. For this figure we use model output from Sect. 4.2. **a:** Linearity test using the regularisation strength suggested by the L-curves and **b:** Linearity test using a stronger regularisation on β.
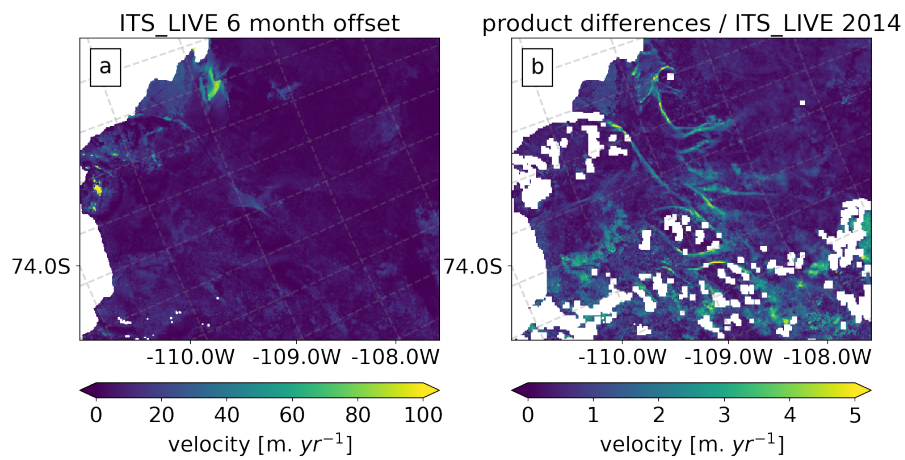
Figure A1. Speed comparisons. **a:** 6 month offset speed change from ITS_LIVE 2014 (Jul-Dec) and **b:** Speed ratio of the difference between the two products (ITS_LIVE 2014 - MEaSUREs acquired from July 2013 to July 2014) and ITS_LIVE 2014. Empty pixels in panel b are due to gaps in the MEaSUREs dataset, which increase when the data is interpolated to the ITS_LIVE grid.
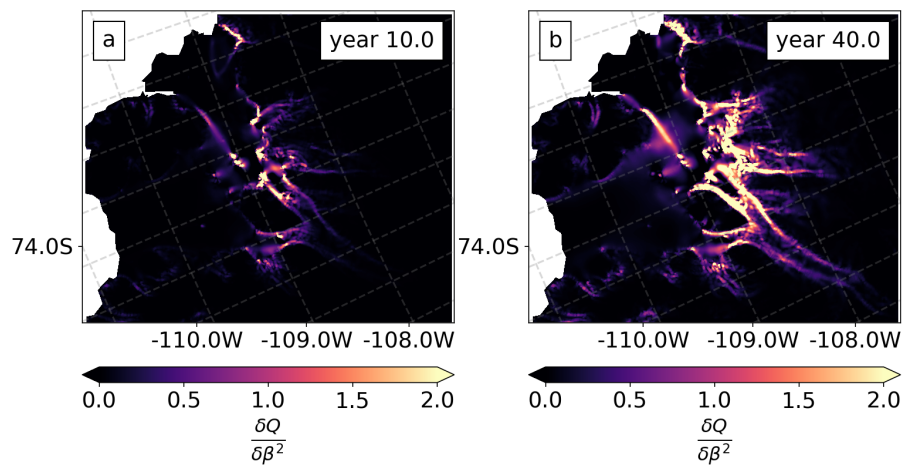
Figure A2. Sensitivity maps of the model's volume above flotation (VAF) to the ice stiffness (B or β^2). Units of B are $\mathrm{Pa\ yr}^{1/3}$
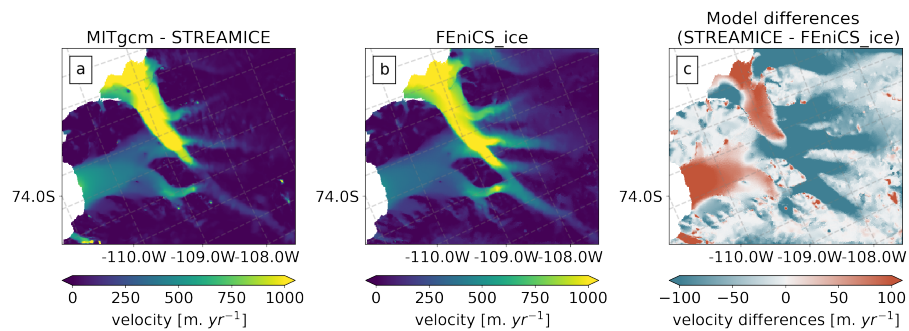
Figure A3. Models surface velocity comparison. Surface velocities are calculated by using `FEniCS_ice` inversions of α and β calibrated with ITS_LIVE and the highlighted weak prior configuration from Table 1. **a:** STREAMICE. **b:** `FEniCS_ice`. **c:** Difference between both models.