https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

An evaluation of multi-fidelity methods for quantifying uncertainty in projections of ice-sheet mass-change

John D. Jakeman[1], Mauro Perego[2], D. Thomas Seidl[2], Tucker A. Hartland[3], Trevor R. Hillebrand[4], Matthew J. Hoffman[4], and Stephen F. Price[4]

[1]Optimization and Uncertainty Quantification, Sandia National Laboratories, Albuquerque, NM, 87123
[2]Scientific Machine Learning, Sandia National Laboratories, Albuquerque, NM, 87123
[3]Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, Livermore, CA, 94550
[4]Fluid Dynamics and Solid Mechanics Group, Los Alamos National Laboratory, Los Alamos, NM, 87544

Correspondence: J. D. Jakeman (jdjakem@sandia.gov)

Abstract. This study investigated the computational benefits of using multi-fidelity uncertainty quantification (MFUQ) algorithms to quantify uncertainty in the mass change of Humboldt Glacier, Greenland, between 2007 and 2100 using a single climate change scenario. The goal of this study was to determine whether MFUQ can use multiple models of varying cost and accuracy to reduce the computational cost of estimating the mean and variance of the projected mass change of an ice sheet. The problem size and complexity were chosen to be representative of future continental scale studies while still facilitating a computationally feasible investigation of MFUQ methods. When quantifying uncertainty introduced by a high-dimensional parameterization of basal friction field, MFUQ was able to reduce the mean-squared error in the estimates of the statistics by well over an order of magnitude when compared to a single fidelity approach that only used the highest fidelity model. This significant reduction in computational cost was achieved despite the low-fidelity models used being incapable of capturing the local features of the ice flow fields predicted by the high-fidelity model. The MFUQ algorithms were able to effectively leverage the high correlation between each model's prediction of mass change, which all responded similarly to perturbations in the model inputs. Consequently, our results suggest that MFUQ could be highly useful for reducing the cost of computing continental scale probabilistic projections of sea-level rise due to ice-sheet mass change.

1 Introduction

The most recent Intergovernmental Panel on Climate Change (IPCC) predicts that the melting of land-based ice sheets will contribute significantly to future rises in sea level (Fox-Kemper et al., 2021), but the amount of sea-level rise is subject to a large degree of uncertainty. For example, estimates of the sea-level rise in 2100, caused by melting of the Greenland Ice Sheet, range from 0.01 m to 0.18 m. Moreover, projections of the Antarctic Ice Sheet's contribution to sea-level rise are subject to even larger uncertainty (Bakker et al., 2017; Fox-Kemper et al., 2021; Edwards et al., 2019). Consequently, there is a strong need to accompany recent improvements in the numerical modeling of ice-sheet dynamics with rigorous methods that quantify uncertainty in model predictions.

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

Accurately quantifying uncertainty in ice-sheet predictions requires estimating the impacts of all sources of model variability. Prediction uncertainty is caused by the inadequacy of the governing equations used by the model to approximate reality, the errors introduced by the numerical discretization used to solve the governing equations, the uncertainty in model inputs used to parameterize future climate forcing, and the current condition of the ice sheet, etc. Several studies have demonstrated that model discretization significantly effects model predictions (Cornford et al., 2013; Durand et al., 2009), but the impact of discretization errors has not been explicitly considered with other sources of uncertainty. In addition, while the comparison of model outputs has been used to approximate model inadequacy (Goelzer et al., 2018), such studies are not guaranteed to estimate the true model inadequacy (Knutti et al., 2010). Consequently, several recent efforts have focused solely on quantifying parametric uncertainty (Edwards et al., 2021; Ritz et al., 2015; Schlegel et al., 2018).

Parametric uncertainty is often estimated using Monte Carlo (MC) uncertainty quantification (UQ) methods that compute statistics, or construct probability densities, using a large number of model simulations evaluated at different random realizations of the uncertain model inputs. However, the substantial computational cost of evaluating ice-sheet models limits the number of model simulations that can be run, and thus the accuracy of uncertainty estimates. Therefore, recent UQ efforts have constructed emulators (also known as surrogates) of the numerical model from a limited amount of simulation data and then sampled the surrogate to quantify uncertainty (Berdahl et al., 2021; Bulthuis et al., 2019; Edwards et al., 2019; Jantre et al., 2024). While surrogates can improve the computational tractability of UQ when uncertainty is parameterized by a small number of inputs, they cannot be applied when there are more than 10-20 variables because the amount of simulation data required to build them grows exponentially with the number of inputs (parameter dimension) (Jakeman, 2023). Consequently, methods for quantifying uncertainty in ice-sheet models with large numbers of uncertain parameters — such as a spatially varying basal friction field — are needed.

Multi-fidelity UQ (MFUQ) methods (Giles, 2015; Peherstorfer et al., 2016; Gorodetsky et al., 2020; Schaden and Ullmann, 2020) have the potential to reduce the computational cost of quantifying uncertainty in ice-sheet models. MFUQ methods utilize models of varying fidelity, that is models with different inadequacy, numerical discretization and computational cost, to efficiently and accurately quantify uncertainty. Specifically, MFUQ methods produce unbiased statistics of a trusted highest-fidelity model by combining a small number of simulations of the high-fidelity model with larger amounts of data from multiple cheaper sources. Furthermore, provided the low-fidelity models are highly correlated with the high-fidelity model and are substantially cheaper to simulate, the mean squared error (MSE) of the MFUQ statistic will often be an order of magnitude smaller than the estimate obtained using solely high-fidelity evaluations, for a fixed computational budget. However, such gains have yet to be realized when quantifying uncertainty in ice-sheet models.

This study investigated the efficacy of using MFUQ methods to reduce the computational cost needed to accurately estimate statistics summarizing the uncertainty in predictions of sea-level rise obtained using ice-sheet models parameterized by large numbers of inputs. To facilitate a computationally feasible investigation, we studied the performance of MFUQ methods when used to quantify the uncertainty in mass-change predictions for a single glacier as a function of a spatially varying basal friction field. We focused on the evolution of the Humboldt Glacier in northern Greenland under a single climate change scenario between 2007 and 2100. Specifically, to estimate glacier mass change, we used 13 different model fidelities, based

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

on different numerical discretizations of mono-layer, higher-order (MOLHO) (Dias dos Santos et al., 2022) and shallow-shelf (SSA) (Morland and Johnson, 1980; Weis et al., 1999) physics approximations. Each model was parameterized by a piece-wise linear discretization of a log-normal basal friction field, and Bayesian inference was used to calibrate the resulting 11,536 dimensional uncertain variable to match available observations of glacier surface velocity.

Our study makes two novel improvements over previously published glaciology literature. First, we quantified the impact of a high-dimensional parameterizations of basal friction on long-term ice-sheet projections. Additionally, we used Bayesian inference to constrain the prior distribution assigned to the friction field on observational data, which in turn reduced the uncertainty in our model predictions. In contrast, previous UQ studies (Nias et al., 2023; Ritz et al., 2015; Schlegel et al., 2018; Jantre et al., 2024) only employed low-dimensional parameterizations despite high-dimensional parameterizations being necessary to calibrate ice-sheet models to observational data (Barnes et al., 2021; Isaac et al., 2015; Johnson et al., 2023; Perego et al., 2014). Moreover, only Johnson et al. (2023) used Bayesian inference to quantify the uncertainty inherent in the calibration. Second, this study represents the first applications of MFUQ methods to quantify uncertainty in ice-sheet projections. Most importantly, our results demonstrate that MFUQ can reduce the serial computational time required for an accurate UQ study of ice-sheet contribution to sea-level rise from years to a month. Our paper also provides a comprehensive discussion of the practical issues that arise when using MFUQ, which are often ignored in the existing MFUQ literature.

This paper is organized as follows. First, Section 2 details the different ice-sheet models considered by this study and the parameterization of uncertainty employed. Second, Section 3 presents the calibration of the ice sheet model and how the posterior samples were generated. Third, Section 4 presents the MFUQ methods that were used to quantify uncertainty. Fourth, Section 5 presents the numerical results of the study and Section 6 presents our findings. Finally, conclusions are drawn in Section 7.

2 Methods

This section presents the model formulations (Section 2.1) and the numerical discretization of these models (Section 2.2) used to model ice-sheet evolution, as well as the sources of model uncertainty we considered (Section 2.3) when quantifying uncertainty in the mass change from Humboldt Glacier between 2007 and 2100.

2.1 Model Formulations

Ice-sheets behave as a shear thinning fluid and can be modeled with the nonlinear Stokes equation (Cuffey and Paterson, 2010). This section details the Stokes equations and two computationally less expensive simplifications, MOLHO (Dias dos Santos et al., 2022) and SSA (Morland and Johnson, 1980; Weis et al., 1999) which were used to quantify uncertainty in predictions of sea-level rise due to melting of the Humboldt Glacier.[1]

[1]We also implemented shallow-ice SIA models (Halfar, 1983), but despite our models passing verification tests the SIA models were not able to robustly simulate samples from the posterior distribution of the basal friction field. Most runs did not complete because the non-linear solver failed to converge, likely due to the fact that SIA models are not designed to model ice-sheets in the presence of low-friction.

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

Let x and y denote the horizontal coordinates and z the vertical coordinate, chosen such that the sea level, assumed to remain constant during the period of interest, corresponds to $z = 0$. We approximated the ice domain at time t as a vertically extruded domain Ω defined as

$$\Omega(t) := \{(x, y, z) \text{ s.t. } (x, y) \in \Sigma, \text{ and } l(x, y, t) < z < s(x, y, t)\},$$

where $\Sigma \subset \mathbb{R}^2$ denotes the horizontal extent of the ice, $\Gamma_l(t) := \{(x, y, z) \text{ s.t. } z = l(x, y, t), (x, y) \in \Sigma\}$ denotes the lower surface of the ice at time t, and $\Gamma_s(t) := \{(x, y, z) \text{ s.t. } z = s(x, y, t), (x, y) \in \Sigma\}$ denotes the upper surface of the ice[2].

The Stokes, MOLHO, and SSA models defined the thickness of the ice $H(x, y, t) = s(x, y, t) - l(x, y, t)$ as the difference between the ice-sheet surface $s(x, y, t)$ and the bottom of the ice-sheet $l(x, y, t)$. The bottom of the ice-sheet was allowed to be both grounded to the bed topography $b(x, y)$, such that $l(x, y, t) = b(x, y)$, or floating such that $l(x, y, t) = -\frac{\rho}{\rho_w} H(x, y, t)$, where ρ and ρ_w are the densities of ice and ocean water, respectively. Different boundary conditions were then applied on the grounded portion Γ_g of the ice bottom and on the floating portion Γ_f of the ice bottom, where $\Gamma_g \cap \Gamma_f = \emptyset$ and the ice bottom was given by $\Gamma_g \cup \Gamma_f$. The lateral boundary of Ω was also partitioned into the ice-sheet margin (either terrestrial or marine margin) Γ_m and an internal (artificial) boundary Γ_d marking the interior extent of the Humboldt Glacier that was considered. The relevant domains of the ice-sheet are depicted in Figure 1.

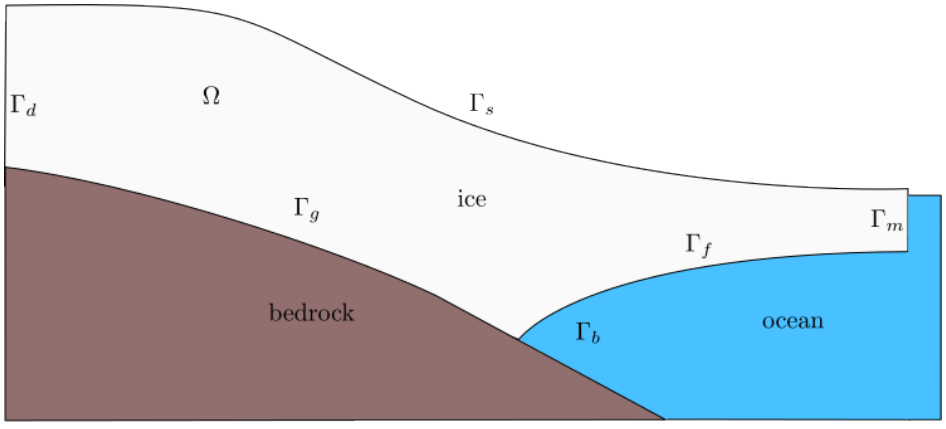


Figure 1. Conceptual model of an ice sheet in the $x - z$ plane.

The Stokes equations model the horizontal ice velocities $(u(x, y, z, t), v(x, y, z, t))$, vertical ice velocity $w(x, y, z, t)$ and thickness $H(x, y, z)$ of an ice-sheet as a function of the three spatial dimensions (x, y, z). In contrast, using the observation that ice-sheets are typically shallow, i.e. their horizontal extent is much greater than their thickness, the MOLHO model neglects the vertical velocity w and only simulates the horizontal velocities $u(x, y, z, t), v(x, y, z, t)$ but still as functions of the three spatial coordinates. Contrasting again, the SSA model makes the additional assumption that the horizontal components of velocity do not vary with thickness (a reasonable approximation in regions where motion is dominated by basal slip) so that the horizontal

[2]For simplicity here we assume that Σ does not change in time. This implies that the ice-sheet cannot extend beyond Σ but it can become thicker or thinner (to the point of disappearing in some regions).

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

velocities $u(x,y,t), v(x,y,t)$ are solved for only as functions of (x,y). In summary, the 3D Stokes model is formulated to simulate all types of ice-flow, whereas the 3D MOLHO model is formulated to simulate the flow of both frozen and thawed beds. Additionally, the simpler 2D SSA model is formulated to simulate grounded ice with significant sliding at the bed or ice shelves floating over the water.

The Stokes, MOLHO, and SSA models all evolve ice thickness $H(x,y,t)$ according to

$$\partial_t H + \nabla \cdot (\bar{\mathbf{u}} H) = f_H, \qquad H \geq 0, \tag{1}$$

where $\bar{\mathbf{u}} := \frac{1}{H} \int_l^s \mathbf{u}\, dz$ is the thickness-integrated velocity and f_H is a forcing term that accounts for accumulation (e.g. snow accumulation) and ablation (e.g. melting) at the upper (s) and lower (l) surfaces of the ice sheet. However, each model determines the velocities of the ice sheet differently. The following three subsections detail how each model computes the velocity of the ice sheet.

2.1.1 Stokes model

This section introduces the Stokes model, which while not used in this study due to its exorbitant computational cost, forms the basis of the other three models used in this study. Specifically, the governing equations of the Stokes model are

$$-\nabla \cdot \sigma = \rho \mathbf{g} \tag{2}$$

$$\nabla \cdot \mathbf{u} = 0. \tag{3}$$

The velocities $\mathbf{u} = (u,v,w)$ are dependent on the pressure p, ρ denotes the density of ice, $\sigma = 2\mu\mathbf{D} - pI$ denotes the stress tensor of the ice and $\mathbf{D}_{ij}(\mathbf{u}) = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)$ denotes the strain rate tensor of the ice; here we used the shorthand $\mathbf{u} = (u,v,w) = (u_1,u_2,u_3)$. The stress tensor is dependent on the non-linear viscosity of the ice which satisfies

$$\mu = \frac{1}{2} A(T)^{-q} D_e(\mathbf{u})^{q-1}, \tag{4}$$

where A is the ice flow factor that depends on the ice temperature T and $q \leq 1$; in our study we set $q = \frac{1}{3}$, which is a typical choice. In addition, the effective strain rate $D_e(\mathbf{u})$ satisfies $D_e(\mathbf{u}) = \frac{1}{\sqrt{2}}|\mathbf{D}(\mathbf{u})|$, where $|\cdot|$ denotes the Frobenius norm.

When used to model ice sheets, the Stokes equation must be accompanied by the following boundary conditions:

$$\begin{cases} \sigma\mathbf{n} = 0 & \text{on } \Gamma_s & \text{stress free, atmospheric pressure neglected} \\ \sigma\mathbf{n} = \rho_w g \min(z,0)\mathbf{n} & \text{on } \Gamma_m & \text{boundary condition at the ice margin} \\ \mathbf{u} = \mathbf{u}_d & \text{on } \Gamma_d & \text{Dirichlet condition at internal boundary (ice-flow divide)} \\ \mathbf{u}\cdot\mathbf{n} = \mathbf{0},\ (\sigma\mathbf{n})_{\|} = \beta\mathbf{u}_{\|} & \text{on } \Gamma_g & \text{impenetrability + sliding condition} \\ \sigma\mathbf{n} = \rho_w g z \mathbf{n} & \text{on } \Gamma_f & \text{hydrostatic pressure of ocean under ice shelves} \end{cases}$$

Here $\beta(x,y)$ is a linearized sliding (or friction) coefficient and $\mathbf{n}$ the unit outward-pointing normal to the boundary. The boundary condition at the margin includes an ocean back-pressure term when the margin is partially submerged ($z < 0$). For a terrestrial margin, $z > 0$, the boundary condition becomes a stress-free condition.

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

2.1.2 Mono-layer higher-order (MOLHO)

The MOLHO model (Dias dos Santos et al., 2022) is based on the Blatter-Pattyn approximation (Dukowicz et al., 2010) which can be derived by neglecting the terms w_x and w_y (the derivatives of w with respect to x and y, respectively) in the strain-rate tensor $\mathbf{D}$ and using the incompressibility condition ($\nabla \cdot \mathbf{u} = 0$) such that w_z can be expressed solely in terms of u_x and v_y and

$$\mathbf{D} = \begin{bmatrix} u_x & \frac{1}{2}(u_y + v_x) & \frac{1}{2}u_z \\ \frac{1}{2}(u_y + v_x) & v_y & \frac{1}{2}u_z \\ \frac{1}{2}u_z & \frac{1}{2}v_z & -(u_x + v_y) \end{bmatrix}. \tag{5}$$

This leads (Jouvet, 2016) to the following elliptic equations for the horizontal velocities (u, v)

$$-\nabla \cdot (2\mu\hat{\mathbf{D}}) = -\rho g \nabla_{xy} s, \tag{6}$$

where $\nabla_{xy} := [\partial_x, \partial_y]^\top$, and

$$\hat{\mathbf{D}} = \begin{bmatrix} 2u_x + v_y & \frac{1}{2}(u_y + v_x) & \frac{1}{2}u_z \\ \frac{1}{2}(u_y + v_x) & u_x + 2v_y & \frac{1}{2}v_z \end{bmatrix}. \tag{7}$$

such that the viscosity μ in (4) has the effective strain rate

$$D_e = \sqrt{u_x^2 + v_y^2 + u_x v_y + \frac{1}{4}(u_y + v_x)^2 + \frac{1}{4}u_z^2 + \frac{1}{4}v_z^2}.$$

MOLHO is derived from the weak form of the Blatter-Pattyn model (6), with the ansatz that the velocity can be expressed as

$$\mathbf{u}(x, y, z) = \mathbf{u}_b(x, y)\,\phi_b + \mathbf{u}_v(x, y)\,\phi_v\left(\frac{s - z}{H}\right), \quad \text{with } \phi_b = 1, \text{ and } \phi_v(\zeta) = 1 - \zeta^{\frac{1}{q}+1},$$

where the functions ϕ_b and ϕ_v are also used to define the test functions of the weak formulation of the MOLHO model. This ansatz allows the Blatter-Pattyn model to be simplified into a system of two two-dimensional partial differential equations (PDEs) for $\mathbf{u}_b$ and $\mathbf{u}_v$ — for a detailed derivation see (Dias dos Santos et al., 2022) — such that the thickness-averaged velocity satisfies $\bar{\mathbf{u}} = \mathbf{u}_b + \frac{(1+q)}{(1+2q)}\,\mathbf{u}_v$, where q is the same coefficient appearing in the viscosity definition (4).

We used the following boundary conditions when using MOLHO to simulate ice-flow

$$\begin{cases} 2\mu\hat{\mathbf{D}}\,\mathbf{n} = 0 & \text{on } \Gamma_s & \text{stress free, atmospheric pressure neglected} \\ 2\mu\hat{\mathbf{D}}\,\mathbf{n} = \psi\mathbf{n} & \text{on } \Gamma_m & \text{boundary condition at at ice margin} \\ \mathbf{u} = \mathbf{u}_d & \text{on } \Gamma_d & \text{Dirichlet condition at internal boundary (ice-flow divide)} \\ 2\mu\hat{\mathbf{D}}\,\mathbf{n} = \beta\mathbf{u}_\parallel & \text{on } \Gamma_g & \text{sliding condition} \\ 2\mu\hat{\mathbf{D}}\,\mathbf{n} = 0 & \text{on } \Gamma_f & \text{free slip under ice shelves.} \end{cases}$$

Additionally, we approximated the term $\psi = \rho g(s - z)\mathbf{n} + \rho_w\, g\, \min(z, 0)\mathbf{n}$ by its thickness-averaged value $\bar{\psi} = \frac{1}{2}gH(\rho - r^2\rho_w)$, where $r = \max\left(1 - \frac{s}{H}, 0\right)$ is the the submerged ratio and $\mathbf{u}_\parallel$ is the component of the velocity $\mathbf{u}$ tangential to the bed.

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

2.1.3 Shallow Shelf Approximation (SSA)

The SSA model (Morland and Johnson, 1980) is a simplification of the Blatter-Pattyn model that assumes the ice-velocity is uniform in z, so $\mathbf{u} = \bar{\mathbf{u}}$ and thus $u_z = 0, v_z = 0$. This simplification yields

$$\mathbf{D} = \begin{bmatrix} u_x & \frac{1}{2}(u_y + v_x) & 0 \\ \frac{1}{2}(u_y + v_x) & v_y & 0 \\ 0 & 0 & -(u_x + v_y) \end{bmatrix}, \quad \hat{\mathbf{D}} = \begin{bmatrix} 2u_x + v_y & \frac{1}{2}(u_y + v_x) & 0 \\ \frac{1}{2}(u_y + v_x) & u_x + 2v_y & 0 \end{bmatrix}, \tag{8}$$

and $D_e = \sqrt{u_x^2 + v_y^2 + u_x v_y + \frac{1}{4}(u_y + v_x)^2}$. Consequently, the SSA is a single two-dimensional PDE in Σ

$$-\nabla \cdot \left(2\mu H \hat{\mathbf{D}}(\bar{\mathbf{u}})\right) + \beta \bar{\mathbf{u}} = -\rho g H \nabla_{xy} s, \quad \text{in } \Sigma,$$

where $\bar{\mu} = \frac{1}{2}\bar{A}(T)^{-q} D_e(\bar{\mathbf{u}})^{q-1}$, and $\bar{A}$ is the thickness-averaged flow factor. This study explored the use of SSA with the boundary conditions

$$\begin{cases} 2\mu \hat{\mathbf{D}}(\bar{\mathbf{u}})\,\mathbf{n} = \bar{\psi}\mathbf{n} & \text{on } \Gamma_m \quad \text{boundary condition at ice margin} \\ \bar{\mathbf{u}} = \bar{\mathbf{u}}_d & \text{on } \Gamma_d \quad \text{Dirichlet condition at internal boundary.} \end{cases}$$

With abuse of notation, here Γ_m and Γ_d denotes subsets of $\partial\Sigma$.

2.2 Numerical discretization

The ability to predict ice-sheet evolution accurately is dictated not only by the governing equations used, but also by the properties of the numerical methods used to solve the governing equations. In this study, we discretized the thickness and the velocity equations of the MOLHO and SSA models using the popular Galerkin-based finite element method with piecewise linear elements, which we implemented in `FEniCS` (Alnæs et al., 2015). Additionally the coupled thickness and velocity equations were solved in a monolitic fashion using a Backward Euler time discretization and the `PETSc` (Balay et al., 1998) SNES nonlinear solver.

Because the thickness H obtained from (1) is not guaranteed to be positive due to the forcing term f_H and that the discretization used is not positivity preserving, we adopted two different approaches to guarantee the positivity of the thickness computed by our finite element models. The first approach involved updating the thickness value at each node so that it was greater than or equal to a minimum thickness value $H_m = 1$ m. The second approach used an optimization-based approach (Bochev et al., 2020) to preserve the thickness constraint ($H \geq H_m$) and guarantee that the total mass change is always consistent with the forcing term in regions where the ice is present and with the boundary fluxes. The first approach is computationally cheaper than the second, but unlike the second method does not conserve mass.

In addition to mass conservation, the number of finite elements and the time-step size both affect the error in the finite element approximation of the governing equations of the MOLHO and SSA models. In this study we investigated the impact of the number of finite elements, which we also refer to as the spatial mesh resolution, and time-step size, on accuracy. Specifically, the MOLHO and SSA models were both used to simulate ice-sheet evolution with four different finite element meshes and four

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

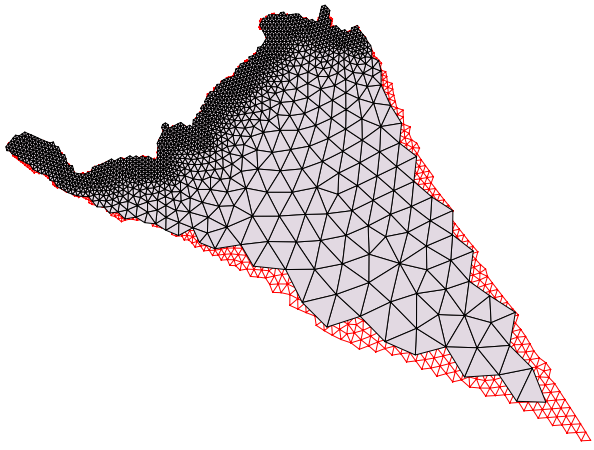


Figure 2. Comparison of the coarsest (solid gray) finite element mesh used to model the Humboldt Glacier, with the second coarsest mesh used (empty red).

different time-step sizes. More details on the spatial mesh and time-step sizes used are given in Section 5.1. Figure 2 compares the coarsest finite element mesh used to model the Humboldt Glacier, with the second coarsest mesh. Due to the differences in the characteristic element size of each mesh, the computational domain of each mesh is different. However, we will show that this did not prevent the use of these meshes in our study.

2.3 Parameterization of uncertainty

Many factors introduce uncertainty into the predictions of ice-sheet models including those associated with atmospheric forcing, ice rheology, basal friction, ice temperature, calving, and submarine melting. While all sources of uncertainty may significantly impact predictions of mass change from ice sheets, this study focused on quantifying uncertainty due to the unknown basal friction, which is considered one of the largest sources of prediction uncertainty. This singular focus was made to improve our ability to assess whether MFUQ is useful for ice-sheet modeling for a very high-dimensional source of uncertainty, which cannot be tractably tackled using most existing UQ methods. This ensures that the conclusions drawn by our study can be plausibly extended to studies considering additional sources of uncertainty.

The uncertainty in basal friction β, which impacts the boundary conditions of the MOLHO and SSA models, can be parameterized in a number of ways. For example, a lumped approach would assign a single scalar random variable to the whole domain or a semi-distributed approach may use different constants in predefined subdomains, e.g. catchments, of the glacier. In this study, we adopted a fully distributed approach that treated the friction as a log-Gaussian random field that is $\theta = \log(\beta)$, where $p(\theta) \sim \mathcal{N}(\mu, \mathcal{C})$; we set $\mu = 0$.

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

Following Isaac et al. (2015), we defined the covariance operator $\mathcal{C}$ to be a infinite-dimensional Laplacian squared operator. Specifically, we used the following finite-dimensional discretization of the operator

$$\Sigma_{\text{prior}}^{-1} = KM^{-1}K,$$

where K and M are finite elements matrices for the elliptic and mass operators, defined as

$$K_{ij} = \gamma \int_{\Gamma_g} \nabla\phi_i(x) \cdot \nabla\phi_j(x)dx + \delta \int_{\Gamma_g} \phi_i(x) \cdot \phi_j(x)dx + \eta \int_{\partial\Gamma_g} \phi_i(x) \cdot \phi_j(x)dx$$

and

$$M_{ij} = \int_{\Gamma_g} \phi_i(x) \cdot \phi_j(x)dx,$$

where ϕ_i are finite element basis functions. The first term in the definition of K is the Laplacian operator, the second term is a mass operator representing a source term, and the last term is a boundary mass operator for Robin boundary conditions. The ratio of the coefficients γ and δ determines the correlation length $l = \sqrt{\frac{\gamma}{\delta}}$ of the covariance. In our simulations, we set $\gamma = 2000$ km, $\delta = 2$ km^{-1} and $\eta = 20$, hence $l \approx 31.6$ km. These values were found to balance the smoothness of realizations of the friction field with the ability to capture the fine scale friction features needed to produce an acceptable match between the model prediction of surface velocity and the observed values. Two random samples from the prior distribution of the log-friction are depicted in Figure 3.

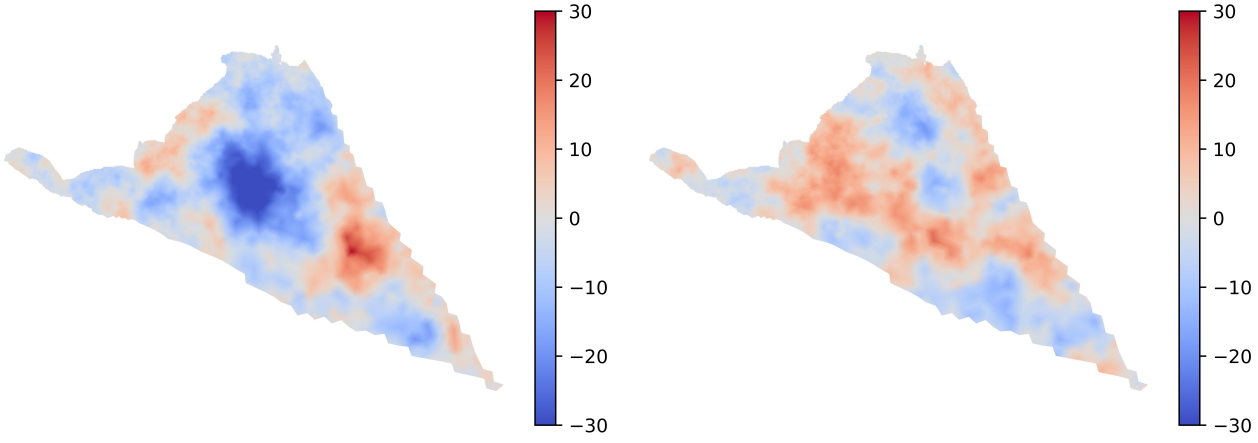


Figure 3. Two random samples from the prior distribution of the log-friction.

The parameterization of the prior we used has two main advantages. First, computationally efficient linear algebra can be used to draw samples from the prior distribution. In this study we drew samples from the prior using

$$\theta = \mu_{\text{prior}} + Ln, \quad n \sim \mathcal{N}(0, I)$$

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

with $\mu_{\text{prior}} = 0$, I is the identity matrix, $L = K^{-1}M^{\frac{1}{2}}$, such that $\Sigma_{\text{prior}} = LL^T$, and we lump the mass matrix M. The second advantage is that this prior enables an efficient procedure for computing the posterior distribution of the friction field on observations, which we present in Section 3.

2.4 Additional model setup

Additional details regarding the model setup are as follows. First, the glacier's bed topography, ice surface elevation, and ice thickness were obtained from observations (refer to Hillebrand et al. (2022) for details) and interpolated onto the finite element mesh. Second, the MIROC5 climate forcing from the CMIP5 for the Representative Concentration Pathway (RCP) 2.6 scenario was used to generate the surface mass balance (difference between ice accumulation and ablation) f_H and drive the ice-sheet evolution from 2007 to 2100. Finally, a fixed calving forcing was applied such that any ice that moved beyond the calving front is assumed to melt. Thus, any explicit ocean forcing is ignored.

3 Calibration

The goal of this study was to investigate uncertainty in predictions of the future mass change of Humboldt Glacier. However, generating realistic predictions with a model requires calibrating that model to available data. Consequently, in this paper we calibrated the basal friction field of our numerical models to measurements of surface velocity of the ice sheet. We processed Humboldt Glacier geometry data and surface velocity observations for year 2007 as detailed in Hillebrand et al. (2022). The data were assumed to represent the initial state of the ice sheet which was itself assumed to be in equilibrium. Thus, we calibrated the friction field by fitting the outputs of a high-resolution steady state flow model to the observational data.

Ice-sheet models are typically calibrated using deterministic optimization methods that find the values of the model parameters that lead to the best match between observations and the model prediction of the observations, e.g MacAyeal (1993); Morlighem et al. (2010); Petra et al. (2012); Perego et al. (2014); Goldberg et al. (2015). However, such approaches ignore the uncertainty in the model parameters due to using a finite amount of noisy observational data. Thus, in this paper we used Bayesian inference to calibrate the model.

Bayesian inference uses Bayes' Theorem to quantify the probability of the parameters conditioned on the data $p(\theta \mid y)$, known as the posterior distribution, as proportional to the conditional probability of observing the data given the parameters $p(y \mid \theta)$, known as the likelihood distribution, multiplied by the prior probability assigned to the parameters $p(\theta)$,

$$p(\theta \mid y) \propto p(y \mid \theta)p(\theta).$$

In this work we assumed that the observational data (surface velocities $y = u_{\text{obs}}$), were corrupted by centered Gaussian noise $\eta \sim \mathcal{N}(0, \Sigma_{\text{noise}})$. Specifically, given a Blatter-Pattyn flow model $g(\theta)$ that maps the logarithm of the basal friction to the computed surface velocity, we assumed $y = g(\theta) + \eta$ such that the likelihood function was given by

$$p(y \mid \theta) = (2\pi|\Sigma_{\text{noise}}|)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(y - g(\theta))^\top \Sigma_{\text{noise}}^{-1}(y - g(\theta))\right).$$

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

Here, $g(\theta)$ denotes the output of the steady-state ice-sheet model at the locations of the observations for a given realization of the model parameters. We also assumed that the observations were uncorrelated and set

$$\Sigma_{\text{noise}} = \alpha\,\text{diag}(u_{\text{rms}})M_s^{-1}\text{diag}(u_{\text{rms}}),$$

where $\text{diag}(u_{\text{rms}})$ is the diagonal matrix containing the root mean square errors u_{rms} of the surface velocities, and M_s is the mass matrix computed on the upper surface Γ_s and α is a scaling term. We set $\alpha = 0.125$ km^2.

Quantifying uncertainty in mass-change projections conditioned on observational data requires drawing samples from the posterior, evaluating the transient model at each sample and computing estimates of statistics summarizing the prediction uncertainty using those evaluations. Typically, samples are drawn using Markov Chain Monte Carlo (Hoffman and Gelman, 2014), however such methods can be computationally intractable for high-dimensional uncertain variables (Bui-Thanh et al., 2013), such as the variable we used to parameterize basal friction. Consequently, we used the two-step method presented in Bui-Thanh et al. (2013); Isaac et al. (2015) to construct a Laplace approximation of the posterior.

First, we performed a PDE-constrained deterministic optimization to compute the maximum a posteriori (MAP) point θ_{MAP}

$$\theta_{\text{MAP}} = \underset{\theta}{\text{argmin}}\,\frac{1}{2}(y-g(\theta))^\top\Sigma_{\text{noise}}^{-1}(y-g(\theta)) + \frac{1}{2}(\theta-\mu_{\text{prior}})^\top\Sigma_{\text{prior}}^{-1}(\theta-\mu_{\text{prior}}), \tag{9}$$

which maximizes the posterior $p(\theta \mid y)$. For linear models and Gaussian priors, the MAP point has close ties with the optimal solution obtained using Tikhonov regularization (Stuart, 2010). Specifically, the first term above minimizes the difference between the model predictions and the observations and the second term penalizes the deviation of the optimal point from the prior mean.

Second, we constructed a low-rank quadratic approximation of the log posterior, centered at the MAP point

$$\log(p(\theta \mid y)) \approx C - \frac{1}{2}(\theta-\theta_{\text{MAP}})^\top\Sigma_{\text{post}}^{-1}(\theta-\theta_{\text{MAP}}),$$

where $\Sigma_{\text{post}}^{-1} = H_{\text{MAP}} + \Sigma_{\text{prior}}^{-1}$ and H_{MAP} is the Hessian of $\frac{1}{2}(y-g(\theta))^\top\Sigma_{\text{noise}}^{-1}(y-g(\theta))$ at $\theta = \theta_{\text{MAP}}$ and C is a constant independent of θ. This resulted in a Gaussian approximation of the posterior $p(\theta \mid y) \approx \mathcal{N}(\theta_{\text{MAP}}, \Sigma_{\text{post}})$, also known as a Laplace approximation of the posterior. Naively computing the posterior covariance using the aforementioned formula for $\Sigma_{\text{post}}^{-1}$ is computationally intractable for the high-dimensional variable we used to parameterize basal friction. Consequently, we used a low-rank Laplace approximation which is detailed in Appendix A.

4 Uncertainty quantification

This study investigated the efficacy of using MFUQ to compute the uncertainty in predictions of future mass change from Humboldt Glacier. We defined mass change to be the difference between the final mass[3] of the ice sheet at $t = 2100$ and the initial mass of the ice sheet at $t = 2007$. While the mass change is a functional of the ice-sheet thickness H, for simplicity the following discussion simply refers to the mass change as a function of only the model inputs, that is $f_\alpha(\theta)$, where α indexes the

[3]In this work, we compute the ice-sheet mass considering only the ice above flotation, which is what contributes to sea-level change

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

model fidelity that was used to simulate the ice sheet. Previous UQ studies computed statistics summarizing the uncertainty in ice-sheet predictions, such as mean and variance, using *single-fidelity* Monte Carlo (SFMC) quadrature, that is MC quadrature applied to a single physics model with a fixed numerical discretization α, for example Ritz et al. (2015); Schlegel et al. (2018). However, in this study we used MFUQ to reduce the computational cost of quantifying uncertainty. Specific details on the MFUQ methods investigated are given in the following subsections.

4.1 Single-fidelity Monte Carlo quadrature

SFMC quadrature is a highly robust procedure, which could be used to compute the mean Q^{μ} and variance Q^{σ^2} of the Humboldt glacier mass change predicted by a single model using a three step procedure. The first step randomly samples N realizations of the the model inputs $\Theta = \{\theta^{(m)}\}_{n=1}^{N}$ from their posterior distribution. The second step simulates the model at each realization of the random variable (friction field) and computed the mass change at the final time $f_{\alpha}^{(n)} = f_{\alpha}(\theta^{(n)})$. The third step approximates the mean and variance using the unbiased estimators

$$Q_{\alpha}^{\mu}(\Theta) = N^{-1}\sum_{n=1}^{N} f_{\alpha}^{(n)} \qquad Q_{\alpha}^{\sigma^2}(\Theta) = (N-1)^{-1}\sum_{n=1}^{N}\left(f_{\alpha}^{(n)} - Q_{\alpha}^{\mu}(\Theta)\right)^2. \tag{10}$$

MC estimators converge to the true mean and variance of f_{α} as the number of samples tends to infinity, but using a finite number of samples N introduces an error into the MC estimator that depends on the sample realizations used to compute the estimators. Consequently, any MC estimator Q is a random variable and the mean-squared error (MSE)

$$\mathbb{E}\left[\left(Q_{\alpha}(\Theta) - Q\right)^2\right] = \mathbb{E}\left[\left(Q_{\alpha}(\Theta) - \mathbb{E}\left[Q_{\alpha}(\Theta)\right]\right)^2 + \left(\mathbb{E}\left[Q_{\alpha}(\Theta)\right] - Q\right)^2\right] = \underbrace{\mathbb{V}\left[Q_{\alpha}(\Theta)\right]}_{I} + \underbrace{\left(\mathbb{E}\left[Q_{\alpha}(\Theta)\right] - Q\right)^2}_{II} \tag{11}$$

is typically used to quantify the expected performance of an MC estimator.

The MSE of an MC estimator (11) consists of two terms referred to as the estimator variance (I) and the estimator bias (II). The variance of an MC estimator comes from using a finite number of samples and decreases as the number of samples increases. For example, the variances of the estimators of mean and variance are, respectively,

$$\mathbb{V}\left[Q_{\alpha}^{\mu}(\Theta)\right] = \frac{1}{N}\mathbb{V}\left[f_{\alpha}\right], \quad \mathbb{V}\left[Q_{\alpha}^{\sigma^2}(\Theta)\right] = \frac{1}{N}\left(2(N-1)^{-1}\mathbb{V}\left[f_{\alpha}\right]^2 + \mathbb{C}\mathrm{ov}\left[\left(f_{\alpha} - \mathbb{E}\left[f_{\alpha}\right]\right)^2, \left(f_{\alpha} - \mathbb{E}\left[f_{\alpha}\right]\right)^2\right]\right), \tag{12}$$

where the variance and covariances involving f_{α} are exact statistics of the model, which are typically unknown. The bias term of the MSE (11) is caused by using a numerical model, with inadequacy and discretization errors, to compute the mass change.

Constructing a SFMC estimator with a small MSE (10) ensures for any set of model inputs samples the value of the estimator will be likely close to the true value. However, constructing an unbiased estimator with a small MSE using a computationally expensive high-fidelity model is computationally demanding because the contribution of the variance term to the MSE only decreases linearly with the number of samples. In contrast, N can be significantly increased if a cheaper, lower-fidelity model is used but the corresponding decrease in the estimator variance will be offset by an increase in its bias. Consequently, the bias and variance of any estimator should be balanced, but most SFMC analyses do not consider this trade-off explicitly when choosing the fidelity of the model used. In the following section we detail how to use MFUQ to substantially improve the accuracy of estiamted statistics for a fixed computational cost.

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

4.2 Multi-fidelity uncertainty quantification

MFUQ leverages the correlation between models of varying cost and accuracy to reduce the computational cost of constructing MC estimators with a given MSE. While various multi-fidelity estimators have been developed, this study used approximate control variate (ACV) estimation (implemented in PyApprox (Jakeman, 2023)), which include most existing estimators, including Multi-level Monte Carlo (MLMC) (Giles, 2015) and Multi-fidelity Monte Carlo (MFMC) (Peherstorfer et al., 2016), as special cases. [4] In this section we describe how to construct an ACV estimate of the mean of a model using one low-fidelity model. We then introduce the ACV procedure we used to compute the mean and variance of our our highest-fidelity ice-sheet model using an ensemble of 13 models.

4.2.1 Two model approximate control variate Monte Carlo

Given a high-fidelity model $f_0(\theta)$ and a low-fidelity model $f_1(\theta)$, an ACV estimator of the mean of the high-fidelity model is

$$Q_{\text{ACV}}(\Theta_0,\Theta_1) = \sum_{n=1}^{N_0} f_0(\theta_0^{(n)}) + \eta \sum_{n=1}^{N_0}\left(f_1(\theta_0^{(n)}) - \sum_{n=1}^{N_1} f_1(\theta_1^{(n)})\right),$$

where η is a constant that can be optimized to reduce the MSE of the estimator. Using such an estimator to approximate the high-fidelity mean requires two different sets of model evaluations. These evaluations are obtained by first drawing two sets of samples $\Theta_0 = \{\theta_0^{(n)}\}_{n=1}^{N_0}, \Theta_1 = \{\theta_1^{(n)}\}_{n=1}^{N_1}$ from the distribution of the random variables; in our study we drew random samples from the posterior distribution of θ. The high-fidelity model is then evaluated on all the samples in Θ_0 and the low-fidelity model is evaluated on both the sets Θ_0 and Θ_1.

The ACV estimator is an unbiased estimator of the mean high-fidelity model. So the MSE is equal to the variance of the estimator, which setting $N_1 = rN_0, r > 1$ is always

$$\mathbb{V}\left[Q^{\mu}_{\text{ACV}}(\Theta_0,\Theta_1)\right] = \mathbb{V}\left[Q^{\mu}_0(\Theta_0)\right]\left(1 - \frac{r-1}{r}\mathbb{C}\text{orr}[f_0,f_1]^2\right)$$

Thus, for a fixed sample ratio r, if the high and low-fidelity models are highly correlated then the ACV estimator will be much more accurate than the SFMC estimator; see (12). Moreover, the value of r can be optimized to minimize the error of an ACV estimator given a fixed computational budget. In the following section, we provide more details on how to construct ACV estimator using more than one-low fidelity model, including information on how to optimize η and the number of samples used to evaluate each model.

4.2.2 Many model approximate control variate Monte Carlo

Given an ensemble of $M+1$ models $\{f_\alpha(\theta)\}_{\alpha=0}^{M}$ an ACV MC can be used to compute a vector-valued estimator $\mathbf{Q}_0 = [Q_{0,1},\cdots,Q_{0,K}]^\top \in \mathbb{R}^K$ of statistics of the highest fidelity model f_0; the specific instances of the ice-sheet models used by

[4] Multilevel Best Linear Unbiased Estimators] Recently, multilevel best linear unbiased estimators (MLBLUEs) Schaden and Ullmann (2020) were developed as an alternative to ACV estimators to estimate the expectation of a high-fidelity model using an ensemble of models of varying cost and accuracy. However, we did not use MBLUEs in this study because they can only be used to estimate the mass-change mean and not its variance.

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

this study are presented in Section 5.1. The vector $\mathbf{Q}_0$ may be comprised of a single type of statistic computed for multiple QoI, multiple statistics of a single QoI, or a combination of both. For example, in this study we computed the ACV estimator of the mean and variance of the mass change, that is $\mathbf{Q}_0 = [Q_0^{\mu}, Q_0^{\sigma^2}]^\top \in \mathbb{R}^2$.

The ACV estimator of a vector of high-fidelity statistics has the form

$$\mathbf{Q}_{\text{ACV}}(\Theta_0, \Theta_1^*, \Theta_1, \ldots, \Theta_M^*, \Theta_M) = \mathbf{Q}_0(\Theta_0) + \begin{bmatrix} \eta_{1,1} & \cdots & \eta_{1,KM} \\ \eta_{2,1} & \cdots & \eta_{2,KM} \\ \vdots & & \\ \eta_{K,1} & \cdots & \eta_{K,KM} \end{bmatrix} \begin{bmatrix} \mathbf{Q}_1(\Theta_1^*) - \mathbf{Q}_1(\Theta_1) \\ \mathbf{Q}_2(\Theta_2^*) - \mathbf{Q}_2(\Theta_2) \\ \vdots \\ \mathbf{Q}_M(\Theta_M^*) - \mathbf{Q}_M(\Theta_M) \end{bmatrix}$$

or in more compact notation

$$\mathbf{Q}_{\text{ACV}}(\Theta_{\text{ACV}}) = \mathbf{Q}_0(\Theta_0) + \boldsymbol{\eta}\boldsymbol{\Delta}(\Theta_\Delta), \qquad \boldsymbol{\Delta}(\Theta_\Delta) = \begin{bmatrix} \boldsymbol{\Delta}_1(\Theta_1^*, \Theta_1) \\ \vdots \\ \boldsymbol{\Delta}_M(\Theta_M^*, \Theta_M) \end{bmatrix} \in \mathbb{R}^{KM \times 1}, \qquad \boldsymbol{\eta} \in \mathbb{R}^{K \times KM}, \tag{13}$$

where the entries of $\boldsymbol{\eta}$ are called control variate weights, $\Theta_\Delta = \{\Theta_1^*, \Theta_1, \ldots, \Theta_M^*, \Theta_M\}$, and $\Theta_{\text{ACV}} = \{\Theta_0, \Theta_\Delta\}$. This estimator is constructed by evaluating each low-fidelity model at two sets of samples $\Theta_\alpha^* = \{\theta^{(n)}\}_{n=1}^{N_{\alpha^*}}$ and $\Theta_\alpha = \{\theta^{(n)}\}_{n=1}^{N_\alpha}$ where some samples may be shared between sets such that in some cases $\Theta_\alpha^* \cup \Theta_\beta \neq \emptyset$. The highest-fidelity model is only evaluated on one set of samples.

By construction any ACV estimator is an unbiased estimator of $\mathbf{Q}_0$ because $\mathbb{E}[\boldsymbol{\Delta}_\alpha] = 0, \alpha > 0$. Consequently, the MSE of the ACV estimator can be minimized by solely optimizing the determinant of the estimator covariance matrix; when estimating a single statistics ($K = 1$) this is equivalent to minimizing the variance of the estimator. Given sample sets Θ_{ACV}, the determinant of the covariance of an ACV estimator can be minimized using the optimal weights

$$\boldsymbol{\eta}(\Theta_{\text{ACV}}) = -\mathbb{C}\text{ov}[\boldsymbol{\Delta}, \boldsymbol{\Delta}]^{-1}\,\mathbb{C}\text{ov}[\boldsymbol{\Delta}, \mathbf{Q}_0], \tag{14}$$

which produces an ACV estimator with covariance

$$\mathbb{C}\text{ov}[\mathbf{Q}_{\text{ACV}}, \mathbf{Q}_{\text{ACV}}](\Theta_{\text{ACV}}) = \mathbb{C}\text{ov}[\mathbf{Q}_0, \mathbf{Q}_0] - \mathbb{C}\text{ov}[\mathbf{Q}_0, \boldsymbol{\Delta}]\,\mathbb{C}\text{ov}[\boldsymbol{\Delta}, \boldsymbol{\Delta}]^{-1}\,\mathbb{C}\text{ov}[\mathbf{Q}_0, \boldsymbol{\Delta}]^\top, \tag{15}$$

where the dependence of $\boldsymbol{\Delta}$ and $\mathbf{Q}_0$ on the sample sets Θ_Δ and Θ_0 was dropped to improve readability.

Remark 4.1 (Pilot and exploitation phases)**.** *The approximation of model statistics using ACV estimators is broken into two steps. The first step, which we refer to as the pilot study, involves collecting evaluations of each model on a common set of samples, using these evaluations to compute the so called pilot statistics in* (14) *and* (15)*, and using these statistics to find the optimal sample allocation of the best estimator (see Algorithm 1). The second step involves evaluating each model according to the optimal sample allocation and then computing the model statistics using* (13)*. We call this second step the exploitation phase.*

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

4.2.3 Pilot Study

4.2.4 Estimating pilot statistics.

Computing the covariance of an ACV estimator requires estimates of the covariance between the estimator discrepancies $\mathbf{\Delta}$ with each other and the high-fidelity estimator and the covariance of the high-fidelity estimator which, for our case study, have the form

$$\mathbb{C}\mathrm{ov}\left[\mathbf{\Delta}_\alpha, \mathbf{\Delta}_\beta\right](\Theta_{\mathrm{ACV}}) = \begin{bmatrix} \mathbb{C}\mathrm{ov}\left[\Delta_\alpha^\mu, \Delta_\beta^\mu\right] & \mathbb{C}\mathrm{ov}\left[\Delta_\alpha^\mu, \Delta_\beta^{\sigma^2}\right] \\ \mathbb{C}\mathrm{ov}\left[\Delta_\beta^{\sigma^2}, \Delta_\alpha^\mu\right] & \mathbb{C}\mathrm{ov}\left[\Delta_\alpha^{\sigma^2}, \Delta_\beta^{\sigma^2}\right] \end{bmatrix}, \mathbb{C}\mathrm{ov}\left[\mathbf{Q}_0, \mathbf{\Delta}_\alpha\right](\Theta_{\mathrm{ACV}}) = \begin{bmatrix} \mathbb{C}\mathrm{ov}\left[Q_0^\mu, \Delta_\alpha^\mu\right] & \mathbb{C}\mathrm{ov}\left[Q_0^\mu, \Delta_\alpha^{\sigma^2}\right] \\ \mathbb{C}\mathrm{ov}\left[Q_0^{\sigma^2}, \Delta_\alpha^\mu\right] & \mathbb{C}\mathrm{ov}\left[Q_0^{\sigma^2}, \Delta_\alpha^{\sigma^2}\right] \end{bmatrix}, \tag{16}$$

$$\mathbb{C}\mathrm{ov}\left[\mathbf{Q}_0, \mathbf{Q}_0\right](\Theta_0) = \begin{bmatrix} \mathbb{V}\left[Q_0^\mu\right] & \mathbb{C}\mathrm{ov}\left[Q_0^\mu, Q_0^{\sigma^2}\right] \\ \mathbb{C}\mathrm{ov}\left[Q_0^{\sigma^2}, Q_0^\mu\right] & \mathbb{V}\left[Q_0^{\sigma^2}\right] \end{bmatrix} \in \mathbb{R}^{2\times 2}, \tag{17}$$

respectively, where $\Delta_{\alpha,1} = \Delta_\alpha^\mu = Q_\alpha^\mu(\Theta_\alpha^*) - Q_\alpha^\mu(\Theta_\alpha)$ and $\Delta_{\alpha,2} = \Delta_\alpha^{\sigma^2} = Q_\alpha^{\sigma^2}(\Theta_\alpha^*) - Q_\alpha^{\sigma^2}(\Theta_\alpha)$ were computed using the expressions in (10). Again the dependence of $\mathbf{\Delta}$ and $\mathbf{Q}_0$ on the sample sets Θ_Δ and Θ_0 was dropped to improve readability.

In practice, these quantities, which we call *pilot statistics*, are unknown and must be estimated with a pilot study. Specifically, following standard practice, we used MC quadrature with P, so-called, *pilot samples* $\Theta_{\mathrm{pilot}} = \{\theta^{(p)}\}_{p=1}^P$ to compute (16) and (17). For example, we approximated $\mathbb{C}\mathrm{ov}\left[f_\alpha, f_\beta\right]$, needed to compute $\mathbb{C}\mathrm{ov}\left[Q_0^\mu, \Delta_\alpha^\mu\right]$, and $\mathbb{C}\mathrm{ov}\left[\Delta_\alpha^{\sigma^2}, \Delta_\beta^{\sigma^2}\right]$, by

$$\mathbb{C}\mathrm{ov}\left[f_\alpha, f_\beta\right] \approx P^{-1}\sum_{p=1}^{P}\left(f_\alpha(\theta^{(p)}) - Q_\alpha^\mu(\Theta_{\mathrm{pilot}})\right)\left(f_\beta(\theta^{(p)}) - Q_\beta^\mu(\Theta_{\mathrm{pilot}})\right) \tag{18}$$

Similarly, the pilot samples were used to estimate

$$\mathbb{C}\mathrm{ov}\left[(f_\alpha - \mathbb{E}\left[f_\alpha\right])^2, (f_\beta - \mathbb{E}\left[f_\beta\right])^2\right] \text{ and } \mathbb{C}\mathrm{ov}\left[f_\alpha, (f_\beta - \mathbb{E}\left[f_\beta\right])^2\right], \tag{19}$$

which are needed to compute the other quantities in (16) and (17). Please refer to Dixon et al. (2023) to see exactly how these quantities are used to compute the covariance blocks of $\mathbb{C}\mathrm{ov}\left[\Delta_\alpha, \Delta_\beta\right]$ and $\mathbb{C}\mathrm{ov}\left[Q_0, \Delta_\alpha\right]$. Finally, we recorded the CPU time needed to simulate each model at all pilot samples and set the model costs $w^\top = [w_0, w_1, \ldots, w_M]$ to be the the median simulation time of each model.

Unfortunately, using a finite P introduces errors into (16) and (17), which in turn induces error in the ACV estimator covariance. This error can be decreased by using a large P but this would require additional evaluations of expensive numerical models, which we were trying to avoid. Consequently, in this study we investigated the sensitivity of the number of pilot samples on the accuracy of ACV MC estimators.

4.2.5 Optimal computational resource allocation.

The variance of an ACV estimator is dependent on how samples are allocated to the sets $\Theta_\alpha, \Theta_\alpha^*$, which we call the sample allocation $\mathcal{A}$. Specifically, $\mathcal{A}$ specifies: the number of samples in the sets $\Theta_\alpha, \Theta_\alpha^*, \forall\alpha$, denoted by N_α and N_{α^*}, respectively; the

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

number of samples in the intersections of pairs of sets, that is $N_{\alpha\cap\beta} = |\Theta_\alpha \cap \Theta_\beta|$, $N_{\alpha^*\cap\beta} = |\Theta^*_\alpha \cap \Theta_\beta|$, $N_{\alpha^*\cap\beta^*} = |\Theta^*_\alpha \cap \Theta^*_\beta|$; and the number of samples in the union of pairs of sets $N_{\alpha\cup\beta} = |\Theta_\alpha \cup \Theta_\beta|$ and similarly $N_{\alpha^*\cup\beta}$, $N_{\alpha^*\cup\beta^*}$. Thus, the best ACV estimator can be theoretically found by solving the constrained non-linear optimization problem

$$\min_{\mathcal{A}\in\mathbb{A}} \mathrm{Det}\left[\mathbb{C}\mathrm{ov}\left[Q_{\mathrm{ACV}}, Q_{\mathrm{ACV}}\right]\right](\mathcal{A}) \qquad \text{s.t.} \qquad W(w,\mathcal{A}) \le W_{\max}, \tag{20}$$

In the above equation, $\mathbb{A}$ is the set of all possible sample allocations and the constraint ensures that the computational cost of computing the ACV estimator

$$W(w,\mathcal{A}) = \sum_{\alpha=0}^{M} N_{\alpha^*\cup\alpha} w_\alpha$$

is smaller than a computational budget $W_{\max}$, where $w^\top = [w_0, w_1, \ldots, w_M]$ are the computational costs of evaluating each model once.

Unfortunately, a tractable algorithm for solving (20) has not yet been developed in large part due to the extremely high number of possible sample allocations in the set $\mathbb{A}$. Consequently, various ACV estimators have been derived in the literature that simplify the optimization problem, by specifying what we call the sample structure $\mathcal{S}$ that restricts how samples are shared between the sets $\Theta_\alpha, \Theta^*_\alpha$, which in turn limits the size of the search space $\mathbb{A}$.

MLMC (Giles, 2015) and MFMC (Peherstorfer et al., 2016) are two popular examples of ACV estimators proposed in the literature.[5] These methods restrict $\mathbb{A}$ such that (20) can be optimized analytically when the estimators are used to estimate the mean of a single model output. More recently, so called ACVMF and ACVIS (Gorodetsky et al., 2020) and tunable versions of these estimators (Bomarito et al., 2022), all which propose alternative restrictions on the set $\mathbb{A}$. Figure 4 summarizes the restrictions on the sample allocations placed on $\mathbb{A}$ by MLMC, MFMC, and ACVMF.

The performance of different ACV estimators is problem dependent. Consequently, in this paper we investigated the use of a large number of different ACV estimators from the literature. For each estimator we used the general purpose numerical optimization algorithm proposed in Bomarito et al. (2022) to find the optimal sample allocations that minimize the determinant of the estimator covariance.[6][7]

4.2.6 Model and estimator selection.

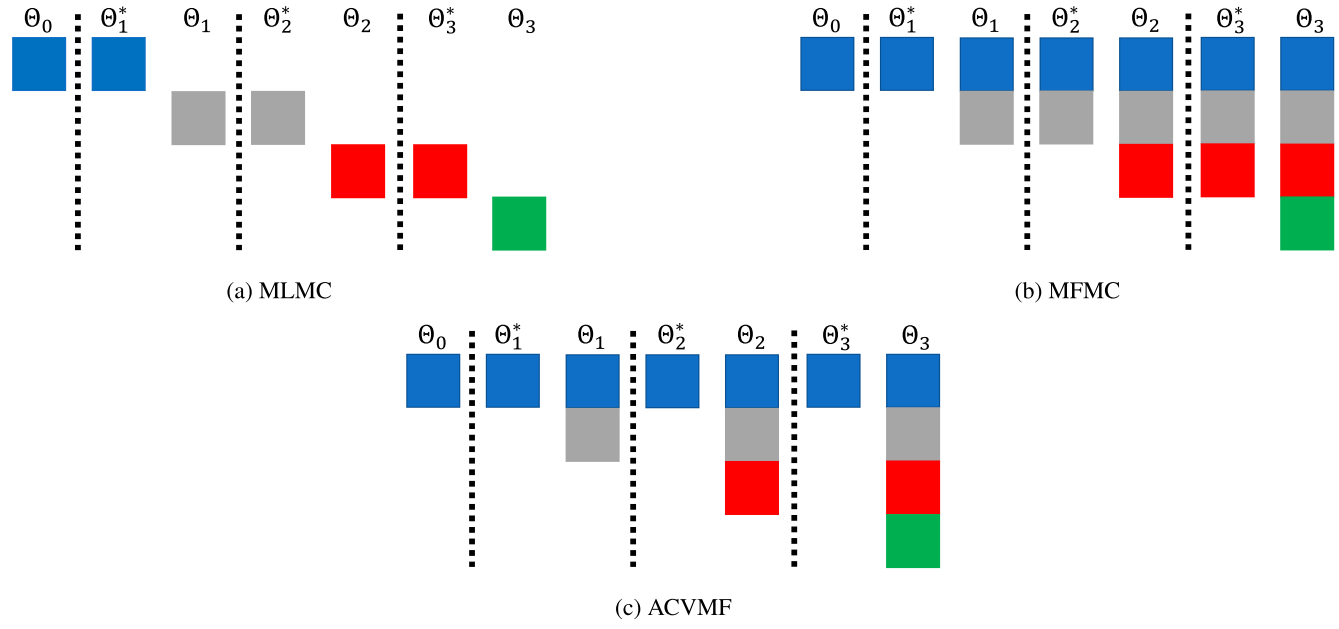
The relative performance of existing estimators is problem dependent. Moreover, constructing ACV estimators with data from all available models may produce an estimator with larger variance than an estimator that is only constructed using a subset of the available models. Consequently, it is difficult to determine the best estimator a priori. However, we can accurately predict

[5] MLMC estimators set all the control variate weights in (13) to $\boldsymbol{\eta} = -1$. Refer to Gorodetsky et al. (2020) for more details on the connections between ACV and MLMC.

[6] The presentation of the optimization algorithms in (Bomarito et al., 2022) focuses on the estimation of a single statistic, but can be trivially be extended to the vector-valued QoI considered here.

[7] Some approaches admit analytical solutions to (20) when minimizing the estimator variance of the mean of a model, e.g. MLMC and MFMC, but such solutions produce sample allocations that are not guaranteed to minimize the estimator variance of other statistics such as variance and thus were not used here.

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024



(a) MLMC (b) MFMC

(c) ACVMF

Figure 4. The fundamental sample allocation structures used by different ACV methods. These structures when combined with $N_\alpha, \alpha = 0, ..., N$ define the the admissible sample allocations $\mathcal{A} \in \mathbb{A}$. Each different colored box represents a set of independent samples. Each ACV sample set Θ_α or Θ_α^* may consist of multiple independent sample sets. The difference between methods arises from restrictions on how the independent sample sets are shared among the sets Θ_α or Θ_α^*.

the relative performance of any ACV estimator using only the model simulations run during the pilot study. Thus, in this study we enumerated a large set of estimators types encoded by the set $\mathbb{A}$ and model models subsets and chose the estimator with the smallest estimator covariance determinant.

Algorithm 1 summarizes the procedure we use to choose the best ACV estimator. Line 8 loops over all model subsets $\mathcal{S}$. In this study, we enumerated all permutations of the sets of models that contained the high-fidelity model and at most 3 low-fidelity models. Line 10 enumerates each parametrically defined estimator E. We enumerated the large sets of parametrically defined generalized multi-fidelity (GMF), generalized independent sample (GIS) and generalized recursive difference (GRD) ACV estimators; these sets of estimators include ACVMF, MFMC, MLMC (with optimized control variate weights) as special cases. For each estimator E and model subset $\mathcal{S}$, line 12 was used to find the optimal sample allocation $\mathcal{A}_E$, using the pilot values $\{f_\alpha(\Theta_{\text{pilot}})\}_{\alpha \in \mathcal{S}}$ when minimizing (20). Lines 13-16 was used to record the best estimator found at each iteration of the outerloops.

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

Algorithm 1 Estimator selection

Input
$\{f_\alpha(\Theta_{\text{pilot}})\}_{\alpha=0}^{M}$ Pilot evaluations of each model
Output
$\mathcal{A}_{\text{best}}$ Best estimator sample allocation
J_{best} Best estimator objective

$J_{\text{best}} \leftarrow \infty$
▷ Loop over all low-fidelity model subsets
for $\mathcal{S} \subseteq \{1, \dots, M\}$ **do**
  ▷ Loop over all MF estimators, e.g MLMC, MFMF, ACVMF, etc
  for $E \in \mathcal{E}$ **do**
    ▷ Compute the optimal estimator objective J_E and sample allocation $\mathcal{A}_E$ for the current estimator and subset of models
    $J_E, \mathcal{A}_E \leftarrow$ Solve (20) using $\mathbb{A}_{E,\mathcal{S}}$ and $\{f_\alpha(\Theta_{\text{pilot}})\}_{\alpha\in\mathcal{S}}$
    if $J_E < J_{\text{best}}$ **then**
      ▷ Update the best estimator
      $\mathcal{A}_{\text{best}} \leftarrow \mathcal{A}_E$
      $J_{\text{best}} \leftarrow J_E$
    end if
  end for
end for

5 Results

This section presents the results of our MFUQ study. First, we describe the ensemble of numerical models we used to solve the governing equations presented in Section 5.1. Second, we summarize the results of our Bayesian model calibration. Third, we present the results of our pilot study, including a comparison of the computational costs of each model and their SFMC estimates of the mean and variance of the mass change computed using the pilot samples in Section 5.3 and the impact of increasing the number of pilot samples in Section 5.4. Finally, we quantify the improvement in the accuracy of MFUQ estimates of mass-change statistics relative to SFMC in Section 5.5. All results were generated with the PyApprox software package (Jakeman, 2023).

5.1 Multi-fidelity model ensemble

In this study we investigated the use of 13 different models of varying computational cost and accuracy to compute ice-sheet mass change. Specifically, we used MFUQ to estimate the mean and variance of a highly-resolved finite element model using an an ensemble of 12 low-fidelity models. We compactly denote the fidelity of each model using the notation

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

PHYSICSNAME$_{dx,dt}$, where PHYSICSNAME refers to the governing equations solved, dx denotes the size of the representative spatial element and dt the size of the time-step.

The highest-fidelity model we considered in this study was a MOLHO-based model denoted by MOLHO$^*_{1km,9days}$ where the star indicates that the model was modified to conserve mass. The low-fidelity model ensemble consisted of four MOLHO-based low-fidelity models, MOLHO$_{1km,36days}$, MOLHO$_{1.5km,36days}$, MOLHO$_{2km,36days}$, MOLHO$_{3km,36days}$, and eight SSA-based low-fidelity models SSA$_{1km,36days}$,SSA$_{1.5km,36days}$, SSA$_{2km,36days}$, SSA$_{3km,36days}$, SSA$_{1km,365days}$, SSA$_{1.5km,365days}$, SSA$_{2km,365days}$, SSA$_{3km,365days}$. No low-fidelity model enforced the conservation of mass.

5.2 Bayesian model calibration

In this study we used the MALI ice-sheet code (see Hoffman et al. (2018); Tezaur et al. (2021)) to calibrate the basal friction field on the finest mesh, as described in Section 3. The MAP point of the posterior, determined using (9), is depicted in the left panel of Figure 5. The pointwise variance of the posterior of the log-friction is depicted in the right-panel of Figure 5. When this figure is compared to the pointwise variance of the prior depicted in the center panel of Figure 5, it is clear that conditioning the prior uncertainty on the surface velocity significantly reduced the uncertainty in the basal friction field. This conclusion is further corroborated by Figure 6 which compares a random sample from the prior and a random sample from the posterior. The minimum and maximum values of the posterior sample of the log-friction are much smaller than the same bounds of the prior sample. However, the posterior sample has much higher-frequency content because the data only informed the lower-frequency modes of the friction field.

Figure 5. (*Left*) Log of the basal friction MAP point. (*Center*) Log of the prior variance. (*Right*) Log of the posterior variance.

To demonstrate a projection to 2100 using a calibrated model, figure 7 depicts the difference between the final (year 2100) and initial (year 2007) ice thickness and the final surface velocity of Humboldt Glacier computed by the highest-fidelity model (MOLHO$^*_{1km,9days}$) for a random posterior realization of the basal friction field. The final ice thickness shown differs substantially from the initial thickness with thickness decreasing substantially at lower elevations of the glacier in the ablation zone where increasingly negative surface mass balance occurs through 2100. In general, the glacier speeds up as it thins, with the largest speedup occurring in the region of fast flow in the north where basal friction is small.

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

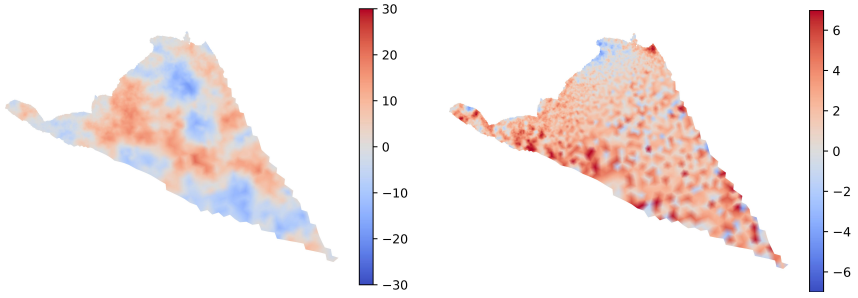


Figure 6. (*Left*) A random samples from the prior distribution of the log-friction. (*Right*) A random sample from the posterior.

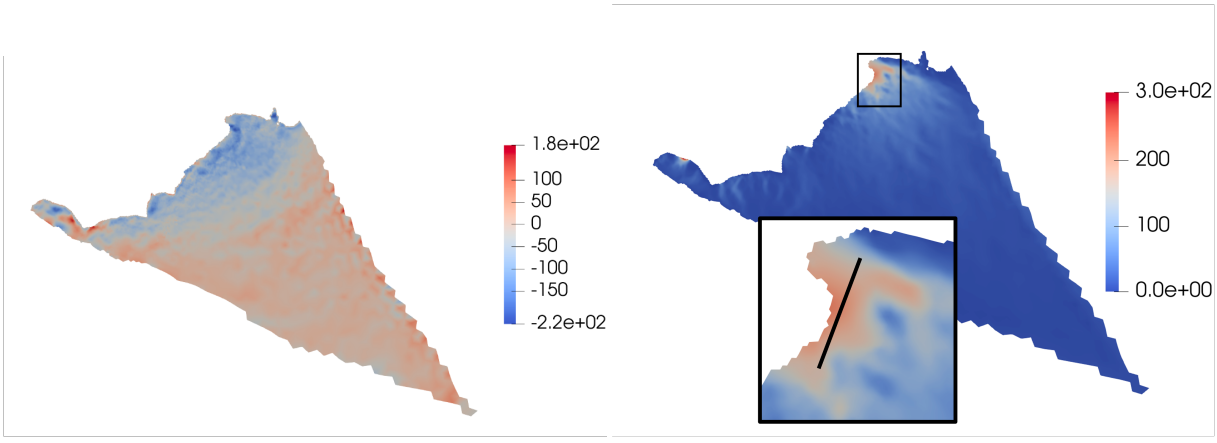


Figure 7. (*Left*) The difference between the final and initial ice thickness in meters and (*Right*) the surface velocity of Humboldt Glacier. The left box is a zoomed in picture of the top right tip of the glacier. The black line in the inset was used to plot cross-sections of the thickness and friction profiles at 2100 in a region with high velocities (see Figure 13). Both the left and right figures were generated using the highest-fidelity model $\mathrm{MOLHO}^*_{1km,9days}$ evaluated at one random realization of the posterior of the basal friction field.

Remark 5.1 (Prior sensitivity)**.** *In this study we used our domain experience to determine the best values of the prior hyper-parameters γ, δ, η reported in Section 2.3. However, varying the hyper-parameters of the prior, would likely change the estimates of uncertainty in ice-sheet predictions produced by this study. Similar to previous studies (Isaac et al., 2015), we did not investigate these sensitivities extensively. However, we did find that increasing δ and γ substantially from the value we ultimately used, prevented the MAP point from capturing the high-frequency content of the basal friction field needed to accurately match the observed surface velocities. Future studies should investigate the sensitivity of mass change to the values of the prior hyper-parameters.*

Remark 5.2 (Interpolating basal friction)**.** *In this study we drew samples from the posterior distribution of the friction parameters defined on the finest spatial mesh. However, a posterior sample defined on the fine mesh cannot be used to predict mass change with a low-fidelity model defined on a coarser mesh. Consequently, before using a low-fidelity model with a coarse*

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

mesh to predict mass change, we first interpolated each sample of the posterior distribution of the basal friction field defined on the finest mesh onto the mesh used by the low-fidelity model.

5.3 Initial Pilot study

This section details the results of the pilot study that we used to obtain the computational cost and the pilot statistics needed to construct ACV estimators. First, we evaluated each of our 13 models at 20 random pilot samples of the model inputs Θ_{pilot}, i.e. 20 different basal friction fields. Second we computed the median computational cost (wall-time) required to solve each model at one pilot sample. The median computational costs are plotted in the top panel of Figure 8 and the total cost of the evaluating the models was approximately 144 hours. Third, using the pilot samples, we computed the SFMC estimators of the mean and standard deviation of the mass change predicted by each of the 13 models considered by this study. The middle and lower panels of Figure 8 show that there are significant differences between the means and standard deviations of each model. However in the next section, we show that despite the statistical differences between models and the differences between the ice-evolution predicted by each model (see Figure 13), MFUQ can be used to effectively increase the accuracy of the mean and variance of the mass change, relative to SFMC.

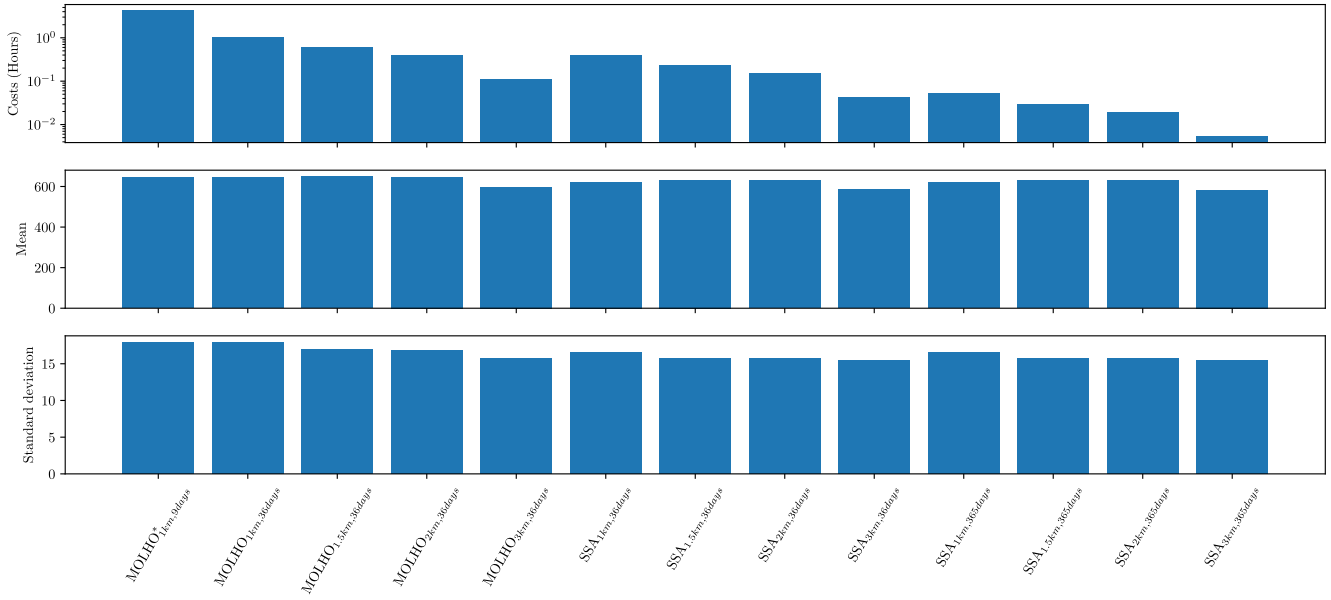


Figure 8. (*Top*) The median computational cost (in hours) of simulating each model used in this study for one realization of the random parameters. (*Middle*) The mean mass loss – negative expected mass change – (gigatons) at 2100. (*Bottom*) The standard deviation of the mass change (gigatons) at 2100. Each quantity was computed using 20 pilot samples.

The exact gain in performance achieved by MFUQ is dependent on the correlations between each model and the other quantities in (19). Figure 9 plots the entries of the correlation matrix and shows that despite the differences between each

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

model's prediction of ice thickness and velocities at the final time (see Figure 13), and the differences in the SFMC estimate of the mean and variance computed using each model, the correlation between each model's prediction of the mass change is high.[8]

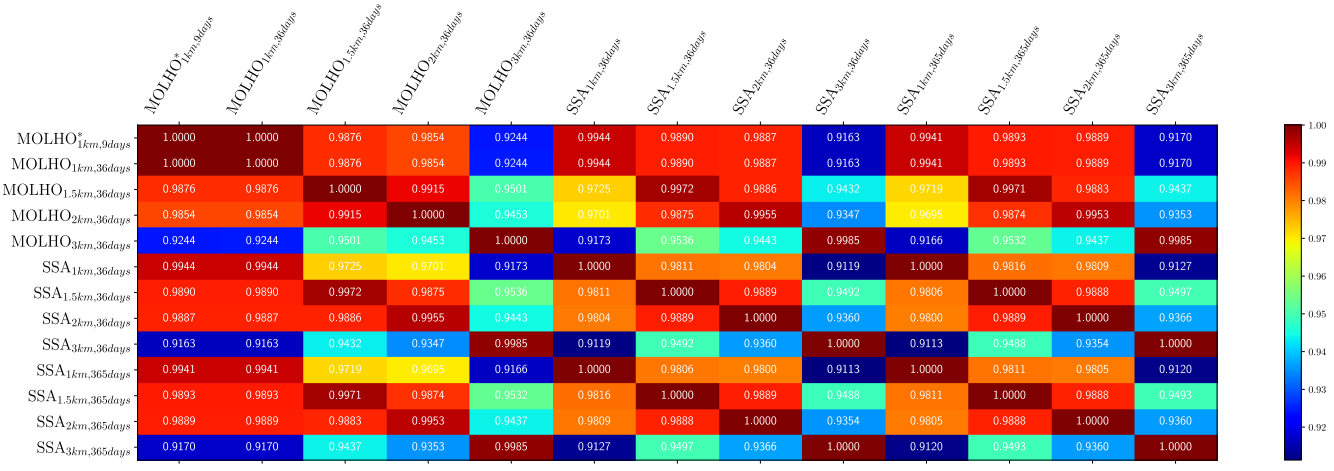


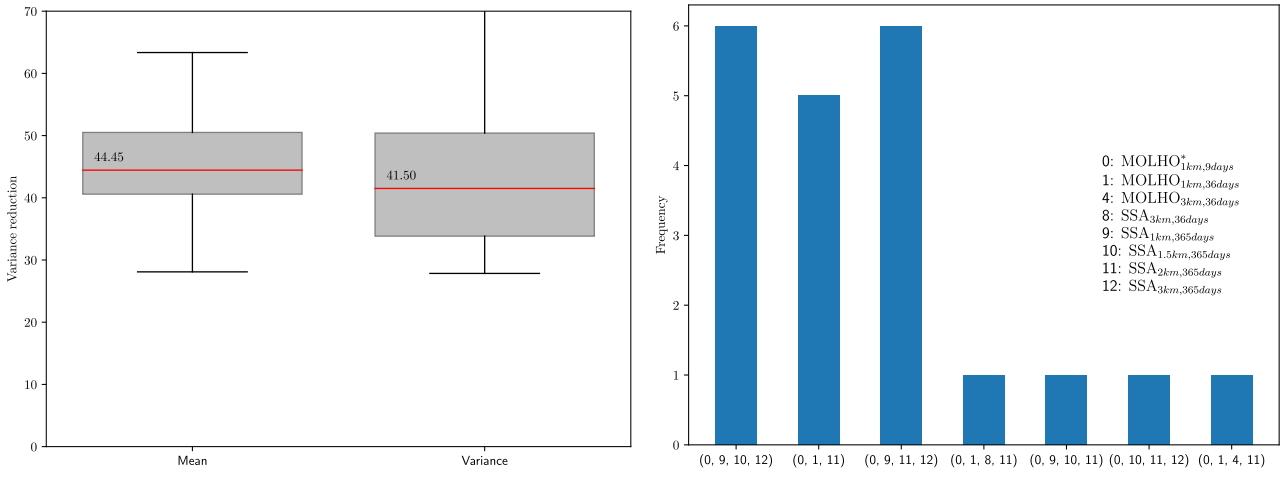
Figure 9. The correlations between the 13 ice-sheet models considered by this study using 20 pilot samples.

Given estimates of the pilot statistics, (18) and (19), we used (20) to predict the determinant of the variance of the ACV estimator of the mean and variance of the mass change. Specifically, we made these predictions assuming that a budget of 160 high-fidelity model evaluations would be allocated to the high-and low-fidelity models. Moreover, this cost was assumed to be additional to the computational cost of simulating each model at the pilot samples. We then computed the ratio of the variance of the best ACV estimator of the mean and variance of the mass change (taken by extracting the diagonal elements of the estimator covariance) to the SFMC estimator variance of those same quantities. For a single statistic, this ratio is often referred to as the *variance reduction* of the ACV estimator. To ensure a fair comparison we compared the ACV estimator variance to the SFMC estimator variance obtained using a computational budget equivalent to 160 high-fidelity evaluations plus the computational cost of collecting the pilot model evaluations.

Existing literature assumes that the pilot statistics used with (20) are exact, however using a small number of pilot samples can effect the accuracy of the predicted estimator covariance. Moreover, we found that the error introduced is not insignificant, yet it is typically ignored in existing literature. Consequently, in Figure 10a[9] we plot the variance reduction of the ACV estimators of the mean and variance of mass loss (diagonal terms of the estimator covariance) for 21 different bootstraps of the 20 pilot samples (1 bootstrap was just the original pilot data). The plot is created by randomly sampling the model evaluations with replacement, computing the pilot statistics with those samples, and solving (20).

[8]Due to rounding, the correlation between $\text{MOLHO}^*_{1km,9days}$ and $\text{MOLHO}_{1km,36days}$ is not exactly 1 as reported in Figure 9.
[9]The upper whisker in the right box plot of Figure 10a is cutoff because its true size would distort the image.

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024



(a) Variance reductions

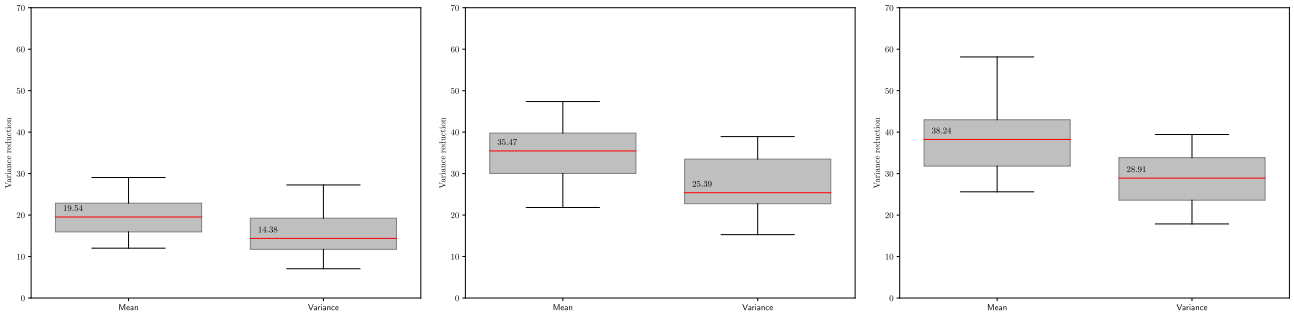
(b) Selected model subsets

Figure 10. (a) The predicted variance reduction of the best ACV estimators of the mean and variance of the mass change obtained using bootstrapping of the initial 20 pilot samples. The red lines represent the median estimator variances. The lower and upper whiskers represent the 5% and 95% quantiles. (b) The model subsets chosen by the bootstrapped estimators using the initial 20 pilot samples.

The median variance reduction is over 40 for the ACV estimators of both the mean and variance of the mass change. In other words, our initial pilot study predicted that using ACV estimators would reduce the cost of estimating uncertainty in projections of the mass change by over a factor of 40 when compared to SFMC estimators that only use the highest-fidelity model. However, the box plots in 10a highlight that using only 20 samples introduces a large degree of uncertainty into the estimated variance reduction. The 5% quantile of the variance reduction for both the mean and variance estimators were slightly below 30.

The estimators obtained by bootstrapping the initial 20 pilot samples not only had different estimator variances (see Figure 10a), they also predicted that different subsets are needed to minimize the estimator variance. Figure 10b plots the model subsets chosen by the bootstrapped estimators and the number of times (frequency) each subset was chosen; the set (0,9,10, 12) was chosen when the original 20 pilot samples (bootstrapping was not used). Not all models were determined useful for reducing the variance of the ACV estimators. Specifically, only eight out of the 13 models considered were chosen at least once by a bootstrapped estimator; models $\text{MOLHO}_{1.5km,36days}$, $\text{MOLHO}_{2km,36days}$, $\text{SSA}_{1km,36days}$, $\text{SSA}_{1.5km,36days}$, $\text{SSA}_{2km,36days}$, were never selected by any of the bootstrapped estimators. Moreover, in some cases only two low-fidelity models were chosen and in other cases three low-fidelity models were chosen. Lastly, not only did the chosen model subsets vary between bootstrapped estimators, the type of estimator chosen also varied. In 7 cases, a hierarchical relationship was identified and in the other 14 cases a non-hierarchical relationship was identified; a non-hierarchical estimator was chosen using the original 20 pilot samples (the 21st estimator). This highlights the difficulty of a priori selecting a model ensemble that is most effectively used by a hierarchical estimator and for which all models will be used to produce the final ACV estimator.

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024



(a) 2 models (b) 3 models (c) 4 models

Figure 11. The predicted ratio of the variance of the best ACV estimators to the SFMC estimator variance (variance reduction) was obtained by bootstrapping the final 30 pilot samples while enforcing a limit on the number of models an estimator can use, including the highest-fidelity model. The red lines indicate the median estimator variances. The lower and upper whiskers represent the 5% and 95% quantiles.

5.4 Secondary pilot study

Upon quantifying the impact of only using 20 pilot samples on the estimator variances and the model subsets chosen, we incremented the number of pilot samples we used to compute the performance of the ACV estimators. To avoid wasting computational resources in our secondary pilot study, we only evaluated the 8 models selected by at least one bootstrapped estimator on an additional 10 pilot samples. The combined cost of the initial and secondary pilot study was approximately 197 hours which equates to the equivalent of approximately 47 simulations of the highest-fidelity model. Note that all models included in the second pilot were simulated 30 times and those only included in the first pilot were simulated 20 times.

Figure 11c illustrates the estimator variance of the mean and variance of mass loss for 21 different bootstraps of the final 30 pilot samples, with the constraint that the number of models an estimator can use does not exceed four. Upon examining Figure 11c, we observed that increasing the number of pilot samples decreased the variability of the estimator variances. However, this increase in pilot samples also led to a higher cost, which in turn reduced the reported median variance reduction. The median variance reduction decreased because ACV estimators utilize the pilot samples solely to compute pilot statistics, such as variance, and do not reuse these samples for calculating the final statistics. In contrast, an equivalent SFMC estimator can leverage both the pilot and exploitation budgets to estimate the final statistics. In other words, the variance of an SFMC estimator decreases linearly with the number of pilot samples, whereas the variance of an ACV estimator does not exhibit the same behavior. The variance of an ACV estimator is only marginally affected by an increase in the number of pilot samples, as the sample allocation becomes more optimal.

While, increasing the number of pilot samples decreased variability, we believed that the benefit of further increasing the number of pilot samples would be outweighed by the resulting drop in the variance reduction. Despite the remaining variability in the variance reduction, we were able to confidently conclude that the final ACV estimator we would construct would be much more accurate than a SFMC estimator of the same cost because even the smallest variance reduction was greater than

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

20. Consequently, we used the unaltered 30 pilot samples to determine the ACV estimator and its optimal sample allocation that we used to construct our final estimates of the mean and variance of the mass change. The best estimator chosen was an MFMC estimator that used the three models $\text{MOLHO}_{1km,36days}$, $\text{SSA}_{1km,36days}$, and $\text{SSA}_{1.5km,365days}$.

5.5 Multi-fidelity sea-level rise projections

The cost of constructing our final estimator was equal to the pilot cost and the exploitation cost, that is $197.13 + (160 \times 4.18)$ hours or approximately 36 days. The number of samples allocated to evaluating each model by the ACV estimator during the exploitation phase are shown in the right panel of Figure 12. Only 2 samples of the high-fidelity model were used, which accounts for approximately 1.25% of the total computational cost budget; these samples ensured the estimators were unbiased. In contrast, many more evaluations of the lower fidelity models were used. The lower computational costs of these models and their high-correlation with each other and the highest-fidelity model were effectively exploited to significantly reduce the MSE of the ACV estimator relative to the SFMC estimator.

We constructed our final estimator of the mean and variance of the mass change by evaluating each model at the number of samples determined by Figure 12. All models were evaluated at the same two samples, the two low-fidelity models were both evaluated at another 351 samples, and the $\text{SSA}_{1.5km,365days}$ model was evaluated at another 10130 samples; the exact structure reported is determined by the properties of the MFMC estimator chosen, however, if another estimator was chosen to use the same models the way samples are shared between models would likely change.[10]

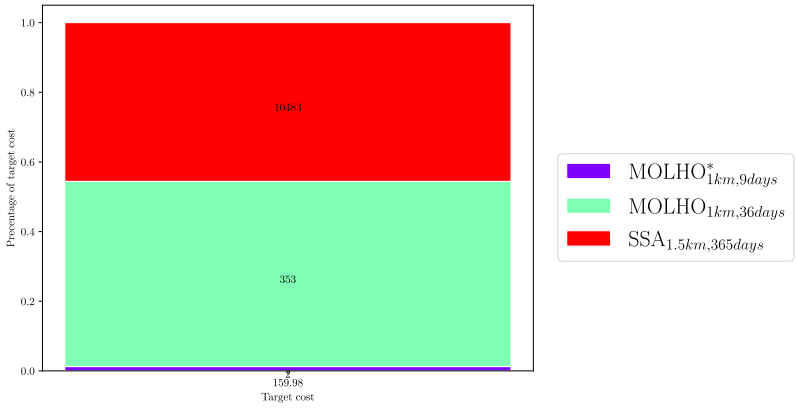


Figure 12. The number of samples allocated to evaluating each model (number inside rectangles) by the ACV estimator

The mean and variance computed using the best ACV estimator are shown in Table 1. It is clear that with our budget we were able to confidently estimate the expected mass change at year 2100. However, our estimates of the standard deviation in the mass change were less accurate. We could improve the estimates of both statistics by further increasing the exploitation budget, however we did not do this because our results serve the purpose of demonstrating that estimating statistics other than the mean

[10] 37 of the 10130 $\text{SSA}_{1.5km,365days}$ model simulations failed so an additional 37 simulations at new random realizations of the friction field were run.

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

of a model is more computationally demanding. Moreover, the accuracy requirements of a UQ study should be determined by the stakeholders that will use the uncertainty estimates to make decisions.

Table 1. The mean and variance of the final mass change in Gigatons from Humboldt glacier between 2007 and 2100. The reported plus/minus corresponds to one standard deviation of the estimator error.

Mean	Std. deviation
-639.06 ± 0.23	17.68 ± 6.67

6 Discussion

The cost of constructing our final estimator was equal to the pilot cost and the exploitation cost, that is $197.13 + (160 \times 4.18)$ hours or approximately 36 days. Given that the median variance reduction obtained by the bootstrapped estimators was 38.24 and 28.91 when estimating the mean and variance of the mass change, respectively, constructing SFMC estimators of the same accuracy when using only the highest fidelity model would require approximately $28.91 \times 160 \times 4.18$ hours $= 805$ days to estimate the mean and variance; we take the worst case variance reduction as samples used to compute the variance via SFMC can also be used to compute the mean. Thus, MFUQ reduced the cost of estimating uncertainty from over two and a half years of CPU time to just over a month assuming the models are evaluated in serial.

While the highest-fidelity model MOLHO was capable of capturing ice-sheet dynamics that the SSA model was not, that is vertical changes in the horizontal velocities, the best ACV estimator was still able to use the simplified physics of SSA to reduce the MSE of the best ACV estimator. Moreover, the best ACV estimator also used evaluations of the SSA model on a coarse mesh that was not able to resolve all the local features of the friction and ice-sheet flow-field (see Figure 13) and did not conserve mass, unlike the highest fidelity model. This result demonstrates that MFUQ can be effective when there is high correlation between the model predictions of a QoI, even when the model states vary differently across time and space for a single realization of the random model inputs. Moreover, future MFUQ studies may benefit from not only using low-fidelity models derived from different physics assumptions and numerical discretizations but also those based on data driven models, such as machine learning operators (He et al., 2023; Lowery et al., 2024). However, if such models are used the computational cost of constructing such models must also be accounted for in the same way we accounted for the pilot cost in this study (Peherstorfer, 2019).

Our study used a high-dimensional representation of the basal friction field that is capable of capturing high-frequency modes, however it has been common in previous studies to use lower-dimensional parameterizations. Consequently, we investigated the impact of using a low-frequency/lower-dimensional representation of the friction field on the efficiency of ACV estimators using ice-sheet models. Specifically, we estimated the mean and variance of the mass change using a 10 dimensional Karhunen Loeve expansion (KLE) to represent the posterior uncertainty of the basal friction field (complete details are presented in B). We found that using the low-dimensional KLE smoothed realizations of the basal friction which in turn drastically improved the variance reduction of MFUQ to over a factor of 200. However, only using 10 modes to represent the basal

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

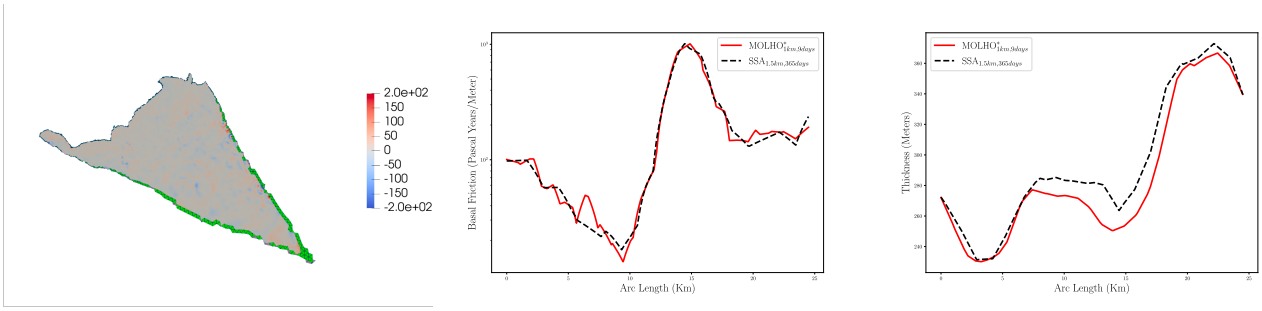


Figure 13. Differences in the final ice thickness at 2100. (*Left*) The difference between the thickness fields simulated by the $\mathrm{MOLHO}^*_{1km,9days}$ model and the $\mathrm{SSA}_{1.5km,365days}$. (*Center*) The basal friction and the ice thickness (*right*) at 2100 along the cross section (black line) depicted in the right panel of Figure 7.

friction led to the variance of the mass change being substantially underestimated. Consequently, while it may be appealing to use low-dimensional representations of friction to enable faster UQ, the results of that UQ may be misleading. Thus, future research is needed to balance the increased bias introduced by the low-dimensional parameterization with the improved variance reduction properties introduced in an ACV estimator.

This study also emphasizes that the best MFUQ algorithm is problem dependent. While theoretically each MFUQ algorithm in the literature has its own advantages and disadvantages, it is often difficult to determine at the beginning of a study which will be the most effective. Indeed, several types of estimators enumerated by this study provided similar reductions in the MSE of the mean and variance of the mass change. For example, Figures 11a, 11b, 11c show that while using two low-fidelity models is clearly better than using one, there is little, if any, marginal benefit, given the size of the box plots, of moving from two to three low-fidelity models. Moreover, it is difficult to determine *a priori* the numerical discretizations and model physics needed by a model ensemble to produce the MFUQ estimates with the smallest MSE. Consequently, we used a small pilot sample to compute the correlation between model outputs and then use the analytical properties of ACV estimators to predict the MSE of each estimator produced by popular MFUQ algorithms.

While pilot studies are required for ACV methods, our results show that using a small number of pilot samples can introduce non-trivial variability into the optimal sample allocation used by ACV estimators. Consequently, we introduced a novel two-step bootstrapping procedure to quantify the impact of a small number of pilot samples. While our two step procedure was able to down select from a large set of possible models, further research is needed to develop algorithms that can efficiently conduct a pilot study when large numbers of models are available. Furthermore, it is essential that new algorithms must balance the computational cost of computing the correlation between models with the impact the error in the estimated correlations has on determining the optimal MSE of an ACV estimator.

Our study predicted the mean and variance of the mass change (in Gigatons) from Humboldt Glacier to be -639.06 and 17.68 respectively. However, the significance of these numbers is impacted by our modeling choices. First, we only quantified uncertainty due to unknown basal friction which ignores other contributions to mass-loss variance arising from uncertain

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

climate, and other ice-sheet processes such as iceberg calving, subglacial hydrology, and submarine melting. Including these processes would have likely affected both the mean and variance of the mass change. Indeed, our predicted mass loss is significantly less than in two recent studies of Humboldt Glacier (Hillebrand et al., 2022; Carr et al., 2024) due to our use of a low-emissions climate scenario and the absence of ocean forcing. However, despite our imperfect description of uncertainty our study reflects the challenges of a more comprehensive study while still facilitating a computationally feasible investigation of MFUQ methods.

Many recent studies have demonstrated formal uncertainty quantification of projections of ice-sheet change, and many of these studies consider numerous sources of uncertainty, such as climate forcing, iceberg calving, basal friction parameters, and ice viscosity. However, these generally deal with scalar parameters, such as a single calving threshold stress (Aschwanden and Brinkerhoff, 2022; Jantre et al., 2024) or scalar adjustment factors to basal friction and ice viscosity fields (Nias et al., 2023; Felikson et al., 2023; Jantre et al., 2024). While the present study is limited in scope due to its investigation of basal friction alone, to our knowledge it is the first study to quantify uncertainty associated with a high-dimensional representation of basal friction, which is a great improvement in realism relative to previous work. Additionally, other UQ studies have primarily relied on a large number of simulations from a single low fidelity model(e.g., Nias et al., 2019; Bevan et al., 2023) — sometimes with informal validation using a small number of higher-fidelity simulations (e.g., Nias et al., 2023) — or on the construction of surrogate models to sufficiently sample the parameter space (e.g., Bulthuis et al., 2019; Berdahl et al., 2021; DeConto et al., 2021; Hill et al., 2021; Aschwanden and Brinkerhoff, 2022; Jantre et al., 2024). Furthermore, another set of studies quantifies the uncertainty associated with the use of many different numerical models — a so-called "ensemble of opportunity" — which includes a wide range of modeling choices that sample parameter values and model fidelity in an unsystematic way (Edwards et al., 2021; Seroussi et al., 2023; Van Katwyk et al., 2023; Yoo et al., 2024). While each of the aforementioned approaches have their benefits and limitations, our results demonstrate that even when low-fidelity ice-sheet models do not capture the flow features predicted by higher-fidelity model, such low-fidelity models can still be used by MFUQ methods to reduce the cost of UQ for ice sheets. Consequently, low-fidelity models when used with MFUQ methods may be able to substantially reduce the computational cost of future attempts to quantify uncertainty in the projection of the mass change from the entire Greenland and Antarctic ice sheets.

7 Conclusions

Melting of land-based ice sheets is anticipated to contribute substantially to sea-level rise in the next century. However, projections of sea-level rise due to ice-sheet mass change are subject to uncertainty, and the impact of this uncertainty on sea-level projections must be quantified in order for these projections to be of use to policy makers and planners. Unfortunately, accurately estimating uncertainty is challenging because it requires numerous simulations of a computationally expensive numerical model. Consequently, we evaluated the efficacy of MFUQ for reducing the computational cost of quantifying uncertainty in projections of mass loss from Humboldt Glacier, Greenland.

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

This study used MFUQ to estimate the mean and the variance of uncertain mass-change projections caused by uncertainty in glacier basal friction using 13 different models of varying computational cost and accuracy. While ice sheets are subject to other sources of uncertainty, focus was given to basal friction because its inherent high-dimensionality renders quantifying the impact of its uncertainty difficult. Yet we found that for a fixed computational budget, MFUQ was able to reduce the MSE in our estimates of the mean and variance of the mass change by over an order-of-magnitude relative to a SFMC based approach that just used simulations from the highest fidelity model.

The MFUQ algorithm we used was able to reduce the MSE error in the statistics by exploiting the correlation between the predictions of the mass change produced by each model. However, using simulations from all of the models was not necessary to reduce the MSE. Indeed, the MFUQ algorithm determined that only 3 models (including the highest-fidelity model) were needed to minimize the MSE in the statistics given our computational budget. The low-fidelity models selected used: 1) simplifications of the high-fidelity model physics, 2) were solved on coarser resolution spatial and temporal meshes, and 3) were solved without the requirement of mass conservation. All three of these simplifications result in significant computational cost savings relative to use of the high-fidelity model alone. This result demonstrated that MFUQ can be effective even when the lower-fidelity models are incapable of capturing the local features of the ice flow fields predicted by the high-fidelity model. Moreover, the utility of the lower-fidelity models ultimately chosen for MFUQ were not clear at the onset of the study.

Finally, this study demonstrated that MFUQ can be used to reduce the computational cost of quantifying uncertainty in projections of a single glacier, which suggests that MFUQ could plausibly be used for continental-scale studies of ice-sheet evolution in Antarctica. Future research should increase the complexity of this study in two directions. First, future studies should include additional sources of ice-sheet uncertainty beyond the basal friction field studied here, for example uncertain surface mass balance and ocean forcing. Second, future studies should include the use of model fidelities that capture additional physical processes such as calving, fracture, and ocean-forced melting. Consequently, while our findings should be interpreted with caution given the aforementioned limitations they encourage future studies to utilize MFUQ for reducing the cost of computing probabilistic projections of sea-level rise due to ice-sheet mass change.

Code availability. The code used to construct ACV estimators has been released in the open-source Python package PyApprox https://github.com/sandialabs/pyapprox.

Appendix A: Low-rank Laplace approximation

Following Bui-Thanh et al. (2013); Isaac et al. (2015) we computed the covariance of the Laplace approximation of the posterior distribution of the friction parameters using

$$\Sigma_{\text{post}} = \left(H + \Sigma_{\text{prior}}^{-1}\right)^{-1} = L\left(L^\top H L + I\right)^{-1} L^\top$$

Drawing samples from this Gaussian posterior is computationally challenging because the posterior covariance Σ_{post} depends on the Hessian H which is a high-dimensional dense matrix. Consequently, following Bui-Thanh et al. (2013) and Isaac et al.

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

(2015) we constructed a low-rank approximation of the prior-preconditioned Hessian $L^\top HL$ using matrix-free randomized methods that requires only multiplications of the Hessian with random vectors. Specifically, computing a spectral decomposition of $L^\top HL = U\Lambda U^\top$, with U orthogonal and Λ diagonal matrices and noting

$$\Sigma_{\text{post}} = L\left(U\Lambda U^\top + I\right)^{-1} L^\top = L\left(U(\Lambda+I)U^\top\right)^{-1} L^\top = LU(\Lambda+I)^{-1}U^\top L^\top$$

we factorized Σ_{post} as

$$\Sigma_{\text{post}} = TT^\top, \qquad T = LU(\Lambda+I)^{-\frac{1}{2}}U^\top = LU\left((\Lambda+I)^{-\frac{1}{2}} - I\right)U^\top + L.$$

Thus a low-rank, or truncated spectral decomposition of the prior preconditioned Hessian is accurate if the matrix $W = U\left((\Lambda+I)^{-\frac{1}{2}} - I\right)U^\top$ is accurate. Consequently, we discarded eigenvalues λ_i such that $\left|1 - \frac{1}{\sqrt{\lambda_i+1}}\right| \ll 1$. The eigenvalues and two eigenvectors of the spectral decomposition we computed are depicted in Figure A1.

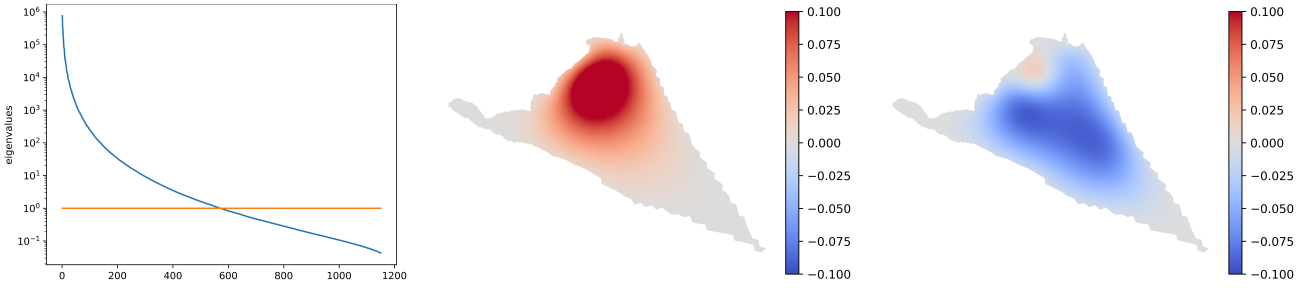


Figure A1. Eigenvalues of the prior-preconditioned Hessian (Left), and its eigenvectors associated to the largest (Center) and third largest (Right) eigenvalue.

We computed the truncated spectral decomposition using randomized algorithms (see Hartland et al. (2023); Halko et al. (2011)) implemented in PyAlbany, see Liegeois et al. (2023). The algorithms used were matrix-free and only required the multiplication of $L^T HL$ with vectors. Moreoever, as described in Hartland et al. (2023); Isaac et al. (2015), the multiplication of the Hessian with a vector required solving two adjoint systems of the flow model. Similarly, the multiplication of the matrix L with a vector required the solution of the two-dimensional linear elliptic system with matrix K. Consequently, we were able to efficiently draw samples from the posterior distribution of the friction parameters using

$$\theta_{\text{post}} = \mu_{\text{MAP}} + Tn, \quad n \sim \mathcal{N}(0,1).$$

Appendix B: Low-dimensional representation of basal friction using a Karhunen-Loeve expansion

In our main study we found that when using a high-dimensional representation of the uncertainty in the basal friction field, bootstrapped ACV estimators rarely chose to use models that had coarse spatial meshes relative to the mesh used by the high-fidelity model. This was likely due to the fact that our high-dimensional bi-Laplacian representation of the friction uncertainty

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

was constructed on the high-fidelity mesh and interpolated onto coarser meshes. To verify this hypothesis we investigated using a lower-dimensional representation of the friction field based on a Karhunen Loeve expansion (KLE) of the friction field that smoothed out the high-frequency variations in the posterior samples of the friction field we used in our main study.

Construction of the KLE

In our investigations we used a KLE

$$\theta = \theta_{MAP} + \sum_{i=1}^{D} \sqrt{\lambda_i}\psi_i \eta_i, \qquad \eta_i \sim \mathcal{N}(0,1) \tag{B1}$$

to provide a low-dimension representation of the Laplace approximation of the posterior of the log basal friction field. We computed the eigenvalues λ_i and the orthonormal eigenvectors ψ_i by solving the eigenvalue problem

$$\Sigma_{\text{post}}\psi_i = \lambda_i \psi_i,$$

using the randomized matrix-free methods Hartland et al. (2023); Halko et al. (2011).

While a KLE basis could have been constructed on any of the four meshes we considered, this study we solved the discretized eigenvalue problem using the finest mesh. The 1st, 2nd, and 10th mode of the KLE used in this study are depicted in Figure B1. The finite element basis on the finest mesh was then used to interpolate the KLE basis from the fine mesh onto the coarser meshes. This procedure ensured that varying the coefficients of the KLE basis (the random inputs to the model) would affect each model similarly regardless of the mesh discretization employed. Similarly to the KLE basis, the mean of the log KLE field (taken to be the mean of the Laplace approximation) was computed on the finest mesh.

Figure B2 compares a realization of the log of the basal friction perturbation (mean zero) drawn from the Laplace approximation of the posterior and a random random realization of the log of the basal friction perturbation computed using the KLE. It is clear that the KLE smooths out much of the high-frequency content present in the realization drawn from the Laplace approximation of the posterior.

Figure B1. From left to right, the 1st, 2nd, and 10th mode of the KLE used in this study.

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

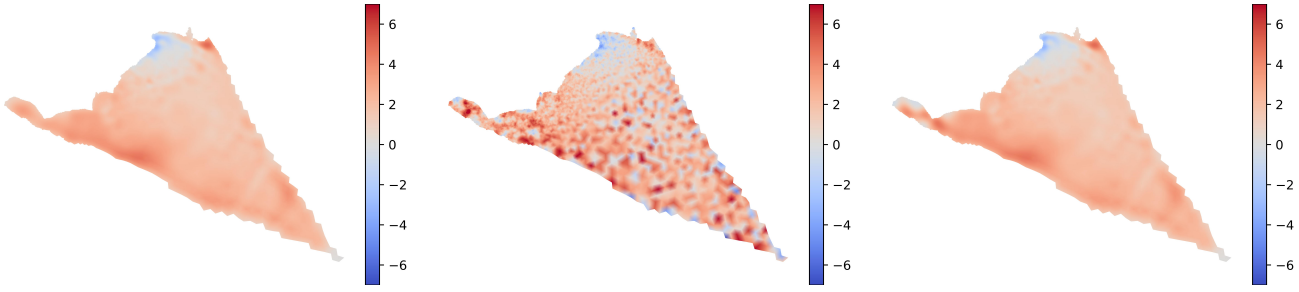


Figure B2. (*Left*) The mean of the log of the basal friction, (*Center*) A random realization of the log of the basal friction drawn from the Laplace approximation of the posterior. (*Right*) A random realization of the log of the basal friction computed using the KLE.

Pilot study

In this section we detail the pilot study we undertook to investigate the impact of using a low-dimensional KLE to represent friction when using MFUQ to estimate statistics of mass change. We did not move beyond the pilot study to compute the values of the statistics to limit the computational cost of this supplementary study that is not the main focus of our study.

First, we evaluated each of our 13 models at 20 random pilot samples of the KLE. Second we computed the pilot statistics needed to find the best MF estimator. Third we bootstrapped the pilot samples to estimate the median and confidence intervals on the variance reduction obtained by the best MFUQ estimator.

The variance bootstrapped variance reduction are depicted in Figure B4. The variance reductions reported are almost an order of magintude larger than those reported for MFUQ based on the Laplace approximation of the posterior. This improved performance is because correlations between the models (Figure B3) are significantly higher than the correlations obtained when sampling from the Laplace approximation of the posterior (Figure 9). However, the KLE representation underestimates the uncertainty in the mass change at 2100. Specifically, the standard deviation of the mass change computed using 20 pilot samples of the highest-fidelity model using the Laplace approximation of the posterior is significantly higher than the standard deviation computed using the KLE.

Author contributions. John Jakeman was responsible for formulation of the overarching research goals and aims (conceptualization), data curation, application of the statistical techniques used to analyze the data (formal analysis), conducting the computer experiments (investigation), developing the methodology , implementing and maintaining the software used (software), oversight and leadership for the research planning and execution (supervision), writing the original draft. Mauro Perego was responsible for investigation, methodology, software and writing the original draft. Tom Seidl was responsible for data curation, investigation, software, and writing the original draft. Tucker Hartland was responsible for software and writing the original draft. Trevor Hillebrand was responsible for developing the data curation, methodology, and writing the original draft. Matthew Hoffman was responsible for developing the conceptualization, methodology, supervision, funding

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

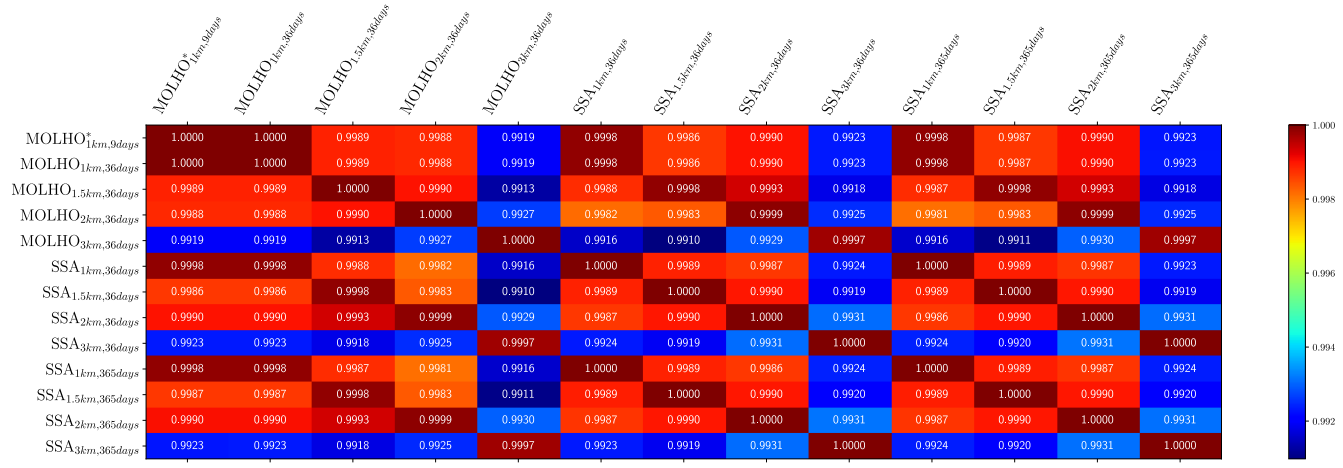


Figure B3. The correlations between the 13 ice-sheet models considered by this study using 20 pilot samples of the KLE.

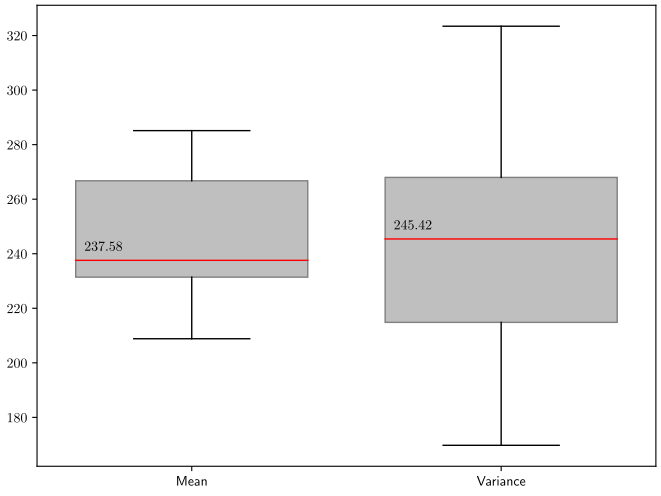


Figure B4. The predicted ratio of the variance of the best ACV estimators and the SFMC estimator variance (variance reduction) obtained using bootstrapping of the initial 20 pilot samples. The red lines represent the median estimator variances. The lower and upper whiskers represent the 5% and 95% quantiles.

aquisition, and writing the original draft. Stephen Price was responsible for developing the conceptualization, funding aquisition, and writing the original draft.

Competing interests. The authors have no competing interests.

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

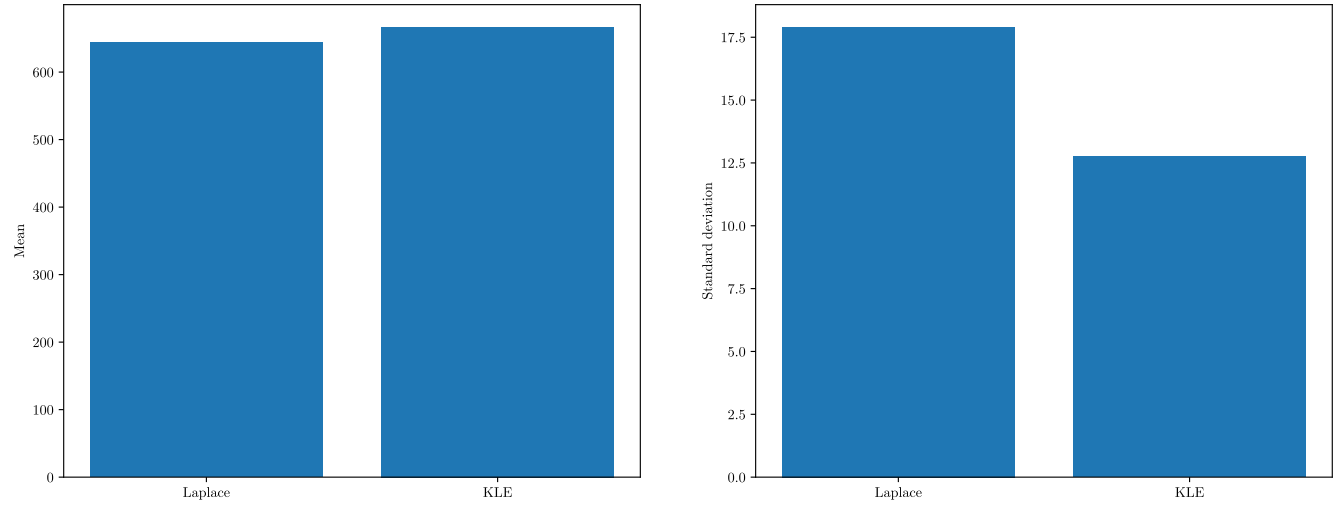


Figure B5. The mean and standard deviation computed using 20 pilot samples from the Laplace approximation of the posterior and the KLE.

Acknowledgements. This work was was sponsored by the Sandia National Laboratories the US Department of Energy's Office of Science Advanced Scientific Computing Research, Biological, and Environmental Research Scientific Discovery through Advanced Computing (SciDAC) programs.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC (NTESS), a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration (DOE/NNSA) under contract DE-NA0003525. This written work is authored by an employee of NTESS. The employee, not NTESS, owns the right, title and interest in and to the written work and is responsible for its contents. Any subjective views or opinions that might be expressed in the written work do not necessarily represent the views of the U.S. Government. The publisher acknowledges that the U.S. Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this written work or allow others to do so, for U.S. Government purposes. The DOE will provide public access to results of federally sponsored research in accordance with the DOE Public Access Plan.

The authors would like to thank Michael Eldred for his helpful discussions on how to practically assess the impact of small number of pilot samples on the predicted variance of ACV estimators. The authors would also like to thank Thomas Dixon and Alex Gorodetsky for their useful discussions on constructing ACV estimators for multiple statistics.

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

References

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, 2015.

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

Bakker, A. M. R., Wong, T. E., Ruckert, K. L., and Keller, K.: Sea-level projections representing the deeply uncertain contribution of the West Antarctic ice sheet, Scientific Reports, 7, 3880, https://doi.org/10.1038/s41598-017-04134-5, 2017.

Balay, S., Gropp, W., McInnes, L. C., and Smith, B. F.: PETSc, the portable, extensible toolkit for scientific computation, Argonne National Laboratory, 2, 1998.

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

Berdahl, M., Leguy, G., Lipscomb, W. H., and Urban, N. M.: Statistical emulation of a perturbed basal melt ensemble of an ice sheet model to better quantify Antarctic sea level rise uncertainties, The Cryosphere, 15, 2683–2699, https://doi.org/10.5194/tc-15-2683-2021, 2021.

Bevan, S., Cornford, S., Gilbert, L., Otosaka, I., Martin, D., and Surawy-Stepney, T.: Amundsen Sea Embayment Ice-Sheet Mass-Loss Predictions to 2050 Calibrated Using Observations of Velocity and Elevation Change, Journal of Glaciology, pp. 1–11, https://doi.org/10.1017/jog.2023.57, 2023.

Bochev, P., Ridzal, D., D'Elia, M., Perego, M., and Peterson, K.: Optimization-based, property-preserving finite element methods for scalar advection equations and their connection to Algebraic Flux Correction, Computer Methods in Applied Mechanics and Engineering, 367, 112 982, https://doi.org/https://doi.org/10.1016/j.cma.2020.112982, 2020.

Bomarito, G., Leser, P., Warner, J., and Leser, W.: On the optimization of approximate control variates with parametrically defined estimators, Journal of Computational Physics, 451, 110 882, https://doi.org/10.1016/j.jcp.2021.110882, 2022.

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, SIAM Journal on Scientific Computing, 35, A2494–A2523, https://doi.org/10.1137/12089586X, 2013.

Bulthuis, K., Arnst, M., Sun, S., and Pattyn, F.: Uncertainty quantification of the multi-centennial response of the Antarctic ice sheet to climate change, The Cryosphere, 13, 1349–1380, https://doi.org/10.5194/tc-13-1349-2019, 2019.

Carr, J. R., Hill, E. A., and Gudmundsson, G. H.: Sensitivity to forecast surface mass balance outweighs sensitivity to basal sliding descriptions for 21st century mass loss from three major Greenland outlet glaciers, The Cryosphere, 18, 2719–2737, 2024.

Cornford, S. L., Martin, D. F., Graves, D. T., Ranken, D. F., Le Brocq, A. M., Gladstone, R. M., Payne, A. J., Ng, E. G., and Lipscomb, W. H.: Adaptive mesh, finite volume modeling of marine ice sheets, Journal of Computational Physics, 232, 529–549, https://doi.org/https://doi.org/10.1016/j.jcp.2012.08.037, 2013.

Cuffey, K. and Paterson, W.: The Physics of Glaciers, Butterworth-Heinneman, Amsterdam, 4th edn., 2010.

DeConto, R. M., Pollard, D., Alley, R. B., Velicogna, I., Gasson, E., Gomez, N., Sadai, S., Condron, A., Gilford, D. M., Ashe, E. L., et al.: The Paris Climate Agreement and future sea-level rise from Antarctica, Nature, 593, 83–89, 2021.

Dias dos Santos, T., Morlighem, M., and Brinkerhoff, D.: A new vertically integrated MOno-Layer Higher-Order (MOLHO) ice flow model, The Cryosphere, 16, 179–195, https://doi.org/10.5194/tc-16-179-2022, 2022.

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

Dixon, T., Warner, J., Bomarito, G., and Gorodetsky, A.: Multi-Fidelity Estimation for Multi-Output Systems using Approximate Control Variates, ARXIV, 2023.

Dukowicz, J. K., Price, S. F., and Lipscomb, W. H.: Consistent approximations and boundary conditions for ice-sheet dynamics from a principle of least action, Journal of Glaciology, 56, 480–496, https://doi.org/10.3189/002214310792447851, 2010.

Durand, G., Gagliardini, O., Zwinger, T., Meur, E. L., and Hindmarsh, R. C.: Full Stokes modeling of marine ice sheets: influence of the grid size, Annals of Glaciology, 50, 109–114, https://doi.org/10.3189/172756409789624283, 2009.

Edwards, T. L., Brandon, M. A., Durand, G., Edwards, N. R., Golledge, N. R., Holden, P. B., Nias, I. J., Payne, A. J., Ritz, C., and Wernecke, A.: Revisiting Antarctic ice loss due to marine ice-cliff instability, Nature, 566, 58–64, https://doi.org/10.1038/s41586-019-0901-4, 2019.

Edwards, T. L., Nowicki, S., Marzeion, B., Hock, R., Goelzer, H., Seroussi, H., Jourdain, N. C., Slater, D. A., Turner, F. E., Smith, C. J., McKenna, C. M., Simon, E., Abe-Ouchi, A., Gregory, J. M., Larour, E., Lipscomb, W. H., Payne, A. J., Shepherd, A., Agosta, C., Alexander, P., Albrecht, T., Anderson, B., Asay-Davis, X., Aschwanden, A., Barthel, A., Bliss, A., Calov, R., Chambers, C., Champollion, N., Choi, Y., Cullather, R., Cuzzone, J., Dumas, C., Felikson, D., Fettweis, X., Fujita, K., Galton-Fenzi, B. K., Gladstone, R., Golledge, N. R., Greve, R., Hattermann, T., Hoffman, M. J., Humbert, A., Huss, M., Huybrechts, P., Immerzeel, W., Kleiner, T., Kraaijenbrink, P., Le clec'h, S., Lee, V., Leguy, G. R., Little, C. M., Lowry, D. P., Malles, J.-H., Martin, D. F., Maussion, F., Morlighem, M., O'Neill, J. F., Nias, I., Pattyn, F., Pelle, T., Price, S. F., Quiquet, A., Radić, V., Reese, R., Rounce, D. R., Rückamp, M., Sakai, A., Shafer, C., Schlegel, N.-J., Shannon, S., Smith, R. S., Straneo, F., Sun, S., Tarasov, L., Trusel, L. D., Van Breedam, J., van de Wal, R., van den Broeke, M., Winkelmann, R., Zekollari, H., Zhao, C., Zhang, T., and Zwinger, T.: Projected land ice contributions to twenty-first-century sea level rise, Nature, 593, 74–82, https://doi.org/10.1038/s41586-021-03302-y, 2021.

Felikson, D., Nowicki, S., Nias, I., Csatho, B., Schenk, A., Croteau, M. J., and Loomis, B.: Choice of observation type affects Bayesian calibration of Greenland Ice Sheet model simulations, The Cryosphere, 17, 4661–4673, https://doi.org/10.5194/tc-17-4661-2023, 2023.

Fox-Kemper, B., Hewitt, H., Xiao, C., Aoalgeirsdottir, G., Drijfhout, S., Edwards, T., Golledge, N., Hemer, M., Kopp, R., Krinner, G., Mix, A., Notz, D., Nowicki, S., Nurhati, I., Ruiz, L., Sallee, J.-B., Slangen, A., and Yu, Y.: Ocean, cryosphere and sea level change., p. 1211–1362, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA,, https://doi.org/10.1017/9781009157896.011, 2021.

Giles, M. B.: Multilevel Monte Carlo methods, Acta Numerica, 24, 259–328, https://doi.org/10.1017/S096249291500001X, 2015.

Goelzer, H., Nowicki, S., Edwards, T., Beckley, M., Abe-Ouchi, A., Aschwanden, A., Calov, R., Gagliardini, O., Gillet-Chaulet, F., Golledge, N. R., Gregory, J., Greve, R., Humbert, A., Huybrechts, P., Kennedy, J. H., Larour, E., Lipscomb, W. H., Le clec'h, S., Lee, V., Morlighem, M., Pattyn, F., Payne, A. J., Rodehacke, C., Rückamp, M., Saito, F., Schlegel, N., Seroussi, H., Shepherd, A., Sun, S., van de Wal, R., and Ziemen, F. A.: Design and results of the ice sheet model initialisation experiments initMIP-Greenland: an ISMIP6 intercomparison, The Cryosphere, 12, 1433–1460, https://doi.org/10.5194/tc-12-1433-2018, 2018.

Goldberg, D. N., Heimbach, P., Joughin, I., and Smith, B.: Committed retreat of Smith, Pope, and Kohler Glaciers over the next 30 years inferred by transient model calibration, The Cryosphere, 9, 2429–2446, https://doi.org/10.5194/tc-9-2429-2015, 2015.

Gorodetsky, A., Geraci, G., Eldred, M., and Jakeman, J.: A generalized approximate control variate framework for multifidelity uncertainty quantification, Journal of Computational Physics, 408, 109 257, https://doi.org/10.1016/j.jcp.2020.109257, 2020.

Halfar, P.: On the dynamics of the ice sheets 2, Journal of Geophysical Research: Oceans, 88, 6043–6051, https://doi.org/https://doi.org/10.1029/JC088iC10p06043, 1983.

Halko, N., Martinsson, P. G., and Tropp, J. A.: Finding Structure with Randomness: Probabilistic Algorithms for Constructing Approximate Matrix Decompositions, SIAM Review, 53, 217–288, https://doi.org/10.1137/090771806, 2011.

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

Hartland, T., Stadler, G., Perego, M., Liegeois, K., and Petra, N.: Hierarchical off-diagonal low-rank approximation of Hessians in inverse problems, with application to ice sheet model initialization, Inverse Problems, 39, 085 006, https://doi.org/10.1088/1361-6420/acd719, 2023.

He, Q., Perego, M., Howard, A. A., Karniadakis, G. E., and Stinis, P.: A hybrid deep neural operator/finite element method for ice-sheet modeling, Journal of Computational Physics, 492, 112 428, https://doi.org/https://doi.org/10.1016/j.jcp.2023.112428, 2023.

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.

Hillebrand, T. R., Hoffman, M. J., Perego, M., Price, S. F., and Howat, I. M.: The contribution of Humboldt Glacier, northern Greenland, to sea-level rise through 2100 constrained by recent observations of speedup and retreat, The Cryosphere, 16, 4679–4700, https://doi.org/10.5194/tc-16-4679-2022, 2022.

Hoffman, M. D. and Gelman, A.: The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo, Journal of Machine Learning Research, 15, 1593–1623, http://jmlr.org/papers/v15/hoffman14a.html, 2014.

Hoffman, M. J., Perego, M., Price, S. F., Lipscomb, W. H., Zhang, T., Jacobsen, D., Tezaur, I., Salinger, A. G., Tuminaro, R., and Bertagna, L.: MPAS-Albany Land Ice (MALI): a variable-resolution ice sheet model for Earth system modeling using Voronoi grids, Geoscientific Model Development, 11, 3747–3780, 2018.

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.

Jakeman, J.: PyApprox: A software package for sensitivity analysis, Bayesian inference, optimal experimental design, and multi-fidelity uncertainty quantification and surrogate modeling, Environmental Modelling & Software, 170, 105 825, https://doi.org/https://doi.org/10.1016/j.envsoft.2023.105825, 2023.

Jantre, S., Hoffman, M. J., Urban, N. M., Hillebrand, T., Perego, M., Price, S., and Jakeman, J. D.: Probabilistic projections of the Amery Ice Shelf catchment, Antarctica, under high ice-shelf basal melt conditions, EGUsphere, 2024, 1–45, https://doi.org/10.5194/egusphere-2024-1677, 2024.

Johnson, A., Aschwanden, A., Albrecht, T., and Hock, R.: Range of 21st century ice mass changes in the Filchner-Ronne region of Antarctica, Journal of Glaciology, p. 1–11, https://doi.org/10.1017/jog.2023.10, 2023.

Jouvet, G.: Mechanical error estimators for shallow ice flow models, Journal of Fluid Mechanics, 807, 40–61, https://doi.org/10.1017/jfm.2016.593, 2016.

Knutti, R., Furrer, R., Tebaldi, C., Cermak, J., and Meehl, G. A.: Challenges in Combining Projections from Multiple Climate Models, Journal of Climate, 23, 2739 – 2758, https://doi.org/https://doi.org/10.1175/2009JCLI3361.1, 2010.

Liegeois, K., Perego, M., and Hartland, T.: PyAlbany: A Python interface to the C++ multiphysics solver Albany, Journal of Computational and Applied Mathematics, 425, 115 037, https://doi.org/https://doi.org/10.1016/j.cam.2022.115037, 2023.

Lowery, M., Turnage, J., Morrow, Z., Jakeman, J., Narayan, A., Zhe, S., and Shankar, V.: Kernel Neural Operators (KNOs) for Scalable, Memory-efficient, Geometrically-flexible Operator Learning, https://arxiv.org/abs/2407.00809, 2024.

MacAyeal, D. R.: A tutorial on the use of control methods in ice-sheet modeling, Journal of Glaciology, 39, 91–98, https://doi.org/10.3189/S0022143000015744, 1993.

Morland, L. W. and Johnson, I. R.: Steady Motion of Ice Sheets, Journal of Glaciology, 25, 229–246, https://doi.org/10.3189/S0022143000010467, 1980.

https://doi.org/10.5194/egusphere-2024-2209
Preprint. Discussion started: 19 July 2024

Morlighem, M., Rignot, E., Seroussi, H., 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, Geophysical Research Letters, 37, https://doi.org/https://doi.org/10.1029/2010GL043853, 2010.

Nias, I. J., Cornford, S. L., Edwards, T. L., Gourmelen, N., and Payne, A. J.: Assessing Uncertainty in the Dynamical Ice Response to Ocean Warming in the Amundsen Sea Embayment, West Antarctica, Geophysical Research Letters, 46, 11 253–11 260, https://doi.org/10.1029/2019GL084941, 2019.

Nias, I. J., Nowicki, S., Felikson, D., and Loomis, B.: Modeling the Greenland Ice Sheet's Committed Contribution to Sea Level During the 21st Century, Journal of Geophysical Research: Earth Surface, 128, e2022JF006 914, https://doi.org/https://doi.org/10.1029/2022JF006914, e2022JF006914 2022JF006914, 2023.

Peherstorfer, B.: Multifidelity Monte Carlo Estimation with Adaptive Low-Fidelity Models, SIAM/ASA Journal on Uncertainty Quantification, 7, 579–603, https://doi.org/10.1137/17M1159208, 2019.

Peherstorfer, B., Willcox, K., and Gunzburger, M.: Optimal Model Management for Multifidelity Monte Carlo Estimation, SIAM Journal on Scientific Computing, 38, A3163–A3194, https://doi.org/10.1137/15M1046472, 2016.

Perego, M., Price, S., and Stadler, G.: Optimal initial conditions for coupling ice sheet models to Earth system models, Journal of Geophysical Research: Earth Surface, 119, 1894–1917, https://doi.org/https://doi.org/10.1002/2014JF003181, 2014.

Petra, N., Zhu, H., Stadler, G., Hughes, T. J., and Ghattas, O.: An inexact Gauss-Newton method for inversion of basal sliding and rheology parameters in a nonlinear Stokes ice sheet model, Journal of Glaciology, 58, 889–903, https://doi.org/10.3189/2012JoG11J182, 2012.

Ritz, C., Edwards, T. L., Durand, G., Payne, A. J., Peyaud, V., and Hindmarsh, R. C. A.: Potential sea-level rise from Antarctic ice-sheet instability constrained by observations, Nature, 528, 115–118, https://doi.org/10.1038/nature16147, 2015.

Schaden, D. and Ullmann, E.: On Multilevel Best Linear Unbiased Estimators, SIAM/ASA Journal on Uncertainty Quantification, 8, 601–635, https://doi.org/10.1137/19M1263534, 2020.

Schlegel, N.-J., Seroussi, H., Schodlok, M. P., Larour, E. Y., Boening, C., Limonadi, D., Watkins, M. M., Morlighem, M., and van den Broeke, M. R.: Exploration of Antarctic Ice Sheet 100-year contribution to sea level rise and associated model uncertainties using the ISSM framework, The Cryosphere, 12, 3511–3534, https://doi.org/10.5194/tc-12-3511-2018, 2018.

Seroussi, H., Verjans, V., Nowicki, S., Payne, A. J., Goelzer, H., Lipscomb, W. H., Abe-Ouchi, A., Agosta, C., Albrecht, T., Asay-Davis, X., et al.: Insights into the vulnerability of Antarctic glaciers from the ISMIP6 ice sheet model ensemble and associated uncertainty, The Cryosphere, 17, 5197–5217, https://doi.org/10.5194/tc-17-5197-2023, 2023.

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

Tezaur, I., Peterson, K., Powell, A., Jakeman, J., and Roesler, E.: Global sensitivity analysis of ultra-low resolution Energy Exascale Earth System Model (E3SM), Journal of Advances in Modeling Earth Systems, submitted, 2021.

Van Katwyk, P., Fox-Kemper, B., Seroussi, H., Nowicki, S., and Bergen, K. J.: A Variational LSTM Emulator of Sea Level Contribution From the Antarctic Ice Sheet, Journal of Advances in Modeling Earth Systems, 15, e2023MS003 899, https://doi.org/10.1029/2023ms003899, 2023.

Weis, M., Greve, R., and Hutter, K.: Theory of shallow ice shelves, Continuum Mechanics and Thermodynamics, 11, 15–50, https://doi.org/10.1007/s001610050102, 1999.

Yoo, M., Gopalan, G., Hoffman, M. J., Coulson, S., Han, H. K., Wikle, C. K., and Hillebrand, T.: Uncertainty-enabled machine learning for emulation of regional sea-level change caused by the Antarctic Ice Sheet, https://arxiv.org/abs/2406.17729, 2024.