West Greenland ice sheet retreat history reveals elevated precipitation during the Holocene thermal maximum

Jacob Downs[1], Jesse Johnson[2], Jason Briner[3], Nicolás Young[4], Alia Lesnek[3], and Josh Cuzzone[5]

[1]Department of Mathematical Sciences, University of Montana, Missoula MT, USA
[2]Department of Computer Science, University of Montana, Missoula MT, USA
[3]Department of Geology, University at Buffalo, Buffalo NY, USA
[4]Lamont-Doherty Earth Observatory, Columbia University, Palisades NY, USA
[5]Department of Earth System Science, University of California Irvine, Irvine CA, USA

Correspondence: Jacob Downs (jacob.downs@umontana.edu)

Abstract. We investigate changing precipitation patterns in the Kangerlussuaq region of west central Greenland during the Holocene thermal maximum, using a new chronology of ice sheet terminus position through the Holocene and a novel inverse modeling approach based on the unscented transform (UT). The UT is applied to estimate changes in annual precipitation in order to reduce the misfit between modeled and observed terminus positions. We demonstrate the effectiveness of the UT for time-dependent data assimilation, highlighting its low computational cost and trivial parallel implementation. Our results indicate that Holocene warming coincided with elevated precipitation, without which modeled retreat in the Kangerlussuaq region is more rapid than suggested by observations. Less conclusive is if high temperatures during the HTM were specifically associated with a transient increase in precipitation, as the results depend on the assumed temperature history. Our results highlight the important role that changing precipitation patterns had in controlling ice sheet extent during the Holocene.

1 Introduction

During the early Holocene (∼11.7 - 8 ka BP), terrestrial and marine climate proxies from the northern hemisphere reveal a warmer than present peak in temperature (Kaufman et al., 2004; Marcott et al., 2013). This period of elevated temperatures, likely initiated by greater than modern insolation, is referred to as the Holocene thermal maximum (HTM). Its onset, duration, and severity were likely spatially variable (Kaufman et al., 2004). Records of HTM warming can be found in Greenland ice core records. For example, temperatures measured in the Dye-3 borehole show a pronounced HTM signal occurring from 7 to 4 ka BP and having value of 2.5° above present temperatures (Dahl-Jensen et al., 1998; Miller et al., 2010), whereas at the GISP2 site, the HTM appears to occur slightly earlier, following the 8.2 ka BP cold event (Kobashi et al., 2017) (Figure 1).

While warming during the HTM is well established, less is known about the regional changes in precipitation that accompanied increased temperatures. Ice core records provide long term estimates of accumulation (Alley et al., 1993), but these point measurements near ice divides are not representative of the precipitation across the ice sheet, particularly at lower elevations near the coast. Because the HTM was accompanied by lower Arctic sea ice extent (Polyak et al., 2010), it is possible that additional moisture was available to the GrIS from open Arctic waters. This is supported by proxy evidence showing an increase

in winter precipitation in western Greenland coincident with HTM warming (Thomas et al., 2016). However, temperature is known with greater certainty than precipitation.

Understanding feedbacks between temperature and precipitation during the HTM has implications for the future of the GrIS. Warming and declining sea ice are projected to cause an increase in Arctic precipitation (Bintanja and Selten, 2014; Singarayer et al., 2006). On a global scale, the moisture content of the atmosphere increases by around 7% for every degree of warming, according to the Clausius-Clapeyron relation. On a regional scale, declining arctic sea ice is expected to cause changes in atmospheric circulation, bringing more moisture to the Arctic (Bintanja and Selten, 2014). While there are important differences between HTM and modern climate, the history of retreat in West Greenland may provide insights into how the GrIS will respond to a warmer and possibly wetter future climate.

Modeling studies indicate that Holocene retreat in land terminating regions of the GrIS were controlled primarily by surface mass balance rather than ice dynamics (Cuzzone et al., 2019; Lecavalier et al., 2014). Given the primary importance of surface mass balance in controlling modeled retreat, we explore the hypothesis that enhanced winter snowfall during the HTM may have slowed retreat by partially offsetting increased surface melt (Thomas et al., 2016). We investigate changes in precipitation in a land terminating sector of the western central GrIS, near Kangerlussuaq, taking advantage of a new chronology of ice sheet terminus position (Young et al., 2019, In Review) and a novel inverse modeling approach based on the unscented transform (UT) (Julier and Uhlmann, 1997). In particular, we use the UT to estimate changes in annual precipitation during the Holocene by reducing the misfit between modeled and observed terminus positions in a flowline ice dynamics model (Brinkerhoff et al., 2017) (section 2.1).

The inverse problem is posed as a Bayesian inference problem, and its solution involves estimating a non-Gaussian posterior probability distribution. Markov Chain Monte Carlo (MCMC) methods, such as Metropolis Hastings method (Chib and Greenberg, 1995), provide one means of solving the inference problem by generating random samples from the posterior distribution. Generating samples from the posterior, however, requires repeatedly running the ice dynamics model with different precipitation histories as input, which is intractable even for a relatively computationally inexpensive flowline model.

The unscented transform provides a computationally efficient and trivially parallelizable alternative to MCMC methods. The basic idea of the unscented transform is to use a small, fixed number of deterministic sample points in order to estimate the statistical moments (e.g. mean and covariance) of the posterior distribution. Sigma points, each of which represents a different precipitation history input to the ice dynamics model, are generated a priori (section 2.5.3). Consequently, all model runs can be performed simultaneously in parallel resulting in a 1000 fold speed up compared to MCMC methods.

2 Numerical methods for inference

2.1 Ice-sheet model

We use the 1D, isothermal, flowline model with higher-order momentum balance described in Brinkerhoff et al. (2017). The momentum conservation equations are simplified using the Blatter-Pattyn approximation, assuming hydrostatic pressure and

negligible vertical resistive stresses (Blatter, 1995; Pattyn, 2003). Default parameter values used in this work are specified in Table 1.

We adopt a linear sliding law of the form

$$\tau_b = \beta^2 N u_b \tag{1}$$

where τ_b is basal shear stress, β^2 is a constant basal traction parameter, N is effective pressure, and u_b is sliding speed. Based on borehole water pressure measurements in Wright et al. (2016), basal water pressure P_w is assumed to be a fixed fraction $P_{frac} = 0.85$ of the ice overburden pressure P_0. Effective pressure is given by

$$N = P_0 - P_{frac} P_0. \tag{2}$$

The basal traction parameter β^2 is tuned to minimize the misfit between modeled and observed surface velocities from Mouginot et al. (2017) for modern Isunnguata Sermia.

2.2 Flowline Selection and Moraine Age Constraints

To define the path followed by ice, we assume that flow follows the modern surface velocity field inland of the present day margin. In ice free regions, the direction of ice flow is inferred from bedrock topography (Figure 1). Since the direction of ice flow is unknown and time varying, we cannot directly quantify the uncertainty introduced by errors in flowline selection. To account for some of this uncertainty, we perform inversions on two plausible, adjacent paleo-flowlines in the Kangerlussuaq area.

The rate of Holocene retreat on each flowline is estimated using constraints on ice sheet terminus position from (Young et al., 2019, In Review) (Figure 1). Terminus position data in Young et al. (2019, In Review) indicates that in the early Holocene (11.6 ka BP), the ice sheet margin was some tens of kilometers inland of the present day coastline. Although the moraine patterns are spatially complex, generally speaking there was a period of moderate retreat ($\sim$10 km on the northern flowline and $\sim$30 km on the southern flowline) from 11.6 to 10.3 ka BP, followed by rapid retreat ($\sim$100 km on both flowlines) from 10.3 to 8.1 ka BP. By 8.1 ka BP, the margin position was within 20 km of its present position on both flowlines (Figure 5). The modern terminus position provides one additional constraint.

For modern bedrock geometry along the flowlines, we use BedMachine v3 (Morlighem et al., 2017). Isostatic uplift and relative sea level changes are accounted for using a Glacial isostatic adjustment model (Caron et al., 2018). This model, combined with the retreat chronology in Young et al. (2019, In Review), indicate that ice remained grounded on both the northern and southern flowlines from 11.6 ka BP onward.

2.3 Positive degree day model

Surface mass balance is estimated using a positive degree day (PDD) model (Johannesson et al., 1995). Annual surface mass balance is constructed in the PDD model using estimates of average monthly precipitation and temperature. Inputs into the PDD model include the unknown ice surface elevation S, modern monthly temperature T_m and precipitation P_m along the

flowlines, as well as the seasonal temperature anomaly ΔT. As in Cuzzone et al. (2019) , modern temperature and precipitation are computed as 30 year averages from 1980-2010 using Box (2013).

To assess the sensitivity of modeled retreat history to temperature, we perform experiments using temperature reconstructions from both Buizert et al. (2018) and Dahl-Jensen et al. (1998). For the spatially explicit Buizert et al. (2018) reconstruction, monthly temperature anomalies are computed as averages along the flowlines. In contrast, Dahl-Jensen et al. (1998) reconstruct temperature only at the GRIP and Dye-3 borehole locations (Figure 1). Since a full Holocene reconstruction is unavailable at the Dye-3 borehole site, which is closer to Kangerlussuaq, we use the temperature reconstruction at GRIP. The HTM is roughly 0.3° C warmer and 1,500 years later in Dahl-Jensen et al. (1998) than Buizert et al. (2018) (Figure 6)

A limitation of the Dahl-Jensen et al. (1998) reconstruction is that it does not resolve seasonal temperatures. To address this, we calculate the difference between monthly and mean annual temperatures in Buizert et al. (2018) and apply those offsets to the mean annual temperature at GRIP from Dahl-Jensen et al. (1998) to obtain monthly temperature anomalies.

Surface temperature T is computed monthly as

$$T = T_m + \Delta T + \alpha(S - S_m) \tag{3}$$

where S_m is the modern surface elevation, and $\alpha = 5\ ^\circ\text{C km}^{-1}$ is the lapse rate (Abe-Ouchi et al., 2007). Following Ritz et al. (2001) and Cuzzone et al. (2019), precipitation P along the flowline is determined by the Clausius-Clapeyron relation

$$P = P_T + \Delta P = P_m \exp\left(\lambda_p (T - T_m)\right) + \Delta P. \tag{4}$$

The term P_T accounts for changes in precipitation solely due to changes in temperature. Here $\lambda_p = 0.07$, which results in a 7% increase in precipitation for every 1° C increase in temperature above modern (Abe-Ouchi et al., 2007; Ritz et al., 2001).

The term P_T does not capture the effects of many unknown climate factors that may have caused dynamic, regional changes in Holocene precipitation. Therefore, we introduce a precipitation anomaly term ΔP, analogous to the temperature anomaly ΔT. This time-dependent function, which has units of meters water equivalent (m.w.e.) a^{-1}, is used to adjust precipitation uniformly across a flowline in order to reduce mismatch between modeled and observed terminus positions. Unlike ΔT, which can be inferred from ice cores, ΔP will be used as a control variable to be determined using the inverse methods detailed in section 2.5.3. Equations 3 and 4 provide a method of accounting for elevation changes through time and downscaling inputs to match the mesh resolution of the model ($\sim$ 1 km).

Positive degree days and snowfall are computed month-by-month based on mean monthly temperature and precipitation (Johannesson et al., 1995). Snow is melted first at a rate of 5×10^{-3} m.w.e. per degree day followed by ice at a rate of 8×10^{-3} m.w.e. per degree day. Snow melt is initially supposed to refreeze in the snowpack as superimposed ice. Runoff begins when the superimposed ice reaches a given fraction (60%) of the snow cover (Reeh, 1991). A listing of ice-flow and PDD model parameters is provided in Table 1, and all data sets used in the model are shown in Table 2.

2.4 Model Limitations

A limitation of our modeling approach is that we do not account for potential ice dynamical effects caused by changes in surface runoff or subglacial hydrology. Modeling melt water runoff would be difficult in a flowline model due to flux of melt

Table 1. Summary of primary model parameters used in this work. Default values are provided where applicable.

Description	Symbol	Value	Units
PDD Model Parameters			
Std. deviation	σ	5.5	C
Std. deviation, accumulation	σ_a	5 C	
Ablation rate, snow	λ_s	5×10^{-3}	m.w.e. C^{-1} d^{-1}
Ablation rate, ice	λ_i	8×10^{-3}	m.w.e. C^{-1} d^{-1}
Precipitation param.	λ_p	7×10^{-2}	C^{-1}
Superimposed ice fraction	p_{max}	0.6	-
Ice Flow Parameters			
Rate factor	A	3.5×10^{-25}	s^{-1} Pa^{-3}
Basal traction	β^2	1.2×10^{-3}	Pa a m^{-1}
Water pressure fraction	P_{frac}	0.85	-
Precipitation Prior			
Prior kernel variance	σ_p^2	5×10^{-4}	-
Prior time Scale	τ	8×10^{3}	-
Measurement Prior			
Prior kernel variance	σ_ℓ^2	1×10^{6}	-

Table 2. Citations for the primary data sets used in this work.

Data	Citation
Terminus position chronology	Young et al. (2019, In Review)
Bedrock elevation	Morlighem et al. (2017)
Modern ice surface velocity	Mouginot et al. (2017)
Modern precipitation	Box (2013)
Temperature reconstructions	Buizert et al. (2018)
	Dahl-Jensen et al. (1998)
Glacial Isostatic Adjustment	Caron et al. (2018)

water in and out of the path of ice flow. Another limitation of our model is that it is isothermal. Unless ice temperature is treated in a vertically averaged sense, resolving temperature would require a 2D mesh, which would considerably increase the computational cost of the model. We consider the consequences of this simplification in section 3.3, where we test sensitivity to the ice hardness parameter.

2.5 Data Assimilation Approach

In order to assess the initial mismatch between modeled and observed retreat histories, we perform a reference experiment with $\Delta P = 0$, and ΔT estimates from both Buizert et al. (2018) and Dahl-Jensen et al. (1998). To improve the fit to observations, we assimilate terminus position data to obtain improved estimates of Holocene precipitation anomalies. Previous modeling studies indicate that Holocene retreat in land-terminating sectors of the GrIS were dominated by surface mass balance rather than ice dynamics (Cuzzone et al., 2019; Lecavalier et al., 2014). Uncertainty in Holocene climate, and consequently surface mass balance, is therefore likely the primary cause of discrepancies between modeled and observed terminus positions.

In principle ΔP, ΔT, or both could be tuned to improve the fit between modeled and observed terminus positions. We focus on precipitation because it is more poorly constrained than temperature. In the upcoming sections, we introduce a framework for time-dependent data assimilation based on the unscented transform (UT). Sections 2.5.1 - 2.5.2 outline the basic tenets of the UT. Sections 2.5.3 - 2.5.7 outline how the UT can be applied to estimate precipitation anomalies.

2.5.1 Overview of the Unscented transform

In what follows, the notation $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{x_0}, P_x)$ means that $\boldsymbol{x}$ is a normally distributed random variable with mean vector $\boldsymbol{x_0}$ and covariance matrix P_x. Suppose that $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{x_0}, P_x)$, and $\mathcal{F} : \mathbb{R}^n \to \mathbb{R}^m$ is a nonlinear function. We would like to estimate the distribution $P(y)$ of the non-Gaussian random variable

$$\boldsymbol{y} = \mathcal{F}(\boldsymbol{x}) + \boldsymbol{\epsilon}. \tag{5}$$

where $\boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, R)$ is the measurement noise.

In general, the non-Gaussian distribution $P(\boldsymbol{y})$ can be approximated using Markov chain Monte Carlo (MCMC) methods such as the Metropolis-Hastings algorithm (Chib and Greenberg, 1995). However, if the nonlinear function is time-consuming to compute, generating thousands of MCMC samples is often intractable. As a computationally efficient alternative to MCMC methods, Julier and Uhlmann (1997) introduced a method for approximating the mean and covariance of $\boldsymbol{y}$ called the unscented transform (UT) [1].

The term unscented transform has been applied somewhat broadly to a family of methods that approximate the statistical moments of a non-Gaussian random variable using a small, deterministic set of sample points called sigma points. It is known primarily in the context of the unscented Kalman filter. However, the UT can be applied more generally as an alternative to traditional MCMC methods. Sigma points and weight sets are designed to accurately estimate moments of a transformed random variable using a minimal number of function evaluations.

A set of vectors, called sigma points, are chosen with the same weighted sample mean and weighted covariance structure as $\boldsymbol{x}$. There are many algorithms for generating sigma points sets with different numbers of points and orders of accuracy. A

[1] According to Jeffrey Uhlmann, the creator of the UT, the term "unscented" was inspired by a stick of deodorant and has no technical significance.

commonly used set of $2n+1$ sigma points is given by

$$\boldsymbol{\chi_i} = \begin{cases} \boldsymbol{x_0} & i = 0 \\ \boldsymbol{x_0} - \sqrt{n+\kappa}[\sqrt{P_x}]_i & i = 1, \cdots, n \\ \boldsymbol{x_0} + \sqrt{n+\kappa}[\sqrt{P_x}]_i & i = n+1, \cdots, 2n \end{cases} \tag{6}$$

with corresponding weights given by

$$w_i^{(m)} = w_i^{(c)} = \begin{cases} \kappa/(n+\kappa) & i = 0 \\ 1/2(n+\kappa) & \text{otherwise} \end{cases}. \tag{7}$$

The notation $\left[\sqrt{P_x}\right]_i$ refers to the i-th row of a matrix square root (typically computed by Cholesky factorization) of P_x, and κ is a free parameter controlling the scaling of the sigma points around the mean. Julier and Uhlmann (1997) recommend a default value of $\kappa = 3 - n$. However, κ can be fine tuned to reduce prediction errors for a given problem.

The nonlinear function $\mathcal{F}$ is applied to each sigma point to yield a set of transformed points

$$\boldsymbol{\mathcal{Y}_i} = \mathcal{F}(\boldsymbol{\chi_i}). \tag{8}$$

The mean $\bar{\boldsymbol{y}}$ and covariance matrix P_y of $\boldsymbol{y}$ are then estimated as weighted sums

$$\bar{\boldsymbol{y}} = \sum_{i=0}^{2n} w_i^{(m)} \boldsymbol{\mathcal{Y}_i} \tag{9}$$

$$P_y = \sum_{i=0}^{2n} w_i^{(c)} (\boldsymbol{\mathcal{Y}_i} - \bar{\boldsymbol{y}})(\boldsymbol{\mathcal{Y}_i} - \bar{\boldsymbol{y}})^T + R. \tag{10}$$

A visual example of this algorithm is shown in Figure 2.

2.5.2 Bayesian Inference using the UT

Given a measurement $\boldsymbol{y_o}$, we would like to estimate the posterior distribution

$$P(\boldsymbol{x}|\boldsymbol{y_o}) \propto P(\boldsymbol{y}|\boldsymbol{x})P(\boldsymbol{x}). \tag{11}$$

Using an approach called statistical linearization (Sarkka, 2013) the joint distribution for $[\boldsymbol{x}, \boldsymbol{y}]^T$ can be approximated by

$$\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{\mu} \end{bmatrix}, \begin{bmatrix} P_x & P_{xy} \\ P_{xy}^T & P_y \end{bmatrix} \right) \tag{12}$$

with

$$\boldsymbol{\mu} = \sum_i w_i^{(m)} \boldsymbol{\mathcal{Y}_i} \tag{13}$$

$$P_y = \sum_i w_i^{(c)} (\boldsymbol{\mathcal{Y}_i} - \boldsymbol{\mu})(\boldsymbol{\mathcal{Y}_i} - \boldsymbol{\mu})^T + R \tag{14}$$

$$P_{xy} = \sum_i w_i^{(c)} (\boldsymbol{\mathcal{X}_i} - \boldsymbol{x_0})(\boldsymbol{\mathcal{Y}_i} - \boldsymbol{\mu})^T. \tag{15}$$

Here, $\boldsymbol{\mathcal{X}_i}$ and $\boldsymbol{\mathcal{Y}_i} = \mathcal{F}(\boldsymbol{\mathcal{X}_i})$ are sigma points and transformed sigma points respectively. We distinguish between between mean and covariance weights, denoted by $w_i^{(m)}$ and $w_i^{(c)}$ respectively, as some methods use distinct weights for each. Matrices P_y and P_{xy} are known as the measurement covariance and cross covariance respectively.

Given a measurement $\boldsymbol{y_0}$, the joint distribution then easily yields a Gaussian approximation of the posterior distribution. Letting

$$K = P_{xy} P_y^{-1} \tag{16}$$

we have

$$\boldsymbol{x'} = \boldsymbol{x_o} + K\left[\boldsymbol{y_o} - \boldsymbol{\mu}\right] \tag{17}$$

$$P' = P_x - K P_y K^T \tag{18}$$

where $\boldsymbol{x'}$ and P' are approximations of the posterior mean and covariance respectively. Readers familiar with Kalman filters might recognize that $\boldsymbol{x'}$ and P' are computed using a Kalman update step given a measurement $\boldsymbol{y_o}$ and Kalman gain K (Sarkka, 2013).

2.5.3 Assimilating Glacier Length Observations

Time-dependent data assimilation using the UT involves running the ice sheet model with a set of different precipitation anomaly histories, each corresponding to a different sigma point. This is followed by a post-processing step, which incorporates the ice sheet terminus chronology data via a correction of the prior mean vector and covariance matrix. Implementation of the unscented transform is straightforward and easily parallelizable since each model run is independent. In the following section, we outline the mathematical details of this process.

We seek to find ΔP histories that match the observed retreat history on both flowlines. An optimal solution should reproduce the observed retreat history within uncertainty, while not overfitting the data. We discretize the problem by estimating the precipitation anomaly at times $t_1, t_2, \cdots, t_n$ with from 11.6 - 0 ka BP. In practice, we use a regular grid of 44 points, spaced roughly 250 years apart. Precipitation anomaly values at these time points are denoted by $\Delta p_1, \Delta p_2, \cdots, \Delta p_n$ respectively and assembled in a vector

$$\boldsymbol{\Delta p} = [\Delta p_1, \Delta p_2, \cdots, \Delta p_n]^T. \tag{19}$$

Given a multivariate Gaussian prior $\boldsymbol{\Delta p} \sim \mathcal{N}(\boldsymbol{\Delta p_0}, P_0)$, which encodes assumptions about the structure of the precipitation anomaly (section 2.5.4), we would like to estimate the mean and covariance of the posterior distribution

$$P(\boldsymbol{\Delta p}|\boldsymbol{\ell_0}) \propto P(\boldsymbol{\ell_0}|\boldsymbol{\Delta p})P(\boldsymbol{\Delta p}). \tag{20}$$

Here, the measurement vector

$$\boldsymbol{\ell_0} = [\ell_1, \ell_2, \cdots, \ell_m]^T \tag{21}$$

contains measured glacier lengths at discrete points in time. Our procedure for defining the measurement mean $\boldsymbol{y_o}$ and covariance R are discussed in section 2.5.5.

We can think of the ice sheet model as a function that maps precipitation anomaly inputs to glacier length outputs with some additive observation noise

$$\boldsymbol{\ell} = \mathcal{F}(\boldsymbol{\Delta p}) + \boldsymbol{\epsilon}. \tag{22}$$

Discrete precipitation anomaly values are linearly interpolated for input into the ice-dynamics model, which has time steps on the order of months. The function $\mathcal{F}$ returns glacier lengths at the same m discrete times as in $\boldsymbol{\ell_0}$. "

To predict the posterior distribution, we use the methods outlined in sections 2.5.1 and 2.5.2. First, sigma points are generated based on the prior distribution for $\boldsymbol{\Delta p}$. To reduce computational costs, we use a minimal set of $n+1 = 45$ sigma points $\boldsymbol{\mathcal{P}_i}$ with corresponding weights $w_i^{(m)} = w_i^{(c)}$ generated using the method presented in Menegaz et al. (2011). Their method includes one free parameter $0 < w_0 < 1$, which can be tuned to reduce prediction errors. While this method has a lower order of accuracy than other methods, we find it often produces comparable results to other larger sigma point sets in practice.

Sigma points are propagated through the model to obtain transformed points $\boldsymbol{\mathcal{L}_i} = \mathcal{F}(\boldsymbol{\mathcal{P}_i})$. In this context, sigma points $\boldsymbol{\mathcal{P}_i}$ correspond to different time-dependent precipitation anomaly histories, while the transformed points $\boldsymbol{\mathcal{L}_i}$ correspond to the resulting glacier length histories given those precipitation anomalies as input (Figure 3). The structure of the sigma points reflects the mathematical formulation of the Menegaz et al. (2011) sigma points. Hence they are not merely random samples from the prior distribution.

Transformed sigma points are computed simultaneously in parallel, using one core per sigma point. After all transformed sigma points have been computed, the mean and covariance of the posterior distribution are estimated as outlined in Section 2.5.2. In parallel, this procedure takes roughly the same amount of time as a single forward model run. Unlike a standard filtering approach to data assimilation, all measurement data is incorporated simultaneously rather than time step by time step. For that reason, the Kalman update step corrects the entire time-dependent precipitation history at once. Moreover, unlike in Kalman smoothing, we approximate the full posterior distribution rather rather than the probability distributions

$$\begin{matrix} P(\Delta p_i|\boldsymbol{\ell}) \\ i = 1, \cdots, n \end{matrix}. \tag{23}$$

Note that the variables Δp_k with $k \neq i$ are marginalized out of the Kalman smoothing distributions. The use of time-dependent sigma points distinguishes our approach from standard Kalman filtering or Kalman smoothing approaches, and does not rely on the assumptions that states (Δp_i's) and measurements (ℓ_i's) satisfy the Markov property.

2.5.4 Gaussian Process Prior for Regularization

We adopt a Gaussian process prior (Rasmussen, 2004) to control the temporal smoothness of ΔP. A Gaussian process can be thought of as a distribution over functions. That is, random samples from a Gaussian process are functions rather than individual points or vectors. A collection of random variables $\{f(t) : t \in \mathcal{T}\}$ is said to be drawn from a Gaussian process with mean function $m(\cdot)$ and covariance function $k(\cdot,\cdot)$ if for any finite set of elements $t_1, \cdots, t_n \in \mathcal{T}$, the random variables $f(t_1), \cdots, f(t_n)$ have the distribution

$$\boldsymbol{f} \sim \mathcal{N}(\boldsymbol{m}, K) \tag{24}$$

with

$$\boldsymbol{f} = [f(t_1), f(t_2), \cdots, f(t_n)]^T, \tag{25}$$

$$\boldsymbol{m} = [m(t_1), m(t_2), \cdots, m(t_n)]^T, \tag{26}$$

and

$$K = \begin{bmatrix} k(t_1,t_1) & \cdots & k(t_1,t_n) \\ \vdots & \ddots & \vdots \\ k(t_n,t_1) & \cdots & k(t_n,t_n) \end{bmatrix}. \tag{27}$$

The set $\mathcal{T}$ is called the index set, and specifies the domain of the Gaussian process. Here the index set consists of numbers on the real line representing points in time.

The prior distribution for $\boldsymbol{\Delta p}$ has mean vector $\boldsymbol{\Delta p_0}$ and covariance matrix $P_0 = K$. We use a squared exponential covariance function

$$k(t,\, t') = \sigma_p^2 \exp\left(\frac{(t-t')^2}{2\tau^2}\right) \tag{28}$$

where σ^2 is a scaling constant, and τ is a characteristic time scale. Variables Δp_i and Δp_k are more highly correlated the closer they are in time. In effect, this acts as a form of temporal regularization, in which smooth precipitation history functions are preferred over less smooth ones. We discuss the choice of the mean $\boldsymbol{\Delta p_0}$ in Section 3.2.

2.5.5 Measurement Mean and Variance

Observations of terminus position are available roughly every 1000 years between 11.6 and 7.2 ka BP, with a gap from 7.2 ka BP to present. In contrast, model time steps are on the order of months. Due to these disparate time scales, we use the following procedure to estimate the measurement mean $\boldsymbol{\ell_0}$ and covariance matrix R on a time scale more appropriate for the ice sheet model.

We define a Gaussian process prior of "candidate" glacier length histories $\ell_p(t)$ as follows. The mean function $\hat{\ell_p}(t)$ is obtained by linearly interpolating between glacier length observations. The Brownian covariance kernel is defined by

$$k_p(t,t') = \sigma_\ell^2 \min(t,t'). \tag{29}$$

Candidate retreat histories are generated by drawing random samples from the Gaussian process. Length histories are resampled so that the mean moraine formation times and uncertainties match the observations in Young et al. (2019, In Review). The average length and variance of these curves is computed at a series of time slices to obtain a plausible measurement mean $\boldsymbol{\ell_0}$ and diagonal measurement covariance matrix R (Figure 4).

2.5.6 Iterative Optimization Procedure

Optimizations are conducted in multiple passes. In the first pass, the measurement covariance matrix R is multiplied by a factor of $1/4$ so that the measurements are initially weighted more than the prior. This produces a reasonable fit to the data, even given a poor initial estimate of ΔP. The optimal precipitation anomaly from a given iteration is used as the prior mean in the next iteration. We use the same prior covariance matrix P_0 for regularization in each iteration. After two to three iterations, modeled and observed terminus positions match within measurement uncertainty (Section 3). In our experience, the results of iteration are not dependent on the choice of prior mean in the first iteration, but we find that convergence can be improved by choosing a sensible initial guess as in Section 3.2.

2.5.7 Approach to Sensitivity Testing

As described, the data assimilation method accounts for measurement but not model uncertainty. It can easily be extended to account for uncertainties in the ice flow and PDD model parameters. We define an augmented state vector

$$\boldsymbol{u} = [\boldsymbol{\Delta p},\, \boldsymbol{\theta}]^T \tag{30}$$

where $\boldsymbol{\theta}$ is a vector of scalar parameters including the natural logarithm of the rate factor for ice, the basal traction parameter, a parameter controlling precipitation scaling with temperature, and the PDD melt rate parameters for ice and snow

$$\boldsymbol{\theta} = [\ln(A),\, \beta^2,\, \lambda_p,\, \lambda_i,\, \lambda_s]^T. \tag{31}$$

The prior distribution for the augmented state vector is given by

$$\boldsymbol{u} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{\Delta p_o} \\ \boldsymbol{\theta_0} \end{bmatrix}, \begin{bmatrix} P_0 & 0 \\ 0 & \Theta \end{bmatrix} \right) \tag{32}$$

where $\boldsymbol{\theta_0}$ is the parameter mean vector and Θ is a diagonal matrix containing parameter variances.

The unscented transform is applied to the augmented function

$$\boldsymbol{\ell} = \hat{\mathcal{F}}(\boldsymbol{u}) + \boldsymbol{\epsilon} \tag{33}$$

to obtain estimates of the joint distribution for $[\boldsymbol{\Delta p}, \boldsymbol{\theta}, \boldsymbol{\ell}]^T$ and the conditional distribution for $[\boldsymbol{\Delta p}, \boldsymbol{\theta}]^T \mid \boldsymbol{\ell_0}$. Since parameters are included as state variables, sigma points reflect a variety of precipitation histories and parameter sets. A model run for a particular sigma point is initialized from an appropriate steady state using the parameter set for that point.

2.6 Model Initialization

Model runs are initialized by tuning the precipitation anomaly to obtain a steady state at 12.6 ka BP, with a margin position 5 km beyond the 11.6 ka BP moraine. We invert for a precipitation anomaly time series that forces a retreat of 5 km over 1000 years to obtain an initial ice sheet configuration with the correct terminus position at 11.6 ka BP. This initialization procedure is intended to ease the ice sheet out of steady state in order to avoid strong transient effects at the beginning of model runs.

3 Results

3.1 Reference Experiment

To assess the initial misfit between modeled and observed Holocene retreat, the model is forced with ΔT reconstructions from Buizert et al. (2018) and Dahl-Jensen et al. (1998) and a zero precipitation anomaly. Precipitation is scaled with temperature according to Equation 4, neglecting possible influences from transient changes in Arctic sea ice cover, atmospheric circulation, or other unknown climate factors. Modeled ice retreat is far more rapid than observed (Figure 7). Colder temperatures during the early Holocene (Figure 7 a) lead to a somewhat more plausible retreat history using the Dahl-Jensen et al. (1998) temperature forcing versus the Buizert et al. (2018) forcing. However, by 8 ka BP, ice has retreated inland of the present day margin in both reconstructions.

3.2 Precipitation Anomaly Inversions

We estimate precipitation anomalies on the northern and southern flowlines using both the Buizert et al. (2018) and Dahl-Jensen et al. (1998) temperature reconstructions. Given the rapid retreat in the reference experiment, we expect that a positive precipitation anomaly will be required to match observed terminus positions in the early Holocene. Therefore, in the first round of optimization, we assume a prior mean of the form

$$\Delta P = \frac{1}{2}(1 - \hat{\tau}) \tag{34}$$

where $\hat{\tau}$ is a rescaled time variable that increases from zero at 11.6 ka BP to one at 0 ka BP. The results of the iterative optimization procedure are insensitive to the prior mean selected in the first iteration. Using a sigma point scaling parameter $w_0 = 0.5$ for the Menegaz et al. (2011) sigma point set ensures that a wide region around the mean is explored in each iteration.

Positive precipitation anomalies are predicted throughout most of the Holocene for both temperature reconstructions (Figure 6). While differences between the northern and southern flowlines are relatively minor, there are significant differences in precipitation between the Buizert et al. (2018) and Dahl-Jensen et al. (1998) inversions. In the Buizert et al. (2018) inversion,

Table 3. Summary of primary model parameters used in this work. Default values are provided where applicable. Prior parameter uncertainties assumed for the sensitivity test are shown in the 2σ column.

Description	Symbol	Prior Mean	2σ Range	Posterior Mean	Units
Ablation rate, snow	λ_s	5×10^{-3}	3 - 7 $\times10^{-3}$	2.8 $\times10^{-3}$	m.w.e. C^{-1} d^{-1}
Ablation rate, ice	λ_i	8×10^{-3}	6 - 10 $\times10^{-3}$	8.9 $\times10^{-3}$	m.w.e. C^{-1} d^{-1}
Precipitation param.	λ_p	7 $\times10^{-2}$	5 - 9 $\times10^{-2}$	9.8 $\times10^{-2}$	C^{-1}
Rate factor	A	3.5 $\times10^{-25}$	2.1 - 5.7 $\times10^{-25}$	2 $\times10^{-25}$	s^{-1} Pa^{-3}
Basal traction	β^2	1.2×10^{-3}	1.1 - 1.3 $\times10^{-3}$	1.2 $\times10^{-3}$	Pa a m^{-1}

the largest precipitation anomalies (up to 1 m.w.e. a^{-1}) occur during the early Holocene. Precipitation remains relatively high during the HTM (10 - 6 ka BP), but dips before the 8.2 ka BP cold event. For the Dahl-Jensen et al. (1998) inversion, in constrast, ΔP is lower during the early Holocene. Unlike the Buizert et al. (2018) inversion, there is also a clear trend between HTM warming from 8.5 - 3 ka BP and increased ΔP (Figure 6 b).

Forcing the model with estimated precipitation anomalies yields plausible retreat histories for all four temperature history / flowline combinations. Modeled and observed retreat chronologies match within within uncertainty (Figure 7). A significant fraction of Holocene precipitation (typically > 90%) falls as snow. Hence, a positive precipitation anomaly can be interpreted directly as increasing snowfall / accumulation. Average HTM snowfall is around 35% higher than modern in both temperature reconstructions (Figure 8). However, overall trends in Holocene snowfall differ between reconstructions.

3.3 Sensitivity Testing

We assess the sensitivity of Holocene precipitation anomalies to modeling uncertainties by performing an HTM inversion using the methodology described in Section 2.5.7. To obtain accurate uncertainty estimates, we use a fifth order accurate sigma point set based on Li et al. (2017) (Appendix A), as we find that second order Menegaz et al. (2011) set likely underestimates covariance. Inversions are performed on the northern flowline using the Buizert et al. (2018) temperature reconstruction. Model runs are initialized from steady states around at 10.5 ka BP, 500 years prior to the Buizert et al. (2018) HTM. Prior and posterior parameter values are reported in Table 3.

Mean HTM precipitation anomalies are within 2 cm.w.e. a^{-1} of each other in the sensitivity test and the inversion in section3.2 (Figure 9). Parameter uncertainties contribute to uncertainty in ΔP. However, temperature uncertainty is far more significant than uncertainty in ice sheet or PDD model parameters. Differences in model initialization do not significantly impact the results of the inversion. Although sensitivity tests are initialized from steady states at 10.5 ka BP, while inversions in Section 3.2 are initialized from a transient state at 11.6 ka BP, the mean HTM precipitation anomalies are nearly identical. Estimated posterior parameter values are similar to the assumed prior values (Table 3).

4 Discussion

We infer changes in Holocene precipitation in the Kangerlussuaq region of western Greenland using a new chronology of ice sheet terminus position from Young et al. (2019, In Review) and an inverse modeling procedure based on the unscented transform. We find that scaling precipitation with temperature via equation (4) results in excessively fast retreat during the early Holocene for both the Buizert et al. (2018) and Dahl-Jensen et al. (1998) temperature reconstructions. Thus, temperature driven changes in precipitation / accumulation alone are not sufficient to reproduce the observed pattern of retreat. Inversions show that adding precipitation throughout the Holocene and specifically during the HTM yields a good fit to observations (Figures 6, 7). Since a large fraction of precipitation falls as snow, a positive precipitation anomaly can be interpreted as increasing the amount of accumulation along a flowline.

There are considerable differences in predicted precipitation anomalies ΔP depending on the assumed temperature history ΔT. Inversions using the Buizert et al. (2018) temperature reconstruction show generally decreasing snowfall through the Holocene. Due to high snowfall in the early Holocene (nearly 100% higher than modern) and a dip in HTM precipitation associated with the 8.2 ka BP cold event, average HTM snowfall roughly matches the overall Holocene average (about 35% above modern). In contrast, inversions using the Dahl-Jensen et al. (1998) temperature reconstruction show a clear trend between HTM warming and increased snowfall (Figure 8). This trend could be interpreted as a transient increase in snowfall due to lower sea ice cover or changes in atmospheric circulation in the Arctic during the HTM.

A large positive ΔP correction during the early Holocene in the Buizert et al. (2018) inversion likely reflects the dependence of precipitation on temperature. Low temperatures during the early Holocene result in low precipitation. Without an additional moisture source to increase snowfall, the ice sheet retreats far more rapidly than observed. Hence, a large ΔP correction, up to 1 m.w.e. a^{-1} at 11.6 ka BP, is necessary to slow retreat (Figure 6). Warming is more gradual during the early Holocene in the Dahl-Jensen et al. (1998) reconstruction. While accumulation is low, it is more closely balanced with ablation, resulting in a smaller precipitation anomaly correction from 11.6 to 10 ka BP. These trends in the early Holocene demonstrate a complex, nonlinear interplay between accumulation and ablation.

When interpreting results of precipitation inversions, it is important to consider that the Buizert et al. (2018) and Dahl-Jensen et al. (1998) reconstructions are obtained using different methodologies. The Greenland-wide Buizert et al. (2018) reconstruction is obtained by merging the TraCE-21ka coupled ocean-atmosphere general circulation model of the last deglaciation **??** with borehole temperature reconstructions at GISP2, NGRIP, and NEEM. In the absence of additional borehole constraints on temperature, the spatial pattern of temperature near the ice sheet margins is based mainly on TraCE-21ka output. The Dahl-Jensen et al. (1998) reconstruction at GRIP, in constrast, offers only a pointwise estimate containing no explicit spatial or seasonal information.

While the Buizert et al. (2018) temperature reconstruction is arguably more suitable for our purposes, since it resolves spatial and seasonal patterns in ΔT, there is still considerable uncertainty in temperature, particularly near the ice margins. Despite this, there is some consensus between the Buizert et al. (2018) and Dahl-Jensen et al. (1998) inversions. Generally speaking,

positive precipitation anomalies are needed to slow retreat through the early to middle Holocene. Although the timing of the HTM differs between reconstructions, average HTM snowfall is roughly 35% higher than modern for both (Figure 8).

Predicted HTM precipitation anomalies are not particularly sensitive to uncertainties in the ice sheet or PDD model parameters, or to the initialization procedure. In the sensitivity test presented in Section 3.3, model runs are initialized using steady states at 10.5 ka BP. Runs in the full Holocene inversions in Section 3.2, which use fixed parameter sets, are initialized from a transient state at 11.6 ka BP. Even considering these differences in model initialization, estimated HTM precipitation anomalies are comparable in all inversions (Figure 9). This result supports previous findings by Lecavalier et al. (2014) and Cuzzone et al. (2019) that modeled Holocene retreat in ground terminating regions is dominated by surface mass balance rather than ice dynamics.

4.1 Conclusions

4.1.1 The Unscented Transform as a Data Assimilation Method

Significant strides have been made in time-dependent data assimilation in glaciology using adjoint based methods. Goldberg and Heimbach (2013) infer the initial thickness and basal conditions for a synthetic ice sheet given snapshots of ice thickness at discrete times. Larour et al. (2014) demonstrate a data assimilation framework within the Ice Sheet System Model (ISSM), capable of obtaining temporal estimates of surface mass balance and basal friction given surface altimetry.

In contrast to adjoint based approaches, the unscented transform (UT) does not require computing the Jacobian or Hessian of an objective function, or special checkpointing code for time-dependent problems. Adjoint based methods are advantageous for extremely high-dimensional problems, as the required number of model runs is independent of the number of parameters. However, the UT provides more accurate uncertainty estimates than linearization (Julier and Uhlmann, 1997). Hessian information can be used to improve uncertainty estimates (Isaac et al., 2015). However, this methodology uses purely local approximations to the nonlinear function around the maximum a posteriori probability (MAP) estimator, which may affect the quality of uncertainty estimates.

The unscented transform has advantages over Markov chain Monte Carlo methods for inference problems with a relatively small number of unknown parameters (< 1000 or so parameters). In this work, we optimize for $n = 44$ parameters representing precipitation anomaly values at discrete points in time. Since function evaluations at each sigma point are independent, the UT is trivially parallelizable. Consequently, on a high end desktop, the iterative optimization process presented in Section 2.5.6 takes roughly the same amount of time as performing three successive forward model runs.

MCMC methods can be parallelized to some extent by utilizing multiple interacting Markov chains running in parallel (e.g. Chowdhury and Jermaine, 2018) or by combining samples from independent chains in a post processing step (e.g. Neiswanger et al., 2013). Nonetheless, computational bottlenecks persist in MCMC methods since individual Markov chains are inherently serial. Consequently, UT approximations to the posterior distribution can be generated hundreds to thousands of times faster for small parameter sets.

A drawback of the UT is that its accuracy is inherently limited by using a predetermined number of sample points. In our case, for example, using a minimal set of $n+1$ sigma points for n unknown parameters yields accurate mean estimates but appears to underestimate covariance when compared to a higher order cubature method. Without testing multiple sigma point sets and scaling parameters, it can be difficult to assess the accuracy of UT estimates of the posterior. MCMC methods, in contrast, can provide arbitrarily accurate estimates of the posterior given sufficient computation time. Moreover, they can resolve the full posterior distribution rather than computing its moments as in the UT.

An alternative to traditional MCMC methods is to use surrogate models or emulators (Gong and Duan, 2017). Here, a computationally inexpensive surrogate model is trained to approximate the output of a more complex model function. The surrogate model can then be used in place of the full model for the purpose of MCMC sampling, significantly reducing the overall computational cost. Learning a surrogate model requires an initial training stage. For inference problems with Gaussian priors, sigma points for UT might provide a reasonable initial set of training points. Exploring surrogate modeling approaches to solving inference problems in glaciology is a promising avenue for future research.

4.1.2 Modeling Conclusions

Our work follows a number of previous observationally constrained paleo ice sheet modeling studies (e.g. Tarasov and Peltier, 2002; Lecavalier et al., 2014; Calov et al., 2015). Perhaps most relevant to this work is Lecavalier et al. (2014), who model the deglaciation of Greenland from the Last Glacial Maximum using a 3-D thermomechanically coupled ice sheet model. Model runs in Lecavalier et al. (2014) are informed by constraints on relative sea level, ice core thinning, and LGM ice sheet extent.

In contrast to previous modeling studies, the computational efficiency of the flowline model outlined in Brinkerhoff et al. (2017) makes time-dependent data-assimilation, sensitivity testing, and robust uncertainty estimation tractable. Sensitivity testing indicates that estimated precipitation is insensitive to parameter uncertainties in the PDD and ice-dynamics models. This conclusion supports earlier findings showing that modeled Holocene retreat in land terminating sectors of the GrIS is more sensitive to surface mass balance than other factors like the flow law or basal sliding (Cuzzone et al., 2019; Lecavalier et al., 2014).

A drawback of our modeling approach is that we cannot account for inherently map plane effects such as changes in ice flow direction, or convergent and divergent flow in or out of the assumed flow path. These factors likely contribute to small discrepancies in estimated precipitation anomalies between the northern and southern flowlines (Figure 6).

In this work, we do not treat temperature as a random variable with its own covariance structure. However, differences between the Buizert et al. (2018) and Dahl-Jensen et al. (1998) inversions indicate that temperature is the dominant source of uncertainty in ΔP. This result underscores the importance of generating improved, regionally specific, temperature reconstructions constrained by proxy records. More regionally specific estimates of temperature would help to decrease uncertainty in the estimated precipitation history.

Despite lingering uncertainties, our modeling results indicate that the Holocene thermal maximum was accompanied by elevated snowfall, which slowed ice retreat in the Kangerlussuaq region of the GrIS. Inversions conducted using both the Buizert et al. (2018) and Dahl-Jensen et al. (1998) temperature reconstructions show average HTM snowfall around 35%

higher than modern. In the Dahl-Jensen et al. (1998) inversion specifically, there is also a clear trend between HTM warming and increased accumulation. This could be interpreted as transient increase in precipitation due to reduced Arctic sea ice or changes in atmospheric circulation during the HTM.

Code availability. The flowline ice sheet model and PDD model used in this work are available at https://github.com/JacobDowns/flow_model.

Appendix A: A Higher Order Method for Estimating Covariance

Estimating the mean and covariance of the posterior distribution requires approximating Gaussian weighted expectation integrals of the form

$$E[\mathcal{F}(\boldsymbol{x})] = \int_{\mathbb{R}^n} \mathcal{F}(\boldsymbol{x})\mathcal{N}(\mathbf{0}, I)\, d\boldsymbol{x} \tag{A1}$$

via numerical integratation rules, also called cubature rules, which approximate the expectation integral as a weighted sum

$$E[\mathcal{F}(\boldsymbol{x})] \approx Q[\mathcal{F}(\boldsymbol{x})] = \sum_{i=1}^{N} w_i \mathcal{F}(\boldsymbol{\chi_i}). \tag{A2}$$

General Gaussian weight functions $\mathcal{N}(\boldsymbol{x_0}, P_x)$ are handled by changing variables. Letting $\sqrt{P_x}$ be a matrix square root of the covariance matrix, we have

$$\begin{aligned} &\int_{\mathbb{R}^n} \mathcal{F}(\boldsymbol{x})\mathcal{N}(\boldsymbol{x_0}, P_x)\, d\boldsymbol{x} \\ &= \int_{\mathbb{R}^n} \mathcal{F}(\boldsymbol{x_0} + \sqrt{P_x}\,\boldsymbol{\xi})\mathcal{N}(\mathbf{0}, I)\, d\boldsymbol{\xi}. \end{aligned} \tag{A3}$$

Cubature rules, including the unscented transform, are constructed to exactly integrate polynomial functions $\mathcal{F}(\boldsymbol{x})$ up to a certain degree d. Suppose that $\boldsymbol{x} = [x_1, x_2, \cdots, x_n]^T$ is a point in $\mathbb{R}^n$. A monomial of degree d refers to a function $x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$ where the exponents are non-negative integers that sum to d. A polynomial of degree d is a linear combination of monomials with highest degree d.

Li et al. (2017) describe a fifth-order cubature rule using fully symmetric sets of sigma points. A set $X = \{\boldsymbol{\chi_1}, \boldsymbol{\chi_2}, \cdots, \boldsymbol{\chi_N}\}$ is fully symmetric if it is closed under the operations of coordinate position and sign permutations. Their cubature rule has the form

$$\begin{aligned} Q[\mathcal{F}(\boldsymbol{x})] &= w_1 \mathcal{F}(0, 0, \cdots, 0) \\ &+ w_2 \sum_{\text{full sym}} \mathcal{F}(\lambda, 0, \cdots, 0) \\ &+ w_3 \sum_{\text{full sym}} \mathcal{F}(\lambda, \lambda, 0, \cdots, 0). \end{aligned} \tag{A4}$$

The notation $\sum_{\text{full sym}} \mathcal{F}(\cdot)$ refers to a sum of the function $\mathcal{F}$ evaluated at all points in the fully symmetric set generated by the given point.

Due to the symmetry of the sigma points and the Gaussian weight function, all moments (that is, integrals of Gaussian weighted monomial functions) containing an odd order exponent are automatically satisfied. Exploiting this fact, and the symmetries of the sigma points, it can be shown that satisfying the remaining moment constraint equations up the fifth order reduces to solving the following system of four equations in four unknowns w_1, w_2, w_3 and λ

$$\begin{aligned} E[1] &= w_1 + 2nw_2 + 2n(n-1)w_3 \\ E[x_i^2] &= 2\lambda^2 w_2 + 4(n-1)\lambda^2 w_3 \\ E[x_i^4] &= 2\lambda^4 w_2 + 4(n-1)\lambda^4 w_3 \\ E[x_i^2 x_j^2] &= 4\lambda^4 w_3. \end{aligned} \tag{A5}$$

By slightly modifying this cubature rule

$$\begin{aligned} Q[\mathcal{F}(\boldsymbol{x})] &= w_1 \mathcal{F}(0, 0, \cdots, 0) \\ &+ w_2 \sum_{\text{full sym}} \mathcal{F}(\lambda_1, 0, \cdots, 0) \\ &+ w_3 \sum_{\text{full sym}} \mathcal{F}(\lambda_2, \lambda_2, 0, \cdots, 0) \end{aligned} \tag{A6}$$

we introduce a new free parameter λ_2 that allows scaling of the sigma points about the mean. The new moment constraint equations become

$$\begin{aligned} E[1] &= w_1 + 2nw_2 + 2n(n-1)w_3 \\ E[x_i^2] &= 2\lambda_1^2 w_2 + 4(n-1)\lambda_2^2 w_3 \\ E[x_i^4] &= 2\lambda_1^4 w_2 + 4(n-1)\lambda_2^4 w_3 \\ E[x_i^2 x_j^2] &= 4\lambda_2^4 w_3. \end{aligned} \tag{A7}$$

Using that $E[1] = 1$, $E[x_i^2] = 1$, $E[x_i^4] = 3$, and $E[x_i^2 x_j^2] = 1$ we obtain

$$\begin{aligned} \lambda_1 &= \frac{\lambda_2 \sqrt{n-4}}{n - \lambda_2^2 - 1} \\ w_2 &= \frac{4-n}{2\lambda_1^4} \\ w_3 &= \frac{1}{4\lambda_2^2} \end{aligned} \tag{A8}$$

$$w_1 = 1 - 2nw_2 - 2n(n-1)w_3$$

with $n > 4$ and $n - \lambda_2^2 - 1 \neq 0$. A drawback of the original cubature rule, as well the modification of it presented here, is that it requires negative weights, which can lead to numerical instability.

Author contributions. JD was the lead author of this manuscript. JJ and JC assisted JD with ice sheet modeling efforts. JB, NY, and AL provided the moraine age data used in this work. JD and JJ wrote the manuscript with input from all authors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. JD and JJ were supported by NSF grant 1504457. JD would like to thank Doug Brinkerhoff for providing code for the flowline ice sheet model used in this work, as well as many elucidating discussions. JD would also like to thank John Bardsley for providing many mathematical insights into inverse problems. Our thanks to reviewers Stephen Price, Lev Tarasov, and editor Kerim Hestnes Nisancioglu, who helped to significantly improve the quality of this manuscript.

References

Abe-Ouchi, A., Segawa, T., and Saito, F.: Climatic conditions for modelling the Northern Hemisphere ice sheets throughout the ice age cycle, https://doi.org/10.5194/cp-3-423-2007, 2007.

Alley, R. B., Meese, D. A., Shuman, C. A., Gow, A. J., Taylor, K. C., Grootes, P. M., White, J. W., Ram, M., Waddington, E. D., Mayewski, P. A., and Zielinski, G. A.: Abrupt increase in Greenland snow accumulation at the end of the Younger Dryas event, Nature, https://doi.org/10.1038/362527a0, 1993.

Bintanja, R. and Selten, F. M.: Future increases in Arctic precipitation linked to local evaporation and sea-ice retreat, Nature, https://doi.org/10.1038/nature13259, 2014.

Blatter, H.: Velocity and stress fields in grounded glaciers: a simple algorithm for including deviatoric stress gradients, Journal of Glaciology, https://doi.org/10.1017/S002214300001621X, 1995.

Box, J. E.: Greenland ice sheet mass balance reconstruction. Part II: Surface mass balance (1840-2010), Journal of Climate, https://doi.org/10.1175/JCLI-D-12-00518.1, 2013.

Brinkerhoff, D., Truffer, M., and Aschwanden, A.: Sediment transport drives tidewater glacier periodicity, Nature Communications, https://doi.org/10.1038/s41467-017-00095-5, 2017.

Buizert, C., Keisling, B. A., Box, J. E., He, F., Carlson, A. E., Sinclair, G., and DeConto, R. M.: Greenland-Wide Seasonal Temperatures During the Last Deglaciation, Geophysical Research Letters, https://doi.org/10.1002/2017GL075601, 2018.

Calov, R., Robinson, A., Perrette, M., and Ganopolski, A.: Simulating the Greenland ice sheet under present-day and palaeo constraints including a new discharge parameterization, Cryosphere, https://doi.org/10.5194/tc-9-179-2015, 2015.

Caron, L., Ivins, E. R., Larour, E., Adhikari, S., Nilsson, J., and Blewitt, G.: GIA Model Statistics for GRACE Hydrology, Cryosphere, and Ocean Science, Geophysical Research Letters, https://doi.org/10.1002/2017GL076644, 2018.

Chib, S. and Greenberg, E.: Understanding the metropolis-hastings algorithm, American Statistician, https://doi.org/10.1080/00031305.1995.10476177, 1995.

Chowdhury, A. and Jermaine, C.: Parallel and Distributed MCMC via Shepherding Distributions, Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics, 84, 1819–1827, http://proceedings.mlr.press/v84/chowdhury18a.html, 2018.

Cuzzone, J. K., Schlegel, N. J., Morlighem, M., Larour, E., Briner, J. P., Seroussi, H., and Caron, L.: The impact of model resolution on the simulated Holocene retreat of the southwestern Greenland ice sheet using the Ice Sheet System Model (ISSM), Cryosphere, https://doi.org/10.5194/tc-13-879-2019, 2019.

Dahl-Jensen, D., Mosegaard, K., Gundestrup, N., Clow, G. D., Johnsen, S. J., Hansen, A. W., and Balling, N.: Past temperatures directly from the Greenland Ice Sheet, Science, https://doi.org/10.1126/science.282.5387.268, 1998.

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

Gong, W. and Duan, Q.: An adaptive surrogate modeling-based sampling strategy for parameter optimization and distribution estimation (ASMO-PODE), Environmental Modelling and Software, 95, 61–75, https://doi.org/10.1016/j.envsoft.2017.05.005, 2017.

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, https://doi.org/10.1016/j.jcp.2015.04.047, 2015.

Johannesson, T., Sigurdsson, O., Laumann, T., and Kennett, M.: Degree-day glacier mass-balance modelling with applications to glaciers in Iceland, Norway and Greenland, Journal of Glaciology, https://doi.org/10.1017/S0022143000016221, 1995.

Julier, S. J. and Uhlmann, J. K.: New extension of the Kalman filter to nonlinear systems, Int Symp AerospaceDefense Sensing Simul and Controls, https://doi.org/10.1117/12.280797, 1997.

Kaufman, D. S., Ager, T. A., Anderson, N. J., Anderson, P. M., Andrews, J. T., Bartlein, P. J., Brubaker, L. B., Coats, L. L., Cwynar, L. C., Duvall, M. L., Dyke, A. S., Edwards, M. E., Eisner, W. R., Gajewski, K., Geirsdóttir, A., Hu, F. S., Jennings, A. E., Kaplan, M. R., Kerwin, M. W., Lozhkin, A. V., MacDonald, G. M., Miller, G. H., Mock, C. J., Oswald, W. W., Otto-Bliesner, B. L., Porinchu, D. F., Rühland, K., Smol, J. P., Steig, E. J., and Wolfe, B. B.: Holocene thermal maximum in the western Arctic (0-180 W), Quaternary Science Reviews, https://doi.org/10.1016/j.quascirev.2003.09.007, 2004.

Kobashi, T., Menviel, L., Jeltsch-Thömmes, A., Vinther, B. M., Box, J. E., Muscheler, R., Nakaegawa, T., Pfister, P. L., Döring, M., Leuenberger, M., Wanner, H., and Ohmura, A.: Volcanic influence on centennial to millennial Holocene Greenland temperature change, Scientific Reports, https://doi.org/10.1038/s41598-017-01451-7, 2017.

Larour, E., Utke, J., Csatho, B., Schenk, A., Seroussi, H., Morlighem, M., Rignot, E., Schlegel, N., and Khazendar, A.: Inferred basal friction and surface mass balance of the Northeast Greenland Ice Stream using data assimilation of ICESat (Ice Cloud and land Elevation Satellite) surface altimetry and ISSM (Ice Sheet System Model), Cryosphere, https://doi.org/10.5194/tc-8-2335-2014, 2014.

Lecavalier, B. S., Milne, G. A., Simpson, M. J., Wake, L., Huybrechts, P., Tarasov, L., Kjeldsen, K. K., Funder, S., Long, A. J., Woodroffe, S., Dyke, A. S., and Larsen, N. K.: A model of Greenland ice sheet deglaciation constrained by observations of relative sea level and ice extent, Quaternary Science Reviews, 102, 54–84, https://doi.org/10.1016/j.quascirev.2014.07.018, 2014.

Li, Z., Yang, W., Ding, D., and Liao, Y.: A Novel Fifth-Degree Cubature Kalman Filter for Real-Time Orbit Determination by Radar, Mathematical Problems in Engineering, https://doi.org/10.1155/2017/8526804, 2017.

Marcott, S. A., Shakun, J. D., Clark, P. U., and Mix, A. C.: A reconstruction of regional and global temperature for the past 11,300 years, Science, 339, 1198–1201, https://doi.org/10.1126/science.1228026, 2013.

Menegaz, H. M., Ishihara, J. Y., and Borges, G. A.: A new smallest sigma set for the Unscented Transform and its applications on SLAM, in: Proceedings of the IEEE Conference on Decision and Control, https://doi.org/10.1109/CDC.2011.6161480, 2011.

Miller, G. H., Brigham-Grette, J., Alley, R. B., Anderson, L., Bauch, H. A., Douglas, M. S., Edwards, M. E., Elias, S. A., Finney, B. P., Fitzpatrick, J. J., Funder, S. V., Herbert, T. D., Hinzman, L. D., Kaufman, D. S., MacDonald, G. M., Polyak, L., Robock, A., Serreze, M. C., Smol, J. P., Spielhagen, R., White, J. W., Wolfe, A. P., and Wolff, E. W.: Temperature and precipitation history of the Arctic, Quaternary Science Reviews, https://doi.org/10.1016/j.quascirev.2010.03.001, 2010.

Morlighem, M., Williams, C. N., Rignot, E., An, L., Arndt, J. E., Bamber, J. L., Catania, G., Chauché, N., Dowdeswell, J. A., Dorschel, B., Fenty, I., Hogan, K., Howat, I., Hubbard, A., Jakobsson, M., Jordan, T. M., Kjeldsen, K. K., Millan, R., Mayer, L., Mouginot, J., Noël, B. P., O'Cofaigh, C., Palmer, S., Rysgaard, S., Seroussi, H., Siegert, M. J., Slabon, P., Straneo, F., van den Broeke, M. R., Weinrebe, W., Wood, M., and Zinglersen, K. B.: BedMachine v3: Complete Bed Topography and Ocean Bathymetry Mapping of Greenland From Multibeam Echo Sounding Combined With Mass Conservation, Geophysical Research Letters, https://doi.org/10.1002/2017GL074954, 2017.

Mouginot, J., Rignot, E., Scheuchl, B., and Millan, R.: Comprehensive annual ice sheet velocity mapping using Landsat-8, Sentinel-1, and RADARSAT-2 data, Remote Sensing, https://doi.org/10.3390/rs9040364, 2017.

Neiswanger, W., Wang, C., and Xing, E.: Asymptotically Exact, Embarrassingly Parallel MCMC, Proceeding UAI'14 Proceedings of the Thirtieth Conference on Uncertainty in Artificial Intelligence, pp. 1–16, http://arxiv.org/abs/1311.4780, 2013.

Pattyn, F.: A new three-dimensional higher-order thermomechanical ice sheet model: Basic sensitivity, ice stream development, and ice flow across subglacial lakes, Journal of Geophysical Research, https://doi.org/10.1029/2002JB002329, 2003.

Polyak, L., Alley, R. B., Andrews, J. T., Brigham-Grette, J., Cronin, T. M., Darby, D. A., Dyke, A. S., Fitzpatrick, J. J., Funder, S., Holland, M., Jennings, A. E., Miller, G. H., O'Regan, M., Savelle, J., Serreze, M., St. John, K., White, J. W., and Wolff, E.: History of sea ice in the Arctic, Quaternary Science Reviews, https://doi.org/10.1016/j.quascirev.2010.02.010, 2010.

Rasmussen, C. E.: Gaussian Processes in machine learning, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2004.

Reeh, N.: Parameterization of Melt Rate and Surface Temperature on the Greenland Ice Sheet, Polarforschung, https://doi.org/10.1007/978-3-7091-0973-1_2, 1991.

Ritz, C., Rommelaere, V., and Dumas, C.: Modeling the evolution of Antarctic ice sheet over the last 420,000 years: Implications for altitude changes in the Vostok region, Journal of Geophysical Research Atmospheres, https://doi.org/10.1029/2001JD900232, 2001.

Sarkka, S.: Bayesian Filtering and Smoothing, Cambridge University Press, https://doi.org/10.1017/CBO9781139344203, 2013.

Singarayer, J. S., Bamber, J. L., and Valdes, P. J.: Twenty-first-century climate impacts from a declining Arctic sea ice cover, Journal of Climate, https://doi.org/10.1175/JCLI3649.1, 2006.

Tarasov, L. and Peltier, W. R.: Greenland glacial history and local geodynamic consequences, Geophysical Journal International, https://doi.org/10.1046/j.1365-246X.2002.01702.x, 2002.

Thomas, E. K., Briner, J. P., Ryan-Henry, J. J., and Huang, Y.: A major increase in winter snowfall during the middle Holocene on western Greenland caused by reduced sea ice in Baffin Bay and the Labrador Sea, Geophysical Research Letters, https://doi.org/10.1002/2016GL068513, 2016.

Wright, P. J., Harper, J. T., Humphrey, N. F., and Meierbachtol, T. W.: Measured basal water pressure variability of the western Greenland Ice Sheet: Implications for hydraulic potential, Journal of Geophysical Research: Earth Surface, https://doi.org/10.1002/2016JF003819, 2016.

Young, N. E., Briner, J. P., Miller, G. H., Crump, S. E., Lesnek, A. J., and Thomas, E. K.: Arctic deglaciation interrupted by abrupt coolings, Science Advances, 2019, In Review.

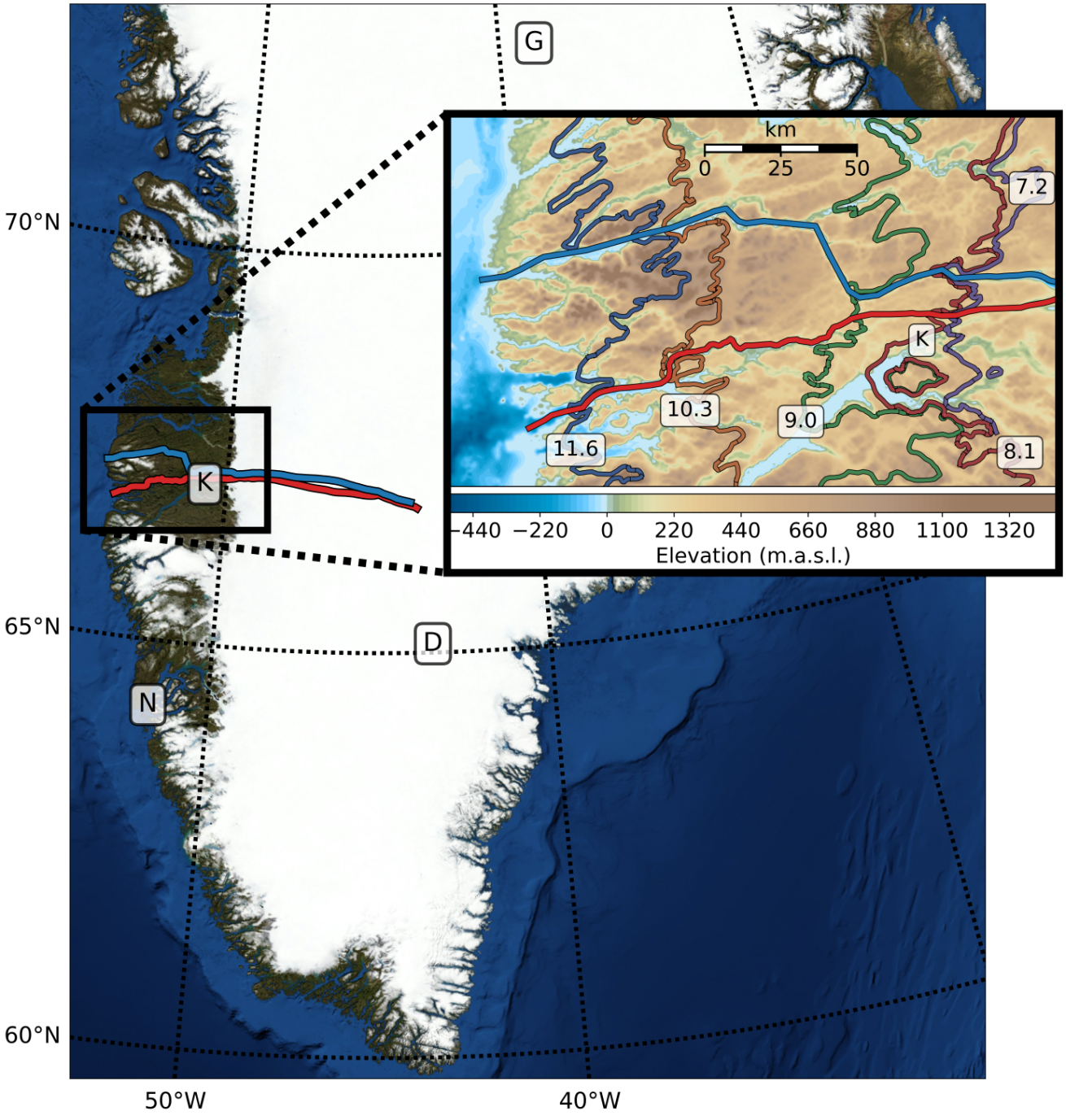
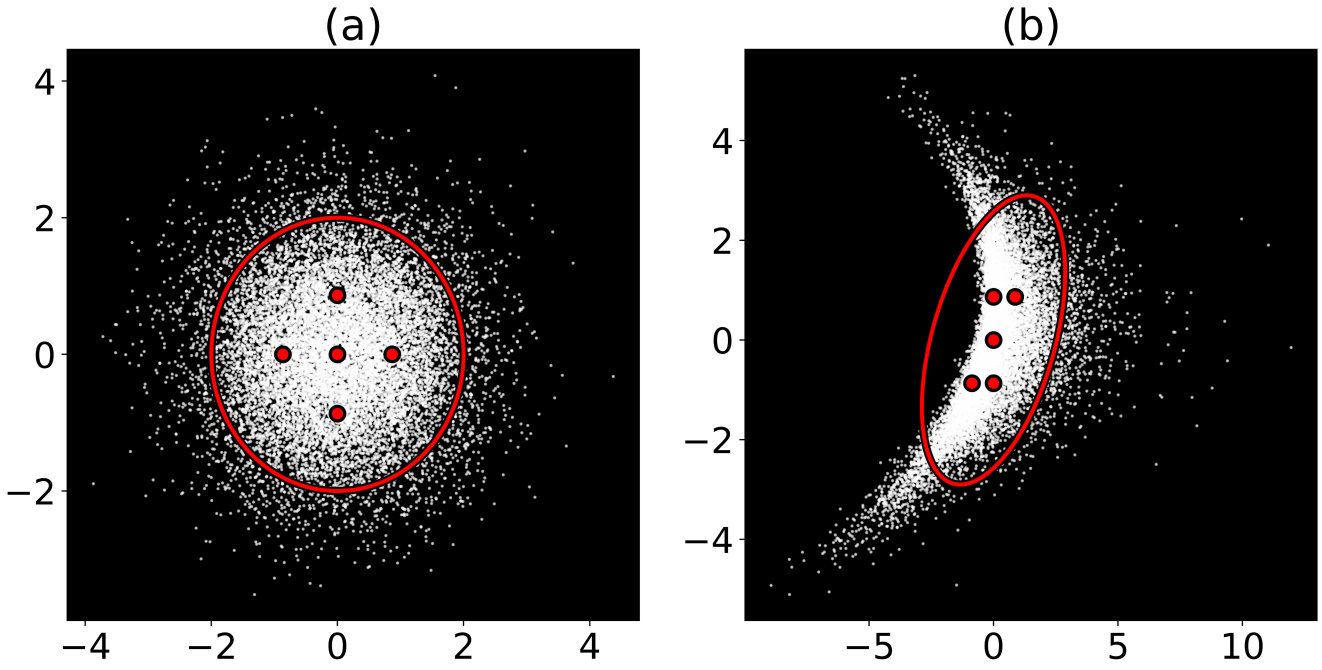


Figure 1. Bedrock geometry in the Kangerlussuaq region of western Greenland in meters above sea level. K and N mark the locations of Kangerlussuaq and Nuuk respectively. D and G mark the locations of the Dye-3 and GISP2 boreholes respectively. Northern and southern paleo-flowlines are shown as blue and red lines running left to right. Historical moraines dating from 11.6 to 7.2 ka BP are shown as colored lines.



$$\boldsymbol{x_0} = \sum_i w_i^{(m)} \boldsymbol{\chi_i} \qquad \boldsymbol{\mathcal{Y}_i} = \mathcal{F}(\boldsymbol{\chi_i}) \qquad \bar{\boldsymbol{y}} \approx \sum_i w_i^{(m)} \boldsymbol{\mathcal{Y}_i}$$

$$P_x = \sum_i w_i^{(c)} (\boldsymbol{\chi_i} - \boldsymbol{x_0})(\boldsymbol{\chi_i} - \boldsymbol{x_0})^T \qquad P_y \approx \sum_i w_i^{(c)} (\boldsymbol{\mathcal{Y}_i} - \bar{\boldsymbol{y}})(\boldsymbol{\mathcal{Y}_i} - \bar{\boldsymbol{y}})^T + R$$

Figure 2. (a) White points are samples from a 2-D Gaussian random variable $\boldsymbol{x} \sim \mathcal{N}(\mathbf{0}, I)$. The red ellipse represents the covariance of the distribution, while red points are sigma points $\boldsymbol{\chi_i}$ used for the UT. (b) White points are samples from the transformed, non-Gaussian distribution for $\boldsymbol{y} = \mathcal{F}(\boldsymbol{x}) + \boldsymbol{\epsilon}$ with $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, R)$. Red points are transformed sigma points $\boldsymbol{\mathcal{Y}_i} = \mathcal{F}(\boldsymbol{\chi_i})$. UT approximations of the mean $\bar{\boldsymbol{y}}$ and covariance P_y of $\boldsymbol{y}$ are estimated via weighted sums of transformed sigma points.

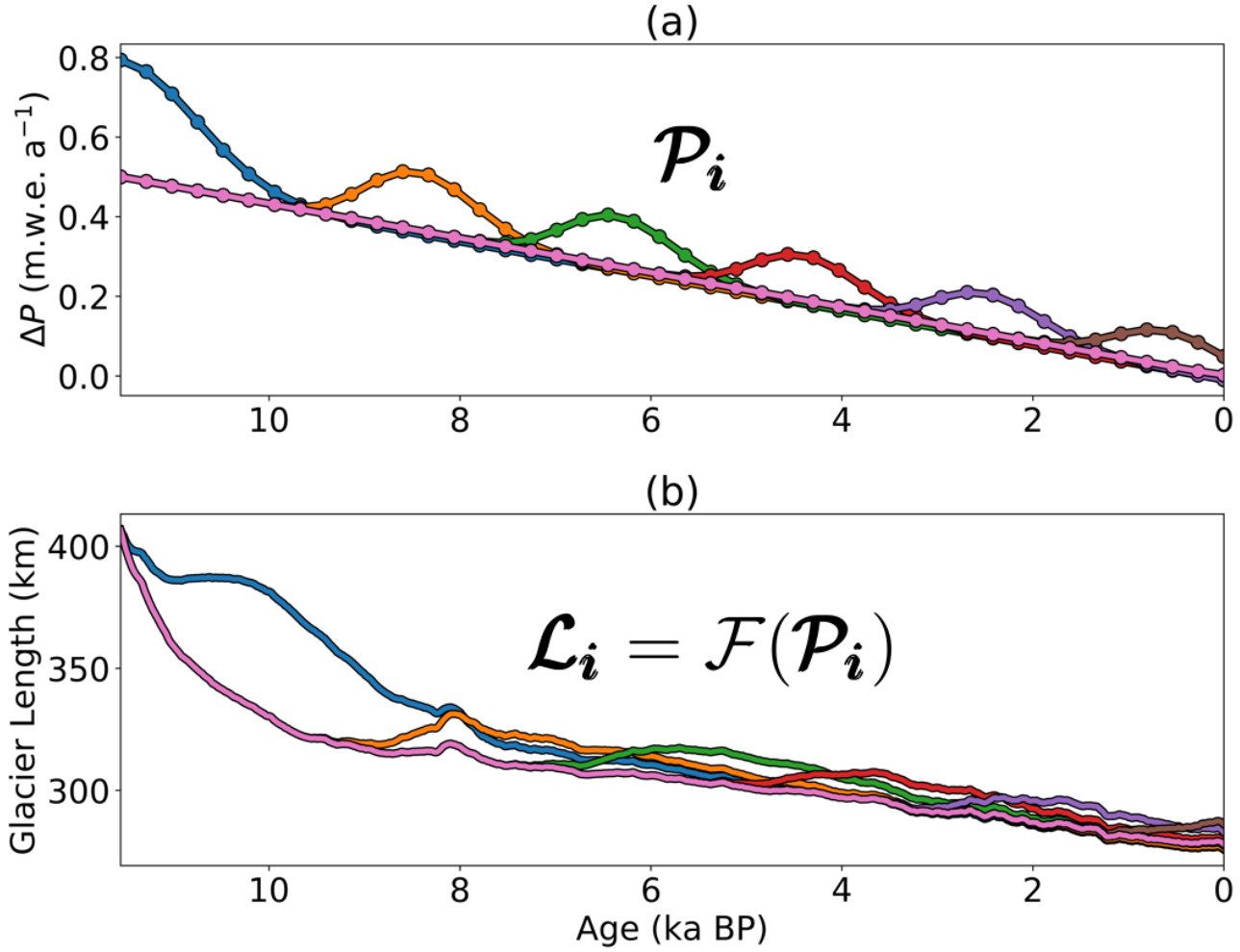


Figure 3. (a) A subset of six, out of a total of 45, Menegaz et al. (2011) sigma points $\boldsymbol{\mathcal{P}_i}$ representing different precipitation anomaly histories. (b) Six corresponding glacier length histories $\boldsymbol{\mathcal{L}_i}$, or transformed sigma points, obtained by inputting the sigma points $\boldsymbol{\mathcal{P}_i}$ through the ice dynamics model $\mathcal{F}$.

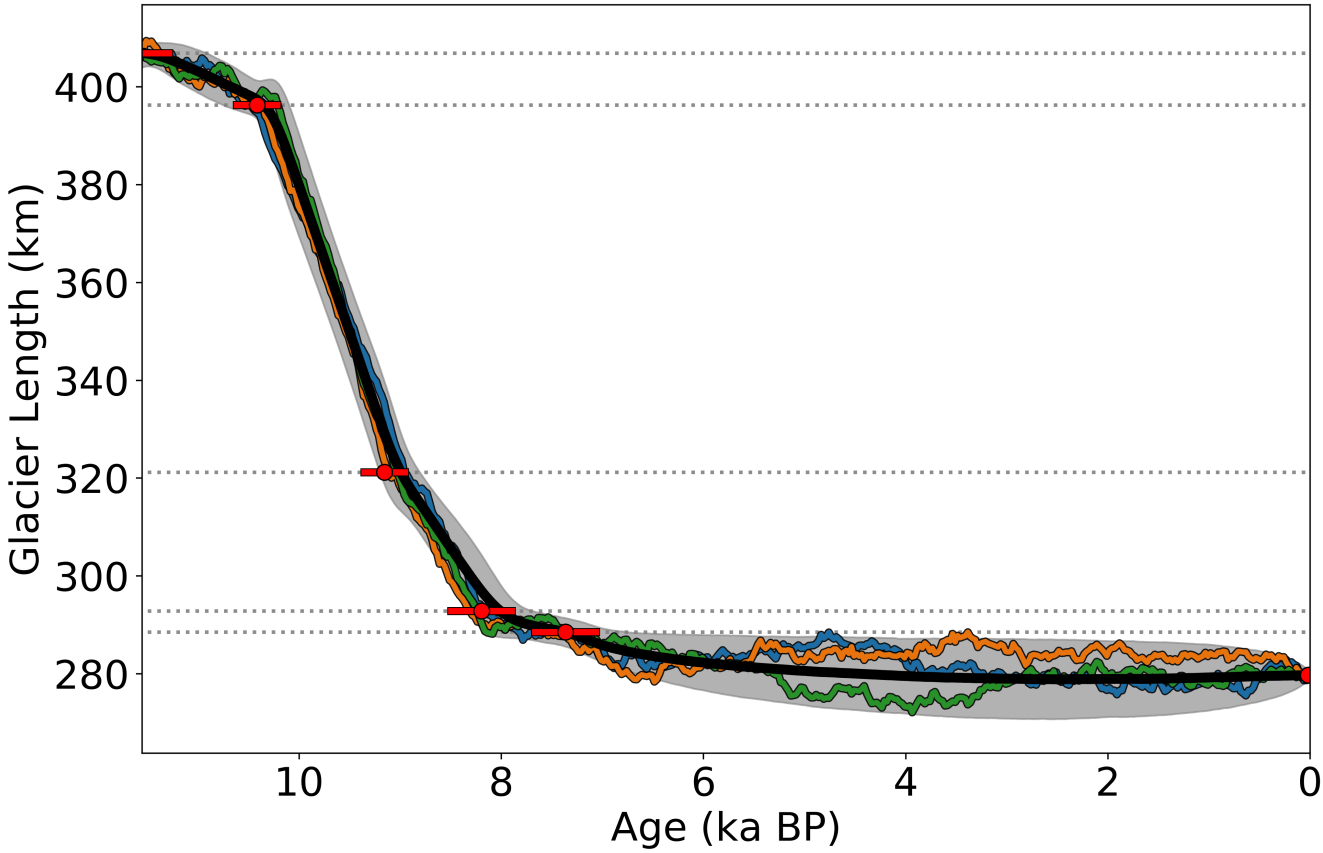


Figure 4. The measurement mean $\boldsymbol{\ell_0}$ (solid black line) and 95% confidence bands (gray shaded region) for the northern flowline are estimated by generating random retreat histories with the same mean moraine formation ages and variances as the observations. The green, blue, and orange lines represent four random plausible retreat histories. Red dots denote the estimated mean moraine formation ages, while red lines show 95% confidence intervals.

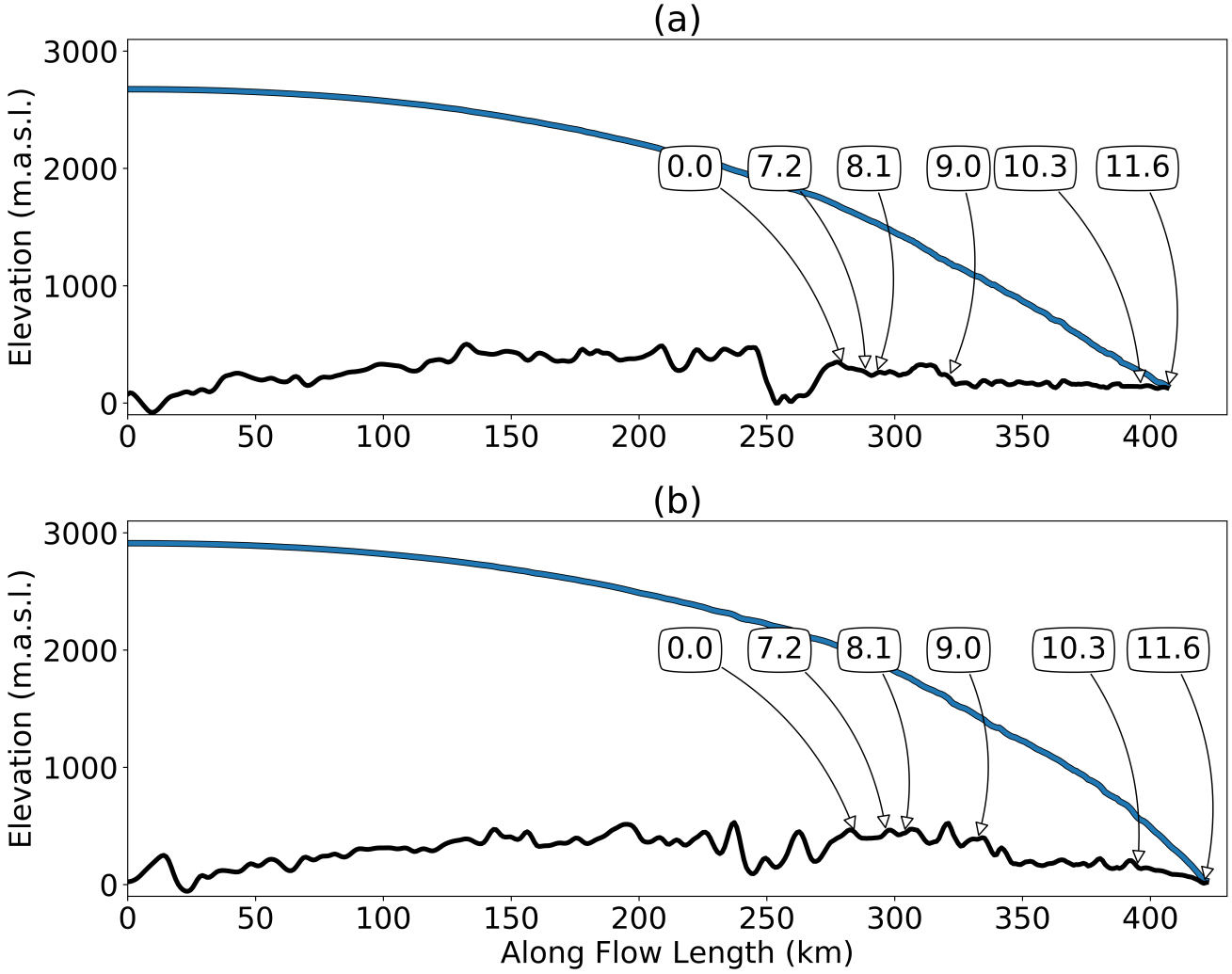


Figure 5. Initial ice sheet configurations for the northern (a) and southern (b) paleo-flowlines. Solid black and blue lines represent the bedrock elevation and ice surface respectively. Moraine positions are indicated by arrows with associated ages expressed in thousands of years before present.

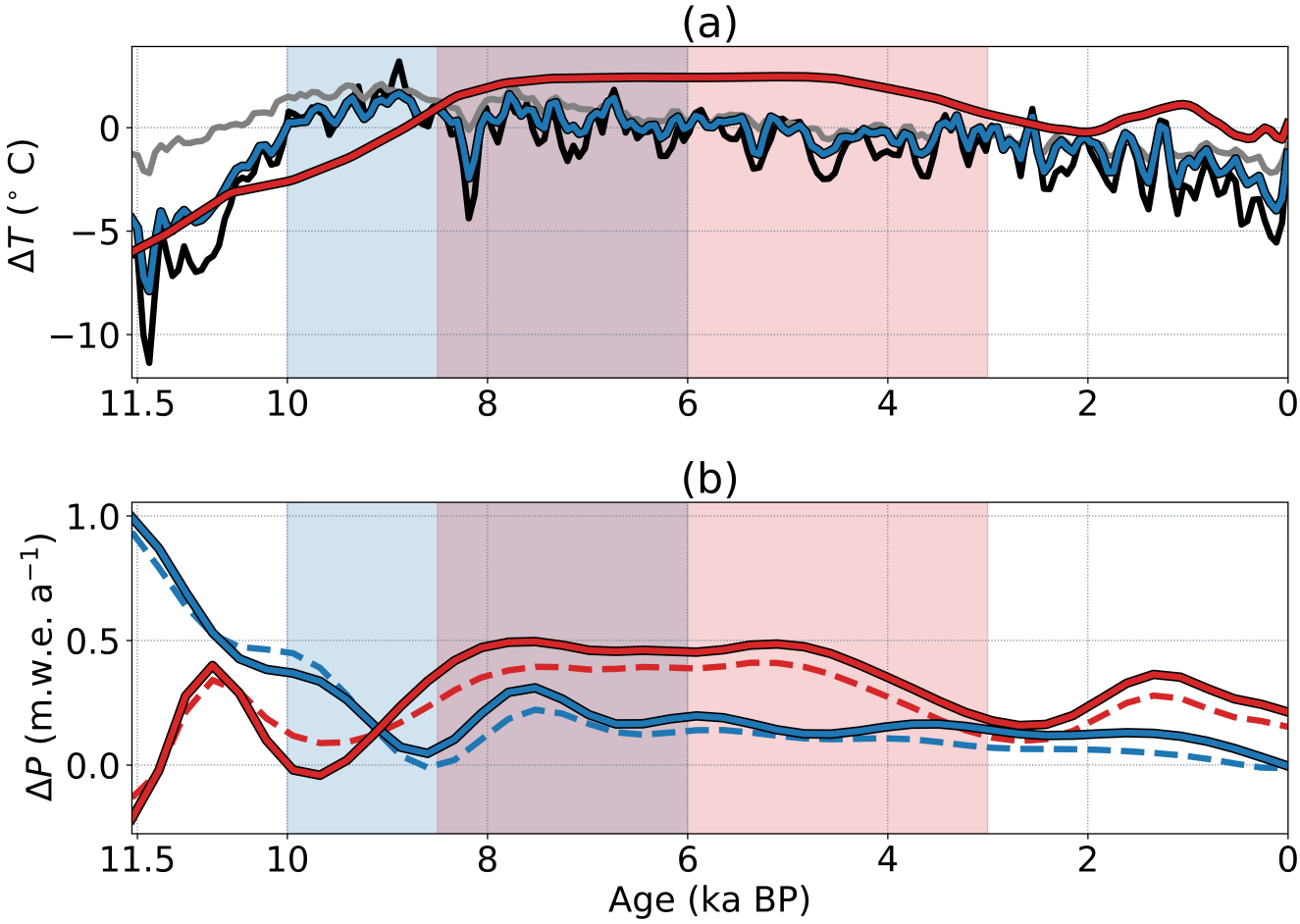


Figure 6. (a) Buizert et al. (2018) and Dahl-Jensen et al. (1998) mean annual temperature anomalies are shown as blue and red lines respectively. Gray and Black lines show average December, January, February and June, July, August temperature anomalies for the Buizert et al. (2018) reconstruction respectively. Shaded blue / pink regions demarcate the extent of the HTM for the Buizert et al. (2018) and Dahl-Jensen et al. (1998) reconstructions. (b) Estimated precipitation anomaly histories for the Buizert et al. (2018) and Dahl-Jensen et al. (1998) inversions are shown by blue and red lines respectively. Solid lines show estimates for the northern flowline, while dashed lines are for the southern flowline.

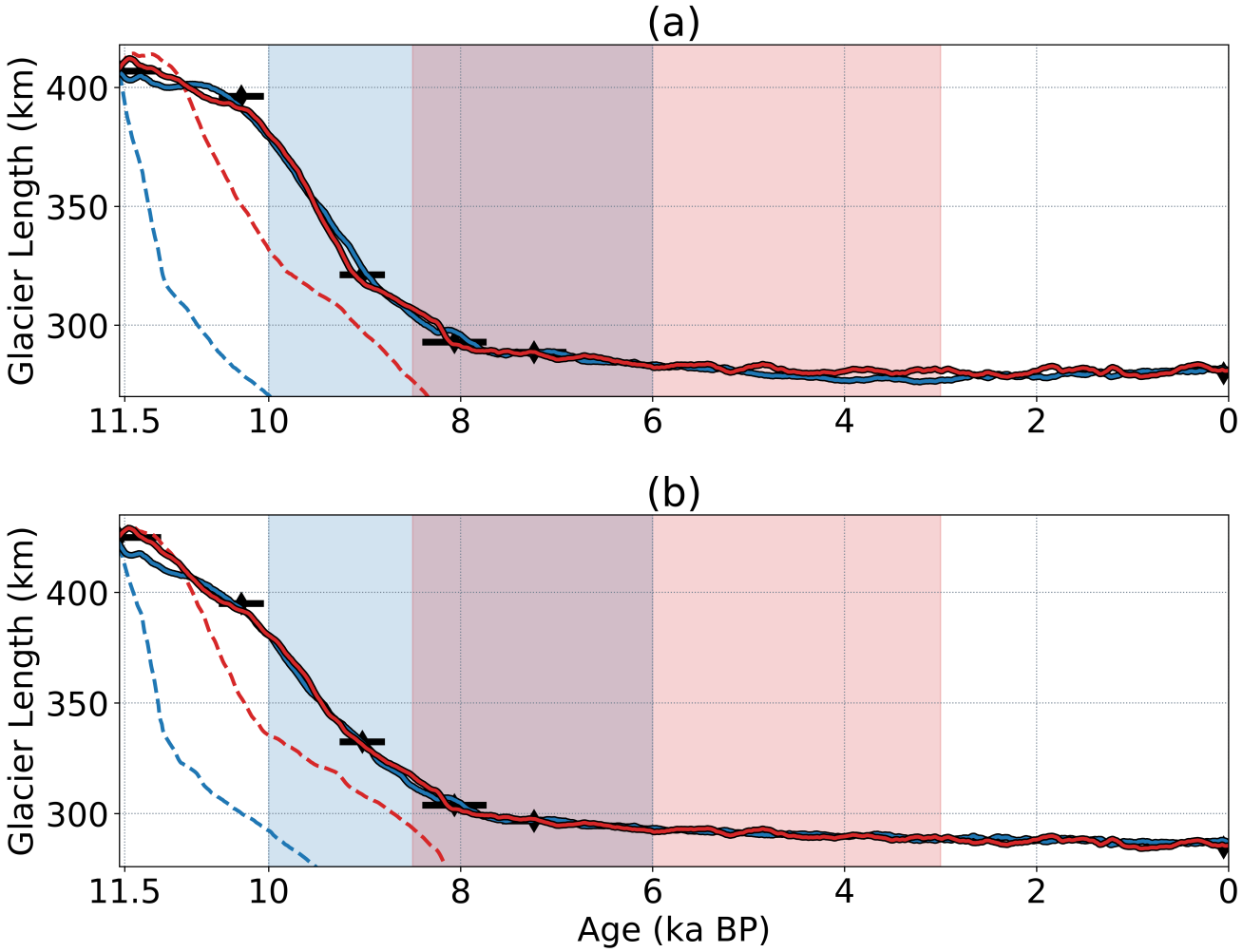


Figure 7. Modeled retreat histories for the northern (a) and southern (b) flowlines using a variety of ΔT and ΔP forcings. Blue and red lines show modeled retreat using Buizert et al. (2018) and Dahl-Jensen et al. (1998) temperature forcings respectively. Dashed lines show the results of forcing the model with a $\Delta P = 0$, while solid lines show the results of using optimal estimated precipitation anomalies. Black diamonds / bars show observed terminus positions along with 95% confidence intervals. Shaded blue / pink regions demarcate the extent of the HTM for the Buizert et al. (2018) and Dahl-Jensen et al. (1998) reconstructions respectively.

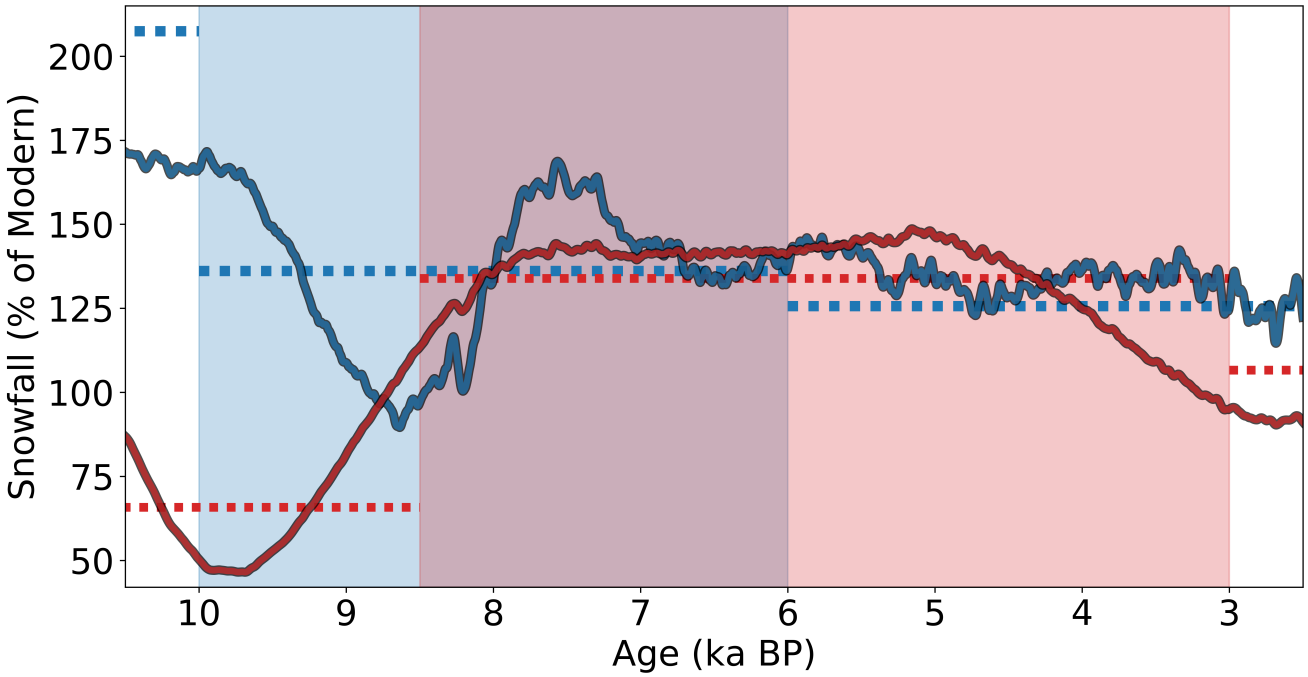


Figure 8. HTM snowfall averaged across the northern and southern flowlines for the Buizert et al. (2018) (solid blue line) and Dahl-Jensen et al. (1998) (solid red line) inversions. Shaded blue / pink regions demarcate the extent of the HTM for the Buizert et al. (2018) and Dahl-Jensen et al. (1998) reconstructions respectively. Dashed lines show pre-HTM, HTM, and post-HTM average snowfall for each inversion.

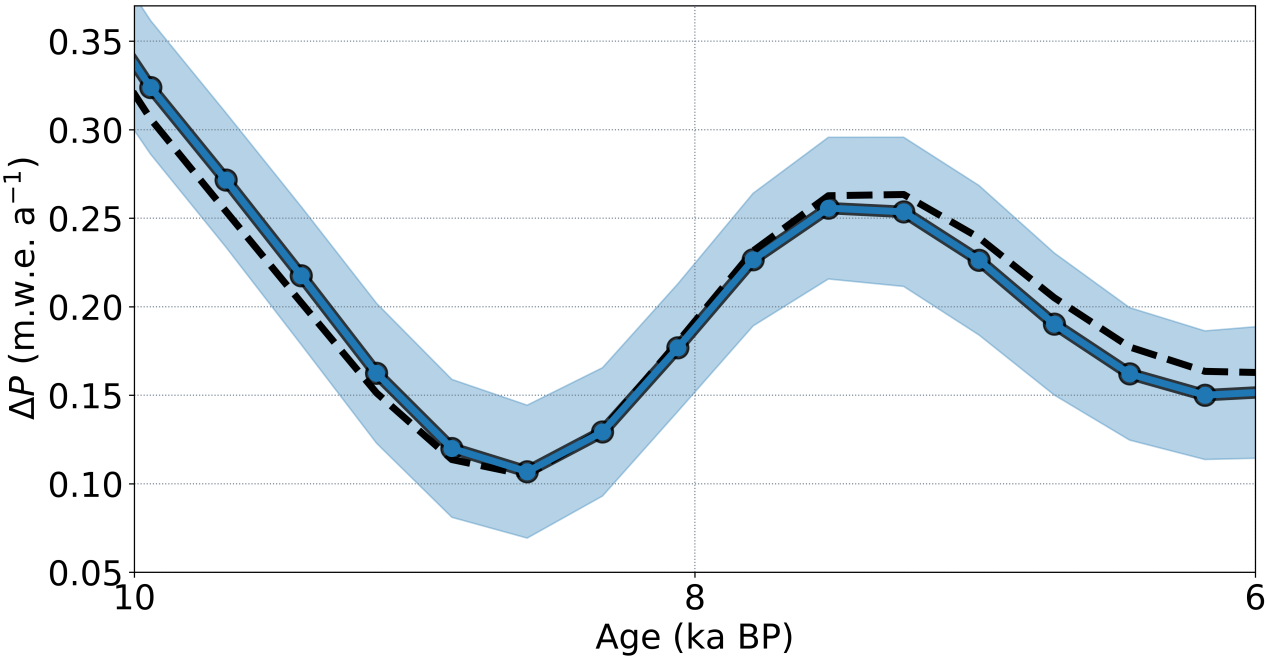


Figure 9. The blue line / shaded region show the estimated mean ΔP and 95% confidence bands for the sensitivity test, accounting for uncertainty in a number of ice-flow and PDD model parameters. The black dashed line shows the mean estimated ΔP from a previous inversion assuming no uncertainty in model parameters.