Supplementary Material to "Joint Bayesian Calibration of Frontal Ablation and Surface Mass Balance in coupled Global Glacier Models"

Ruitang Yang[1], Lizz Ultee[2,3], Kristoffer Aalstad[1], Matvey V. Debolskiy[1], Regine Hock[1], Patrick Schmitt[4], and Tian Li[5]

[1]University of Oslo, Oslo, Norway
[2]NASA Goddard Space Flight Center, Greenbelt, MD, USA
[3]Morgan State University, Baltimore, MD, USA
[4]University of Innsbruck, Innsbruck, Austria
[5]University of Bristol, Bristol, UK

Correspondence: Ruitang Yang (ruitang.yang@geo.uio.no)

Supplementary material

S1 Bayesian approach

Our calibration can be cast as an inverse problem of the form (Sanz-Alonso et al., 2023)

$$\mathbf{y} = \mathcal{G}(\boldsymbol{\theta}) + \boldsymbol{\epsilon}\,, \tag{S1}$$

where $\mathcal{G}(\cdot)$ denotes the forward (data generating) model, in this case the coupled PyGEM-OGGM-SERMeQ glacier model, and the noise term $\boldsymbol{\epsilon}$ represents observation error. For simplicity we have here made the typical strong constraint assumption (see Evensen et al., 2022) that the forward model is perfect other than errors in the parameters (including forcing) $\boldsymbol{\theta}$, which can be justified at least to a first order in both atmospheric (Hersbach et al., 2020) and cryospheric applications (Alonso-González et al., 2022; Willmes et al., 2025; Cao et al., 2025). Conceptually, calibration is then tantamount to solving the inverse problem in (S1): identify the parameters $\boldsymbol{\theta}$ that may have generated the observations $\mathbf{y}$. Due to the nonlinearity of $\mathcal{G}$ and noise $\boldsymbol{\epsilon}$ this is a challenging and ill-posed inverse problem with a non-unique solution (Sanz-Alonso et al., 2023). As such, the notion of a universally objectively optimal solution for the parameters $\boldsymbol{\theta}$ is ill-conceived and we instead adopt a Bayesian approach where we seek to infer plausible parameter sets given the noisy observations.

Adopting the Bayesian approach to calibration via inverse modeling, we infer these parameters by conditioning on data stored in the observation vector $\mathbf{y}$ through Bayes' rule (Evensen et al., 2022; Sanz-Alonso et al., 2023; Murphy, 2023)

$$p(\boldsymbol{\theta} \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid \boldsymbol{\theta}) p(\boldsymbol{\theta})}{p(\mathbf{y})}\,, \tag{S2}$$

where the target posterior probability density $p(\boldsymbol{\theta} \mid \mathbf{y})$ of interest is a measure of epistemic uncertainty concerning possible parameter vector settings $\boldsymbol{\theta}$ given the data. This posterior is obtained by updating the prior probability density $p(\boldsymbol{\theta})$, which encodes the initial parameter uncertainty before considering the data, with the likelihood $p(\mathbf{y} \mid \boldsymbol{\theta})$, a function of the parameters $\boldsymbol{\theta}$ that measures the goodness of fit of the modeled observables $\widehat{\mathbf{y}} = \mathcal{G}(\boldsymbol{\theta})$ with parameters $\boldsymbol{\theta}$ to the noisy observations $\mathbf{y}$, all divided by the evidence $p(\mathbf{y})$ which here plays the role of a normalizing constant. In theory, inferring the posterior via Equation (S2) is (up to a normalizing constant) a simple distributed multiplication problem that can be approximated up to a desired accuracy by discretizing the parameter vector $\boldsymbol{\theta}$ over a sufficiently fine grid. This grid approximation is rarely feasible due to the large number of likelihood evaluations required, where each such evaluation runs the costly forward model to obtain $\widehat{\mathbf{y}} = \mathcal{G}(\boldsymbol{\theta})$ for each parameter set $\boldsymbol{\theta}$, due to the curse of dimensionality whereby the number of parameter settings (grid points) grows exponentially with the number of parameters N_p (Murphy, 2023).

The challenge of inferring the posterior in (S2), especially when $\mathcal{G}$ is nonlinear, has spurred the development of tractable computational methods for approximate Bayesian inference that can be split into two main approaches: variational and Monte Carlo methods (see Murphy, 2023, for an overview). In geophysics, this challenging inference task falls under the umbrella of inverse modeling and, more specifically, the field of data assimilation (DA) when applied to dynamical systems (Evensen et al., 2022; Sanz-Alonso et al., 2023). Geophysical DA methods can be directly mapped onto the aforementioned computational methods: ensemble-based DA methods are Monte Carlo approximations and optimization-based variational DA methods are

plugin variational approximations (cf. Evensen et al., 2022; Murphy, 2023). On the one hand, variational methods have a long and promising track-record in glaciology (MacAyeal, 1993; Brinkerhoff, 2022; Schmitt et al., 2025), but they require gradients of the form $\partial_{\boldsymbol{\theta}}\mathcal{G}$ which are often not available in complex geophysical models, including our own. On the other hand, ensemble-based methods have also recently shown promise in glaciology (Navari et al., 2021; Landmann et al., 2021; Herrmann et al., 2025; Cao et al., 2025), with the advantage that they are often more uncertainty-aware (Evensen et al., 2022), but also gradient-free and thus more generally applicable to models like ours that are not (yet) differentiable. As such, herein we employ an ensemble-based data assimilation method to perform Bayesian calibration of our glacier model.

S2 Model Performance Evaluation

We assess model performance during the calibration (2000–2009) and validation (2010–2019) periods at two complementary scales: individual-glacier (glacier scale) and whole-region (regional scale). At the glacier scale we quantify probabilistic accuracy and ensemble reliability, and at the regional scale we evaluate distributional agreement.

S2.1 Glacier-scale metrics

S2.1.1 Normalized Continuous Ranked Probability Score

For each glacier the observed value is denoted by x_{obs} with associated observational standard deviation σ_{obs}, so that observational uncertainty is represented by $y \sim \mathcal{N}\big(x_{\mathrm{obs}}, \sigma_{\mathrm{obs}}\big)$, and the model output is represented by an ensemble $\{x_i\}_{i=1}^{N}$ (here $N = 200$).

To quantify probabilistic model performance at the glacier scale we use the Continuous Ranked Probability Score (CRPS). We employ the bias-corrected form

$$\mathrm{CRPS} = \mathbb{E}_{x,y}\big[\,|x-y|\,\big] - \tfrac{1}{2}\,\mathbb{E}_{x,x'}\big[\,|x-x'|\,\big], \tag{S3}$$

where x and x' denote independent draws from the empirical ensemble distribution (i.e., from $\{x_i\}$ with equal probability) and y is drawn from $\mathcal{N}(x_{\mathrm{obs}}, \sigma_{\mathrm{obs}})$. The first term, $\mathbb{E}_{x,y}[|x-y|]$, measures the expected model–observation mismatch; the second term, $\frac{1}{2}\mathbb{E}_{x,x'}[|x-x'|]$, is one half of the ensemble self-spread and is subtracted to isolate model–observation error.

To enable comparison across glaciers with differing observational uncertainty we normalize CRPS by the observational standard deviation to obtain the normalized CRPS (nCRPS):

$$\mathrm{nCRPS} \;=\; \frac{\mathrm{CRPS}}{\sigma_{\mathrm{obs}}}. \tag{S4}$$

Lower nCRPS indicates better probabilistic agreement after accounting for observational uncertainty.

In practice we estimate CRPS and nCRPS with discrete (Monte Carlo) approximations. In our implementation we use $M = 2000$ independent samples $y_j \sim \mathcal{N}(x_{\mathrm{obs}}, \sigma_{\mathrm{obs}})$ and the ensemble members $\{x_i\}_{i=1}^{N}$. The expectations in (S3) are estimated by

$$\widehat{\mathbb{E}}_{x,y}\big[\,|x-y|\,\big] = \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}\big|x_i - y_j\big|. \tag{S5}$$

$$\widehat{\mathbb{E}}_{x,x'}\big[\,|x-x'|\,\big] = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{i'=1}^{N}\big|x_i - x_{i'}\big|. \tag{S6}$$

where the double sum in (S6) corresponds to independent draws x and x' from the empirical ensemble (with replacement). Substituting (S5) and (S6) into (S3) yields the estimated CRPS, and (S4) gives the estimated nCRPS. The glacier-wise median nCRPS reported in the tables (Tables S2 and S3) is the median of the nCRPS values across all glaciers; lower values indicate better probabilistic performance after accounting for observational uncertainty.

S2.1.2 Fractional overlap

To provide a simple glacier-scale diagnostic of consistency between observational uncertainty and ensemble spread, we compute the fractional overlap between the observational $\pm 1\sigma$ interval and the ensemble range. For a given glacier let x_{obs} denote the observed value with standard deviation σ_{obs}, and let $\{x_i\}_{i=1}^{N}$ be the model ensemble. The observational interval is $I_{\mathrm{obs}} = \big[x_{\mathrm{obs}} - \sigma_{\mathrm{obs}},\, x_{\mathrm{obs}} + \sigma_{\mathrm{obs}}\big]$, and the ensemble range is taken as the central 95% credible interval, $I_{\mathrm{ens}} = \big[P_{2.5}(x),\, P_{97.5}(x)\big]$, where $P_p(x)$ denotes the pth percentile of the ensemble $\{x_i\}$.

The fractional overlap is the length of the intersection of these two intervals normalized by the length of the observational interval.

$$\mathrm{overlap}_{\mathrm{frac}} = \frac{|I_{\mathrm{obs}} \cap I_{\mathrm{ens}}|}{|I_{\mathrm{obs}}|} = \frac{|I_{\mathrm{obs}} \cap I_{\mathrm{ens}}|}{2\,\sigma_{\mathrm{obs}}}. \tag{S7}$$

By construction $\mathrm{overlap}_{\mathrm{frac}} = 1$ when the observational interval is fully contained in the ensemble range and $\mathrm{overlap}_{\mathrm{frac}} = 0$ when the intervals do not intersect; intermediate values quantify the fraction of the observational $\pm 1\sigma$ interval covered by the ensemble range.

For reporting, per-glacier overlaps are expressed as percentages, $\mathrm{overlap}_{\%} = 100 \times \mathrm{overlap}_{\mathrm{frac}}$. The summary statistic used in Tables S2 and S3, "Median fractional overlap" is $\mathrm{median}\big(\mathrm{overlap}_{\%,1}, \ldots, \mathrm{overlap}_{\%,N_{\mathrm{g}}}\big)$, where $\mathrm{overlap}_{\%,k}$ is the overlap percentage for glacier k and N_{g} is the total number of glaciers.

S2.2 Regional-scale metric

Regional agreement is assessed by comparing empirical cumulative distribution functions (CDFs) of observations and model output aggregated across all glaciers.

S2.2.1 Observed CDF with observational uncertainty

For each glacier $k = 1, \ldots, N_{\mathrm{g}}$ let $x_{\mathrm{obs},k}$ denote the observed value and $\sigma_{\mathrm{obs},k}$ its standard deviation. To propagate observational uncertainty into the regional distribution, we draw N_{MC} Monte Carlo realizations (here $N_{\mathrm{MC}} = 2000$):

$$x_k^{(m)} \sim \mathcal{N}\big(x_{\mathrm{obs},k},\, \sigma_{\mathrm{obs},k}\big), \qquad m = 1, \ldots, N_{\mathrm{MC}}.$$

For each Monte Carlo replicate m we form the empirical CDF across glaciers,

$$F_{\mathrm{obs},m}(x) = \frac{1}{N_{\mathrm{g}}} \sum_{k=1}^{N_{\mathrm{g}}} \mathbb{I}\big(x_k^{(m)} \le x\big), \tag{S8}$$

where $\mathbb{I}(\cdot)$ is the indicator function. At each x the ensemble $\{F_{\mathrm{obs},m}(x)\}_{m=1}^{N_{\mathrm{MC}}}$ is summarized by its median, denoted $F_{\mathrm{obs}}(x)$, and by the 2.5th and 97.5th percentiles, which define an uncertainty envelope for the observed regional CDF.

S2.2.2 Modeled CDF

Let $x_{k,j}$ denote the modeled value for glacier $k = 1, \ldots, N_{\mathrm{g}}$ and ensemble member $j = 1, \ldots, N_{\mathrm{ens}}$ (here $N_{\mathrm{ens}} = 200$). For each ensemble member j we form the empirical CDF across glaciers,

$$F_{\mathrm{mod},j}(x) = \frac{1}{N_{\mathrm{g}}} \sum_{k=1}^{N_{\mathrm{g}}} \mathbb{I}\big(x_{k,j} \le x\big). \tag{S9}$$

At each x the collection $\{F_{\mathrm{mod},j}(x)\}_{j=1}^{N_{\mathrm{ens}}}$ is summarized by its median, denoted $F_{\mathrm{mod}}(x)$, with the 2.5th and 97.5th percentiles providing the modeled uncertainty envelope.

S2.2.3 CDF overlap metric

Regional agreement is quantified by an overlap metric between the median observed and modeled CDFs on a common evaluation domain $[x_{\min}, x_{\max}]$. The continuous form is

$$\text{CDF overlap} = 1 - \frac{1}{x_{\max} - x_{\min}} \int_{x_{\min}}^{x_{\max}} \big|F_{\mathrm{obs}}(x) - F_{\mathrm{mod}}(x)\big| \, \mathrm{d}x, \tag{S10}$$

which yields a value in $[0, 1]$ (1 for identical median CDFs, 0 for maximal discrepancy). In practice the integral is evaluated numerically on a discrete grid $\{x_i\}$ covering $[x_{\min}, x_{\max}]$:

$$\text{CDF overlap} = 1 - \frac{1}{x_{\max} - x_{\min}} \sum_i \big|F_{\mathrm{obs}}(x_i) - F_{\mathrm{mod}}(x_i)\big| \, \Delta x_i, \qquad \Delta x_i = x_{i+1} - x_i. \tag{S11}$$

The CDF overlap is reported as a percentage by multiplying the result by 100 in Tables S2 and S3. The higher CDF overlap (%) indicates closer regional agreement between observed and modeled distributions.

S3 Delta Normalized RMSE

In this study, we assess the predictive accuracy of our model outputs for each glacier, by calculating the delta normalized Root Mean Square Error (RMSE) with respect to observational data, both before (prior) and after (posterior) the integration of observational information.

S3.1 Normalized RMSE

The normalized RMSE (NRMSE) is calculated for both the posterior and prior model predictions using the following formula:

$$\text{NRMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\text{Obs}_i - \text{Model}_i}{\sigma_i}\right)^2} \tag{S12}$$

where Obs_i represents the observational values, Model_i denotes the model output (either posterior or prior), σ_i is the uncertainty associated with each observation, and n is the number of years and i denotes the specific year. This normalization adjusts the RMSE based on the observational uncertainties, providing a more meaningful comparison of model performance.

S3.2 Delta Normalized RMSE

After obtaining the normalized RMSE for both the posterior and prior model results, we compute the delta normalized RMSE (ΔNRMSE):

$$\Delta\text{NRMSE} = \text{NRMSE}_{\text{posterior}} - \text{NRMSE}_{\text{prior}} \tag{S13}$$

The resulting ΔNRMSE is a unitless metric that quantifies the change in model performance following data assimilation. If $\Delta\text{NRMSE} < 0$, the posterior model demonstrates improved accuracy compared to the prior model, indicating effective assimilation of new data while accounting for uncertainty. Conversely, if $\Delta\text{NRMSE} = 0$, it suggests no improvement in model performance. If $\Delta\text{NRMSE} > 0$, the posterior model is less accurate than the prior model, indicating a deterioration in performance due to data assimilation. The result in this study is shown in Figure S2.

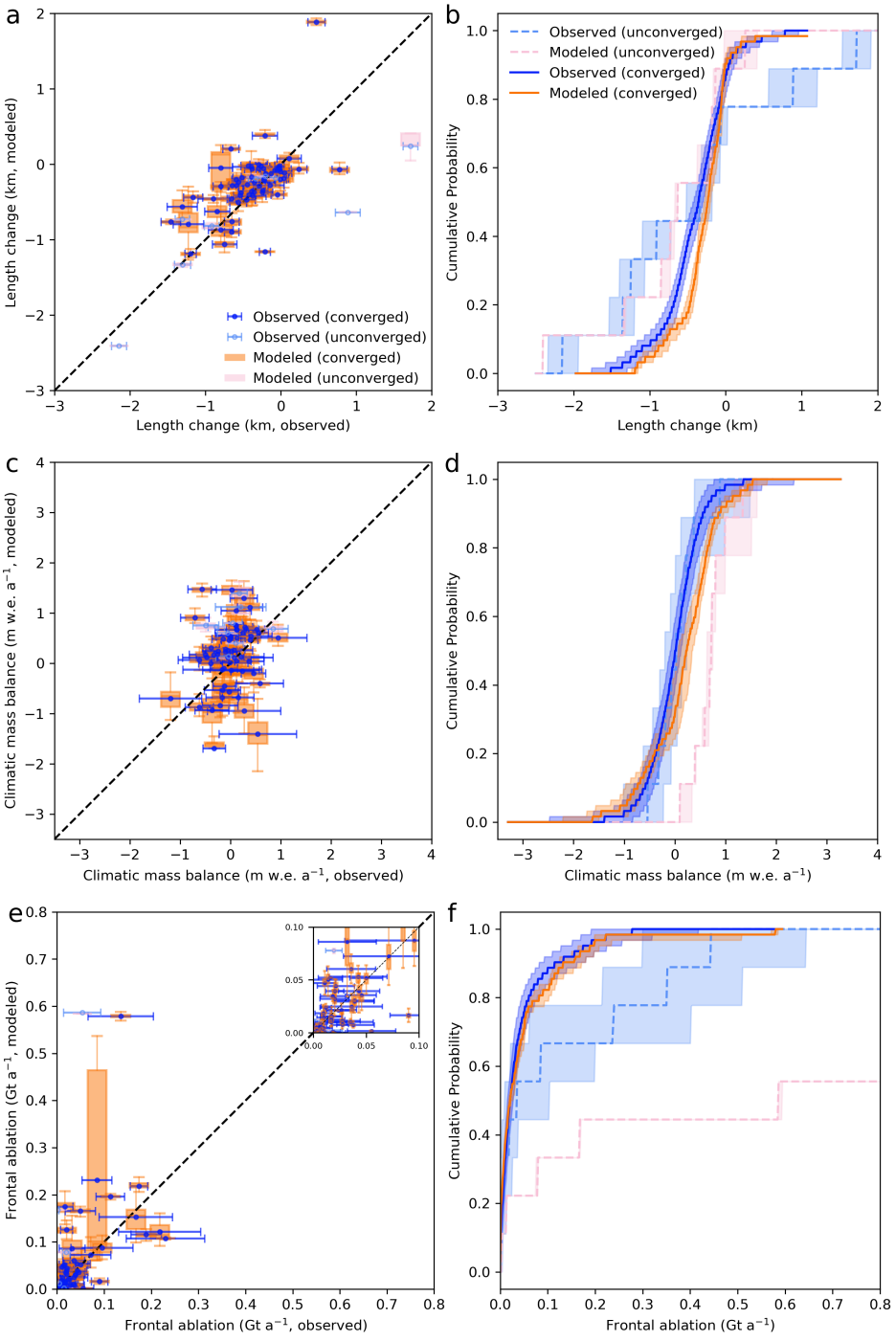

Figure S1. Modeled results against observations for 71 glaciers in Svalbard during the calibration period 2000 - 2009. (**a, b**) Total length change along the glacier centerline; (**c, d**) Mean climatic mass balance; (**e, f**) Mean frontal ablation. Panels (**a, c, e**) show observations (blue error bar) and modeled ensemble means (orange box represents the interquartile range (IQR) of the ensemble of each glacier). Panels (**b, d, f**) illustrates the empirical cumulative distribution functions (CDFs) for observations (blue line ± 95% confidence interval via Monte Carlo sampling) and model results (orange line ± ensemble range).

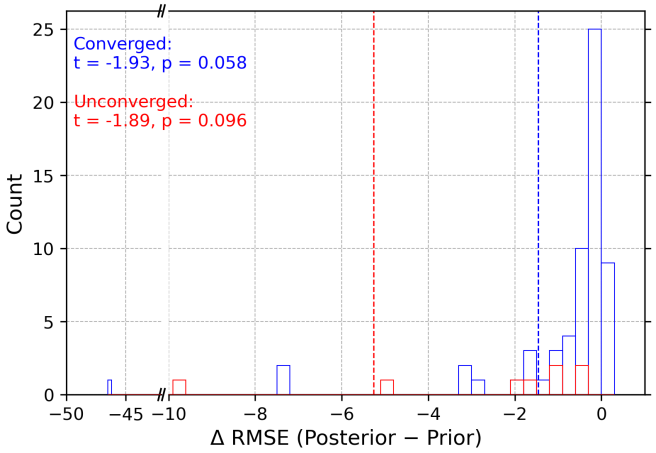

Figure S2. Histogram of Δ Normalized RMSE (ΔNRMSE) for 71 glaciers in Svalbard. The blue bars represent converged cases, while the red bars indicate unconverged cases. Dashed vertical lines mark the mean ΔNRMSE for each group. Positive ΔNRMSE values reflect a deterioration in model performance, whereas negative values signify an improvement. The annotated t-statistics and p-values are derived from a one-sample Student's t-test that evaluates whether the mean ΔNRMSE for each group significantly differs from zero, indicating whether the model has significantly improved. A diagonal slash is included to denote a break in the x-axis.

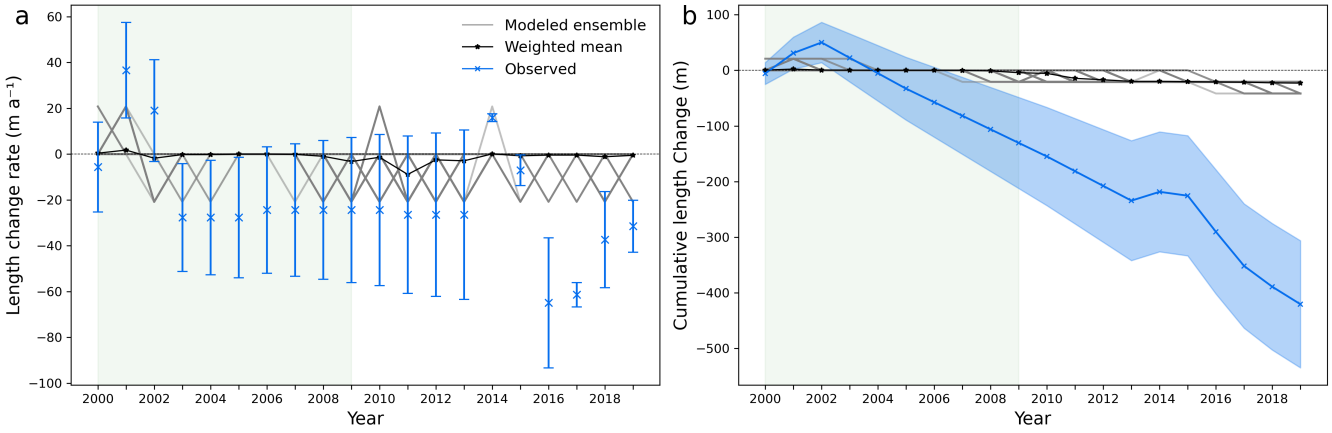

Figure S3. Comparison of modeled results with observations from 2000 to 2020 for (a) annual length change rate, (b) cumulative length change of Sabinebreen, which illustrate results from OGGM's flowline model. The observed values in blue include the associated uncertainties. The light green shading refers to the calibration period.

Table S1. List of studied glaciers and associated model convergence information. RGIIDs are listed without the common "RGI60-0" prefix. In the column "Iterations to converge" (Iter.), "Converged" denotes convergence within 10 iterations; numerical values indicate convergence achieved after more than 10 but fewer than 100 iterations; "–" indicates no convergence within 100 iterations. Superscript † denotes cases with physical inconsistencies.

RGIID	Area (km^2)	Iter. (≤ 100)	RGIID	Area (km^2)	Iter. (≤ 100)	RGIID	Area (km^2)	Iter. (≤ 100)
7.00025	1095.9	55	7.00471	21.3	Converged	7.00892	222.6	Converged
7.00026	153.5	Converged	7.00474	165.2	Converged	7.00896	34.0	16
7.00027	1226.4	17	7.00502	126.6	Converged	7.00903	69.9	Converged
7.00029	662.9	Converged	7.00551	356.5	Converged	7.00910	41.6	Converged
7.00031	218.6	Converged	7.00552	328.3	Converged	7.00945	93.0	21
7.00036	67.4	Converged	7.00558	88.7	Converged	7.01328†	14.1	Converged
7.00042	109.8	Converged	7.00645	50.7	Converged	7.01353	9.5	Converged
7.00060	79.9	Converged	7.00659	28.1	Converged	7.01449	331.5	Converged
7.00061	84.7	Converged	7.00681	641.0	Converged	7.01458	163.4	21
7.00064	69.1	Converged	7.00682	129.1	Converged	7.01464	295.5	Converged
7.00224†	15.8	Converged	7.00714	10.4	Converged	7.01487	326.5	Converged
7.00250	52.8	Converged	7.00715	36.6	Converged	7.01488	200.8	Converged
7.00251	16.6	Converged	7.00724	3.3	Converged	7.01492	76.3	Converged
7.00276	24.4	Converged	7.00727	190.8	Converged	7.01493	38.1	Converged
7.00277	29.4	–	7.00733	41.8	Converged	7.01494	394.3	Converged
7.00301	312.5	Converged	7.00789	3.6	Converged	7.01499	179.5	Converged
7.00315	29.8	Converged	7.00790	16.7	Converged	7.01500	318.8	Converged
7.00425	472.6	Converged	7.00791†	43.2	Converged	7.01503	55.4	Converged
7.00448	42.1	Converged	7.00794	55.6	Converged	7.01506†	963.9	–
7.00463	38.1	Converged	7.00832	98.5	Converged	7.01507	77.2	Converged
7.00464	110.3	Converged	7.00854	288.5	Converged	7.01559	876.6	60
7.00465	119.1	Converged	7.00859	117.6	14	7.01560	594.9	Converged
7.00468	29.3	Converged	7.00874	108.7	Converged	7.01561	76.9	Converged
7.00470	46.6	Converged	7.00883	35.9	Converged			

Table S2. Model performance for 62 converged glaciers during calibration (2000–2010) and validation (2010–2020) periods. Glacier-scale metrics: "Median nCRPS" denotes the median normalized Continuous Ranked Probability Score (lower is better, dimensionless); "Mean fractional overlap" denotes the mean fractional overlap between observational uncertainty ($\pm 1\sigma$) and the model ensemble range. Regional-scale metric: "CDF overlap" denotes overlap between observed and modeled empirical cumulative distribution functions (percent). ΔL denotes total length change over the period; B_{clim} denotes period-averaged climatic mass balance; A_{f} denotes period-averaged frontal ablation.

Period	Median nCRPS			Mean fractional overlap (%)			CDF overlap (%)		
	ΔL	B_{clim}	A_{f}	ΔL	B_{clim}	A_{f}	ΔL	B_{clim}	A_{f}
2000–2010	1.5	1.6	1.1	39	34	31	96	96	98
2010–2020	3.2	1.9	4.4	40	39	24	97	95	96

Table S3. Model performance summary for the nine unconverged glaciers during calibration (2000–2010) and validation (2010–2020) periods. See Table S2 for metric definitions and notation.

Period	Median nCRPS			Mean fractional overlap (%)			CDF overlap (%)		
	ΔL	B_{clim}	A_{f}	ΔL	B_{clim}	A_{f}	ΔL	B_{clim}	A_{f}
2000–2010	2.4	3.2	11.1	18	10	0.1	91	82	71
2010–2020	5.3	3.7	3.6	11	0	9	96	86	91

Table S4. Overview of 12 GCMs used in the study for the period 2000–2100 over scenarios SSP1-2.6 and SSP5-8.5 (24 projections total). All models use the CMIP6 variant label r1i1p1f1.

GCMs	Variant Label
BCC-CSM2-MR	r1i1p1f1
CESM2	r1i1p1f1
CESM2-WACCM	r1i1p1f1
EC-Earth3	r1i1p1f1
EC-Earth3-Veg	r1i1p1f1
FGOALS-f3-L	r1i1p1f1
GFDL-ESM4	r1i1p1f1
INM-CM4-8	r1i1p1f1
INM-CM5-0	r1i1p1f1
MPI-ESM1-2-HR	r1i1p1f1
MRI-ESM2-0	r1i1p1f1
NorESM2-MM	r1i1p1f1

References

Alonso-González, E., Aalstad, K., Baba, M. W., Revuelto, J., López-Moreno, J. I., Fiddes, J., Essery, R., and Gascoin, S.: The Multiple Snow Data Assimilation System (MuSA v1.0), Geosci. Model Dev., https://doi.org/10.5194/gmd-15-9127-2022, 2022.

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

Cao, W., Aalstad, K., Schmidt, L., Westermann, S., and Schuler, T.: Bayesian data assimilation on an Arctic glacier: Learning from large ensemble twin experiments, Journal of Glaciology, pp. 1–44, https://doi.org/10.1017/jog.2025.10101, 2025.

Evensen, G., Vossepoel, F. C., and van Leeuwen, P. J.: Data Assimilation Fundamentals, Springer, https://doi.org/10.1007/978-3-030-96709-3, 2022.

Herrmann, O., Groos, A. R., Tabone, I., Jouvet, G., and Fürst, J. J.: A Kalman filter-based framework for assimilating remote sensing observations into a surface mass balance model, Annals of Glaciology, 66, e23, https://doi.org/10.1017/aog.2025.10020, 2025.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/https://doi.org/10.1002/qj.3803, 2020.

Landmann, J. M., Künsch, H. R., Huss, M., Ogier, C., Kalisch, M., and Farinotti, D.: Assimilating near-real-time mass balance stake readings into a model ensemble using a particle filter, The Cryosphere, 15, 5017–5040, https://doi.org/10.5194/tc-15-5017-2021, 2021.

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.

Murphy, K.: Probabilistic Machine Learning: Advanced Topics, MIT Press, http://probml.github.io/book2, 2023.

Navari, M., Margulis, S. A., Tedesco, M., Fettweis, X., and van de Wal, R. S. W.: Reanalysis Surface Mass Balance of the Greenland Ice Sheet Along K-Transect (2000–2014), Geophysical Research Letters, 48, e2021GL094 602, https://doi.org/https://doi.org/10.1029/2021GL094602, 2021.

Sanz-Alonso, D., Stuart, A. M., and Taeb, A.: Inverse problems and data assimilation, vol. 107, Cambridge, https://doi.org/10.1017/9781009414319, 2023.

Schmitt, P., Maussion, F., Goldberg, D. N., and Gregor, P.: The Open Global Glacier Data Assimilation Framework (AGILE) v0.1, EGUsphere, 2025, 1–35, https://doi.org/10.5194/egusphere-2025-3401, 2025.

Willmes, C., Aalstad, K., and Westermann, S.: Assimilating high-resolution satellite snow cover data in a permafrost model, EGUsphere, https://doi.org/10.5194/egusphere-2025-3142, 2025.