

Regional Glacier Melt Modeling: Insights from Surface Energy Balance and Positive-Degree-Day Comparisons

Hannah PHELPS,¹ Valentina RADIC,¹ Scott N. WILLIAMSON,^{2,3}

¹*Department of Earth, Ocean and Atmospheric Sciences, University of British Columbia, Vancouver, Canada*

²*Polar Knowledge Canada, Canadian High Arctic Research Station, Cambridge Bay, Nunavut, Canada*

³*School of Environmental Science, Simon Fraser University, Burnaby, Canada*

Correspondence: Hannah Phelps <hphelps@eoas.ubc.ca>

ABSTRACT.

Large-scale glacier mass balance models often rely on positive-degree-day (PDD) melt models, which have known limitations. This study evaluates a relatively simple, elevation-dependent surface energy balance (SEB) model that requires minimal downscaling of climate input data to simulate glacier melt. Using ERA5 reanalysis data and multi-year mass balance observations from 23 glaciers across Canada, we compare mass balance models incorporating SEB and PDD components under various calibration scenarios. Initial tests with the uncalibrated SEB model highlight the importance of accurate ERA5 inputs, particularly lapse-rate corrections for 2 m air temperature. Mass balance simulations with the SEB model that includes calibrated corrections for precipitation and albedo match or outperform those with the PDD model, especially when using a machine learning-derived albedo trained on remote sensing albedo data, which tends to underestimate summer albedo in accumulation zones. Seasonal calibration further improves accuracy of the mass balance simulations by addressing biases in summer melt and winter accumulation. Despite its simplicity, the SEB model provides a good balance of performance and computational efficiency, emphasizing its utility for regional-scale applications when calibrated appropriately.

INTRODUCTION

Over the past decade, large-scale simulations of glacier mass changes predominantly use positive-degree-day (PDD) models (e.g. Clarke and others, 2015; Huss and Hock, 2015; Maussion and others, 2019; Radi and Hock, 2014; Rounce and others, 2023). These models, which estimate glacier melt based on the empirical relationship between temperature and melt, are relatively straightforward to apply but have several known limitations: they are often not transferable across different spatial and temporal scales (MacDougall and others, 2011; Rounce and others, 2020), exhibit excessive sensitivity to temperature shifts (Pellicciotti and others, 2005; Huss and others, 2008), and lack the capacity to provide process-based insights into glacier responses to climate change. As a result, large-scale projections of glacier mass change still contain significant uncertainties, particularly at the scale of individual glaciers and glacierized watersheds, where the impacts of deglaciation on local freshwater supplies are most critical (Anderson and Radi, 2020). While advances in non-linear melt models, such as deep learning approaches (Bolibar and others, 2020), overcome some of the limitations of PDD models, they also fall short in delivering process-based understanding of glacier-climate interactions. Although progress continues to be made in enhancing ice-dynamics modeling with more physics-based approaches (Rounce and others, 2023), the melt modeling component continues to rely on empirical methods, highlighting the need for physics-based melt models, such as surface energy balance (SEB) models.

Using SEB models in large-scale glacier simulations presents substantial challenges due to the limited availability of crucial climatic inputs, including shortwave and longwave radiation, wind speed, specific humidity, and surface characteristics like albedo and roughness length, all of which need to be resolved accurately at the glacier scale. Consequently, large-scale SEB models are rare and often depend on semi-empirical methods that require calibration with mass balance observations (Giesen and Oerlemans, 2013; Huss and Hock, 2018; Gunnarsson and others, 2023) or statistical downscaling techniques, which are challenging to generalize across varying timescales (Shannon and others, 2019; Noël and others, 2018). Even regional applications of complex SEB models, such as COSIPY (Sauter and others, 2020), require calibration of parameters such as albedo with mass balance observations to achieve comparable performance to PDD models (Temme and others, 2023). Until recently, mass balance observations for individual glaciers covered fewer than 0.1 percent of glaciers worldwide. The availability of glacier-specific remote sensing estimates from 2000 onward (Hugonnet and others, 2021) has facilitated broader model calibration over multi-year timescales (e.g., Rounce and others, 2023; Temme and others, 2023; Zekollari and others, 2024;

Hanus and others, 2024; Schuster and others, 2023; Compagno and others, 2022). While these estimates align well with in-situ mass balance observations for long-term trends, significant discrepancies persist at annual and sub-annual timescales (Hugonnet and others, 2021).

Recent advances in climate reanalysis products and the growing availability of remote sensing data have created new opportunities to address some limitations in large-scale SEB modeling. For instance, the ECMWF Reanalysis v5 (ERA5) reanalysis dataset (Hersbach and others, 2020) reasonably represents key meteorological inputs for SEB models at glaciers in western Canada (Draeger and others, 2024). However, while ERA5 provides surface albedo data, it poorly represents albedo at the glacier scale (Draeger and others, 2024). In contrast, Moderate Resolution Imaging Spectroradiometer (MODIS) data (Hall and Riggs, 2021) enables regional mapping of snow and ice albedo and are commonly used in SEB studies on ice caps (Gascoine and others, 2017; Gunnarsson and others, 2023) and ice sheets (Stroeve and others, 2005; Box and others, 2012). However, for mountain glaciers, MODIS-derived albedo presents larger errors and more frequent data gaps (Davaze and others, 2018; Xiao and others, 2022), and its direct application in regional-scale SEB models remains unexplored.

This study builds on recent advances by leveraging ERA5 climatic data and MODIS-derived albedo to investigate physics-based melt modeling with minimal downscaling. Specifically, we evaluate an elevation-dependent SEB model, with varying calibration levels, for glacier mass balance simulations and compare its performance against the widely-used elevation-dependent PDD model. A central focus is on representing glacier albedo in the SEB model, contrasting a simple empirical albedo model with a novel machine learning approach trained on MODIS data. With the limitations of mass balance observations at sub-annual scales such as those provided in the Hugonnet and others (2021) database in mind, we investigate the impact of using seasonal versus annual mass balance data for model calibration. We evaluate model performance across a diverse set of Canadian glaciers with in-situ mass balance observations, spanning varied climatic regimes from small temperate valley glaciers in the Canadian Rockies and the maritime climate of the Coast Mountains to large polythermal glaciers in the Arctic.

DATA AND METHODS

Study glaciers and their mass balance observations

We select 23 glaciers across Canada (Figure 1, Table 1), each with a minimum of three years of elevation-based mass balance records since 1979 (WGMS, 2022). On average, these glaciers have 13 years of mass

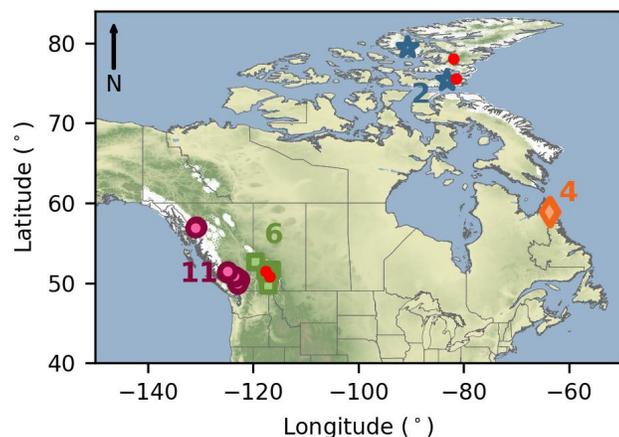


Fig. 1: Geographic distribution of glaciers included in this study, organized by subregion. Marker colors and shapes represent subregions: red circles for West Coast, green squares for Rockies, orange diamonds for East Coast, and blue stars for Arctic glaciers. The number of glaciers in each subregion is indicated by corresponding colored labels. Red dots denote locations of on-glacier meteorological observation sites used for ERA5 data comparison. Map projection: WGS84.

balance data post-1979, with individual records ranging from 3 to 39 years. The glaciers are clustered into four geographic subregions: West Coast, Rockies, East Coast and Arctic (Figure 1). Devon Ice Cap North West in the Arctic is the largest glacier in our sample at 765.44 km², though it represents an outlier in size, with the next largest, Bridge Glacier, covering 81.98 km². The smallest glacier, Abraham Glacier, located on Canada's East Coast, has an area of 0.503 km². Of the 23 glaciers 17 also have seasonal mass balance measurements and mass balance profiles, allowing for more precise calibration of the elevation-dependent mass balance model and a more comprehensive comparison of model outputs than possible with annual mass balance data alone. Further details for each glacier are available in Table S3.

Climate data and downscaling

We use surface-level, hourly ERA5 reanalysis climate data at a 30 × 30 km grid resolution (Hersbach and others, 2020) from 1979 to 2019 in this study. For each glacier, we select the reanalysis data from the nearest ERA5 grid point to its central coordinates. Depending on the variable, hourly data is either averaged or summed to produce daily values. The PDD model utilizes mean daily near-surface temperature, while the SEB model relies on mean daily values for near-surface temperature, dewpoint temperature, wind speed, surface pressure, as well as daily totals for incoming shortwave and longwave radiation, and precipitation. Both daily precipitation and mean near-surface temperature are used in the accumulation model. To demonstrate difference in these variables among the four subregions, we show their timeseries as area-

weighted mean across all glaciers per subregion (Figure S1).

Among the meteorological inputs used in this study, we apply downscaling only to temperature by lapse-rate correcting ERA5 temperature to account for elevation differences between the ERA5 grid cell and the glacier site. Following Draeger and others (2024), we calculate the lapse rate using multi-level monthly ERA5 temperature data between the surface and 400 hPa, spanning 17 levels. This monthly lapse rate is linearly interpolated to obtain a daily lapse rate, lr ($^{\circ}\text{C m}^{-1}$), which we then apply to correct the temperature at each elevation band of each glacier:

$$T_h = T_{ERA} + lr(h - h_{ERA}), \quad (1)$$

where T_h ($^{\circ}\text{C}$) is the corrected near-surface temperature for the given elevation, h (m a.s.l.), while T_{ERA} ($^{\circ}\text{C}$) is ERA5 temperature at the elevation of the grid cell, h_{ERA} (m a.s.l.).

Albedo data

We use the MOD10A1 MODIS/Terra Snow Cover Daily L3 Global 500 m SIN Grid, Version 6 dataset (Hall and Riggs, 2021) to derive daily snow albedo values for our 23 study glaciers during the summer months (June to September) from 2000 to 2019. The MODIS Terra satellite, operating in a near-polar, sun-synchronous orbit, captures up to 9 daily observations of high Arctic glaciers (Williamson and others, 2020). Data processing utilizes the Normalized-Difference Snow Index (Hall and others, 1995), which calculates albedo based on the visible-to-infrared radiation ratio and identifies snow-covered surfaces. The algorithm dynamically assigns snow albedo values to snow/ice-covered grid cells, leaving non-snow-covered cells with null values. Daily snow albedo data is provided on a 500×500 m grid, with the highest quality observation each day selected based on illumination, satellite angle, cloud cover, and fractional snow cover (Hall and Riggs, 2021).

To ensure data quality, we filter the MODIS albedo values by removing physically unrealistic observations, specifically those greater than 0.99 and less than 0.05 (Williamson and others, 2020). As our model operates at an elevation-band level, we convert gridded daily MODIS data into an average daily albedo for each elevation band of each glacier (detailed in the next section). Due to cloud cover, the dataset is often incomplete; not all days contain observations, and available data may cover only select elevation bands. For some bands, no albedo observations are available. Across the study glaciers, albedo observations are present in at least one elevation band 17% to 77% of the time, with an average observation coverage of

Fig. 2: Daily MODIS-derived albedo (α) compared to in-situ albedo from radiometer measurements at on-glacier sites over multiple summer seasons for four glaciers (listed in the legend; see text for site details).

44%.

Comparison between MODIS-derived and in-situ albedo

MODIS-derived albedo often differs from in-situ albedo measurements at glacier sites (Davaze and others, 2018), but such comparisons are limited due to the scarcity of field observations. To estimate these differences for our study glaciers, we compiled in-situ albedo observations from four glacier sites in or near our study area (see Figure 2). Radiometer data, which record incoming and reflected shortwave radiation, are available for two study glaciers in the Rockies, covering multiple summer seasons (Fitzpatrick and others, 2017, 2019). For the Arctic subregion, radiometer observations are accessible from nearby glaciers, specifically Belcher Glacier (two summer seasons; Sharp and others (2011)) and Prince of Wales Glacier (one summer season; S. Marshall, personal communication, May 2020). We derive daily average albedo values from these radiometer data by calculating the ratio of integrated hourly reflected to incoming shortwave radiation, excluding hours with solar zenith angles exceeding 65° to minimize error and bias (Fitzpatrick and others, 2017). Additional details on the radiometer measurements are provided in Table S1. In processing MODIS data to derive an average albedo for each glacier elevation band, we select the MODIS-derived albedo from the band nearest to each glaciers site with daily in-situ albedo values (Figure 2, Table 2). The comparison shows that MODIS-derived albedo consistently underestimates albedo at all four sites, with a mean bias error of -0.14 across sites.

Models

To simulate mass changes of our study glaciers, we use an elevation-dependent mass balance model, commonly used for large-scale simulations (e.g., Radi and Hock, 2014; Huss and Hock, 2015; Rounce and others, 2023). Our focus is on modeling the climatic or reference mass balance, which assumes constant glacier area and hypsometry over time, set to the values from the initial year of the mass balance observational record for each glacier. Consequently, the model does not account for ice flow contributions to glacier geometry changes or the feedback between changes in glacier thickness, area, and mass balance. For consistency with the mass balance data, the hypsometry is sourced from the WGMS dataset, where glacier area and mass balance are provided for 100 m elevation bands.

For each elevation band, we calculate daily specific mass balance (in m water equivalent, w.e.), as the difference between modeled daily accumulation and ablation. The glacier-wide mass balance is then computed as the area-weighted mean of specific mass balance across all elevation bands for each glacier. Summer mass balance is derived as the integrated mass balance over the summer season, defined as the period from 01 May to 01 October, which roughly corresponds to the reported observation dates for the study glaciers (WGMS, 2022). Winter mass balance is calculated over the remaining months of the year. We maintain a constant definition for summer and winter periods to allow for fair comparison between glaciers and over the model period, as well as to maintain consistency with other regional/global approaches to mass balance modeling (e.g., Radi and Hock, 2014; Shannon and others, 2019).

Accumulation model

We adopt a relatively simple accumulation model, commonly used in previous large-scale glacier mass balance simulations (e.g., Radi and Hock, 2014; Rounce and others, 2023), where daily accumulation, c (in m w.e.), at elevation h (m a.s.l.), the center elevation of each elevation band, is calculated as:

$$c(h) = \begin{cases} 0 & T_h > 2 \\ k_p \times (P_{ERA} + d_p(h - h_{ERA})) & T_h \leq 2, \end{cases} \quad (2)$$

where T_h ($^{\circ}\text{C}$) is the lapse-rate corrected ERA5 temperature for the given elevation band, P_{ERA} (m day $^{-1}$) is ERA5 precipitation, and k_p and d_p are calibration-dependent parameters: a multiplicative precipitation factor (k_p) to correct the ERA5 precipitation, which often underestimates precipitation over

mountainous regions (Tarek and others, 2020), and a precipitation gradient (d_p , 100m^{-1}) to account for orographic precipitation increases not captured by ERA5's original spatial resolution. Despite not being mass conserving, this correction is often necessary to achieve a better match between modeled and observed profiles of winter glacier mass balance (Radi and Hock, 2014). Both parameters are calibrated against the mass balance observations for each of our study glaciers; further details are provided later in this section.

Melt model: PDD approach

We use a PDD melt model following Radi and Hock (2014), which assumes an empirical relationship between temperature and glacier's daily melt, a (in m w.e.), given by:

$$a(h) = f_{ice/snow} \int \max(T_h, 0) dt, \quad (3)$$

where $f_{ice/snow}$ is the calibration-dependent melt factor for ice or snow (m w.e. $\text{day}^{-1} \text{ } ^\circ\text{C}^{-1}$); both are constant values that are individually calibrated for each glacier. The melt factor for snow (f_{snow}) is used above the assumed firn-line elevation regardless of snow cover, while below the firn-line elevation we apply the melt factor for ice (f_{ice}) when snow depth is zero and f_{snow} when the snow depth is greater than zero. Following Huss and Hock (2015) and Rounce and others (2020), the firn-line elevation is set to the elevation where the five-year running mean of each elevation band's annual mass balance changes from positive to negative.

Melt model: SEB approach

We use a simple SEB model, keeping only the key contributors to melt energy as is done in Draeger and others (2024). Daily surface melt, a (in m w.e.) is calculated for each elevation band as:

$$a(h) = \frac{Q_{melt}}{L_f \rho_w}, \quad (4)$$

where Q_{melt} is the total daily energy available for surface melt (J m^{-2}), L_f is the latent heat of fusion (334000 J kg^{-1}) and ρ_w is the density of water (1000 kg m^{-3}).

To calculate the energy available for melt, we use:

$$Q_{melt} = Q_S^\downarrow (1 - \alpha) + Q_L^\downarrow - Q_L^\uparrow + Q_H + Q_E. \quad (5)$$

Q_S^\downarrow represents incoming shortwave radiation, α is surface albedo, simulated by albedo models described below, Q_L^\downarrow and Q_L^\uparrow represent incoming and outgoing longwave radiation, respectively, and Q_H and Q_E denote the sensible and latent heat fluxes, respectively. All fluxes are integrated over one day, thus expressed in J m^{-2} . We intentionally neglect: (i) empirical correction schemes applied to shortwave radiation fluxes, such as separating direct and diffuse components, (ii) conductive heat flux into the ground, and (iii) energy supplied by rain. These terms generally rely on calibration-dependent parameterizations and typically play a minimal role in determining seasonal melt, depending on glacier type and location (Hock, 2005; Fitzpatrick and others, 2017, 2019). Our rationale for excluding these terms is to simplify the SEB model while retaining the most impactful terms for surface melt.

Q_L^\downarrow is taken directly from ERA5 without modifications. Q_L^\uparrow is computed using the Stefan-Boltzmann law for blackbody radiation, assuming the ice/snow surface remains at a constant temperature of 0°C year-round (i.e., the ice/snow surface is consistently at the melting point) (Hock, 2005; Cuffey and Paterson, 2010; Fitzpatrick and others, 2017). We also tested an approach where the ice/snow surface temperature is set to the lapse-rate-corrected 2 m air temperature (T_h), when it falls below freezing, but observed negligible effects on the modeled annual Q_{melt} relative to the former approach.

The turbulent fluxes, Q_H and Q_E , are calculated using the bulk aerodynamic method, specifically following the C_{\log} approach described by Fitzpatrick and others (2017):

$$\begin{aligned} Q_H(h) &= \frac{c_p \rho_a C_H p u (T_h - T_0)}{p_0} \\ Q_E(h) &= \frac{0.622 l_v \rho_a C_E p u (e_h - e_0)}{p_0}, \end{aligned} \quad (6)$$

where c_p is the specific heat capacity of air ($1005 \text{ J kg}^{-1} \text{ K}^{-1}$), p is the surface air pressure (hPa), p_0 is the air pressure at sea level (1013.25 hPa), ρ_a is the density of air at sea level (1.29 kg m^{-3}), l_v is the latent heat of vaporization of water ($2.514 \times 10^6 \text{ J kg}^{-1}$), and T_0 is the temperature of the ice/snow surface (assumed to be at 0°C over a summer season).

The mean daily wind speed, u (m s^{-1}), is taken directly from ERA5 data ($u = u_{\text{ERA}}$) and is also tested with a calibration-dependent multiplicative wind correction factor, k_u , such that $u = k_u \times u_{\text{ERA}}$. The vapor pressures e_0 and e_h represent the surface vapor pressure (assumed to be 6.1078 hPa at 0°C) and the vapor pressure 2 m above the surface, respectively. The latter is calculated using ERA5 data via the August-Roche-Magnus formula (Alduchov and Eskridge, 1996), with relative humidity derived from ERA5 temperature and dewpoint temperature.

The bulk transfer coefficients for heat (C_H) and moisture (C_E) are dimensionless and are parameterized following Fitzpatrick and others (2017), without stability corrections. These coefficients depend on the roughness lengths for momentum ($z_{0,v}$), temperature ($z_{0,T}$), and humidity ($z_{0,q}$) at the glacier surface. For this study, we use roughness length values of $z_{0,v} = 10^{-3}$ m, $z_{0,T} = 10^{-5}$ m, and $z_{0,q} = 10^{-5}$ m, consistent with typical observations at mid-latitude glaciers (Hock, 2005; Fitzpatrick and others, 2019).

Albedo models

During the winter months, the albedo for each glacier is assigned its specific snow albedo value, calculated as the mean annual maximum MODIS albedo over the summer months during the 2000–2019 period (see Table S2 for individual glacier values). Across the 23 study glaciers, the average MODIS-derived albedo values are 0.85 for snow, 0.52 for firn, and 0.23 for ice. The evolution of summer albedo in each glacier's elevation bands is simulated using two distinct modeling approaches. The first approach, denoted as α_{Dec} , is based on the empirical model of Hirose and Marshall (2013), developed for a glacier in the Canadian Rockies. This model employs a logarithmic expression to represent the decay of albedo following snowfall events. The daily summer albedo, $\alpha_{Dec}(t)$, is simulated for each elevation band as:

$$\alpha_{Dec}(t) = \alpha_{snow} - A \ln \sum \text{PDD}(t), \quad (7)$$

where α_{snow} is the albedo of fresh snow, A is a decay factor calibrated individually for each glacier, and $\text{PDD}(t)$ represents the cumulative positive degree days (in °C) up to time t . The sum of $\text{PDD}(t)$ is calculated daily, starting from the onset of the melt season and resetting to zero after each fresh snowfall. Hirose and Marshall (2013) calibrate this model against in-situ albedo observations, and find that $A = 0.05668$. We optimize A using the Nelder-Mead method (Nelder and Mead, 1965) to minimize the root-mean-square-error (RMSE) between the MODIS-derived and modeled daily albedo over the 2000–2019 period. Optimization of A is performed individually for each glacier, with RMSE calculated using two metrics: the daily glacier-wide albedo and the mean summer albedo for each elevation band, both weighted equally. The optimized A values for each glacier are provided in Table S2. To ensure physical realism, we constrain albedo in the model to avoid values below a minimum threshold: above the firn-line, this threshold is set to firn albedo, and below the firn-line, it is set to ice albedo. For most glaciers, we find A to be below 0.5, except for the Arctic glaciers (see Table S2).

The second albedo model, denoted as α_{MLP} , uses a neural network multilayer perceptron (MLP) to

simulate daily summer albedo. The MLP model is trained on MODIS-derived albedo values for each elevation band, aggregated across glaciers within the same subregion rather than on individual glaciers. This approach assumes that a model trained on a broader dataset within a subregion is more robust, benefiting from the increased amount of training data, as machine learning models typically improve with larger datasets (Hsieh, 2009). The MLP is a feed-forward neural network consisting of an input layer, one or more hidden layers, and an output layer, with neurons in each layer connected via non-linear activation functions. During training, the model's parameters (weights) are optimized to minimize the mean square error between the predicted and observed MODIS-derived daily albedo values for each elevation band.

For each subregion, the MLP model is trained using data from any study glacier with available MODIS albedo observations in a given elevation band, implicitly incorporating elevation information. When multiple glaciers have observations in the same elevation band, their data are combined for model training. The model inputs include the day of the year, as well as ERA5 temperature and precipitation data for the target glacier, spanning the current day and the preceding six days. Each input variable is normalized to a range of 0 to 1 based on its minimum and maximum values. The paired input data and MODIS albedo observations are randomly shuffled and split into training and testing datasets, ensuring data from each glacier within the subregion are represented in both. We use 80% of the data for model training and validation and reserve 20% for testing. Within the training set, 80% is allocated for training and 20% for validation.

For the model runs, we test various configurations with 1 to 2 hidden layers, each containing between 1 and 15 neurons. The model with the best performance (i.e., lowest mean square error) is selected. This process is repeated 20 times using different initial weights, creating an ensemble of the 20 best-performing MLP models for each elevation band and subregion. We experimented with several model architectures, including training separate models for each glacier and using different combinations of input parameters. Ultimately, we chose the described model setup as it minimized the need for interpolation and effectively reproduced the MODIS observations. The average RMSE across all glaciers between modeled and MODIS-derived albedo is 0.11.

The model is implemented in Python (Van Rossum and Drake, 2009), utilizing the Scikit-Learn library (Pedregosa and others, 2011). We use the Rectified Linear Activation Unit (ReLU) as the activation function and optimize the model with the Adaptive Moment Estimation (Adam) method. Adam combines the advantages of both moving average gradients and adaptive learning rates to update the neural network

weights, effectively minimizing the objective function (Kingma and Ba, 2014; Sun and others, 2019). For further details on these algorithms, refer to Kingma and Ba (2014).

After calibrating or training both albedo models, the resulting albedo values are incorporated into the SEB model. Given the known biases between MODIS-derived and in-situ albedo, as well as discrepancies in ERA5 incoming shortwave radiation compared to observations at glacier surfaces (Draeger and others, 2024), we introduce an additional calibration parameter as an additive correction to the modeled albedo. The final albedo is then computed as:

$$\alpha = \alpha_{mod} + \Delta\alpha, \quad (8)$$

where α_{mod} represents the albedo value from either the empirical model (α_{Dec}) or the machine learning model (α_{MLP}), and $\Delta\alpha$ is the additive correction factor. The melt model that uses the empirical representation of albedo is denoted SEB_{Dec} and the one that uses the MLP albedo representation is denoted SEB_{MLP} .

Since neither of the two albedo models fully capture short-term albedo variations due to fresh snowfall events during the summer, we account for this process separately for each study glacier. On days when daily snowfall, as estimated from the accumulation model (Equation 2), exceeds 0.5 mm in a given elevation band, we override the modeled albedo by setting it to the snow albedo value derived from MODIS data. This adjustment is applied regardless of the albedo value predicted by the empirical or machine learning models. Finally, to ensure physical plausibility, the albedo time series for each elevation band is constrained such that the values do not exceed 0.9 or fall below 0.1.

Model calibration

As described in previous sections, several calibration-dependent parameters are optimized during the calibration of the mass balance model. These include the precipitation correction parameters (k_p and d_p) in the accumulation model and the melt factors (f_{snow} and f_{ice}) in the PDD model. While the SEB model is originally designed to be calibration-independent, we also test model runs that incorporate calibration-dependent wind correction (k_u) and albedo correction ($\Delta\alpha$). Since the melt and accumulation models must operate together to generate mass balance estimates, it is not possible to optimize them independently. All calibration procedures aim to minimize the RMSE between the model output and observed mass balance data (WGMS) for each study glacier. The RMSE is computed from both the seasonal mass balance time

series and the average seasonal mass balance profile, with both metrics given equal weight. For comparison, we also perform a calibration where the RMSE is calculated from annual mass balance observations instead of seasonal ones.

The parameters are optimized using the Nelder-Mead optimization scheme (Nelder and Mead, 1965). The bounds for the precipitation correction parameter (d_p) range from 0 to 3% per 100 m elevation increase, and for the precipitation factor (k_p), they range from 1 to 3, in line with the mass balance model used by Radi and Hock (2014) for Canadian glaciers. Following Hock (2003), the bounds for the snow melt factor (f_{snow}) are set between 1 and 9 mm w.e. day⁻¹ °C⁻¹, and for the ice melt factor (f_{ice}), between 3 and 11 mm w.e. day⁻¹ °C⁻¹, with the condition that f_{ice} must be greater than f_{snow} . The bounds for the albedo correction parameter ($\Delta\alpha$) are between 0 and 0.3, with the upper bound determined by the difference between the maximum physically reasonable albedo for snow (0.9) and the lowest MODIS-derived snow albedo (0.6 for the Nordic glacier). The multiplicative wind correction factor (k_u) is bound between 0 and 4, based on the work of Shannon and others (2019).

Model intercomparison and sensitivity tests

Our primary goal is to evaluate and compare the performance of the models in replicating the observed (WGMS) inter-annual variability of glacier-wide summer, winter, and net mass balance, as well as the averaged mass balance profile over the observational period. Given the limited sample size, we do not apply calibration-validation methods (e.g., leave-one-out cross-validation); thus, our assessment focuses on how well each model fits the available data. We begin by analyzing model runs where the SEB model parameters remain uncalibrated. We then proceed to evaluate runs with an optimized albedo correction ($\Delta\alpha$), and finally, we assess runs that include both optimized albedo and wind corrections.

We also evaluate the impact of temperature perturbations on the simulated mass balance over the 1980–2019 period. To do this, we apply uniform temperature adjustments to the ERA5 input temperatures for each glacier and across all models, including the albedo models. The temperature perturbations are set at -2°C, -1°C, +1°C, and +2°C for the entire 40-year period. The impact is quantified by calculating the mass balance sensitivity ($\frac{\partial B}{\partial T}$) for each glacier, defined as the difference in the mean modeled mass balance between the adjusted temperature scenario and the original scenario, divided by the temperature perturbation.

RESULTS

Calibrated parameters

Here we provide a summary of the calibrated parameter values across the study glaciers, with detailed values available in Table S2. The precipitation correction parameters (k_p and d_p) are calibrated separately for each of the three melt models (SEB_{MLP} , SEB_{Dec} , and PDD). The multiplicative factor k_p had mean values (with relative variance in parentheses) of 1.76 (19%) for the SEB_{MLP} model, 1.94 (16%) for the SEB_{Dec} model, and 1.87 (27%) for the PDD model. The precipitation gradient d_p averaged 1.11 (48%) 100m^{-1} for the SEB_{MLP} model, 0.90 (72%) 100m^{-1} for the SEB_{Dec} model, and 0.85 (75%) 100m^{-1} for the PDD model.

In the PDD model, we observe substantial variability in the melt factors (f_{snow} and f_{ice}) across the study region, though some consistency is noted within certain subregions. On the West Coast, glaciers generally displayed minimal differences between f_{snow} and f_{ice} , with a mean f_{snow} of 3.68 mm w.e. $\text{day}^{-1} \text{ } ^\circ\text{C}^{-1}$ (5.8% variance) and a mean f_{ice} of 3.84 mm w.e. $\text{day}^{-1} \text{ } ^\circ\text{C}^{-1}$ (1.4% variance). In contrast, East Coast glaciers exhibited greater variability and larger differences between melt factors, with a mean f_{snow} of 1.75 mm w.e. $\text{day}^{-1} \text{ } ^\circ\text{C}^{-1}$ (16.7% variance) and a mean f_{ice} of 4.84 mm w.e. $\text{day}^{-1} \text{ } ^\circ\text{C}^{-1}$ (69.0% variance). For Arctic glaciers, where only two glaciers are included in the analysis, the f_{ice} values spanned a wide range, from 3.77 to 10.79 mm w.e. $\text{day}^{-1} \text{ } ^\circ\text{C}^{-1}$.

In the SEB models, the albedo correction factor ($\Delta\alpha$) ranged from 0 to 0.3 across the study glaciers. The mean $\Delta\alpha$ is 0.11 for the MLP model and 0.19 for the logistic decay model. Analyzing the subregional differences, the mean $\Delta\alpha$ values (with relative variance) for the MLP model are as follows: 0.06 (141.4%) in the Arctic, 0.22 (66.7%) on the East Coast, 0.09 (74.6%) on the West Coast, and 0.08 (85.1%) in the Rockies. For the logistic decay model, the mean values are 0.18 (89.1%) in the Arctic, 0.28 (11.1%) on the East Coast, 0.16 (51.1%) on the West Coast, and 0.18 (36.1%) in the Rockies.

Modeled versus observed albedo

To assess the performance of the two albedo models in simulating seasonal albedo evolution, we compare timeseries of modeled and MODIS-derived daily glacier-wide albedo, averaged over all available MODIS data days and then averaged across glaciers in each subregion (Figure 3, a-d). Both the MLP and logistic decay models successfully capture the gradual decrease in albedo during the early melt season, showing the

Fig. 3: a-d: Time series of daily average albedo, shown as a 5-day running mean, averaged across glaciers within each subregion. Values are derived from MODIS observations (black), the machine learning (MLP) model (blue), and the logistic decay model (green), calculated as the mean of all days with MODIS observations during the 2000-2019 period. e-f: Average summer (June-August) albedo versus glacier elevation for each subregion, derived from MODIS and the two models, calculated as the mean of all days with MODIS observations for each elevation of the study glaciers during the same period.

best agreement with MODIS-derived albedo for Arctic and East Coast glaciers. However, for glaciers in western Canada, the MLP model more accurately follows the seasonal pattern of MODIS-derived albedo compared to the logistic decay model, which exhibits a faster decline in albedo at the onset of the melt season. In contrast, the MLP model tends to produce higher albedo values in late summer relative to MODIS-derived albedo.

We further compare modeled and MODIS-derived average albedo across elevation bands in each subregion, averaged over all available MODIS observation days (Figure 3, e-h). Both the modeled and MODIS-derived albedo generally display the expected increase with elevation. However, relatively low absolute values are observed at the highest elevation bands in both MODIS data and the models. Except for Arctic glaciers, the average summer albedo for the highest elevations rarely exceeds 0.5, whereas a range of 0.6 to 0.7 would be more physically realistic for the accumulation zones of glaciers without debris cover (Hock, 2005; Cuffey and Paterson, 2010). More detailed results of the elevation profiles for individual glaciers are provided in Figure S2, where these subregional patterns are consistently confirmed.

The relatively low MODIS-derived albedo in glacier accumulation zones is further supported by comparisons with in-situ albedo observations from multiple glacier sites (Figure 4). We closely examine the timeseries of daily albedo values from both in-situ measurements and MODIS observations, alongside outputs from the two albedo models, for two glaciers in the Rockies where in-situ albedo data are available. The modeled albedo includes a calibrated correction factor ($\Delta\alpha$) and adjustments for fresh snowfall events. Both albedo models successfully capture the gradual decline in albedo during the early melt season and the sharp increase as the melt season ends. Although neither model fully replicates short-term albedo variations due to fresh snowfall, they generally indicate the presence of these events, albeit with some discrepancies in timing compared to observations. Across all sites, the RMSE between modeled and in-situ albedo timeseries ranges from 0.12 to 0.16, suggesting a reasonable fit. However, MODIS-derived albedo is observed to be consistently lower than in-situ measurements, particularly in the accumulation zone (e.g., Station 3 on Conrad Glacier), where MODIS-derived albedo is on average 0.2 units lower than the in-situ observations over the observational period.

Finally, we examine whether the calibrated albedo correction factor ($\Delta\alpha$) is associated with the tendency of both MODIS data and the models to underestimate albedo values at higher glacier elevations. Our analysis shows a strong, statistically significant negative correlation between the magnitude of the calibrated albedo correction and the maximum MODIS-derived summer albedo across the elevation range typically

Fig. 4: Timeseries of daily albedo at multiple sites on two glaciers (Conrad and Nordic) in the Rockies, shown for various summer seasons. The albedo values are derived from: MODIS observations (black), automatic weather station (AWS) measurements (grey), the logistic decay model (green), and the machine learning (MLP) model (blue). Both the logistic decay and MLP models include the calibrated albedo correction factor, $\Delta\alpha$ (refer to text for calibration details). MODIS albedo values are taken from the grid cell nearest to the AWS location, while modeled albedo values correspond to the elevation band closest to each AWS site. Data from Conrad Glacier are labeled as: Conrad Stn 1 (a and b), Conrad Stn 2 (c), and Conrad Stn 3 (d), while data from Nordic Glacier are labeled as Nordic Stn (e). Further site details are available in Table S1.

Fig. 5: Albedo correction ($\Delta\alpha$) from the SEB_{MLP} model versus the maximum MODIS-derived albedo along the glacier elevation range for each study glacier. Points are color-coded by subregion, with the linear regression slope and Pearson's correlation coefficient shown in the top right corner.

in the accumulation zone for most glaciers (Figure 5). This finding suggests that lower albedo values in the accumulation zone necessitate a larger correction to align the modeled mass balance with observations during the calibration process.

Model intercomparison

We compare the performance of the PDD and SEB modeling approaches using box-and-whisker plots that illustrate the distribution of normalized root-mean-square-error (NRMSE) across the study glaciers (Figure 6). The NRMSE is calculated by dividing the RMSE, which measures the discrepancy between observed and modeled mass balance timeseries as well as the averaged mass balance profiles (using the same RMSE metric applied during model calibration), by the range of observed values (i.e., the difference between the maximum and minimum values). This normalization provides a standardized error metric, enabling direct comparison across glaciers of different sizes. We evaluate NRMSE separately for annual (net), summer, and winter mass balances. The annual NRMSE is calculated for all 23 study glaciers, while the seasonal NRMSE is evaluated for a subset of 17 glaciers.

First, we examine the results from the uncalibrated SEB models and the calibrated PDD model, where melt factors are individually optimized for each glacier (Figure 6). In these runs, no precipitation corrections are applied (i.e., $d_p = 0$ and $k_p = 1$). The performance for winter mass balance is similarly poor across all three models, with median NRMSE values of 51.5%, 54.0%, and 42.4% for the SEB_{MLP} , SEB_{Dec} , and PDD models, respectively. Statistical analysis using an independent two-sample t-test reveals no

significant differences among these distributions, as shown by the box plots. In contrast, the PDD model shows significantly better performance than the SEB models for summer and net mass balance. For annual (net) mass balance, the median NRMSE of the PDD model is 17.8%, substantially lower than the 47.4% and 72.7% recorded by the SEB_{MLP} and SEB_{Dec} models, respectively.

Next, we consider the model intercomparison using SEB models calibrated with varying sets of calibration-dependent parameters (Figure 6). In these runs, all models include calibrated precipitation correction parameters (k_p and d_p). Compared to the previous results without precipitation correction, incorporating these parameters leads to lower median NRMSE values for winter mass balance across the study glaciers. For summer mass balance, we observe a median NRMSE reduction of 9.2% for the SEB_{MLP} model, 2.2% for the SEB_{Dec} model, and 0.5% for the PDD model. Notably, the NRMSE distributions for summer and annual mass balance in the SEB models, when only precipitation factors are calibrated, show statistically significant differences from those of the PDD model, as indicated by an independent two-sample t-test.

Incorporating the albedo correction into the SEB models results in the most significant improvement in summer mass balance performance, reducing the median NRMSE from 24.6% to 13.0% for the SEB_{MLP} model and from 44.9% to 13.1% for the SEB_{Dec} model. Similarly, for annual mass balance, the albedo correction decreases the median NRMSE from 18.5% to 11.4% in the SEB_{MLP} model and from 38.5% to 10.7% in the SEB_{Dec} model. With this correction, the NRMSE distributions of the SEB models are no longer significantly different from those of the PDD model. Adding a wind correction alongside precipitation and albedo corrections provides only minor, statistically insignificant improvements, lowering the median summer NRMSE by 1.3% for the SEB_{MLP} model and by 1.6% for the SEB_{Dec} model.

Across all subregions, we observe a consistent reduction in NRMSE values with the introduction of calibrated parameters. In each subregion, incorporating the albedo correction significantly reduces NRMSE for both the SEB_{MLP} and SEB_{Dec} models, while adding the wind correction yields more modest improvements. The East Coast glaciers typically have the highest mean annual NRMSE values, starting at 30.3% for the uncalibrated SEB_{MLP} model, which decreases to 25.6% after applying the albedo correction, and results in a mean annual NRMSE of 21.8% for the PDD model. In contrast, the Rockies glaciers exhibit the lowest NRMSE values, with the SEB_{MLP} model achieving a mean annual NRMSE of 16.2%, which further decreases to 11.9% with the albedo correction applied, while the PDD model has a comparable mean annual NRMSE of 12.3%.

At the scale of individual glaciers, we find that the PDD and SEB_{MLP} models, with calibrated pre-

Fig. 6: Box plots of normalized root-mean-square error (NRMSE) for annual (a), summer (b), and winter (c) mass balance across 23 glaciers, comparing SEB_{MLP} (blue), SEB_{Dec} (green), and PDD (pink) models under different calibration scenarios. Scenarios include: no calibration (SEB_{MLP} , SEB_{Dec}), precipitation-only ($SEB_{MLP} P$, $SEB_{Dec} P$), precipitation and albedo ($SEB_{MLP} P,\alpha$, $SEB_{Dec} P,\alpha$), and precipitation, albedo, and wind corrections ($SEB_{MLP} P,\alpha,u$, $SEB_{Dec} P,\alpha,u$). PDD model results are shown for melt factor calibration (PDD) and calibration of melt factor and precipitation (PDD P). Solid lines represent median NRMSE, dashed lines indicate mean NRMSE, and median values are labeled above the whiskers. Boxes span the interquartile range, with whiskers extending to the 5th and 95th percentiles.

Fig. 7: Normalized root-mean-square error (NRMSE) for summer mass balance: SEB_{MLP} vs. PDD model (a) and SEB_{Dec} vs. PDD model (b). SEB models include calibrated precipitation and albedo corrections, while the PDD model uses calibrated melt factors and precipitation. Results are shown for 17 glaciers with seasonal observations, colored by subregion.

precipitation and albedo corrections, produce similarly good fits to the mass balance observations (annual comparison of NRMSE is shown in Figure S3). Specifically, the NRMSE for summer mass balance across the 17 glaciers is strongly and significantly correlated between the two models ($r = 0.88$; Figure 7). This correlation in NRMSE is stronger for the SEB_{MLP} model than for the SEB_{Dec} model, which exhibits substantially larger NRMSE values than the PDD model for several glaciers. Overall, the correlation of NRMSE between the SEB and PDD models is stronger for winter (r in the range 0.86 to 0.89) than for summer mass balance (r in the range 0.70 to 0.88). Notably, the SEB_{MLP} model consistently achieves substantially lower NRMSE values than the PDD model for winter mass balance across multiple glaciers (Figure S4).

Finally, we examine the effects of calibrating the models using annual versus seasonal mass balance observations (Figure 8). For consistency, we limit this analysis to the 17 glaciers with available seasonal mass balance data. We compare model runs for the SEB models with calibrated precipitation and albedo corrections, and the PDD model that has calibrated precipitation and melt factors. For all three models (SEB_{MLP} , SEB_{Dec} , and PDD), calibration based on annual mass balance observations leads to lower median NRMSE values for annual mass balance; however, the differences are not statistically significant. In contrast, for seasonal mass balance, median NRMSE values are either similar or higher, with a notably larger spread in NRMSE across the 17 glaciers. When calibrated solely on annual mass balance, the models

Fig. 8: Box plots of normalized root-mean-square error (NRMSE) for annual, summer, and winter mass balance across 17 glaciers with seasonal observations. Results compare SEB_{MLP} (blue), SEB_{Dec} (green), and PDD (pink) models under two calibration scenarios: seasonal (“S”) and annual (“A”) mass balance observations. SEB models include calibrated precipitation factors and albedo corrections, while the PDD model includes calibrated melt and precipitation factors. Solid lines indicate median NRMSE, dashed lines show mean NRMSE, and median values are labeled above the whiskers. Boxes represent the interquartile range, with whiskers extending to the 5th and 95th percentiles.

for many glaciers tend to overestimate both summer melt and winter accumulation, balancing these errors to achieve a better fit to the net mass balance (see Figure S5). In contrast, calibration using seasonal mass balance observations effectively reduces biases in winter and summer mass balances, leading to more accurate seasonal simulations.

Modeled versus observed in-situ mass balance

In addition to evaluating NRMSE, we assess how well the models replicate the in-situ (stake) mass balance observations at the study glaciers. Although these in-situ measurements are not directly used for model calibration, their aggregates, such as glacier-wide mass balance timeseries and average mass balance profiles, are considered. For each glacier, we evaluate the “goodness of fit” between modeled and observed (WGMS) seasonal and annual in-situ mass balances, using the modeled value from the elevation band nearest to the stake measurement. The performance metrics include RMSE, mean bias error (MBE), Pearson correlation coefficient (r), and Nash-Sutcliffe efficiency coefficient (NSE), as detailed in Table 3. For conciseness, we present results only for the SEB model runs incorporating calibrated precipitation and albedo corrections.

Across all subregions and models, there is a statistically significant correlation (at the 95% confidence level) between modeled and observed in-situ annual mass balance (Table 3). For summer mass balance, the correlation coefficients are consistently higher than for winter mass balance, with all correlations statistically significant (p -value < 0.05), except for the winter mass balance in the Arctic subregion. Despite this, the RMSE for winter mass balance is generally smaller than for summer mass balance, likely due to the lower inter-annual variability in winter mass balance. When comparing RMSE and correlation coefficients, the PDD model gives a slightly better fit to observations than the SEB_{MLP} model in the East Coast subregion. In contrast, the SEB_{MLP} model either performs similarly to or marginally better than the PDD model in the other subregions. A detailed comparison of the modeled versus observed in-situ mass balance for each study glacier is provided in Figures S7 and S8.

Sensitivity to temperature perturbations

To quantify mass balance sensitivity to temperature changes, the models were first run with unperturbed ERA5 data over the 1980–2019 period, followed by runs in which ERA5 temperatures were systematically perturbed. When forced with unperturbed ERA5 data, the PDD and SEB_{MLP} models produced closely resembling mass balance reconstructions across all study glaciers. Both models captured a consistent

Fig. 9: Modeled summer, winter, and annual (net) mass balance from 1980–2019, averaged across all study glaciers, comparing SEB_{MLP} , SEB_{Dec} , and PDD models. a: Results using unperturbed ERA5 temperature data. b: Results with a $+1^{\circ}\text{C}$ temperature perturbation. SEB model results include calibrated precipitation and albedo correction factors, while the PDD model incorporates calibrated precipitation and melt factors.

negative long-term trend over the 40-year period, with the SEB_{MLP} model generally exhibiting a slower decline in mass balance compared to the PDD model (Figure S6). To assess whether these differences in long-term trends were statistically significant, we performed an analysis of covariance (ANCOVA). The results indicate that none of the glacier-specific trends differ significantly at the 95% confidence level (noting that ANCOVA considers both trend slopes and intercepts).

In addition to mass balance trends, both models yielded similar interannual variability in mass balance. On average, the variance of the PDD models mass balance time series was higher than that of the SEB model, with a mean variance ratio ($\sigma_{PDD}^2/\sigma_{SEB}^2$) of 1.49 ± 0.36 across all study glaciers. However, when averaged across glaciers, differences in trends and interannual variability among SEB_{MLP} , SEB_{Dec} , and PDD models were negligible (Figure 9).

When a $+1^{\circ}\text{C}$ temperature perturbation was introduced to ERA5 data, the PDD model exhibited a substantially more negative mass balance response than the SEB model (Figure 9). Across all models, winter mass balance was less affected by the temperature perturbation compared to summer mass balance. These results are further summarized in box plots for all temperature perturbation scenarios, illustrating the sensitivity of annual, winter, and summer mass balance separately (Figure 10).

The results for winter mass balance sensitivity to a 2°C temperature increase show that the SEB models have a median $\frac{\partial B}{\partial T}$ value of $-0.03 \text{ m w.e. yr}^{-1} \text{ }^{\circ}\text{C}^{-1}$, whereas the PDD model shows a higher sensitivity with a median value of $-0.11 \text{ m w.e. yr}^{-1} \text{ }^{\circ}\text{C}^{-1}$ (Figure 10). In contrast, the summer mass balance exhibits

Fig. 10: Box plots of the temperature sensitivity of mass balance ($\frac{\partial B}{\partial T}$) for annual (a), summer (b), and winter (c) mass balance across the 23 study glaciers under different temperature perturbations. The sensitivity $\frac{\partial B}{\partial T}$ is determined by comparing the mass balance timeseries from the original model runs (SEB_{MLP}, SEB_{Dec}, and PDD) to those from model runs with applied temperature perturbations. Results are shown for the SEB models with calibrated precipitation and albedo correction factors, and for the PDD model with calibrated precipitation and melt factors. Solid horizontal lines indicate the median values, while dashed lines represent the mean values of $\frac{\partial B}{\partial T}$. The median value is labeled above the top whisker of each distribution. Box edges denote the first and third quartiles, and whiskers represent the 5th and 95th percentiles.

Fig. 11: Mass balance sensitivity to a 1°C^{-1} temperature increase ($\frac{\partial B}{\partial T}$) derived from the PDD model compared to the SEB_{MLP} model (a) and the SEB_{Dec} model (b). Results are shown for all 23 study glaciers, with points colored by subregion. SEB models incorporate calibrated precipitation factors and albedo corrections, while the PDD model includes calibrated melt and precipitation factors.

greater sensitivity to temperature increases across all models. For the SEB_{MLP} model, the median $\frac{\partial B}{\partial T}$ values range from -0.20 to -0.21 m w.e. $\text{yr}^{-1} \text{ }^{\circ}\text{C}^{-1}$, while for the SEB_{Dec} model, they range from -0.28 to -0.29 m w.e. $\text{yr}^{-1} \text{ }^{\circ}\text{C}^{-1}$. The PDD model, however, shows significantly higher sensitivity, with median values ranging from -0.50 to -0.53 m w.e. $\text{yr}^{-1} \text{ }^{\circ}\text{C}^{-1}$ across the temperature perturbations. In all cases, the PDD model's sensitivity is significantly different from that of the SEB models (as confirmed by an independent two-sample t-test) and exhibits a wider spread of $\frac{\partial B}{\partial T}$ values compared to the SEB models.

At the scale of individual glaciers, the PDD model consistently produces higher annual mass balance sensitivities to a $+1^{\circ}\text{C}$ temperature change compared to the SEB models (Figure 11). The annual mass balance sensitivities across the 23 glaciers are significantly correlated between the PDD and SEB_{MLP} models ($r = 0.54$) and between the PDD and SEB_{Dec} models ($r = 0.61$). While the SEB models maintains relatively stable sensitivities across different magnitudes of temperature perturbations, the PDD model exhibits substantial variability. Specifically, the lowest sensitivity for all models, but particularly the PDD model, occurs under a -2°C perturbation and increases progressively with temperature perturbations up to $+2^{\circ}\text{C}$. This effect is most pronounced for glaciers with a large difference between the snow melt factor (f_{snow}) and ice melt factor (f_{ice}), where the PDD model shows the most significant shifts in mass balance sensitivity, especially for winter mass balance.

The mass balance sensitivity to temperature change varies significantly across the four subregions. The Arctic region shows the lowest mean annual $\frac{\partial B}{\partial T}$ values for all three models (PDD, SEB_{MLP} , and SEB_{Dec}), with values ranging from -0.02 to -0.11 m w.e. $\text{yr}^{-1} \text{ } ^\circ\text{C}^{-1}$ for the SEB_{MLP} model and -0.03 to -0.18 m w.e. $\text{yr}^{-1} \text{ } ^\circ\text{C}^{-1}$ for the PDD model. In contrast, the West Coast glaciers exhibits the highest mean annual mass balance sensitivity, with $\frac{\partial B}{\partial T}$ values ranging from -0.21 to -0.61 m w.e. $\text{yr}^{-1} \text{ } ^\circ\text{C}^{-1}$ for the SEB_{MLP} model and from -0.24 to -0.76 m w.e. $\text{yr}^{-1} \text{ } ^\circ\text{C}^{-1}$ for the PDD model.

DISCUSSION

Our results demonstrate that the PDD model, calibrated with glacier-specific melt factors, initially outperforms the SEB models when no additional parameters are adjusted. Introducing calibrated precipitation correction factors (d_p and k_p) to the models leads to significant improvements in simulating winter and net mass balance. However, even with these corrections, the SEB models underperform relative to the PDD model. It is only after incorporating a calibrated albedo correction in the SEB models that their performance matches or slightly exceeds that of the PDD model. The median NRMSE values for net mass balance are 12.3% for SEB_{MLP} and 11.6% for SEB_{Dec} , compared to 12.8% for the PDD model. These findings highlight the pivotal role of accurately simulating glacier albedo, as net shortwave radiation is the dominant contributor to melt energy at our sites, accounting for approximately 70% of the total melt energy.

A key focus of this study is the use of MODIS-derived albedo data to train the MLP model and calibrate the logistic decay model. However, for many study glaciers, the MODIS-derived albedo exhibits relatively low summer values, including unrealistic estimates (below 0.5) in glacier accumulation zones. These underestimated albedo values, when used as inputs in the SEB models, lead to an overestimation of melt and consequently higher NRMSE values. The calibrated albedo correction mitigates this bias by increasing the albedo by an average of 0.11, significantly improving the model's fit to observed mass balance. Introducing a calibrated wind speed correction, in addition to the precipitation and albedo corrections, yields only marginal enhancements in model performance.

Our analysis also reveals a consistent tendency for MODIS to underestimate albedo when compared with in-situ albedo observations from several glacier sites. This finding is consistent with other studies that have compared MODIS-derived albedo with in-situ snow albedo measurements across various mountainous regions, where results often vary widely (e.g., Sorman and others, 2007; Calleja and others, 2019; Stroeve

and others, 2006; Williamson and others, 2016). These discrepancies highlight the site-specific performance of MODIS in assessing snow and ice albedo. According to the MODIS user guide, snow albedo estimates are generally within 10% of in-situ observations under optimal conditions. However, the guide also notes that errors can be significantly higher over steep mountainous terrain (Hall and Riggs, 2021). Our findings underscore the need for applying albedo corrections to MODIS data or to models trained on this data, particularly given the variability in MODIS performance across different regions and local conditions. We observe that the greater the underestimation of summer albedo in the glacier accumulation zone, the larger the required albedo correction.

Additional factors to consider when using MODIS-derived albedo data include atmospheric conditions, shadowing effects, and surface type changes. Shadowing from local terrain and variations in surface type primarily impacts grid cells near glacier edges, introducing potential errors in the observed MODIS albedo (Williamson and others, 2020). Furthermore, extensive cloud cover often coincides with snowfall events, preventing albedo measurements from being recorded. Consequently, MODIS observations may miss the rapid, short-term increases in albedo over summer following fresh snowfall on exposed glacier surfaces. To address this discrepancy, we adjust the albedo models by increasing albedo to fresh snow values on days with modeled snowfall during summer. This adjustment has a minimal impact on the simulated mass balance, resulting in an average difference of less than 2% in the NRMSE of summer mass balance across our study glaciers. While the frequency of fresh snowfall events has been shown to significantly influence summer albedo and, consequently, summer mass balance (Oerlemans and Klok, 2004), our findings suggest that accurately resolving the average summer albedo across the glacier elevation range is more critical than capturing short-term albedo variability from individual snowfall events.

While the albedo data simulated by both models MLP and logistic decay ultimately result in similar simulations of summer mass balance, there are critical differences between these models that can affect their suitability for regional and global-scale applications. Notably, without the calibrated albedo correction, the SEB_{MLP} model showed a significantly better fit to mass balance observations compared to the SEB_{Dec} model, with a median summer NRMSE of 24.6%, compared to 44.9% for the SEB_{Dec} . One key difference lies in their sensitivity to temperature: the logistic decay albedo model relies solely on temperature as an input, making it more responsive to temperature perturbations, as evidenced by our mass balance sensitivity tests. In contrast, the MLP albedo model incorporates additional features and is designed to be elevation-specific within subregions rather than glacier-specific, potentially enhancing its spatial transferability. Moreover,

while we utilized a neural network (MLP) approach for regional albedo simulation in this study, this is just one of many possible machine learning strategies. More advanced architectures, such as recurrent neural networks (RNNs) or long short-term memory (LSTM) models, have yet to be explored and could offer further improvements in simulating albedo on regional scales.

Since our SEB model is elevation-dependent, it does not explicitly account for the effects of glacier slope and aspect on incoming shortwave radiation. In contrast, a fully distributed SEB model (e.g., Hock and Holmgren, 2005; Hill and others, 2020) that incorporates these topographic factors would likely offer a more accurate depiction of the net radiative balance but would introduce substantially greater model complexity. Instead, the calibrated albedo correction we apply to help mitigate biases in ERA5 incoming radiation (Draeger and others, 2024) and inaccuracies in MODIS-derived albedo data, also indirectly compensates for the combined effects of glacier slope and aspect on shortwave radiation.

Recognizing the limitations of sub-annual mass balance observations, particularly from remote sensing data such as those provided in the Hugonnet and others (2021) database, we explore the effects of using seasonal versus annual mass balance data for model calibration. Our analysis shows that calibrating both the PDD and SEB models using annual mass balance observations results in a similarly good fit to annual mass balance as when seasonal observations are used. However, relying solely on annual data for calibration introduces notable biases in the seasonal components; overestimated (or underestimated) winter accumulation is compensated by overestimated (or underestimated) summer melt, effectively masking the errors in net mass balance simulation. In contrast, calibrating the models with seasonal mass balance data helps to optimize the accumulation and melt processes separately, leading to more accurate simulations of both winter and summer mass balance. These findings underscore the critical importance of in-situ seasonal mass balance measurements, especially as the field increasingly relies on remote sensing data (Rounce and others, 2023; Zekollari and others, 2024; Compagno and others, 2022; Hanus and others, 2024), which currently have limited accuracy at annual and sub-annual scales. Accurate seasonal calibration ensures that simple models are better tuned to represent the distinct processes driving glacier mass changes, a crucial consideration for reliable mass balance projections in future climate scenarios. Our results highlight the importance of seasonal-scale calibration while recognizing the value of glacier-specific calibration based on long-term trends, as implemented in Rounce and others (2023) and Zekollari and others (2024). These findings highlight the need for further improvements in global-scale glacier modeling.

We attribute the strong performance of the SEB models (with calibrated precipitation and albedo

corrections) in fitting mass balance observations to the use of ERA5 reanalysis data, which reasonably captures local meteorological conditions at multiple glacier sites in western Canada (Draeger and others, 2024). The only bias correction applied to the ERA5 input data is a lapse rate adjustment for 2 m air temperature. This correction proved to be crucial: omitting it increased the median NRMSE by 24.6% for the SEB_{MLP} model and by 49.9% for the PDD model using the same melt factors. The PDD model shows a greater sensitivity to the lapse rate correction because temperature directly drives its melt calculations. In contrast, the SEB model is less affected by temperature biases; near-surface temperature primarily influences turbulent heat fluxes, which contribute less than 30% to seasonal melt energy at the study sites. Additionally, temperature had a minimal effect on simulated albedo evolution in the MLP model, resulting in an average difference of 0.08 in summer albedo across the study glaciers.

The higher sensitivity of the PDD model to temperature changes is further evident in our sensitivity tests, where temperature is uniformly adjusted across the entire observational period. The PDD model consistently exhibits a greater response, with median summer $\frac{\partial B}{\partial T}$ values around -0.52 m w.e. $\text{yr}^{-1} \text{ } ^\circ\text{C}^{-1}$, compared to approximately -0.20 m w.e. $\text{yr}^{-1} \text{ } ^\circ\text{C}^{-1}$ for the SEB_{MLP} model. The spatial pattern, with the lowest sensitivities observed in Arctic glaciers and the highest in West Coast glaciers, aligns with previous findings showing greater sensitivities for maritime glaciers and lower sensitivities for Arctic glaciers (Radi and Hock, 2011). Additionally, our sensitivity values fall within the range reported for static mass balance sensitivities across various glaciers globally (e.g., Woul and Hock, 2005). The temperature sensitivity of the PDD model is also influenced by the difference between f_{snow} and f_{ice} ; glaciers with larger differences between these melt factors show a more pronounced, non-uniform response to varying magnitudes of temperature changes. Given that many large-scale glacier models primarily rely on projections of temperature and precipitation (e.g., Marzeion and others, 2020; Rounce and others, 2023), the differing temperature sensitivities of the PDD and SEB models could lead to substantially different predictions of future glacier mass changes in climate projections where temperature is the only changing variable. Unlike the PDD model, which responds solely to temperature, the SEB model accounts for multiple climatic variables, providing additional degrees of freedom in its response. Consequently, it exhibits lower sensitivity to temperature alone. While these findings do not invalidate the use of PDD models for glacier melt projections, they highlight the limitations of relying exclusively on temperature as the sole meteorological input for melt calculations.

Given our focus on evaluating a relatively simple SEB model, we deliberately excluded components that

would require more complex and calibration-intensive frameworks. This choice was guided by the principle that increasing model complexity does not necessarily yield better regional-scale glacier simulations (Temme and others, 2023) and can complicate the diagnostic assessment of model performance. However, some of our simplified assumptions may not be universally applicable across glaciers from different climatic regimes. For instance, processes such as meltwater refreezing within the snowpack and heat flux into the glacier surface are more critical for SEB and mass balance in Arctic glaciers compared to those in mid-latitudes (Woodward and others, 1997; Hock, 2005). Similarly, rain heat flux has been shown to contribute significantly (up to 20%) to daily melt energy during extreme rainfall events on some glaciers in the Rockies (Fitzpatrick and others, 2017). However, the uncertainty in the models used to estimate rain heat flux remains relatively high (Fitzpatrick and others, 2017).

Despite its simplicity, our SEB model enhanced with precipitation and albedo corrections achieves a “goodness of fit” comparable to or better than those reported in studies using more complex, heavily parameterized SEB models, such as Shannon and others (2019), which are part of global-scale simulations. Specifically, our model consistently produces lower RMSE values (0.20-0.64 m w.e. yr⁻¹ compared to 0.96-1.73 m w.e. yr⁻¹ in Shannon and others (2019)), higher correlation coefficients, lower mean bias errors, and Nash-Sutcliffe efficiency coefficients closer to one metrics evaluated using approximately the same in-situ mass balance observations across Canadian glaciers. While differences in model complexity, climate forcing, and downscaling approaches exist between our study and Shannon and others (2019), the superior fit of our model to in-situ mass balance observations underscores its potential for large-scale mass balance modeling.

CONCLUSIONS

Large-scale glacier mass balance models commonly rely on PDD melt models, which, while widely used, have well-recognized limitations. This study investigated the ability of a relatively simple, elevation-dependent SEB model, requiring minimal downscaling of climate input data, to replicate seasonal and annual mass balance observations. We compared mass balance models utilizing SEB and PDD melt components against observations from 23 glaciers across Canada, spanning diverse climatic regimes. The models were driven by ERA5 data, with downscaling limited to lapse-rate corrections of 2 m air temperature for better local representation. We initially assessed SEB models without calibrated parameters and then introduced calibrated correction factors for precipitation, albedo, and wind to enhance the alignment between modeled

and observed mass balance. A central aspect of this study was the improved representation of glacier albedo within the SEB model, contrasting a simple empirical albedo approach (SEB_{Dec}) with a machine learning-based method trained on MODIS-derived albedo data (SEB_{MLP}). Our principal conclusions are:

The accuracy of ERA5 input data, especially the lapse rate correction for 2 m temperature, were crucial for the performance of both SEB and PDD models. Initially, the PDD-based mass balance model outperformed the SEB-based model when no calibrated parameters or only precipitation correction factors were applied, emphasizing the efficiency of glacier-specific melt factor calibration in PDD models.

Incorporating a calibrated albedo correction significantly improved the performance of SEB models, bringing them on par with or marginally better than the PDD model in simulating glacier mass balance. The calibrated SEB_{MLP} model outperformed the SEB_{Dec} model in most scenarios, demonstrating the importance of advanced albedo parameterization.

MODIS-derived albedo data tended to underestimate summer albedo at the study glaciers, particularly in the accumulation zones, resulting in overestimated melt. Calibrated albedo corrections in the SEB models significantly improved their alignment with observed mass balance, underscoring the importance of calibrating albedo or, more broadly, net shortwave radiative fluxes to enhance model accuracy.

Calibration using seasonal mass balance data, as opposed to annual mass balance, led to more accurate simulations of summer and winter mass balance. This finding underscores the importance of using detailed seasonal data for calibrating mass balance models to avoid compensating biases in accumulation and melt during the calibration process.

Although the calibrated SEB and PDD models produced similar reconstructions of glacier mass balance over the 1980–2019 period in terms of long-term trends and interannual variability, the PDD model exhibited twice the sensitivity to temperature changes when temperature was the only altered climate variable. Since climate change affects multiple climatic variables, these findings highlight the limitations of the PDD model in relying solely on temperature as the primary input for melt calculations.

While a more complex, fully distributed SEB model incorporating glacier slope and aspect could yield more accurate simulations, the simpler, elevation-dependent SEB model assessed in this study provided

a good balance between computational efficiency and model performance. This suggests that for regional applications, simplified models like the SEB model can be effective without unnecessary complexity, as long as key parameters such as albedo are properly calibrated.

SUPPLEMENTARY MATERIAL

The supplementary material for this article can be found at XXX.

DATA

All data used in this study was sourced from publicly available data repositories (see Methods for further details).

ACKNOWLEDGEMENTS

We thank the two anonymous reviewers, as well as the editors David Rounce and Hester Jiskoot, for their helpful comments and feedback throughout the review process. We acknowledge funding from the Natural Sciences and Engineering Research Council of Canada Discovery Grant and the Polar Continental Shelf Program to Valentina Radi and the General Sir John Monash Foundation scholarship to Hannah Phelps.

REFERENCES

- Alduchov O and Eskridge R (1996) Improved Magnus form approximation of saturation vapor pressure. Technical report, Department of Commerce, Asheville, NC (United States) (doi: 10.2172/548871)
- Anderson S and Radi V (2020) Identification of local water resource vulnerability to rapid deglaciation in Alberta. *Nature Climate Change*, **10**, 933938 (doi: 10.1038/s41558-020-0863-4)
- Bolibar J, Rabatel A, Gouttevin I, Galiez C, Condom T and Sauquet E (2020) Deep learning applied to glacier evolution modelling. *The Cryosphere*, **14**(2), 565–584 (doi: 10.5194/tc-14-565-2020)
- Box J, Fettweis X, Stroeve J, Tedesco M, Hall D and Steffen K (2012) Greenland ice sheet albedo feedback: thermodynamics and atmospheric drivers. *The Cryosphere*, **6**(4), 821–839 (doi: 10.5194/tc-6-821-2012)
- Calleja JF, Corbea-Pérez A, Fernández S, Recondo C, Peón J and de Pablo M (2019) Snow albedo seasonality and trend from MODIS sensor and ground data at Johnsons Glacier, Livingston Island, maritime Antarctica. *Sensors*, **19**, 3569–3592 (doi: 10.3390/s19163569)

- Clarke G, Jarosch A, Anslow F, Radi V and Menounos B (2015) Projected deglaciation of western Canada in the twenty-first century. *Nature Geoscience*, **8**, 327–377 (doi: 10.1038/ngeo2407)
- Compagno L and 7 others (2022) Modelling supraglacial debris-cover evolution from the single-glacier to the regional scale: an application to High Mountain Asia. *The Cryosphere*, **16**(5), 1697–1718 (doi: 10.5194/tc-16-1697-2022)
- Cuffey K and Paterson W (2010) *The physics of glaciers*. Elsevier, Inc., 4th edition, ISBN 978-0-12-369461-4 (doi: 10.1016/B978-0-12-369461-4.00000-0)
- Davaze L and 6 others (2018) Monitoring glacier albedo as a proxy to derive summer and annual surface mass balances from optical remote-sensing data. *The Cryosphere*, **12**(1), 271–286 (doi: 10.5194/tc-12-271-2018)
- Draeger C, Radi V, White R and Tessema M (2024) Evaluation of reanalysis data and dynamical downscaling for surface energy balance modeling at mountain glaciers in western Canada. *The Cryosphere*, **18**, 17–42 (doi: 10.5194/tc-18-17-2024)
- Fitzpatrick N, Radi V and Menounos B (2017) Surface energy balance closure and turbulent flux parameterization on a mid-latitude mountain glacier, Purcell Mountains, Canada. *Frontiers in Earth Science*, **5**, 67 (doi: 10.3389/feart.2017.00067)
- Fitzpatrick N, Radić V and Menounos B (2019) A multi-season investigation of glacier surface roughness lengths through in situ and remote observation. *The Cryosphere*, **13**(3), 1051–1071 (doi: 10.5194/tc-13-1051-2019)
- Gascoïn S and 6 others (2017) Evaluation of MODIS albedo product over ice caps in iceland and impact of volcanic eruptions on their albedo. *Remote Sensing*, **9**(5), 399 (doi: 10.3390/rs9050399)
- Giesen R and Oerlemans J (2013) Climate-model induced differences in the 21st century global and regional glacier contributions to sea-level rise. *Climate Dynamics*, **41**, 32833300 (doi: 10.1007/s00382-013-1743-7)
- Gunnarsson A, Gardarsson S and Pálsson F (2023) Modeling of surface energy balance for Icelandic glaciers using remote-sensing albedo. *The Cryosphere*, **17**(9), 3955–3986 (doi: 10.5194/tc-17-3955-2023)
- Hall D and Riggs G (2021) MODIS/Terra snow cover daily L3 global 500m SIN grid, version 61 (doi: 10.5067/MODIS/MOD10A1.061)
- Hall D, Salomonson V and Riggs G (1995) Development of methods for mapping global snow cover using moderate resolution imaging spectroradiometer data. *Remote Sensing of Environment*, **54**, 127–140 (doi: 10.1016/0034-4257(95)00137-P)
- Hanus S, Schuster L, Burek P, Maussion F, Wada Y and Viviroli D (2024) Coupling a large-scale glacier and hydrological model (OGGM v1.5.3 and CWatM V1.08) towards an improved representation of mountain water resources in global assessments. *Geoscientific Model Development*, **17**, 5123–5144 (doi: 10.5194/gmd-17-5123-2024)

- Hersbach H and 19 others (2020) The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146**, 1999–2049 (doi: 10.1002/qj.3803)
- Hill T, Dow C, Bash E and Copland L (2020) Application of an improved surface energy balance model to two large valley glaciers in the St. Elias Mountains, Yukon. *Journal of Glaciology*, **67**, 1–16 (doi: 10.1017/jog.2020.106)
- Hirose J and Marshall S (2013) Glacier meltwater contributions and glaciometeorological regime of the Illecillewaet River Basin, British Columbia, Canada. *Atmosphere-Ocean*, **51**(4), 416435 (doi: 10.1080/07055900.2013.791614)
- Hock R (2003) Temperature index melt modelling in mountain areas. *Journal of Hydrology*, **282**, 104–115 (doi: 10.1016/S0022-1694(03)00257-9)
- Hock R (2005) Glacier melt: A review of processes and their modelling. *Progress in Physical Geography*, **27**, 1914 (doi: 10.1191/0309133305pp453ra)
- Hock R and Holmgren B (2005) A distributed surface energy-balance model for complex topography and its application to Storglaciären, Sweden. *Journal of Glaciology*, **51**, 25–36 (doi: 10.3189/172756505781829566)
- Hsieh W (2009) *Machine Learning Methods in the Environmental Sciences: Neural Networks and Kernels*. Cambridge University Press (doi: 10.1017/CBO9780511627217)
- Hugonnet R and 10 others (2021) Accelerated global glacier mass loss in the early twenty-first century. *Nature*, **592**, 726–731 (doi: 10.1038/s41586-021-03436-z)
- Huss M and Hock R (2015) A new model for global glacier change and sea-level rise. *Frontiers in Earth Sciences*, **3**, 54 (doi: 10.3389/feart.2015.00054)
- Huss M and Hock R (2018) Global-scale hydrological response to future glacier mass loss. *Nature Climate Change*, **8**, 135140 (doi: 10.1038/s41558-017-0049-x)
- Huss M, Farinotti D, Bauder A and Funk M (2008) Modelling runoff from highly glacierized alpine drainage basins in a changing climate. *Hydrological Processes*, **22**, 3888 – 3902 (doi: 10.1002/hyp.7055)
- Kingma D and Ba J (2014) Adam: A method for stochastic optimization. In *International Conference on Learning Representations* (doi: <https://doi.org/10.48550/arXiv.1412.6980>)
- MacDougall A, Wheler B and Flowers G (2011) A preliminary assessment of glacier melt-model parameter sensitivity and transferability in a dry subarctic environment. *The Cryosphere*, **5**(4), 1011–1028 (doi: 10.5194/tc-5-1011-2011)
- Marzeion B and 16 others (2020) Partitioning the uncertainty of ensemble projections of global glacier mass change. *Earth's Future*, **8**, 1–25 (doi: 10.1029/2019EF001470)

- Maussion F and 14 others (2019) The Open Global Glacier Model (OGGM) v1.1. *Geoscientific Model Development*, **12**, 909–931 (doi: 10.5194/gmd-12-909-2019)
- Nelder J and Mead R (1965) A simplex method for function minimization. *The Computer Journal*, **7**, 308–313 (doi: 10.1093/comjnl/7.4.308)
- Noël B, Berg W, Lhermitte S, Wouters B, Schaffer N and Van den Broeke M (2018) Six decades of glacial mass loss in the Canadian Arctic Archipelago. *Journal of Geophysical Research: Earth Surface*, **123**, 14301449 (doi: 10.1029/2017JF004304)
- Oerlemans J and Klok L (2004) Effect of summer snowfall on glacier mass balance. *Annals of Glaciology*, **38**, 97100 (doi: 10.3189/172756404781815158)
- Pedregosa F and 15 others (2011) Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, **12**, 2825–2830 (doi: 10.5555/1953048.2078195)
- Pellicciotti F, Brock B, Strasser U, Burlando P, Funk M and Corripio J (2005) An enhanced temperature-index glacier melt model including the shortwave radiation balance: development and testing for Haut Glacier d'Arolla, Switzerland. *Journal of Glaciology*, **51**(175), 573587 (doi: 10.3189/172756505781829124)
- Radi V and Hock R (2011) Regionally differentiated contribution of mountain glaciers and ice caps to future sea-level rise. *Nature Geoscience*, **4**, 9194 (doi: 10.1038/ngeo1052)
- Radi V and Hock R (2014) Glaciers in the earth's hydrological cycle: Assessments of glacier mass and runoff changes on global and regional scales. *Surveys in Geophysics*, **35**, 813837 (doi: 10.1007/s10712-013-9262-y)
- Rounce D, Khurana T, Short M, Hock R, Shean D and Brinkerhoff D (2020) Quantifying parameter uncertainty in a large-scale glacier evolution model using Bayesian inference: Application to high mountain Asia. *Journal of Glaciology*, **66**, 175–187 (doi: 10.1017/jog.2019.91)
- Rounce D and 12 others (2023) Global glacier change in the 21st century: Every increase in temperature matters. *Science*, **379**, 78–83 (doi: 10.1126/science.abo1324)
- Sauter T, Arndt A and Schneider C (2020) COSIPY v1.3 – an open-source coupled snowpack and ice surface energy and mass balance model. *Geoscientific Model Development*, **13**(11), 5645–5662 (doi: 10.5194/gmd-13-5645-2020)
- Schuster L, Rounce D and Maussion F (2023) Glacier projections sensitivity to temperature-index model choices and calibration strategies. *Annals of Glaciology*, **64**(92), 293308 (doi: 10.1017/aog.2023.57)
- Shannon S and 9 others (2019) Global glacier volume projections under high-end climate change scenarios. *The Cryosphere*, **13**, 325–350 (doi: 10.5194/tc-13-325-2019)

- Sharp M, Duncan A, Wolken G and Boon S (2011) Glaciodyn (Canada) IPY project: Weather station data collected on the Belcher Glacier, Devon Island Ice Cap, 2008-2009 (doi: <https://doi.org/10.5443/10933>)
- Sorman U, Akyurek Z, Sensoy A, Sorman A and Tekeli A (2007) Commentary on comparison of MODIS snow cover and albedo products with ground observations over the mountainous terrain of Turkey. *Hydrology and Earth System Sciences*, **11**(4), 1353–1360 (doi: [10.5194/hess-11-1353-2007](https://doi.org/10.5194/hess-11-1353-2007))
- Stroeve J and 6 others (2005) Tracking the Arctic's shrinking ice cover: Another extreme September minimum in 2004. *Geophysical Research Letters*, **32**(4), L04501 (doi: [10.1029/2004GL021810](https://doi.org/10.1029/2004GL021810))
- Stroeve J, Box J and Haran T (2006) Evaluation of the MODIS (MOD10A1) daily snow albedo product over the Greenland ice sheet. *Remote Sensing of Environment*, **105**, 155–171 (doi: [10.1016/j.rse.2006.06.009](https://doi.org/10.1016/j.rse.2006.06.009))
- Sun S, Cao Z, Zhu H and Zhao J (2019) A survey of optimization methods from a machine learning perspective. *IEEE Transactions on Cybernetics*, **50**(8), 3668–3681 (doi: [10.1109/TCYB.2019.2950779](https://doi.org/10.1109/TCYB.2019.2950779))
- Tarek M, Brissette F and Arsenault R (2020) Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. *Hydrology and Earth System Sciences*, **24**, 2527–2544 (doi: [10.5194/hess-24-2527-2020](https://doi.org/10.5194/hess-24-2527-2020))
- Temme F and 9 others (2023) Strategies for regional modeling of surface mass balance at the Monte Sarmiento Massif, Tierra del Fuego. *The Cryosphere*, **17**(6), 2343–2365 (doi: [10.5194/tc-17-2343-2023](https://doi.org/10.5194/tc-17-2343-2023))
- Van Rossum G and Drake F (2009) *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA, ISBN 1441412697 (doi: [10.5555/1593487](https://doi.org/10.5555/1593487))
- WGMS (2022) Fluctuations of glaciers database (doi: [10.5904/wgms-fog-2022-09](https://doi.org/10.5904/wgms-fog-2022-09))
- Williamson S, Copland L and Hik D (2016) The accuracy of satellite-derived albedo for northern alpine and glaciated land covers. *Polar Science*, **10**(3), 262–269 (doi: [10.1016/j.polar.2016.06.006](https://doi.org/10.1016/j.polar.2016.06.006))
- Williamson S, Copland L, Thomson L and Burgess D (2020) Comparing simple albedo scaling methods for estimating Arctic glacier mass balance. *Remote Sensing of Environment*, **246**, 111858 (doi: [10.1016/j.rse.2020.111858](https://doi.org/10.1016/j.rse.2020.111858))
- Woodward J, Sharp M and Arendt A (1997) The influence of superimposed-ice formation on the sensitivity of glacier mass balance to climate change. *Annals of Glaciology*, **24**, 186190 (doi: [10.3189/S0260305500012155](https://doi.org/10.3189/S0260305500012155))
- Wool MD and Hock R (2005) Static mass-balance sensitivity of arctic glaciers and ice caps using a degree-day approach. *Annals of Glaciology*, **42**, 217224 (doi: [10.3189/172756405781813096](https://doi.org/10.3189/172756405781813096))
- Xiao Y, Ke C, Fan Y, Shen X and Cai Y (2022) Estimating glacier mass balance in high mountain asia based on moderate resolution imaging spectroradiometer retrieved surface albedo from 2000 to 2020. *International Journal of Climatology*, **42**(16), 9931–9949 (doi: <https://doi.org/10.1002/joc.7873>)

Zekollari H and 10 others (2024) Twenty-first century global glacier evolution under CMIP6 scenarios and the role of glacier-specific observations. *The Cryosphere*, **18**, 5045–5066 (doi: 10.5194/tc-18-5045-2024)

Table 1: Details of glaciers used in the study

Name	RGI ID	Latitude (°)	Longitude (°)	Period of observations	Seasonal data	Glacier area (km ²)	Subregion
White	RGI60-03.04539	79.45	-90.7	1960-2017	No	40.51	Arctic
Devon Ice Cap NW	RGI60-03.02435	75.42	-83.25	1961-2016	Yes	765.44	Arctic
Tiedemann	RGI60-02.04379	51.33	-125.05	1981-1985	Yes	58.91	W Coast
Yuri	RGI60-01.05211	56.98	-130.68	1981-1984	Yes	1.62	W Coast
Alexander	RGI50-01.05355	57.1	-130.82	1979-1990	Yes	4.03	W Coast
Andrei	RGI60-01.04591	56.93	-130.97	1978-1990	Yes	71.75	W Coast
Place	RGI60-02.01104	50.42	-122.6	1965-2016	Yes	3.02	W Coast
Sentinel	RGI60-02.00296	49.89	-122.98	1966-1989	Yes	2.03	W Coast
Helm	RGI60-02.00377	49.96	-122.99	1975-2015	Yes	0.98	W Coast
Zavisha	RGI60-02.02107	50.79	-123.41	1976-1985	Yes	6.15	W Coast
Bridge	RGI60-02.02636	50.82	-123.57	1981-1985	Yes	81.98	W Coast
Bench	RGI60-02.04503	51.43	-124.92	1981-1990	Yes	9.56	W Coast
Sykora	RGI60-02.02631	50.87	-123.58	1976-1985	Yes	24.66	W Coast
Superguksoak	RGI60-04.07387	58.95	-63.78	1982-1984	No	1.26	E Coast
Abraham	RGI60-04.07381	58.94	-63.55	1982-1984	No	0.5	E Coast
Hidden	RGI60-04.07380	58.93	-63.56	1982-1984	No	0.58	E Coast
Minaret	RGI60-04.07329	58.89	-63.68	1982-1984	No	0.73	E Coast
Peyto	RGI60-02.05098	51.66	-116.56	1966-2016	Yes	9.7	Rockies
Kokanee	RGI60-02.00147	49.75	-117.14	2013-2018	Yes	1.8	Rockies
Nordic	RGI60-02.04264	51.43	-117.71	2014-2019	Yes	4.42	Rockies
Zillmer	RGI60-02.07780	52.67	-119.58	2014-2018	Yes	6.49	Rockies
Conrad	RGI60-02.02171	50.81	-116.93	2015-2019	Yes	16.9	Rockies
Illecillewaet	RGI60-02.03688	51.24	-117.43	2009-2017	No	5.94	Rockies

Table 2: Performance metrics (root-mean-square error (RMSE), mean bias error (MBE), and correlation coefficient (r)) evaluating daily MODIS-derived albedo against in-situ albedo observations at four glaciers across multiple summer seasons. The number of observations per glacier is listed in the final column. The last row summarizes metrics calculated across all sites combined.

Glacier	RMSE	MBE	r	Number of obs
Conrad	0.19	-0.15	0.73	99
Nordic	0.30	-0.28	0.40	24
Belcher	0.18	-0.14	0.69	68
POW	0.09	-0.03	0.56	31
All sites	0.19	-0.14	0.73	222

Table 3: Performance metrics for simulations of annual (bold), summer, and winter mass balance from the PDD model and two SEB models, summarized for each subregion. Results are shown for SEB models with calibrated precipitation and albedo corrections and the PDD model with calibrated precipitation and melt factors. Metrics include root-mean-square error (RMSE), mean bias error (MBE), correlation coefficient (r), and Nash-Sutcliffe efficiency (NS). Reference data are in-situ specific mass balance measurements from WGMS (WGMS, 2022). The number of observations per subregion is listed in the final column.

Model type	Region	RMSE			MBE			r			NS			Number of obs		
		(m w.e. yr ⁻¹)			(m w.e. yr ⁻¹)											
PDD	Rockies	0.66	0.60	0.28	0.12	0.02	0.03	0.92	0.91	0.85	0.83	0.80	0.69	571	377	377
	Arctic	0.29	0.18	0.06	0.08	0.15	-0.04	0.87	0.96	0.06	0.76	0.84	-10.28	984	54	54
	West Coast	0.61	0.51	0.39	0.13	0.27	-0.16	0.94	0.92	0.89	0.88	0.81	0.76	716	690	690
	East Coast	0.47			-0.06			0.84			0.71			72	0	0
SEB _{MLP}	Rockies	0.64	0.56	0.28	0.14	0.02	0.08	0.92	0.92	0.84	0.82	0.83	0.57	571	377	377
	Arctic	0.20	0.10	0.06	0.06	0.04	-0.04	0.93	0.97	0.06	0.81	0.93	-10.28	984	54	54
	West Coast	0.58	0.49	0.29	0.15	0.21	-0.08	0.94	0.92	0.91	0.87	0.80	0.82	716	690	690
	East Coast	0.62			0.03			0.70			0.19			72	0	0
SEB _{Dec}	Rockies	0.60	0.56	0.26	0.06	-0.03	0.03	0.93	0.92	0.84	0.87	0.85	0.65	571	377	377
	Arctic	0.25	0.25	0.06	0.04	0.14	-0.04	0.90	0.91	0.06	0.81	0.75	-10.28	984	54	54
	West Coast	0.62	0.53	0.29	0.13	0.22	-0.12	0.94	0.92	0.92	0.87	0.81	0.82	716	690	690
	East Coast	0.52			0.01			0.77			0.32			72	0	0