

## SUPPLEMENTARY MATERIAL

### Estimating snow melt prior to field data collection

Positive temperatures and clear skies occurred between 11–16 May 2016, which we suspect resulted in melt occurring on Glacier 13. The snow in the lower part of the ablation area of Glacier 13 was isothermal and showed clear signs of melt and metamorphosis. To estimate the total amount of melt we use a degree-day factor for melting snow (Braithwaite, 2008) and temperatures from a high-elevation site that have been scaled for the minimum elevation of Glacier 13 (2054 m). The high-elevation weather station is located close to the head of the Kaskawulsh Glacier (UTM Zone 7 565512 E 6729971 N), approximately 50 km from Glacier 13, at an elevation of 3050 m. Data for this weather station is publicly available at [https://datagarrison.com/users/300034012631040/300234060337540/temp/DataGarrison\\_Sat\\_Station\\_015.txt](https://datagarrison.com/users/300034012631040/300234060337540/temp/DataGarrison_Sat_Station_015.txt). A lapse rate of  $-6.5 \text{ K km}^{-1}$  elevation is applied to the temperature recorded at the high-elevation weather station to obtain approximate air temperature at the toe of Glacier 13. We find that a total of six days throughout the accumulation season experienced positive temperature and five of those days occurred between May 11–15. To obtain hourly melt the hourly temperatures that exceed  $0^\circ\text{C}$  are converted to equivalent days and then multiplied by a degree-day factor for melting snow equal to  $4 \text{ mm day}^{-1} \text{ K}^{-1}$  (Braithwaite, 2008). The sum of all melt is found to be  $0.006 \text{ m w.e.}$  This amount of melt is estimated for the toe of Glacier 13 so melt at higher elevations is likely smaller. Since the estimated melt is small compared to observed  $b_w$  and estimated  $B_w$  on Glacier 13 (Table 4), no corrections were made.

### Topographic parameters

Topographic parameters are easy-to-calculate proxies for physical processes, such as orographic precipitation, solar radiation effects, wind redistribution and preferential deposition. We derive all parameters (Table S2) for our study from a SPOT-5 DEM ( $40 \times 40 \text{ m}$ ) (Korona and others, 2009). Two DEMs are stitched together to cover the Donjek Range. An iterative 3D-coregistration algorithm (?) is used to correct the horizontal ( $\sim 2 \text{ m E}$ ,  $\sim 4 \text{ m N}$ ) and vertical ( $5.4 \text{ m}$ ) discrepancy between the two DEMs before stitching. See Pulwinski (2017) for details regarding DEM stitching. See Pulwinski (2017) for full details on topographic parameter calculation.

### DEM smoothing

Visual inspection of the curvature fields calculated using the DEM indicated that the spatial patterns of curvature were noisy and did not vary smoothly. Olaya (2009) states that the curvature calculation is sensitive to noisy data and a smoothing filter must often be applied to the DEM prior to calculation. Curvature, as well as slope, aspect and northness, are all sensitive to noise because calculating these parameters involves calculating the first and second derivatives of the elevation which are highly dependent on the size of the DEM cell. To minimize the effect of noise on these four parameters, a smoothing filter was applied to the DEM and this smoothed DEM was used to calculate curvature, slope, aspect and northness. The unsmoothed DEM was used to determine elevation and  $S_x$  because these parameters do not depend on a topographic length scale and their values are not as sensitive to the size of the DEM gridcell.

To choose a smoothing algorithm and window size, we applied a number of smoothing algorithms and chose the combination that resulted in the highest correlation between topographic parameters and point-scale winter balance values. Window sizes of  $3 \times 3$ ,  $5 \times 5$ ,  $7 \times 7$  and  $9 \times 9$  gridcells were used. For all sizes, inverse-distance weighted smoothing and Gaussian smoothing were poorly correlated with  $B_w$ . The smoothing algorithm that used  $7 \times 7$  window resulted in the highest correlation between curvature (second derivative) and  $B_w$  as well as slope (first derivative) and  $B_w$ . The window size that produced the highest correlation of  $B_w$  values and curvature for each glacier differed, but for all  $B_w$  values taken together, the  $7 \times 7$  window resulted in the highest correlation. For slope, all three glaciers showed the highest correlation with a  $7 \times 7$  smoothing window, but not when combined. To maintain consistency between parameters, the  $7 \times 7$  smoothing window was chosen and applied to the DEM for calculation of curvature ( $\kappa$ ), slope ( $m$ ), aspect ( $\alpha$ ) and “northness” ( $N$ ).

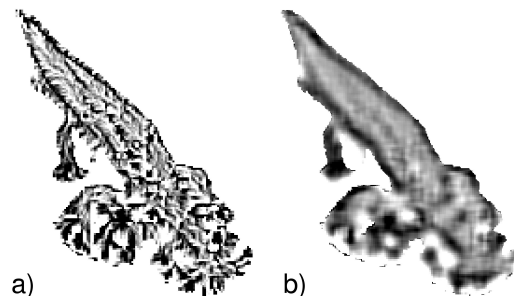


Fig. S1: (a) Curvature found using the original DEM. (b) Curvature found using the smoothed ( $7 \times 7$  window moving average) DEM.

Table S1: Values of azimuth ( $A$ ) and maximum search distance ( $d_{\max}$ ), that correspond to the  $Sx$  that had the highest absolute correlation to observed  $B_w$ .

	$A$ ( $^{\circ}$ from North)	$d_{\max}$ (m)	Correlation Coefficient
Glacier 4	85	300	-0.26
Glacier 2	330	300	0.56
Glacier 13	280	200	0.28

### Wind redistribution parameter

$Sx$  represents wind exposure/shelter and requires the identification of a cell within a certain angle and distance from the cell of interest that has the greatest upward slope relative to the cell of interest (Winstral and others, 2002). The identified cell is referred to as the maximum upwind slope. Negative  $Sx$  values represent exposure relative to the shelter-defining pixel, which means that the cell of interest is higher than the cell with greatest upwind slope. Conversely, positive values represent sheltered cells. To determine  $Sx$  values, the following equation is used

$$Sx_{A,d_{\max}}(x_i, y_i) = \max \left[ \tan^{-1} \left( \frac{z(x_v, y_v) - z(x_i, y_i)}{[(x_v - x_i)^2 + (y_v - y_i)^2]^{1/2}} \right) \right], \quad (1)$$

where  $A$  is the azimuth of the search direction,  $(x_i, y_i)$  are the coordinates of the cell of interest and  $(x_v, y_v)$  are the set of all cell coordinates located along the search vector defined relative to  $(x_i, y_i)$  and by the azimuth ( $A$ ) and maximum search distance ( $d_{\max}$ ). Code for this calculation was provided by Adam Winstral (2016, personal communication). As done by McGrath and others (2015), we compute  $Sx$  at  $5^{\circ}$  azimuth increments for  $d_{\max}$  distances of 100, 200 and 300 m. These values are then correlated (Pearson correlation) with observed values of point-scale winter balance. The  $Sx$  values from the combination of azimuth and  $d_{\max}$  input values that have the highest correlation are used for subsequent analysis (Table S1). The code for calculating  $Sx$  requires a UTM raster formatted to ASCII in ArcGIS.

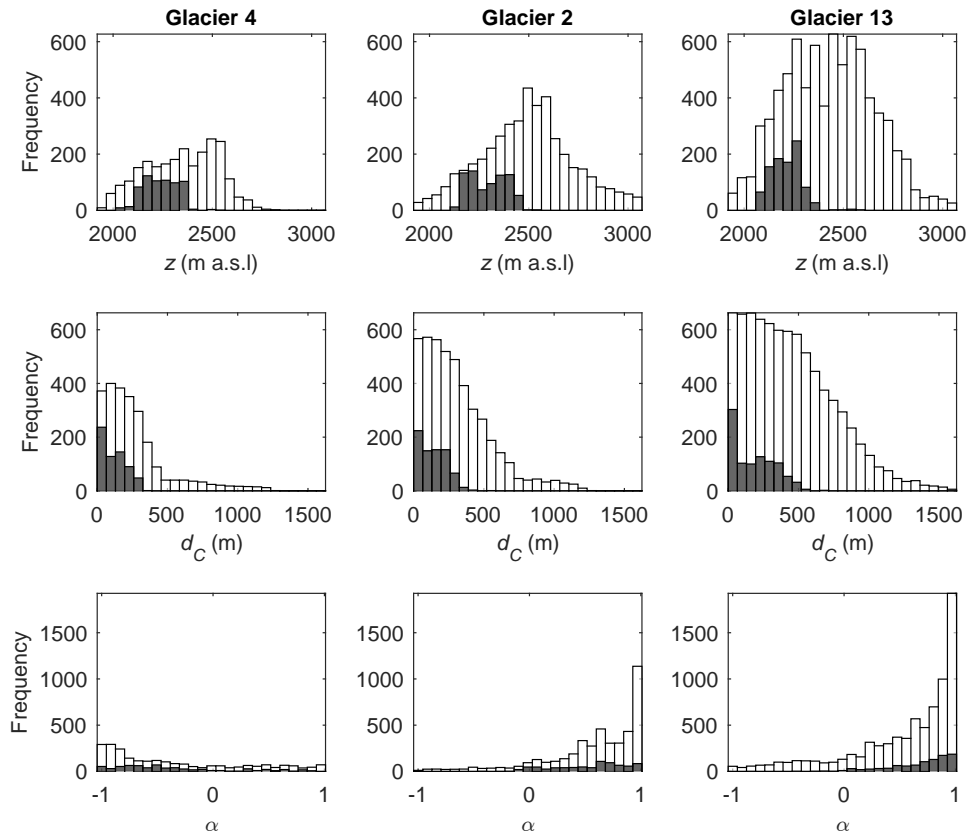


Fig. S2: Distribution of sampled (gray bars) and all (white bars) topographic parameters for over Glacier 4 (left column), Glacier 2 (middle column) and Glacier 13 (right column). From top to bottom, topographic parameters are elevation ( $z$ ), distance from centreline ( $d_C$ ), aspect ( $\alpha$ ), slope ( $m$ ), northness ( $N$ ), mean curvature ( $\kappa$ ), and wind redistribution ( $Sx$ ).

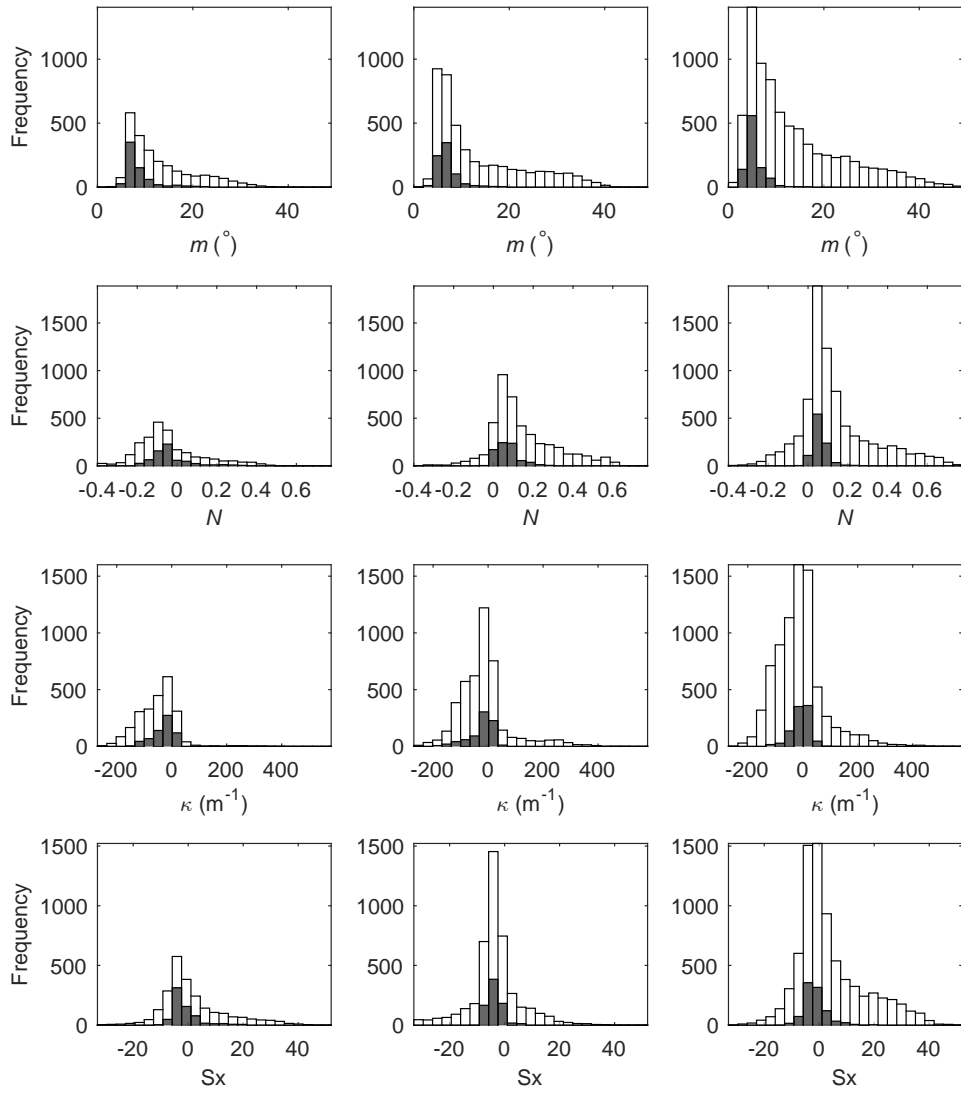


Fig. S2 (Cont.): Distribution of sampled (gray bars) and all (white bars) topographic parameters for Glacier 4 (left column), Glacier 2 (middle column) and Glacier 13 (right column). From top to bottom, topographic parameters are elevation ( $z$ ), distance from centreline ( $d_C$ ), aspect ( $\alpha$ ), slope ( $m$ ), northness ( $N$ ), mean curvature ( $\kappa$ ), and wind redistribution ( $S_x$ ).

Table S2: Description of topographic parameters used in the linear regression.

Topographic parameter	Definition	Calculation method	Notes	Source
<b>Elevation (<math>z</math>)</b>	Height above mean sea level	Values taken directly from DEM		
<b>Distance from centreline (<math>d_C</math>)</b>	Linear distance from user-defined glacier centreline	Minimum distance between the Easting and Northing of the northwest corner of each grid cell and a manually defined centreline		
<b>Slope (<math>m</math>)</b>	Angle between a plane tangential to the surface (gradient) and the horizontal	<code>r.slope.aspect</code> module in GRASS GIS software run through QGIS		Mitášová and Hofierka (1993); Hofierka and others (2009); Olaya (2009)
<b>Aspect (<math>\alpha</math>)</b>	Dip direction of the slope	<code>r.slope.aspect</code> module in GRASS GIS software run through QGIS	$\sin(\alpha)$ , a linear quantity describing a slope as north/south facing, is used in the regression	Mitášová and Hofierka (1993); Hofierka and others (2009); Olaya (2009)
<b>Mean curvature (<math>\kappa</math>)</b>	Average of profile (direction of the surface gradient) and tangential (direction of the contour tangent) curvature	<code>r.slope.aspect</code> module in GRASS GIS software run through QGIS	(+) mean-concave terrain and (-) mean-convex terrain	Mitášová and Hofierka (1993); Hofierka and others (2009); Olaya (2009)
<b>“Northness” (<math>N</math>)</b>	A value of $-1$ represents a vertical, south facing slope, a value of $+1$ represents a vertical, north facing slope, and a flat surface yields $0$	Product of the cosine of aspect and sine of slope		(Molotch and others, 2005)
<b>Wind exposure/shelter parameter (<math>S_x</math>)</b>	Proxy for snow deposition due to wind redistribution	Executable obtained from Adam Winstral that follows the procedure outlined in Winstral and others (2002)	Calculation based on selecting a cell within a certain angle and distance from the cell of interest that has the greatest upward slope relative to the cell of interest	(Winstral and others, 2002)

## Additional results

Table S3: Snow density values used for density assignment methods. Density values derived from snow pit (SP) densities and Federal Sampler (FS) densities. Four interpolation methods are chosen: (1) using a mean snow density for all three glaciers (S1 or F1), (2) using a mean density for each glacier (S2 or F2), (3) using a regression between density and elevation (S3 or F3), and (4) inverse-distance weighted mean density (not shown, S4 or F4). Standard deviation (STD) is given for S1/F1 and S2/F2 values and  $R^2$  values are given for density–elevation regressions (S3/F3).

	SP-derived density ( $\text{kg m}^{-3}$ )		FS-derived density ( $\text{kg m}^{-3}$ )		
	Mean	STD or $R^2$	Mean	STD or $R^2$	
<b>S1 or F1</b>	342	26	318	42	
<b>S2 or F2</b>	G4	348	13	355	18
	G2	333	26	286	34
	G13	349	38	316	41
<b>S3 or F3</b>	G4	$0.03z + 274$	0.16	$-0.16z + 714$	0.53
	G2	$-0.14z + 659$	0.75	$0.24z - 282$	0.72
	G13	$-0.20z + 802$	>0.99	$0.12z + 33$	0.21

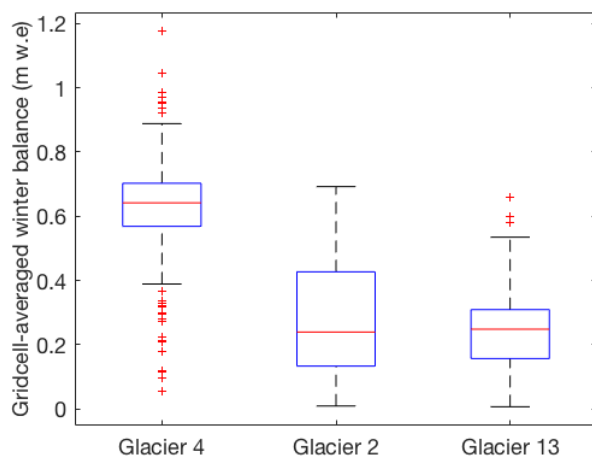


Fig. S4: Boxplot of gridcell-averaged winter balance on three study glaciers. Red line indicates median, blue box shows first quartiles, bars indicate minimum and maximum values (excluding outliers) and red crosses show outliers, which are defined as being outside of the range of 1.5 times the quartiles (approximately  $\pm 2.7\sigma$ ).

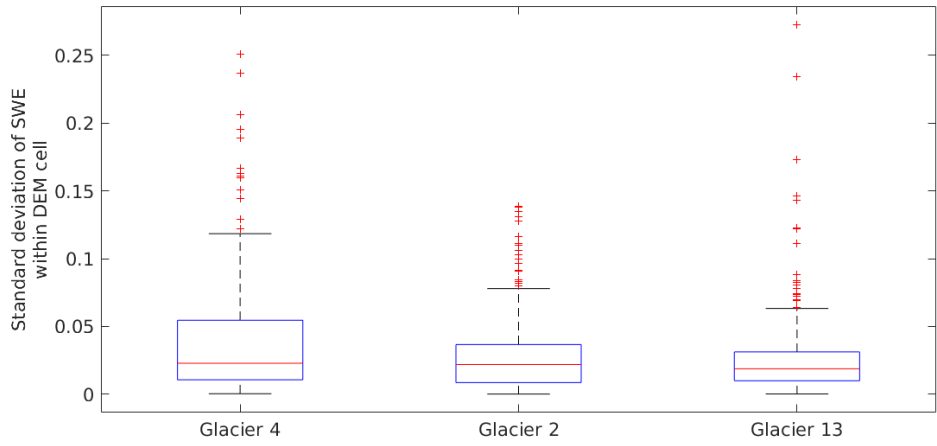


Fig. S5: Boxplot of the standard deviation of measured winter balance in one DEM gridcell. Red line indicates median, blue box shows first quartiles, bars indicate minimum and maximum values (excluding outliers) and red crosses show outliers, which are defined as being outside of the range of 1.5 times the quartiles (approximately  $\pm 2.7\sigma$ ).

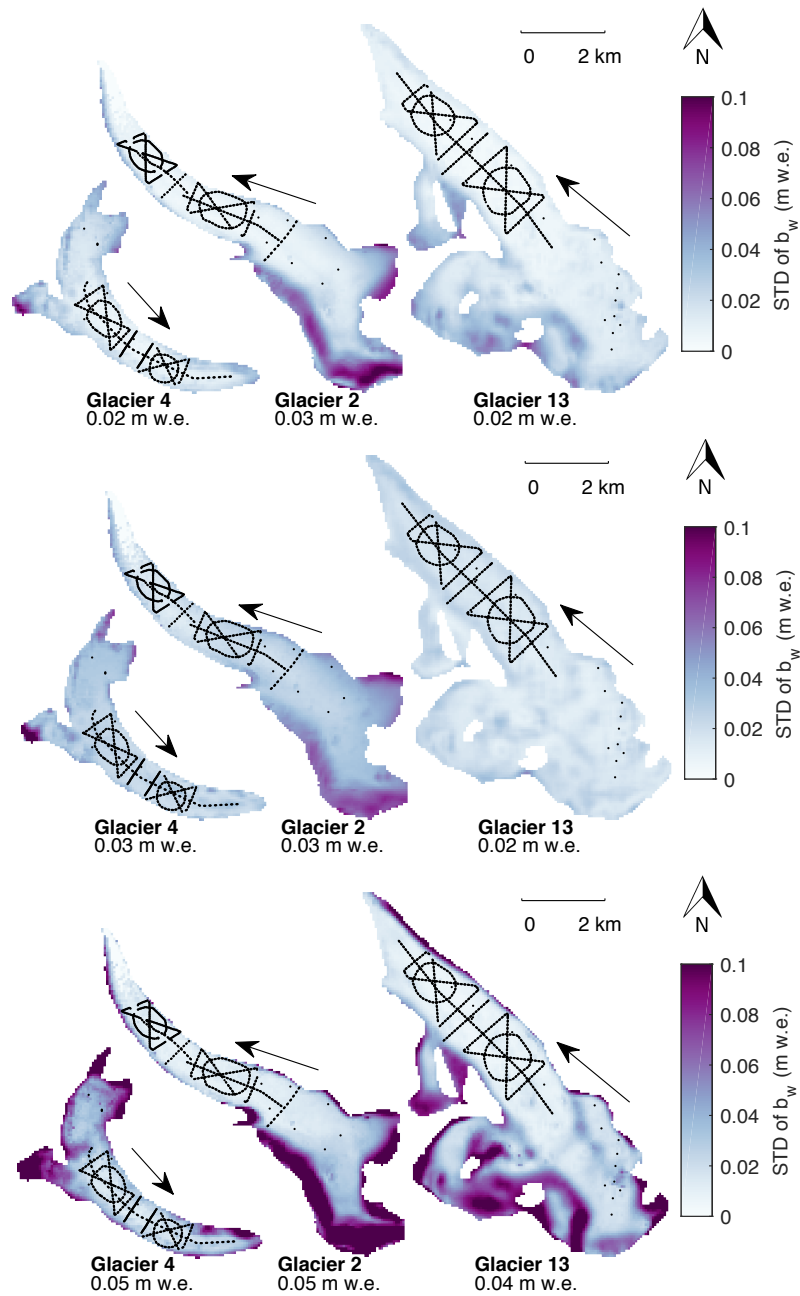


Fig. S6: Standard deviation (STD) of distributed winter balance ( $b_w$ ) arising from (top row) gridcell-scale variability, (middle row) method of density assignment and (bottom row) interpolation uncertainty found using linear regression. Ice-flow directions are indicated by arrows.



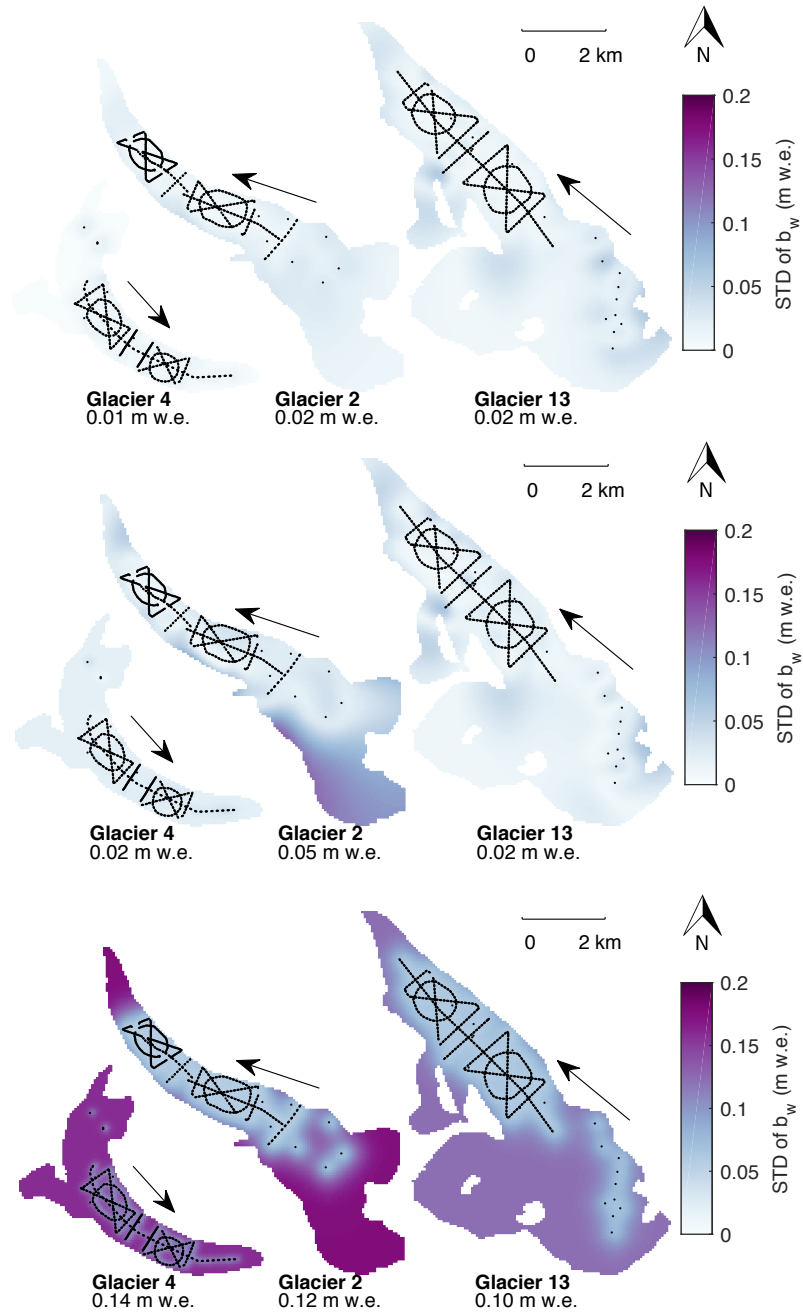


Fig. S7: Standard deviation (STD) of distributed winter balance ( $b_w$ ) arising from (top row) gridcell-scale variability, (middle row) method of density assignment and (bottom row) interpolation uncertainty found using ordinary kriging. Ice-flow directions are indicated by arrows.

## REGRESSION KRIGING

### Background

Regression kriging (RK) estimates values between measurement locations by combining a regression estimate with a kriged estimate of the regression residuals (Hengl and others, 2007). First, the regression estimate is determined using auxiliary variables (e.g. topographic parameters). Then, OK is used to interpolate regression residuals, which have an assumed mean of zero. The two surface estimates are then added. The final estimate can be written as

$$\hat{z}(x_0) = \hat{m}(x_0) + \hat{e}(x_0) \quad (2)$$

$$= \sum_{k=0}^p \hat{\beta}_k \cdot q_k(x_0) + \sum_{i=1}^n \lambda_i \cdot e(x_i), \quad (3)$$

where  $\hat{m}(x_0)$  is the regression estimate and  $\hat{e}(x_0)$  is the interpolated residual,  $\hat{\beta}_k$  are the estimated regression coefficients,  $q_k(x_0)$  are the regressors,  $p$  is the number of regressors,  $\lambda_i$  are the kriging weights for the residuals and  $e(x_i)$  is the residual at  $x_i$ .

RK can be thought of as an intermediate between pure kriging (no auxiliary variables) and pure regression (small residuals) and can be more strongly skewed to either end-member based on the strength of the regression correlation (Hengl and others, 2007). Separating the trend analysis and kriging steps has the advantage of being able to test interpolation methods that go beyond a basic linear trend. Kriging combined with regression has been found to produce better estimates of spatial fields when compared to OK and co-kriging (e.g. Knotters and others, 1995).

### Methods

Distributed  $b_w$  values are first estimated with a LR of gridcell-averaged  $b_w$  on topographic parameters, as described in the manuscript. Second, the residuals at each measurement location are calculated and the distributed residuals are found using OK, as described above. The LR and OK-estimated residuals are then added together to obtain the final distributed  $b_w$ .

### Results

The range of residual values is highest on Glacier 4 and lowest on Glacier 13 (Figure S8). Extreme values are located in the accumulation area of Glacier 4 with both strongly negative and strongly positive residuals located within a kilometre of each other. The low sampling density in the accumulation area biases the interpolation of residuals to fit the over- and under-estimated values of  $B_w$  at the two uppermost sampling locations. Residuals show less variation on Glacier 2, although relatively large residuals of approximately  $\pm 0.4$  m.w.e. are present in the upper ablation area along the ice margins. Glacier 13 has the smallest range of residuals. The glacier-wide value of distributed residuals is positive for Glacier 4, indicating that the distributed residuals will increase the overall estimate of  $B_w$ . Conversely, the glacier-wide residual for Glacier 2 is negative and will decrease the estimated  $B_w$ .

Spatial patterns in estimated  $B_w$  found using RK are similar to those found with both OK and regression (Figure S9) for Glaciers 2 and 13. The RK estimate is closer to the linear regression for Glaciers 2 and 13 due to the relatively high explanatory power of the regressions for these glaciers. The distributed  $B_w$  on Glacier 4 found using RK is somewhat uniform in the central region of the glacier, but the accumulation area has considerably larger estimates of  $B_w$  compared to OK and regression. The spatial patterns of distributed  $B_w$  on Glacier 4 more strongly resemble those of the distributed residuals (Figure S8). This supports the idea that  $B_w$  estimates are more strongly affected by the kriged residuals because the regression had low explanatory power. RK applied to Glacier 4 is therefore closer to OK. The marked alteration of the regression estimate by the kriged residuals arises from the large and opposing residual values at the upper-most measurement locations. Higher  $B_w$  values in the accumulation area result in a much higher estimate of glacier-wide  $B_w$ . The large effect of two measurement locations on the glacier-wide  $B_w$  highlights the over representation of individual data points in under-sampled regions.

Both OK and RK produce a gradient in glacier-wide  $B_w$  across the mountain range (Figure S9). Glacier 4 has the highest glacier-wide  $B_w$  and Glacier 13 has the lowest glacier-wide  $B_w$ . However, the  $B_w$  gradient is steeper for OK estimates when compared with RK estimates. Glacier 4 has a similar glacier-wide  $B_w$  between the two methods but glacier-wide  $B_w$  on Glaciers 2 and 13 are much lower than on Glacier 4.

RK produces  $R^2$  values almost as large as those of OK (Figure S10, Fig. 6). The correlation coefficient for Glacier 2 is especially high, with more than 80% of the observed variance explained by the RK model. Glacier 4 has the lowest correlation coefficient but this value is considerably higher than that arising from topographic regression alone.

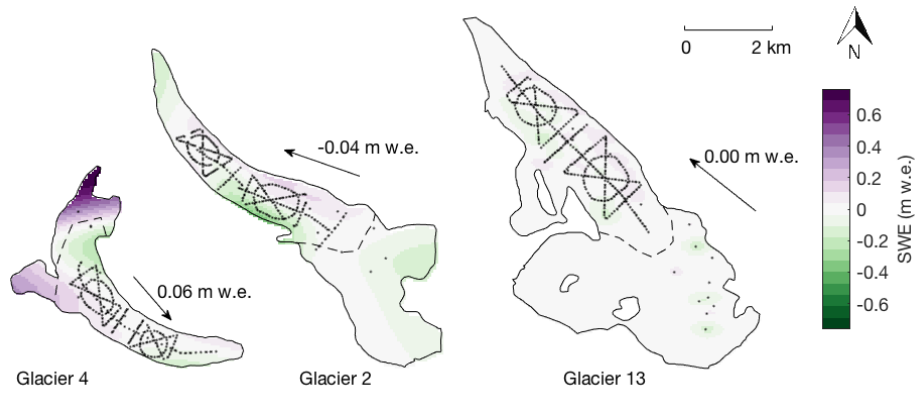


Fig. S8: Distributed Bayesian model averaging residuals estimated by regression kriging (RK). Arrows indicate glacier flow direction and black dots show snow-depth sampling locations. Dashed line indicates approximate ELA.

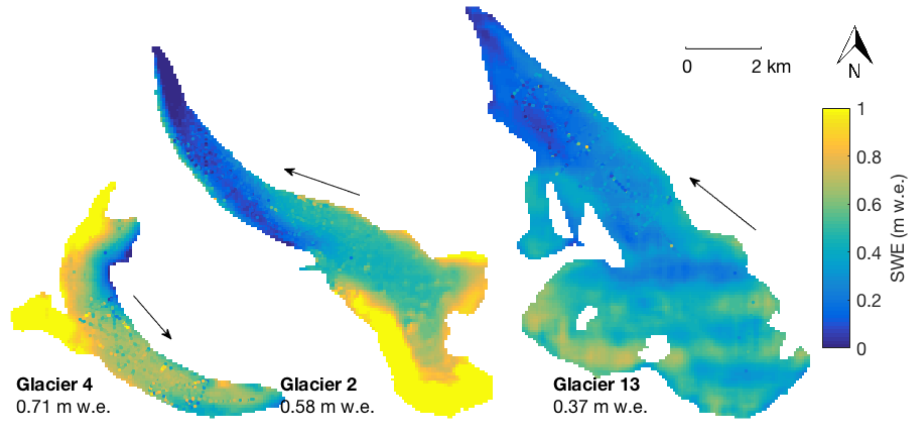


Fig. S9: Winter balance estimated by adding kriged regression residuals to winter balance estimated using topographic regression with Bayesian model averaging. Arrows indicate glacier flow direction and observed values of winter balance are overlain on the maps.

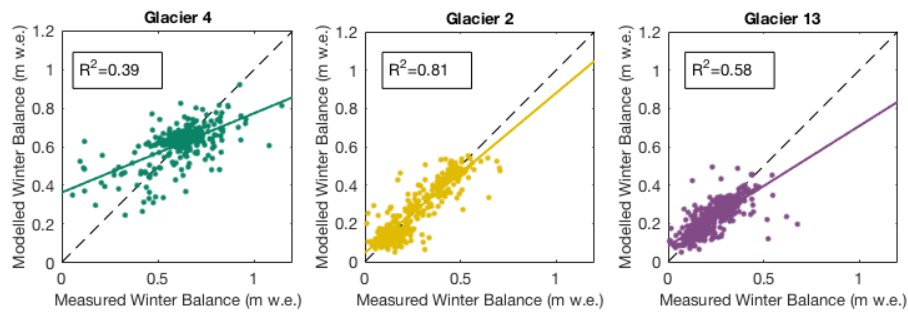


Fig. S10: Comparison of estimated (regression kriging, RK) and observed winter balance ( $B_w$ ) for study glaciers. The  $B_w$  values are calculated using inverse-distance weighted snow-pit densities (S4).

**REFERENCES**

- Braithwaite RJ (2008) Temperature and precipitation climate at the equilibrium-line altitude of glaciers expressed by the degree-day factor for melting snow. *J. Glaciol.*, **54**(186), 437–444 (doi: 10.3189/002214308785836968)
- Hengl T, Heuvelink GB and Rossiter DG (2007) About regression-kriging: From equations to case studies. *Comput. Geosci.*, **33**(10), 1301–1315 (doi: 10.1016/j.cageo.2007.05.001)
- Hofierka J, Mitášová H and Neteler M (2009) Geomorphometry in GRASS GIS. *Dev. Soil Sci.*, **33**, 387–410 (doi: 10.1016/S0166-2481(08)00017-2)
- Knotters M, Brus D and Voshaar JO (1995) A comparison of kriging, co-kriging and kriging combined with regression for spatial interpolation of horizon depth with censored observations. *Geoderma*, **67**(3), 227–246 (doi: 10.1016/0016-7061(95)00011-C)
- Korona J, Berthier E, Bernard M, Rémy F and Thouvenot E (2009) SPIRIT SPOT 5 stereoscopic survey of Polar Ice: Reference images and topographies during the fourth International Polar Year (2007–2009). *J. Photogram. Rem. Sens.*, **64**(2), 204–212 (doi: 10.1016/j.isprsjprs.2008.10.005)
- McGrath D, Sass L, O’Neel S, Arendt A, Wolken G, Gusmeroli A, Kienholz C and McNeil C (2015) End-of-winter snow depth variability on glaciers in Alaska. *J. Geophys. Res.-Earth*, **120**(8), 1530–1550 (doi: 10.1002/2015JF003539)
- Mitášová H and Hofierka J (1993) Interpolation by regularized spline with tension: II. Application to terrain modeling and surface geometry analysis. *Math. Geol.*, **25**(6), 657–669 (doi: 10.1007/BF00893172)
- Molotch N, Colee M, Bales R and Dozier J (2005) Estimating the spatial distribution of snow water equivalent in an alpine basin using binary regression tree models: the impact of digital elevation data and independent variable selection. *Hydrol. Process.*, **19**(7), 1459–1479 (doi: 10.1002/hyp.5586)
- Olaya V (2009) Basic land-surface parameters. *Dev. Soil Sci.*, **33**, 141–169 (doi: 10.1016/S0166-2481(08)00006-8)
- Pulwicky A (2017) *Multi-scale investigation of winter balance on alpine glaciers*. Master’s thesis, Simon Fraser University
- Winstral A, Elder K and Davis RE (2002) Spatial snow modeling of wind-redistributed snow using terrain-based parameters. *J. Hydrometeorol.*, **3**(5), 524–538 (doi: 10.1175/1525-7541(2002)0030524:SSMOWR2.0.CO;2)