Estimating winter balance and its uncertainty from direct measurements of snow depth and density on alpine glaciers

Alexandra PULWICKI,[1] Gwenn E. FLOWERS,[1] Valentina RADIĆ,[2]

[1] *Department of Earth Sciences, Faculty of Science, Simon Fraser University, Burnaby, BC, Canada*

[2] *Department of Earth, Ocean and Atmospheric Sciences, Faculty of Science, University of British Columbia, Vancouver, BC, Canada*

Correspondence: Alexandra Pulwicki <apulwick@sfu.ca>

ABSTRACT. Accurately estimating winter surface mass balance on glaciers is central to assessing glacier health and predicting glacier runoff. However, measuring and modelling snow distribution is inherently difficult in mountainous terrain. Here we explore rigorous statistical methods of estimating winter balance and its uncertainty from multiscale measurements of snow depth and density. In May 2016 we collected over 9000 manual measurements of snow depth across three glaciers in the St. Elias Mountains, Yukon, Canada. Linear regression, combined with cross correlation and Bayesian model averaging, as well as simple kriging are used to interpolate point-scale values to glacier-wide estimates of winter balance. Elevation and a wind-redistribution parameter exhibit the highest correlations with winter balance, but the relationship varies considerably between glaciers. A Monte Carlo analysis reveals that the interpolation itself introduces more uncertainty than the assignment of snow density or the representation of grid-scale variability. For our study glaciers, the winter balance uncertainty from all assessed sources ranges from 0.03 m w.e. (8%) to 0.15 m w.e. (54%). Despite the challenges associated with estimating winter balance, our results are consistent with a regional-scale winter-balance gradient.

INTRODUCTION

Winter surface mass balance, or "winter balance", is the net accumulation and ablation of snow over the winter season (?), which constitutes glacier mass input. Winter balance is half of the seasonally resolved mass balance, initializes summer ablation conditions and must be estimated to simulate energy and mass exchange between the land and atmosphere (e.g. ??). Effectively representing the spatial distribution of snow is also central to monitoring surface runoff and its downstream effects (e.g. ?).

Winter balance (WB) is notoriously difficult to estimate. Snow distribution in alpine regions is highly variable with short correlation length scales (e.g. ????????) and is influenced by dynamic interactions between the atmosphere and complex topography, operating on multiple spatial and temporal scales (e.g. ???). Simultaneously extensive, high resolution and accurate snow distribution measurements on glaciers are therefore difficult to obtain (e.g. ??). Physically based models are computationally intensive and require detailed meteorological data to drive them (?). As a result, there is significant uncertainty in estimates of winter balance, thus limiting the ability of models to represent current and projected glacier conditions.

Studies that have focused on obtaining detailed estimates of WB have used a wide range of observational techniques, including direct measurement of snow depth and density (e.g. ?), lidar or photogrammerty (e.g. ?) and ground-penetrating radar (e.g. ???). Spatial coverage of direct measurements is generally limited and comprises an elevation transect along the glacier centreline (e.g. ?). Measurements are often interpolated using linear regression on only a few topographic parameters (e.g. ?), with elevation being the most common. Other established techniques include hand contouring (e.g. ?), kriging (e.g. ?) and attributing measured winter balance values to elevation bands (e.g. ?). Physical snow models have been used to estimate spatial patterns of winter balance (e.g. ???) but availability of the required meteorological data generally prohibits their widespread application. Error analysis is rarely undertaken and few studies have thoroughly investigated uncertainty in spatially distributed estimates of winter balance (c.f. ?).

More sophisticated snow-survey designs and statistical models of snow distribution are widely used in the field of snow science. Surveys described in the snow science literature are generally spatially extensive and designed to measure snow depth and density throughout a basin, ensuring that all terrain types are sampled. A wide array of measurement interpolation methods are used, including linear (e.g. ?) and non-linear regressions (e.g. ?) that include numerous terrain parameters, as well as geospatial interpolation (e.g. ?) including various forms of kriging. Different interpolation methods are also combined; for example, regression kriging adds kriged residuals to a field obtained with linear regression (e.g. ?). Physical snow models such

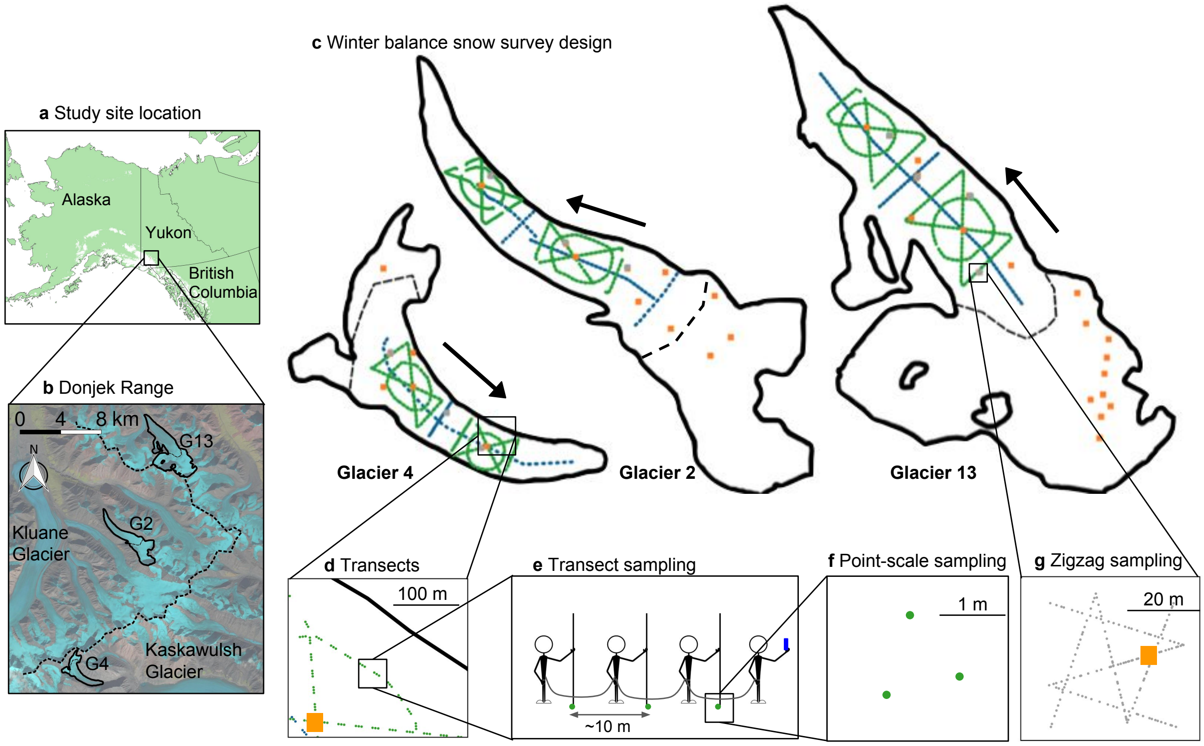

Fig. 1. Study area location and sampling design for Glaciers 4, 2 and 13. (a) Study region in the Donjek Range of the St. Elias Mountains of Yukon, Canada. (b) Study glaciers located along a southwest-northeast transect through the Donjek Range. The local topographic divide is shown as a dashed line. Imagery from Landsat8 (5 September 2013, data available from the U.S. Geological Survey). (c) Details of the snow-survey sampling design, with centreline and transverse transects (blue dots), hourglass and circle designs (green dots) and locations of snow density measurements (orange squares). Arrows indicate ice-flow directions. Approximate location of ELA on each glacier is shown as a black dashed line. (d) Close up of linear and curvilinear transects. (e) Configuration of navigator and observers. (f) Point-scale snow-depth sampling. (g) Linear-random snow-depth measurements in 'zigzag' design (grey dots) with one density measurement (orange square) per zigzag.

as Alpine3D (**?**) and SnowDrift3D (**?**) are widely used, and errors in estimating snow distribution have been examined from theoretical (e.g. **?**) and applied perspectives (e.g. **???**).

The goals of this study are to (1) critically examine methods of converting direct snow depth and density measurements to distributed estimates of winter balance and (2) identify sources of uncertainty, evaluate their magnitude and assess their combined contribution to uncertainty in glacier-wide winter balance. We focus on commonly applied, low-complexity methods of measuring and estimating winter balance in the interest of making our results broadly applicable.

Table 1. Physical characteristics of the study glaciers and May 2016 winter-balance survey details, including number of snow-depth measurement locations along transects (n_T), total length of transects (d_T), number of combined snow pit and Federal Sampler density measurement locations (n_ρ) and number of zigzag surveys (n_zz).

	Location	**Elevation (m a.s.l)**			**Slope (°)**	**Area**	**Survey**	**Survey Details**			
	UTM Zone 7	*Mean*	*Range*	*ELA*	*Mean*	**(km^2)**	**Dates**	n_T	d_T (km)	n_ρ	n_zz
Glacier 4	595470 E 6740730 N	2344	1958–2809	∼2500	12.8	3.8	4–7 May 2016	649	13.1	10	3
Glacier 2	601160 E 6753785 N	2495	1899–3103	∼2500	13.0	7.0	8–11 May 2016	762	13.6	11	3
Glacier 13	604602 E 6763400 N	2428	1923–3067	∼2380	13.4	12.6	12–15 May 2016	941	18.1	20	4

STUDY SITE

The St. Elias Mountains (Fig. 1a) rise sharply from the Pacific Ocean, creating a significant climatic gradient between coastal maritime conditions, generated by Aleutian–Gulf of Alaska low-pressure systems, and interior continental conditions, driven by the Yukon–Mackenzie high-pressure system (?). The boundary between the two climatic zones is generally aligned with the divide between the Hubbard and Kaskawulsh Glaciers, approximately 130 km from the coast. Research on snow distribution and glacier mass balance in this area is limited. A series of research programs, including Project "Snow Cornice" and the Icefield Ranges Research Project, were operational in the 1950s and 60s (??) and in the last 30 years, there have been a few long-term studies on selected alpine glaciers (e.g. ?) as well as several regional studies of glacier mass balance and dynamics (e.g. ???).

We carried out winter balance surveys on three unnamed glaciers in the Donjek Range of the St. Elias Mountains. The Donjek Range is located approximately 40 km to the east of the regional mountain divide and has an area of about $30 \times 30\,\mathrm{km}^2$. Glacier 4, Glacier 2 and Glacier 13 (labelling adopted from ?) are located along a southwest-northeast transect through the range (Fig. 1b, Table 1). These small alpine glaciers are generally oriented southeast-northwest, with Glacier 4 having a predominantly southeast aspect and Glaciers 2 and 13 have generally northwest aspects. The glaciers are situated in valleys with steep walls and have simple geometries. Based on a detailed study of Glacier 2 (?) and related theoretical modelling (?) we suspect all of the study glaciers to be polythermal.

METHODS

Estimating glacier-wide winter balance involves transforming measurements of snow depth and density into values of winter balance distributed across a defined grid. We do this in four steps. (1) Obtain direct measurements of snow depth and density in the field. (2) Assign density values to all depth-measurement locations to calculate point-scale values of WB at each location. (3) Average all point-scale values within each gridcell of a digital elevation model (DEM) to obtain the gricell-averaged WB. (4) Interpolate and extrapolate these gridcell-averaged WB values to obtain estimates of WB (in m w.e.) in each gridcell across the domain. For brevity, we refer to these four steps as (1) field measurements, (2) density assignment, (3) gridcell-averaged WB and (4) distributed WB. Detailed methodology for each step is outlined below.

Field measurements

Sampling design

The snow surveys were designed to capture variability in snow depth at regional, basin, gridcell and point scales (**?**). To capture variability at the regional scale we chose three glaciers along a transect aligned with the dominant precipitation gradient (Fig. 1) (**?**). To account for basin-scale variability, snow depth was measured along linear and curvilinear transects on each glacier (Fig. 1c) with a sample spacing of 10–60 m (Fig. 1d). Sample spacing was constrained by protocols for safe glacier travel, while survey scope was constrained by the need to complete all surveys within the period of peak accumulation. We selected centreline and transverse transects as the most commonly used survey designs in winter balance studies (e.g. **??**) as well as an hourglass pattern with an inscribed circle, which allows for sampling in multiple directions and easy travel (personal communication from C. Parr, 2016). To capture variability at the grid scale, we densely sampled up to four gridcells on each glacier using a linear-random sampling design we term a ‘zigzag’. To capture point-scale variability, each observer made 3–4 depth measurements within ∼1 m (Fig. 1e) at each transect measurement location. In total, we collected more than 9000 snow-depth measurements throughout the study area (Table 1).

Snow depth: transects

Winter balance can be estimated as the product of snow depth and depth-averaged density. Snow depth is generally accepted to be more variable than density (**???**) so we chose a sampling design that resulted in a high ratio (∼55:1) of snow depth to density measurements. Our sampling campaign involved four people and occurred between 5–15 May 2016, which falls within the period of historical peak snow accumulation in southwestern Yukon (Yukon Snow Survey Bulletin and Water Supply Forecast, May 1, 2016). While roped-up for glacier travel with fixed distances between observers, the lead observer used a single-frequency GPS unit (Garmin GPSMAP 64s) to navigate between predefined transect measurement locations (Fig. 1e). The remaining three observers used 3.2 m graduated aluminum avalanche probes to make snow-depth measurements. The locations of each set of depth measurements, made by the second, third and fourth observers, are estimated using the recorded location of the first observer, the approximate distance between observers and the direction of travel.

Snow-depth sampling was concentrated in the ablation area to ensure that only snow from the current accumulation season was measured. The boundary between snow and firn in the accumulation area can be difficult to detect and often misinterpreted, especially when using an avalanche probe (**??**). We intended to use a firn corer to measure winter balance in the accumulation area, but cold snow combined with positive air temperatures led to cores being unrecoverable. Successful snow depth and density measurements within the accumulation area were made either in snow pits or using a Federal Sampler (described below) to unambiguously identify the snow–firn transition.

Snow depth: zigzags

To capture snow-depth variability within a single DEM gridcell, we implemented a linear-random zigzag sampling design (**?**). We measured depth at random intervals of 0.3–3.0 m along two 'Z'-shaped patterns, resulting in 135–191 measurements per zigzag, within three to four 40×40 m gridcells (Fig. 1g) per glacier. Random intervals were machine-generated from a uniform distribution in sufficient numbers that each survey was unique. Zigzag locations were randomly chosen within the upper, middle and lower regions of the ablation area of each glacier. A fourth zigzag was measured on Glacier 13 in the central ablation area at $\sim$2200 m a.s.l.

Snow density

Snow density was measured using a Snowmetrics wedge cutter in three snow pits on each glacier, as well as with a Geo Scientific Ltd. metric Federal Sampler. Within the snow pits (SP), we measured a vertical density profile (in 5 cm increments) with the $5 \times 10 \times 10$ cm wedge-shaped cutter (250 cm^3) and a Presola 1000 g spring scale (e.g. **??**). Wedge-cutter error is approximately $\pm$6% (e.g. **??**). Uncertainty in estimating density from snow-pit measurements also stems from incorrect assignment of density to layers that cannot be sampled (e.g. ice lenses and hard layers). We attempt to quantify this uncertainty by varying estimated ice-layer thickness by $\pm$1 cm ($\leq$100%) of the recorded thickness, ice layer density between 700 and 900 kg m^{-3}

and the density of layers identified as being too hard to sample (but not ice) between 600 and 700 kg m^{-3}. When considering all three sources of uncertainty, the range of integrated density values is always less than 15% of the reference density. Depth-averaged densities for shallow pits (<50 cm) that contain ice lenses are particularly sensitive to changes in prescribed density and ice-lens thickness.

While snow pits provide the most accurate measure of snow density, digging and sampling a snow pit is time and labour intensive. Therefore, a Federal Snow Sampler (FS) (?), which directly measures depth-integrated snow-water equivalent, was used to augment the snow pit measurements. A minimum of three Federal Sampler measurements were taken at each of 7–19 locations on each glacier and an additional eight Federal Sampler measurements were co-located with each snow pit profile. Measurements for which the snow core length inside the sampling tube was less than 90% of the snow depth were discarded. Densities at each measurement location (eight at each snow pit, three elsewhere) were then averaged, with the standard deviation taken to represent the uncertainty.

During the field campaign there were two small accumulation events. The first, on 6 May 2016, also involved high winds so accumulation could not be determined. The second, on 10 May 2016, resulted in 0.01 m w.e accumulation measured at one location on Glacier 2. 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. The total amount of accumulation and melt during the study period could not be estimated so no corrections were made.

Density assignment

Measured snow density must be interpolated or extrapolated to estimate point-scale winter balance at each snow-depth sampling location. We employ four commonly used methods to interpolate and extrapolate density (Table 2): (1) calculate mean density over an entire mountain range (e.g. ?), (2) calculate mean density for each glacier (e.g. ??), (3) linear regression of density on elevation for each glacier (e.g. ??) and (4) calculate mean density using inverse-distance weighting (e.g. ?) for each glacier. Densities derived from snow-pit (SP) measurements and the Federal Sampler (FS) are treated separately, for reasons explained below, resulting in eight possible methods of assigning density.

Gridcell-averaged winter balance

We average one to six (mean of 2.1 measurements) point-scale values of WB within each 40 × 40 m DEM gridcell to obtain the gricell-averaged WB. The locations of individual measurements have uncertainty due to

Table 2. Eight methods used to estimate snow density at unmeasured locations. Total number of resulting density values given in parentheses, with n_T the total number of snow-depth measurement locations along transects (Table 1).

Method code	**Source of measured snow density**		**Density assignment method**
	Snow pit	*Federal Sampler*	
S1	■		Mean of measurements
F1		■	across all glaciers (1)
S2	■		Mean of measurements
F2		■	for each glacier (3)
S3	■		Regression of density on
F3		■	elevation for a glacier (n_T)
S4	■		Inverse distance weighted
F4		■	mean for a glacier (n_T)

the error in the horizontal position given by the GPS unit and the estimation of observer location based on the recorded GPS positions of the navigator. This location uncertainty could result in the incorrect assignment of a point-scale WB to a particular gridcell. However, this source of error is not further investigated because we assume that the uncertainty in gridcell-averaged WB is captured in the zigzag measurements described below. Uncertainty due to having multiple observers was also evaluated. There are no significant differences between snow-depth measurements made by observers along any transect ($p>0.05$), with the exception of the first transect on Glacier 4 (51 measurements).

Distributed winter balance

Gridcell-averaged values of WB are interpolated and extrapolated across each glacier using linear regression (LR) and simple kriging (SK). The regression relates gridcell-averaged WB and various topographic parameters, as this method is simple and has precedent for success (e.g. **?**). Instead of a basic regression however, we use cross-validation and model averaging to test all combinations of the topographic parameters. We compare the regression approach with simple kriging (SK), a data-driven interpolation method free of any physical interpretation (e.g. **?**).

Linear regression

In the regression, we use commonly applied topographic parameters as in ?, including elevation, slope, aspect, curvature, "northness" and a wind-redistribution parameter; we add distance-from-centreline as an additional parameter. Our sampling design ensured that the ranges of topographic parameters associated with our measurement locations represent more than 70% of the total area of each glacier (except elevation on Glacier 2, where our measurements captured only 50%). Topographic parameters are standardized and then weighted by a set of fitted regression coefficients (β_i) calculated by minimizing the sum of squares of the vertical deviations of each datum from the regression line (?). For details on data and methods used to estimate the topographic parameters see the Supplementary Material.

To avoid overfitting the data and to incorporate every possible combination of topographic parameters, cross-validation and model averaging are implemented. First, cross-validation is used to obtain a set of β_i values that have the greatest predictive ability. We randomly select 1000 subsets of the data (2/3 of the values) to fit the LR and use the remaining data (1/3 of the values) to calculate a root mean squared error (RMSE) (?). From the 1000 sets of β_i values, we select the set that results in the lowest RMSE. Second, we use model averaging to account for uncertainty when selecting predictors and to maximize the model's predictive ability (?). Models are generated by calculating a set of β_i as described above for all possible combinations of topographic parameters (2^7 models). Using a Bayesian framework, model averaging involves weighting all models by their posterior model probabilities (?). To obtain the final regression coefficients, the β_i values from each model are weighted according to the relative predictive success of the model, as assessed by the value of the Bayesian Information Criterion (BIC) (?). BIC penalizes more complex models which further reduces the risk of overfitting. The distributed WB is then obtained by applying the resulting regression coefficients to the topographic parameters associated with each gridcell.

Simple kriging

Simple kriging (SK) is a data-driven method of estimating variables at unsampled locations by using the isotropic spatial correlation (covariance) of measured values to find a set of optimal weights (??). Simple kriging assumes spatial correlation between sampling locations that are distributed across a surface and then applies the correlation to interpolate between these locations. We used the `DiceKriging` R package (?) to calculate the maximum likelihood covariance matrix, as well as the range distance (θ) and nugget for gridcell-averaged values of winter balance. The range distance is a measure of data correlation length and the nugget is the residual that encompasses sampling-error variance as well as the spatial variance at distances less than the minimum sample spacing (?). Unlike topographic regression, simple kriging is not useful for generating hypotheses to explain the physical controls on snow distribution, nor can it be used to estimate winter balance on unmeasured glaciers.

Uncertainty analysis

To quantify the uncertainty on estimates of glacier-wide WB we conduct a Monte Carlo analysis, which uses repeated random sampling of input variables to calculate a distribution of output variables (?). We repeat the random sampling process 1000 times, resulting in a distribution of values of the glacier-wide WB based on uncertainties associated with the four steps outlined above. We use the standard deviation of this distribution as a useful metric of uncertainty on the glacier-wide WB. Three sources of uncertainty are considered separately: the uncertainty due to (1) grid-scale variability of WB (σ_{GS}), (2) the assignment of snow density (σ_ρ) and (3) interpolating and extrapolating gridcell-averaged values of WB (σ_{INT}). These individual sources of uncertainty are propagated through the conversion of snow depth and density measurements to glacier-wide WB. Finally, the combined effect of all three sources of uncertainty on the glacier-wide WB is quantified. We calculate a relative uncertainty as the normalized sum of differences between every pair of one hundred distributed WB estimates including σ_{GS} and σ_{INT}.

Grid-scale uncertainty (σ_{GS})

We make use of the zigzag surveys to quantify the true variability of WB at the grid scale. Our limited data do not permit a spatially-resolved assessment of grid-scale uncertainty, so we characterize this uncertainty as uniform across each glacier and represent it by a normal distribution. The distribution is centred at zero and has a standard deviation equal to the mean standard deviation of all zigzag measurements for each glacier. For each iteration of the Monte Carlo, WB values are randomly chosen from the distribution and added to the values of gridcell-averaged WB. These perturbed gridcell-averaged values of WB are then used in the interpolation. We represent uncertainty in glacier-wide WB due to grid-scale uncertainty (σ_{GS}) as the standard deviation of the resulting distribution of glacier-wide WB estimates.

Density assignment uncertainty (σ_ρ)

We incorporate uncertainty due to the method of density assignment by carrying forward all eight density interpolation methods (Table 2) when estimating glacier-wide WB. By choosing to retain even the least plausible options, as well as the questionable FS data, this approach results in a generous assessment of uncertainty. We represent the glacier-wide WB uncertainty due to density assignment uncertainty (σ_ρ) as the standard deviation of glacier-wide WB estimates calculated using each density assignment method.

Interpolation uncertainty (σ_{INT})

We represent the uncertainty due to interpolation of gridcell-averaged WB in different ways for LR and SK. LR interpolation uncertainty is represented by a multivariate normal distribution of possible regression coefficients (β_i). The standard deviation of each distribution is calculated using the covariance of regression coefficients as outlined in **?**, which ensures that regression coefficients are internally consistent. The β_i distributions are randomly sampled and used to calculate gridcell-estimated WB.

SK interpolation uncertainty is represented by the 95% confidence interval for gridcell-estimated values of WB generated by the `DiceKriging` package. From this confidence interval, the standard deviation of each gridcell-estimated WB is then calculated. The standard deviation of glacier-wide WB is then found by taking the square root of the average variance of each gridcell-estimated WB. The final distribution of glacier-wide WB values is centred at the glacier-wide WB estimated with SK. For simplicity, the standard deviation of glacier-wide WB values that result from either LR or SK interpolation uncertainty is referred to as σ_{INT}.

RESULTS AND DISCUSSION

Field measurements

Snow depth

Mean snow depth varied systematically across the study region, with Glacier 4 having the highest mean snow depth and Glacier 13 having the lowest (Fig. 2). At each measurement location, the median range of measured depths (3–4 points) as a percent of the mean local depth is 2%, 11% and 12%, for Glaciers 4, 2 and 13, respectively. While Glacier 4 has the lowest point-scale variability, as assessed above, it also has the highest proportion of outliers, indicating a more variable snow depth across the glacier. The average standard deviation of all zigzag depth measurements is 0.07 m, 0.17 m and 0.14 m, for Glaciers 4, 2 and 13, respectively. When converted to values of WB using the local FS-derived density measurement, the average standard deviation is 0.027 m w.e., 0.035 m w.e. and 0.040 m w.e. WB data for each zigzag are not normally distributed (Fig. 3).

Snow density

Contrary to expectation, co-located FS and SP measurements are found to be uncorrelated ($R^2 = 0.25$, Fig. 2b). The Federal Sampler appears to oversample in deep snow and undersample in shallow snow.

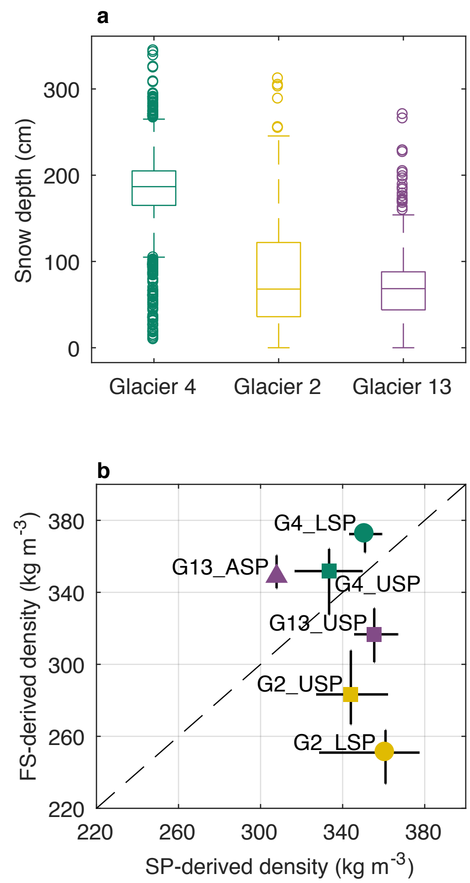

Fig. 2. Measured snow depth and density. (a) Boxplot of measured snow depth on Glaciers 4, 2 and 13 with the first quartiles (box), median (line within box), minimum and maximum values excluding outliers (bar) and outliers (circles), which are defined as being outside of the range of 1.5 times the quartiles (approximately $\pm 2.7\sigma$). (b) Comparison of depth-averaged densities estimated using Federal Sampler (FS) measurements and a wedge cutter in a snow pit (SP) for Glacier 4 (G4), Glacier 2 (G2) and Glacier 13 (G13). Labels indicate snow pit locations in the accumulation area (ASP), upper ablation area (USP) and lower ablation area (LSP). Error bars for SP-derived densities are calculated by varying the thickness and density of layers that are too hard to sample, and error bars for FS-derived densities are the standard deviation of measurements taken at one location. One-to-one line is dashed.

Oversampling by small-diameter (3.2–3.8 cm) sampling tubes has been observed in previous studies, with a percent error between 6.8% and 11.8% (e.g. **???**). Studies that use Federal Samplers often apply a 10% correction to all measurements for this reason (e.g. **?**). Oversampling has been attributed to slots "shaving" snow into the tube as it is rotated (e.g. **?**) and to snow falling into the slots, particularly for snow samples with densities >400 kg m^{-3} and snow depths >1 m (e.g. **?**). Undersampling is likely to occur due to loss of snow from the bottom of the sampler (**?**). Loss by this mechanism may have occurred in our study, given the isothermal and melt-affected snow conditions observed over the lower reaches of Glaciers 2 and 13. Relatively

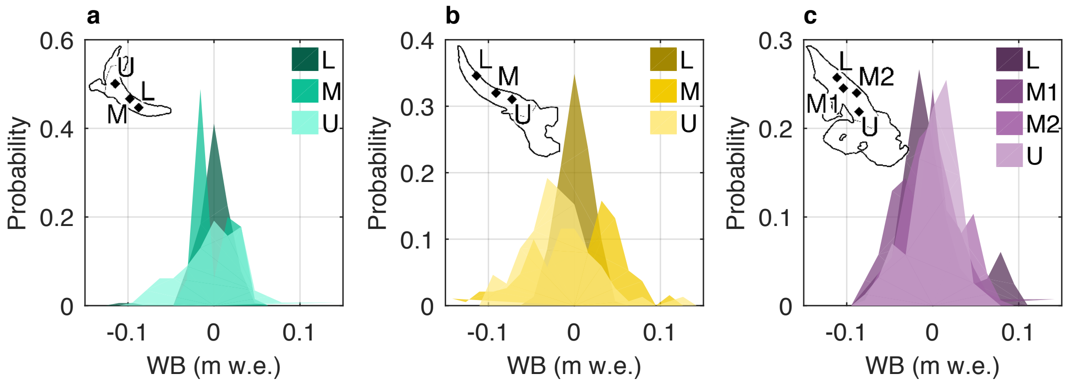

Fig. 3. Distributions of estimated winter-balance values for each zigzag survey in lower (L), middle (M) and upper (U) ablation areas (insets). Local mean has been subtracted. (a) Glacier 4. (b) Glacier 2. (c) Glacier 13.

poor Federal Sampler spring-scale sensitivity also calls into question the reliability of measurements for snow depths <20 cm.

Our FS-derived density values are positively correlated with snow depth ($R^2 = 0.59$). This relationship could be a result of physical processes, such as compaction in deep snow and preferential formation of depth hoar in shallow snow, but is more likely a result of measurement artefacts for a number of reasons. First, the total range of densities measured by the Federal Sampler seems improbably large (227–431 kg m^{-3}) given the conditions at the time of sampling. Moreover, the range of FS-derived values is much larger than than of SP-derived values when co-located measurements are compared. Second, compaction effects of the magnitude required to explain the density differences between SP and FS measurements would not be expected at the measured snow depths (up to 340 cm). Third, no linear relationship exists between depth and SP-derived density ($R^2 = 0.05$). These findings suggest that the Federal Sampler measurements have a bias for which we have not identified a suitable correction.

Density assignment

Given the lack of correlation between co-located SP- and FS-derived densities (Fig. 2), we use the densities derived from these two methods separately (Table 2). SP-derived regional (S1) and glacier-mean (S2) densities are within one standard deviation of the corresponding FS-derived densities (F1 and F2) although SP-derived density values are larger (see Supplementary Material, Table S2). For both SP- and FS-derived densities, the mean density for any given glacier (S2 or F2) is within one standard deviation of the mean across all glaciers (S1 or F1). Correlations between elevation and SP- and FS-derived densities are generally high ($R^2 > 0.5$) but vary between glaciers (Supplementary material, Table S2). For any given glacier, the standard deviation

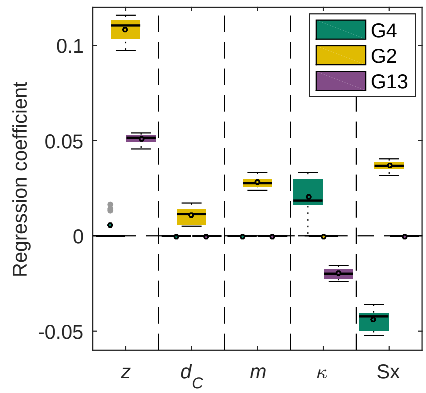

Fig. 4. Distribution of coefficients (β_i) determined by linear regression of gridcell-averaged WB on DEM-derived topographic parameters for the eight different density assignment methods (Table 2). Coefficients are calculated using standardized data, so values can be compared across parameters. Regression coefficients that are not significant are assigned a value of zero. Topographic parameters include elevation (z), distance from centreline (d_C), slope (m), curvature (κ) and wind redistribution (Sx). Aspect (α) and "northness" (N) are not shown because coefficient values are zero in every case. The box plot shows first quartiles (box), median (line within box), mean (circle within box), minimum and maximum values excluding outliers (bars) and outliers (gray dots), which are defined as being outside of the range of 1.5 times the quartiles (approximately $\pm 2.7\sigma$).

of the 3–4 SP- or FS-derived densities is <13% of the mean of those values (S2 or F2) (Supplementary material, Table S2). We adopt S2 (glacier-wide mean of SP-derived densities) as the reference method of density assignment. Though the method described by S2 does not account for known basin-scale spatial variability in snow density (e.g. **?**), it is commonly used in winter balance studies(e.g. **???**).

Gridcell-averaged winter balance

The distributions of gridcell-averaged WB values for the individual glaciers are similar to those in Fig. 2a but with fewer outliers. The standard deviations of WB values determined from the zigzag surveys are almost twice as large as the mean standard deviation of point-scale WB values within a gridcell measured along transects. However, a small number of gridcells sampled in transect surveys have standard deviations in WB that exceed 0.25 m w.e. ($\sim$10 times greater than those for zigzag surveys). We nevertheless assume that the gridcell uncertainty is captured with dense sampling in zigzag gridcells.

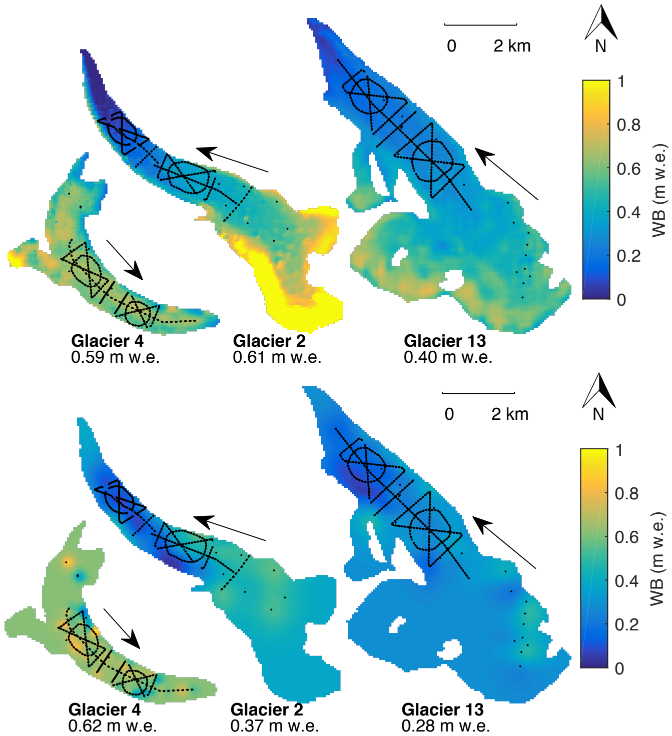

Fig. 5. Spatial distribution of winter balance (WB) estimated using linear regression (top row) and simple kriging (bottom row) with densities assigned as per S2 (Table 2). Locations of snow-depth measurements are shown as black dots. Ice-flow directions are indicated by arrows. Values of glacier-wide WB are given below labels.

Distributed winter balance

Linear Regression

Of the topographic parameters in the linear regression, elevation (z) is the most significant predictor of gridcell-averaged WB for Glaciers 2 and 13, while wind redistribution (Sx) is the most significant predictor for Glacier 4 (Fig. 5, Fig. 4). As expected, gridcell-averaged WB is positively correlated with elevation

Table 3. Glacier-wide winter balance (WB, m w.e.) estimated using linear regression and simple kriging for the three study glaciers. Root mean squared error (RMSE, m w.e.) is computed as the average of all RMSE values between gridcell-averaged values of WB (the data) that were randomly selected and excluded from interpolation (1/3 of all data) and those estimated by interpolation. RMSE as a percent of the glacier-wide WB is shown in brackets.

	Linear regression		**Simple kriging**	
	WB	RMSE	WB	RMSE
G4	0.58	0.15 (26%)	0.62	0.13 (21%)
G2	0.58	0.10 (17%)	0.37	0.07 (19%)
G13	0.38	0.08 (21%)	0.27	0.07 (26%)

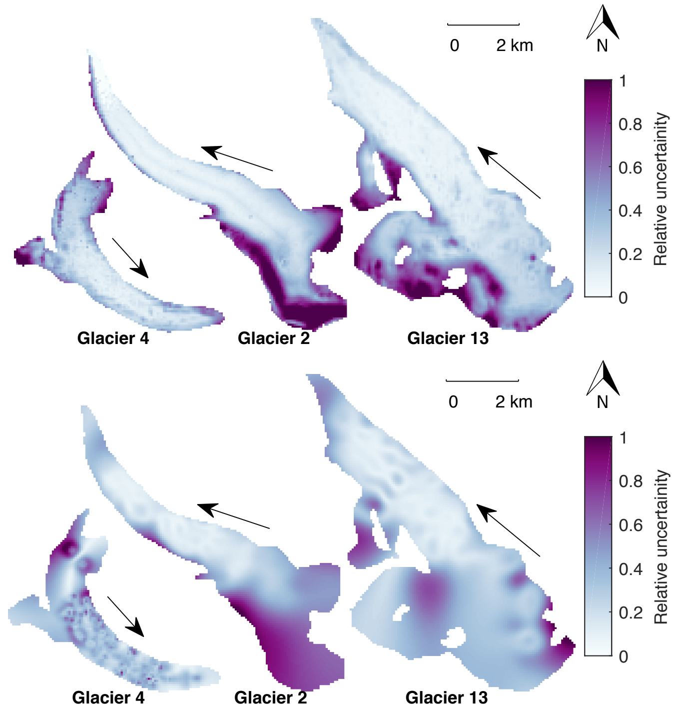

Fig. 6. Relative uncertainty in distributed winter balance (WB) (Fig. 4) found using linear regression (top row) and simple kriging (bottom row). Values closer to one indicate higher relative uncertainty. Ice-flow directions are indicated by arrows.

where the correlation is significant. It is possible that the elevation correlation was accentuated due to melt onset for Glacier 13 in particular. Many studies have found elevation to be the most significant predictor of winter-balance data (e.g. **??**). However, WB–elevation gradients vary considerably between glaciers (e.g. **?**) and other factors, such as glacier shape and orientation relative to dominant wind direction, are strong predictors of the winter-balance distribution (**??**). Some studies find no significant correlation between WB on glaciers and topographic parameters, with highly variable distributions of snow attributed to complex interactions between topography and the atmosphere that could not be easily quantified (e.g. **??**). Extrapolating relationships to unmeasured locations, especially the accumulation area, is susceptible to large uncertainties (Fig. 7). This extrapolation has a considerable effect on values of glacier-wide WB, as the highest values of WB are typically found in the accumulation area (Fig. 4).

Gridcell-averaged WB is negatively correlated with Sx on Glacier 4, counter-intuitively indicating less snow in what would be interpreted as sheltered areas. Gridcell-averaged WB is positively correlated with Sx on Glaciers 2 and 13. Similarly, gridcell-averaged WB is positively correlated with curvature on Glacier 4 and negatively correlated on Glaciers 2 and 13. Wind redistribution and preferential deposition of snow are

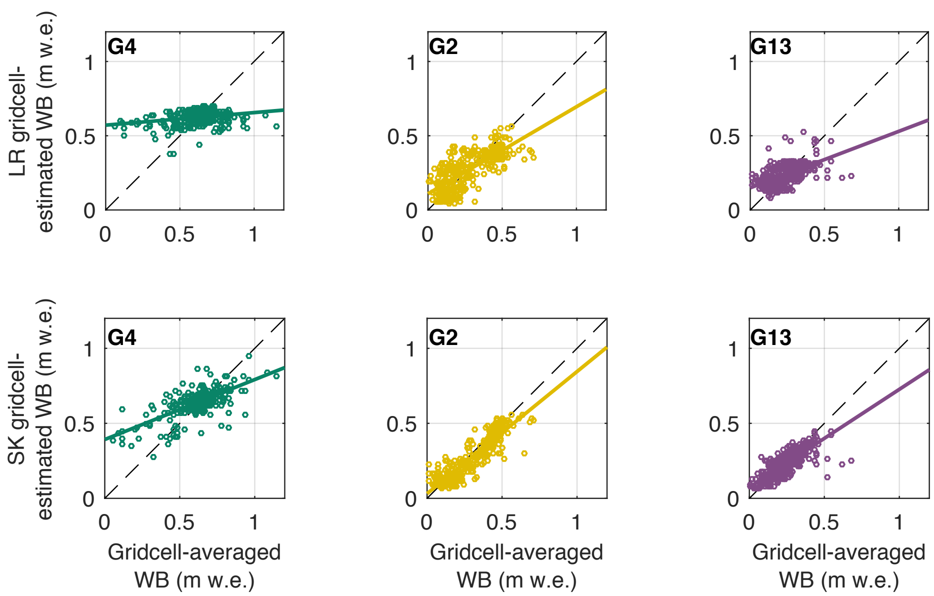

Fig. 7. Winter balance (WB) estimated by linear regression (LR, top row) or simple kriging (SK, bottom row) versus the grid-cell averaged WB data for Glacier 4 (left), Glacier 2 (middle) and Glacier 13 (right). Each circle represents a single gridcell. Best-fit (solid) and one-to-one (dashed) lines are shown.

known to have a large influence on snow distribution at sub-basin scales (e.g. **???**). Our results point to wind having an impact on snow distribution, but the wind redistribution parameter (Sx) may not adequately capture these effects at our study sites. For example, Glacier 4 is located in a curved valley with steep side walls, so specifying a single cardinal direction for wind may not be adequate. Further, the scale of deposition may be smaller than the resolution of the Sx parameter estimated from the DEM. Our results corroborate those of **?** in a study of six glaciers in Alaska (DEM resolutions of 5 m) where elevation and Sx were the only significant parameters for all glaciers; Sx regression coefficients were smaller than elevation regression coefficients, and in some cases, negative. In addition to wind redistribution, sublimation from blowing snow has also been shown to be an important mechanism of mass loss from ridges (e.g. **?**). Incorporating such losses, as well as redistribution and preferential deposition, may be important for improving representations of distributed winter balance.

We find that transfer of LR coefficients between glaciers results in large estimation errors. Regression coefficients from Glacier 4 produce the highest root mean squared error (0.38 m w.e. on Glacier 2 and 0.40 m w.e. on Glacier 13, see Table 3 for comparison) and glacier-wide WB values are the same for all glaciers (0.64 m w.e.) due to the dominance of the regression intercept. Even if the regression is performed with WB values from all glaciers combined, the resulting coefficients produce large root mean squared errors when applied to individual glaciers (0.31 m w.e., 0.15 m w.e. and 0.14 m w.e. for Glaciers 4, 2 and 13, respectively). Our results are consistent with those of **?**, who found that local statistical models cannot be transferred across basins and that regional-scale models are not able to explain the majority of observed variance in winter balance.

Simple kriging

Fitted kriging parameters, including the nugget and spatial correlation length, can provide insight into important scales of winter-balance variability. The model fitted to the gridcell-averaged values of WB for Glacier 4 has a short correlation length (90 m) and large nugget (see Supplementary Material Table S3), suggesting variability in winter balance at smaller scales. Conversely, Glaciers 2 and 13 have longer correlation lengths (∼450 m) and smaller nuggets, suggesting variability at larger scales. Additionally, simple kriging is better able to estimate values of WB for Glaciers 2 and 13 than for Glacier 4 (Fig. 6). Due to a paucity of data, simple kriging produces almost uniform gridcell-estimated values of winter balance in the accumulation area of each glacier, inconsistent with observations described in the literature (e.g. **??**). Extrapolation using simple kriging leads to large uncertainty (Fig. 7), further emphasizing the need for spatially distributed point-scale measurements.

LR and SK comparison

Glacier-wide WB estimates found using both LR and SK are ∼0.58 m w.e. for Glacier 4 but both are poor predictors of WB in measured gridcells (Table 3). For Glaciers 2 and 13, SK estimates are more than 0.1 m w.e. (up to 40%) lower than LR estimates (Table 3). RMSE as a percentage of the glacier-wide WB are comparable between LR and SK (Table 3) with an average RMSE of 22%. This comparability is interesting, given that all of the data were used to generate the SK model, while only 2/3 were used in the LR. Gridcell-estimated values of WB found using LR and SK differ markedly in the upper accumulation areas of Glaciers 2 and 13 (Fig. 4), where observations are sparse and topographic parameters, such as elevation, vary considerably. The influence of elevation results in substantially higher LR-estimated values of WB at high elevation, whereas SK-estimated values approximate the nearest data. Estimates of ablation-area-wide WB differ by <7% between LR and SK on each glacier, further emphasizing the combined role of interpolation method and measurement scarcity in the accumulation area on glacier-wide WB estimates.

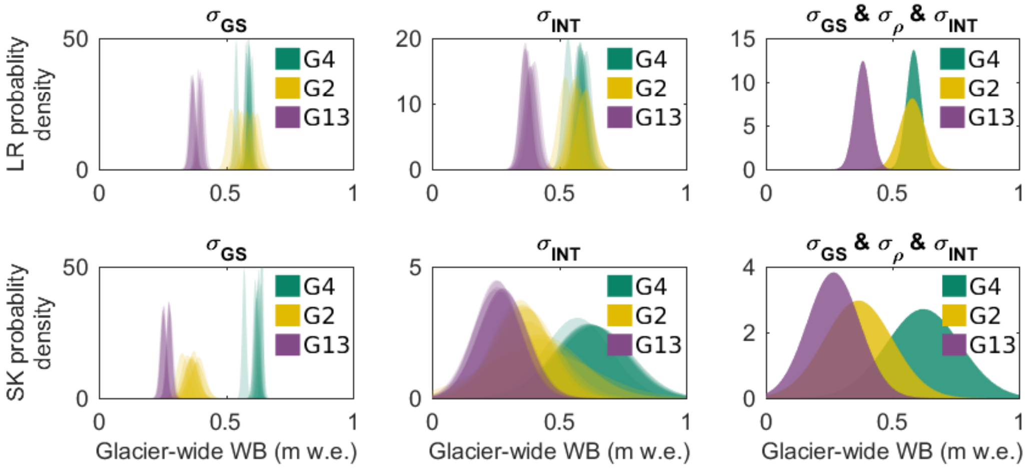

Fig. 8. Distributions of glacier-wide winter balance (WB) for Glaciers 4 (G4), 2 (G2) and 13 (G13) that arise from various sources of uncertainty. WB distribution arising from grid-scale uncertainty (σ_{GS}) (left column). WB distribution arising from interpolation uncertainty (σ_{INT}) (middle column). WB distribution arising from a combination of σ_{GS}, σ_{INT} and density assignment uncertainty (σ_ρ) (right column). Results are shown for interpolation by linear regression (LR, top row) and simple kriging (SK, bottom row). Left two columns include eight distributions per glacier (colour) and each corresponds to a density assignment method (S1–S4 and F1–F4).

Uncertainty analysis

Glacier-wide winter balance is affected by uncertainty introduced by the representativeness of gridcell-averaged values of WB (σ_{GS}), choosing a method of density assignment (σ_ρ), and interpolating WB values across the domain (σ_{INT}). Using a Monte Carlo analysis, we find that interpolation uncertainty contributes more to WB uncertainty than grid-scale uncertainty or density assignment method. In other words, the distribution of glacier-wide WB that arises from grid-scale uncertainty and the differences in distributions between methods of density assignment are smaller than the distribution that arises from interpolation uncertainty (Fig. 8 and Table 4). The WB distributions obtained using LR and SK overlap for a given glacier, but the distribution modes differ (Fig. 8). For reasons outlined above, SK-estimated values of WB in the accumulation area are generally lower, which lowers the glacier-wide WB estimate. The uncertainty in SK-estimated values of WB is large, and unrealistic glacier-wide values of WB of 0 m w.e. can be estimated (Fig. 8). Our results caution strongly against including extrapolated values of WB in comparisons with remote sensing- or model-derived estimates of WB. If possible, such comparisons should be restricted to point-scale data.

Table 4. Standard deviation ($\times 10^{-2}$ m w.e.) of glacier-wide winter balance distributions arising from uncertainties in grid-scale WB (σ_{GS}), density assignment (σ_ρ), interpolation (σ_{INT}) and all three sources combined (σ_{ALL}) for linear regression (left columns) and simple kriging (right columns)

	Linear regression				**Simple kriging**			
	σ_{GS}	σ_ρ	σ_{INT}	σ_{ALL}	σ_{GS}	σ_ρ	σ_{INT}	σ_{ALL}
Glacier 4	0.86	1.90	2.13	2.90	0.85	2.15	14.05	14.72
Glacier 2	1.80	3.37	3.09	4.90	2.53	2.03	13.78	13.44
Glacier 13	1.12	1.68	2.80	3.20	1.15	1.27	9.65	10.43

Grid-scale uncertainty (σ_{GS}) is the smallest assessed contributor to overall WB uncertainty. This result is consistent with the generally smoothly-varying snow depths encountered in zigzag surveys, and previously reported ice-roughness lengths on the order of centimetres (e.g. **?**) compared to snow depths on the order of decimetres to metres. Given our assumption that zigzags are an adequate representation of grid-scale variability, the low WB uncertainty arising from σ_{GS} implies that subgrid-scale sampling need not be a high priority for reducing overall uncertainty. Our assumption that the 3–4 zigzag surveys can be used to estimate glacier-wide σ_{GS} may be flawed, particularly in areas with debris cover, crevasses and steep slopes.

Our analysis did not include uncertainty arising from a number of sources, which we assume either to be encompassed by the sources investigated or to be negligible contributors. These sources of uncertainty include density measurement errors associated with the Federal Sampler, wedge cutters and spring scales, vertical and horizontal errors in the DEM and error associated with estimating measurement locations.

The values of glacier-wide WB for our study glaciers (using LR and S2 density assignment method), with an uncertainty equal to one standard deviation of the distribution found with Monte Carlo analysis, are: 0.59 ± 0.03 m w.e. for Glacier 4, 0.61 ± 0.05 m w.e. for Glacier 2 and 0.40 ± 0.03 m w.e. for Glacier 13. The glacier-wide WB uncertainty from combined sources of uncertainty ranges from 0.03 m w.e (5%) to 0.05 m w.e (8%) for linear-regression estimates and from 0.10 m w.e (37%) to 0.15 m w.e (24%) for simple-kriging estimates (Table 3).

Context and caveats

Regional winter-balance gradient

Although we find considerable inter- and intra-basin variability in winter balance, our results are consistent with a regional-scale winter-balance gradient for the continental side of the St. Elias Mountains (Fig. 9). Winter-balance data are compiled from **?**, the three glaciers presented in this paper and two snow pits

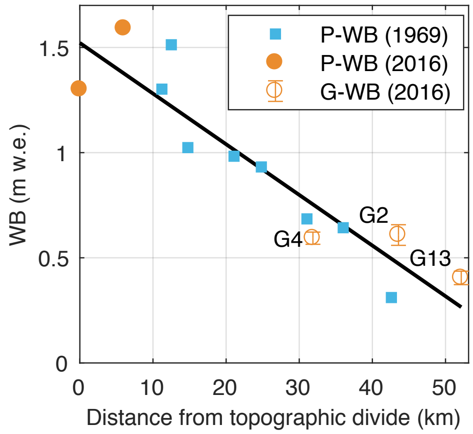

Fig. 9. Relationship between winter balance (WB) and linear distance from the regional topographic divide between the Kaskawulsh and Hubbard Glaciers in the St. Elias Mountains. Point-scale values of WB from snow-pit data reported by **?** (blue boxes, P-WB). LR-estimated glacier-wide WB calculated using density assignment S2 for Glaciers 4 (G4), 2 (G2) and 13 (G13) with errors bars calculated as the standard deviation of Monte Carlo-derived WB distributions (this study) (open orange circles, G-WB). Point-scale WB estimated from snow-pit data at two locations in the accumulation area of the Kaskawulsh Glacier, collected in May 2016 (unpublished data, SFU Glaciology Group) (filled orange dots, P-WB). Black line indicates best fit ($R^2 = 0.85$).

we analyzed near the head of the Kaskawulsh Glacier between 20–21 May 2016. The data show a linear decrease of 0.024 m w.e. km^{-1} ($R^2 = 0.85$) in winter balance with distance from the regional topographic divide between the Kaskawulsh and Hubbard Glaciers, as identified by **?**. While the three study glaciers fit the regional trend, the same relationship would not apply if just the Donjek Range were considered. We hypothesize that interaction between meso-scale weather patterns and large-scale mountain topography is a major driver of regional-scale winter balance. Further insight into regional-scale patterns of winter balance in the St. Elias Mountains could be gained by investigating moisture source trajectories and the contribution of orographic precipitation.

Limitations and future work

The potential limitations of our work include the restriction of our data to a single year, minimal sampling in the accumulation area, the problem of uncorrelated SP- and FS-derived densities, a sampling design that could not be optimized a priori, the assumption of spatially uniform subgrid variability and lack of more finely resolved DEMs.

Inter-annual variability in winter balance is not considered in our study. A number of studies have found temporal stability in spatial patterns of snow distribution and that statistical models based on topographic parameters could be applied reliably between years (e.g. **?**). For example, **?** analyzed more than 40 years of winter balance recorded on two Norwegian glaciers and found that snow distribution is spatially heterogeneous yet exhibits robust temporal stability. Contrary to this, **?** found that snow distribution in Iceland differed considerably between years and depended primarily on the dominant wind direction over the course of a winter. Therefore, multiple years of snow depth and density measurements, that are not necessarily consecutive, are needed to better understand inter-annual variability in winter-balance distribution within the Donjek Range.

There is a conspicuous lack of data in the accumulation areas of our study glaciers. With increased sampling in the accumulation area, interpolation uncertainties would be reduced where they are currently greatest and the linear regression would be better constrained. Although certain regions of the glaciers remain inaccessible for direct measurements, other methods of obtaining winter-balance, including ground-penetrating radar and DEM differencing with photogrammetry or lidar, could be used in conjunction with manual probing to increase the spatial coverage of measurements.

The lack of correlation between SP- and FS-derived densities needs to be reconciled. Contrary to our results, most studies that compare SP- and FS-derived densities report minimal discrepancy (e.g. **?**, and sources within). Additional co-located density measurements are needed to better compare the two methods of obtaining density values. Comparison with other Federal Samplers would also be informative. Even with this limitation, density assignment was, fortunately, not the largest source of uncertainty in estimating glacier-wide winter balance.

Our sampling design was chosen to achieve broad spatial coverage of the ablation area, but is likely too finely resolved along transects for many mass-balance surveys to replicate. An optimal sampling design would minimize uncertainty in winter balance while reducing the number of required measurements. Analysis of the estimated winter balance obtained using subsets of the data is underway to make recommendations on optimal transect configuration and along-track spacing of measurements. **?** found that 200–400 observations are needed within a non-glacierized alpine basin ($6\,\mathrm{km}^2$) to obtain accurate and robust snow distribution models. Similar guidelines would be useful for glacierized environments.

In this study, we assume that the subgrid variability of winter balance is uniform across a given glacier. Contrary to this assumption, **?** found greater variability of winter-balance values close to the terminus. Testing our assumption could be a simple matter of prioritizing the labour-intensive zigzags surveys. To ensure consistent quantification of subgrid variability, zigzag survey measurements could also be tested against other measurements methods, such as lidar.

DEM gridcell size is known to influence values of computed topographic parameters (**????**). The relationship between topographic parameters and winter balance is, therefore, not independent of DEM gridcell size. For example, **?** and **?** found that a decrease in spatial resolution of the DEM results in a decrease in the importance of curvature and an increase in the importance of elevation in regressions of snow distribution on topographic parameters in non-glacierized basins. The importance of curvature in our study is affected by the DEM smoothing that we applied to obtain a spatially continuous curvature field (see Supplementary Material). A comparison of regression coefficients from high-resolution DEMs obtained from various sources and sampled with various gridcell sizes could be used to characterize the dependence of topographic parameters on DEMs, and therefore assess the robustness of inferred relationships between winter balance and topographic parameters.

CONCLUSION

We estimate winter balance for three glaciers (termed Glacier 2, Glacier 4 and Glacier 13) in the St. Elias Mountains, Yukon, Canada from multiscale snow depth and density measurements. Linear regression and simple kriging are used to obtain estimates of distributed winter balance. We use Monte Carlo analysis to evaluate the contributions of interpolation, the assignment of snow density and grid-scale variability of winter balance to uncertainty in glacier-wide winter balance.

Values of glacier-wide winter balance estimated using linear regression and simple kriging differ by up to 0.24 m w.e. ($\sim$50%). We find that interpolation uncertainty is the largest assessed source of uncertainty in glacier-wide winter balance (5% for linear-regression estimates and 32% for simple-kriging estimates). Uncertainty resulting from the method of density assignment is comparatively low, despite the wide range of methods explored. Given our representation of grid-scale variability, the resulting winter balance uncertainty is small indicating that extensive subgrid-scale sampling is not required to reduce overall uncertainty.

Our results suggest that processes governing distributed winter balance differ between glaciers, highlighting the importance of regional-scale winter-balance studies. The estimated distribution of winter balance on Glacier 4 is characterized by high variability, as indicated by the poor correlation between estimated and observed values and large number of data outliers. Glaciers 2 and 13 appear to have lower spatial variability, with elevation being the dominant predictor of gridcell-averaged winter balance. A wind-redistribution parameter is found to be a weak but significant predictor of winter balance, though conflicting relationships between glaciers make it difficult to interpret. Although challenges persist when estimating winter balance, our data are consistent with a regional-scale winter-balance gradient for the continental side of the St. Elias Mountains.

AUTHOR CONTRIBUTION STATEMENT

AP planned and executed the data collection, performed all calculations and drafted the manuscript. GF conceived of the study, contributed to field planning and data collection, oversaw all stages of the work and edited the manuscript. VR provided guidance with statistical methods and edited the manuscript.

ACKNOWLEDGEMENTS

We thank the Kluane First Nation (KFN), Parks Canada and the Yukon Territorial Government for granting us permission to work in KFN Traditional Territory and Kluane National Park and Reserve. We are grateful for financial support provided by the Natural Sciences and Engineering Research Council of Canada, Simon Fraser University and the Northern Scientific Training Program. We kindly acknowledge Kluane Lake Research Station, Sian Williams, Lance Goodwin and Trans North pilot Dion Parker for facilitating field logistics. We are grateful to Alison Criscitiello and Coline Ariagno for all aspects of field assistance and Sarah Furney for assistance with data entry. Thank you to Etienne Berthier for providing us with the SPIRIT SPOT-5 DEM and for assistance in DEM correction. We are grateful to Derek Bingham and Michael Grosskopf for assistance with the statistics, including simple kriging. Luke Wonneck, Leif Anderson and Jeff Crompton all provided thoughtful and constructive comments on drafts of the manuscript.