

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

Alexandra PULWICKI,<sup>1</sup> Gwenn E. FLOWERS,<sup>1</sup> Valentina RADIC,<sup>2</sup> Derek BINGHAM<sup>3</sup>

<sup>1</sup> *Department of Earth Sciences, Simon Fraser University, Burnaby, BC, Canada*

<sup>2</sup> *Department of Earth, Ocean and Atmospheric Sciences, University of British Columbia, Vancouver, BC, Canada*

<sup>3</sup> *Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, BC, Canada*

*Correspondence: Alexandra Pulwinski <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 ordinary 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 (Cogley and others, 2011), which constitutes glacier mass input. Winter balance ( $B_w$ ) 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. Hock, 2005; Réveillet and others, 2016). Effectively representing the spatial distribution of snow on glaciers is also central to monitoring surface runoff and its downstream effects (e.g. Clark and others, 2011).

Winter balance is notoriously difficult to estimate (e.g. Dadić and others, 2010; Cogley and others, 2011). Snow distribution in alpine regions is highly variable with short correlation length scales (e.g. Anderton and others, 2004; Egli and others, 2011; Grünewald and others, 2010; Helbig and van Herwijnen, 2017; López-Moreno and others, 2011, 2013; Machguth and others, 2006; Marshall and others, 2006) and is influenced by dynamic interactions between the atmosphere and complex topography, operating on multiple spatial and temporal scales (e.g. Barry, 1992; Liston and Elder, 2006; Clark and others, 2011; Scipión and others, 2013). Simultaneously extensive, high resolution and accurate snow distribution measurements on glaciers are therefore difficult to acquire (e.g. Cogley and others, 2011; McGrath and others, 2015) and obtaining such measurements is further complicated by the inaccessibility of many glacierized regions during the winter. Use of physically based models to estimate winter balance is computationally intensive and requires detailed meteorological data to drive the models (Dadić and others, 2010). 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  $B_w$  have used a wide range of observational techniques, including direct measurement of snow depth and density (e.g. Cullen and others, 2017), lidar or photogrammetry (e.g. Sold and others, 2013) and ground-penetrating radar (e.g. Machguth and others, 2006; Gusmeroli and others, 2014; McGrath and others, 2015). Spatial coverage of direct measurements is generally limited and often comprises an elevation transect along the glacier centreline (e.g. Kaser and others, 2003). Measurements are typically interpolated using linear regression on only a few topographic parameters (e.g. MacDougall and Flowers, 2011), with elevation being the most common. Other established techniques include

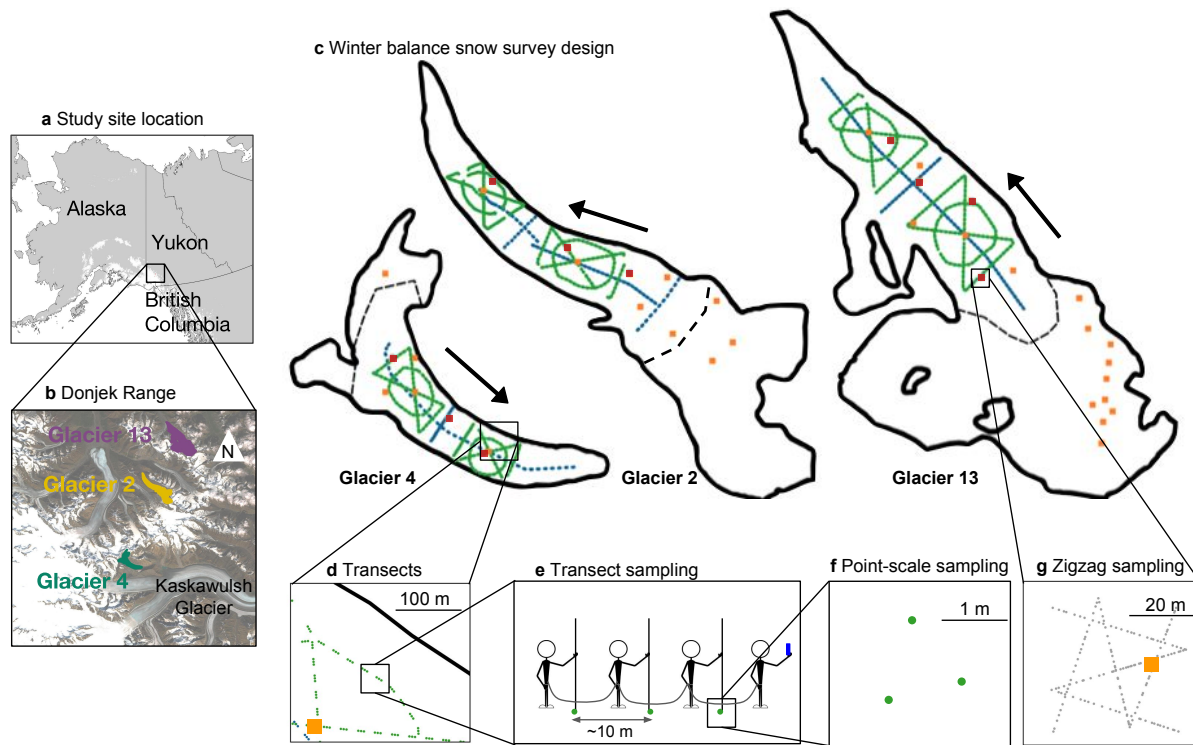
55 hand contouring (e.g. Tangborn and others, 1975), kriging (e.g. Hock and Jensen, 1999) and attributing  
56 measured winter balance values to elevation bands (e.g. Thibert and others, 2008). Physical snow models  
57 have been used to estimate spatial patterns of winter balance (e.g. Mott and others, 2008; Schuler and others,  
58 2008; Dadić and others, 2010), but availability of the required meteorological data generally prohibits their  
59 widespread application. Error analysis is rarely undertaken and few studies have thoroughly investigated  
60 uncertainty in spatially distributed estimates of winter balance (c.f. Schuler and others, 2008).

61 More sophisticated snow-survey designs and statistical models of snow distribution are widely used in  
62 the field of snow science. Surveys described in the snow science literature are generally spatially extensive  
63 and designed to measure snow depth and density throughout a basin, ensuring that all terrain types are  
64 sampled. A wide array of measurement interpolation methods are used, including linear (e.g. López-Moreno  
65 and others, 2010) and non-linear regressions (e.g. Molotch and others, 2005) that include numerous terrain  
66 parameters, as well as geospatial interpolation (e.g. Erxleben and others, 2002; Cullen and others, 2017)  
67 including various forms of kriging. Different interpolation methods are also combined; for example, regression  
68 kriging (see Supplementary Material) adds kriged residuals to a field obtained with linear regression (e.g.  
69 Balk and Elder, 2000). Physical snow models such as SnowTran-3D (Liston and Sturm, 1998), Alpine3D  
70 (Lehning and others, 2006) and SnowDrift3D (Schneiderbauer and Prokop, 2011) are widely used, and errors  
71 in estimating snow distribution have been examined from theoretical (e.g. Trujillo and Lehning, 2015) and  
72 applied perspectives (e.g. Turcan and Loijens, 1975; Woo and Marsh, 1978; Deems and Painter, 2006).

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

## 78 **STUDY SITE**

79 The St. Elias Mountains (Fig. 1a) rise sharply from the Pacific Ocean, creating a significant climatic gradient  
80 between coastal maritime conditions, generated by Aleutian–Gulf of Alaska low-pressure systems, and interior  
81 continental conditions, driven by the Yukon–Mackenzie high-pressure system (Taylor-Barge, 1969). The  
82 boundary between the two climatic zones is generally aligned with the divide between the Hubbard and  
83 Kaskawulsh Glaciers, approximately 130 km from the coast. Research on snow distribution and glacier mass  
84 balance in this area is limited. A series of research programs, including Project “Snow Cornice” and the



**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 (red dots) with one density measurement (orange square) per zigzag.

Icefield Ranges Research Project, were operational in the 1950s and 60s (Wood, 1948; Danby and others, 2003) and in the last 30 years, there have been a few long-term studies on selected alpine glaciers (e.g. Clarke, 2014) as well as several regional studies of glacier mass balance and dynamics (e.g. Arendt and others, 2008; Berthier and others, 2010; Burgess and others, 2013; Waechter and others, 2015).

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  $30 \times 30 \text{ km}^2$ . Glacier 4, Glacier 2 and Glacier 13 (labelling adopted from Crompton and Flowers (2016)) are located along a southwest–northeast transect through the range (Fig. 1b, Table 1).

**Table 1.** Physical characteristics of the study glaciers.

	<b>Location</b>		<b>Elevation (m a.s.l.)</b>			<b>Slope (°)</b>	<b>Area</b>
	UTM Zone 7		<i>Mean</i>	<i>Range</i>	<i>ELA</i>	<i>Mean</i>	(km <sup>2</sup> )
<b>Glacier 4</b>	595470 E	6740730 N	2344	1958–2809	~2500	12.8	3.8
<b>Glacier 2</b>	601160 E	6753785 N	2495	1899–3103	~2500	13.0	7.0
<b>Glacier 13</b>	604602 E	6763400 N	2428	1923–3067	~2380	13.4	12.6

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 (Wilson and others, 2013) and related theoretical modelling (Wilson and Flowers, 2013) we suspect all of the study glaciers to be polythermal.

## METHODS

Estimating glacier-wide winter balance ( $B_w$ ) involves transforming measurements of snow depth and density into values of winter balance distributed across a defined grid ( $b_w$ ). 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  $b_w$  at each location. Winter balance, measured in units of metres water equivalent (m w.e.), can be estimated as the product of snow depth and depth-averaged density. (3)

**Table 2.** Details of the May 2016 winter-balance survey, 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 ( $n_\rho$ ), number of zigzag surveys ( $n_{zz}$ ), number (as percent of total number of gridcells, and of the number of gridcells in the ablation area) of gridcells sampled ( $n_S$ ) and the elevation range (as percent of total elevations range and of ablation-area elevation range).

	<b>Date</b>	$n_T$	$d_T$ (km)	$n_\rho$	$n_{zz}$	$n_S$	<b>Elevation range (m a.s.l.)</b>
<b>Glacier 4</b>	4–7 May 2016	649	13.1	7	3	295 (12%, 21%)	2015–2539 (62%, 97%)
<b>Glacier 2</b>	8–11 May 2016	762	13.6	7	3	353 (8%, 14%)	2151–2541 (32%, 47%)
<b>Glacier 13</b>	12–15 May 2016	941	18.1	19	4	468 (6%, 14%)	2054–2574 (45%, 62%)

104 Average all point-scale values of  $b_w$  within each gridcell of a digital elevation model (DEM) to obtain the  
 105 gridcell-averaged  $b_w$ . (4) Interpolate and extrapolate these gridcell-averaged  $b_w$  values to obtain estimates of  
 106  $b_w$  in each gridcell across the domain.  $B_w$  is then calculated by taking the average of all gridcell-averaged  
 107  $b_w$  values for each glacier. For brevity, we refer to these four steps as (1) field measurements, (2) density  
 108 assignment, (3) gridcell-averaged  $b_w$  and (4) distributed  $b_w$ . Detailed methodology for each step is outlined  
 109 below. We use the SPIRIT SPOT-5 DEM (40×40 m) from 2005 (Korona and others, 2009) throughout this  
 110 study.

## 111 **Field measurements**

112 Our sampling campaign involved four people and occurred between 5–15 May 2016, which falls within the  
 113 period of historical peak snow accumulation in southwestern Yukon (Yukon Snow Survey Bulletin and Water  
 114 Supply Forecast, May 1, 2016). Snow depth is generally accepted to be more variable than density (Elder  
 115 and others, 1991; Clark and others, 2011; López-Moreno and others, 2013) so we chose a sampling design  
 116 that resulted in a high ratio ( $\sim 55:1$ ) of snow depth to density measurements. In total, we collected more  
 117 than 9000 snow-depth measurements and more than 100 density measurements throughout the study area  
 118 (Table 1).

119 During the field campaign there were two small accumulation events. The first, on 6 May 2016, also involved  
 120 high winds so accumulation could not be determined. The second, on 10 May 2016, resulted in 0.01 m w.e  
 121 accumulation measured at one location on Glacier 2. Assuming both accumulation events contributed a  
 122 uniform 0.01 m w.e accumulation to all study glaciers then our survey did not capture  $\sim 3\%$  and  $\sim 2\%$  of  
 123 estimated  $B_w$  on Glaciers 4 and 2, respectively. We therefore assume that these accumulation events were  
 124 negligible and apply no correction. Positive temperatures and clear skies occurred between 11–16 May 2016,  
 125 which we suspect resulted in melt occurring on Glacier 13. The snow in the lower part of the ablation area  
 126 of Glacier 13 was isothermal and showed clear signs of melt and metamorphosis. The total amount of melt  
 127 during the study period was estimated using a degree-day model and found to be small ( $\leq 0.01$  m w.e., see  
 128 Supplementary Material) so no corrections were made.

## 129 *Sampling design*

130 The snow surveys were designed to capture variability in snow depth at regional, basin, gridcell and point  
 131 scales (Clark and others, 2011). To capture variability at the regional scale we chose three glaciers along  
 132 a transect aligned with the dominant precipitation gradient (Fig. 1b) (Taylor-Barge, 1969). To account for  
 133 basin-scale variability, snow depth was measured along linear and curvilinear transects on each glacier (Fig.

134 1c) with a sample spacing of 10–60 m (Fig. 1d). Sample spacing was constrained by protocols for safe glacier  
 135 travel, while survey scope was constrained by the need to complete all surveys within the period of peak  
 136 accumulation. We selected centreline and transverse transects as the most commonly used survey designs  
 137 in winter balance studies (e.g. Kaser and others, 2003; Machguth and others, 2006) as well as an hourglass  
 138 pattern with an inscribed circle, which allows for sampling in multiple directions and easy travel (personal  
 139 communication from C. Parr, 2016). To capture variability at the grid scale, we densely sampled up to four  
 140 gridcells on each glacier using a linear-random sampling design (Shea and Jamieson, 2010) we term a ‘zigzag’.  
 141 To capture point-scale variability, each observer made 3–4 depth measurements within  $\sim 1$  m (Fig. 1f) at  
 142 each transect measurement location.

#### 143 *Snow depth: transects*

144 While roped-up for glacier travel with fixed distances between observers, the lead observer used a single-  
 145 frequency GPS unit (Garmin GPSMAP 64s) to navigate between predefined transect measurement locations  
 146 (Fig. 1e). The remaining three observers used 3.2 m graduated aluminum avalanche probes to make snow-  
 147 depth measurements (Kinar and Pomeroy, 2015). The locations of each set of depth measurements, made by  
 148 the second, third and fourth observers, are estimated using the recorded location of the first observer, the  
 149 approximate distance between observers and the direction of travel. The 3–4 point-scale depth measurements  
 150 are averaged to obtain a single depth measurement at each transect measurement location. When considering  
 151 snow variability at the point scale as a source of uncertainty in snow depth measurements, we find that the  
 152 mean standard deviation of point-scale snow depth measurements is  $<7\%$  of the mean snow depth for all  
 153 study glaciers.

154 Snow-depth sampling was concentrated in the ablation area to ensure that only snow from the current  
 155 accumulation season was measured. The boundary between snow and firn in the accumulation area can be  
 156 difficult to detect and often misinterpreted, especially when using an avalanche probe (Grünewald and others,  
 157 2010; Sold and others, 2013). We intended to use a firn corer to measure winter balance in the accumulation  
 158 area, but cold snow combined with positive air temperatures led to cores being unrecoverable. Successful  
 159 snow depth measurements within the accumulation area were made either in snow pits or using a Federal  
 160 Sampler (described below) to unambiguously identify the snow–firn transition.

#### 161 *Snow depth: zigzags*

162 We measured depth at random intervals of 0.3–3.0 m along two ‘Z’-shaped patterns (Shea and Jamieson,  
 163 2010), resulting in 135–191 measurements per zigzag, within three to four  $40 \times 40$  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. Extra time in the field allowed us to measure a fourth zigzag 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. Within the snow pits (SP), we measured a vertical density profile (in 10 cm increments) with the  $5 \times 5 \times 10$  cm wedge-shaped cutter ( $250 \text{ cm}^3$ ) and a Presola 1000 g spring scale (e.g. Gray and Male, 1981; Fierz and others, 2009; Kinar and Pomeroy, 2015). Wedge-cutter error is approximately  $\pm 6\%$  (e.g. Proksch and others, 2016; Carroll, 1977). Uncertainty in estimating density from SP 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  $900 \text{ kg m}^{-3}$  and the density of layers identified as being too hard to sample (but not ice) between  $600$  and  $700 \text{ kg m}^{-3}$ . 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 SP provide the most accurate measure of snow density, digging and sampling a SP is time and labour intensive. Therefore, a Geo Scientific Ltd. metric Federal Sampler (FS) (Clyde, 1932) with a 3.2–3.8 cm diameter sampling tube, which directly measures depth-integrated snow-water equivalent, was used to augment the SP measurements. A minimum of three FS measurements were taken at each of 7–19 locations on each glacier and an additional eight FS measurements were co-located with two SP profiles for each glacier. 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 SP, three elsewhere) were then averaged, with the standard deviation taken to represent the uncertainty. The mean standard deviation of FS-derived density was  $\leq 4\%$  of the mean density for all glaciers.

### **Density assignment**

Measured snow density must be interpolated or extrapolated to estimate point-scale  $b_w$  at each snow-depth sampling location. We employ four commonly used methods to interpolate and extrapolate density (Table 3): (1) calculate mean density over an entire mountain range (e.g. Cullen and others, 2017), (2) calculate



**Table 3.** 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
	<i>Snow pit</i>	<i>Federal</i>	
		<i>Sampler</i>	
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 each glacier ( $n_T$ )
S4	■		Inverse distance weighted
F4		■	mean for each glacier ( $n_T$ )

mean density for each glacier (e.g. Elder and others, 1991; McGrath and others, 2015), (3) linear regression of density on elevation for each glacier (e.g. Elder and others, 1998; Molotch and others, 2005) and (4) calculate mean density using inverse-distance weighting (e.g. Molotch and others, 2005) for each glacier. Densities derived from SP and FS measurements 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  $b_w$  within each DEM gridcell to obtain the gridcell-averaged  $b_w$ . The locations of individual measurements have uncertainty due to 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  $b_w$  measurement to a particular gridcell. However, this source of error is not further investigated because we assume that the uncertainty resulting from incorrect locations of point-scale  $b_w$  values is captured in the uncertainty derived from zigzag measurements, as described below. Error due to having multiple observers is also evaluated by conducting an analysis of variance (ANOVA) of snow-depth measurements along a transect (amounting to 23 hypothesis tests, one for each transect) and testing for differences between observers. We find 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), where snow depth measurements collected by one observer were, on average, greater than the snow depth measurements taken by the other two observers ( $p < 0.01$ ). Since this was the first transect and the only one to show differences by observer, this difference can be considered an anomaly. We conclude that observer bias is not an important effect in this study and therefore apply no observer corrections to the data.

## Distributed winter balance

Gridcell-averaged values of  $b_w$  are interpolated and extrapolated across each glacier using linear regression (LR) and ordinary kriging (OK). The LR relates gridcell-averaged  $b_w$  to various topographic parameters. We use this method because it is simple and has precedent for success (e.g. McGrath and others, 2015). Instead of a basic LR however, we implement cross-validation to prevent data overfitting as well as model averaging to allow for all combinations of the chosen topographic parameters. We compare the LR approach with OK, an interpolation method free of physical interpretation beyond the premise of spatial correlation in the data (e.g. Hock and Jensen, 1999).

### Linear regression

In the LR, we use commonly applied topographic parameters as in McGrath and others (2015), including elevation, slope, aspect, curvature, “northness” and a wind-redistribution parameter ( $Sx$  from Winstral and others (2002)); we add distance-from-centreline as an additional parameter. Topographic parameters are standardized for use in the LR. For details on data and methods used to estimate the topographic parameters see the Supplementary Material and Pulwinski (2017). 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%).

The goal of the LR is to obtain a set of fitted regression coefficients ( $\beta_i$ ) that correspond to each topographic parameter and to a model intercept. The LR implemented in this study is an extension of a basic multiple linear regression; we use cross-validation to avoid overfitting the data and model averaging to incorporate every possible combination of topographic parameters.

First, cross-validation is used to obtain a set of  $\beta_i$  that have the greatest predictive ability (Kohavi and others, 1995). We randomly select 1000 subsets of the data (2/3 of the values) and fit a basic multiple linear regression (implemented in MATLAB) to the data subsets, thus obtaining 1000 sets of  $\beta_i$ . The basic multiple linear regression calculates a set of  $\beta_i$  by minimizing the sum of squares of the vertical deviations of each datum from the regression line (Davis and Sampson, 1986). Distributed  $b_w$  is then calculated using each

240 set of  $\beta_i$  by weighting topographic parameters by their corresponding  $\beta_i$  values for all DEM gridcells. We  
 241 then use the remaining data (1/3 of the values) to calculate a root mean squared error (RMSE) between the  
 242 estimated  $b_w$  and the observed  $b_w$  for corresponding locations. From the 1000 sets of  $\beta_i$  values, we select the  
 243 set that results in the lowest RMSE.

244 Second, we use model averaging to account for uncertainty when selecting predictors and to maximize the  
 245 model's predictive ability (Madigan and Raftery, 1994). Models are generated by calculating a set of  $\beta_i$  (as  
 246 described above) for all possible combinations of topographic parameters, resulting in  $2^7$  models (i.e.  $2^7$  sets  
 247 of  $\beta_i$  with the greatest predictive ability for each linear combination of topographic parameters). Using a  
 248 Bayesian framework, model averaging involves weighting all models by their posterior model probabilities  
 249 (Raftery and others, 1997). We weight the models according to their relative predictive success, as assessed  
 250 by the value of the Bayesian Information Criterion (BIC) (Burnham and Anderson, 2004). BIC penalizes  
 251 more complex models, which further reduces the risk of overfitting. The final set of  $\beta_i$  is then the weighted  
 252 sum of  $\beta_i$  from all models. Distributed  $b_w$  is obtained by applying the final set of  $\beta_i$  to the topographic  
 253 parameters associated with each gridcell.

#### 254 *Ordinary kriging*

255 Kriging is a method of estimating variables at unsampled locations by using the spatial correlation  
 256 (covariance) of measured values to find a set of optimal weights (Davis and Sampson, 1986; Li and Heap,  
 257 2008). Kriging assumes spatial correlation between sampling locations that are distributed across a surface  
 258 and then applies the correlation to interpolate between these locations. Many forms of kriging have been  
 259 developed to accommodate different data types (e.g. Li and Heap, 2008, and sources within). Ordinary  
 260 kriging (OK) is the most basic form of kriging where the mean of the estimated field is unknown. Unlike LR,  
 261 OK is not useful for generating hypotheses to explain the physical controls on snow distribution, nor can it  
 262 be used to estimate winter balance on unmeasured glaciers. However, we chose to use OK because it does  
 263 not require external inputs and is therefore a means of obtaining  $B_w$  that is free of physical interpretation  
 264 beyond the information contained in the covariance matrix.

265 We used the `DiceKriging` R package (Roustant and others, 2012) to calculate the maximum likelihood  
 266 covariance matrix, as well as the range distance ( $\theta$ ) and nugget for gridcell-averaged values of winter balance.  
 267 The range distance is a measure of data correlation length and the nugget is the residual that encompasses  
 268 sampling-error variance as well as the spatial variance at distances less than the minimum sample spacing

(Li and Heap, 2008). A Matérn covariance function with  $\nu=5/2$  is used to define a stationary and isotropic covariance and covariance kernels are parameterized as in Rasmussen and Williams (2006).

### Uncertainty analysis using a Monte Carlo approach

Three sources of uncertainty are considered separately: the uncertainty due to (1) grid-scale variability of  $b_w$  ( $\sigma_{GS}$ ), (2) the assignment of snow density ( $\sigma_\rho$ ) and (3) interpolating and extrapolating gridcell-averaged values of  $b_w$  ( $\sigma_{INT}$ ). To quantify the uncertainty of grid-scale and interpolation uncertainty on estimates of  $B_w$  we conduct a Monte Carlo analysis that uses repeated random sampling of input variables to calculate a distribution of output variables (Metropolis and Ulam, 1949). We repeat the random sampling process 1000 times, resulting in a distribution of values of  $B_w$  based on uncertainties associated with the four steps associated with deriving  $B_w$  from distributed snow-depth and density measurements. Individual sources of uncertainty are propagated through the conversion of snow depth and density measurements to  $B_w$ . Finally, the combined effect of all three sources of uncertainty on  $B_w$  is quantified. We use the standard deviation of the distribution of  $B_w$  as a useful metric of  $B_w$  uncertainty. Density assignment uncertainty is calculated as the standard deviation of the eight resulting values of  $B_w$ . We calculate a relative uncertainty, equal to the normalized sum of differences between every pair of one hundred distributed  $b_w$  estimates including  $\sigma_{GS}$  and  $\sigma_{INT}$ , to investigate the spatial patterns in  $b_w$  uncertainty.

#### *Grid-scale uncertainty ( $\sigma_{GS}$ )*

We make use of the zigzag surveys to quantify the true variability of  $b_w$  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,  $b_w$  values are randomly chosen from the distribution and added to the values of gridcell-averaged  $b_w$ . These perturbed gridcell-averaged values of  $b_w$  are then used in the interpolation. We represent uncertainty in  $B_w$  due to grid-scale uncertainty ( $\sigma_{GS}$ ) as the standard deviation of the resulting distribution of  $B_w$  estimates.

#### *Density assignment uncertainty ( $\sigma_\rho$ )*

We incorporate uncertainty due to the method of density assignment by carrying forward all eight density interpolation methods (Table 3) when estimating  $B_w$ . 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  $B_w$  uncertainty due to density assignment uncertainty ( $\sigma_\rho$ ) as the standard deviation of  $B_w$  estimates calculated using each density assignment method.

### Interpolation uncertainty ( $\sigma_{\text{INT}}$ )

We represent the uncertainty due to interpolation/extrapolation of gridcell-averaged  $b_w$  in different ways for LR and OK. LR interpolation uncertainty is represented by a multivariate normal distribution of possible regression coefficients ( $\beta_i$ ). The standard deviation of each distribution is calculated using the covariance of  $\beta_i$  as outlined in Bagos and Adam (2015), which ensures that  $\beta_i$  are internally consistent. The  $\beta_i$  distributions are randomly sampled and used to calculate gridcell-estimated  $b_w$ .

OK interpolation uncertainty is represented by the standard deviation for each gridcell-estimated value of  $b_w$  generated by the `DiceKriging` package. The standard deviation of  $B_w$  is then found by taking the square root of the average variance of each gridcell-estimated  $b_w$ . The final distribution of  $B_w$  values is centred at the  $B_w$  estimated with OK. For simplicity, the standard deviation of  $B_w$  values that results from either LR or OK interpolation/extrapolation uncertainty is referred to as  $\sigma_{\text{INT}}$ .

## RESULTS

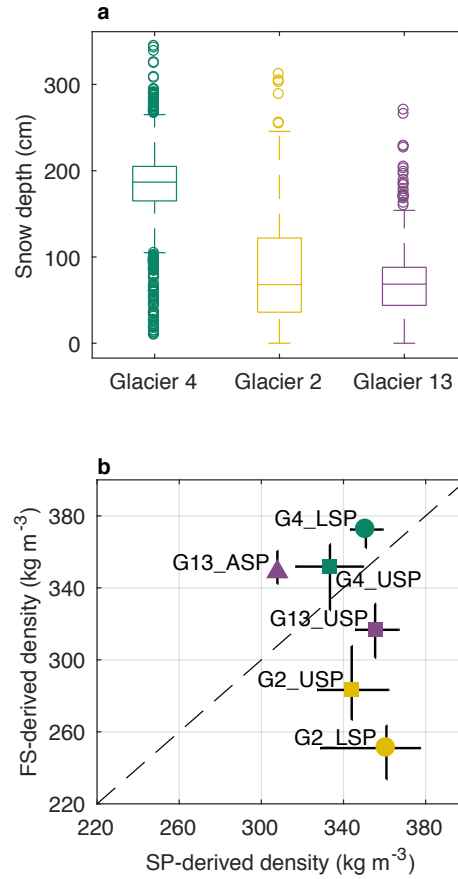
### Field measurements

#### *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. 2a). 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  $b_w$  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. Winter-balance data for each zigzag are not normally distributed (Fig. 3).

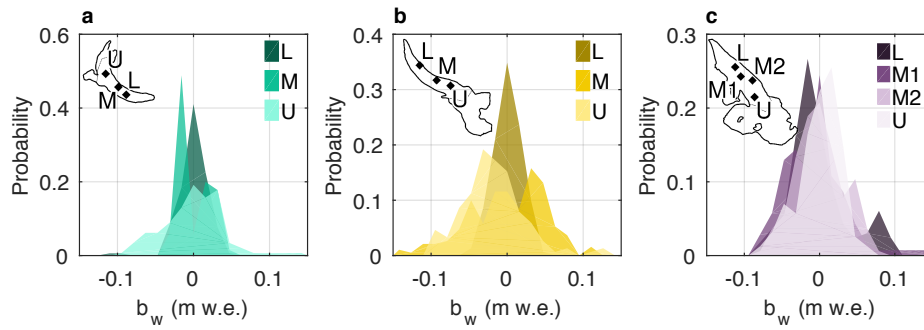
#### *Snow density*

Contrary to expectation, co-located FS and SP measurements are found to be uncorrelated ( $R^2 = 0.25$ , Fig. 2b). The FS appears to oversample in deep snow and undersample in shallow snow. Oversampling by small-diameter sampling tubes has been observed in previous studies, with a percent error between 6.8%



**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 SP 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.

327 and 11.8% (e.g. Work and others, 1965; Farnes and others, 1982; Conger and McClung, 2009). Studies that  
 328 use FS often apply a 10% correction to all measurements for this reason (e.g. Molotch and others, 2005).  
 329 Oversampling has been attributed to slots “shaving” snow into the tube as it is rotated (e.g. Dixon and Boon,  
 330 2012) and to snow falling into the slots, particularly for snow samples with densities  $>400 \text{ kg m}^{-3}$  and snow  
 331 depths  $>1 \text{ m}$  (e.g. Beaumont and Work, 1963). Undersampling is likely to occur due to loss of snow from the  
 332 bottom of the sampler (Turcan and Loijens, 1975). Loss by this mechanism may have occurred in our study,  
 333 given the isothermal and melt-affected snow conditions observed over the lower reaches of Glaciers 2 and 13.



**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.

334 Relatively poor FS spring-scale sensitivity also calls into question the reliability of measurements for snow  
 335 depths  $<20$  cm.

336 Our FS-derived density values are positively correlated with snow depth ( $R^2 = 0.59$ ). This relationship  
 337 could be a result of physical processes, such as compaction in deep snow and preferential formation of depth  
 338 hoar in shallow snow, but is more likely a result of measurement artefacts for a number of reasons. First,  
 339 the total range of densities measured by the FS seems improbably large ( $227\text{--}431\text{ kg m}^{-3}$ ). At the time of  
 340 sampling, the snowpack had little new snow, few ice lenses and was not saturated; the range of measured  
 341 densities is therefore difficult to explain with physical conditions. Moreover, the range of FS-derived values is  
 342 much larger than that of SP-derived values when co-located measurements are compared. Second, compaction  
 343 effects of the magnitude required to explain the density differences between SP and FS measurements would  
 344 not be expected at the measured snow depths (up to 340 cm). Third, no linear relationship exists between  
 345 depth and SP-derived density ( $R^2 = 0.05$ ). These findings suggest that the FS measurements have a bias for  
 346 which we have not identified a suitable correction. Despite this bias, we use FS-derived densities to generate  
 347 a range of possible  $b_w$  estimates and to provide a generous estimate of uncertainty arising from density  
 348 assignment.

### 349 Density assignment

350 Given the lack of correlation between co-located SP- and FS-derived densities (Fig. 2), we use the densities  
 351 derived from these two methods separately (Table 3). SP-derived regional (S1) and glacier-mean (S2) densities  
 352 are within one standard deviation of the corresponding FS-derived densities (F1 and F2) although SP-derived  
 353 density values are larger (see Supplementary Material, Table S3). For both SP- and FS-derived densities, the  
 354 mean density for any given glacier (S2 or F2) is within one standard deviation of the mean across all glaciers

**Table 4.** Glacier-wide winter balance ( $B_w$ , m w.e.) estimated using linear regression and ordinary 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  $b_w$  (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  $B_w$  is shown in brackets.

	Linear regression		Ordinary kriging	
	$B_w$	RMSE	$B_w$	RMSE
<b>G4</b>	0.58	0.15 (26%)	0.62	0.11 (18%)
<b>G2</b>	0.58	0.10 (17%)	0.35	0.06 (18%)
<b>G13</b>	0.38	0.08 (21%)	0.27	0.06 (21%)

(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 S3). For any given glacier, the standard deviation of the 3–4 SP- or FS-derived densities is <13% of the mean of those values (S2 or F2) (Supplementary material, Table S3). 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. Wetlaufer and others, 2016), it is commonly used in winter balance studies (e.g. Elder and others, 1991; McGrath and others, 2015; Cullen and others, 2017).

### Gridcell-averaged winter balance

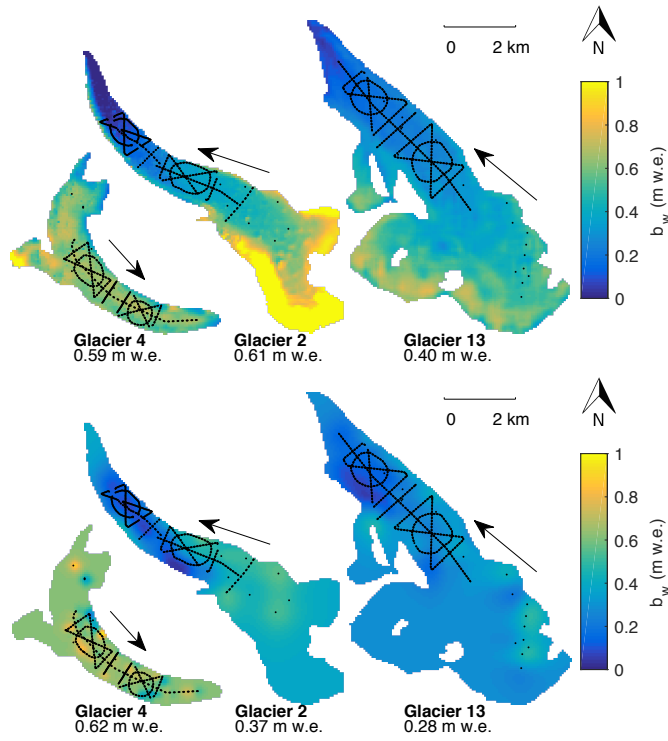
The distributions of gridcell-averaged  $b_w$  values for the individual glaciers are similar to those in Fig. 2a but with fewer outliers (see Supplementary Material, Fig. S4). The standard deviations of  $b_w$  values determined from the zigzag surveys are almost twice as large as the mean standard deviation of point-scale  $b_w$  values within a gridcell measured along transects (see Supplementary Material, Fig. S5). However, a small number of gridcells sampled in transect surveys have standard deviations in  $b_w$  that exceed 0.25 m w.e. ( $\sim 10$  times greater than those for zigzag surveys).

### Distributed winter balance

#### Linear regression

The highest values of estimated  $b_w$  are found in the upper portions of the accumulation areas of Glaciers 2 and 13 (Fig. 4). These areas also correspond to large values of elevation, slope, and wind redistribution. Extrapolation of the positive relation between  $b_w$  and elevation, as well as slope and  $Sx$  for Glacier 2, results in high  $b_w$  estimates and large relative uncertainty in these estimates (Fig. 5). On Glacier 4, the distributed



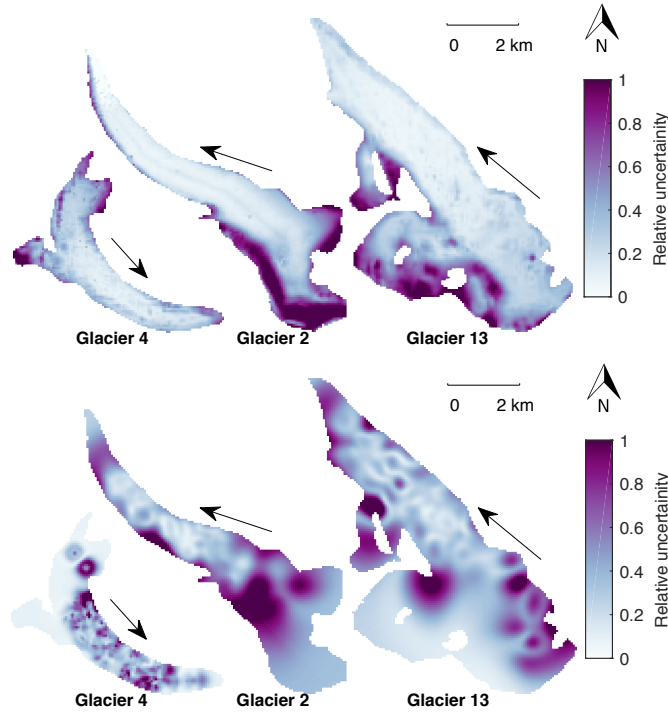


**Fig. 4.** Spatial distribution of winter balance ( $b_w$ ) estimated using linear regression (top row) and ordinary kriging (bottom row) with densities assigned as per S2 (Table 3). The linear regression (LR) method involves multiplying regression coefficients, found using cross validation and model averaging, by topographic parameters for each gridcell. Ordinary kriging (OK) uses the covariance of measured values to find a set of optimal weights for estimating values at unmeasured locations. Locations of snow-depth measurements made in May 2016 are shown as black dots. Ice-flow directions are indicated by arrows. Values of  $B_w$  are given below labels.

375  $b_w$  and the relative uncertainty are almost uniform (Fig. 4) due to the small regression coefficients for  
 376 all topographic parameters. The variance explained by the LR-estimated  $b_w$  differs considerably between  
 377 glaciers (Fig. 6), with the best correlation between modelled and observed  $b_w$  occurring for Glacier 2. LR is  
 378 an especially poor predictor of  $b_w$  on Glacier 4, where  $B_w$  can be estimated equally well using the mean of  
 379 the data. RMSE is also highest for Glacier 4 (Table 4).

#### 380 Ordinary kriging

381 For all three glaciers, large areas that correspond to locations far from measurements have  $b_w$  estimates  
 382 equal to the kriging mean. Distributed  $b_w$  estimated with OK on Glacier 4 is mostly uniform except for local  
 383 deviations close to measurement locations (Fig. 4) and relative uncertainty is highest close to measurement  
 384 locations. Distributed  $b_w$  varies more smoothly on Glaciers 2 and 13 (Fig. 4). Glacier 2 has a distinct region  
 385 of low estimated  $b_w$  ( $\sim 0.1$  m w.e.) in the lower part of the ablation area, which corresponds to a wind-scoured

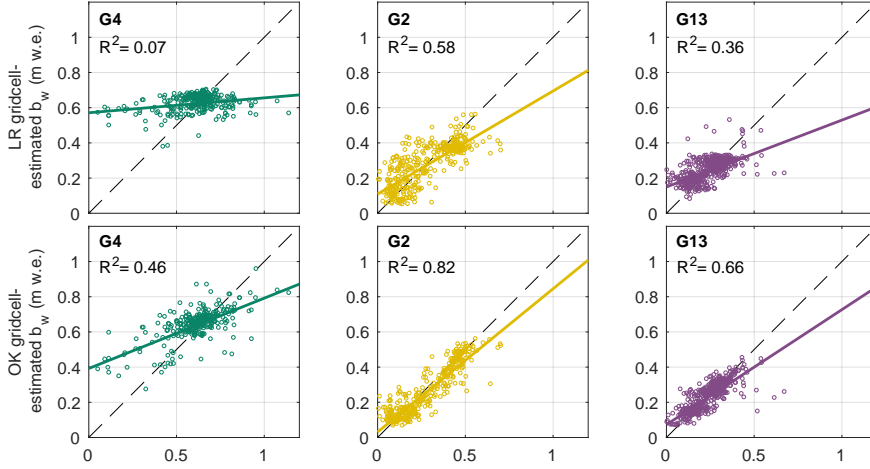


**Fig. 5.** Relative uncertainty in distributed winter balance ( $b_w$ ) (Fig. 4) found using linear regression (top row) and ordinary kriging (bottom row). Values closer to one indicate higher relative uncertainty. Ice-flow directions are indicated by arrows.

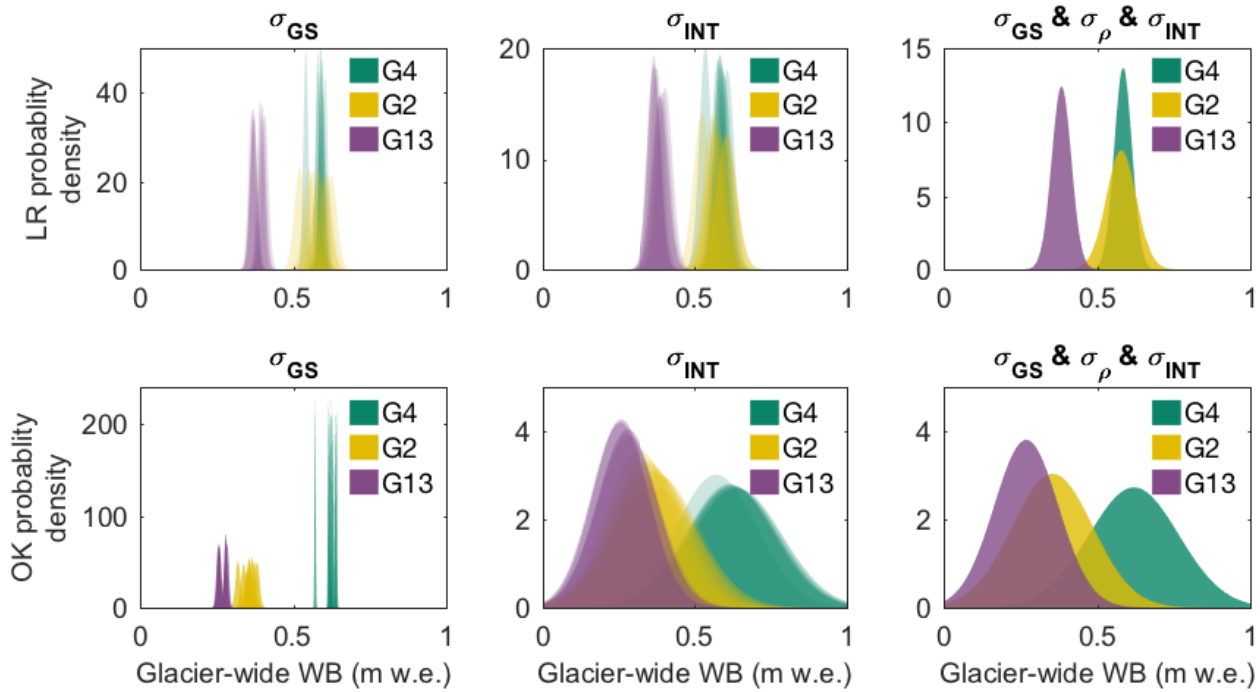
region of the glacier. Glacier 13 has the lowest estimated mean  $b_w$  and only small deviations from this mean at measurement locations (Fig. 4). Relative uncertainty varies considerably across the three study glaciers with the greatest uncertainty just outside of the region with observed  $b_w$  (Fig. 5). As expected, explained variance of OK-estimated  $b_w$  is high for both Glaciers 2 and 13 (Fig. 6) because OK is a data-fitting algorithm. However, explained variance (Fig. 6) for Glacier 4 is relatively low and RMSE is relatively high (Table 4), indicating a highly variable distribution of  $b_w$ .

### Uncertainty analysis using a Monte Carlo approach

Estimates of  $B_w$  are affected by uncertainty introduced by the representativeness of gridcell-averaged values of  $b_w$  ( $\sigma_{GS}$ ), choosing a method of density assignment ( $\sigma_\rho$ ), and interpolating/extrapolating  $b_w$  values across the domain ( $\sigma_{INT}$ ). Using a Monte Carlo analysis, we find that interpolation uncertainty contributes more to  $B_w$  uncertainty than grid-scale uncertainty or the method of density assignment. In other words, the distribution of  $B_w$  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. 7 and Table 5). The  $B_w$  distributions obtained using LR and OK overlap for a given glacier, but the distribution modes



**Fig. 6.** Winter balance ( $b_w$ ) estimated by linear regression (LR, top row) or ordinary kriging (OK, bottom row) versus the grid-cell averaged  $b_w$  data for Glacier 4 (left), Glacier 2 (middle) and Glacier 13 (right). Each circle represents a single gridcell. Explained variance ( $R^2$ ) is provided. Best-fit (solid) and one-to-one (dashed) lines are shown.



**Fig. 7.** Distributions of glacier-wide winter balance ( $B_w$ ) for Glaciers 4 (G4), 2 (G2) and 13 (G13) that arise from various sources of uncertainty.  $B_w$  distribution arising from grid-scale uncertainty ( $\sigma_{GS}$ ) (left column).  $B_w$  distribution arising from interpolation uncertainty ( $\sigma_{INT}$ ) (middle column).  $B_w$  distribution arising from a combination of  $\sigma_{GS}$ ,  $\sigma_{INT}$  and density assignment uncertainty ( $\sigma_\rho$ ) (right column). Results are shown for interpolation by linear regression (LR, top row) and ordinary kriging (OK, bottom row). Left two columns include eight distributions per glacier (colour) and each corresponds to a density assignment method (S1–S4 and F1–F4).

**Table 5.** Standard deviation ( $\times 10^{-2}$  m w.e.) of glacier-wide winter balance ( $B_w$ ) distributions arising from uncertainties in grid-scale  $b_w$  ( $\sigma_{GS}$ ), density assignment ( $\sigma_\rho$ ), interpolation ( $\sigma_{INT}$ ) and all three sources combined ( $\sigma_{ALL}$ ) for linear regression (left columns) and ordinary kriging (right columns)

	Linear regression				Ordinary kriging			
	$\sigma_{GS}$	$\sigma_\rho$	$\sigma_{INT}$	$\sigma_{ALL}$	$\sigma_{GS}$	$\sigma_\rho$	$\sigma_{INT}$	$\sigma_{ALL}$
<b>Glacier 4</b>	0.86	1.90	2.13	2.90	0.18	2.16	14.35	14.64
<b>Glacier 2</b>	1.80	3.37	3.09	4.90	0.80	2.06	12.65	13.14
<b>Glacier 13</b>	1.12	1.68	2.80	3.20	0.57	1.30	9.74	10.48

differ (Fig. 7). OK-estimated values of  $b_w$  in the accumulation area are generally lower (Fig. 4), which lowers the  $B_w$  estimate. The uncertainty in OK-estimated values of  $B_w$  is large, and unrealistic (e.g.  $B_w = 0$  m w.e.) values of  $B_w$  are possible (Fig. 7).

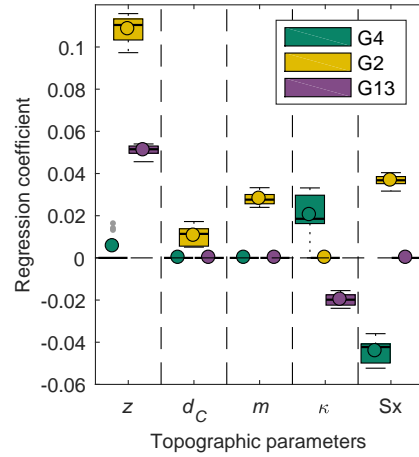
The values of  $B_w$  for our study glaciers (using LR and S2 density assignment method), with an uncertainty equal to one standard deviation of the distribution found with Monte Carlo analysis, are:  $0.59 \pm 0.03$  m w.e. for Glacier 4,  $0.61 \pm 0.05$  m w.e. for Glacier 2 and  $0.40 \pm 0.03$  m w.e. for Glacier 13. The  $B_w$  uncertainty from the three investigated sources of uncertainty ranges from 0.03 m w.e (5%) to 0.05 m w.e (8%) for LR estimates and from 0.10 m w.e (37%) to 0.15 m w.e (24%) for ordinary-kriging estimates (Table 4).

## DISCUSSION

### Distributed winter balance

#### Linear regression

Of the topographic parameters in the LR, elevation ( $z$ ) is the most significant predictor of gridcell-averaged  $b_w$  for Glaciers 2 and 13, while wind redistribution ( $Sx$ ) is the most significant predictor for Glacier 4 (Fig. 8). As expected, gridcell-averaged  $b_w$  is positively correlated with elevation where the correlation is significant. It is possible that the elevation correlation was accentuated due to melt onset for Glacier 13 in particular. Glacier 2 had little snow at the terminus likely due to steep slopes and wind-scouring but the snow did not appear to have been affected by melt. Our results are consistent with many studies that have found elevation to be the most significant predictor of seasonal snow accumulation data (e.g. Machguth and others, 2006; Grünwald and others, 2014; McGrath and others, 2015). The  $b_w$ -elevation gradient is  $13 \text{ mm } 100 \text{ m}^{-1}$  on Glacier 2 and  $7 \text{ mm } 100 \text{ m}^{-1}$  on Glacier 13. These gradients are consistent with those reported for a few glaciers in Svalbard (Winther and others, 1998) but are considerably smaller than many reported  $b_w$ -



**Fig. 8.** Distribution of coefficients ( $\beta_i$ ) determined by linear regression of gridcell-averaged  $b_w$  on DEM-derived topographic parameters for the eight different density assignment methods (Table 3). 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 ( $\kappa$ ) and wind redistribution ( $Sx$ ). Aspect ( $\alpha$ ) 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  $\pm 2.7\sigma$ ).

elevation gradients, which range from about 60 to 240 mm 100 m<sup>-1</sup> (e.g. Hagen and Liestøl, 1990; Tveit and Killingtveit, 1994; Winther and others, 1998). Extrapolating linear relationships to unmeasured locations typically results in considerable estimation error, as seen by the large  $b_w$  values (Fig. 4) and large relative uncertainty (Fig. 5) in the high-elevation regions of Glaciers 2 and 13. The low correlation between  $b_w$  and elevation on Glacier 4 is consistent with Grabiec and others (2011) and López-Moreno and others (2011), who conclude that highly variable distributions of snow can be attributed to complex interactions between topography and the atmosphere that cannot be easily quantified. The snow on Glacier 4 also did not appear to have been affected by melt and it is hypothesized that significant wind-redistribution of snow, which was not captured by the  $Sx$  parameter, covered ice-topography and produced a relatively uniform snow depth across the glacier.

Gridcell-averaged  $b_w$  is negatively correlated with  $Sx$  on Glacier 4, counter-intuitively indicating less snow in what would be interpreted as sheltered areas. Gridcell-averaged  $b_w$  is positively correlated with  $Sx$  on

Glaciers 2 and 13. Our results corroborate those of McGrath and others (2015) 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. While our results point to wind having an impact on snow distribution, the wind redistribution parameter ( $Sx$ ) may not adequately capture these effects at our study sites. For example, Glacier 4 has a curvilinear plan-view profile and is surrounded by steep valley 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. Creation of a parametrization for sublimation from blowing snow, which has been shown to be an important mechanism of mass loss from ridges (e.g. Musselman and others, 2015), may also increase the explanatory power of LR for our study sites.

We find that transfer of LR coefficients between glaciers results in large estimation errors. Regression coefficients from Glacier 4 produce the highest RMSE (0.38 m w.e. on Glacier 2 and 0.40 m w.e. on Glacier 13, see Table 4 for comparison) and  $B_w$  values are the same for all glaciers (0.64 m w.e.) due to the dominance of the regression intercept. Even if the LR is performed with  $b_w$  values from all glaciers combined, the resulting coefficients produce large RMSE 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 Grünewald and others (2013), 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.

#### Ordinary kriging

Due to a paucity of data, ordinary kriging produces almost uniform gridcell-estimated  $b_w$  in the accumulation area of each glacier, inconsistent with observations described in the literature (e.g. Machguth and others, 2006; Grabiec and others, 2011). Glacier 4 has the highest estimated mean with large deviations from the mean at measurement locations. The longer correlation lengths of the data for Glaciers 2 and 13 result in a more smoothly varying distributed  $b_w$ . As expected, extrapolation using OK leads to large uncertainty (Fig. 5), further emphasizing the need for spatially distributed point-scale measurements.

#### LR and OK comparison

LR and OK produce similar estimates of distributed  $b_w$  (Fig. 5) and  $B_w$  ( $\sim 0.60$  m w.e., Table 4) for Glacier 4 but both are relatively poor predictors of  $b_w$  in measured gridcells (Fig. 6). For Glaciers 2 and 13, OK estimates are more than  $\sim 0.22$  m w.e. (39%) and  $\sim 0.11$  m w.e. (30%) lower than LR estimates, respectively (Table 4). RMSE as a percentage of the  $B_w$  is lower for OK than LR only for Glacier 4 but the absolute

RMSE of OK is  $\sim 0.03$  m w.e. lower for all glaciers, likely because OK is a data-fitting interpolation method (Table 4). RMSEs as a percentage of the glacier-wide  $B_w$  are comparable between LR and OK (Table 4) with an average RMSE of 22%. This comparability is interesting, given that all of the data were used to generate the OK model, while only 2/3 were used in the LR. Tests in which only 2/3 of the data were used in the OK model yielded similar results to those in which all data were used. Gridcell-estimated values of  $b_w$  found using LR and OK differ markedly in the upper accumulation areas of Glaciers 2 and 13, where observations are sparse and topographic parameters, such as elevation, attain their highest values. The influence of elevation results in substantially higher LR-estimated values of  $b_w$  at high elevation, whereas OK-estimated values are more uniform. Estimates of ablation-area-wide  $B_w$  differ by  $<6\%$  between LR and OK on each glacier, further emphasizing the combined influence of interpolation method and measurement scarcity in the accumulation area on  $B_w$  estimates.

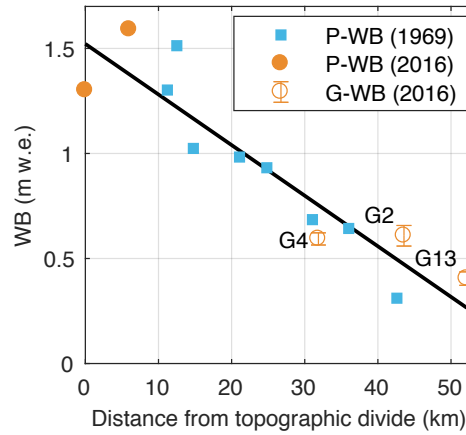
#### Uncertainty analysis using a Monte Carlo approach

Interpolation/extrapolation of  $b_w$  data is the largest contributor to  $b_w$  uncertainty in our study. These results caution strongly against including values of  $b_w$  in comparisons with remote sensing- or model-derived estimates of  $B_w$ . If possible, such comparisons should be restricted to point-scale data. Grid-scale uncertainty ( $\sigma_{GS}$ ) is the smallest assessed contributor to overall  $B_w$  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. Hock, 2005) 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  $B_w$  uncertainty arising from  $\sigma_{GS}$  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  $\sigma_{GS}$  may be flawed, particularly in areas with debris cover, crevasses and steep slopes.

Our analysis did not include uncertainty arising from density measurement errors associated with the FS, wedge cutters and spring scales, from vertical and horizontal errors in the DEM or from error associated with estimating measurement locations based on the GPS position of the lead observer. We assume that these sources of uncertainty are either encompassed by the sources investigated or are negligible.

#### 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 Taylor-Barge (1969), the three glaciers presented in this paper and



**Fig. 9.** Relationship between winter balance and linear distance from the regional topographic divide between the Kaskawulsh and Hubbard Glaciers in the St. Elias Mountains. Point-scale values of winter balance from snow-pit data reported by Taylor-Barge (1969) (blue boxes, P-WB). LR-estimated glacier-wide winter balance ( $B_w$ ) 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  $B_w$  distributions (this study) (open orange circles, G-WB). Point-scale winter balance 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 ( $R^2 = 0.85$ ).

two SP we analyzed near the head of the Kaskawulsh Glacier between 20–21 May 2016. The data show a linear decrease of  $0.024 \text{ m.w.e. km}^{-1}$  ( $R^2 = 0.85$ ) in winter balance with distance from the regional topographic divide between the Kaskawulsh and Hubbard Glaciers, as identified by Taylor-Barge (1969). 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.



506 Inter-annual variability in winter balance is not considered in our study. A number of studies have found  
507 temporal stability in spatial patterns of snow distribution and that statistical models based on topographic  
508 parameters could be applied reliably between years (e.g. Grünwald and others, 2013). For example, Walmsley  
509 (2015) analyzed more than 40 years of winter balance recorded on two Norwegian glaciers and found that  
510 snow distribution is spatially heterogeneous yet exhibits robust temporal stability. Contrary to this, Crochet  
511 and others (2007) found that snow distribution in Iceland differed considerably between years and depended  
512 primarily on the dominant wind direction over the course of a winter. Therefore, multiple years of snow depth  
513 and density measurements, that are not necessarily consecutive, are needed to better understand inter-annual  
514 variability of winter balance within the Donjek Range.

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

521 The lack of correlation between SP- and FS-derived densities needs to be reconciled. Contrary to our  
522 results, most studies that compare SP- and FS-derived densities report minimal discrepancy (e.g. Dixon and  
523 Boon, 2012, and sources within). Additional co-located density measurements are needed to better compare  
524 the two methods of obtaining density values. Comparison with other FS would also be informative. Even  
525 with this limitation, density assignment was, fortunately, not the largest source of uncertainty in estimating  
526 glacier-wide winter balance.

527 Our sampling design was chosen to achieve broad spatial coverage of the ablation area, but is likely too  
528 finely resolved along transects for many mass-balance surveys to replicate. An optimal sampling design would  
529 minimize uncertainty in winter balance while reducing the number of required measurements. Analysis of  
530 the estimated winter balance obtained using subsets of the data is underway to make recommendations on  
531 optimal transect configuration and along-track spacing of measurements. López-Moreno and others (2010)  
532 found that 200–400 observations are needed within a non-glacierized alpine basin (6 km<sup>2</sup>) to obtain accurate  
533 and robust snow distribution models. Similar guidelines would be useful for glacierized environments.

534 In this study, we assume that the subgrid variability of winter balance is uniform across a given glacier.  
535 Contrary to this assumption, McGrath and others (2015) 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 (Zhang and Montgomery, 1994; Garbrecht and Martz, 1994; Guo-an and others, 2001; López-Moreno and others, 2010). The relationship between topographic parameters and winter balance is, therefore, not independent of DEM gridcell size. For example, Kienzle (2004) and López-Moreno and others (2010) 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 LR 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, Fig. S1). 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 ordinary kriging are used to obtain estimates of distributed winter balance ( $b_w$ ). We use Monte Carlo analysis to evaluate the contributions of interpolation, assignment of snow density and grid-scale variability of winter balance to uncertainty in estimates of glacier-wide winter balance ( $B_w$ ).

Values of  $B_w$  estimated using linear regression and ordinary 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  $B_w$  (7% for linear-regression estimates and 34% for ordinary-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  $B_w$  uncertainty is small indicating that extensive subgrid-scale sampling is not required to reduce overall uncertainty.

Our results suggest that processes governing distributed  $b_w$  differ between glaciers, highlighting the importance of regional-scale winter-balance studies. The estimated distribution of  $b_w$  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  $b_w$ . A wind-redistribution parameter is found to be a weak but significant predictor of  $b_w$ , though conflicting relationships between glaciers make it difficult to interpret. The major limitations of our work include the restriction of our data to a single year and minimal sampling in the accumulation area. 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 the methods of data analysis and edited the manuscript. DB contributed to the statistical analysis 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 Michael Grosskopf for assistance with the statistics, including ordinary kriging. Luke Wonneck, Leif Anderson and Jeff Crompton all provided thoughtful and constructive comments on drafts of the manuscript. We appreciate the thoughtful suggestions and edits from two anonymous reviewers.

## REFERENCES

- Anderton S, White S and Alvera B (2004) Evaluation of spatial variability in snow water equivalent for a high mountain catchment. *Hydrological Processes*, **18**(3), 435–453 (doi: 10.1002/hyp.1319)
- Arendt AA, Luthcke SB, Larsen CF, Abdalati W, Krabill WB and Beedle MJ (2008) Validation of high-resolution GRACE mascon estimates of glacier mass changes in the St Elias Mountains, Alaska, USA, using aircraft laser altimetry. *Journal of Glaciology*, **54**(188), 778–787 (doi: 10.3189/002214308787780067)

- 594 Bagos PG and Adam M (2015) On the Covariance of Regression Coefficients. *Open Journal of Statistics*, **5**, 680–701  
595 (doi: 10.4236/ojs.2015.57069)
- 596 Balk B and Elder K (2000) Combining binary decision tree and geostatistical methods to estimate snow distribution  
597 in a mountain watershed. *Water Resources Research*, **36**(1), 13–26 (doi: 10.1029/1999WR900251)
- 598 Barry RG (1992) *Mountain weather and climate*. Cambridge University Press, 3rd edition
- 599 Beaumont RT and Work RA (1963) Snow sampling results from three samplers. *International Association of Scientific*  
600 *Hydrology. Bulletin*, **8**(4), 74–78 (doi: 10.1080/02626666309493359)
- 601 Berthier E, Schiefer E, Clarke GK, Menounos B and Rémy F (2010) Contribution of Alaskan glaciers to sea-level  
602 rise derived from satellite imagery. *Nature Geoscience*, **3**(2), 92–95
- 603 Burgess EW, Forster RR and Larsen CF (2013) Flow velocities of Alaskan glaciers. *Nature communications*, **4**,  
604 2146–2154 (doi: 10.1038/ncomms3146)
- 605 Burnham KP and Anderson DR (2004) Multimodel Inference: Understanding AIC and BIC in Model Selection.  
606 *Sociological Methods & Research*, **33**(2), 261–304 (doi: 10.1177/0049124104268644)
- 607 Carroll T (1977) A comparison of the CRREL 500 cm<sup>3</sup> tube and the ILTS 200 and 100 cm<sup>3</sup> box cutters used for  
608 determining snow densities. *Journal of Glaciology*, **18**(79), 334–337 (doi: 10.1017/S0022143000021420)
- 609 Clark MP, Hendrikx J, Slater AG, Kavetski D, Anderson B, Cullen NJ, Kerr T, Örn Hreinsson E and Woods RA  
610 (2011) Representing spatial variability of snow water equivalent in hydrologic and land-surface models: A review.  
611 *Water Resources Research*, **47**(7) (doi: 10.1029/2011WR010745)
- 612 Clarke GK (2014) A short and somewhat personal history of Yukon glacier studies in the Twentieth Century. *Arctic*,  
613 **37**(1), 1–21
- 614 Clyde GD (1932) Circular No. 99-Utah Snow Sampler and Scales for Measuring Water Content of Snow. *UAES*  
615 *Circulars*, Paper 90
- 616 Cogley J, Hock R, Rasmussen L, Arendt A, Bauder A, Braithwaite R, Jansson P, Kaser G, Möller M, Nicholson L  
617 and others (2011) *Glossary of glacier mass balance and related terms*. UNESCO-IHP, Paris
- 618 Conger SM and McClung DM (2009) Comparison of density cutters for snow profile observations. *Journal of*  
619 *Glaciology*, **55**(189), 163–169 (doi: 10.3189/002214309788609038)
- 620 Crochet P, Jóhannesson T, Jónsson T, Sigurðsson O, Björnsson H, Pálsson F and Barstad I (2007) Estimating  
621 the Spatial Distribution of Precipitation in Iceland Using a Linear Model of Orographic Precipitation. *Journal of*  
622 *Hydrometeorology*, **8**(6), 1285–1306 (doi: 10.1175/2007JHM795.1)
- 623 Crompton JW and Flowers GE (2016) Correlations of suspended sediment size with bedrock lithology and glacier  
624 dynamics. *Annals of Glaciology*, **57**(72), 1–9 (doi: 10.1017/aog.2016.6)

- 625 Cullen NJ, Anderson B, Sirguey P, Stumm D, Mackintosh A, Conway JP, Horgan HJ, Dadić R, Fitzsimons SJ  
626 and Lorrey A (2017) An 11-year record of mass balance of Brewster Glacier, New Zealand, determined using a  
627 geostatistical approach. *Journal of Glaciology*, **63**(238), 199–217 (doi: 10.1017/jog.2016.128)
- 628 Dadić R, Mott R, Lehning M and Burlando P (2010) Parameterization for wind-induced preferential deposition of  
629 snow. *Journal of Geophysical Research: Earth Surface*, **24**(14), 1994–2006 (doi: 10.1029/2009JF001261)
- 630 Danby RK, Hik DS, Slocombe DS and Williams A (2003) Science and the St. Elias: an evolving framework  
631 for sustainability in North America’s highest mountains. *The Geographical Journal*, **169**(3), 191–204 (doi:  
632 10.1111/1475-4959.00084)
- 633 Davis JC and Sampson RJ (1986) *Statistics and data analysis in geology*. Wiley New York et al., 2nd edition
- 634 Deems JS and Painter TH (2006) Lidar measurement of snow depth: accuracy and error sources. In *Proceedings of*  
635 *the International Snow Science Workshop*
- 636 Dixon D and Boon S (2012) Comparison of the SnowHydro snow sampler with existing snow tube designs. *Hydrological*  
637 *Processes*, **26**(17), 2555–2562 (doi: 10.1002/hyp.9317)
- 638 Egli L, Griessinger N and Jonas T (2011) Seasonal development of spatial snow-depth variability across different  
639 scales in the Swiss Alps. *Annals of Glaciology*, **52**(58), 216–222 (doi: 10.3189/172756411797252211)
- 640 Elder K, Dozier J and Michaelsen J (1991) Snow accumulation and distribution in an alpine watershed. *Water*  
641 *Resources Research*, **27**(7), 1541–1552 (doi: 10.1029/91WR00506)
- 642 Elder K, Rosenthal W and Davis RE (1998) Estimating the spatial distribution of snow water equiv-  
643 alence in a montane watershed. *Hydrological Processes*, **12**(1011), 1793–1808 (doi: 10.1002/(SICI)1099-  
644 1085(199808/09)12:10<111793::AID-HYP6953.0.CO;2-K)
- 645 Erxleben J, Elder K and Davis R (2002) Comparison of spatial interpolation methods for estimating snow distribution  
646 in the Colorado Rocky Mountains. *Hydrological Processes*, **16**(18), 3627–3649 (doi: 10.1002/hyp.1239)
- 647 Fames PE, Peterson N, Goodison B and Richards RP (1982) Metrication of Manual Snow Sampling Equipment. In  
648 *Proceedings of the 50th Western Snow Conference*, 120–132
- 649 Fierz C, Armstrong RL, Durand Y, Etchevers P, Greene E, McClung DM, Nishimura K, Satyawali PK and Sokratov  
650 SA (2009) *The international classification for seasonal snow on the ground*. UNESCO/IHP, unesco/ihp paris  
651 edition
- 652 Garbrecht J and Martz L (1994) Grid size dependency of parameters extracted from digital elevation models.  
653 *Computers & Geosciences*, **20**(1), 85–87 (doi: 10.1016/0098-3004(94)90098-1)
- 654 Grabiec M, Puczek D, Budzik T and Gajek G (2011) Snow distribution patterns on Svalbard glaciers derived from  
655 radio-echo soundings. *Polish Polar Research*, **32**(4), 393–421 (doi: 10.2478/v10183-011-0026-4)
- 656 Gray DM and Male DH (1981) *Handbook of snow: principles, processes, management & use*. Pergamon Press, 1st  
657 edition

- Grünewald T, Schirmer M, Mott R and Lehning M (2010) Spatial and temporal variability of snow depth and ablation rates in a small mountain catchment. *Cryosphere*, **4**(2), 215–225 (doi: 10.5194/tc-4-215-2010)
- Grünewald T, Stötter J, Pomeroy J, Dadic R, Moreno Baños I, Marturià J, Spross M, Hopkinson C, Burlando P and Lehning M (2013) Statistical modelling of the snow depth distribution in open alpine terrain. *Hydrology and Earth System Sciences*, **17**(8), 3005–3021 (doi: 10.5194/hess-17-3005-2013)
- Grünewald T, Bühler Y and Lehning M (2014) Elevation dependency of mountain snow depth. *The Cryosphere*, **8**(6), 2381–2394 (doi: 10.5194/tc-8-2381-2014)
- Guo-an T, Yang-he H, Strobl J and Wang-qing L (2001) The impact of resolution on the accuracy of hydrologic data derived from DEMs. *Journal of Geographical Sciences*, **11**(4), 393–401 (doi: 10.1007/BF02837966)
- Gusmeroli A, Wolken GJ and Arendt AA (2014) Helicopter-borne radar imaging of snow cover on and around glaciers in Alaska. *Annals of Glaciology*, **55**(67), 78–88 (doi: 10.3189/2014AoG67A029)
- Hagen JO and Liestøl O (1990) Long-term glacier mass-balance investigations in Svalbard, 1950–88. *Annals of Glaciology*, **14**(1), 102–106
- Helbig N and van Herwijnen A (2017) Subgrid parameterization for snow depth over mountainous terrain from flat field snow depth. *Water Resources Research*, **53**(2), 1444–1456 (doi: 10.1002/2016WR019872)
- Hock R (2005) Glacier melt: a review of processes and their modelling. *Progress in Physical Geography*, **29**(3), 362–391 (doi: 10.1191/0309133305pp453ra)
- Hock R and Jensen H (1999) Application of kriging interpolation for glacier mass balance computations. *Geografiska Annaler: Series A, Physical Geography*, **81**(4), 611–619 (doi: 10.1111/1468-0459.00089)
- Kaser G, Fountain A, Jansson P and others (2003) *A manual for monitoring the mass balance of mountain glaciers*. ICSI/UNESCO
- Kienzle S (2004) The Effect of DEM Raster Resolution on First Order, Second Order and Compound Terrain Derivatives. *Transactions in GIS*, **8**(1), 83–111 (doi: 10.1111/j.1467-9671.2004.00169.x)
- Kinar N and Pomeroy J (2015) Measurement of the physical properties of the snowpack. *Reviews of Geophysics*, **53**(2), 481–544 (doi: 10.1002/2015RG000481)
- Kohavi R and others (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, 1137–1145
- 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). *ISPRS Journal of Photogrammetry and Remote Sensing*, **64**(2), 204–212 (doi: 10.1016/j.isprsjprs.2008.10.005)
- Lehning M, Völksch I, Gustafsson D, Nguyen TA, Stähli M and Zappa M (2006) ALPINE3D: a detailed model of mountain surface processes and its application to snow hydrology. *Hydrological Processes*, **20**(10), 2111–2128 (doi: 10.1002/hyp.6204)

- Li J and Heap AD (2008) A review of spatial interpolation methods for environmental scientists. *Geoscience Australia*, Record 2008/23
- Liston GE and Elder K (2006) A distributed snow-evolution modeling system (SnowModel). *Journal of Hydrometeorology*, **7**(6), 1259–1276 (doi: 10.1175/JHM548.1)
- Liston GE and Sturm M (1998) A snow-transport model for complex terrain. *Journal of Glaciology*, **44**(148), 498–516
- López-Moreno J, Latron J and Lehmann A (2010) Effects of sample and grid size on the accuracy and stability of regression-based snow interpolation methods. *Hydrological Processes*, **24**(14), 1914–1928 (doi: 10.1002/hyp.7564)
- López-Moreno J, Fassnacht S, Heath J, Musselman K, Revuelto J, Latron J, Morán-Tejeda E and Jonas T (2013) Small scale spatial variability of snow density and depth over complex alpine terrain: Implications for estimating snow water equivalent. *Advances in Water Resources*, **55**, 40–52 (doi: 10.1016/j.advwatres.2012.08.010)
- López-Moreno JJ, Fassnacht S, Beguería S and Latron J (2011) Variability of snow depth at the plot scale: implications for mean depth estimation and sampling strategies. *The Cryosphere*, **5**(3), 617–629 (doi: 10.5194/tc-5-617-2011)
- MacDougall AH and Flowers GE (2011) Spatial and temporal transferability of a distributed energy-balance glacier melt model. *Journal of Climate*, **24**(5), 1480–1498 (doi: 10.1175/2010JCLI3821.1)
- Machguth H, Eisen O, Paul F and Hoelzle M (2006) Strong spatial variability of snow accumulation observed with helicopter-borne GPR on two adjacent alpine glaciers. *Geophysical Research Letters*, **33**(13), 1–5 (doi: 10.1029/2006GL026576)
- Madigan D and Raftery AE (1994) Model Selection and Accounting for Model Uncertainty in Graphical Models Using Occam’s Window. *Journal of the American Statistical Association*, **89**(428), 1535–1546
- Marshall HP, Koh G, Sturm M, Johnson J, Demuth M, Landry C, Deems J and Gleason J (2006) Spatial variability of the snowpack: Experiences with measurements at a wide range of length scales with several different high precision instruments. In *Proceedings International Snow Science Workshop*, 359–364
- 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. *Journal of Geophysical Research: Earth Surface*, **120**(8), 1530–1550 (doi: 10.1002/2015JF003539)
- Metropolis N and Ulam S (1949) The Monte Carlo Method. *Journal of the American Statistical Association*, **44**(247), 335–341
- 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. *Hydrological Processes*, **19**(7), 1459–1479 (doi: 10.1002/hyp.5586)
- Mott R, Faure F, Lehning M, Löwe H, Hynek B, Michlmayer G, Prokop A and Schöner W (2008) Simulation of seasonal snow-cover distribution for glacierized sites on Sonnblick, Austria, with the Alpine3D model. *Annals of Glaciology*, **49**(1), 155–160 (doi: 10.3189/172756408787814924)

- Musselman KN, Pomeroy JW, Essery RL and Leroux N (2015) Impact of windflow calculations on simulations of alpine snow accumulation, redistribution and ablation. *Hydrological Processes*, **29**(18), 3983–3999 (doi: 10.1002/hyp.10595)
- Proksch M, Rutter N, Fierz C and Schneebeli M (2016) Intercomparison of snow density measurements: bias, precision, and vertical resolution. *The Cryosphere*, **10**(1), 371–384 (doi: 10.5194/tc-10-371-2016)
- Pulwinski A (2017) *Multi-scale investigation of winter balance on alpine glaciers*. Master’s thesis, Simon Fraser University
- Raftery AE, Madigan D and Hoeting JA (1997) Bayesian Model Averaging for Linear Regression Models. *Journal of the American Statistical Association*, **92**(437), 179–191 (doi: 10.1080/01621459.1997.10473615)
- Rasmussen CE and Williams CK (2006) *Gaussian processes for machine learning*. MIT press Cambridge
- Réveillet M, Vincent C, Six D and Rabatel A (2016) Which empirical model is best suited to simulate glacier mass balances? *Journal of Glaciology*, **63**(237), 1–16 (doi: 10.1017/jog.2016.110)
- Roustant O, Ginsbourger D and Deville Y (2012) DiceKriging, DiceOptim: Two R packages for the analysis of computer experiments by kriging-based metamodeling and optimization. *Journal of Statistical Software*, **21**, 1–55
- Schneiderbauer S and Prokop A (2011) The atmospheric snow-transport model: SnowDrift3D. *Journal of Glaciology*, **57**(203), 526–542 (doi: 10.3189/002214311796905677)
- Schuler TV, Crochet P, Hock R, Jackson M, Barstad I and Jóhannesson T (2008) Distribution of snow accumulation on the Svartisen ice cap, Norway, assessed by a model of orographic precipitation. *Hydrological Processes*, **22**(19), 3998–4008 (doi: 10.1002/hyp.7073)
- Scipión DE, Mott R, Lehning M, Schneebeli M and Berne A (2013) Seasonal small-scale spatial variability in alpine snowfall and snow accumulation. *Water Resources Research*, **49**(3), 1446–1457, ISSN 1944-7973 (doi: 10.1002/wrcr.20135)
- Shea C and Jamieson B (2010) Star: an efficient snow point-sampling method. *Annals of Glaciology*, **51**(54), 64–72 (doi: 10.3189/172756410791386463)
- Sold L, Huss M, Hoelzle M, Anderegg H, Joerg PC and Zemp M (2013) Methodological approaches to infer end-of-winter snow distribution on alpine glaciers. *Journal of Glaciology*, **59**(218), 1047–1059 (doi: 10.3189/2013JoG13J015)
- Tangborn WV, Krimmel RM and Meier MF (1975) A comparison of glacier mass balance by glaciological, hydrological and mapping methods, South Cascade Glacier, Washington. *International Association of Hydrological Sciences Publication*, **104**, 185–196
- Taylor-Barge B (1969) *The summer climate of the St. Elias Mountain region*. Montreal: Arctic Institute of North America, Research Paper No. 53



- Thibert E, Blanc R, Vincent C and Eckert N (2008) Glaciological and volumetric mass-balance measurements: error analysis over 51 years for Glacier de Sarennes, French Alps. *Journal of Glaciology*, **54**(186), 522–532
- Trujillo E and Lehning M (2015) Theoretical analysis of errors when estimating snow distribution through point measurements. *The Cryosphere*, **9**(3), 1249–1264 (doi: 10.5194/tc-9-1249-2015)
- Turcan J and Loijens H (1975) Accuracy of snow survey data and errors in snow sampler measurements. In *32nd Eastern Snow Conference*
- Tveit J and Killingtveit Å (1994) Snow surveys for studies of water budget on Svalbard 1991–1994. In *Proceedings of the 10th International Northern Research Basins Symposium and Workshop, Spitsbergen, Norway. SINTEF Report*, volume 22, A96415
- Waechter A, Copland L and Herdes E (2015) Modern glacier velocities across the Icefield Ranges, St Elias Mountains, and variability at selected glaciers from 1959 to 2012. *Journal of Glaciology*, **61**(228), 624–634 (doi: 10.3189/2015JoG14J147)
- Walmsley APU (2015) *Long-term observations of snow spatial distributions at Hellstugubreen and Gråsubreen, Norway*. Master’s thesis, University of Oslo
- Wetlaufer K, Hendrikx J and Marshall L (2016) Spatial Heterogeneity of Snow Density and Its Influence on Snow Water Equivalence Estimates in a Large Mountainous Basin. *Hydrology*, **3**(3), 1–17 (doi: 10.3390/hydrology3010003)
- Wilson N and Flowers G (2013) Environmental controls on the thermal structure of alpine glaciers. *The Cryosphere*, **7**(1), 167–182 (doi: 10.5194/tc-7-167-2013)
- Wilson NJ, Flowers GE and Mingo L (2013) Comparison of thermal structure and evolution between neighboring subarctic glaciers. *Journal of Geophysical Research: Earth Surface*, **118**(3), 1443–1459 (doi: 10.1002/jgrf.20096)
- Winstral A, Elder K and Davis RE (2002) Spatial snow modeling of wind-redistributed snow using terrain-based parameters. *Journal of Hydrometeorology*, **3**(5), 524–538 (doi: 10.1175/1525-7541(2002)0030524:SSMOWR2.0.CO;2)
- Winther J, Bruland O, Sand K, Killingtveit A and Marechal D (1998) Snow accumulation distribution on Spitsbergen, Svalbard, in 1997. *Polar Research*, **17**, 155–164 (doi: 10.3402/polar.v17i2.6616)
- Woo MK and Marsh P (1978) Analysis of Error in the Determination of Snow Storage for Small High Arctic Basins. *Journal of Applied Meteorology*, **17**(10), 1537–1541 (doi: 10.1175/1520-0450(1978)0171537:AOEITD2.0.CO;2)
- Wood WA (1948) Project “Snow Cornice”: the establishment of the Seward Glacial research station. *Arctic*, **1**(2), 107–112
- Work R, Stockwell H, Freeman T and Beaumont R (1965) *Accuracy of field snow surveys*. Cold Regions Research & Engineering Laboratory
- Zhang W and Montgomery DR (1994) Digital elevation model grid size, landscape representation, and hydrologic simulations. *Water Resources Research*, **30**(4), 1019–1028 (doi: 10.1029/93WR03553)