Uncertainties in estimating winter balance 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 (WB) on glaciers is central to assessing glacier health and predicting glacier runoff. However, measuring and modelling snow distribution is inherently difficult in mountainous terrain, resulting in high uncertainties in estimates of WB. Our work focuses on uncertainty attribution within the process of converting direct measurements of snow depth and density to estimates of WB. We collected more than 9000 direct measurements of snow depth across three glaciers in the St. Elias Mountains, Yukon, Canada in May 2016. Linear regression (LR) and simple kriging (SK), combined with cross correlation and Bayesian model averaging, are used to interpolate point-scale WB estimates. Snow distribution patterns differ considerably between glaciers, highlighting strong inter- and intra-basin variability. Elevation is found to be the dominant control of the spatial distribution of WB, but the relationship varies considerably between glaciers. A simple parameterization of wind redistribution is also a small but statistically significant predictor of point-scale WB. Through a Monte Carlo analysis, we find that the interpolation of estimated values of WB is a larger source of uncertainty than the assignment of snow density or than the representation of WB value within a terrain model grid cell. For our study glaciers, the total WB uncertainty ranges from 0.03 (8%) to 0.15 (54%) m w.e. depending primarily on the interpolation method. Despite the challenges associated with accurately and precisely estimating WB, our results are consistent with the previously reported regional WB gradient. (244 words)

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. Accurate estimation of winter surface mass balance is critical for correctly simulating the summer and overall mass balance of a glacier (e.g. Hock, 2005). Effectively representing the spatial distribution of snow is also important for simulating snow and ice melt as well as energy and mass exchange between the land and atmosphere, which allows for better monitoring of surface runoff and its downstream effects (e.g. Clark and others, 2011). Snow distribution is sensitive to a number of complex process that partially depend on glacier location, topography, and orientation (e.g. Blöschl and others, 1991; Mott and others, 2008; Clark and others, 2011; Sold and others, 2013). Current models are not able to fully represent these processes so there is a significant source of uncertainty that undermines the ability of models to represent current and projected glacier conditions (Réveillet and others, 2016).

Winter balance is notoriously difficult to estimate. Snow distribution in alpine regions is highly variable and 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). As a result, snow distribution is highly variable with short correlation length scales (e.g. Anderton and others, 2004; Egli and others, 2011; Grunewald and others, 2010; Helbig and van Herwijnen, 2017; López-Moreno and others, 2011, 2013; Machguth and others, 2006; Marshall and others, 2006). As a result, extensive, high resolution and accurate snow distribution measurements on glaciers are almost impossible to achieve (e.g. Cogley and others, 2011; McGrath and others, 2015).

Those studies that have focused on obtaining detailed estimates of winter balance have used a wide range of techniques to measure snow water equivalent, including direct measurement of snow depth and density (e.g. Cullen and others, 2017), lidar/photogrammerty (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 measurements is often limited and often consists of an elevation transect along the glacier centreline (e.g. Kaser and others, 2003; Machguth and others, 2006). Interpolation of these measurements is primarily done with a linear regression that includes only a few topographic parameters (e.g. MacDougall and Flowers, 2011), with elevation being the most common. Other established techniques include hand contouring (e.g.

Tangborn and others, 1975), kriging (e.g. Hock and Jensen, 1999) and attributing measured accumulation values to elevation bands (e.g. Thibert and others, 2008). Physical snow models have been applied on only a few glaciers (e.g. Mott and others, 2008; Dadic and others, 2010) but a lack of detailed meteorological data generally prohibits their wide spread application. Error analysis is rarely undertaken and to our knowledge, no studies have thoroughly investigated uncertainty in spatially distributed estimates of winter balance estimates.

More sophisticated models and measurement techniques of snow distribution are available and 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 and ensure that all terrain types are sampled. A wide array of measurement interpolation methods are used, including linear (e.g. López-Moreno and others, 2010) and non-linear regressions (e.g. Molotch and others, 2005) that include numerous terrain parameters as well as geospatial interpolation (e.g. Erxleben and others, 2002) such as kriging, and interpolation methods are often combined (e.g. regression kriging) to yield improved fit (e.g. Balk and Elder, 2000). Physical snow models such as Alpine3D (Lehning and others, 2006) and SnowDrift3D (Schneiderbauer and Prokop, 2011) are continuously being improved and tested within the snow science literature. Error analysis has been considered from both a theoretical (e.g. Trujillo and Lehning, 2015) and applied perspective (e.g. Turcan and Loijens, 1975; Woo and Marsh, 1978; Deems and Painter, 2006).

The precision and accuracy of winter balance estimates can likely be improved by incorporating more sophisticated tools and interpolation methodologies, and by gaining a more comprehensive understanding of inherent uncertainties. The overall goals of our work are to (1) critically examine methods of moving from direct snow depth and density measurements to estimating winter balance and to (2) identify sources of uncertainty, evaluate their magnitude and assess their combined contribution to uncertainty in winter balance. We focus on commonly applied low-complexity methods of measuring and estimating winter balance with the hope of making our results broadly applicable.

STUDY SITE

Winter balance surveys were conducted on three glaciers in the Donjek Range of the St. Elias Mountains, located in south western Yukon, Canada (Fig. 1, Table 1). The Donjek Range is approximately 30 × 30 km and Glacier 4, Glacier 2 and Glacier 13 (labelling adopted from Crompton and Flowers (2016)) are located along a SW-NE transect through the range. These small, polythermal alpine glaciers are generally oriented SE-NW, with Glacier 4 predominantly southeast facing and Glaciers 2 and 13 generally northwest facing. The glaciers have simple geometries and have steep head and valley walls. The St. Elias Mountains 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 (Taylor-Barge, 1969). The boarder between the two climatic zones is generally aliged with the divide between Hubbard and Kaskawulsh Glaciers, approximately 13 km from the ocean. The Donjek Range is located approximately 40 km to the east of the divide between the Hubbard and Kaskawulsh Glaciers (Taylor-Barge, 1969). Research on snow distribution and glacier mass balance in this area is limited. A series of research programs were operational in the 1960s (Wood, 1948; Danby and others, 2003) and long-term studies on a few alpine glaciers have arisen in the last 30 years (e.g. Flowers and others, 2014).

METHODS

Estimating glacier midwinter balance involves transforming measurements of snow depth and density into distributed estimates of snow water equivalent (SWE). We do this in four steps: (1) We obtain direct measurements of snow depth and density in the field. (2) We interpolate density measurements to all depth-measurement locations in order to calculate the SWE at each of these locations. This is necessary because we measure density at relatively few locations of depth. (3) We average all SWE values within each gridcell of a digital elevation model (DEM). (4) We interpolate and extrapolate these gridcell averaged SWE values to obtain a spatially distributed winter balance (in m w.e.) across the glacier surface. We choose to use a linear regression between gridcell average SWE and topographic parameters, as well as simple kriging for this process of interpolation and extrapolation. The specific winter balance is then calculated as the areally-averaged SWE. For brevity, we refer to these four steps as (1) field measurements, (2) distributed snow density, (3) gridcell average SWE and (4) distributed SWE. 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 spatial scales (Clark and others, 2011). To capture variability at the regional scale we choose three glaciers along the precipitation gradient in the St. Elias Mountains, Yukon (Fig. 1b) (Taylor-Barge, 1969). To account for basin-scale variability, snow depth was measured along linear and curvilinear transects on each glacier (Fig. 1c) with sample spacing of 10 − 60 m (Fig. 1d). Sample spacing was restricted by glacier travel and the need to complete surveys on all three glaciers within the period of peak accumulation. We selected centreline and transverse transects because they are commonly used for winter balance estimates (e.g. Kaser and others, 2003; Machguth and others, 2006) as well as an hourglass pattern with an inscribed circle, which allows for sampling in multiple directions and easy travel (Parr, C., 2016 personal communication). To capture point-scale variability, we took 3–4 depth measurements within ∼1 m of each other (Fig. 1e) at each transect measurement

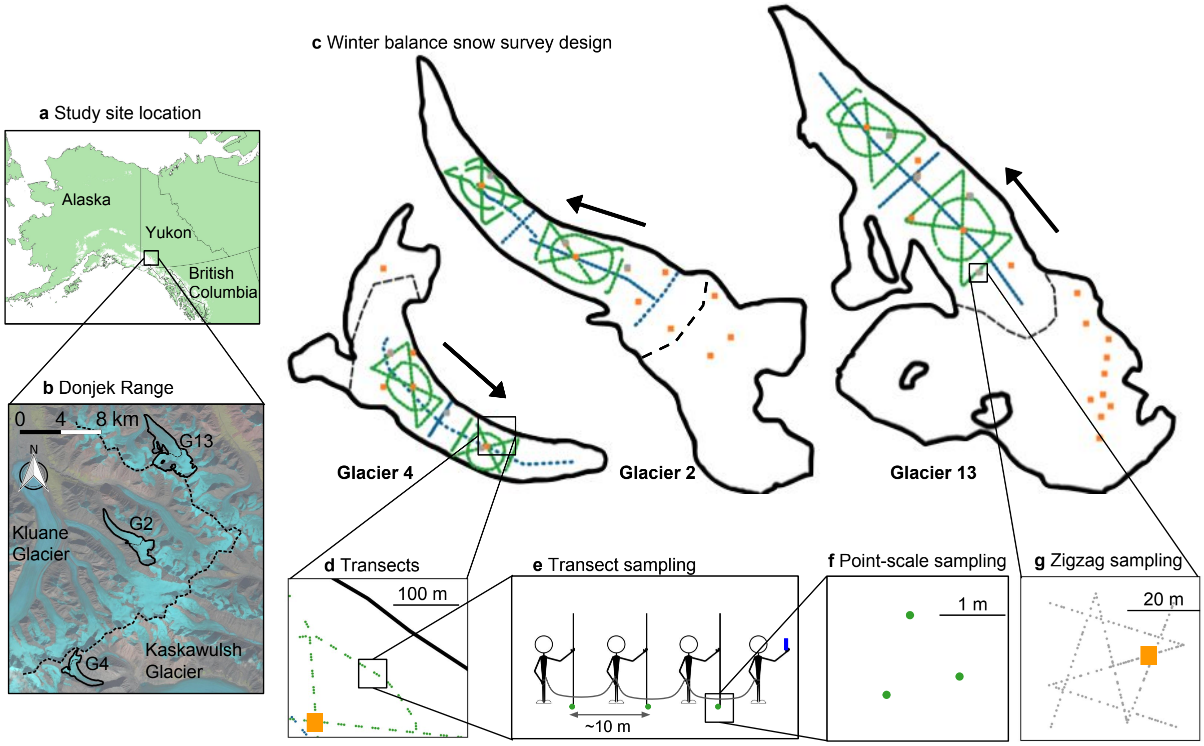

Fig. 1. Study area location and sampling design for Glaciers 4, 2 and 13. (a) The study region is located in the Donjek Range of the St. Elias Mountains of Yukon, Canada. (b) Study glaciers are located along a SW-NE transect through the Donjek Range. The local topographic divide is shown as a dashed line. Imagery from Landsat8 (5 September 2013, data available from the U.S. Geological Survey). (c) Details of the snow survey sampling design. Centreline and transverse transects are shown in blue dots, hourglass and circle design are shown in green dots. Orange squares are locations of snow density measurements. Arrows indicate glacier flow direction and the approximate location of each ELA is shown as a black dashed line. (d) Linear and curvilinear transects typically consist of sets of three measurement locations, (e) spaced ∼10 m apart. (f) At each location, three snow-depth measurements are made. (f) Linear-random snow-depth measurements in 'zigzag' design are shown as grey dots.

location. To capture variability at the gridcell scale, we densely sample up to four gridcells on each glacier using a linear-random sampling design termed 'zigzag'. In total, we collected more than 9000 snow depth measurements throughout the study area (Table 1).

Snow depth: transects

SWE can be estimated as the product of the snow depth and depth-averaged density. Snow depth is generally accepted to be more variable than density (Elder and others, 1991; Clark and others, 2011; López-Moreno and others, 2013) so we chose a sampling design that resulted in a ratio of approximately 55:1 snow depth to snow density measurements. Our sampling campaign involved four people and occurred between 5–15 May, 2016, which corresponds to the historical peak seasonal snow accumulation in Yukon (Yukon Snow Survey Bulletin and Water Supply Forecast, May 1, 2016). While roped-up for glacier travel at 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 aluminium avalanche probes to make snow depth measurements. The location of each set of depth measurements, taken by the second, third and fourth observers, was approximated based on the recorded location of the first observer 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, especially when using an avalanche probe, can be difficult to detect and often misinterpreted (Grunewald and others, 2010; Sold and others, 2013). We intended to use a firn corer to measure SWE 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 to unambiguously identify the snow–firn transition.

Snow depth: zigzags

To capture variability within a single DEM gridcell, we implemented a linear-random sampling design (Shea and Jamieson, 2010), termed 'zigzag'. We measured depth at random intervals ($0.3-3.0$ m) along two 'Z'-shaped transects within three to four 40×40 m gridcells (Fig. 1g) resulting in $135-191$ measurements in each zigzag. Zigzag locations were randomly chosen within the upper, middle, and lower portions of the ablation area of each glacier. We

Table 1. Physical characteristics of study glaciers and May 2016 winter balance survey details for Glacier 4 (G4), Glacier 2 (G2), and Glacier 13 (G13), including number of snow-depth measurement locations along transects (n_T), total length of transects (d_T), number of combined snow pit (SP) and Federal Sampler (FS) density measurement locations (n_ρ) and number of zigzag surveys (n_zz).

	Location	**Elevation (m a.s.l)**			**Slope (°)**	**Area**	**Date**	**Survey Details**			
	UTM Zone 7	*Mean*	*Range*	*ELA*	*Mean*	**(km)**		n_T	d_T (km)	n_ρ	n_zz
G4	595470 E 6740730 N	2344	1958–2809	∼2500	12.8	3.8	4–7 May 2016	649	13.1	7	3
G2	601160 E 6753785 N	2495	1899–3103	∼2500	13.0	7.0	8–11 May 2016	762	13.6	7	3
G13	604602 E 6763400 N	2428	1923–3067	∼2380	13.4	12.6	12–15 May 2016	941	18.1	19	4

were able to measure a fourth zigzag on Glacier 13 that was located in the central ablation area (∼2200 m a.s.l.).

Snow density

Snow density was measured using a wedge cutter in three snow pits on each glacier as well as a using a Federal Sampler. Within the snow pit (SP), we measured a vertical density profile by inserting a 5 × 10 × 10 cm wedge-shaped cutter (250 cm^3) in 5 cm increments and then weighing the samples with a spring scale (e.g. Gray and Male, 1981; Fierz and others, 2009). Uncertainty in estimating density from snow pits stems from incorrect assignment of density to layers that could not be sampled (i.e. ice lenses and hard layers). We attempt to quantify this uncertainty by varying three values: ice layer thickness by ±1 cm (≤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. Density values for shallow pits 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) (Clyde, 1932), which directly measures depth-integrated SWE, was used to augment the snow pit 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 each snow pit profile. Measurements where the snow core length inside the FS was less than 90% of the snow depth were discarded. Density values at each measurement location were then averaged and error is taken to be the standard deviation of these measurements.

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.

Table 2. Eight methods used to estimate snow density at unmeasured locations for purpose of converting measured snow depth to SWE.

Method code	**Source of measured snow density** *snow pit*	*Federal Sampler*	**Density assignment method**
S1	■		Mean of measurements across all glaciers
F1		■	
S2	■		Mean of measurements within a given glacier
F2		■	
S3	■		LR of density on elevation within a given glacier
F3		■	
S4	■		Inverse distance weighted mean
F4		■	

Distributed snow density

Measured snow density must be interpolated or extrapolated to estimate SWE at each snow-depth sampling location. We employ four separate methods that are commonly used to interpolate density (Table 2): (1) calculating mean density over an entire mountain range (e.g. Cullen and others, 2017), (2) calculating 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) inverse-distance weighted density (e.g. Molotch and others, 2005). SP- and FS-derived densities are treated separately, for reasons explained below, resulting in eight possible methods of assigning density.

Gridcell average SWE

We average SWE values within each 40×40 m DEM gridcell. Each measured gridcell contains one to six measurements (mean of 2.1 measurements) that are averaged to give SWE gridcell values. The locations of measurements have considerable uncertainty both from the error in the horizontal position given by the GPS unit ($2.7-4.6$ m) and the estimation of observer location based on the recorded GPS positions of the navigator. These errors could result in the incorrect assignment of a SWE measurement to a particular gridcell. However, this source of error is not further investigated because we assume that SWE uncertainty is captured in the zigzag measurements described below. We are able to combine data from different observers because there are no significant differences between snow depth measurements made by observers along a transect ($p>0.05$), with the exception of the first transect on Glacier 4. No corrections to the data based on observer differences are therefore applied.

Distributed SWE

Linear regression

SWE values as determined above are interpolated and extrapolated across each glacier using linear regression (LR) as well as simple kriging (SK). We use LRs to relate observed SWE to gridcell values of DEM-derived topographic parameters, which include elevation, distance from centreline, slope, aspect, curvature, "northness" and a wind-redistribution parameter (e.g. McGrath and others, 2015). 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 for the elevation range on Glacier 2 which is 50%). Topographic parameters are weighted by a set of fitted regression coefficients (β_i) calculated by minimizing the sum of squares of the vertical deviations of each data point from the regression line (Davis and Sampson, 1986). For details on data and methods used to estimate the topographic parameters see the Supplementary Material.

To avoid overfitting the data, cross-validation and model averaging are implemented. First, cross-validation is used to obtain a set of β_i values that have greater predictive ability. We select 1000 random subsets (2/3 values) of the data to fit the LR and the remaining data (1/3 values) are used to calculate a root mean squared error (RMSE) (Kohavi and others, 1995). Regression coefficients resulting in the lowest RMSE are selected. Second, we use model averaging to take into account uncertainty when selecting predictors and to also maximize predictive ability (Madigan and Raftery, 1994). Models are generated by calculating a set of β_i for all possible combinations of predictors. Following a Bayesian framework, model averaging involves weighting all models by their posterior model probabilities (Raftery and others, 1997). 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 Bayesian Information Criterion (BIC) value (Burnham and Anderson, 2004). BIC penalizes more complex models, which further reduces the risk of overfitting. Spatially distributed SWE is then estimated by applying the resulting regression coefficients to the topographic parameters associated with each gridcell. Specific winter balance is calculated as the areally-averaged, integrated SWE for each glacier (m w.e.).

Simple kriging

Simple kriging (SK) is a data-driven method of estimating values at unsampled locations by using the isotropic spatial correlation (covariance) of measured values to find a set of optimal weights (Davis and Sampson, 1986; Li and Heap, 2008). SK assumes spatial correlation between sampling points that are distributed across a surface and then applies the correlation to interpolate between sampling points. We used the `DiceKriging` R package (Roustant and others, 2012) to calculate the maximum likelihood covariance matrix, as well as range distance (θ) and nugget for gridcell-averaged winter balance values. 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 (Li and Heap, 2008). SK cannot be used to understand physical processes that may be controlling snow distribution and in the absence of data, cannot be used to estimate winter balance on an unmeasured, neighbouring glacier.

Uncertainty analysis

To quantify the uncertainty on the estimated glacier-wide winter balance, we conduct a Monte Carlo analysis, which uses repeated random sampling of input variables to calculate a distribution of output variables (Metropolis and Ulam, 1949). This random sampling process is done 1000 times, resulting in a distribution of possible glacier-wide winter balance values based on uncertainties associated with the four steps outlined above. We use the standard deviation of the distribution as a useful metric of uncertainty of the winter balance. Three sources of uncertainty are considered separately: (1) gridcell uncertainty, (2) density uncertainty and (3) interpolation uncertainty. These individual sources of uncertainty are propagated through the conversion of snow depth and density measurements to winter balance. Finally, the cumulative effect of all three sources of uncertainty on the winter balance is quantified.

Gridcell uncertainty ($\sigma_{\rm GS}$)

To estimate glacier-wide winter balance, we make use of the grid-scale zigzag surveys to represent the uncertainty in estimating the gridcell-average SWE. For simplicity, we assume uniform SWE uncertainty between gridcells for each glacier and represent this uncertainty by a normal distribution. The normal distribution is centred at zero and has a standard deviation equal to the mean standard deviation of all zigzags on each glacier. For each iteration of the Monte Carlo, a set of SWE values is randomly chosen from the distribution and added to the original SWE values. These perturbed SWE values are then used in the interpolation. Uncertainty in the winter balance due to uncertainty in estimating the gridcell SWE ($\sigma_{\rm GS}$) is represented as the standard deviation of the resulting distribution of winter balance estimates.

Density uncertainty (σ_ρ)

We incorporate uncertainty in interpolating density measurements by carrying forward all eight density interpolation methods when estimating winter balance. Using multiple density interpolation methods results in a generous estimate of density uncertainty. The density measurement and interpolation methods used in our study encompass a broad spectrum of possible density values. The winter balance uncertainty due to density uncertainty (σ_ρ) is calculated as the standard deviation of glacier-wide winter balance estimates calculated using each density interpolation method.

Interpolation uncertainty (σ_{INT})

We represent the uncertainty due to interpolation of SWE values to observed data in different ways for LR and SK. LR 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 Bagos and Adam (2015). The β_i distributions are randomly sampled and used to estimate winter balance.

SK uncertainty is represented by the 95% confidence interval for SWE in each gridcell generated by the `DiceKriging` package. The standard deviation of SWE in each gridcell is then calculated and the standard deviation of glacier-wide winter balance is found by taking the square root of the average variance in each gridcell. The final distribution of glacier-wide winter balance values is centred at the SK winter balance estimate and has a standard deviation equal to the glacier-wide standard deviation. For consistency, the standard deviation of winter balance values that result from either LR or SK interpolation uncertainty is referred to as σ_{INT}.

RESULTS AND DISCUSSION

Field measurements

Snow depth

We observed a wide range of snow depth on all three study glaciers (Fig. 2). Glacier 4 has the highest mean snow depth and a high proportion of outliers, indicating a more variable snow depth overall. At each measurement location, the median range of measured depths ($3-4$ points) as a percent of the mean depth at that location is 2%, 11%, and 12%, for Glaciers 4, 2 and 13, respectively. The average standard deviation of all zigzags on Glacier 4 is $\sigma_{\mathrm{G4ZZ}} = 0.027$ m w.e., on Glacier 2 is $\sigma_{\mathrm{G2ZZ}} = 0.035$ m w.e. and on Glacier 13 is $\sigma_{\mathrm{G13ZZ}} = 0.040$ m w.e. SWE measurements for each zigzag are not normally distributed about the mean SWE (Fig. 3).

Snow density

The standard deviation of glacier-wide mean density is less than 10% of the mean density.

The mean SP densities are within one standard deviation between glaciers, whereas mean FS densities are not.

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 (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. Work and others, 1965; Fames and others, 1982; Conger and McClung, 2009). Studies that use Federal Samplers often apply a 10% correction to all measurements for this reason (e.g. Molotch and others, 2005). Oversampling has been attributed to slots "shaving" snow into the tube as it is rotated (e.g. Dixon and Boon, 2012) and to snow falling into the slots, particularly for snow

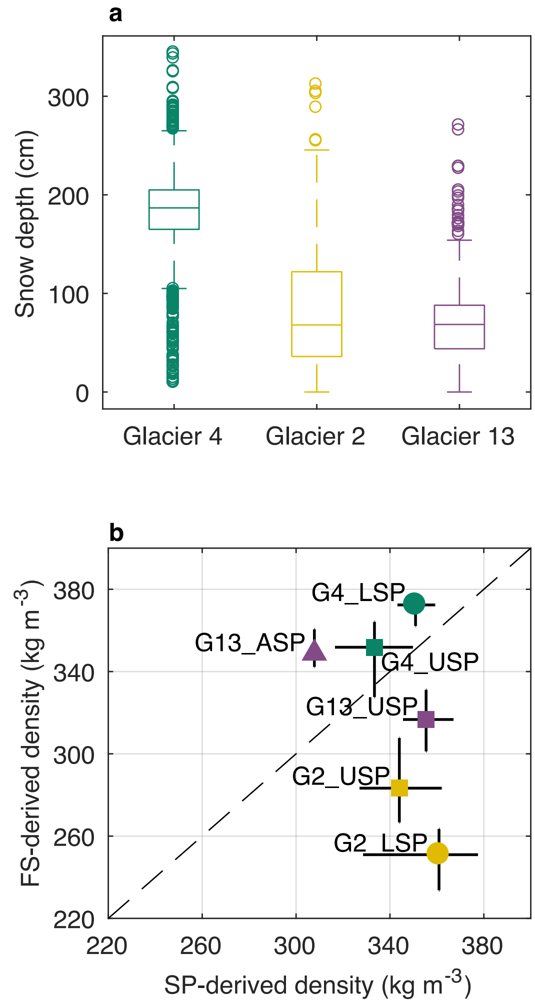

Fig. 2. Snow depth and density data. (a) Boxplot of measured snow depth on Glaciers 4, 2 and 13. The box shows first quartiles, the line within the box indicates the median, bars indicate minimum and maximum values (excluding outliers) and circles show outliers, which are defined as being outside of the range of 1.5 times the quartiles (approximately $\pm 2.7\sigma$). (b) Comparison of integrated density estimated using a vertical profile sampled in 5 cm increments using a wedge cutter in a snow pit (SP) and density estimated using Federal Sampler measurements (FS) 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 are determined differently for SP and FS densities (see text).

samples with densities >400 kg m^{-3} and snow depths >1 m (e.g. Beaumont and Work, 1963). Undersampling is likely to occur due to snow falling out of the bottom of the sampler (Turcan and Loijens, 1975), which likely occurred in our study since a large portion of the lower elevation snow on both Glaciers 2 and 13 was melt affected and weak, allowing for easier lateral displacement of the snow as the sampler was extracted. Relatively poor FS spring-scale sensitivity also made it difficult to obtain accurate measurements for snow depth <20 cm.

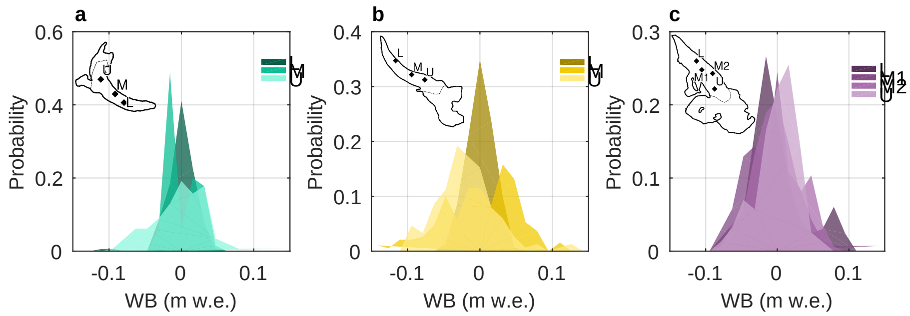

Fig. 3. Distributions of estimated SWE values for each zigzag survey. Local mean has been subtracted. (a) Glacier 4 zigzag surveys. (b) Glacier 2 zigzag surveys. (c) Glacier 13 zigzag surveys. ZZigzag locations shown in insets in lower (L), middle (M, M1, M2) and upper (U) ablation areas.

Additionally, FS density values are positively correlated with snow depth ($R^2 = 0.59$, $p<0.01$). This positive relationship could be a result of physical processes, such as compaction, but is more likely a result of measurement artefacts for a number of reasons. First, the range of densities measured by the Federal Sampler is large (227–431 kg m^{-3}) and the extreme values seem unlikely given the conditions of our study region at the time of sampling, which experiences a continental snow pack with minimal mid-winter melt events. Second, compaction effects of a magnitude able of explaining density differences between SP and FS would not be expected at the measured depths (up to 340 cm). Third, no linear relationship exists between depth and SP-derived density ($R^2 = 0.05$). Together, these findings indicate that the FS measurements have a bias which is challenging to correct for.

Distributed snow density

Since we find no correlation between co-located SP and FS densities (Fig. 2), SP- and FS-derived densities are used separately (Table 2). SP-derived regional (S1) and glacier mean (S2) densities are within one standard deviation of FS-derived densities (F1 and F2) although SP-derived density values are larger (see Supplementary Material, Table 6). Density gradient with elevation differs when using SP (S3) versus FS (F3) densities (see Supplementary Material Table, 6) and density-elevation correlations are generally low. We adopt glacier-wide mean of SP-derived densities (S2) for our reference case. This is consistent with most winter balance studies, which assume a uniform density for individual glaciers and measure snow density profiles at multiple locations in a study basin (e.g. Elder and others, 1991; McGrath and others, 2015; Cullen and others, 2017). However, S2 density interpolation method does not account for known basin-scale spatial variability in snow density (e.g. **?**).

Gridcell-averaged SWE

The distributions of gridcell SWE values for the individual glaciers are similar to those in Fig. 2a but with fewer outliers. The zigzag standard deviation is almost twice as large as the mean standard deviation within gridcells measured along transects. However, a small number of gridcells sampled in transect surveys have standard deviations in SWE that exceed 0.25 m w.e. We nevertheless assume that the gridcell uncertainty is captured with dense sampling in zigzag gridcells. There is therefore little need to take multiple measurements within a gridcell along a transect. As a result, along-track transect spacing can be decreased to allow for greater spatial coverage of transects to better capture basin-scale variability.

Table 3. Specific 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 estimated and averaged gridcell SWE values that were randomly selected and excluded from interpolation (1/3 of all data), is also shown. RMSE as a percent of the glacier-wide WB is shown in brackets.

	Linear regression		**Simple kriging**	
	WB	RMSE	WB	RMSE
G4	0.582	0.153 (26%)	0.616	0.134 (22%)
G2	0.577	0.102 (18%)	0.367	0.073 (20%)
G13	0.381	0.080 (21%)	0.271	0.068 (25%)

Distributed SWE

Linear Regression

Analysis of topographic parameters reveals that elevation (z) is the most significant predictor of SWE for Glaciers 2 and 13, while wind distribution (Sx) is the most significant predictor for Glacier 4 (Fig. 5). SWE is positively correlated with elevation. It is possible that the elevation correlation was accentuated due to melt onset (1–2 week early), especially on Glacier 13 (Yukon Snow Survey Bulletin and Water Supply Forecast, May 1, 2016). The southwestern Yukon winter snow pack in 2016 was also well below average, possibly emphasizing effects of early melt onset. Many winter balance studies have found elevation to be the most significant predictor of SWE (e.g. Machguth and others, 2006; McGrath and others, 2015). However, SWE–elevation gradients vary considerably between glaciers (e.g.

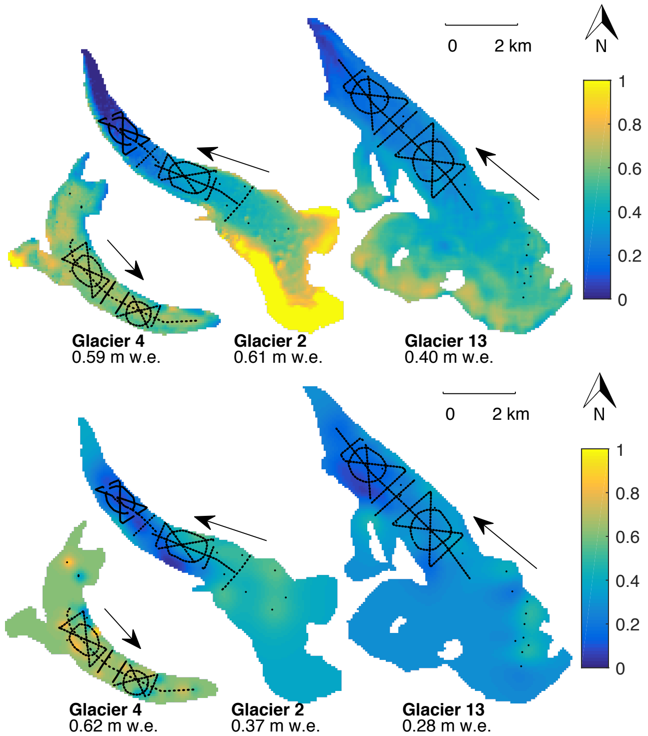

Fig. 4. Spatial distribution of SWE estimated using linear regression (top row) and simple kriging (bottom row). Locations of point-scale winter balance values are shown as black dots. Gridcell-averaged winter balance values are calculated using glacier-wide mean SP-derived densities (S2, Table 2). Glacier flow directions are indicated by arrows. Values of specific winter balance given below labels.

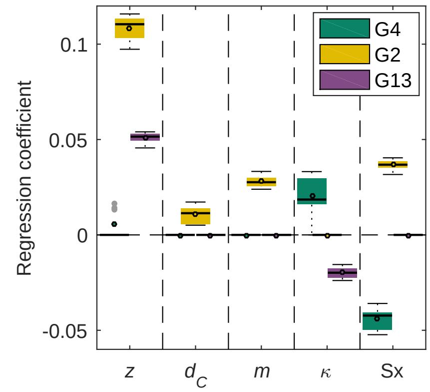

Fig. 5. Distribution of coefficients determined by linear regression of SWE on gridcell topographic parameters for the eight different density interpolation methods (Table 2) on each study glaciers. Coefficients are calculated using normalized data resulting in directly comparable coefficient values. Topographic parameters include elevation (z), distance from centreline (d_C), slope (m) , curvature (κ) and wind exposure (Sx). Regression coefficients that are not significant are assigned a value of zero. Aspect (α) and "northeness" (N) are not shown because coefficient values are zero for all interpolation methods. The box shows first quartiles, the line within the box indicates the median, circle within the box indicated mean, bars indicate minimum and maximum values (excluding outliers) and gray dots show outliers, which are defined as being outside of the range of 1.5 times the quartiles (approximately $\pm 2.7\sigma$).

Winther and others, 1998) and other factors, such as glacier orientation relative to dominant wind direction and glacier shape, have been noted to affect SWE distribution (Machguth and others, 2006; Grabiec and others, 2011). There are also a number of studies that find no significant correlation between SWE on glaciers and topographic parameters, with highly variable distribution of snow was attributed to complex interactions between topography and the atmosphere (e.g. Grabiec and others, 2011; López-Moreno and others, 2011). Linearly extrapolating relationships into unmeasured locations, especially the accumulation area, is most susceptible to large errors (Fig. 7). This area typically also has the highest SWE values (Fig. 4), affecting the specific winter balance estimated for the glacier. In our study, the dependence of SWE on elevation results in $\sim$1% of the area of Glacier 2 with SWE estimates >1.5 m w.e.

SWE is negatively correlated with Sx on Glacier 4, counter-intuitively indicating less snow in 'sheltered' areas, while SWE is positively correlated with Sx on Glaciers 2 and 13. Similarly, SWE is positively correlated with curvature for Glacier 4 and negatively correlated for the other two glaciers. Wind redistribution and preferential deposition of snow is known to have a large influence on accumulation at sub-basin scales (e.g. Dadic and others, 2010; Winstral and others, 2013; Gerber and others, 2017). Our results indicate that wind likely has an impact on snow distribution but that the wind redistribution parameter may not adequately represent wind effects as applied to our study glaciers. For example, Glacier 4 is located in a curved valley with steep side walls, so specifying a single cardinal direction for wind may not b adequate. Further, the scale of deposition may be smaller than the resolution of the Sx parameter estimated from our DEM. Our results corroborate McGrath and others (2015), who undertook a winter-balance study on six glaciers in Alaska (DEM resolutions of 5 m) and found that Sx was the only other significant parameter, besides elevation, for all glaciers. Sx regression coefficients were smaller than elevation regression coefficients and in some cases, negative. Sublimation from blowing snow has also been shown to be an important mechanism mass loss from ridges (e.g. Musselman and others, 2015). Incorporating snow loss, as well as redistribution and preferential deposition, may be needed for accurate representations of distributed SWE.

While a LR have been used to predict distributed SWE in other basins, we find that transfer of LR coefficients between glaciers results in large estimation error. The lowest root mean squared error (0.21 m w.e.) results from estimating a LR using all available observations. Our results are consistent with Grünewald and others (2013), who found that local statistical models are able to perform relatively

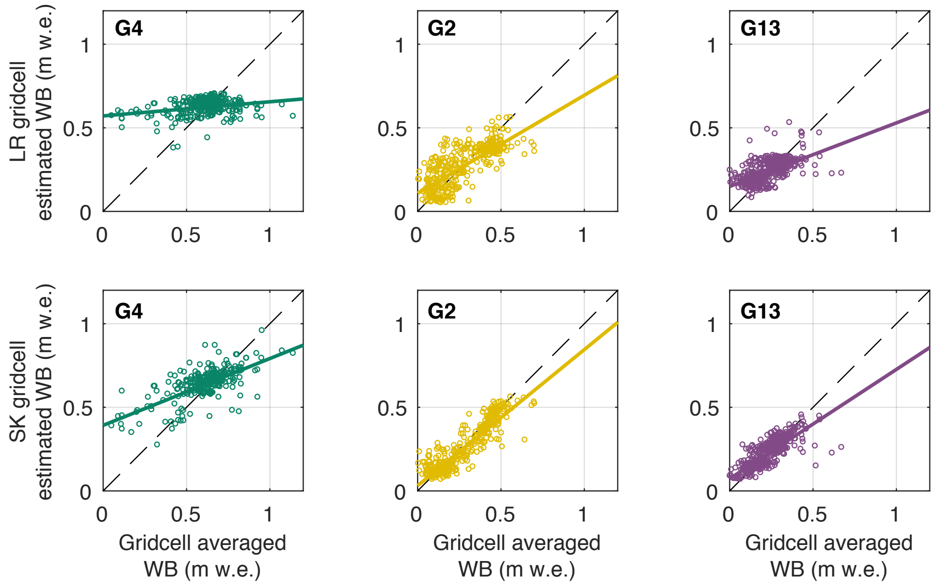

Fig. 6. Estimated gridcell-averaged SWE found using linear regression (LR, top row) and simple kriging (SK, bottom row) plotted against observed values of SWE along with best fit regression lines for Glacier 4 (left), Glacier 2 (middle) and Glacier 13 (right).

well but they cannot be transferred to different regions and that regional-scale models are not able to explain the majority of observed variance.

Simple kriging

Since simple kriging (SK) is a data-driven interpolation method, the RMSE of gridcell-estimated WB values is lower for SK than LR (Fig. 6 and Table 3). However, the uncertainty in winter balance that arises from using SK is large, and unrealistic glacier-wide winter balance values of 0 m w.e. can be estimated. Further, our observations are generally limited to the ablation area, so SK produces SWE estimates in the accumulation area that are almost uniform, which is inconsistent with observations described in the literature (e.g. Machguth and others, 2006; Grabiec and others, 2011). Extrapolation using SK leads to large uncertainty (Fig. 7) in estimating winter balance, further emphasing the need for spatially distributed SWE observations in a glacierized basin.

Fitted kriging parameters, including the nugget and spatial correlation length, can provide insight into important scales of variability. Glaciers 2 and 13 have longer correlation lengths (∼450 m) and smaller nuggets, indicating variability at large scales (see Supplementary Material Table 7). Conversely, the model fitted to the SWE data for Glacier 4 has a short correlation length (90 m) and large nugget, indicating that accumulation variability occurs at small scales.

LR and SK comparison

LR and SK estimate a winter balance of 0.6 m w.e. for Glacier 4 but both are poor predictors of SWE at measurement locations (Table 3). For Glaciers 2 and 13, SK estimates are more than 0.1 m w.e. (up to 40%) lower than LR estimates due to differences in extrapolation (Table 3). SWE estimated with LR and SK differs considerably in the upper accumulation areas of Glaciers 2 and 13 (Fig. 4), where observations are sparse and topographic parameters like elevation vary dramatically. The significant influence of elevation in the LR results in substantially higher SWE values at high elevation, whereas the accumulation area of the SK estimates approximate the mean observed SWE. However, when only the ablation area is considered, LR and SK produce winter balance estimates that differ by less than 7% for all glaciers. Choice of interpolation method therefore affects how SWE data is extrapolated, which has a large effect on winter balance estimates on our study glaciers.

Uncertainty analysis

Specific winter balance is affected by uncertainty introduced when interpolating density (σ_ρ), when calculating gridcell SWE values (σ_{GS}), and when interpolating WB data (σ_{INT}). We find that when using LR and SK, σ_{INT} has a larger effect on winter balance uncertainty than σ_{GS} or σ_ρ. In other words, the distribution of winter balance values that arises from σ_{GS} is much narrower than the distribution that arises from σ_{INT} (Fig. 8 and Table 4). The winter balance distributions obtained using LR and SK overlap for each glacier, but the distribution modes differ (Fig. 8). SK generally estimates lower winter balance in the accumulation area, which lowers the overall winter balance estimate. For both LR and SK, the greatest uncertainty in estimated SWE occurs in the accumulation area (Fig. 7). Our results caution against that using extrapolated data to compare with winter balance estimates from remote sensing or modelling studies may produce misleading results. If possible, comparison studies should use observed SWE data rather than interpolated winter balance values.

Grid-scale uncertainty is the smallest contributor to winter balance uncertainty. This is likely due to the

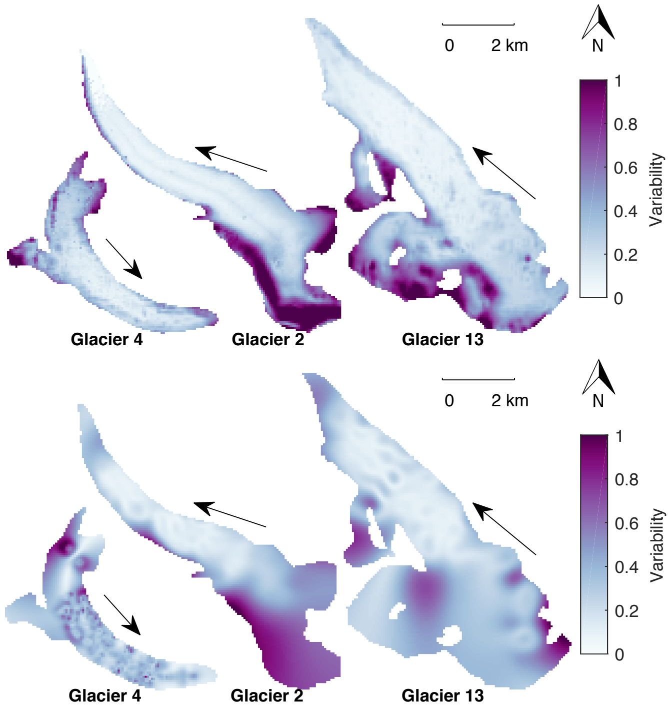

Fig. 7. Relative uncertainty in distributed SWE estimated using linear regression (top row) and simple kriging (bottom row). Relative uncertainty is calculated as the sum of differences between every pair of one hundred estimates of distributed winter balance that include gridcell uncertainty and interpolation uncertainty. The sum is then normalized for each glacier. Values closer to one indicate higher relative uncertainty. Results for density interpolation method S2 are shown. Glacier flow directions are indicated by arrows.

fact that nany parts of a glacier are characterized by a relatively smooth surface, with roughness lengths on the order of centimetres (e.g. Hock, 2005). Low glacier-wide WB uncertainty arising from σ_{GS} implies that obtaining the most accurate value of SWE to represent a gridcell does not need to be a priority when designing a snow survey. However, we assume that the gridcells selected for zigzag surveys are representative of the uncertainty across each glacier, which is likely not true for areas with debris cover, crevasses and steep slopes.

Our Monte Carlo analysis did not include uncertainty arising from a number of data sources, which we assume to contribute negligibly to the uncertainty in winter balance or to be encompassed by investigated sources of uncertainty. These sources of uncertainty include error associated with SP and FS density measurement, DEM vertical and horizontal error and error associated with estimating measurement locations.

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

	Linear Regression			**Simple Kriging**		
	σ_ρ	σ_{GS}	σ_{INT}	σ_ρ	σ_{GS}	σ_{INT}
G4	1.90	0.86	2.13	2.15	0.85	14.05
G2	3.37	1.80	3.09	2.03	2.53	13.78
G13	1.68	1.12	2.80	1.27	1.15	9.65

Context and caveats

Regional winter balance gradient

The glacier-wide WBs of our three study glaciers (S2 density assignment method), with an uncertainty equal to one standard deviation of the distribution found with Monte Carlo analysis, are: 0.593 ± 0.029 m w.e. on Glacier 4, 0.608 ± 0.049 m w.e. on Glacier 2 and 0.404 ± 0.029 m w.e. on Glacier 13. Although we find considerable inter- and intra-basin variability in WB estimates, our data fit are consistent with an accumulation gradient for the continental side of the St. Elias Mountains (Fig. 9). Accumulation data are compiled from Taylor-Barge (1969), the three glaciers presented in this paper, as well as two snow pits we dug near the head of the Kaskawulsh Glacier in May 2016. The data show a linear decrease (-0.024 m w.e. km^{-1}, $R^2 = 0.85$) in observed SWE with distance from the regional topographic divide between Kaskawulsh and Hubbard Glaciers in the St. Elias Mountains, as identified by Taylor-Barge (1969). While the three study glaciers fit the regional relationship, the same relationship would not apply when just the Donjek Range is considered. Interaction between meso-scale weather patterns and mountain topography is a major driver of regional winter balance. Further insight into regional-scale winter balance trends can be gained by investigating moisture source trajectories and and the contribution of orographic precipitation to winter balance.

Limitations and future work

Extensions to this work could include investigating experimental design, examining the effects of DEM grid size on winter balance and resolving temporal variability. Our sampling design was chosen to extensively sample the ablation area and is likely too finely resolved for many future mass balance surveys to replicate. Determining a sampling design that minimizes error and reduces the number of measurements, known as data efficiency thresholds, would contribute to optimizing snow surveys in mountainous regions. For example, López-Moreno and others (2010) concluded that $200 - 400$ observations are needed to obtain accurate and robust snow distribution models within an alpine basin.

DEM gridcell size is known to significantly affect computed topographic parameters and the ability for a DEM to resolve important hydrological features (i.e. drainage pathways) in the landscape (Zhang and Montgomery, 1994; Garbrecht and Martz, 1994; Guo-an and others, 2001; López-Moreno and others, 2010), which can have implications when using topographic parameters in a LR. Zhang and Montgomery (1994) found that a 10 m gridcell size is an optimal compromise between resolution and data volume. Further, the correlation between topographic parameters and SWE is correlated with DEM gridcell size, whereby a decrease in spatial resolution of the DEM results in a decrease in the importance of curvature and

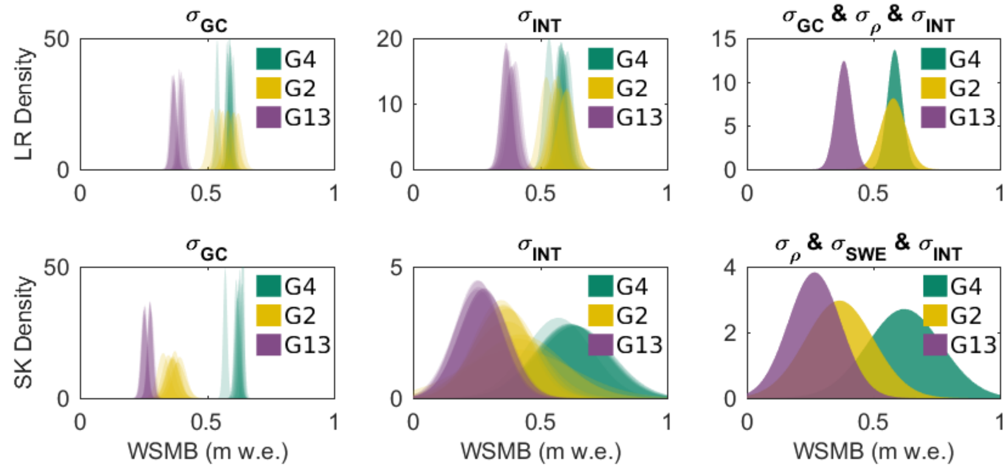

Fig. 8. Distributions of glacier-wide winter balances that arise from various sources of uncertainty. (Left column) WB distribution arising from gridcell-scale uncertainty ($\sigma_{\rm GS}$). (Middle column) WB distribution arising from interpolation uncertainty (σ_{INT}). (Right column) WB distribution arising from a combination of $\sigma_{\rm GS}$, $\sigma_{\rm INT}$ and density assignment uncertainty (σ_{ρ}). Results are shown for interpolation by (top row) linear regression and (bottom row) simple kriging. Distributions for each density assignment method are plotted within each panel for Glacier 4 (G4), Glacier 2 (G2) and Glacier 13 (G13).

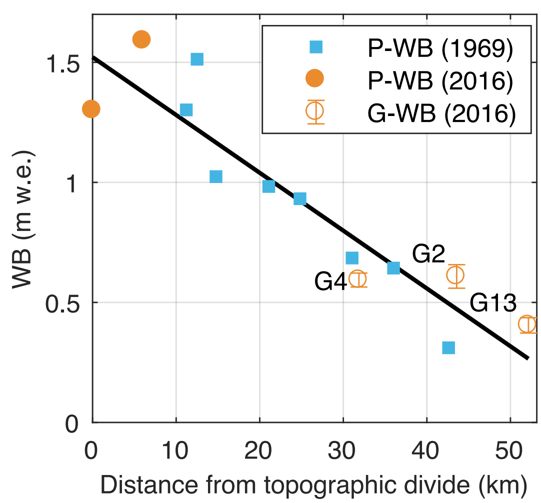

Fig. 9. Relation between winter balance (WB) and linear distance from the regional topographic divide between Kaskawulsh and Hubbard Glaciers in the St. Elias Mountains. Blue squares are point-scale WBs from snow-pit data reported by Taylor-Barge (1969). Open orange circles, labelled G4, G2 and G13, are glacier-wide WBs estimated with LR and density assignment S2 for Glaciers 4, 2 and 13, with errors bars calculated as the standard deviation of Monte Carlo-derived WB distributions. Filled orange dots are point-scale WBs from snow-pit data at two locations in the accumulation area of the Kaskawulsh Glacier, collected in May 2016 (unpublished data, SFU Glaciology Group). Black line indicates line of best fit ($R^2 = 0.85$).

an increase in the importance of elevation (e.g. Kienzle, 2004; López-Moreno and others, 2010). A detailed and ground controlled DEM is therefore needed to identify features that drive winter balance variability. Even with a high resolution DEM, small-scale snow variability created by microtopography cannot be resolved. For example, the lower part of Glacier 2 has an undulating ice surface (5 m horizontal displacement and 0.5 m vertical displacement) that results in large variability in snow depth.

Temporal variability in accumulation is not considered in our study. While this limits our conclusions, a number of studies have found temporal stability in spatial patterns of snow accumulation and that terrain-based model could be applied reliable between years (e.g. Grünewald and others, 2013). For example, Walmsley (2015) analyzed more than 40 years of accumulation recorded on two Norwegian glaciers and found that snow accumulation is spatially heterogeneous yet exhibits robust time stability in its distribution.

CONCLUSION

We estimate winter balance (WB) for three glaciers (termed as Glacier 2, Glacier 4 and Glacier 13) in the St. Elias Mountains from direct snow depth and density sampling. Our objectives are to (1) critically examine methods of moving from direct snow depth and density measurements to estimating winter balance and to (2) identify sources of uncertainty, evaluate their magnitude and assess their combined contribution to uncertainty in winter balance.

We find that the method used to interpolate observations has a large effect on WB estimates. On Glacier 4, WB estimates are consistent between linear regression (LR) and simple kriging (SK) but both explain only a small portion of the observed variance. This highlights that although the winter balance estimates are relatively precise they may not necessarily be accurate. On Glaciers 2 and 13, LR and SK are better able to estimate observed SWE values but winter balance estimates differ considerably between the two interpolation methods due to extrapolation into the accumulation area. Snow distribution patterns are found to differ considerably between glaciers, highlighting strong intra- and inter-basin variability and accumulation drivers acting on multiple scales. SWE distribution on Glacier 4 is highly variable, as indicated by shorter range distance, higher nugget value and lower explained variance. Glaciers 2 and 13 have lower SWE variability and elevation is the primary control of observed variation.

For our study glaciers, the glacier-wide WB uncertainty ranges from 0.03 m w.e (8%) to 0.15 m w.e (54%), depending primarily on the interpolation method. Uncertainty within the interpolation method is the largest source of glacier-wide WB uncertainty when compared to uncertainty in grid-scale WB and density assignment method. Future studies could reduce WB uncertainty by increasing the spatial distribution of snow depth sampling rather than the number of measurements within a single gridcell. In our work, increased sampling within the accumulation area would better constrain SWE extrapolation and decrease uncertainty. Despite challenges in accurately estimating winter balance, our data are consistent with a previously reported linear decrease in SWE with increased distance from the main topographic divide along the continental side of the St. Elias Mountains.

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)

Bagos PG and Adam M (2015) On the Covariance of Regression Coefficients. *Open Journal of Statistics*, **5**(07), 680 (doi: 10.4236/ojs.2015.57069)

Balk B and Elder K (2000) Combining binary decision tree and geostatistical methods to estimate snow distribution in a mountain watershed. *Water Resources Research*, **36**(1), 13–26 (doi: 10.1029/1999WR900251)

Barry RG (1992) *Mountain weather and climate*. Psychology Press

Beauont RT and Work RA (1963) Snow sampling results from three sampler. *International Association of Scientific Hydrology. Bulletin*, **8**(4), 74–78 (doi: 10.1080/02626666309493359)

Berthier E, Arnaud Y, Kumar R, Ahmad S, Wagnon P and Chevallier P (2007) Remote sensing estimates of glacier mass balances in the Himachal Pradesh (Western Himalaya, India). *Remote Sensing of Environment*, **108**(3), 327–338

Blöschl G, Kirnbauer R and Gutknecht D (1991) Distributed snow melt simulations in an alpine catchment. *Water Resources Research*, **27**(12), 3171–3179

Burnham KP and Anderson DR (2004) Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods & Research*, **33**(2), 261–304 (doi: 10.1177/0049124104268644)

Clark MP, Hendrikx J, Slater AG, Kavetski D, Anderson B, Cullen NJ, Kerr T, Örn Hreinsson E and Woods RA (2011) Representing spatial variability of snow water equivalent in hydrologic and land-surface models: A review. *Water Resources Research*, **47**(7) (doi: 10.1029/2011WR010745)

Clyde GD (1932) Circular No. 99-Utah Snow Sampler and Scales for Measuring Water Content of Snow

Cogley J, Hock R, Rasmussen L, Arendt A, Bauder A, Braithwaite R, Jansson P, Kaser G, Möller M, Nicholson L and others (2011) Glossary of glacier mass balance and related terms

Conger SM and McClung DM (2009) Comparison of density cutters for snow profile observations. *Journal of Glaciology*, **55**(189), 163–169

Crompton JW and Flowers GE (2016) Correlations of suspended sediment size with bedrock lithology and glacier dynamics. *Annals of Glaciology*, 1–9 (doi: 10.1017/aog.2016.6)

Cullen NJ, Anderson B, Sirguey P, Stumm D, Mackintosh A, Conway JP, Horgan HJ, Dadic R, Fitzsimons SJ and Lorrey A (2017) An 11-year record of mass balance of Brewster Glacier, New Zealand, determined using a geostatistical approach. *Journal of Glaciology*, **63**(238), 199–217 (doi: 10.1017/jog.2016.128)

Dadic R, Mott R, Lehning M and Burlando P (2010) Parameterization for wind-induced preferential deposition of snow. *Journal of Geophysical Research: Earth Surface (2003–2012)*, **115** (doi: 10.1029/2009JF001261)

Danby RK, Hik DS, Slocombe DS and Williams A (2003) Science and the St. Elias: an evolving framework for sustainability in North America's highest mountains. *The Geographical Journal*, **169**(3), 191–204 (doi: 10.1111/1475-4959.00084)

Davis JC and Sampson RJ (1986) *Statistics and data analysis in geology*, volume 646. Wiley New York et al.

Deems JS and Painter TH (2006) Lidar measurement of snow depth: accuracy and error sources. In *Proceedings of the International Snow Science Workshop*, 1–6

Dixon D and Boon S (2012) Comparison of the SnowHydro snow sampler with existing snow tube designs. *Hydrological Processes*, **26**(17), 2555–2562, ISSN 1099-1085 (doi: 10.1002/hyp.9317)

Egli L, Griessinger N and Jonas T (2011) Seasonal development of spatial snow-depth variability across different scales in the Swiss Alps. *Annals of Glaciology*, **52**(58), 216–222 (doi: 10.3189/172756411797252211)

Elder K, Dozier J and Michaelsen J (1991) Snow accumulation and distribution in an alpine watershed. *Water Resources Research*, **27**(7), 1541–1552 (doi: 10.1029/91WR00506)

Elder K, Rosenthal W and Davis RE (1998) Estimating the spatial distribution of snow water equivalence in a montane watershed. *Hydrological Processes*, **12**(1011), 1793–1808 (doi: 10.1002/(SICI)1099-1085(199808/09)12:10/11¡1793::AID-HYP695¿3.0.CO;2-)

Erxleben J, Elder K and Davis R (2002) Comparison of spatial interpolation methods for estimating snow distribution in the Colorado Rocky Mountains. *Hydrological Processes*, **16**(18), 3627–3649 (doi: 10.1002/hyp.1239)

Fames PE, Peterson N, Goodison B and Richards RP (1982) Metrication of Manual Snow Sampling Equipment. In *Proceedings of the 50th Western Snow Conference*, 120–132

Fierz C, Armstrong RL, Durand Y, Etchevers P, Greene E, McClung DM, Nishimura K, Satyawali PK and Sokratov SA (2009) *The international classification for seasonal snow on the ground*, volume 25. UNESCO/IHP Paris

Flowers GE, Copland L and Schoof CG (2014) Contemporary Glacier Processes and Global Change: Recent Observations from Kaskawulsh Glacier and the Donjek Range, St. Elias Mountains. *Arctic*, **67**, 22–34, copyright - Copyright Arctic Institute of North America 2014; Document feature - Maps; Illustrations; Graphs; ; Last updated - 2015-04-07

Garbrecht J and Martz L (1994) Grid size dependency of parameters extracted from digital elevation models. *Computers & Geosciences*, **20**(1), 85–87, ISSN 0098-3004 (doi: 10.1016/0098-3004(94)90098-1)

Gerber F, Lehning M, Hoch SW and Mott R (2017) A close-ridge small-scale atmospheric flow field and its influence on snow accumulation. *Journal of Geophysical Research: Atmospheres*, n/a–n/a, ISSN 2169-8996 (doi: 10.1002/2016JD026258), 2016JD026258

Grabiec M, Puczko D, Budzik T and Gajek G (2011) Snow distribution patterns on Svalbard glaciers derived from radio-echo soundings. *Polish Polar Research*, **32**(4), 393–421 (doi: 10.2478/v10183-011-0026-4)

Gray DM and Male DH (1981) *Handbook of snow: principles, processes, management & use*. Pergamon Press

Grunewald 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)

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, ISSN 1861-9568 (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)

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, ISSN 0043-1397 (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)

Hofierka J, Mitášová H and Neteler M (2009) Geomorphometry in GRASS GIS. *Developments in Soil Science*, **33**, 387–410 (doi: 10.1016/S0166-2481(08)00017-2)

Kaser G, Fountain A, Jansson P and others (2003) *A manual for monitoring the mass balance of mountain glaciers*. Unesco Paris

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, ISSN 1467-9671 (doi: 10.1111/j.1467-9671.2004.00169.x)

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*, volume 14, 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

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

Li J and Heap AD (2008) A review of spatial interpolation methods for environmental scientists No. Record 2008/23. *Geoscience Australia*

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)

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, ISSN 1099-1085 (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, ISSN 0309-1708 (doi: 10.1016/j.advwatres.2012.08.010), snow–Atmosphere Interactions and Hydrological Consequences

López-Moreno JI, 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) (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, ISSN 01621459

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 ISSW*, 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, ISSN 01621459

Mitášová H and Hofierka J (1993) Interpolation by regularized spline with tension: II. Application to terrain modeling and surface geometry analysis. *Mathematical Geology*, **25**(6), 657–669 (doi: 10.1007/BF00893172)

Molotch N, Colee M, Bales R and Dozier J (2005) Estimating the spatial distribution of snow water equivalent in an alpine basin using binary regression tree models: the impact of digital elevation data and independent variable selection. *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)

Olaya V (2009) Basic land-surface parameters. *Developments in Soil Science*, **33**, 141–169

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)

Réveillet M, Vincent C, Six D and Rabatel A (2016) Which empirical model is best suited to simulate glacier mass balances? *Journal of Glaciology*, 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)

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, Andereggen 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. Technical report, DTIC Document

Thibert E, Blanc R, Vincent C and Eckert N (2008) Instruments and Methods 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*, 2–11

Walmsley APU (2015) Long-term observations of snow spatial distributions at Hellstugubreen and Gråsubreen, Norway

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

Winstral A, Marks D and Gurney R (2013) Simulating wind-affected snow accumulations at catchment to basin scales. *Advances in Water Resources*, **55**, 64–79, ISSN 0309-1708 (doi: 10.1016/j.advwatres.2012.08.011), snow–Atmosphere Interactions and Hydrological Consequences

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

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)017¡1537:AOEITD¿2.0.CO;2)

Wood WA (1948) Project "Snow Cornice": the establishment of the Seward Glacial research station. *Arctic*, 107–112

Work R, Stockwell H, Freeman T and Beaumont R (1965) Accuracy of field snow surveys. Technical report

Zhang W and Montgomery DR (1994) Digital elevation model grid size, landscape representation, and hydrologic simulations. *Water Resources Research*, **30**(4), 1019–1028, ISSN 1944-7973 (doi: 10.1029/93WR03553)

SUPPLEMENTARY MATERIAL

Topographic parameters

First, cross-validation is used to obtain a set of β_i values that have greater predictive ability. We select 1000 random subsets (2/3 values) of the data to fit the LR and the remaining data (1/3 values) are used to calculate a root mean squared error (RMSE) (Kohavi and others, 1995). Regression coefficients resulting in the lowest RMSE are selected. Second, we use model averaging to take into account uncertainty when selecting predictors and to also maximize predictive ability (Madigan and Raftery, 1994). Models are generated by calculating a set of β_i for all possible combinations of predictors. Following a Bayesian framework, model averaging involves weighting all models by their posterior model probabilities (Raftery and others, 1997). 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 Bayesian Information Criterion (BIC) value (Burnham and Anderson, 2004). BIC penalizes more complex models, which further reduces the risk of overfitting.

Topographic parameters are easy to calculate proxies for physical processes, such as orographic precipitation, solar radiation effects, wind redistribution and preferential deposition. We derive all parameters (Table 5) for our study from a SPOT-5 DEM (40×40 m) (Korona and others, 2009). Two DEMs are stitched together to encompass the Donjek Range. An iterative 3D-coregistration algorithm (Berthier and others, 2007) is used to correct the horizontal (∼2 m E, ∼4 m N) and vertical (5.4 m) discrepancy between the two DEMs before stitching.

Visual inspection of the curvature fields calculated using the full DEM shows a noisy spatial distribution that did not vary smoothly. To smooth the DEM, various smoothing algorithms and window sizes are applied and the combination that produces the highest correlation between topographic parameters and SWE is chosen. Inverse-distance weighted, Gaussian and gridcell averaging smoothing all with window sizes of 3×3, 5×5, 7×7 and 9×9 are used. gridcell average smoothing with a 7×7 window resulted in the highest overall correlation between curvature (second derivative) and SWE as well as slope (first derivative) and SWE. We use the smoothed DEM to calculate curvature, slope, aspect and “northness”.

Table 5. Description of topographic parameters used in the linear regression.

Topographic parameter	Definition	Calculation method	Notes	Source
Elevation (z)		Values taken directly from DEM		
Distance from centreline (d_C)		Minimum distance between the Easting and Northing of the northwest corner of each gridcell and a manually defined centreline		
Slope (m)	Angle between a plane tangential to the surface (gradient) and the horizontal	`r.slope.aspect` module in GRASS GIS software run through QGIS		Mitášová and Hofierka (1993); Hofierka and others (2009); Olaya (2009)
Aspect (α)	Dip direction of the slope	`r.slope.aspect` module in GRASS GIS software run through QGIS	$\sin(\alpha)$, a linear quantity describing a slope as north/south facing, is used in the regression	Mitášová and Hofierka (1993); Hofierka and others (2009); Olaya (2009)
Mean curvature (κ)	Average of profile (direction of the surface gradient) and tangential curvature (direction of the contour tangent)	`r.slope.aspect` module in GRASS GIS software run through QGIS	mean-concave (positive values) terrain with relative accumulation and mean-convex (negative values) terrain with relative scouring	Mitášová and Hofierka (1993); Hofierka and others (2009); Olaya (2009)
“Northness” (N)	-1 represents a vertical, south facing slope, a value of $+1$ represents a vertical, north facing slope, and a flat surface yields 0	Product of the cosine of aspect and sine of slope		Molotch and others (2005)
Wind exposure/shelter parameter (Sx)		Executable obtained from Adam Winstral that follows the procedure outlined in Winstral and others (2002)	Calculation based on selecting a cell within a certain angle and distance from the cell of interest that has the greatest upward slope relative to the cell of interest	Winstral and others (2002)

Table 6. Snow density values used for interpolating density based on snow pit (SP) densities and Federal Sampler (FS) densities. Four interpolation methods are chosen: (1) using a mean snow density for all three glaciers (Range mean density), (2) using a mean density for each glacier (Glacier mean density), (3) using a regression between density and elevation (Elevation regression), and (4) inverse-distance weighted mean density (not shown).

		SP density (kg m^{-3})	FS density (kg m^{-3})
Range mean density		342	316
Glacier mean density	G4	348	327
	G2	333	326
	G13	349	307
Elevation regression	G4	$0.03z + 274$	$-0.16z + 714$
	G2	$-0.14z + 659$	$0.24z - 282$
	G13	$-0.20z + 802$	$0.12z + 33$

Table 7. Range and nugget values for simple kriging interpolation

	Range (m)	Nugget ($\times 10^3$m w.e.)
G4	90	10.5
G2	404	3.6
G13	444	4.8

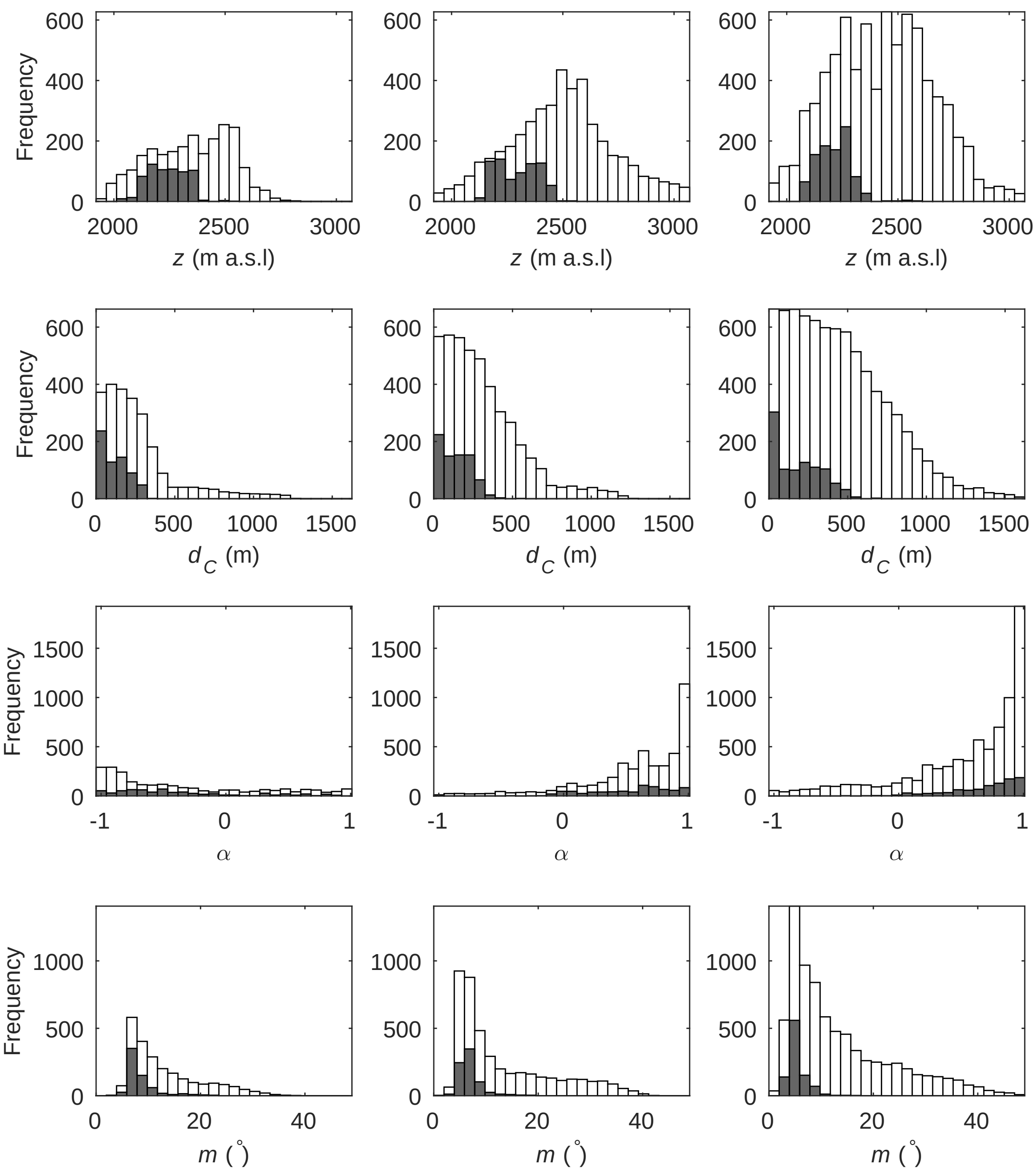

Fig. 10. Distribution of topographic parameters over Glacier 4 (left), Glacier 2 (middle) and Glacier 13 (right) are shown in white. Distribution of topographic parameter values from sampled gridcells in shown in gray. Topographic parameters include elevation (z), distance from centreline (d_C), aspect (α), slope (m), northness (N), mean curvature (κ), and winter redistribution (Sx).

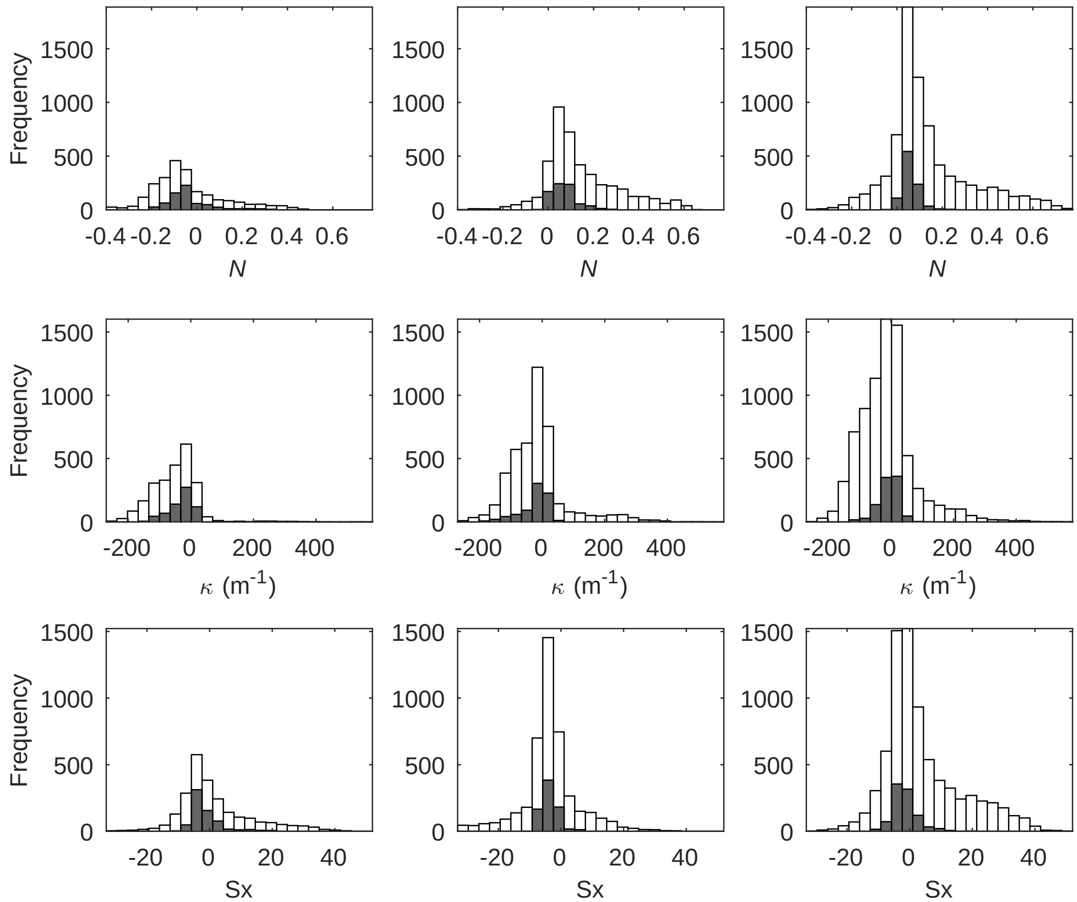

Fig. 11. See Fig. 10