Estimating winter surface mass balance

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

[1] *Simon Fraser University, Burnaby, BC, Canada*

[2] *University of British Columbia, Vancouver, BC, Canada*

Correspondence: Alexandra Pulwicki <apulwick@sfu.ca>

ABSTRACT. Accurately estimating winter surface mass balance for a glacier is central to ??. However, measuring and modelling snow distribution and variability is inherently difficult resulting in high uncertainty. The goal of this paper is to provide a comprehensive sweep of choices and assumptions present in moving from snow observations to winter mass balance estimates to better understand how interactions between snow variability, data error and our methodological choices contribute to uncertainty. We extensively measure snow depth and density, at various spatial scales, on three glaciers in the St. Elias Mountains, Yukon. Elevation is found to be the dominant driver of accumulation variability but results vary between glaciers. Our results also suggest that wind redistribution and preferential deposition affects snow distribution but that more complex parametrization is need to fully capture wind effects. We use a Monte Carlo method to quantify the effects of variability due to density interpolation method, snow water equivalent observations as well as observation interpolation on estimates of winter surface mass balance. The largest source of uncertainty stems from calculating parameters for interpolation using linear regression as well as simple kriging. Spatially extensive measurements in the accumulation area are needed, at the expense of detailed ablation area measurements, to better constrain interpolation models and reduce uncertainity.

INTRODUCTION

Accurate estimation of winter surface mass balance is critical for correctly simulating the summer and overall mass balance of a glacier (Réveillet and others, 2016). Effectively representing 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 to better monitor surface runoff and its downstream effects (Clark and others, 2011). Snow distribution is sensitive to a number of complex process that partially depend on glacier location, topography, and orientation (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 the distribution of snow in remote, mountainous locations is not well known. This is a significant source of uncertainty that undermines the ability of models to represent current glacier conditions and make predictions of glacier response to a warming climate.

Winter mass balance is notoriously difficult to measure. 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 (Liston and Elder, 2006). Extensive, high resolution and accurate accumulation measurements on glaciers are almost impossible to achieve due to cost benefits of the various methods used to quantify snow water equivalent (McGrath and others, 2015). For example, snow probes obtain accurate point observations but have negligible spatial coverage. Conversely, gravimetric methods obtain extensive measurements of mass change but cannot capture relevant spatial variability of snow. Glacierized regions are also generally remote and challenging to access during the winter due to poor travelling conditions.

Predicting winter mass balance is a further challenge. Physically-based dynamic models are able to capture the intricate interactions between the atmosphere and local topography but they are operationally complex and computationally expensive, and require a diverse set of detailed observations (Dadic and others, 2010). Empirical models that rely on statistical relationships between proxy parameters and measured accumulation are widely applied and simple to execute but most are unable to explain the majority of variance observed or lack insight into processes that affect snow distribution (e.g. Grabiec and others, 2011; López-Moreno and others, 2011)..

There is clearly a strong need for a more comprehensive understanding of uncertainties inherent when estimating accumulation on glaciers. Ultimately, we need a thorough knowledge of the processes that affect spatial and temporal snow variability and an effective method to predict snow accumulation. The contribution

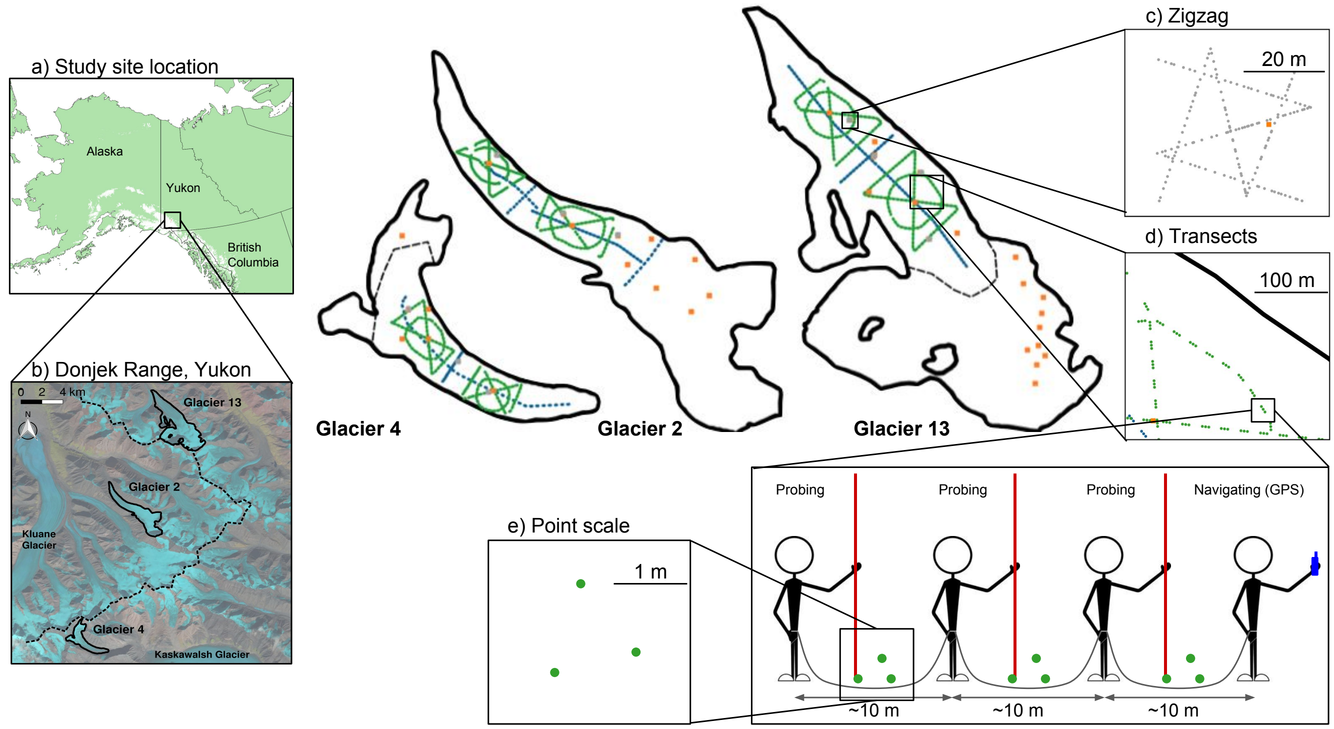

Fig. 1. Sampling design for Glaciers 4, 2 and 13, located in the Donjek Range, Yukon (a,b). Centreline and transverse transects are shown in blue dots, hourglass and circle design are shown in green dots. (d) Linear and curvilinear transects typically consist of sets of three measurement locations, spaced ∼10 m apart. (e) At each measurement location, three snow depth observation are made. (c) Linear-random snow depth measurements in 'zigzag' design are shown as grey dots. Orange squares are locations of snow density measurements.

of our work toward this goal is to (1) conduct a comprehensive sweep of choices and assumptions made when moving from snow measurements to estimating accumulation and (2) show how snow variability, data error and our methodological choices interact to create uncertainty in our estimate of accumulation. We focus on simple and commonly applied methods of measuring and predicting accumulation with the hope of making our results broadly applicable to past and current winter mass balance programs.

METHODS

Estimating accumulation involves transforming snow depth and density measurements to distributed estimates of snow water equivalent (SWE). We use four main processing steps that obtain (1) measurements, (2) distributed density, (3) grid cell average and (4) interpolated SWE. To estimate the specific winter surface mass balance (WSMB) we calculate the mean SWE for a grid cell from the estimated distributed SWE.

Measurements

The estimated SWE is the product of the snow depth and 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 with relatively small measurement spacing along transects that resulted in a ratio of approximately 55:1 snow depth to snow density measurements. The sampling design attempted to capture depth variability at multiple spatial scales and to account for known variation with elevation. Our sampling design is created to avoid bias, be space filling within the ablation area and minimize distance travelled (Shea and Jamieson, 2010).

We measured accumulation at three glaciers along the precipitation gradient in the St. Elias Mountains, Yukon (Taylor-Barge, 1969) in an attempt to account for range-scale variability (Clark and others, 2011). We measured accumulation on Glaciers 4, 2, and 13 (naming adopted from Crompton and Flowers (2016)), which are located increasingly far from the head of the Kaskawalsh Glacier (Figure 1b). Snow depth was measured along linear and curvilinear transects to encompass basin-scale variability. At each measurement location, three values of snow depth were recorded to account for point-scale variability (Clark and others, 2011). We selected centreline and transverse transects with sample spacing of $10-60$ m (Figure 1d) to capture previously established correlations between elevation and accumulation (Machguth and others, 2006; Walmsley, 2015) as well as accumulation differences between ice-marginal and center accumulation. We also implemented an hourglass and circle design (Figure 1), which allows for sampling in all directions and easy travel (Parr, C., 2016 personal communication). At each measurement location, we took $3-4$ depth measurements (Figure 1e), resulting in more than 9,000 snow depth measurements throughout the study area.

Our sampling campaign involved four people and occurred between May 5 and 15, 2015, which corresponds to the historical peak accumulation in the Yukon (Yukon Snow Survey Bulletin and Water Supply Forecast, May 1, 2016). While roped-up for glacier travel, the lead person used a hand-held GPS (Garmin GPSMAP 64s) to navigate as close to the predefined transect measurement locations as possible (Figure 1). The remaining three people used 3.2 m aluminium avalanche probes to take $3-4$ snow depth measurements within $\sim$1 m of each other. Each observer was approximately 10 m behind the person ahead of them along the transect line. 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 person.

Snow depth sampling was primarily done in the ablation area to ensure that only snow from the current accumulation season was measured. Determining the boundary between snow and firn in the accumulation area, especially when using an avalanche probe, is difficult and often incorrect (Grunewald and others, 2010; Sold and others, 2013). We intended to use a firn corer to extract snow cores in the accumulation area but due to technical issues we were unable to obtain cohesive cores. The recorded accumulation area measurements were done either in a snow pit or with a Federal Sampler so that we could identify the snow-firn transition based on a change in snow crystal size and density.

To capture variability at spatial scales smaller than a DEM grid cell, we implemented a linear-random sampling design, termed 'zigzag' (Shea and Jamieson, 2010). We measured depth at random intervals ($0.3-3.0$ m) along two 'Z'-shaped transects within three to four 40×40 m squares (Figure 1c) resulting in $135-191$ measurement points for each zigzag. Zigzag locations were randomly chosen within the upper ($\sim$2350 m a.s.l.), middle ($\sim$2250 m a.s.l.), and lower portions ($\sim$2150 m a.s.l.) of the ablation area of each glacier. We were able to measure a fourth zigzag on Glacier 13, which was located in the middle ablation area ($\sim$2200 m a.s.l.).

Snow density was measured using a wedge cutter in three snowpits on each glacier. We collected a continuous density profile by inserting a $5 \times 5 \times 10$ cm (250 cm^3) wedge-shaped cutter in 5 cm increments to extract snow samples and then weighed 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 measurement errors and incorrect assignment of density to layers that could not be sampled (i.e. ice lenses and 'hard' layers).

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 measures bulk SWE, was used to augment the spatial extent of density measurements. A minimum of three measurements were taken at $7-19$ locations on each glacier and eight FS measurements were co-located with each snow pit profile. Measurements where the tube snow length was less than 90% of the snow depth were assumed to be an incorrect sample and were excluded. Density values were then averaged for each location.

During the field campaign there were two small accumulation events. The first, on May 6, also involved high winds so accumulation could not be determined. The second, on May 10, resulted in 0.01 m w.e accumulation at one location on Glacier 2. High temperatures and clear skies occurred between May 11 and 16, which we believed resulted in significant melt occurring on Glacier 13. The snow in the lower part of the ablation area was isothermal and showed clear signs of melt and snow metamorphosis. Total amount of accumulation and melt during the study period could not be estimated so no corrections were made.

Distributed density

Measured density is interpolated to estimate SWE at each depth sampling location. We chose four separate methods that are commonly applied to interpolate density: (1) mean density over an entire range (e.g. Cullen and others, 2017), (2) mean density for each glacier (e.g. Elder and others, 1991; McGrath and others, 2015), (3) linear regression of density with elevation (e.g. Elder and others, 1998; Molotch and others, 2005) and (4) inverse-distance weighted density (e.g. Molotch and others, 2005).

When designing the sampling campaign we assumed that SP and FS densities could be combined so that we could have a more spatially distributed density data set. However, there is no correlation between co-located SP and FS densities (Figure 3). Therefore, SP and FS measurements were used independently for each interpolation method, resulting in eight density interpolation options.

Grid cell average

We average SWE values within each DEM-aligned grid cell. The locations of measurements have considerable uncertainty both from the error of the GPS unit ($2.7-4.6$ m) and the estimation of observer location based on the GPS unit. These errors could easily result in the incorrect assigning of a SWE measurement to a certain grid but this source of variability was not further investigated because we assume that SWE variability is captured in the zigzag measurements described below. There are no differences between observers ($p>0.05$), with the exception of the first transect on Glacier 4, so no corrections to the data based on observer are applied.

Interpolated SWE

SWE data were interpolated for each glacier using linear regression (LR) as well as simple kriging (SK). Linear regressions relate observed SWE to grid cell values of DEM-derived topographic parameters (Davis and Sampson, 1986). We chose to include elevation, distance from centreline, slope, aspect, curvature, "northness" and a wind redistribution parameter in the LR. Topographic parameters were weighted by a set of fitted regression coefficients (β_i). Regression coefficients are calculated by minimizing the sum of squares of the vertical deviations of each data point from the regression line (Davis and Sampson, 1986). The distributed estimate of SWE was found by using regression coefficients to estimate SWE at each grid cell. Specific WSMB was calculated as the mean SWE for each glacier ([m w.e.]).

The goal of generating a LR is to predict SWE at unsampled grid cells and to tease out dominant relationships between accumulation and topographic parameters. Since snow depth data is highly variable, there is a possibility for the LR to fit to this data noise, a process known as overfitting. To prevent overfitting, cross-validation and model averaging were implemented. Cross-validation was used to obtain a set of β_i values that have greater predictive ability. We selected 1000 random subsets (2/3 values) of the data to fit the LR and the remaining data (1/3 values) was used to calculate a root mean squared error (RMSE) (Kohavi and others, 1995). Regression coefficients resulting in the lowest RMSE were selected. Model averaging takes into account uncertainty when selecting predictors and also maximizes predictive ability (Madigan and Raftery, 1994). Models were 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 were 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 derived all parameters for our study from a SPOT-5 DEM (40×40 m) (Korona and others, 2009). Elevation (z) values were taken from the SPOT-5 DEM directly. Distance from centreline (d_C) was calculated as the minimum distance between the Easting and Northing of the northwest corner of each grid cell and a manually defined centreline. Slope, aspect and curvature were calculated using the `r.slope.aspect` module in GRASS GIS software run through QGIS as described in Mitášová and Hofierka (1993) and Hofierka and others (2009). Slope (m) is defined as the angle between a plane tangential to the surface (gradient) and the horizontal (Olaya, 2009). Aspect (α) is the dip direction of the slope and $\sin(\alpha)$, a linear quantity describing a slope as north/south facing, is used in the regression. Mean curvature (κ) is found by taking the average of profile and tangential curvature. Profile curvature is the curvature in the direction of the surface gradient and it describes the change is slope angle. Tangential curvature represents the curvature in the direction of the contour tangent. Curvature differentiates between mean-concave (positive values) terrain with relative accumulation and mean-convex (negative values) terrain with relative scouring (Olaya, 2009). "Northness" (N) is defined as the product of the cosine of aspect and sine of slope (Molotch and others, 2005). A value of -1 represents a vertical, south facing slope, a value of $+1$ represents a vertical, north facing slope, and a flat surface yields 0. The wind exposure/shelter parameter (Sx) is 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). Sx was calculated using an executable obtained from Adam Winstral that follows the procedure outlined in Winstral and others (2002).

Visual inspection of the curvature fields calculated using the DEM showed a noisy spatial distribution that did not vary smoothly. To minimize the effect of noise on parameters sensitive to DEM grid cell size, we applied a 7×7 grid cell smoothing window to the DEM, which was then used to calculate curvature, slope, aspect and "northness".

Our sampling design ensured that the ranges of topographic parameters covered by the measurements represented more than 70% of the total area of each glacier (except for the elevation range on Glacier 2, which was 50%). However, were were not able to sample at locations with extreme parameter values and the distribution of the sampled parameters generally differed from the full distribution.

Simple kriging (SK) estimates SWE values at unsampled locations by using the isotropic spatial correlation (covariance) of measured SWE to find a set of optimal weights (Davis and Sampson, 1986; Li and Heap, 2008). SK assumes that if sampling points are distributed throughout a surface, the degree of spatial correlation of the observed surface can be determined and the surface can then be interpolated 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. 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).

Quantifying effects of variability

We identify three major sources of variability within the process of translating snow measurements to WSMB. These variability sources encompass error and uncertainty within each processing step. When calculating distributed density, choice of density interpolation method is the largest source of variability. We therefore carry all density interpolation options forward in the estimation of WSMB. When calculating a grid cell average SWE, variability stems from a distribution of SWE values within each grid cell, which is assumed to be caused by random effects that are unbiased and unpredictable (Watson and others, 2006). We therefore choose to encompass the SWE variability by generating a normal distribution of SWE values for each measured grid cell. The normal distribution has a mean equal to the grid cell average SWE and a standard deviation equal to the mean standard deviation of all zigzags on each glacier. When obtaining interpolated SWE, the best fit interpolation itself has variability based on the data that is used to fit the regression line or kriging surface. LR variability is represented by obtaining a multivariate normal distribution of possible β_i

values. The standard deviation of each distribution is calculated using the covariance of regression coefficients as outlined in Bagos and Adam (2015). SK variability is calculated using the DiceKriging package and is returned as an upper and lower 95% confidence interval for SWE at each grid cell. We refer to the three variability sources as (1) density variability, (2) SWE variability and (3) interpolation variability.

To quantify the effects of the three variability sources on the final WSMB estimate, we conduct a Monte Carlo experiment, which uses repeated random sampling to calculate a numerical solution (Metropolis and Ulam, 1949). In our study, we randomly sample the distributions for SWE variability and interpolation variability and carry these values through the data processing steps to obtain a value of WSMB. First, random values from the distribution of SWE values for each grid cell are independently chosen. Then, LR or SK is used to interpolate these SWE values. With the LR, a set of β_i values and their distributions are calculated and the β_i distributions are randomly sampled. These new β_i values are used to calculate WSMB. With SK, a distribution of WSMB is calculated from the 95% confidence interval kriging surfaces. Density variability is accounted for by repeating the process for each density interpolation method. This random sampling process is done 1000 times, which results in a distribution of possible WSMB values based on variability within the data processing steps.

RESULTS

Measurements

A wide range of snow depth is observed on all three study glaciers (Figure 2). Glacier 4 has the highest mean snow depth and a high proportion of outliers, indicating a more variable snow depth overall. Glacier 13 has the lowest mean snow depth and a narrower distribution of observed values. 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.

Mean SP and FS density values are within one standard deviation of each other for each glacier and over all three glaciers. The standard deviation of glacier-wide mean density is less than 10% of the mean density. However, FS densities have a larger range of values ($227-431$kg m^{-3}) when compared to SP densities ($299-381$kg m^{-3}). The mean SP densities are within one standard deviation between glaciers, whereas mean FS densities are not.

Uncertainty in SP density is largely due to sampling error of exceptionally dense snow layers. We quantify this uncertainty by varying three values. Ice layer density is varied between 700 and 900 kg m^{-3}, ice layer

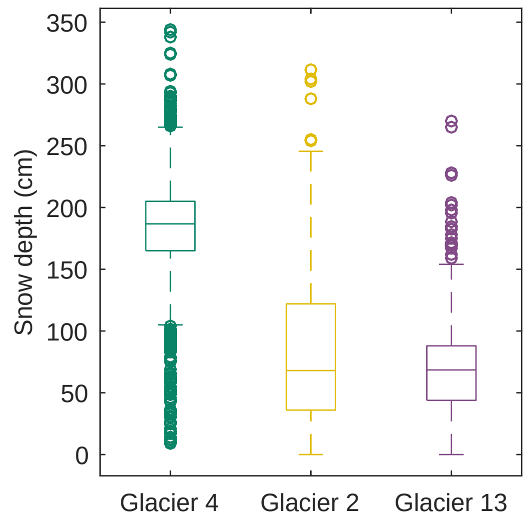

Fig. 2. Boxplot of measured snow depth on Glaciers 4, 2 and 13. The box shows first quartiles, the line within the box indicates data median, bars indicate minimum and maximum values (excluding outliers), and circles show outliers, which are defined as being outside of the range of 1.5 times the quartiles (approximately $\pm 2.7\sigma$).

thickness is varied by ± 1 cm of the observed thickness, and the density of layers identified as being too hard to sample (but not ice) is varied between 600 and 700 kg m^{-3}. The range of integrated density values is always less than 15% of the reference density, with the largest ranges present on Glacier 2. Density values for shallow pits that contain ice lenses are particularly sensitive to changes in density and ice lens thickness.

Distributed density

There is no correlation between co-located SP and FS densities (Figure 3) so each set of density values is used for all four density interpolation options. Range and glacier mean densities are higher when SP densities are used (Table 1). The magnitude and slope of a linear regression of density with elevation differs between SP and FS densities (Table 1). At Glaciers 2 and 13, SP density decreases with elevation, likely indicating melt at lower elevations. SP density is independent of elevation on Glacier 4. FS density increases with elevation on Glacier 2 and there is no relationship with elevation on Glaciers 4 and 13.

There is a positive linear relation ($R^2 = 0.59$, $p<0.01$) between measured snow density and depth for all FS measurements. No correlation exists between SP density and elevation.

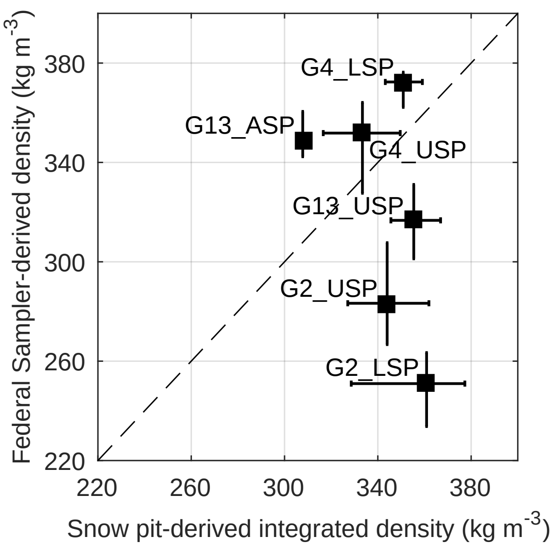

Fig. 3. Comparison of integrated density estimated using wedge cutters in a snow pit and density estimated using Federal Sampler measurements for Glacier 4 (G04), Glacier 2 (G02) and Glacier 13 (G13). Snow pits were distributed in the accumulation area (ASP), upper ablation area (USP) and lower ablation area (LSP). Error bars are minimum and maximum values.

Grid cell average

SWE observations within a DEM grid cell are averaged. Between one and six measurement locations are in each measured grid cell. The distribution of grid-cell SWE values for each glacier is similar to that of Figure 2 but with fewer outliers.

SWE measurements for each zigzag are not normally distributed about the mean SWE (Figure 4). The average standard deviation of all zigzags on Glacier 4 is $\sigma_{\mathrm{G4}} = 0.027$ m w.e., on Glacier 2 is $\sigma_{\mathrm{G2}} = 0.035$ m w.e. and on Glacier 13 is $\sigma_{\mathrm{G13}} = 0.040$ m w.e.

Interpolated SWE

The choice of interpolation method affects the specific WSMB (Table 2). SK produces the highest WSMB on Glacier 4 and the lowest WSMB on Glacier 13. WSMB estimated by SK is ~30% lower than WSMB estimated by LR on Glaciers 2 and 13. When using LR, the WSMB on Glaciers 4 and 2 are similar in magnitude.

The predictive ability of SK and LR differ on the study glaciers. Generally, SK is better able to predict SWE at observed grid cells (Figure 5) and RMSE for all glaciers is lower for SK estimates (Table 2). Glacier

Table 1. 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 ($\mathbf{kg\ m^{-3}}$)	**FS density ($\mathbf{kg\ m^{-3}}$)**
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	$-0.20z + 802$	$0.12z + 33$

13 has the lowest RMSE regardless of interpolation method, indicating lower SWE variability. The highest RMSE and the lowest correlation between estimated and observed SWE is seen on Glacier 4 ($R^2 = 0.12$) , which emphasizes the highly variable snow pack. The highest correlation between estimated and observed SWE is on Glacier 2 when SK is used for interpolation (R^2 =0.84) (Figure 5). Residuals using LR and SK for all glaciers are normally distributed.

The importance of topographic parameters in the LR differs for the three study glaciers (Figure 6). The most important topographic parameter for Glacier 4 is wind redistribution. However, the wind redistribution

Table 2. Specific winter surface mass balance (WSMB [m w.e.]) estimated using linear regression and simple kriging interpolation for study glaciers. Average root mean squared error (RMSE [m w.e.]) between estimated and observed grid cells that were randomly selected and excluded from interpolation.

	Linear Regression		**Simple Kriging**	
	WSMB	RMSE	WSMB	RMSE
Glacier 4	0.582	0.153	0.616	0.134
Glacier 2	0.577	0.102	0.367	0.073
Glacier 13	0.381	0.080	0.271	0.068

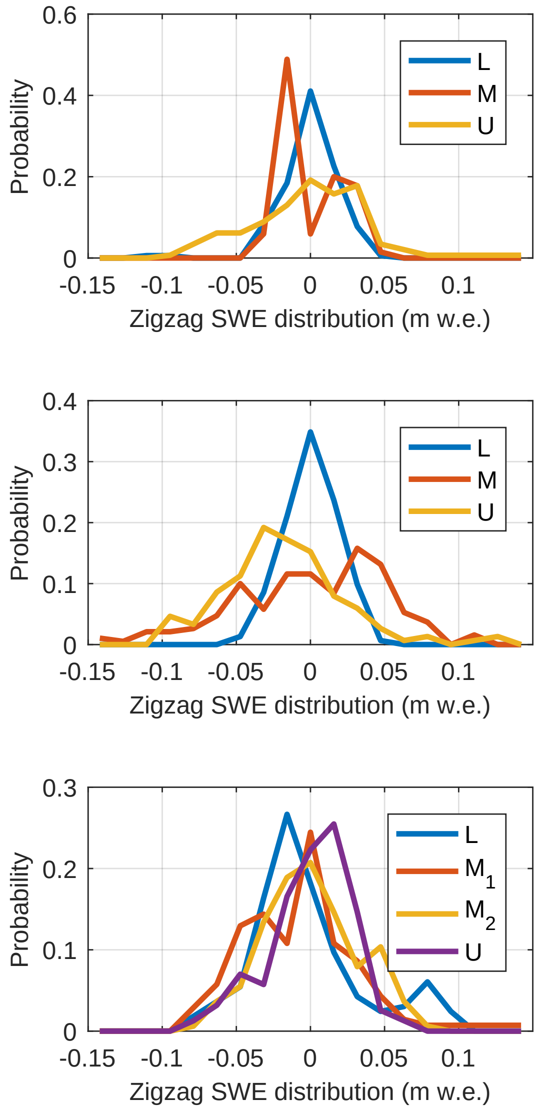

Fig. 4. Distribution of zigzag SWE values about the local mean on Glacier 4 (upper panel), Glacier 2 (middle panel) and Glacier 13 (lower panel). Zigzags are distributed throughout the ablation area of each glacier, with one located in the lower portion (L), one in the middle portion (M), and one in the upper portion (U). There were two zigzags in the middle ablation area of Glacier 13.

coefficient is negative, which indicates less snow in 'sheltered' areas. Curvature is also a significant predictor of accumulation and the positive correlation indicates that concave areas are more likely to have higher

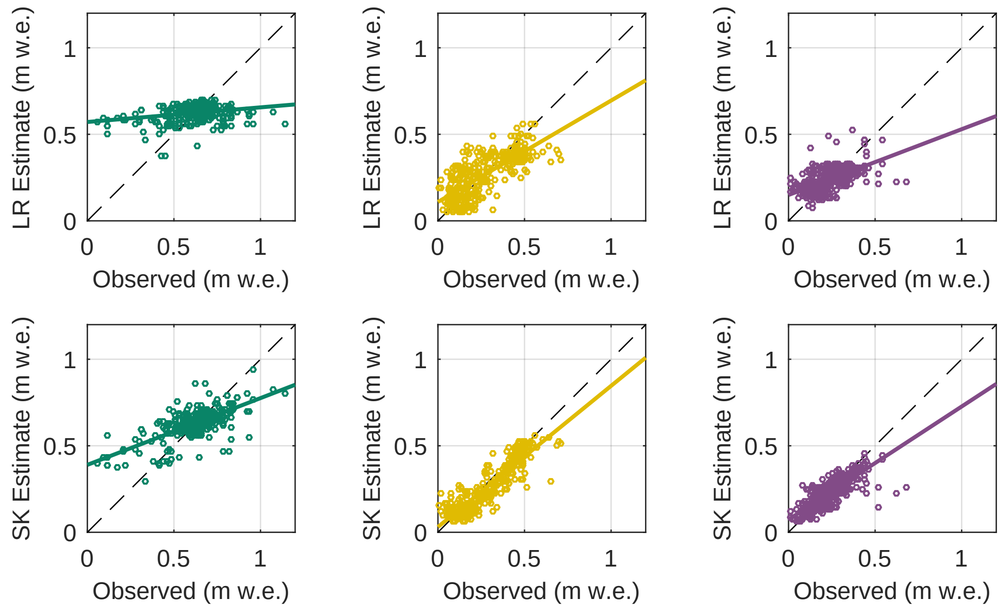

Fig. 5. Estimated grid cell SWE found using linear regression (LR) and simple kriging (SK) plotted against observed values of SWE on Glacier 4 (left), Glacier 2 (middle) and Glacier 13 (right). Line of best fit between estimated and observed SWE is also plotted.

SWE. For Glacier 2, the most important topographic parameter is elevation, which is positively correlated with elevation. Wind redistribution is the second most important topographic parameter and has a positive correlation, which indicates that 'sheltered' areas are likely to have high accumulation. The most important topographic parameter for Glacier 13 is elevation. The coefficient is positive, which means that cells at higher elevation have higher SWE. Curvature is also a significant topographic parameter but the correlation is negative, indicating less accumulation in concave areas. Most of the topographic parameters are not significant predictors of accumulation on Glacier 13. Aspect and "northness" are not significant predictors of accumulation on all study glaciers.

Spatial patterns of SWE found using LR are similar between Glaciers 2 and 13 and differ considerably for Glacier 4 (Figure 7). Estimated SWE on Glacier 4 is relatively uniform, which results from the low predictive ability of the LR. Areas with high wind redistribution values (sheltered), especially in the accumulation area, have the lowest values of SWE. The map of modelled SWE on Glacier 2 closely matches that of elevation, which highlights the strong dependence of SWE on elevation. Glacier 2 has the largest range of estimated SWE ($0-1.92$ m w.e). The area of high estimated accumulation in the southwest region of the glacier results

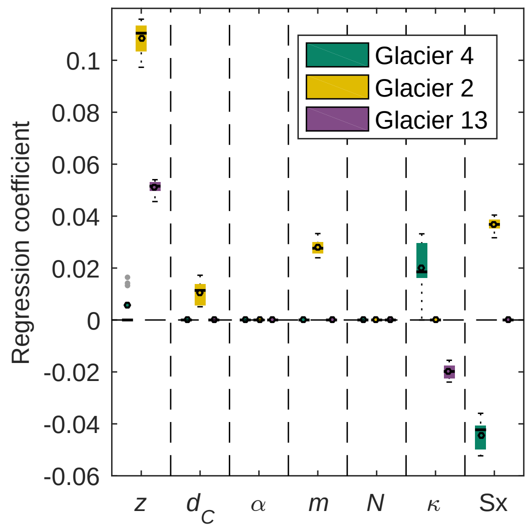

Fig. 6. Distribution of regression coefficients for linear regression of grid cell topographic parameters and SWE calculated using eight density options on study glaciers. Topographic parameters include elevation (z), distance from centreline (d_C), slope (m), aspect (α), curvature (κ), "northness" (N) and wind exposure (Sx). Regression coefficients that were not significant were assigned a value of zero.

from the combination of high elevation and Sx values. The low SWE values at the terminus arise from low elevation and Sx values close to zero. The map of estimated SWE on Glacier 13 also closely follows elevation. However, the lower correlation between SWE and elevation results in a relatively small range of distributed SWE values.

There are large differences in spatial patterns of estimated WSMB for the three study glaciers found using SK (Figure 7). On Glacier 4, the isotropic correlation length is considerably shorter (90 m) compared to Glacier 2 (404 m) and Glacier 13 (444 m), which results in a relatively uniform SWE distribution over the glacier with small deviations at measured grid cells. Nugget values for the study glaciers also differ, with the nugget of Glacier 4 (0.0105 m w.e.) more than twice as large as that of Glacier 2 (0.0036 m w.e.) and Glacier 13 (0.0048 m w.e.). Glacier 2 has two distinct and relatively uniform areas of estimated accumulation. The lower ablation area has low SWE (∼0.1 m w.e.) and the upper ablation and accumulation areas have higher SWE values (∼0.6 m w.e.). Glacier 13 does not appear to have any strong patterns and accumulation is generally low ($\sim 0.1 - 0.5$ m w.e.).

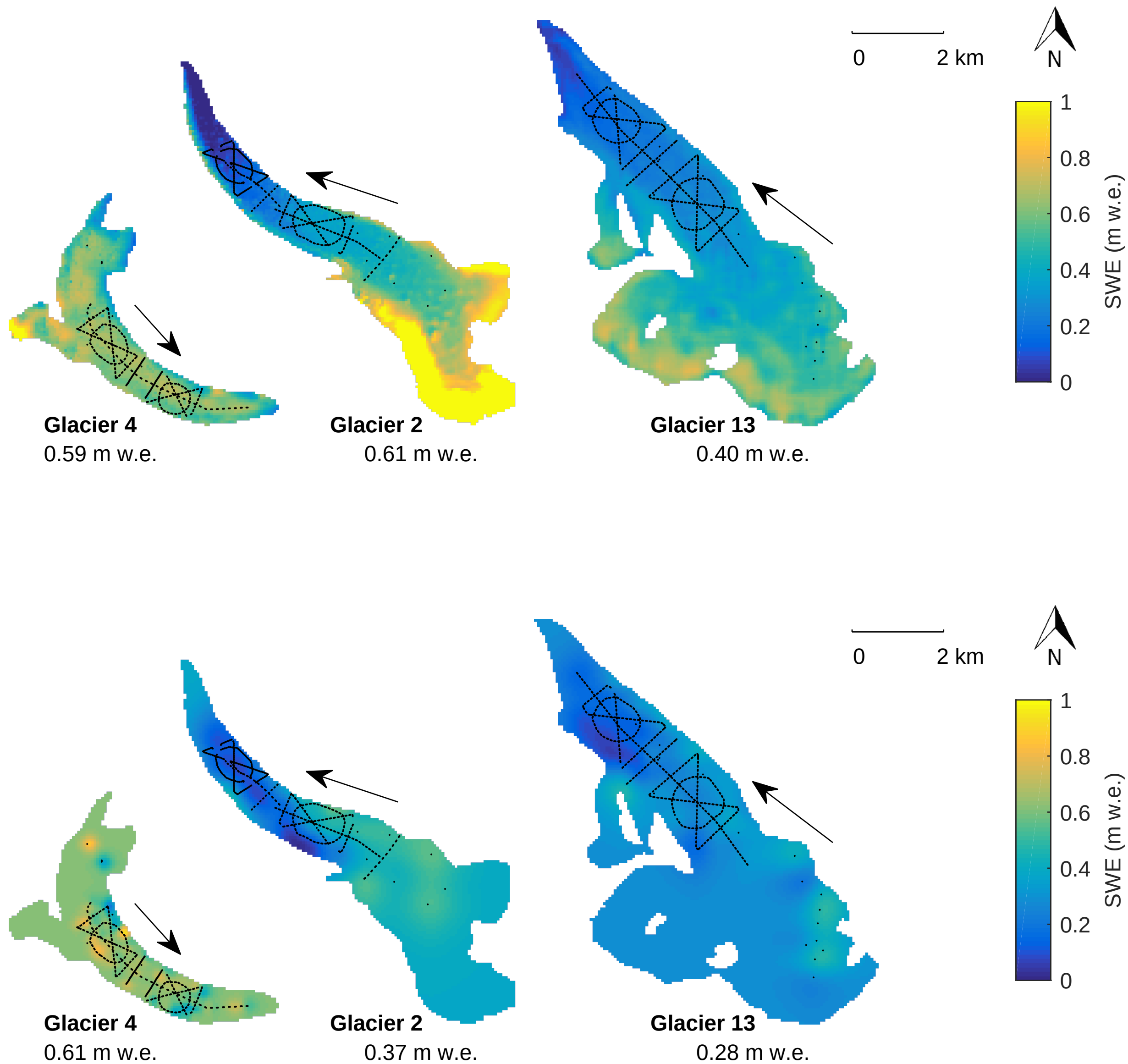

Fig. 7. Spatial distribution of SWE estimated using linear regression (upper) and simple kriging (lower). Grid-cell SWE observations are calculated using glacier wide mean snow pit density and are shown as black dots. Glacier flow directions are indicated by arrows. Specific WSMB values are also shown.

SWE estimated with LR and SK differ considerably in the upper accumulation areas of Glaciers 2 and 13. 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.

Table 3. Standard deviation ([m w.e.]) of specific winter surface mass balance estimated using linear regression (LR) and simple kriging (SK) when variability is introduced. Density variability (σ_ρ) is the standard deviation of WSMB estimated using SWE data with different density interpolation methods. SWE variability (σ_{SWE}) is approximated by a normal distribution about the local SWE value with standard deviation equal to the glacier-wide mean zigzag standard deviation. LR interpolation variability (σ_β) is accounted for by varying the regression coefficients with a normal distribution with standard deviation calculated from regression covariance. SK interpolation variability (σ_{KRIG}) is taken from the range of distributed SWE estimates calculated by the DiceKriging package. Result for Glacier 4 (G4), Glacier 2 (G2) and Glacier 13 (G13) are shown.

	Linear Regression			**Simple Kriging**		
	σ_ρ	σ_{SWE}	σ_β	σ_ρ	σ_{SWE}	σ_{KRIG}
G4	0.0190	0.0086	0.0213	0.0215	0.0085	0.1405
G2	0.0337	0.0180	0.0309	0.0203	0.0253	0.1378
G13	0.0168	0.0112	0.0280	0.0127	0.0115	0.0965

Transferring LR coefficients between glaciers results in a high RMSE across the mountain range. The lowest overall RMSE (0.2051 m w.e.) results from calculating a LR using all available observations. Elevation is the only significant topographic predictor for a range-scale LR ($\beta_z = 0.0525$).

Quantifying effects of variability

Specific WSMB is affected by variability introduced when interpolating density (density variability), when calculating grid cell SWE values (SWE variability), and when interpolating observations (interpolation variability). We find that when using a LR, interpolation variability has a larger effect on WSMB uncertainty than density variability or SWE variability. The probability density function (PDF) that arises from SWE variability is much narrower than the PDF that arises from interpolation variability (Figure 8 and Table 3). WSMB uncertainty found with SK interpolation is dominated by interpolation variability (Table 3).

The total WSMB uncertainty from SK interpolation is 3 to 5 times greater than uncertainty from LR interpolation. The PDFs overlap between the two interpolation methods although the PDF peaks are lower when SK is used for Glaciers 2 and 13 and higher for Glacier 4. SK results in WSMB distributions that overlap between glaciers and there is also a small probability of estimating a WSMB value of 0 m w.e. for Glaciers 2 and 13. LR results in overlapping WSMB distributions for Glaciers 2 and 4, with the PDF peak of Glacier 4 being slightly higher than that of Glacier 2.

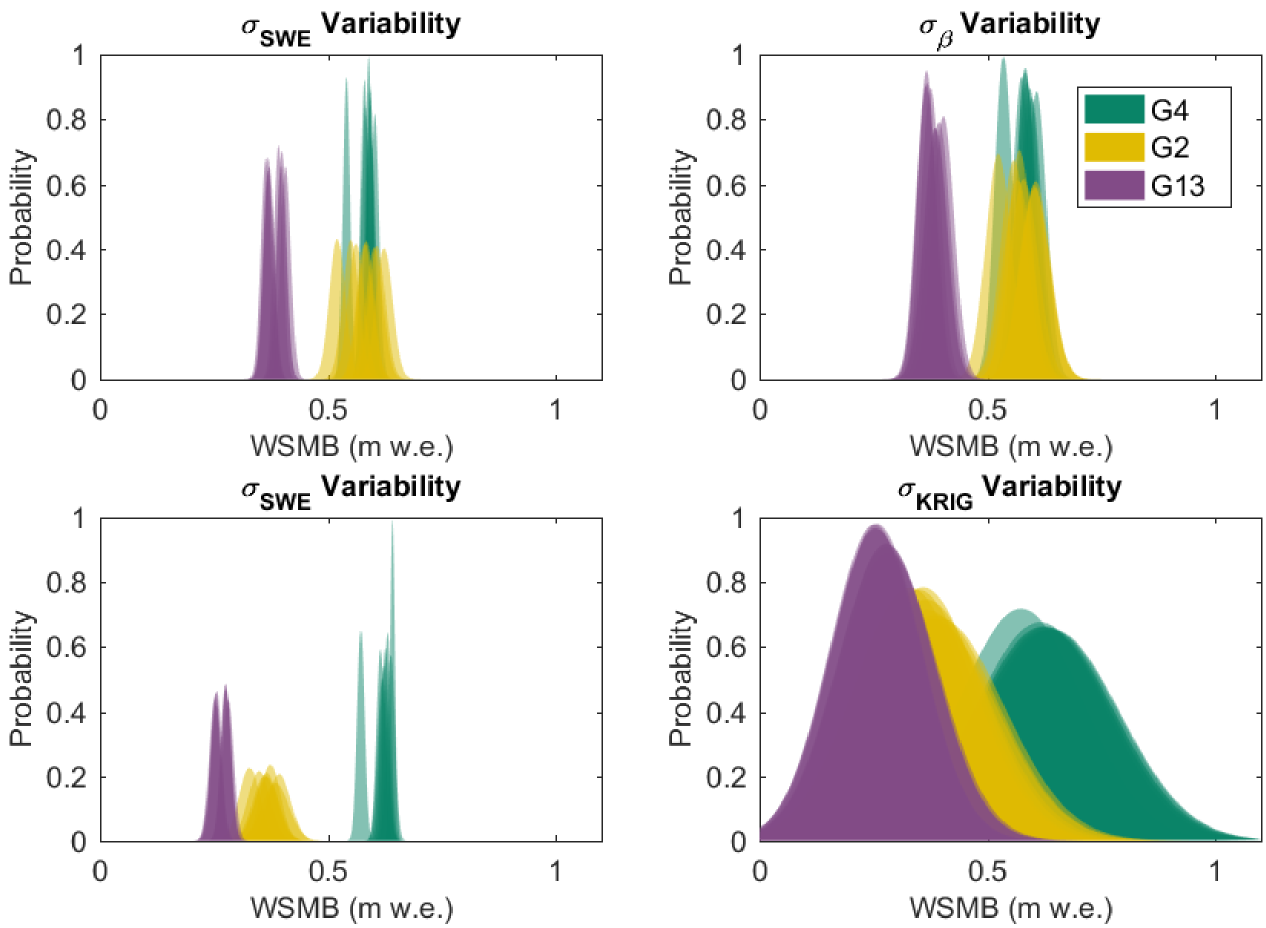

Fig. 8. Probability density functions (PDFs) fitted to distributions of specific winter surface mass balance (WSMB) values that arise from SWE variability (σ_{ZZ}) and interpolation variability (σ_β or σ_{KRIG}). Results from a linear regression interpolation (top panels) and simple kriging (bottom panels) are shown. Each PDF is calculated using one of eight density interpolation methods for Glacier 4 (G4), Glacier 2 (G2) and Glacier 13 (G13).

The spatial patterns of WSMB uncertainty are affected by density, SWE, and interpolation variability (Figure 10). For both LR and SK, the greatest variability in estimated SWE occurs in the accumulation area. When LR is used, estimated SWE is highly sensitive to the elevation regression parameter. In the case of SK, variability is greatest in areas far from observed SWE, which consist of the upper accumulation area on Glaciers 2 and 13. Variability is greatest on Glacier 4 when LR interpolation is used at the upper edges of the accumulation area, which correspond to the locations with extreme values of the wind redistribution parameter. When SK is used for interpolation on Glacier 4, variability is greatest at the measured grid

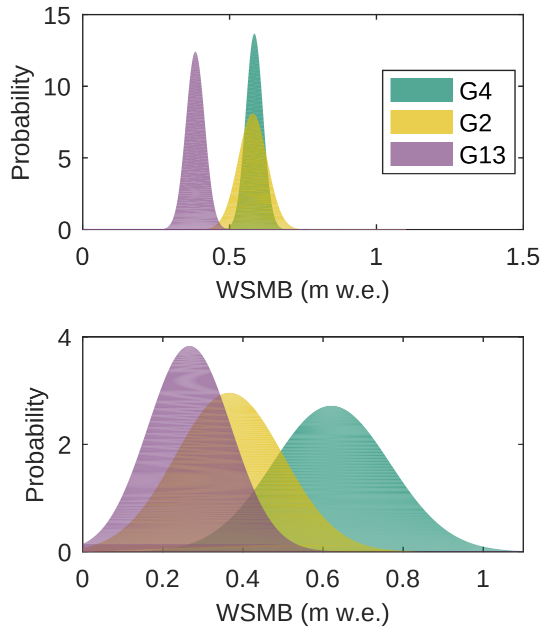

Fig. 9. Probability density functions (PDFs) fitted to distributions of specific winter surface mass balance (WSMB) values estimated using linear regression (top) or simple kriging (bottom). Each PDF includes density variability, SWE variability and interpolation variability for Glacier 4 (G4), Glacier 2 (G2) and Glacier 13 (G13).

cells, which highlights the short correlation length and the large effect of density interpolation on the SK accumulation estimate.

DISCUSSION

The goal of this study is to compile a comprehensive sweep of choices and assumptions present in the process of translating snow measurements to winter mass balance. The discussion focuses on evaluating the choices we made within the four main steps needed to estimate accumulation. We then discuss the relative importance of sources of variability when estimating specific WSMB.

Measurements

Snow probing is the simplest and oldest method used to determine accumulation. Direct measurement of snow depth means that no data processing or corrections are needed and depth uncertainty is simple to quantify by taking multiple depth measurements close together (Sold and others, 2013). However, probing is time consuming and this limits the number of measurements that can be made. Further, measurement is limited

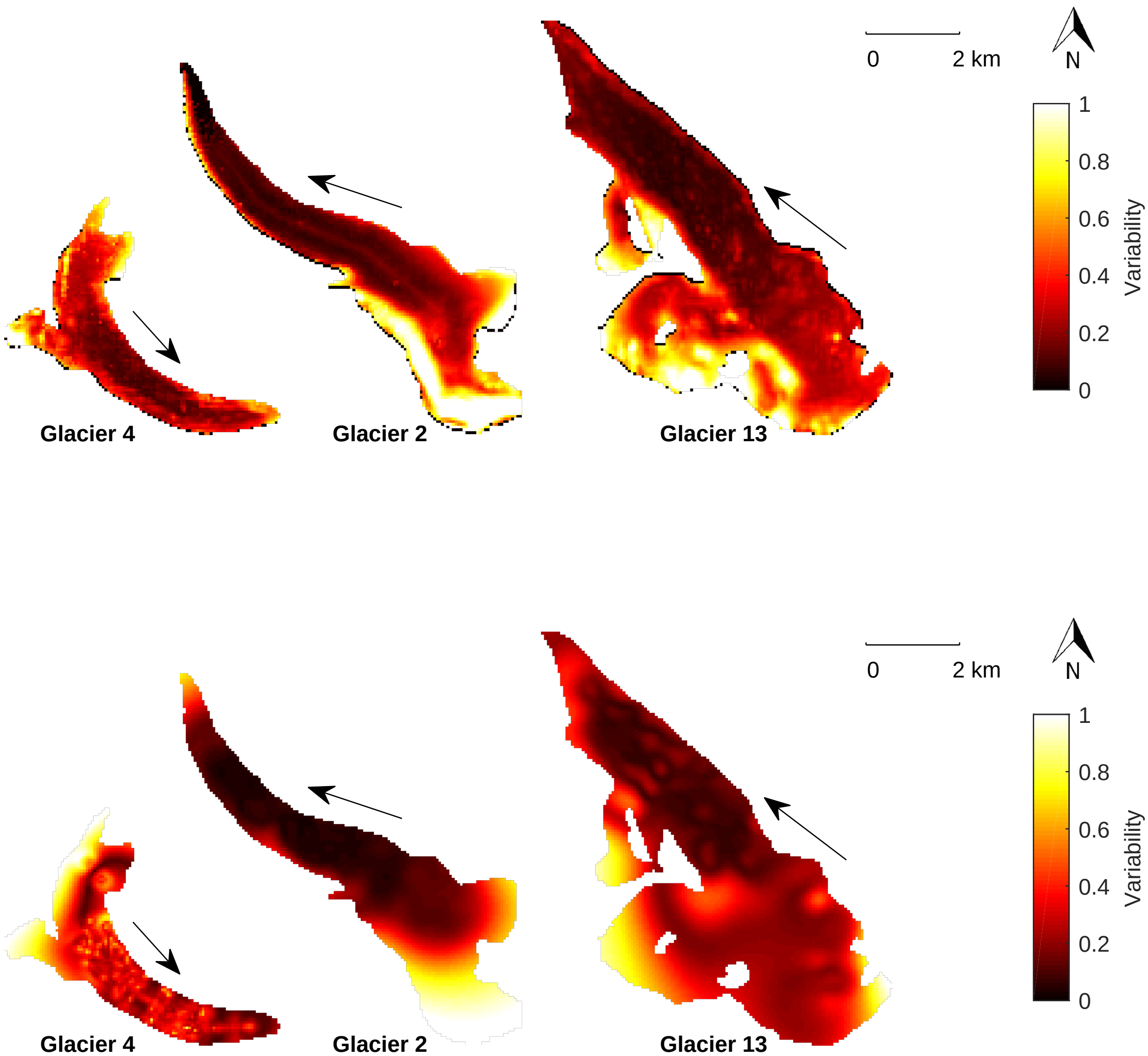

Fig. 10. Variability of SWE estimated using linear regression (top) and simple kriging (bottom). Variability is a relative quantity measured by taking the sum of differences between one hundred estimates of distributed WSMB that include SWE variability and, in the case of linear regression, regression variability. The sum is then normalized for each glacier. Glacier flow directions are indicated by arrows.

to areas that are both accessible and safe for researchers. In complex terrain many areas cannot be surveyed, resulting in data gaps (Deems and Painter, 2006; Sold and others, 2014). Sold and others (2013) noted that this systematic bias can result in incorrect values of glacier-wide accumulation, particularly because inaccessible areas such as cliffs and ridges have relatively shallow accumulations (due to wind erosion), while heavily crevassed areas can accumulate deep snow packs. Despite these limitations, we chose to use snow probing for this study to minimize cost, simplify field logistics and reduce data processing time. By focusing on simple field methods that are easy to execute, we hope to make our conclusions and recommendations for estimating WSMB more broadly applicable and reproducible.

Most contemporary studies that investigate glacier accumulation use ground penetrating radar (GPR), either airborne or ground-based, to obtain continuous and extensive snow depth profiles (e.g. Winther and others, 1998; Machguth and others, 2006; Gusmeroli and others, 2014; McGrath and others, 2015). GPR snow surveys, especially when airborne, are able to quickly collect data over large areas and terrain accessibility does not hamper data collection. The main limitation of GPR is the misinterpretation of radargram layers, especially in areas where the snow-ice boundary is ill-defined such as the accumulation area or heavily crevassed terrain (Machguth and others, 2006; Gusmeroli and others, 2014; McGrath and others, 2015). Complications also arise when radar wave speed is altered due to varying snow density and liquid water content. Further, there is no universal procedure for obtaining snow depth data so methodology is difficult to reproduce. Results therefore depend on available equipment, selection of processing parameters and radargram processing algorithms (Sold and others, 2013).

DEM differencing has also been used to estimate glacier-wide accumulation (Deems and Painter, 2006; Nolan and others, 2015). This method allows for maximal spatial data coverage. However, DEM differencing requires knowledge of glacier dynamics to account for surface changes and data collection, either by lidar or photogrammetry, is subject to considerable errors and noise (Deems and Painter, 2006; Nolan and others, 2015).

Our study suffers from lack of data in the accumulation area. Snow probing cannot be used in the accumulation area because the snow-firn transition is often difficult to determine and ice lenses can be misinterpreted. Both GPR and DEM differencing are also not reliable in the accumulation area. Observing the snow-firn transition using GPR can be difficult because the density difference between snow and firn can be small. Obtaining an accurate snow surface and correlating two DEMs for differencing can also be difficult in the accumulation area because camera sensor noise and poor lighting can result in significant topographic noise. Measuring SWE in the accumulation area is difficult and subject to large errors regardless of the data collection method.

We measured snow density by sampling a snow pit (SP) and by using a Federal Sampler (FS). We found that FS and SP measurements are not correlated and that FS density values are positively correlated with snow depth. This positive relationship could be a result of physical processes, such as compaction, and/or artefacts during data collection. However, it seems more likely that this trend is a result of measurement artefacts for a number of reasons. First, the range of densities measured by the Federal sampler is large (225–410 kg m^{-3}) and the extreme values seem unlikely to exist at these study glaciers, which experience a continental snow pack with minimal mid-winter melt events. Second, compaction effects would likely be small at these study glaciers because of the relatively shallow snow pack (deepest measurement was 340 cm). Third, no linear relationship exists between depth and SP density ($R^2 = 0.05$). Together, these reasons lead us to conclude that the Federal Sampler measurements are biased.

The FS appears to oversample in deep snow and undersample in shallow snow. Oversampling by small diameter (area of 10–12 cm^2) sampling tubes has been observed in previous studies, with a percent error between +6.8% and 11.8% (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 (e.g. Molotch and others, 2005). Dixon and Boon (2012) attributed oversampling to slots "shaving" snow into the tube as it is rotated as well as cutter deign forcing snow into the tube. Beauont and Work (1963) found that only when snow samples had densities greater than 400 kg m^{-3} and snow depth greater than 1 m, the FS oversampled due to snow falling into the greater area of slots. Undersampling is likely to occur due to snow falling out of the bottom of the sampler (Turcan and Loijens, 1975). It is likely that this occurred during our study since a large portion of the lower elevation snow on both Glaciers 2 and 13 was melt affected and thin, allowing for easier lateral displacement of the snow as the sampler was inserted. For example, on Glacier 13 the snow surface had been affected by radiation melt (especially at lower elevations where the snow was shallower) and the surface would collapse when the sampler was inserted into the snow. It is also difficult to measure the weight of the sampler and snow with the spring scale when there was little snow because the weight was at the lower limit of what could be detected by the scale.

Distributed density

We choose four different density interpolation methods and keep SP and FS measurements separate. Despite the wide range of measured density values and variety in density interpolation, density does not appear to strongly affect WSMB estimates and is usually not the dominant source of WSMB uncertainty. We have relatively few density measurements throughout the study glaciers, as is common in many snow surveys, and we believe our FS measurements to be biased. Therefore, our preferred density interpolation is to use a glacier-wise mean of SP densities. This method employs common snow density measurement techniques and is easily transferable to other study areas. While using a glacier-wide mean snow density omits spatial variability in snow density (Wetlaufer and others, 2016), it does not assume unmeasured spatial correlation or trends in density.

Wetlaufer and others (2016) found that distributed density from snow depth and density results in more variability than directly measuring SWE using a FS. Since SWE is more time consuming to measure than snow depth, future studies could consider decreasing the number of sample locations but directly measuring SWE to reduce the variability in distributed density at a measurement location. A detailed investigation of FS error is needed to contrain the variability introduced when using FS to directly measure SWE.

Grid cell average

López-Moreno and others (2011) completed an extensive survey of snow depth variability at the plot scale (10×10 m) in the Spanish Pyrenees Mountains. The authors concluded that at least five measurement points are needed in each plot to ensure estimation error is <10% for plot averaged SWE. Their suggestion amounts to at least 80 measurement points for the grid cells in this study (40×40 m). Rather than gridded or random sampling, as executed by the authors, we suggest a zigzag sampling scheme. The zigzag offered a comprehensive estimation of snow depth variability in a grid cell. Shea and Jamieson (2010) proposed this linear-random sampling scheme and showed that it performs as well as pure-random sampling in detecting spatial correlations and is considerably easier to execute.

Since such a large number of points are needed to characterize the variability in a grid cell there is little advantage to measuring and then averaging snow depth at multiple measurement locations. Rather, time should be spent extensively characterizing grid-cell variability in a few locations and to then decrease the spacing of transect measurements to extend their spatial coverage over the glacier. In our study, the grid cell variability appeared to be captured with dense sampling in select grid cells but the basin-scale variability was not captured because sampling was limited to the ablation area. By decreasing transect spacing, grid cells would only have one or two measurements but more grid cells could be measured.

Interpolated SWE

Linear regression (LR) is chosen for this study because topographic parameters can be used as proxies for physical processes that affect snow distribution. Elevation was the only topographic parameter that offered relevant insight into topographic controls on accumulation. Even so, elevation had little predictive ability for Glacier 4 and the correlation was moderate on Glacier 13. Elevation affects snow distribution through melt at lower elevation due to higher temperatures, as well as increased precipitation and preservation of snow at higher elevation. It is possible that the elevation correlation was accentuated during the field campaign due to warmer than normal temperatures and an early ($1-2$ weeks) start to the melt season (Yukon Snow Survey Bulletin and Water Supply Forecast, May 1, 2016). The southwestern Yukon winter snow pack in 2015 was also well below average, likely resulting in the effects of early melt onset to be emphasized. Glacier 4 had deeper snow and cloudier conditions during the field campaign so perhaps a correlation between SWE and elevation had not manifested.

Our mixed insights into dominant predictors of accumulation are consistent with the conflicting results present in the literature. Many snow accumulation studies have found elevation to be the most significant predictor of SWE (e.g. Machguth and others, 2006; McGrath and others, 2015). However, accumulation-elevation gradients vary considerably between glaciers (Winther and others, 1998) and other factors, such as orientation relative to dominant wind direction and glacier shape, have been noted to affect accumulation distribution (Machguth and others, 2006; Grabiec and others, 2011). Machguth and others (2006), Grünewald and others (2014) and Kirchner and others (2014) observed elevation trends in snow accumulation for the lower parts of their study basins but no correlation or even a decrease in SWE with elevation for the upper portion of their basins. Helbig and van Herwijnen (2017) suggest that an increase in accumulation with elevation can better be approximated by a power law. There are also a number of accumulation studies on glaciers that found no significant correlation between accumulation and topographic parameters and the highly variable snow distribution was attributed to complex local conditions (e.g. Grabiec and others, 2011; López-Moreno and others, 2011).

Wind redistribution and preferential deposition of snow is known to have a large influence on accumulation at sub-basin scales. Dadic and others (2010) used a dynamic model to show that variations in snow depth are caused by preferential deposition, which is well correlated with mean wind speed. Interactions between local wind fields and complex topography create uplift and down drafts that affect snow deposition. Pomeroy and others (1999) looked at snow mass balance in a non-glacierized alpine basin within the St. Elias and found that up to 79% of the snow was redistributed from alpine areas to (primarily) hillsides, where accumulation was tripled. In the study basin, measured accumulation ranged from 54% to 419% of the actual snowfall. The wind redistribution parameter used in the study is found to be a small but significant predictor of accumulation on Glacier 4 (negative correlation) and Glacier 2 (positive correlation). This result indicates that wind likely has an impact on snow distribution but that the wind redistribution parameter is perhaps not the most appropriate way to characterize the effect of wind on our study glaciers. For example, Glacier

4 is located in a curved valley with steep side walls so having a single cardinal direction for wind may be inappropriate. Examining wind redistribution parameter values that assume wind moving up or down glacier and changing direction to follow the valley could allow the wind redistribution parameter to explain more of the variance in SWE. Further work with dynamic modelling that uses high resolution weather modelling and considers small scale mountain topography is also needed to better understand relevant scales of snow deposition and to aid in developing more appropriate wind parameterizations. In our study, the scale of deposition may be smaller than the resolution of the Sx parameter in the relatively large DEM grid cells. An investigation of the wind redistribution parameter with finer DEM resolution is also needed. Accounting for wind in snow distribution models is especially important because it plays a dominant role in spatial patterns of accumulation (Winstral and others, 2013). A universal predictor of distributed SWE therefore continues to elude researchers and accumulation variability due to complex interactions between topography and the atmosphere needs to be considered when estimating winter mass balance.

Since we were unable to measure SWE in grid cells that corresponded to the extreme values of all topographic parameters, we must extrapolate linear relationships. The accumulation area, where there are few observations, is most susceptible to extrapolation errors. This area typically also has the highest SWE values, affecting the specific WSMB estimated for the glacier. In our study, the dependence of SWE on elevation, especially on Glacier 2, means that LR extrapolation results in almost 2 m w.e. estimated in the parts of the accumulation area. This exceptionally large estimate of SWE is unlikely for a continental snow pack. As described above, snow in the accumulation area has been shown to have no correlation or a negative correlation with elevation and wind effects have been observed. Therefore, extrapolating a LR that is fitted to predominantly ablation area SWE values is likely erroneous. Future studies need to focus on collecting SWE observations in the accumulation area, even if it means collecting fewer observations in the ablation area. Observations in the accumulation area can be used both to characterize accumulation patterns in the upper portions on a glacierized basin and to generally increase the spatial extent and topographic parameter range coverage of observations.

While a LR can be used to predict distributed SWE in other basins, we found that transfer of LR coefficients between glaciers results in large estimation error. The LR fitted to all observed data produced the best overall predictor of SWE in the Donjek Range, so transferability of LR is also limited in our study area. Our results are consistent with Grünewald and others (2013), who found that local statistical models are able to perform well but they cannot be transferred to different regions and that regional-scale models are not able to explain the majority of variance. Therefore, if the intent of a study is to estimate range-scale accumulation it is perhaps best to sparsely sample many glaciers and to make assumptions about variability within the basin rather than conducting a detailed study of one basin. The inter-basin variability in our study range is greater than the intra-basin variability.

For all study glaciers, simple kriging (SK) is a better predictor of observed SWE. However, the WSMB uncertainty that arises from using SK is large, and unrealistic values of 0 m w.e. WSMB can be estimated. Such a large uncertainty is undesirable when estimating WSMB. Our observations are generally limited to the ablation area so SK estimates an almost uniform distribution of SWE in the accumulation areas of the study glaciers, which is inconsistent with observations described in the literature. Extrapolation using SK is erroneous and leads to large uncertainty in estimating WSMB, which further emphasis the need for SWE observations in the accumulation area.

SK cannot be used to understand physical processes that may be controlling snow distribution and cannot be used to estimate accumulation beyond the study area. However, fitted kriging parameters, including the nugget and spatial correlation length, can provide insight into important scales of variability. Glaciers 2 and 13 have long correlation lengths and small nuggets indicating variability at large scales. Conversely, Glacier 4 has a short correlation length and large nugget, indicating that accumulation variability occurs at small scales. Using a higher resolution sampling design and DEM may allow us to capture more of the variability on Glacier 4 and to perhaps improve the predictive ability of both LR and SK interpolation.

A number of studies that relate SWE to topographic parameters have found success when using a regression tree interpolation model, which is a non-linear regression method (e.g. Elder and others, 1998; Erickson and others, 2005; López-Moreno and others, 2010). Many relationships between accumulation and topographic parameters have been observed to be non-linear so regression tree are valuable in snow modelling and may yield improved results (Erxleben and others, 2002; Molotch and others, 2005).

Quantifying effects of variability

Interpolation variability is the greatest contributor to WSMB uncertainty for both SK and LR. This uncertainty arises from extrapolation beyond the sampled region, which results in highly variability in estimated SWE in the accumulation area. To reduce WSMB uncertainty, emphasis must therefore be placed on sampling in the accumulation area and generally obtaining measurements throughout the study basin.

SWE variability is the smallest contributor to WSMB uncertainty. Therefore, obtaining the most accurate value of SWE to represent a grid cell, even a relatively large grid cell, does not need to be a priority when designing a snow survey. Extensively measuring SWE variability in a few locations using a zigzag design appears to be a good constraint on SWE variability. Many parts of a glacier though are characterized by a relatively smooth surface, with roughness lengths on the order of centimeters (Hock, 2005) resulting in low snow depth variability. However, we assume that the sampled grid cells are representative of the variability across the entire glacier, which is likely not true for areas with debris cover, crevasses and steep slopes. Snow depth variability can be large and thus exert a dominant control on snow distribution in these area (McGrath and others, 2015). Effects of SWE variability in either smaller or larger grid cells could also be different so further investigation is needed.

Using a Monte Carlo experiment to propagate variability allowed us to quantify effects of variability on estimates of WSMB. However, our analysis did not include variability arising from a number of data sources. Error associated with SP and FS density measurement is not included but we believe that this error is likely to be encompassed in the wide range of density interpolation methods. DEM vertical and horizontal error are not considered in the Monte Carlo experiment mainly because there is no DEM validation data at our study location. Error associated with estimating measurement locations, which is a combination of hand-held GPS error, distance of observers from GPS and travel along a straight line, is also not considered. However, we feel that this source of error is encompassed in the variability estimated from zigzag measurements.

While quantifying WSMB uncertainty is an important feature of accumulation studies, we also need to consider how much uncertainty we are willing to accept. At what point do we say that we are not able to make an accurate estimate of WSMB? In our study, are we able to say that our most probable estimate of WSMB found using SK is appropriate to report when the uncertainty is so large? Further, is our assumption that we have captured the majority of uncertainty in our variability analysis sufficient?

Mountain range accumulation gradient

An accumulation gradient is observed for the continental side of the St. Elias Mountains (Figure 11). Accumulation data is compiled from Taylor-Barge (1969), the three glaciers presented in this paper, as well as two snow pits we dug at the head of the Kaskawalsh Glacier in May 2016. The data show a linear decrease in observed SWE as distance from the main mountain divide (identified by Taylor-Barge (1969)) increases, with a gradient of -0.024 m w.e. km^{-1}. This relationship indicates that glacier location within a mountain range also affects glacier-wide WSMB. Interaction between meso-scale weather patterns and mountain topography is a major driver of glacier-wide accumulation. Further insight into mountain-scale accumulation trends

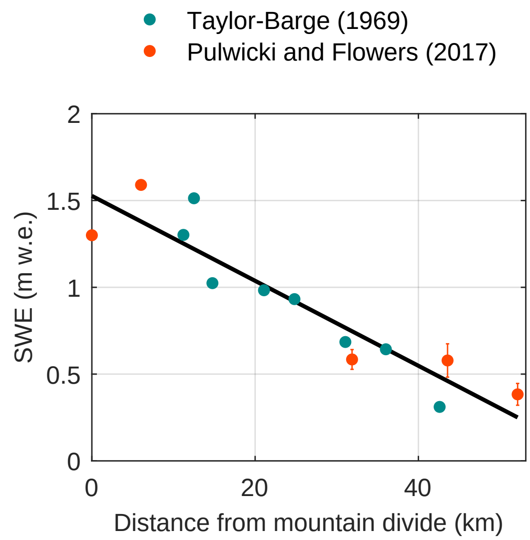

Fig. 11. Relation between SWE and linear distance from St. Elias mountain divide, located at the head of the Kaskawalsh Glacier. Blue dots are snow pit derived SWE values from (Taylor-Barge, 1969). Orange dots farthest from the divide are mean WSMB from Glaciers 4, 2 and 13, with 95% confidence interval using a linear regression interpolation. Orange dots close to the divide are snow pit derived SWE value at two locations in the accumulation area of the Kaskawalsh Glacier collect in May 2016. Black line indicates line of best fit ($R^2 = 0.85$).

can be achieved by investigating moisture source trajectories and orographic precipitation contribution to accumulation.

Limitations and future work

Extensions to this work could include an investigation of experimental design, examining implications of a non-linear SWE elevation trend, examining the effects of DEM grid size on WSMB 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. Therefore, it is valuable to investigate how best to reduce our sampling design and measurement spacing while maintaining a reasonable estimate of distributed WSMB. López-Moreno and others (2010) examined data reduction in a ~6 km^2 basin and found a non-linear response of model stability and accuracy to sample size. The authors concluded that $200-400$ observations are needed to obtain accurate and robust models. 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.

A non-linear SWE-elevation trend has been documented in a number of studies so it would be valuable to further investigate this relationship. Although more observations in the accumulation area are needed to confirm this relationship on our study glaciers, we could apply a variety of non-linear elevation trends to investigate their effects on WSMB estimates.

DEM grid cell size had a large influence on the resolution of topographic features (López-Moreno and others, 2010), which can have implications for calculating a LR for SWE data. DEM grid cell 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). Zhang and Montgomery (1994) found that simulating geomorphic and hydrological process for many landscapes is best accomplished with a 10-m grid cell size, which is an optimal compromise between increasing resolution and large data volumes. The authors found that a 30- and 90-m grid cell size were insufficient in resolving terrain features in a moderate to steep gradient topography. López-Moreno and others (2010) state that a grid cell size of 5 m is need to reliably represent terrain and to accurately identity solar radiation, curvature and slope. The authors conclude that relevant topographic parameters in their $\sim$6 km^2 basin are completely lost at grid sizes greater than 55×55 m making DEMs with a coarse resolution inappropriate for modelling snow pack. Further, the importance of topographic parameters in predicting SWE was correlated with DEM grid size. A decrease in spatial resolution of the DEM resulted in a decrease in the importance of curvature and an increase in the importance of elevation and, to a lesser degree, solar radiation. These results corroborated Kienzle (2004), who found that curvature was the main predictor of SWE with a high resolution DEM. To further confound the use of DEMs to estimate SWE, Molotch and others (2005) found that estimated SWE distributions were dependent on the DEM chosen. Even different DEMs with similar spatial resolutions can generate significantly different topographic parameters and resulting SWE distributions. A detailed and ground controlled DEM is therefore needed to identify the features that drive accumulation variability.

Future studies could also evaluate the effects of DEM uncertainty on elevation and derived topographic parameters. Wechsler and Kroll (2006) used a Monte Carlo experiment to quantify deviation of topographic parameters due to DEM error. The authors found that elevation did not significantly deviate but slope and other hydrological parameters such as catchment area and topographic index were significantly affected.

Guo-an and others (2001) also conducted an DEM error analysis and found that the accuracy of hydrological topographic parameters was closely related to the the vertical resolution of the DEM. Errors were especially large in smooth plain areas with slope less than 4 degrees.

It appears then that topographic parameters included in a LR and the uncertainty in estimating WSMB are dependant on the resolution of DEM grid cells. Future accumulation investigations should therefore focus on obtaining a high resolution DEM and quantifying effects of DEM variability on WSMB. There is a strong need for a better understanding of the effects of DEM error and grid size on glacier accumulation. The majority of published studies focus on hydrological modelling and the study areas are non-glacierized. Glaciers present different accumulation patterns and surface topography so the DEM resolution and uncertainty may also differ.

Temporal variability in accumulation is not considered in our study. While this limits the extent of 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) analysed 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. Reliability maps were then used to reduce the sampling scheme to one index site as well as a transect with 50 m elevation intervals for each glacier and winter balance was estimated to within 0.15 m w.e. However, the temporal transferability of terrain-based parametrization is not always reliable. Walmsley (2015) also found several strongly irregular snow spatial distribution years that were inconsistent with the overall reduced sampling schemes. Revuelto and others (2014) also noted that snow distribution variability could not be explained by their model in low snow years.

CONCLUSION

We estimate spatial accumulation patterns and specific winter surface mass balance (WSMB) for three glaciers in the St. Elias mountains from extensive snow depth and density sampling. Range scale accumulation is sampled by selecting three glaciers along a precipitation gradient found on the continental side of the mountain range. We sample basin scale accumulation by measuring snow depth along linear and curvilinear transects throughout the ablation area of each glacier. Snow depth variability within a DEM grid cell is sampled using a linear-random design. Point scale accumulation is sampled by taking three to four snow depth measurements at each measurement location. Snow density is measured using a wedge cutter in snow pits in three locations on each glacier as well as a Federal Sampler in a number of locations throughout the glacier. Snow water equivalent (SWE) is then calculated by interpolating the measured density values. Four interpolation methods are used for the snow pit and Federal Sampler density measurements, which are found to be uncorrelated. An average SWE value for each measured grid cell is then calculated. The grid cell values of SWE are interpolated to estimate distributed accumulation. Two interpolation methods are used. Liner regression (LR) relates SWE values to topographic parameters, which are derived from a DEM and serve as proxies for physical processes that affect snow distribution. We choose to include elevation, distance from centreline, slope, aspect, curvature, "northness" and a wind redistribution parameter as topographic parameters. Cross-validation and model averaging are used to reduce overfitting of the LR. Simple kriging (SK) is also used to interpolate SWE. SK assumes spatial correlation of the quantity being interpolated and fitted kriging parameters, including the correlation length and nugget, can provide insight into scales of spatial variability. WSMB for each glacier is then calculated as the average SWE for a grid cell.

Overall, elevation is the dominant driver of SWE distribution but results vary between glaciers. Accumulation spatial patterns and scales of variability are considerably different on Glacier 4 when compared to Glaciers 2 and 13. Glaciers 2 and 13 have a dominant elevation-accumulation trend and long spatial correlation lengths. No topographic parameters were able to explain snow distribution on Glacier 4 and a short correlation length and large nugget indicate variability at shorter length scales. Our results also suggest that wind redistribution and preferential distribution are significant drivers of SWE distribution but these effects are not captured by the wind redistribution parameter used. Improved modelling of wind effects on accumulation through modification of the wind redistribution parameter as well as increased physical modelling are needed. A LR applied to our study glaciers resulted in little insight into dominant physical processes indicating that accumulation is controlled by complex interactions between topography and the atmosphere and that a finer resolution DEM is needed to resolve SWE distribution and potentially relevant topographic parameters, such as curvature and wind redistribution.

Glacier accumulation is strongly affected by interactions between topography and atmospheric processes at the basin- and range-scale. Although we could not conclusively identify processes at the basin scale due to low predictive ability of the LRs, there is a dominant trend in accumulation at the range scale. We identify a clear linear decrease in SWE with increased distance from the main topographic divide along the continental side of the St. Elias Mountains. This trend indicates that glacier location within a mountain range has a large influence on WSMB. Further investigation of meso-scale weather patterns could provide insight into relevant processes that affect accumulation at the range scale.

We also quantify the effects of variability from density interpolation, grid cell SWE calculation as well as interpolation method on uncertainty in estimating WSMB. We conduct a Monte Carlo experiment to propagate variability through the process of estimating accumulation from snow measurements. The largest source of uncertainty in our study stems from variability in interpolation method, both within and between methods. We find that SK results in up to five times greater uncertainty than LR and the distribution encompasses unrealistic estimates of WSMB. Spatial distribution of interpolation variability indicates that the accumulation area is the greatest area of uncertainty. This large variability is a result of the accumulation area being poorly sampled, sensitive to estimates of dominant regression coefficients, and having the largest values of estimated SWE within the glacier. To better constrain WSMB estimates, future studies should focus on obtaining snow measurements in the accumulation area at the expense of collecting less data overall. Density and SWE variability are found to be small contributors to WSMB uncertainty. We conclude that the choice of interpolation method in combination with sampling design, especially in the accumulation area, has a major impact on the uncertainty in WSMB estimates.

References

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)

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)

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

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)

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)

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-)

Erickson TA, Williams MW and Winstral A (2005) Persistence of topographic controls on the spatial distribution of snow in rugged mountain terrain, Colorado, United States. *Water Resources Research*, **41**(4) (doi: 10.1029/2003WR002973)

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

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)

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)

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, 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)

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)

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)

Kirchner PB, Bales RC, Molotch NP, Flanagan J and Guo Q (2014) LiDAR measurement of seasonal snow accumulation along an elevation gradient in the southern Sierra Nevada, California. *Hydrology and Earth System Sciences*, **18**(10), 4261–4275, ISSN 1027-5606

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

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-ÂĂÂŘbased 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)

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

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)

Nolan M, Larsen C and Sturm M (2015) Mapping snow-depth from manned-aircraft on landscape scales at centimeter resolution using Structure-from-Motion photogrammetry. *The Cryosphere Discussions*, **9**, 333–381 (doi: 10.5194/tcd-9-333-2015)

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

Pomeroy J, Hedstrom N and Parviainen J (1999) The snow mass balance of Wolf Creek, Yukon: effects of snow sublimation and redistribution. *Wolf Creek, Research Basin: Hydrology, Ecology, Environment, edited by: Pomeroy, JW and Granger RJ*, 15–30

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)

Revuelto J, López-Moreno JI, Azorín-Molina C and Vicente Serrano SM (2014) Topographic control of snowpack distribution in a small catchment in the central Spanish Pyrenees: Intra-and inter-annual persistence (doi: 10.5194/tc-8-1989-2014)

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

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)

Sold L, Huss M, Eichler A, Schwikowski M and Hoelzle M (2014) Recent accumulation rates of an alpine glacier derived from firn cores and repeated helicopter-borne GPR. *The Cryosphere Discussions*, **8**(4), 4431–4462 (doi: 10.5194/tcd-8-4431-201)

Taylor-Barge B (1969) The summer climate of the St. Elias Mountain region. Technical report, DTIC Document

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

Watson FG, Anderson TN, Newman WB, Alexander SE and Garrott RA (2006) Optimal sampling schemes for estimating mean snow water equivalents in stratified heterogeneous landscapes. *Journal of Hydrology*, **328**(3), 432–452 (doi: 10.1016/j.jhydrol.2005.12.032)

Wechsler SP and Kroll CN (2006) Quantifying DEM Uncertainty and its Effect on Topographic Parameters. *Photogrammetric Engineering & Remote Sensing*, **72**(9), 1081–1090, ISSN 0099-1112

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**(1), 3 (doi: 10.3390/hydrology3010003)

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

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)