METHODS

Estimating accumulation from measured values of snow depth and density requires a number of processing steps, each entailing a set of assumptions. In this study, the steps include (1) measuring snow depth and density, (2) interpolating snow density to estimate snow water equivalent (SWE), (3) averaging measurements within one digital elevation model (DEM) grid cell and (4) interpolating grid-cell SWE values to estimate distributed SWE. To estimate the specific winter surface mass balance (WSMB) we calculate the mean SWE for a grid cell from the estimated distributed SWE.

Measuring snow depth and density

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 avoids bias, allow for the greatest variability to be measured, and minimize distance travelled (Shea and Jamieson, 2010).

We measured accumulation at three glaciers to account for range-scale variability (Clark and others, 2011). Snow depth was measured along linear and curvilinear transects to encompass basin-scale variability and at each measurement location, three values of snow depth were recorded to account for point-scale variability (Clark and others, 2011). To exploit the precipitation gradient in the St. Elias Mountains, Yukon (Taylor-Barge, 1969) 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). We selected centreline and transverse transects with sample spacing of $10 - 60$ m (Figure 1d) to capture previously established correlations between elevation and accumulation (e.g. 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 observer, 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.

When estimating accumulation, snow depth variability at scales less than the grid-size of satellite derived elevation models is assumed to be caused by random effects that are unbiased and unpredictable (Watson and others, 2006). To capture grid-scale variability, 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) aligned with randomly selected DEM grid cells distributed throughout the ablation zone.

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 the weighted the sampled 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.

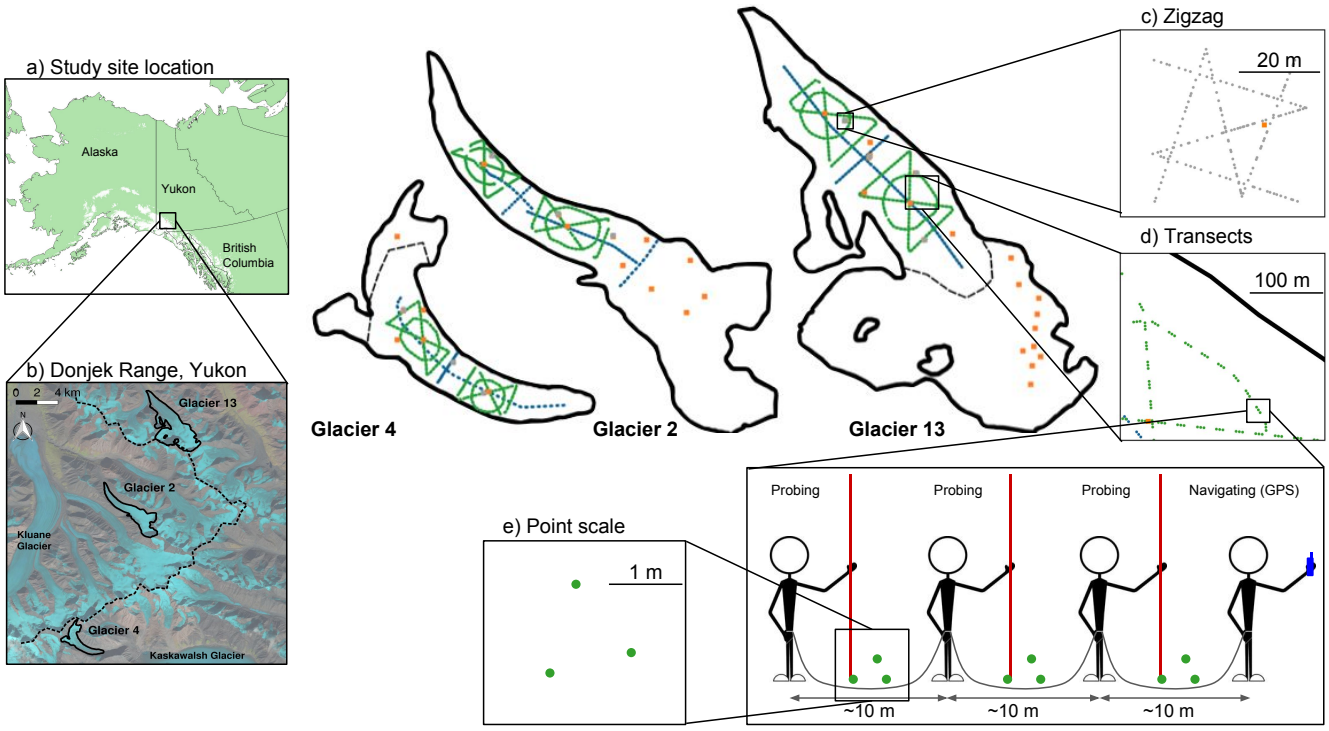

Fig. 1. Sampling design for Glaciers 4, 2 and 13, located in the the Donjek Range, Yukon (a,b). Centreline and transverse transects are shown in blue dots, hourglass and circle design are shown in green dots and zigzag measurements are shown as grey dots (c). Linear and curvilinear transects typically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

Estimating SWE

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 averaging

We average SWE values within each SPOT-5 DEM-aligned grid cell (Korona and others, 2009). 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.

To encompass variability at spatial scales smaller than a DEM grid cell, we measured snow depth extensively (135 – 191 points) using a 'zigzag' configuration (Figure 1c). Zigzag locations were randomly chosen within the upper (∼2350 m a.s.l.), middle (∼2250 m a.s.l.), and lower portions (∼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 (∼2200 m a.s.l.). SWE variability is assumed to be normally distributed about the mean SWE at a measured grid cell with a standard deviation equal to the average standard deviation of all zigzags on a glacier.

Interpolation

SWE data were interpolated for each glacier using linear regression (LR), simple kriging (SK), as well as regression kriging (RK). 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 wind exposure/shelter 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 (**?**). 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).

Topographic parameters were derived from a SPOT-5 DEM (40×40 km) (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 direct 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).

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.

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 cells smoothing window to the DEM, which was then used to calculate curvature, slope, aspect and "northness".

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 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).

Regression kriging (RK) (Hengl and others, 2007) estimates were found by first calculating the residuals from the LR estimate at measurement locations. Then, distributed residuals were estimated using SK, and the linear regression SWE and kriged residuals were added to obtain a RK estimate of distributed SWE. Regression kriging can be thought of as an intermediate between pure kriging (no correlation with topographic parameters and large residuals) and pure regression (high correlation with topographic parameters and small residuals) and can be more strongly skewed to either end-member based on the strength of the regression correlation (Hengl and others, 2007).

Quantifying effects of variability

To provide insight on the effects of variability from (1) density interpolation, (2) observed SWE as well as (3) regression estimation on integrated winter surface mass balance, we use a Monte Carlo experiment (Metropolis and Ulam, 1949) to estimate a WSMB probability density function (PDF). For all eight density options, normally distributed random variability (mean of zero and standard deviation taken as the mean standard deviation of zigags on a glacier) is introduced to grid-cell values of SWE. LR and SK are used to estimate WSMB and the process is repeated 1000 times. Variability in regression estimation is accounted for by sampling, with covariance, a multivariate normal distribution of calculated regression coefficients. The covariance of regression coefficients is found according to **?**. The process is repeated 1000 times and adjusted β values are used to calculate winter surface mass balance.

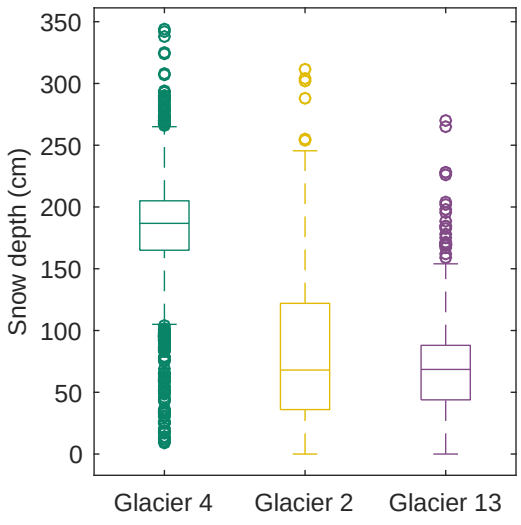

Fig. 2. Boxplot of measured snow depth on Glaciers 4, 2 and 13. The box shows first quartiles, the line within the box indicates data median, bars indicate minimum and maximum values (excluding outliers), and circles show outliers, which are defined as being outside of the range of 1.5 times the quartiles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

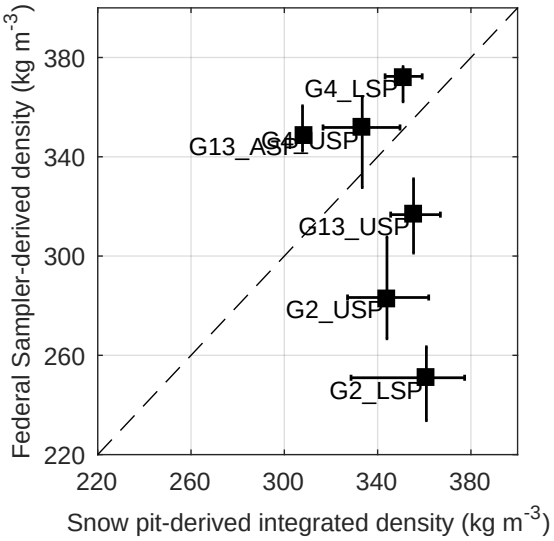

Fig. 3. Comparison of integrated density estimated using wedge cutters in a snow pit and density estimated using Federal Sampler measurements for Glacier 4 (G04), Glacier 2 (G02) and Glacier 13 (G13). Snow pits were distributed in the accumulation area (ASP), upper ablation area (USP) and lower ablation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

RESULTS

Measuring snow depth and density

A wide range of snow depth is observed on all three study glaciers (Figure). 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 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.

Estimating SWE

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.

Grid-cell averaging

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 but with fewer outliers.

SWE measurements for each zigzag are not normally distributed about the mean SWE (Figure 4). The average

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 (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	$-0.20z + 802$	$0.12z + 33$

standard deviation of all zigzags on Glacier 4 is $\sigma_{G4} = 0.027$ m w.e., on Glacier 2 is $\sigma_{G2} = 0.035$ m w.e. and on Glacier 13 is $\sigma_{G13} = 0.040$ m w.e.

Interpolation

The importance of the various topographic parameters differs for the three study glaciers (Table 4.2). The regression for Glacier 2 explains a large portion of the variance (R2=0.66), although the RMSE is higher than that of Glacier 13, for which the regression explains less variance (R2=0.40) (Figure 4.1). Glacier 4 has the least variance explained by the regression (R2=0.12) and the highest RMSE. The intercepts of the regression are similar for Glaciers 2 and 13 (âĹij0.25 m w.e.) and these are much lower than the intercept for Glacier 4 (0.62 m w.e.). The discrepancy between intercept values is a result of the poor fit of the Glacier 4 regression - the value of the intercept approaches that of the data mean (0.63 m w.e.). The residuals for Glacier 4 have a larger range than those of Glacier 2 and 13 (Figure 4.3).

- residuals are normally distributed

The most important regressor for Glacier 4 is Sx. The Sx regression coefficient is a factor of five larger than those of the remaining parameters for all density options (Table 4.3) and for the mean of all density options (Table 4.2). Sx has the highest mean semi- partial correlation (0.034) and raw correlation (0.065) (Table 4.2). The Sx coefficient is negative, which indicates less snow in âĂŸshelteredâĂŹ areas. The negative correlation is counter intuitive so it is surprising that Sx is the best predictor for accumulation.

For Glacier 2, the most important regressor is elevation (Table 4.2). This coefficient is positive, which means that SWE will increase with elevation. The elevation regression coefficient is an order of magnitude larger than the other coefficients and has the highest semi-partial R2 and raw correlation (Table 4.2). Sx is the second most important regresor and has a positive correlation, which indicates that âĂŸshelteredâĂŹ areas are likely to have high accumulation.

The most important regressor for Glacier 13 is elevation (Table 4.2). The coefficient is positive, which means that cells at higher elevation had higher values of SWE. Despite a low value of raw correlation between elevation and SWE, the semi-partial R2 value is the largest between the glaciers. The high semi-partial R2 value indicates that when elevation is added to the regression the total variance

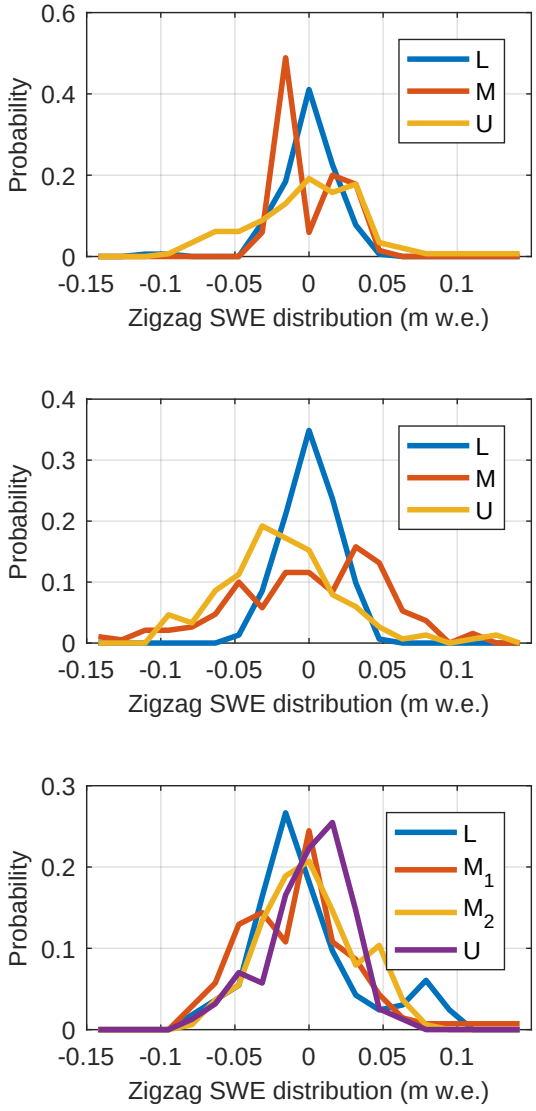

Fig. 4. Distribution of zigzag SWE values about the local mean on Glacier 4 (upper panel), Glacier 2 (middle panel) and Glacier 13 (lower panel). Zigzags are distributed throughout the ablation area of each glacier, with one located in the lower portion (L), one in the middle portion (M), and one in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

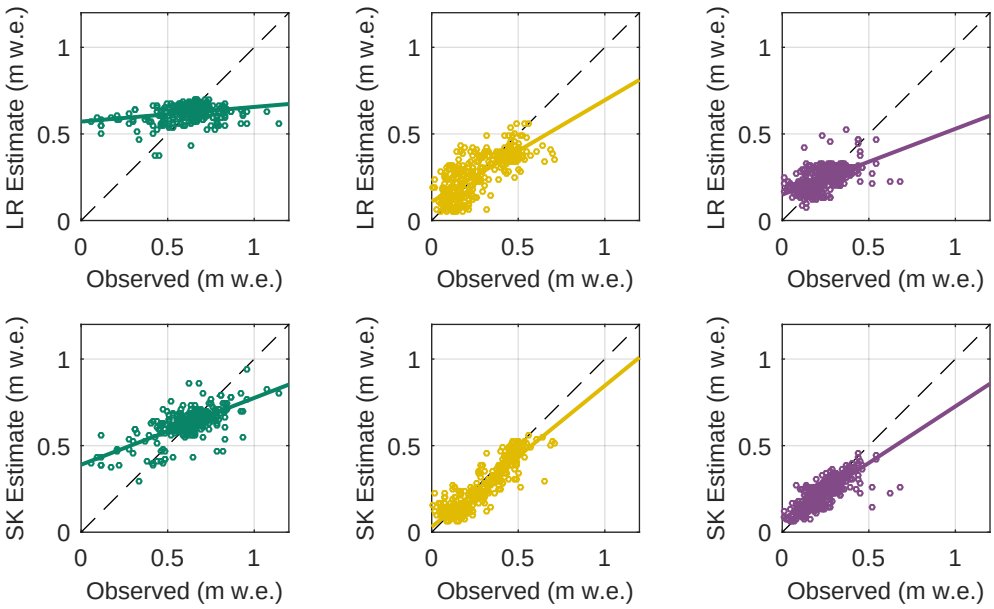

Fig. 5.

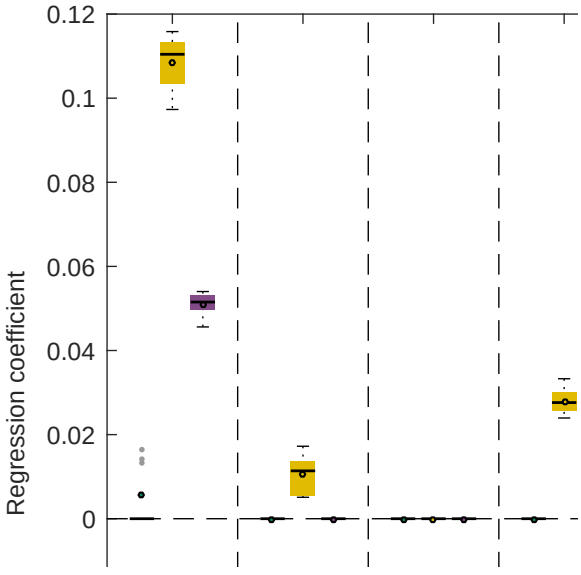

Fig. 6.

explained increases considerably because the remaining regressors are not important.

The map of estimated SWE for the entire glacier shows a relatively uniform SWE distribution over Glacier 4 (Figure 4.8), due to the large influence of the intercept on the regression. Areas with high Sx values (sheltered), especially in the accumulation area, have the lowest values of SWE. This regression indicates that the wind plays a role in snow distribution but since the valley in which the glacier sits is steep walled and curved, perhaps having a single cardinal direction for wind is inappropriate. Examining Sx values that assume wind moving up or down glacier and changing direction to follow the valley could allow the Sx parameter to explain more of the variance.

The map of modelled SWE on Glacier 2 closely matches that of elevation (Figure 3.39), which highlights the strong dependence of SWE on elevation. The range of predicted SWE is largest for Glacier 2 and it also has the highest SWE (1.92 m w.e) and the lowest SWE (0 m w.e.) values (Figure 4.2). Both extremes are perhaps unexpected on this glacier and are likely an artefact from extrapolating from the regression, which largely depends on elevation. The southwest region of the accumulation area with high estimated accumulation results from the combination of high elevation and Sx values. The low SWE values at the terminus are a result of low elevation values and Sx values that are close to zero. The map of estimated SWE on Glacier 13 (Figure 4.8) closely follows elevation although the range of SWE values is relatively small so the elevation effect is less pronounced than on Glacier 2.

spatial patterns

There are large differences in spatial patterns of interpolated SWE distributions for the three study glaciers found using simple kriging (Figure 4.26). The lower half of Glacier 4 has a relatively uniform SWE, while the accumulation area has both low and high values of SWE. The low density

of sampling points in the accumulation area of Glacier 4 result in large gradients in SWE. Glacier 2 has two distinct and relatively uniform areas âĂŤ the lower ablation area has low SWE (âĹij 0.1 m w.e.) and the upper ablation and accumulation areas have higher SWE values (âĹij 0.6 m w.e.). The boundary between two these zones closely follows the outline of the ice dune area observed during field data collection. Glacier 13 does not appear to have any strong patterns and accumulation is generally low (âĹij 0.1âĂŞ0.5 m w.e.).

The accumulation distribution estimated by simple kriging and by regression is qualitatively similar for Glacier 13 but differs con- siderably for Glacier 4 and 2. Glacier 4 has large SWE values in the accumulation area when simple kriging is used, which contrasts strongly with the mostly uniform accumula- tion estimated with regression. The converse is observed on Glacier 2, where the simple kriging estimate has a small SWE range and two distinct regions of SWE values while the regression estimate has a large range and a significant elevation gradient in accumulation. These spatial patterns highlight that kriging is sensitive to individual observation values in areas with sparse sampling.

The choice of interpolation method affects the mean winter balance (Figure 4.32). Kriging interpolation pro- duces the highest mean value of SWE on Glacier 4. The estimates of SWE in the accumulation area are greatest when kriging is used because there is a single single high SWE value in the accumulation area. Kriging is sensitive to outliers in areas with sparse sampling. However, the winter balance is similar between interpolation methods and mean of observed data for Glacier 4. This similarity arises from the low correlation coefficient for all methods, resulting in values closer to the data mean. Relative differences in mean SWE between the three interpolation methods are similar for Glacier 2 and 13, with topographic regression producing the highest mean SWE and kriging producting the lowest. Kriging estimates lower SWE in the accumulation area of both glaciers because elevation is not incorporated into the model.

For all glaciers, the topographic regression results in the lowest mean variance explained (Figure 4.33). The mean correlation coefficients for kriging and regression kriging are similar for all glaciers, with regression kriging being slightly lower than kriging. Variance explained on Glacier 4 is consistently the lowest, indicating that observed SWE values are highly variable. The converse is seen on Glacier 2, where correlation coefficients are consistently high regardless of the interpolation method.

The glacier-wide winter mass balance on each glacier (Bw), calculated as the mean SWE, that is found using simple kriging is considerably different to that found using a topographic regression (Figure 4.21). Accumulation on Glacier 4 is 0.09 m w.e. (15%) higher and has larger spatial gradients in SWE when simple kriging is used. Conversely, simple kriging estimates are 0.13 m w.e. (23%) and 0.08 m w.e. (22%) lower for Glaciers 2 and 13, respec- tively, and the range of SWE values is smaller.

Quantifying effects of variability

The largest difference in estimated SWE between the various density options is found on Glacier 2 in the upper part of the accumulation area (Figure 4.9). Glacier 2 also has the lowest difference in estimated SWE, which result from all density options estimating values of 0 m w.e. at the terminus. The strong relationship between elevation and SWE makes the estimation of glacier-wise SWE sensitive to this parameter.

The spatial patterns and specific winter surface mass balance (WSMB) are affected by variability introduced when interpolating density, estimating grid-cell SWE values, and when interpolating observations.

DISCUSSION

[77] conducted an airborne GRP survey of two adjacent glaciers in Switzerland. The lower part of the larger valley glacier showed a clear correlation between altitude and snow accumulation. The upper part of the glacier and the adjacent smaller glacier had no alti- tudinal trend and the fluctuations in depth were large. Additionally, the accumulation was 40% lower on the smaller glacier. The altitudinal trend is a well documented pattern and was thought to be a result of melt that occurred during warmer weather, which is more pronounced at lower elevations. Spatial variability of precipitation and redistribution of snow were believed to have resulted in the high spatial variability in higher parts of the study area. Since the majority of the precipitation events originated from one direction and the large glacier was on the lee side of a ridge, it experienced preferential deposition. Mean- while, the smaller glacier was further along the storm track so it received less precipitation. Overall, [77] showed that snow distribution on glaciers is not simply a function of altitude, which corroborated research done in other alpine catchments.

In most cases, the resolution of measurements over a large area is insufficient to approximate the true variability [15, 32].

extrapolation of regression models will likely result in large errors. These errors are especially relevant in the accumulation area, which has extreme values for all parameters. Errors in the accumulation area are especially important to acknowledge because this area has the highest values of SWE and is likely to heavily influence final winter mass balance values. Improvements to this study could include using an air-borne GPR to collect a dense network of SWE measurements in difficult to access areas [e.g. 82] (see Section 1.3.2 for more details).

and both under and over sampling is believed to have occurred using the FS (more details or just sources??).

Lopez 2013 for small scale var

Winter snow pack in southwestern Yukon was well below normal in 2015 (Yukon Snow Survey Bulletin and Water Supply Forecast, May 1, 2016). Temperatures were generally warmer than normal and the melt season began $1-2$ weeks early.

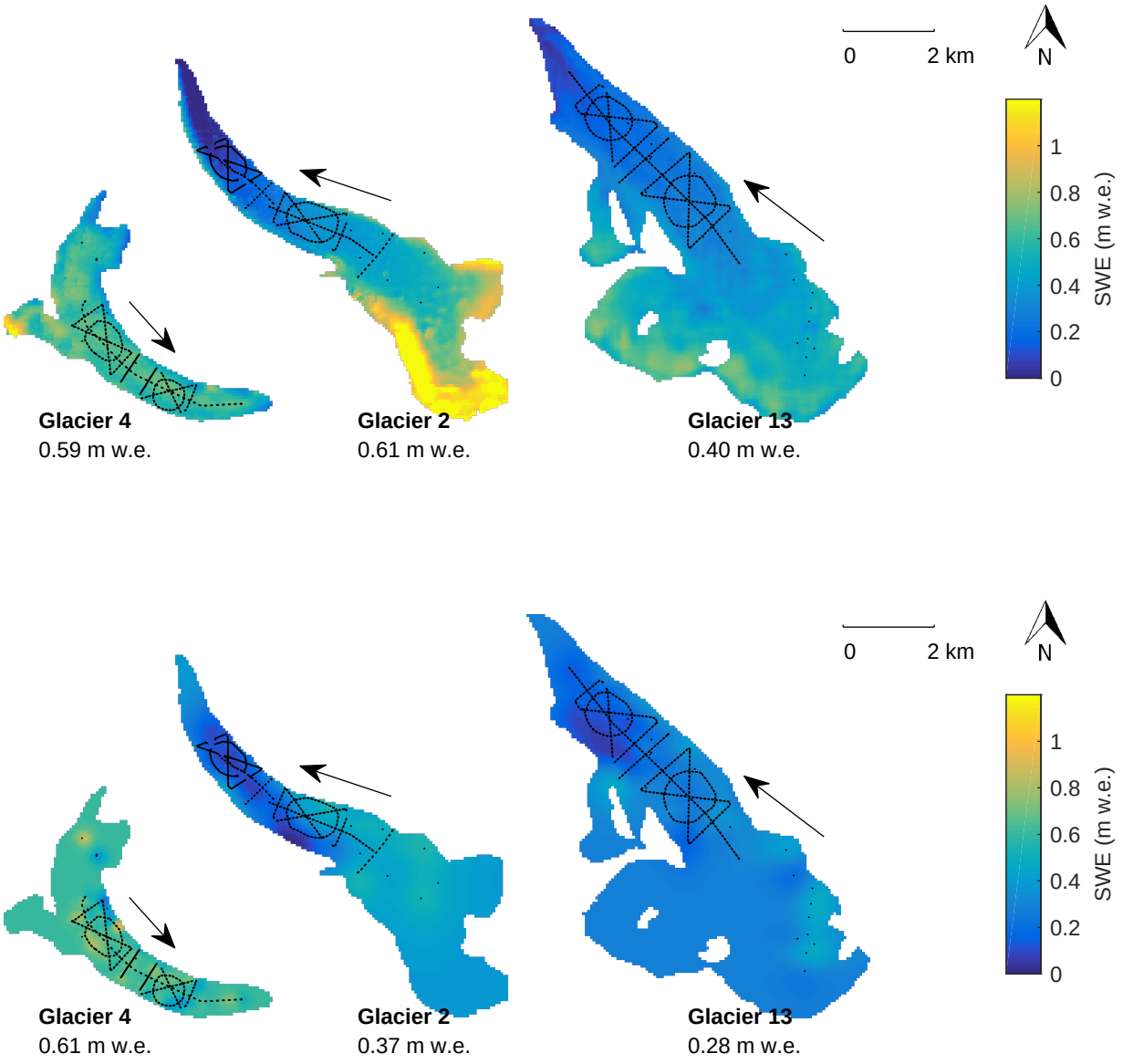

Fig. 7.

Field sampling (also called direct glaciological method) is known to be biased towards small alpine glaciers with simple topography.

INTRODUCTION

Objective: (1) Discuss choices made when moving from measurement to accumulation and (2) show how system variability and our choices interact to create uncertainty in our estimate of accumulation

- snow distribution in alpine regions is not uniform or static, but rather highly variable and influenced by diverse and dynamic processes operating on multiple spatial and temporal scales -¿ topographic effects (crevasses, surface topo, elevation aspect, precip grad across range), snow drift and preferential deposition - [22] note that studies of snow water equivalent (SWE) that have been conducted in alpine environments vary considerably in the extent and spacing of their measurements. - Snow accumulation is spatially variable on point scales (¡5 m), hillslope scales (1âĂŞ100 m), basin scales (100âĂŞ10,000 m) and regional scales (10âĂŞ1000 km) [22]. -Point-scale variability is generally associated with surface roughness effects and the presence of small obstacles. -¿ take three measures Many parts of a glacier though are characterized by a relatively smooth surface, with roughness lengths on the order of centimeters

[57]. In these areas, point-scale variability of snow depth is low. However, in heavily crevassed regions, point-scale variability can be large and thus exert a dominant control on snow distribution in the area [82]. -Hillslope-scale variability is caused by variations in the surface topography of the glacier. The curvature and slope of the surface as well as the presence of local ridges or depressions can affect where snow is located [15, 115]. Avalanching can also redistribute snow, especially on the margins of a glacier [17, 89]. Watershed-scale variability results mainly from the effects of changing elevation and aspect on atmopsheric conditions [22]. In particular, orographic lifting and shading can result in higher elevation and north-facing areas of the glacier having more snow than other areas [89, 115]. Gradients in temperature from elevation changes also affect the freezing level, which determines whether precipitation falls as snow or rain [17]. For example, [77] found a strong influence of elevation in determining accumulation on Findel Glacier in Switzerland. Regional variability occurs when areas within a mountain range have differing amounts of snow. Often, this results from horizontal precipitation gradients and rain shadows forming on the lee side of topographic divides. Areas with large, steep mountains are especially affected by these processes.

---

derived accumulation estimated winter surface mass balance distributed snow water equivalent

References

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

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)

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

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-)

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

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)

Hengl T, Heuvelink GB and Rossiter DG (2007) About regression-kriging: From equations to case studies. *Computers & Geosciences*, **33**(10), 1301–1315 (doi: 10.1016/j.cageo.2007.05.001)

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)

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*

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

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)

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)

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)

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)

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

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

Wasserman L (2000) Bayesian Model Selection and Model Averaging. *Journal of Mathematical Psychology*, **44**(1), 92–107, ISSN 0022-2496 (doi: 10.1006/jmps.1999.1278)

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)

Williams CKI (1998) *Prediction with Gaussian Processes: From Linear Regression to Linear Prediction and Beyond*, 599–621. Springer Netherlands, Dordrecht, ISBN 978-94-011-5014-9 (doi: 10.1007/978-94-011-5014-9$_2$3)

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

Zeugner S and Feldkircher M (2015) Bayesian Model Averaging Employing Fixed and Flexible Priors: The BMS Package for R. *Journal of Statistical Software*, **68**(4), 1–37