

Alexandra Pulwicki  
February 2, 2017

## Results Basin Scale

### Overview

This document describes the sampled and full ranges of topographic parameters and then discusses the multiple linear regression and Bayesian model averaging that were used to explain SWE with topographic parameters. Distributed SWE is found for the three study glaciers using the regressions.

### Contents

<b>1 Topographic parameters</b>	<b>2</b>
1.1 Obtaining digital elevation models (DEMs) for study glaciers . . . . .	2
1.2 Calculating topographic parameters . . . . .	5
1.3 DEM smoothing . . . . .	10
1.4 Parameter correlations . . . . .	11
1.5 Maps of topographic parameters and distribution of parameters sampled .	12
<b>2 Linear Regressions</b>	<b>22</b>
2.1 Background . . . . .	22
2.2 Importance of variables in regression models . . . . .	23
2.3 Snow density estimation methods . . . . .	24
2.4 Multiple Linear Regression (MLR) . . . . .	25
2.4.1 Background . . . . .	25
2.4.2 Methods . . . . .	26
2.4.3 Results and Discussion . . . . .	26
2.5 Bayesian Model Averaging (BMA) . . . . .	35
2.5.1 Background . . . . .	35
2.5.2 Methods . . . . .	38
2.5.3 Results and Discussion . . . . .	38
2.6 MLR and BMA comparison . . . . .	45
<b>3 Summary</b>	<b>50</b>
<b>4 Appendix</b>	<b>51</b>
4.1 Topographic parameters from QGIS to Matlab . . . . .	51
4.2 MLR software . . . . .	52
4.3 BMA software . . . . .	53

# 1 Topographic parameters

## 1.1 Obtaining digital elevation models (DEMs) for study glaciers

Topographic parameters can be derived from a digital elevation model (DEM) of the study area. The DEM used in this project was created from imagery collected by the SPOT-5 satellite and it was provided at no cost by the French Space Agency (CNES) through the SPIRIT International Polar Year project (Korona and others, 2009). The DEM has a cell size of  $40 \times 40$  m. The DEM was created using a set of correlation parameters fit for steeper slopes (E. Berthier personal communication, 2016).

Two DEMs were available for the Donjek Range. The first DEM (GES 08-029) was made from images collected on September 3, 2007 and the second DEM (GES 07-044) was made from images collected on September 13, 2007. The snow extent on September 13, 2007, as imaged by a Landsat 7 satellite can be seen in Figure 2. Since the images were collected in September, the surface would likely be at a seasonal minimum with minimum snow cover. Therefore, the surface described by the DEM in the ablation area represents the topography below the snow. A limitation in using this DEM is that the DEM is from 2007 and there have almost certainly been changes in the end-of-summer glacier surface. However, the SPOT-5 DEM is the best resolution and most current DEM available for the study area.

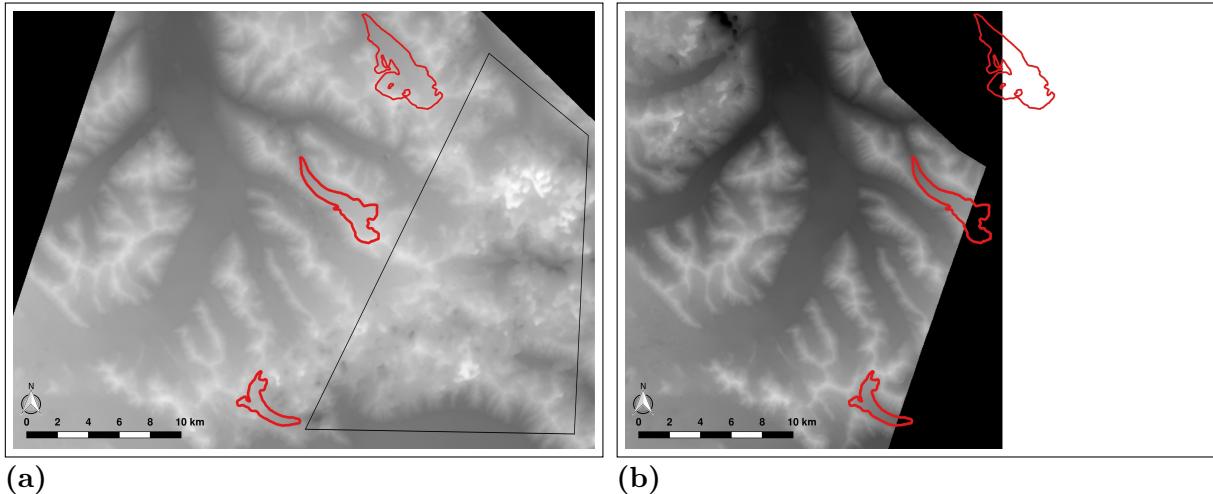


Figure 1: SPOT-5 DEMs available for the Donjek Range. Study glaciers are shown as red outlines. The DEM made from imagery collected on September 3, 2007 (GES 08-029) is shown in (a) and the DEM made from imagery collected on September 13, 2007 (GES 07-044) is shown in (b). Imagery that contains cloud cover result in a distorted DEM, as seen in the boxed area of (a).

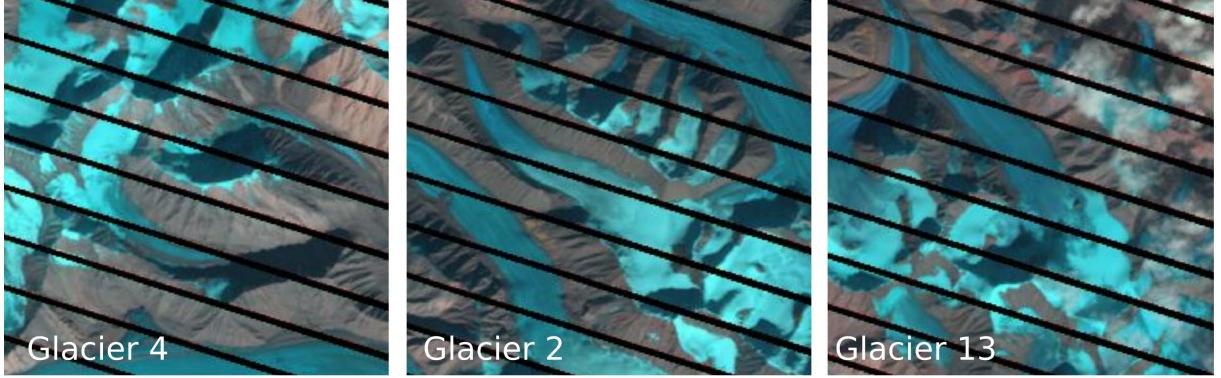


Figure 2: Landsat 7 ETM images of study glaciers on September 13, 2007. Snow cover is shown as light blue and ice is shown as dark blue.

The GES 08-029 DEM covers all three study glaciers (see Figure 1a) but a large part of Glacier 4 and some areas of Glacier 2 were masked by clouds and/or showed limited contrast in the original stereoimage pairs, resulting in incorrect elevation data (E. Berthier personal communication, 2016). The cloudy areas appear as black regions on the DEM mask (not shown) and as distortion in the DEM, as seen in the boxed area of Figure 1a. The second DEM (GES 07-044) spans only part of the Donjek Range, covering most of Glacier 4 and  $\sim 60\%$  of Glacier 2 (see Figure 1b). This DEM had no masked areas over Glaciers 2 and 4. The two DEMs were therefore merged to create a cloud free DEM that spanned all three glaciers.

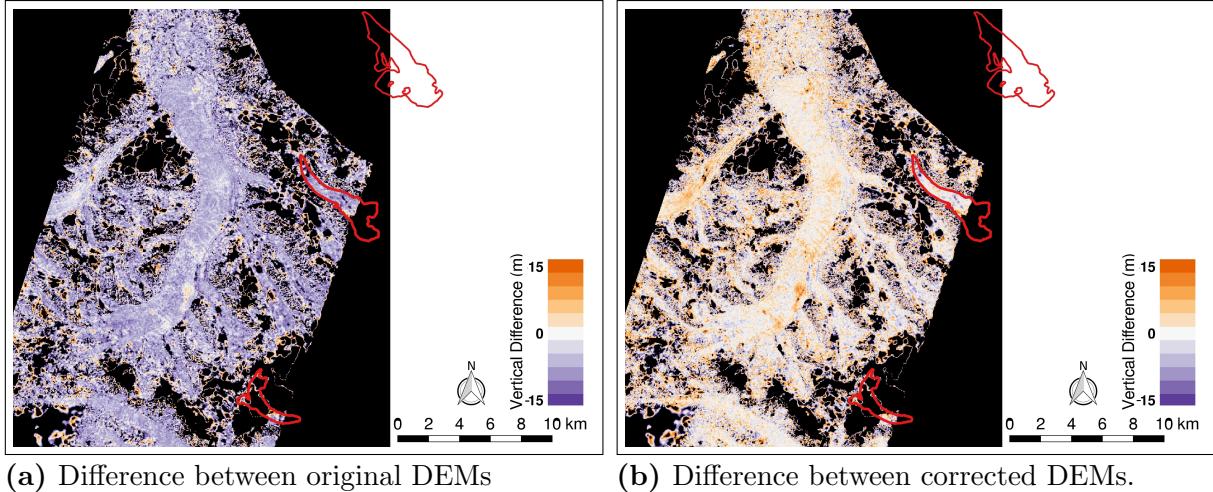


Figure 3: Vertical difference between DEMs in overlapping area. Difference was found by subtracting GES 08-029 from GES 07-044. Positive values indicate that GES 07-044 values are higher than GES 08-029 values.

The merging process was complicated by the fact that there was a horizontal and vertical discrepancy between the two DEMs. Although the discrepancy was not consistent

throughout the study area, the GES 07-044 DEM was generally higher than the first, as can be seen by the overall purple colour in Figure 3a and the positive skew of the vertical difference between the two DEMs in Figure 4. The mean vertical difference is +6.3 m.

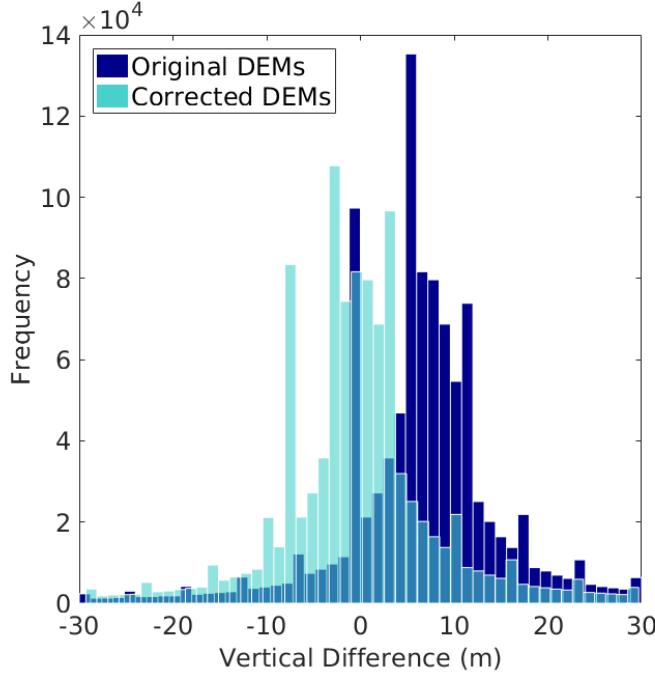


Figure 4: Histogram of the vertical difference between GES 08-029 and GES 07-044 before (dark blue) and after (light blue) correction.

The discrepancy was corrected by E. Bertier (2016, personal communication) using an iterative 3D-coregistration algorithm (Berthier and others, 2007). The GES 07-044 DEM was arbitrarily chosen to be used as the reference DEM. Note that the absolute value of the elevation is not necessarily important for the topographic regression, as long as the relative elevations are correct. The reference DEM, GES 07-044, was first shifted vertically by +5.4 m, estimated using ICESat data (Berthier and others, 2010). Then, the mean horizontal and vertical (X, Y, Z) shift between the reference DEM and the GES 08-029 DEM was found by minimizing the standard deviation of the elevation differences between the DEMs. Using this correlation, the GES 08-029 DEM was shifted  $\sim 2$  m east,  $\sim 4$  m north, and  $\sim 1.9$  m vertically. The GES 08-029 DEM was then reprojected in the same projection as the reference DEM (GES 07-044). The difference map between the two shifted DEMs is shown in Figure 3b. Difference values are not uniform but do show both positive and negative values. The distribution of vertical difference values (Figure 4) after correction is centred at zero with a mean difference of -0.2 m.

Merging of the corrected DEMs was completed in QGIS. First, the rasters were cropped to overlap by a few cells. The crop line was chosen by hand to include as much of the reference DEM as possible (fewer areas of poor data) but was a relatively small distance from the edge of the DEM (see Figure 5). The second DEM was cropped to follow the

same merge line but overlap with the first DEM by a few cells. This was done to avoid gaps in cell values that arise from cropping across a DEM cell. The merging was completed using the built in QGIS tool ‘Merge’ and in areas where the two DEMs overlapped, the GES 08-029 DEM values were chosen.

The final DEM used for subsequent analysis can be seen in Figure 4. Despite the corrections, there were still discrepancies along the intersection of the DEMs, which can be seen as sharp boundaries in the contour lines. However, these discrepancies are not present on the study glaciers and are located more than 250 m from the edge of the glacier so this DEM was used as the final version for the Donjek Range.

## 1.2 Calculating topographic parameters

Topographic parameters are used to describe characteristics of the local topography that may affect snow distribution and can act as proxies for physical processes that determine snow deposition and redistribution. Topographic parameters used in snow accumulation studies on glaciers include elevation ( $z$ ), distance from centreline ( $d_C$ ), slope ( $m$ ), tangential ( $\kappa_T$ ) and profile ( $\kappa_P$ ) curvature, “northness” ( $N$ ), aspect ( $\alpha$ ) and Sx, which is a proxy for wind redistribution (Basist and others, 1994; Revuelto and others, 2014; McGrath and others, 2015).

A number of programs are used to calculate topographic parameters from the DEM. Distance from centreline and “northness” were calculated in Matlab. Sx was determined using a executable obtained from Adam Winstral that follows the procedure outlined in Winstral and others (2002). The remaining parameters 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). Note that topographic parameters were calculated using the full DEM and then trimmed to Randolph Glacier Inventory defined glacier outlines so as to avoid errors that arise at the edge of the DEM when taking derivatives.

Details about the calculation of topographic parameters are described below.

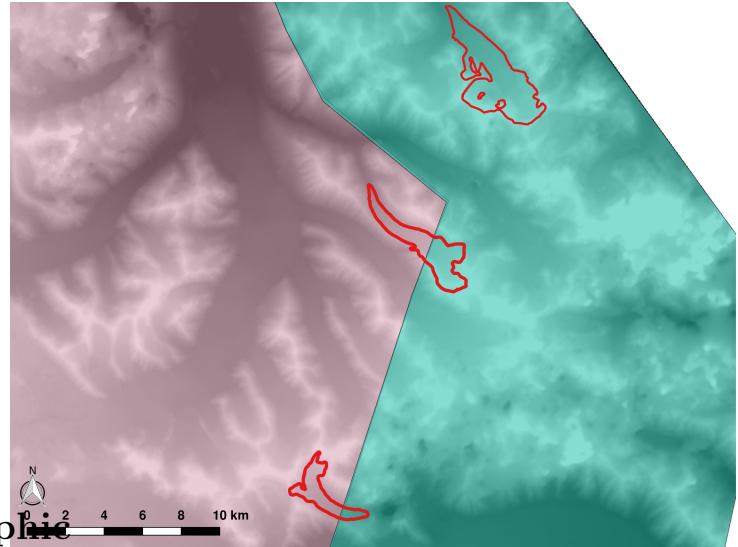


Figure 5: Outlines of the cropped GES 07-044 DEM (pink, left) and cropped GES 08-029 DEM (blue, right) used for merging. There is a slight overlap between the two DEMs that cannot be seen at this scale.

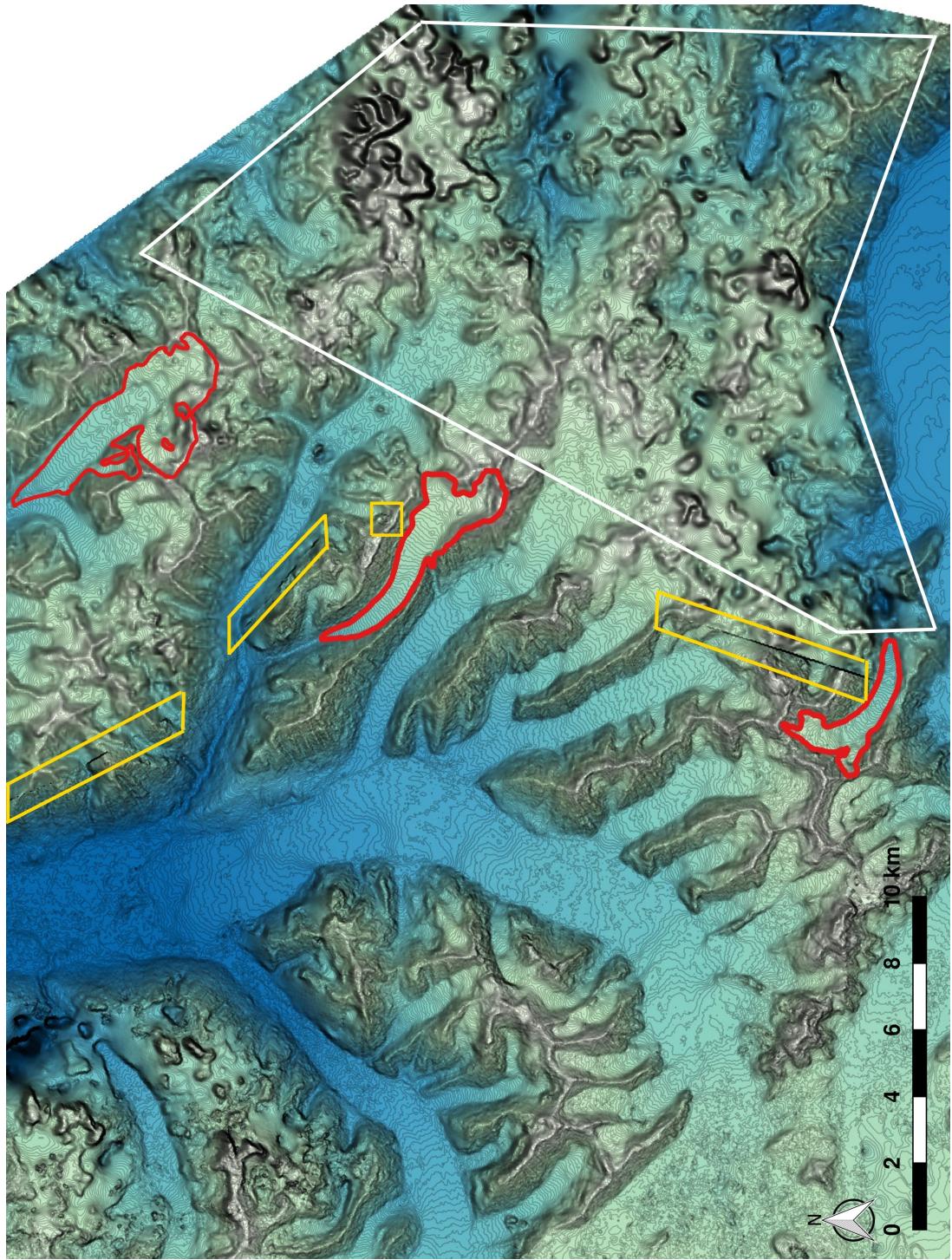


Figure 6: Merged DEM of the Donjek range from two corrected SPOT-5 DEMs, plotted with 10 m contour lines. Study glaciers are shown in red. Discrepancies between the DEMs along the merge line can be seen as anomalous linear features in the contour map (yellow boxes). Distorted contours in the eastern regions (white box) are a result of errors in the DEM.

## Elevation

Elevation ( $z$ ) values were taken from the (corrected) SPOT-5 DEMs directly (Figure 9).

## Distance from centreline

Distance from centreline ( $d_C$ ) was calculated as the minimum distance between the Easting and Northing of the northwest corner of each cell and a manually defined centreline (Figure 11).

## Slope

Slope ( $m$ ) is the maximal rate of change of elevation and is defined as the angle between a plane tangential to the surface (gradient) and the horizontal (Olaya, 2009) (Figure 15). Slope ( $m$ ) is calculated according to

$$m = \arctan \sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}, \quad (1)$$

where the partial derivatives can be approximated by (Mitášová and Hofierka, 1993; Neteler and Mitášová, 2008; Hofierka and others, 2009)

$$\begin{aligned} \frac{\partial z}{\partial x} &\approx \frac{(z_7 - z_9) + (2z_4 - 2z_6) + (z_1 - z_3)}{8\Delta x}, \\ \frac{\partial z}{\partial y} &\approx \frac{(z_7 - z_1) + (2z_8 - 2z_2) + (z_9 - z_3)}{8\Delta y}. \end{aligned} \quad (2)$$

Here,  $z_k$  refers to one of the grid cells surrounding the cell of interest, which is located at row  $i$  and column  $j$  of the DEM. So  $z_3 = z_{i+1,j+1}$ ,  $z_7 = z_{i-1,j-1}$ , and so on (Figure 7). The grid spacing (resolution) of the DEM is  $\Delta x$  and  $\Delta y$  in the east-west and north-south direction, respectively (Neteler and Mitášová, 2008).

## Curvature

Curvature describes the convexity or concavity of a surface. The curvature of a surface is different in different directions, so there are various types of curvature. Profile and tangential curvature are the most common types to consider in geophysical systems. For this study, the mean curvature ( $\kappa$ ), found by taking the average of profile and tangential curvature, is used (Figure ??). The mean curvature emphasizes mean-concave (positive values) areas with relative accumulation and mean-convex (negative values) terrain with relative scouring (Olaya, 2009).

Profile curvature is the curvature in the direction of the surface gradient and it describes the change in slope angle. The equation for profile curvature ( $\kappa_p$  [ $\text{m}^{-1}$ ]) is (Neteler and Mitášová, 2008)

$$\kappa_p = \frac{\frac{\partial^2 z}{\partial x^2} \left(\frac{\partial z}{\partial x}\right)^2 + 2 \frac{\partial^2 z}{\partial x \partial y} \frac{\partial z}{\partial x} \frac{\partial z}{\partial y} + \frac{\partial^2 z}{\partial y^2} \left(\frac{\partial z}{\partial y}\right)^2}{\left[\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2\right] \sqrt{\left[\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 + 1\right]^3}} \quad (3)$$

$z_1$	$z_2$	$z_3$
$z_4$	$z_5$	$z_6$
$z_7$	$z_8$	$z_9$

Figure 7: Labelling of DEM grid cells surrounding the cell of interest. The eight surrounding cells are used for estimating topographic parameters in QGIS. The cell of interest, which is located at row  $i$  and column  $j$  of the DEM, is shown as a shaded cell and is labelled  $z_5$ .

where first-order partial derivatives are found using Equation 2. Second-order partial derivatives can be approximated by (Hofierka and others, 2009; Neteler and Mitášová, 2008)

$$\begin{aligned} \frac{\partial^2 z}{\partial x^2} &\approx \frac{z_1 - 2z_2 + z_3 + 4z_4 - 8z_5 + 4z_6 + z_7 - 2z_8 + z_9}{6(\Delta x)^2}, \\ \frac{\partial^2 z}{\partial y^2} &\approx \frac{z_1 - 4z_2 + z_3 - 2z_4 - 8z_5 - 2z_6 + z_7 + 4z_8 + z_9}{6(\Delta y)^2}, \\ \frac{\partial^2 z}{\partial x \partial y} &\approx \frac{(z_7 - z_9) - (z_1 - z_3)}{4\Delta x \Delta y}. \end{aligned} \quad (4)$$

Tangential Curvature represents the curvature in the direction of the contour tangent. The equation for profile curvature ( $\kappa_t$  [ $\text{m}^{-1}$ ]) is (Neteler and Mitášová, 2008)

$$\kappa_t = \frac{\frac{\partial^2 z}{\partial x^2} \left( \frac{\partial z}{\partial y} \right)^2 - 2 \frac{\partial^2 z}{\partial x \partial y} \frac{\partial z}{\partial x} \frac{\partial z}{\partial y} + \frac{\partial^2 z}{\partial y^2} \left( \frac{\partial z}{\partial x} \right)^2}{\left[ \left( \frac{\partial z}{\partial x} \right)^2 + \left( \frac{\partial z}{\partial y} \right)^2 \right] \sqrt{\left( \frac{\partial z}{\partial x} \right)^2 + \left( \frac{\partial z}{\partial y} \right)^2 + 1}}, \quad (5)$$

where first- and second-order partial derivatives are approximated using Equation 2 and 4.

### “Northness”

“Northness” ( $N$ ) is a solar radiation parameter that has been shown to increasingly affect accumulation distribution during the spring (Revuelto and others, 2014). It is also likely that this parameter may be related to sun induced snow metamorphosis and/or sun crusts, both of which affect SWE (McGrath and others, 2015). “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.

## Aspect

Aspect ( $\alpha$ ) represents the orientation of the steepest slope, with  $0^\circ$  defined as North and no value given to cells that have zero slope. The equation for aspect in degrees is (Neteler and Mitášová, 2008)

$$\alpha = \arctan\left(\frac{\partial z}{\partial y} / \frac{\partial z}{\partial x}\right), \quad (6)$$

where the partial derivatives are approximated by Equations 2. Here,  $\alpha = 0$  is in the west direction but the computed values were transformed to reflect  $0^\circ$  as north (clockwise).

Aspect is a circular parameter ( $0^\circ$  is the same as  $360^\circ$ ) but regressions (Section 2) require that a parameter is linear. Therefore, only the sine of aspect was used in topographic analysis. The sine of aspect is representative of the relative amount of direct solar radiation incident on a slope, which can affect SWE by metamorphosis of snow.

## Sx

Sx represents wind exposure/shelter and 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). This cell is referred to as the maximum upwind slope. Negative Sx values represent exposure relative to the shelter-defining pixel, which means that the cell of interest is higher than the cell with greatest upward slope. Conversely, positive values represent sheltered cells. To determine Sx values, we use the equation

$$Sx_{A,d\max}(x_i, y_i) = \max \left[ \tan^{-1} \left( \frac{z(x_v, y_v) - z(x_i, y_i)}{[(x_v - x_i)^2 + (y_v - y_i)^2]^{1/2}} \right) \right], \quad (7)$$

where  $A$  is the azimuth of the search direction,  $(x_i, y_i)$  are the coordinates of the cell of interest, and  $(x_v, y_v)$  are the set of all cell coordinates located along the search vector defined by  $(x_i, y_i)$ , the azimuth ( $A$ ), and maximum search distance ( $d\max$ ). Code for this calculation was provided by Adam Winstral (2016, personal communication). As done by McGrath and others (2015), we compute Sx at  $5^\circ$  azimuth increments for  $d\max$  distances of 100, 200 and 300 m. These values are then correlated (Pearson correlation) with observed values of SWE and the Sx values from the combination of azimuth and  $d\max$  input values that have the highest correlation are used for subsequent analysis (Table 1). The code for calculating Sx requires a UTM raster formatted to ASCII in ArcGIS.

Table 1: Values of azimuth ( $A$ ) and maximum search distance ( $d\max$ ), that correspond to the Sx that had the highest absolute correlation to observed SWE.

	$A$ ( $^\circ$ from North)	$d\max$ (m)	Correlation Coefficient
Glacier 4	85	300	-0.26
Glacier 2	330	300	0.56
Glacier 13	280	200	0.28

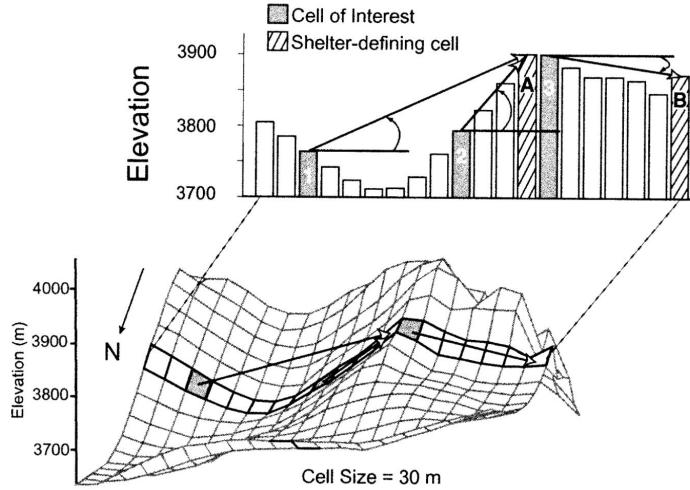


Figure 8: Image and description from (Winstral and others, 2002). Example of Sx calculations for three cells of interest along a 270° search vector. As depicted, with dmax set equal to 300 m, the shelter-defining pixel for cells 1 and 2 is cell A, producing positive Sx values. The shelter-defining cell for cell 3 is cell B, producing a negative Sx. Had dmax been equal to 100 m, the search for the shelter- defining pixel for cell 1 would not extend across the valley, thus producing a negative Sx for cell 1, while Sx for cell 2 would remain the same and that for cell 3 would be slightly lower.

### 1.3 DEM smoothing

Visual inspection of the curvature fields calculated using the DEM showed noisy spatial distribution that did not vary smoothly. Olaya (2009) states that the curvature calculation is sensitive to noisy data and a smoothing filter often needs to be applied to the DEM prior to calculation. Curvature, as well as slope, aspect and “northness”, are all sensitive to noise because calculating these parameters involves calculating the first and second derivatives of the elevation, which are highly dependent on the size of the DEM cell. To minimize the effect of noise on these four parameters, a smoothing filter was applied to the DEM and this smoothed DEM was used to calculate curvature, slope, aspect and “northness”. The non-smoothed DEM was used to determine elevation and Sx because these parameters do not depend on a topographic length scale and their values are not sensitive to the size of the DEM cell size.

To choose a smoothing algorithm and window size, a number of smoothing algorithms was applied and the one that resulted in topographic parameters that had the highest correlation with SWE was chosen. Window sizes of  $3 \times 3$ ,  $5 \times 5$ ,  $7 \times 7$  and  $9 \times 9$  were used. For all sizes, inverse-distance weighted smoothing and Gaussian smoothing were poorly correlated with SWE. An average value smoothing with a  $7 \times 7$  window resulted in the highest correlation between curvature (second derivative) and SWE as well as slope (first derivative) and SWE. The window size that produced the highest correlation of SWE values and curvature for each glacier differed, but for all SWE values taken together, the  $7 \times 7$  window resulted in the highest correlation. For slope, the highest correlation for individ-

ual glaciers was from the DEM with a  $7 \times 7$  smoothing window but the overall correlation was not the highest with this window. To maintain consistency between parameters, the  $7 \times 7$  smoothing window was chosen and applied to the DEM for calculation of curvature, slope, aspect and “northness”.

## 1.4 Parameter correlations

The correlation between topographic parameters at sampling locations on each glacier is shown in Table 2. Correlation values are generally low, with the exception of the correlation between northness and aspect on Glacier 2 and northness and Sx on Glacier 13, which were both larger than 0.7. Since there is little correlation between parameters and the correlations vary between glaciers, the use of a linear regression with these topographic parameters as predictor variables is warranted.

Table 2: Pearson correlation coefficients between topographic parameters at sampled locations. Topographic parameters are distance from centreline ( $d_C$ ), elevation ( $z$ ), aspect ( $\alpha$ ), slope ( $m$ ), northness ( $N$ ), curvature ( $\kappa$ ), and Sx.

	$z$	$d_C$	$\alpha$	$m$	$N$	$\kappa$	Sx	
Glacier 4	$z$	1	0.16	-0.57	-0.08	-0.51	0.17	0.43
	$d_C$		1	0.13	0.56	0.16	-0.43	0.35
	$\alpha$			1	0.57	0.95	-0.61	-0.58
	$m$				1	0.64	-0.58	-0.10
	$N$					1	-0.59	-0.59
	$\kappa$						1	0.05
	Sx							1
Glacier 2	$z$	1	0.06	-0.52	-0.58	-0.62	0.45	0.57
	$d_C$		1	0.06	0.13	0.11	-0.27	0.01
	$\alpha$			1	0.33	0.86	-0.42	-0.45
	$m$				1	0.74	-0.67	-0.41
	$N$					1	-0.67	-0.48
	$\kappa$						1	0.28
	Sx							1
Glacier 13	$z$	1	0.15	0.19	-0.15	0.10	0.02	0.27
	$d_C$		1	-0.05	0.18	0.10	-0.45	0.06
	$\alpha$			1	-0.07	0.68	<0.01	0.45
	$m$				1	0.63	-0.22	-0.22
	$N$					1	-0.21	0.23
	$\kappa$						1	-0.30
	Sx							1

## 1.5 Maps of topographic parameters and distribution of parameters sampled

Elevation maps (Figure 9) show that both Glacier 2 and 13 have small areas with high elevation, which correspond to steep headwalls. Mean elevation is the same for all glaciers for the full and sampled distribution within one standard deviation (Table 3). However, maximum elevation values are lower for Glacier 4 than the other two glaciers (Figure 10) and the sampled elevation means are approximately 200 m less than that of the full distribution. Standard deviations are smaller for sampled ranges for all glaciers. The skewness of sampled and full distributions is different for all glaciers. Elevation full distributions are similar for the study glaciers, with kurtosis for all distributions, except sampled elevation on Glacier 13, being less than 3 (value for normal distribution). Kurtosis of sampled distributions show that Glacier 13 had a broader distribution and Glacier 2 has a narrower distribution.

The distribution of sampled distance from centreline (Figure 12) is different from that of the full distribution. Generally, large distances were not sampled. Larger values of skewness and kurtosis in the full distribution indicate that these distributions are broader and span a larger range of values (Table 3). This is also seen in the mean and standard deviation values, which are also larger for the full distribution. Large values of distance from centreline are located at the edges of the glacier in the accumulation area (Figure 11), which constitute steep, inaccessible terrain. Within the ablation area, the hourglass sampling pattern allowed for locations across the whole width of the glacier to be measured. Note that Glacier 13 has two centrelines in the accumulation area because of the confluence of two major arms of the glacier.

The aspect of Glaciers 2 and 13 indicate that the majority of these two glaciers is north facing, while the majority of Glacier 4 is south facing (Figure 13). This is also highlighted in the mean values of aspect, which are positive for Glaciers 2 and 13 and negative for Glacier 4 (Table 3). Sampled mean aspect is similar to the full distribution, although the standard deviation of sampled aspects is much lower. Further, the skewness and kurtosis of the sampled distributions differs considerably from the full distribution and there are many aspects that were not sampled (Figure 14).

Slope of the three study glaciers (Figure 15) is generally less than 20°, with only the margins of the accumulation area and a few steps on Glacier 13 having steep slopes. The full and sampled distributions of slope are similar between glaciers (Table 3), with mean values of ~13° for the full distribution and 5° to 8° for the sampled distribution. The sampled distributions are all different than the full range, as indicated by the lower means, standard deviations, and skewness, as well as larger kurtosis. This shows that the sampled distributions are generally narrower than full distributions and severely under sample steep slopes (Figure 16).

The mean “northness” values for all glaciers were close to zero and the majority of cells have values close to zero, which is likely due to their low slope values (Figure 17). Even for Glacier 4, which is largely south facing and should thus have lower values of “northness”, has a distribution with a mean close to zero (Table 3). The low slope values mean that the values of “northness” were determined largely by the aspect, which can

be seen by the resemblance between the “northness” map and the aspect map and their high correlation (Table 2). Sampled distributions of “northness” did not resemble the full distributions. Although the mean values were similar, the skewness and kurtosis values were higher for all glaciers, indicating that the sampling was biased.

Curvature values on all glaciers are largely negative, indicating that concave topography is more prevalent than convex topography (Figures 19 and 20). The sampled distribution of curvature is a poor representation of the full distribution as shown by the dramatically different values of skewness, which are positive for the full distribution and negative for the sampled distribution (Table 3).

$S_x$  maps over the study glaciers are shown in Figure 21 and the wind direction and maximum search distance with the highest correlation to SWE for each glacier are shown in Table 1. For Glacier 4, an approximately east wind and 300 m search distance were most strongly correlated. The correlation was negative, which means that negative values of  $S_x$  (exposure) correspond to areas with larger SWE (more snow). This is counter intuitive and perhaps indicates that  $S_x$  is not an appropriate topographic parameter to correlate with SWE on Glacier 4. Despite this,  $S_x$  was retained in future analysis for consistency between glaciers. For Glacier 2, a north wind with a 300 m search distance was most strongly correlated and for Glacier 13, a west wind with a 200 m search distance produced the strongest correlation. Both of these correlations were positive, so a more positive  $S_x$  value (sheltered) corresponds to higher values of SWE (more snow). The correlation values for Glacier 4 and 2 are low ( $<0.3$ ) which indicates that  $S_x$  will likely be insignificant in estimating snow accumulation. The correlation between  $S_x$  and SWE on Glacier 2 is higher at 0.56.

The full distribution of  $S_x$  (Figure 22) is different for each glacier and differs greatly between sampled and full distributions. Glacier 2 has a mean less than zero, indicating that a large portion of the glacier has exposed topography (Table 3). Glacier 4 and 13 have positive mean values of  $S_x$ , indicating more sheltered topography. Extreme values of  $S_x$  are generally located along the edges of the accumulation areas (Figure 21). The sampled distribution mean for Glacier 2 is close to that of its full distribution, while the sampled distribution means of Glacier 4 and 13 are different (negative) than that of the full distribution means (positive). The sampled distribution are more sharply peaked, as indicated by the smaller standard deviation values and larger kurtosis when compared to the full distributions. The skewness also differs for Glaciers 4 and 2, with the sampled distributions skewed more to the right than full distributions. Overall, the sampled distribution of  $S_x$  is a poor representation of the full distribution.

Overall, the sampled topographic parameters are poor representatives of the full distribution of parameters. Extreme values of all parameters are grossly under sampled and the distribution of the sampled parameters generally differs from the full distribution. This was largely do to dangerous travel conditions and an inability to quickly and accurately measure snow depth in the accumulation area. As a result, 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. McGrath and others, 2015) (see Section ?? for more details). Generally, the sampled values do not fully capture the variance in topographic parameters but a regression is still valuable since topographic regressions are common when estimating winter mass balance. Acknowledging that this methodology is limited and prone to extrapolation errors is paramount to capturing uncertainty in accumulation studies.

Table 3: Descriptive statics of topographic parameter full and sampled distribution. Mean and standard deviation are in units of meters for distance from centreline ( $d_C$ ) and elevation ( $z$ ), in units of  $\text{m}^{-1}$  for profile ( $\kappa_P$ ) and tangential ( $\kappa_T$ ) curvature, and are unitless for cosine of aspect ( $\alpha$ ), “northness” ( $N$ ), slope ( $m$ ), and Sx. Skewness is a measure of the data asymmetry about the mean, with positive values indicating data that are more spread to the right of the mean and zero indicating a perfectly symmetric distribution. Kurtosis is a measure of how prone a distribution is to outliers. A normal distribution has a kurtosis value of 3 and larger values indicate distributions that are more prone to outliers.

		Full			Sampled				
		Mean	Standard Deviation	Skewness	Kurtosis	Mean	Standard Deviation	Skewness	Kurtosis
<b>Glacier 4</b>	$z$	2343.76	178.36	-0.17	2.13	2242.63	89.52	0.00	2.59
	$d_C$	258.95	233.49	1.82	6.52	124.33	89.56	0.32	2.04
	$\alpha$	-0.35	0.61	0.75	2.23	-0.37	0.53	0.85	2.57
	$m$	12.81	7.03	1.11	3.37	8.32	3.17	2.19	8.10
	$N$	-0.05	0.17	0.52	4.04	-0.04	0.10	1.35	5.04
	$\kappa$	-53.32	70.30	0.76	7.81	-33.34	38.33	-0.80	2.95
<b>Glacier 2</b>	$Sx$	1.56	11.46	1.03	4.28	-1.06	5.17	1.62	6.21
	$z$	2494.71	233.03	0.09	2.82	2306.90	93.74	-0.02	1.93
	$d_C$	304.95	236.97	1.22	4.51	140.65	98.84	0.37	2.74
	$\alpha$	0.59	0.43	-1.43	4.83	0.52	0.33	-0.42	2.14
	$m$	13.02	9.48	1.08	3.00	6.54	2.24	1.97	9.12
	$N$	0.14	0.16	0.74	3.68	0.06	0.05	1.01	4.10
<b>Glacier 13</b>	$\kappa$	-18.40	91.35	1.30	6.83	-22.94	43.91	-1.44	4.75
	$Sx$	-3.82	9.34	-0.18	6.09	-3.63	2.59	0.81	4.81
	$z$	2427.62	225.15	0.13	2.45	2219.45	82.25	0.92	5.67
	$d_C$	443.88	308.48	0.76	3.21	181.91	152.29	0.46	2.33
	$\alpha$	0.55	0.49	-1.30	3.89	0.69	0.28	-1.31	4.91
	$m$	13.36	10.13	1.15	3.51	5.10	1.98	2.20	12.99
<b>Glacier 13</b>	$N$	0.13	0.18	1.01	4.02	0.06	0.04	1.58	9.63
	$\kappa$	-13.31	89.16	1.24	6.65	-6.93	32.44	-0.77	4.51
	$Sx$	3.69	12.08	0.97	3.68	-1.62	3.94	0.90	4.95

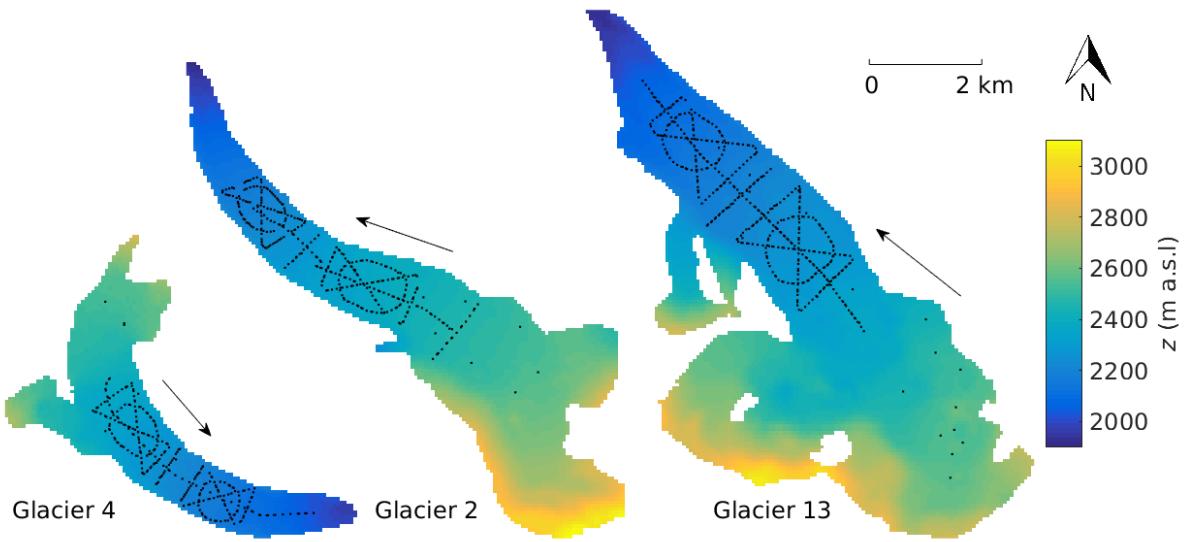


Figure 9: Distributions of elevation ( $z$ ) used in the topographic regressions for the study glaciers. This DEM is derived from a SPOT5 satellite image and has a grid size of  $40 \times 40$  m. Subsequent topographic parameters were derived from this DEM. Arrows indicate glacier flow direction and black dots show snow depth sampling locations.

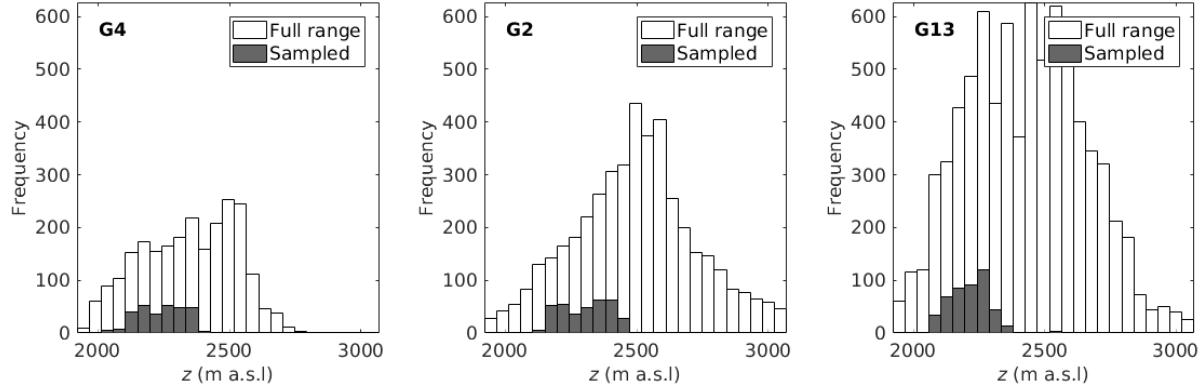


Figure 10: Histograms of elevation ( $z$ ) sampled (black) as compared to total range of elevation (white) of study glaciers.

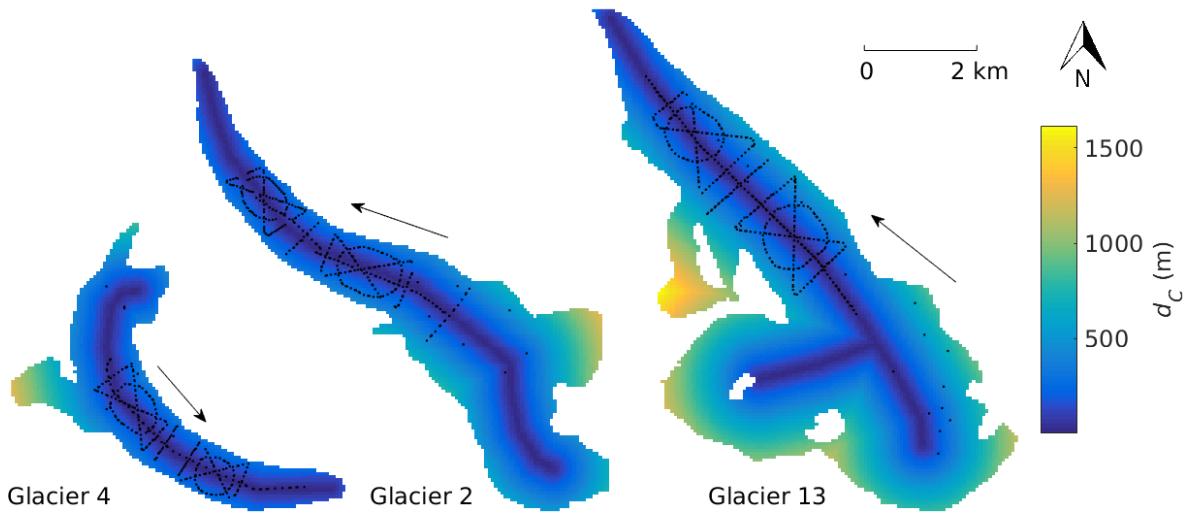


Figure 11: Distributions of distance from centreline ( $d_C$ ) used in the topographic regressions for the study glaciers. Centreline was drawn by hand in QGIS. Arrows indicate glacier flow direction and black dots show snow depth sampling locations.

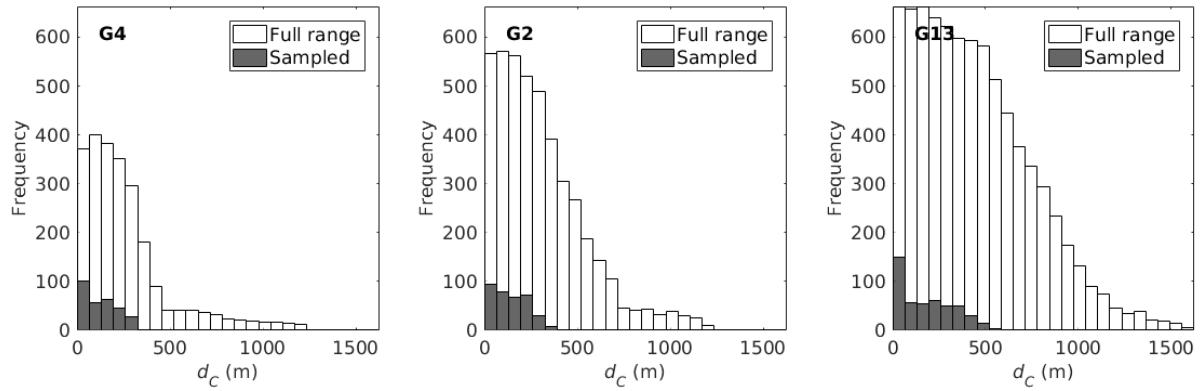


Figure 12: Histograms of distance from centreline ( $d_c$ ) sampled (black) as compared to total range (white) of distance from centreline of study glaciers.

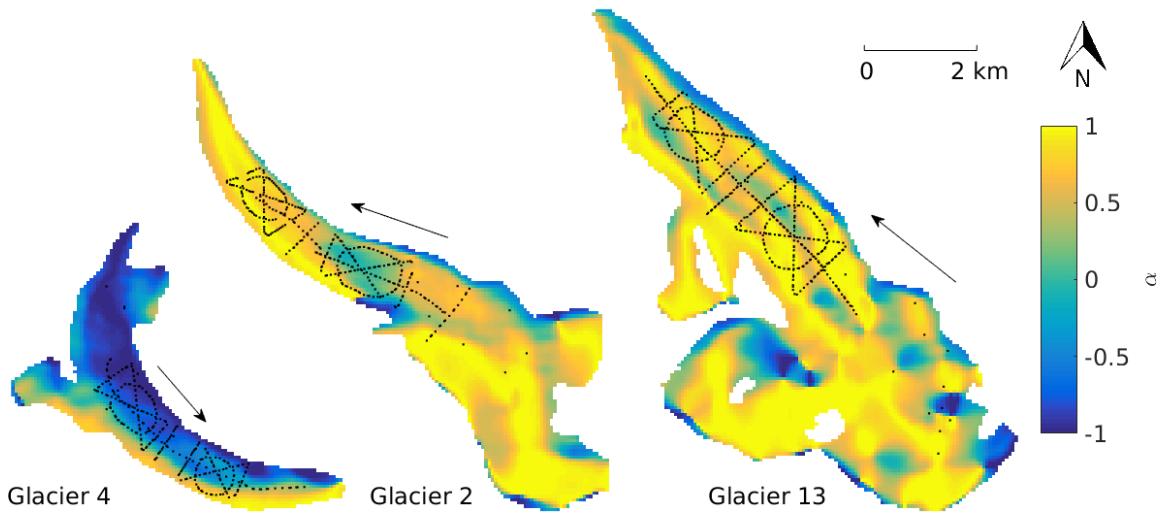


Figure 13: Distributions of the sine of aspect ( $\alpha$ ), which indicates north-south component of a slope (+1 defined as north), used in the topographic regressions for the study glaciers. Values are derived from a smoothed DEM. Arrows indicate glacier flow direction and black dots show snow depth sampling locations.

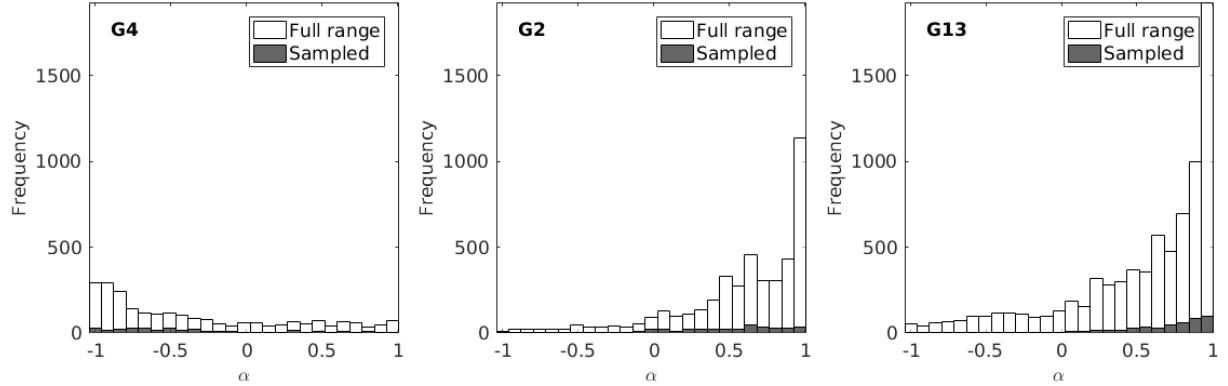


Figure 14: Histograms of aspect ( $\alpha$ ) sampled (black) as compared to total range (white) of aspect of study glaciers.

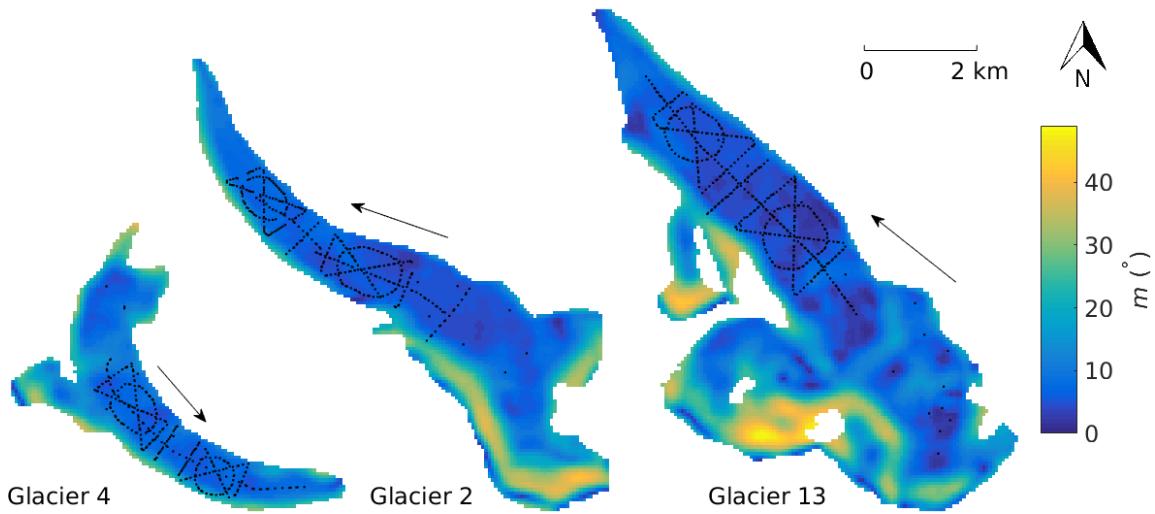


Figure 15: Distributions of slope ( $m$ ) used in the topographic regressions for the study glaciers. Values were derived from a smoothed DEM (grid size of  $40 \times 40$  m) in QGIS. Arrows indicate glacier flow direction and black dots show snow depth sampling locations.

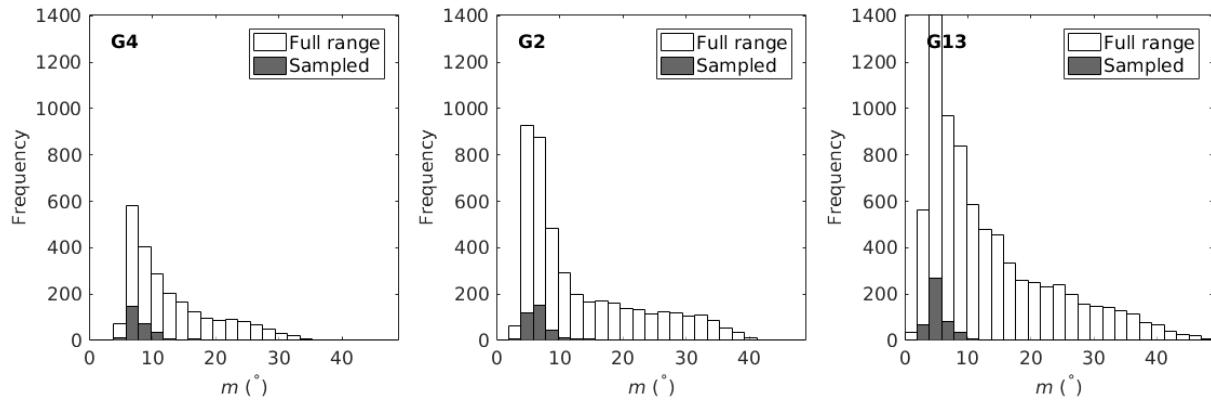


Figure 16: Histograms of slope ( $m$ ) sampled (black) as compared to total range (white) of slope of study glaciers.

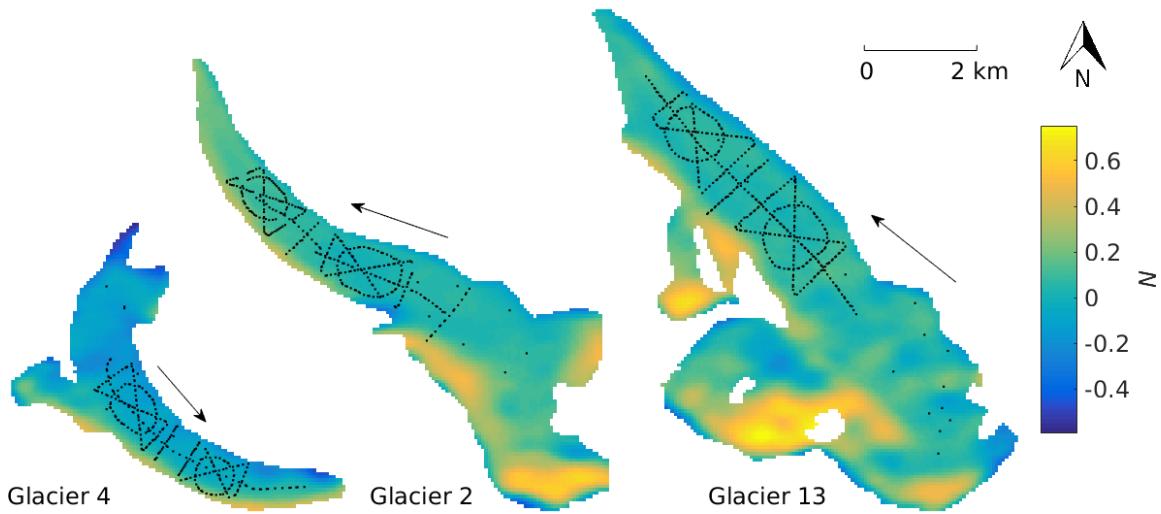


Figure 17: Distributions of “northness” ( $N$ ) used in the topographic regressions for the study glaciers. “Northness” is defined as the product of the cosine of aspect and sine of slope. A value of -1 represents a steep, south facing slope, a value of +1 represents a steep, north facing slope, and flat surfaces yield 0. Values were derived from a smoothed DEM (grid size of  $40 \times 40$  m) in QGIS. Arrows indicate glacier flow direction and black dots show snow depth sampling locations.

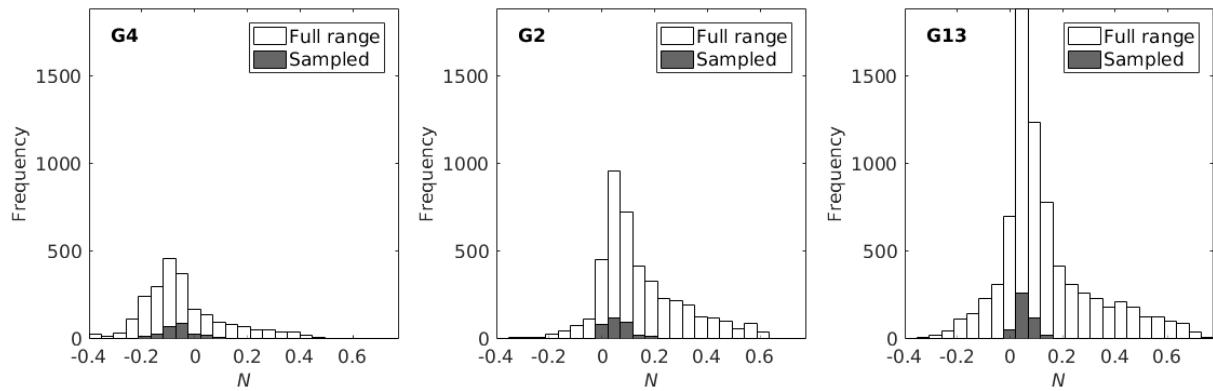


Figure 18: Histograms of “northness” ( $N$ ) sampled (black) as compared to total range (white) of “northness” of study glaciers.

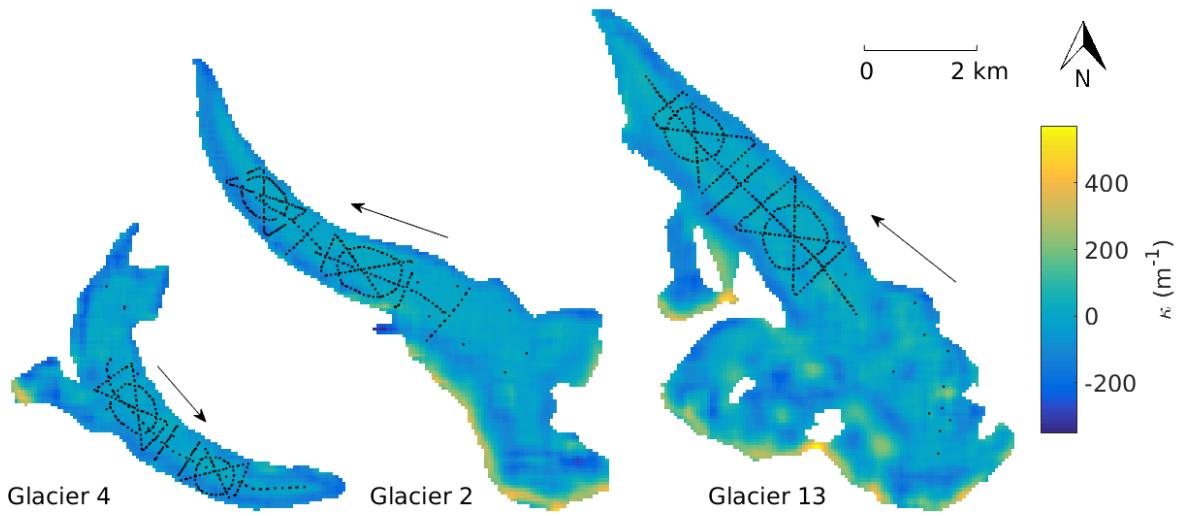


Figure 19: Distributions of curvature ( $\kappa$ ) used in the topographic regressions for the study glaciers. Values were derived from a smoothed DEM (grid size of  $40 \times 40$  m) in QGIS. Colour axis has been scaled to better resolve values close to zero. Arrows indicate glacier flow direction and black dots show snow depth sampling locations.

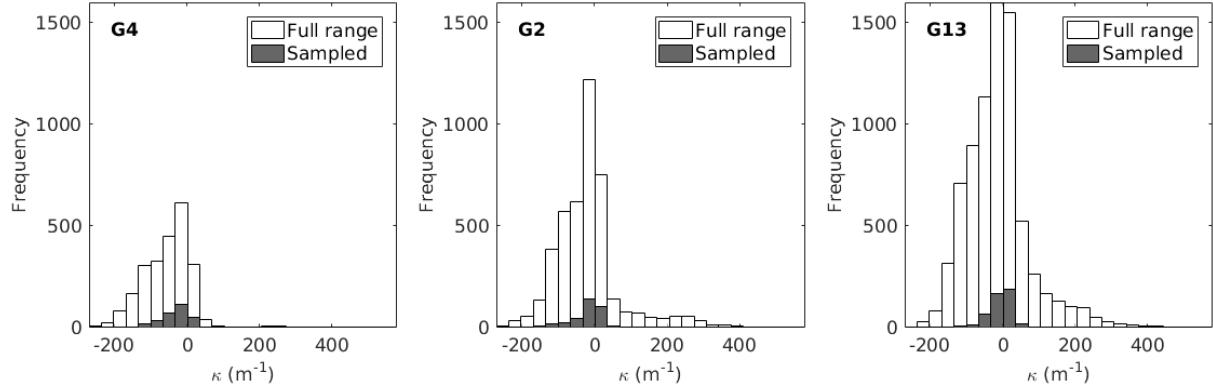


Figure 20: Histograms of curvature ( $\kappa$ ) sampled (black) as compared to total range (white) of profile curvature of study glaciers.

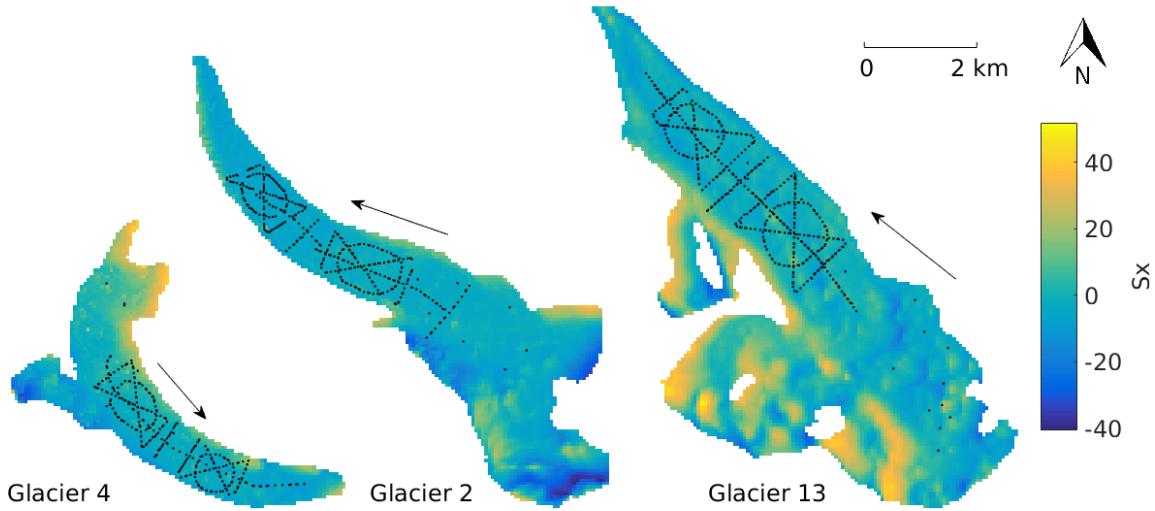


Figure 21: Distributions of  $S_x$ , which is a wind redistribution parameter, used in the topographic regressions for the study glaciers. See section 1.2 and the original paper by Winstral and others (2002) for more details on calculation. See Table 1 for values of best correlated azimuth and maximum search distance for each glacier. Arrows indicate glacier flow direction and black dots show snow depth sampling locations.

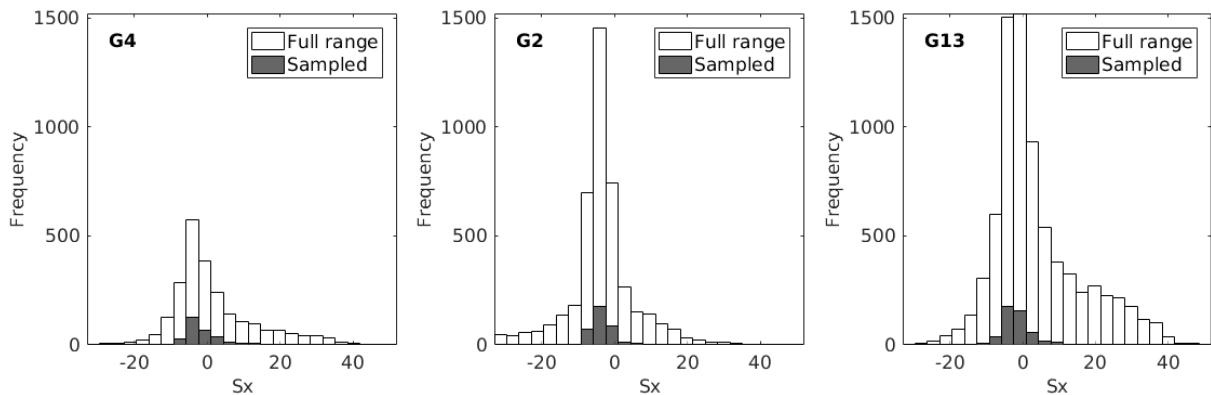


Figure 22: Histograms of  $S_x$  sampled (black) as compared to total range (white) of  $S_x$  of study glaciers.

## 2 Linear Regressions

### 2.1 Background

Relating snow accumulation and terrain parameters to better predict accumulation within a basin has been employed for decades (e.g. Woo and Marsh, 1978; Molotch and others,

2005; McGrath and others, 2015). The most common type of relation between topographic parameters and accumulation is a linear regression, where the observed snow water equivalent (SWE) is related to a linear combination of topographic parameters at each measurement location.

A linear regression takes the form

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (8)$$

where the matrix  $\mathbf{X}$  contains the set of independent regressors  $\mathbf{x}_i$  used to explain the dependant variable  $\mathbf{y}$  (e.g. Davis and Sampson, 1986). The regression coefficient for each regressor is given by  $\boldsymbol{\beta}$  and the error of the system is given by  $\boldsymbol{\varepsilon}$ . Applied to this study, the matrix of independent regressors ( $\mathbf{X}$ ) contains the topographic parameters at the sampling locations, the dependent variable  $\mathbf{y}$  contains the observed SWE, and the  $\boldsymbol{\beta}$  values are determined using a fitting model. While there are many types of fitting models, the ones employed in this study are multiple linear regression (MLR) and Bayesian model averaging (BMA).

To prevent over fitting of the data, regressions are calculated using cross validation. This means that for each regression, a randomly selected portion of the data is used to estimate regression coefficients and the coefficients are used to predict values that correspond to the remaining data (Kohavi and others, 1995). The root mean squared error (RMSE) between the estimated and observed data is then calculated. This process was repeated 1000 times and the regression coefficients that resulted in the lowest RMSE are then chosen for that model.

In this study, regressions for all possible combinations of topographic parameters are calculated. The total number of models is  $2^n$ , where  $n$  is the number of topographic parameters. Eight topographic parameters are used, resulting in  $2^8 = 256$  models. Model averaging is then used to determine the final regression coefficients. Model averaging is described in more details for multiple linear regressions (MLRs) in Section 2.4 and for Bayesian Model Averaging (BMA) in Section 2.5.

Once  $\boldsymbol{\beta}$  values have been estimated, they can then be used to predict values of the dependent variable in other locations where regressors are known (Davis and Sampson, 1986). For each grid cell, known values of topographic parameters can be multiplied by their respective  $\boldsymbol{\beta}$  coefficients and added together to obtain the modelled or predicted value of SWE.

## 2.2 Importance of variables in regression models

Regressions are used not only for estimating the response variable but also for assessing the relative importance of the regressors. Most types of regression models, including the ones used in this study, cannot be used to directly determine variable importance, so many additional metrics have been developed to address this need. Grömping (2015) lists seven simple metrics for assessing variable importance in a univariate regression (one dependant variable). Two of these metrics, raw correlation and semi-partial correlation, are chosen for this study. For both metrics, a larger value indicates a larger influence of a regressor in the model.

The first metric is the square of the raw correlations between regressors and response variable (Grömping, 2015). Raw correlation values can aid in identifying important variables for explaining and interpreting results because it is ignorant of which other regressors are included in the model (Darlington, 1968).

The second metric is the square of the semi-partial (or part) correlations for each regressor variable (Grömping, 2015). Semi-partial correlations are the correlation between the response variable and the residuals of the regression between each regressor and the remaining regressors. The value can be interpreted as the unique variance accounted for by  $\mathbf{x}_1$  in the presence of other predictors  $\mathbf{x}_2, \dots, \mathbf{x}_k$  (Darlington, 1968; Bring, 1996). Semi-partial correlations are helpful in identifying a small number of regressors that have the most influence in the regression (Grömping, 2009).

An assumption for semi-partial correlation is that regressors are independent of each other. When regressors are independent, the semi-partial correlations of all regressors sum to the total coefficient of determination ( $R^2$ ) of the regression (Grömping, 2015). In this study, the regressors are all independent ( $R^2 < 0.35$  for all glaciers) except for aspect and northness, which have  $R^2$  values of 0.79, 0.83 and 0.64 for Glaciers 4, 2 and 13, respectively. Despite this correlation, semi-partial correlation is used because of its simplicity.

## 2.3 Snow density estimation methods

In this study, snow density was not measured at every location where snow depth was measured. Therefore, snow density values need to be estimated at snow depth measurement locations in order to estimate observed SWE values. There are a number of different methods for interpolating between snow density measurement locations. Four methods were chosen for this study and include (1) using a constant value, the mean of all observed snow density values, for all measurement locations, (2) using a constant value (mean) for each glacier, (3) using a linear regression of elevation and snow density to interpolate for each glacier, and (4) taking an inverse-distance weighted mean of density observations. As discussed in Section ??, the snowpit-derived densities and Federal Sampler-derived densities are inconsistent so these two data set are kept separate for the analysis. In total there are eight different methods for estimating density, as summarized in Table 4.

Table 4: Description of density interpolation methods used to calculate SWE used in the topographic regression. Abbreviations with ‘S’ used snowpit-derived densities and abbreviations with an ‘F’ used Federal Sampler-derived densities.

	<b>Source of snow density</b>	<b>Estimation method</b>
	<i>Snowpit</i> <i>Federal Sampler</i>	
S1	■	Mean of all glaciers
F1		
S2	■	Glacier mean
F2		
S3	■	Linear regression of elevation and density for each glacier
F3		
S4	■	Inverse distance weighted mean
F4		

## 2.4 Multiple Linear Regression (MLR)

### 2.4.1 Background

Perhaps the most basic and well used method for relating SWE and topographic parameters is a multiple linear regression (MLR) (e.g. Cohen and others, 2013). The best fit line of one type of MLR is the one described by coefficients that minimize the sum of squares of the vertical deviations of each data point ( $Y_i$ ) from the estimated value according to the equation ( $\hat{Y}_i$ ) (Davis and Sampson, 1986)

$$\sum_{i=1}^n (\hat{Y}_i - Y_i)^2 = \text{minimum.} \quad (9)$$

Note that if a point falls on the line then the deviation is zero and the positive and negative deviations from the line do not cancel because the values are first squared and then summed. The residuals are simply the differences between the estimated and observed data values. To prevent over fitting of the data, cross validation is done for each MLR, as described in Section 2.

In this study, there are  $2^8$  models that encompass all possible linear combinations of topographic parameters. There is no reason to favour any of the models so a weighted sum of all models is used to estimate regression coefficients. The Bayesian information criterion (BIC) value is used to assess the relative predictive success of each model.

A BIC value is found according to

$$\text{BIC} = -2 \ln L(\hat{\theta}_k | y) + k \ln(n), \quad (10)$$

where the values of  $\hat{\theta}_k$ , which are the model parameters, maximize the likelihood function for data  $y$  (Burnham and Anderson, 2004). The likelihood function,  $L(\hat{\theta}_k | y)$ , is the probability of the model parameters occurring given the data. The number of data points is

$n$  and the number of regressors is  $k$ . BIC values are used to assess the relative predictive success of models while penalizing for overfitting of data. While the absolute BIC value is meaningless, models can be selected or averaged using the relative BIC values, with lower values indicating a better model (Burnham and Anderson, 2004).

The BIC value for each model ( $BIC_i$ ) is used to determine the normalized weight of each model ( $w_i$ ) relative to the best model (lowest BIC value indicated by  $BIC_{\min}$ ) as defined by the equation (Burnham and Anderson, 2004)

$$w_i = \frac{e^{-0.5(BIC_i - BIC_{\min})}}{\sum_{i=1}^R e^{-0.5(BIC_i - BIC_{\min})}}. \quad (11)$$

Parameters not included in a particular model are assigned coefficients of zero. The sum of the weighted coefficients gives the final  $\beta$  values.

#### 2.4.2 Methods

The MLR of SWE values and topographic parameters is done in Matlab (Appendix 4.2). The best set of regression coefficients for each model is selected using cross-validation of a linear regression, with the coefficients chosen by minimizing the vertical sum of squares. Then, the regression coefficients from all models are weighted according to the BIC value of the model.

#### 2.4.3 Results and Discussion

The importance of the various topographic parameters differs for the three study glaciers (Table 5). The regression for Glacier 2 explains a large portion of the variance ( $R^2=0.66$ ), although the RMSE is higher than that of Glacier 13, for which the regression explains less variance ( $R^2=0.40$ ) (Figure 23). Glacier 4 has the least variance explained by the regression ( $R^2=0.12$ ) and the highest RMSE. The intercepts of the regression are similar for Glaciers 2 and 13 ( $\sim 0.25$  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 25).

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 6) and for the mean of all density options (Table 5). Sx has the highest mean semi-partial correlation (0.034) and raw correlation (0.065) (Table 5). The Sx coefficient is negative, which indicates less snow in ‘sheltered’ areas. The negative correlation is counter intuitive so it is surprising that Sx is the best predictor for accumulation.

The map of estimated SWE for the entire glacier shows a relatively uniform SWE distribution over Glacier 4 (Figure 28), 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,

Table 5: Mean MLR and BMA coefficients for topographic regression between measured SWE and standardized topographic parameters. Topographic parameters are distance from centreline ( $d_C$ ), elevation ( $z$ ), aspect ( $\alpha$ ), slope ( $m$ ), northness ( $N$ ), curvature ( $\kappa$ ), and  $S_x$ . Since parameters are standardized, the magnitude of the coefficients is representative of their importance in predicting SWE. The root-mean-squared error (RMSE) between modelled SWE using those coefficients and observed SWE is also given. Semi-partial correlation is a metric that describes the increase in  $R^2$  from the addition of a parameter to a regression that contains all other regressors. Raw correlation is the square of the Pearson correlation between a parameter and SWE.

	$z$	$d_C$	$\alpha$	$m$	$N$	$\kappa$	$S_x$	Intercept	RMSE
<b>Glacier 4</b>	MLR Coefficient	0.008	-0.001	-0.012	-0.004	-0.002	0.016	-0.051	0.619
	BMA Coefficient	0.006	-0.001	-0.010	-0.007	-0.003	0.016	-0.050	0.619
	Semi-partial $R^2$	0.015	<0.001	<0.001	0.007	0.001	0.010	0.034	—
	Raw correlation	0.003	0.044	0.001	0.027	0.001	0.041	0.065	—
<b>Glacier 2</b>	MLR Coefficient	0.110	0.008	-0.010	0.026	0.011	0.001	0.036	0.262
	BMA Coefficient	0.111	0.008	-0.011	0.029	0.012	0.002	0.036	0.261
	Semi-partial $R^2$	0.205	0.014	0.004	<0.001	0.003	0.004	0.028	—
	Raw correlation	0.586	0.031	0.211	0.100	0.193	0.078	0.333	—
<b>Glacier 13</b>	MLR Coefficient	0.054	<0.001	<0.001	0.001	0.001	-0.021	0.003	0.229
	BMA Coefficient	0.054	<0.001	-0.001	<0.001	0.001	-0.019	0.003	0.228
	Semi-partial $R^2$	0.245	0.003	0.007	0.002	0.005	0.008	0.009	—
	Raw correlation	0.347	0.051	0.012	<0.001	0.022	0.040	0.083	—

Table 6: MLR coefficients for topographic regression between measured SWE and standardized topographic parameters. Topographic parameters are distance from centreline ( $d_C$ ), elevation ( $z$ ), aspect ( $\alpha$ ), slope ( $m$ ), northness ( $N$ ), curvature ( $\kappa$ ), and  $S_x$ . Since parameters are standardized, the magnitude of the coefficients is representative of their importance in predicting SWE. The root-mean-squared error (RMSE) between modelled SWE using those coefficients and observed SWE is also given. See Table 4 for description of density options.

	Parameter	<b>S1</b>	<b>F1</b>	<b>S2</b>	<b>F2</b>	<b>S3</b>	<b>F3</b>	<b>S4</b>	<b>F4</b>
<b>Glacier 4</b>	$z$	0.011	0.006	0.008	0.009	0.018	<0.001	0.008	0.003
	$d_C$	<0.001	-0.002	<0.001	-0.002	-0.001	-0.001	<0.001	<0.001
	$\alpha$	-0.021	-0.006	-0.024	-0.013	-0.016	-0.002	-0.008	-0.010
	$m$	-0.002	-0.004	-0.002	-0.005	-0.013	-0.002	-0.004	-0.003
	$N$	-0.002	-0.002	-0.002	-0.002	-0.004	-0.001	-0.001	-0.002
	$\kappa$	0.009	0.021	0.012	0.018	0.009	0.028	0.016	0.015
	$S_x$	-0.059	-0.046	-0.055	-0.053	-0.057	-0.051	-0.045	-0.042
<b>Glacier 2</b>	Intercept	0.617	0.567	0.626	0.631	0.621	0.642	0.618	0.633
	RMSE	0.144	0.134	0.146	0.149	0.146	0.148	0.144	0.147
	$z$	0.119	0.109	0.114	0.099	0.102	0.120	0.109	0.109
	$d_C$	0.009	0.021	0.008	0.003	0.007	0.002	0.011	0.001
<b>Glacier 13</b>	$\alpha$	-0.011	-0.007	-0.019	-0.009	-0.013	-0.003	-0.011	-0.005
	$m$	0.030	0.024	0.023	0.023	0.025	0.029	0.022	0.029
	$N$	0.012	0.008	0.022	0.009	0.012	0.003	0.011	0.008
	$\kappa$	0.003	0.002	0.002	0.001	0.001	0.001	0.001	0.001
	$S_x$	0.040	0.036	0.040	0.031	0.035	0.027	0.040	0.037
	Intercept	0.287	0.263	0.275	0.235	0.273	0.240	0.282	0.238
	RMSE	0.096	0.089	0.093	0.081	0.093	0.082	0.096	0.084
<b>Glacier 13</b>	$z$	0.058	0.052	0.055	0.052	0.045	0.058	0.054	0.055
	$d_C$	<0.001	<0.001	0.001	0.001	0.001	<0.001	<0.001	<0.001
	$\alpha$	<0.001	<0.001	<0.001	-0.002	<0.001	<0.001	<0.001	<0.001
	$m$	0.001	<0.001	<0.001	0.001	0.001	0.001	<0.001	<0.001
	$N$	0.002	<0.001	<0.001	0.001	0.001	<0.001	<0.001	<0.001
	$\kappa$	-0.024	-0.020	-0.021	-0.019	-0.023	-0.018	-0.021	-0.020
	$S_x$	0.001	0.002	0.001	0.007	0.004	0.008	0.001	0.001
Glacier 13	Intercept	0.236	0.220	0.241	0.217	0.250	0.208	0.246	0.211
	RMSE	0.078	0.073	0.080	0.072	0.084	0.068	0.081	0.071

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.

For Glacier 2, the most important regressor is elevation (Table 5). 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  $R^2$  and raw correlation (Table 5). Sx is the second most important regressor and has a positive correlation, which indicates that ‘sheltered’ areas are likely to have high accumulation.

The map of modelled SWE on Glacier 2 closely matches that of elevation (Figure 9), 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 (Table ??). 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 most important regressor for Glacier 13 is elevation (Table 5). 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  $R^2$  value is the largest between the glaciers. The high semi-partial  $R^2$  value indicates that when elevation is added to the regression the total variance explained increases considerably because the remaining regressors are not important. The map of estimated SWE on Glacier 13 (Figure 28) closely follows elevation although the range of SWE values is relatively small so the elevation effect is less pronounced than on Glacier 2.

Qualitatively, there is little variation in the fit between modelled and observed winter balance between the various density options for all glaciers and the residuals display a similar distribution between the density options (Figure 25). The semi-partial  $R^2$  and raw correlation for each regressor also varies little between different density options (Figure 27).

The choice of density measurements and interpolation techniques does not affect the relative importance of regressors. Although there is a range of coefficient values that result from the choice of density options (Figure 26), the relative dominance of Sx for Glacier 4 and elevation for Glacier 2 and 13 is consistent. The largest range in semi-partial correlation is in elevation for Glaciers 2 and 13 (Figure 27). Estimating SWE using the linear elevation regression and Federal Sampler-derived density (F3) resulted in the highest semi-partial correlation for Glaciers 2 and 13 (0.24 and 0.30, respectively). Since the Federal Sampler derived densities are correlated with snow depth, which is correlated with elevation, it is likely that these density values amplify the elevation component of the regression.

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 29). 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 difference in estimated SWE expressed as a percent (Figure 30) is relatively consistent on each glacier. The mean percent difference is highest on Glacier 2 (25%) and slightly lower on Glaciers 4 and 13 (18% and 20%, respectively). Extreme values in percent difference are located in areas with low values of estimated SWE (small denominator), including the terminus of Glaciers 2 and 13 as well as parts of the accumulation area on Glacier 4.

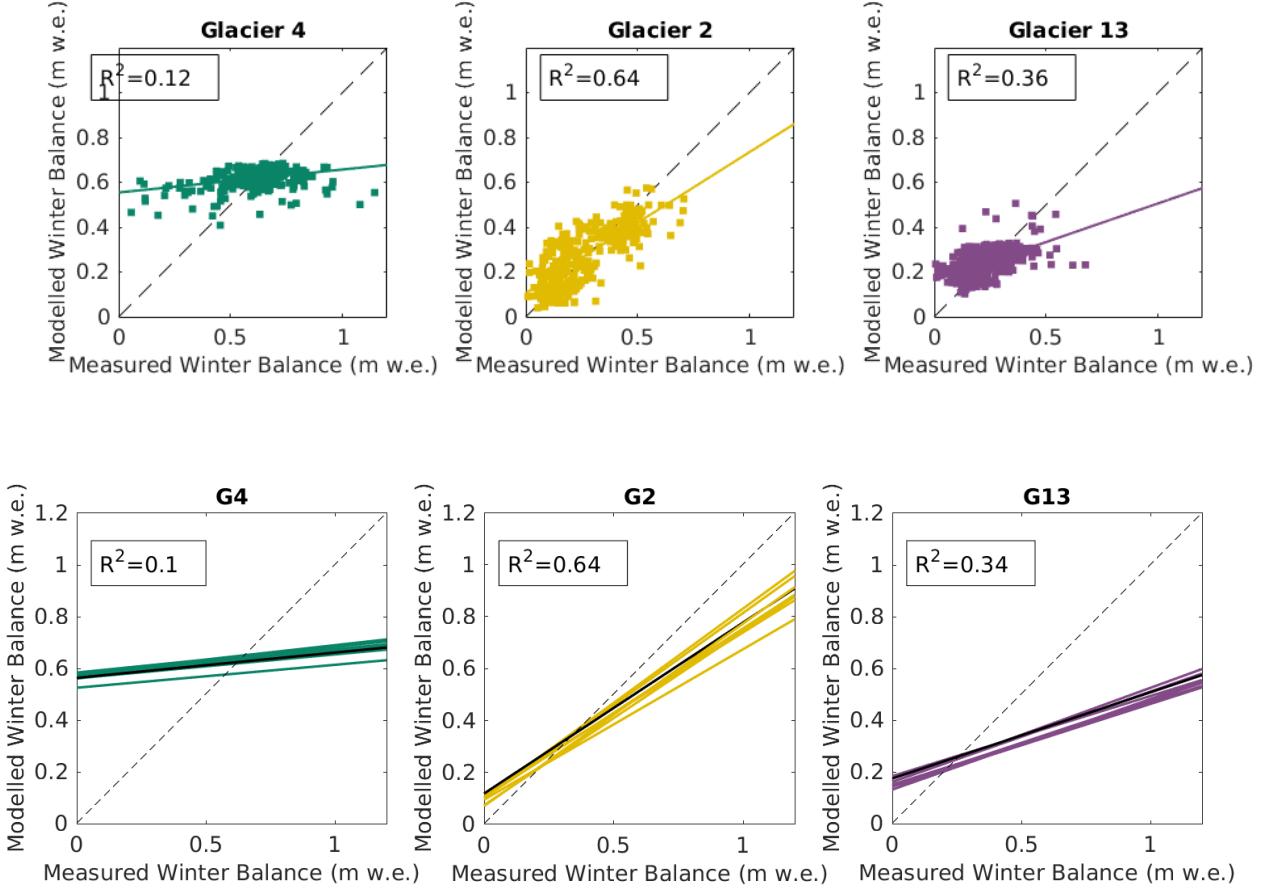


Figure 23: Top panel shows comparison of estimated (MLR) and observed (original) snow water equivalent (SWE) for three study glaciers. The SWE values were calculated using inverse-distance weighted snowpit densities (S4). Bottom panel shows plots of all linear fits between estimated and observed SWE using eight options for calculating density. Mean  $R^2$  value is shown for each sub-plot and a reference 1:1 line is also provided. Black line highlights the S4 option from the top panel. See Figure 31 for a plot of all estimated SWE values.

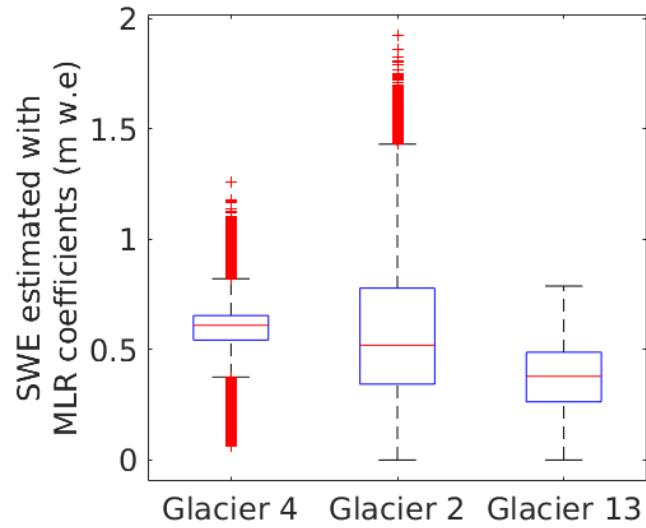


Figure 24: Summary of estimated SWE values found using MLR coefficients. Red line indicates median, blue box shows first quantiles, bars indicate minimum and maximum values (excluding outliers), and red crosses show outliers, which are defined as being outside of the range of 1.5 times the quartiles (approximately  $\pm 2.7\sigma$ ).

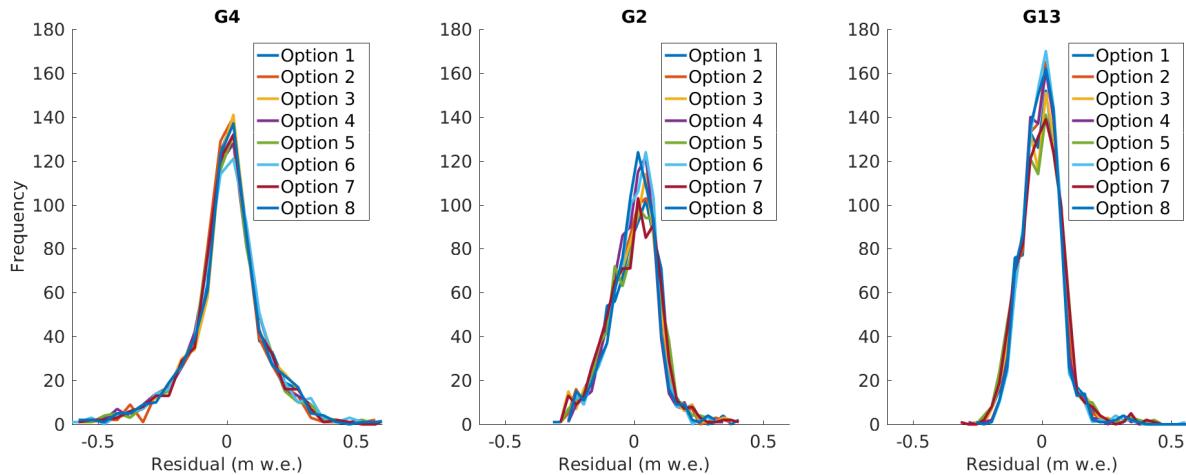


Figure 25: Residuals of SWE predicted using MLR for all options of estimating density.

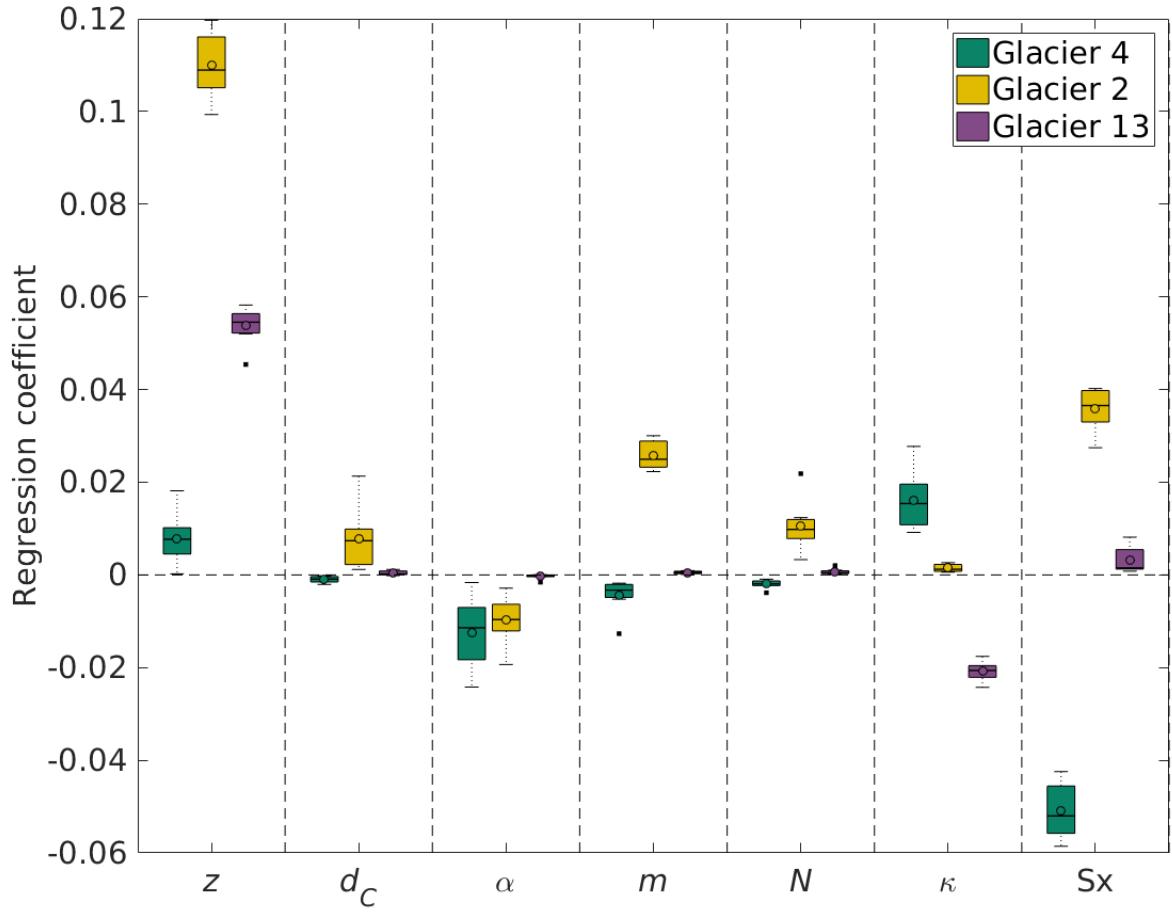


Figure 26: Boxplot showing the range of regressor coefficients explained by each topographic parameter for each option of estimating snow water equivalent (SWE). Topographic parameters are distance from centreline ( $d_C$ ), elevation ( $z$ ), aspect ( $\alpha$ ), slope ( $m$ ), northness ( $N$ ), curvature ( $\kappa$ ), and Sx. Within each box, the mean is shown as a circle, the median as a horizontal line, the interquartile range (IQR) as a coloured box, two times the IQR as dashed lines beyond the box, and outliers as single points.

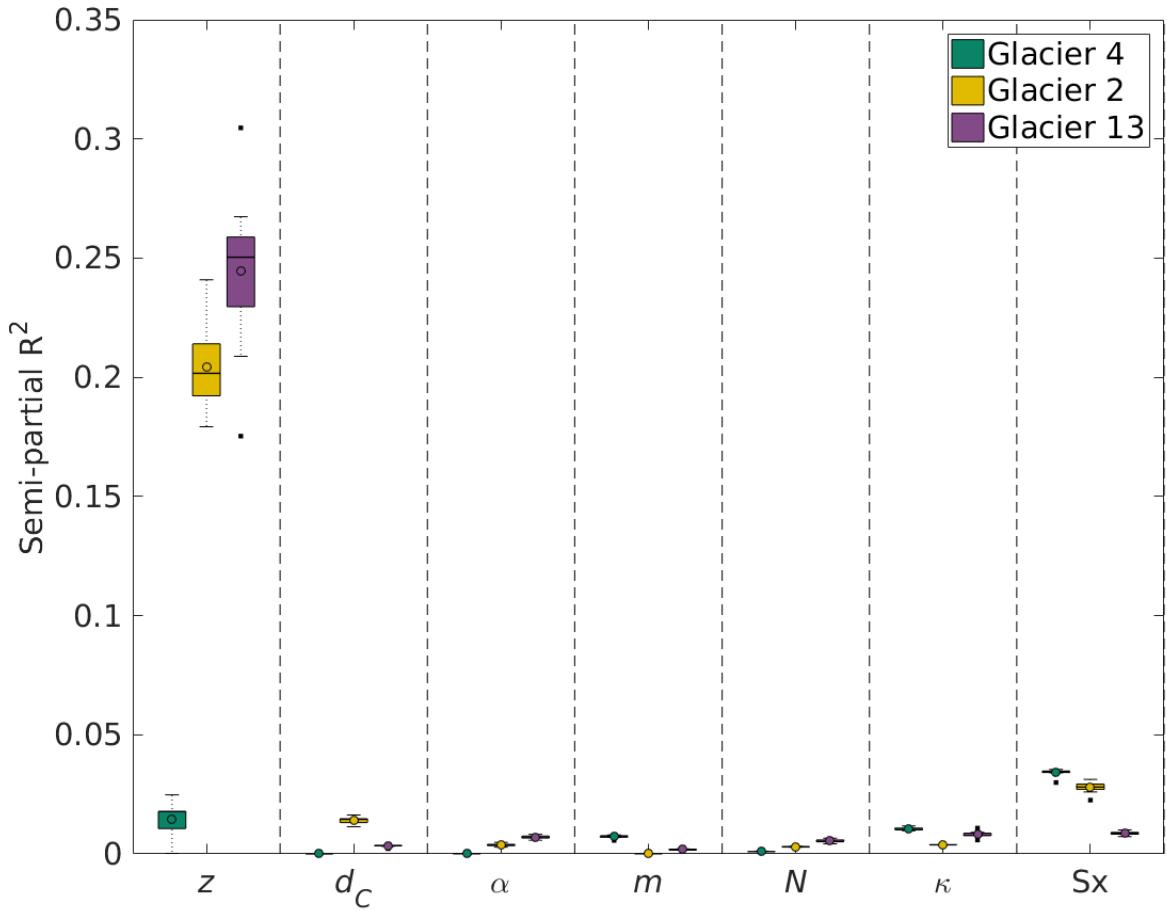


Figure 27: Boxplot showing the range of semi-partial correlation explained by each topographic parameter for each option of estimating snow water equivalent (SWE). Topographic parameters are distance from centreline ( $d_C$ ), elevation ( $z$ ), aspect ( $\alpha$ ), slope ( $m$ ), northness ( $N$ ), curvature ( $\kappa$ ), and Sx. Within each box, the mean is shown as a circle, the median as a horizontal line, the interquartile range (IQR) as a coloured box, two times the IQR as dashed lines beyond the box, and outliers as single points.

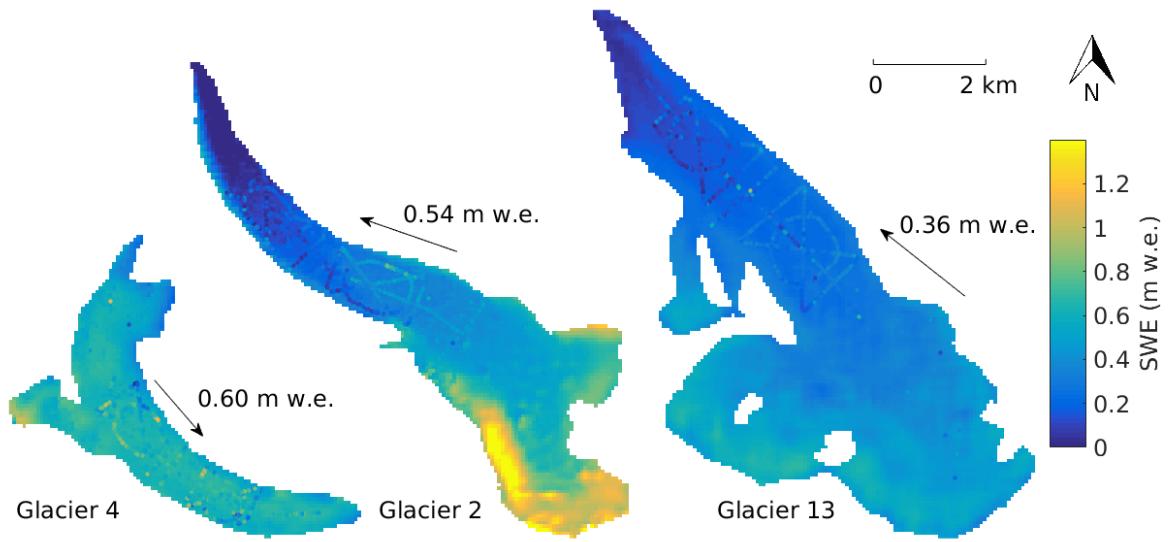


Figure 28: Modelled SWE using coefficients determined using MLR and density interpolated with inverse-distance weighting from snowpits (S4). Observed SWE values are overlaid on the maps.

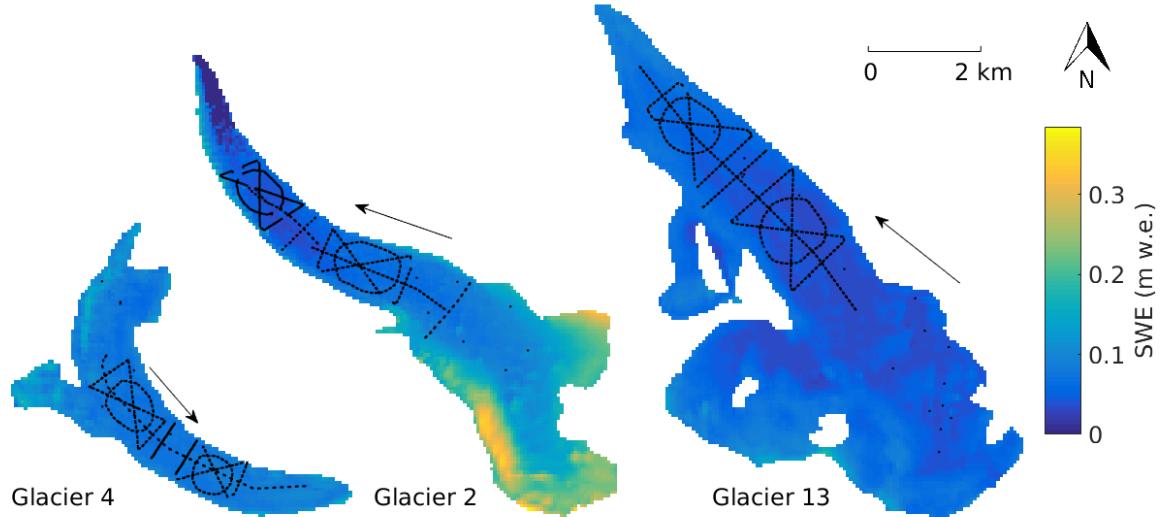


Figure 29: Map of the difference between maximum and minimum SWE values for each DEM cell between all density options using MLR coefficients. Observed SWE values are overlaid on the maps.

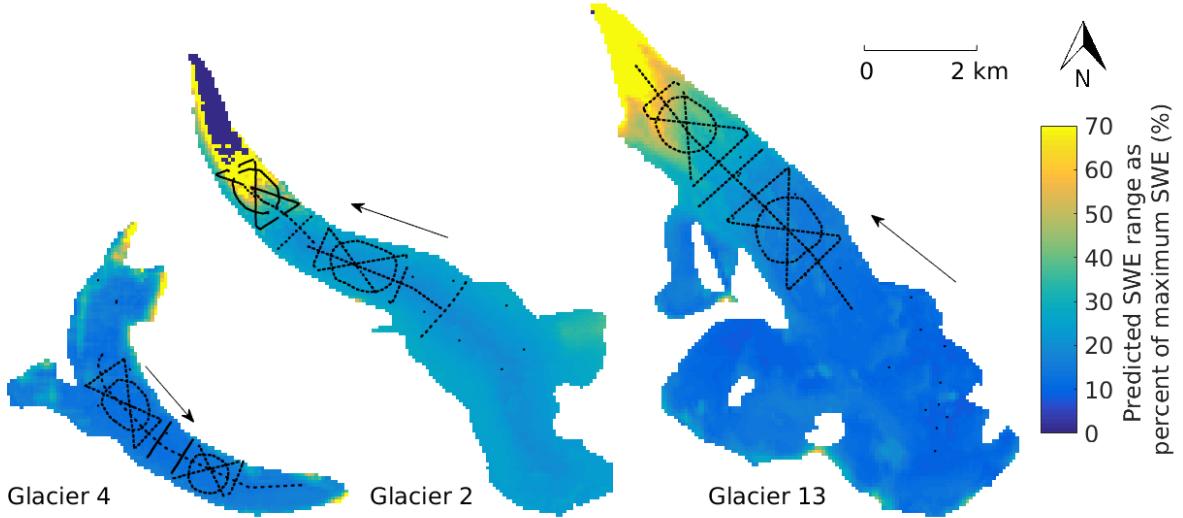


Figure 30: Map of the difference between maximum and minimum SWE values, expressed as a percent of the maximum SWE, for each DEM cell between all density options using MLR coefficients. The colours have been scaled to highlight difference in the main part of the glaciers. Values of zero are found where the minimum estimated SWE is zero m w.e. Observed SWE values are overlain on the maps.

## 2.5 Bayesian Model Averaging (BMA)

### 2.5.1 Background

Bayesian model averaging (BMA) is a method of estimating all possible linear combinations of predictor variables, in this case topographic parameters, and then averaging over all models (Raftery and others, 1997; Wasserman, 2000; Raftery and others, 2005). This method is based on Bayesian principals in which the probability of an outcome is determined based on an initial probably distribution that is determined by the researcher, as well as the data provided. Given that the predictive outcome has a probability distribution of  $x$  given  $y$ , written as  $P(x|y)$ , we use Bayes' theorem to to write this as

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)}. \quad (12)$$

$P(x|y)$  is often called the posterior model probability (PMP). The quantity  $P(y|x)$  is the likelihood function, which determines unknown parameters from a known outcome (i.e. observed data). The term  $P(x)$  is an observer determined prior probability distribution (typically just called a *prior*) and it reflects the prior knowledge of the system (Raftery and others, 1997). Choosing an appropriate prior is one of the most challenge components of Bayesian probability theory (Wasserman, 2000). The  $P(y)$  term can be obtained by integrating  $P(y|x)P(x)$  over all  $x$  and is thus a constant for all models that is typically discarded (Wasserman, 2000).

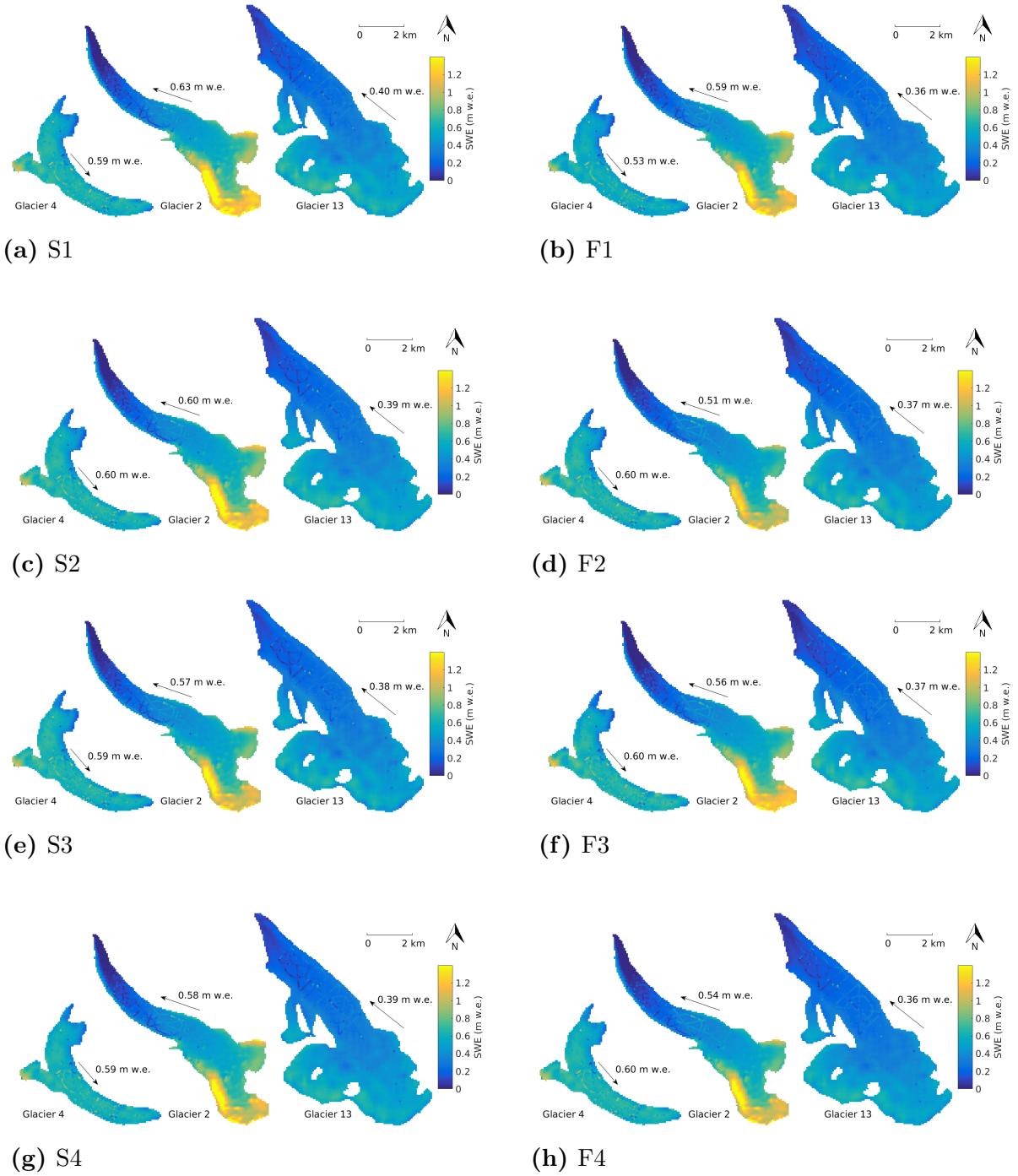


Figure 31: Map of modelled SWE using the MLR coefficient values for all density options. Measured SWE is plotted as overlaid filled circles. Glacier flow directions are indicated by black arrows and mean estimated SWE values for each glacier are shown.

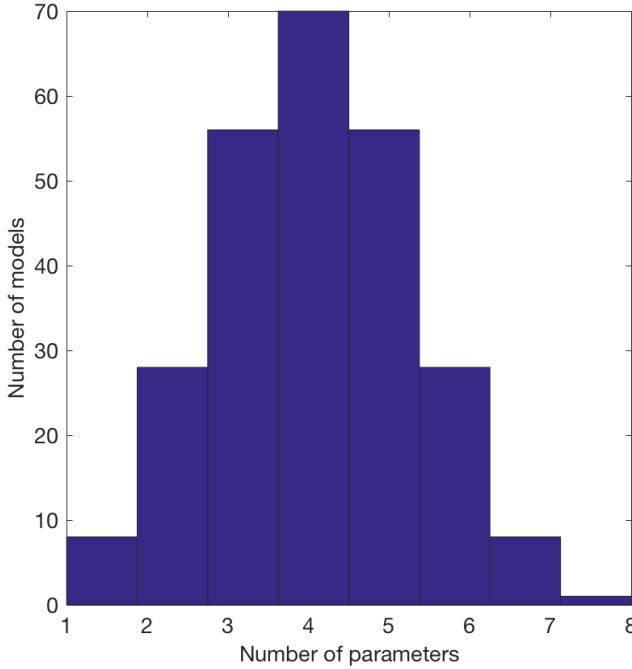


Figure 32: Uniform model prior for eight topographic regressors used in BMA.

of data then the prior will have little effect on the posterior. The final coefficients for the linear combination of predictor variables is often reported as posterior distribution means or values that maximize the log-likelihood.

Within BMA, Bayes' theorem is used to find the posterior model probability and the PMP is used as a weight when averaging over all models (Wasserman, 2000). The model weighted posterior distribution for the coefficients  $\beta$  of  $k$  models after normalization is given by

$$P(\beta|y, X) = \sum_{i=1}^{2^k} P(M_i|X, y)P(\beta|M_i, y, X), \quad (14)$$

where the responding variable is given by  $y$  and the matrix of variables is given by  $X$  (Raftery and others, 1997). Here, the model prior is  $P(M_i|X, y)$  and the likelihood of the  $\beta$  coefficient is  $P(\beta|M_i, y, X)$ .

There are a number of different priors to describe model size distribution that have been applied in BMA. A commonly used prior is the ‘uniform’ prior, which assumes a normal model distribution with a total of  $2^n$  models, where  $n$  is the number of regressors (Wasserman, 2000). This model prior states that the observer has no knowledge of the system and all models are equally likely. The uniform prior has a prior model probability of the form  $P(M_i) = 2^{-n}$  (Figure 32), which is symmetric about the mean  $n/2$  (Zeugner and Feldkircher, 2015). This prior inherently favours models of an intermediate size.

Other types of priors include those that are skewed to favour smaller models, ones

Together then, the posterior model probability is a function of both the model prior, specified by the researcher, as well as the distribution of the observed data — the PMP is the transformation of the prior as a result of considering collected data (Wasserman, 2000). This can be loosely expressed as

$$\text{posterior} \propto \text{prior} \times \text{likelihood}. \quad (13)$$

If the prior is uninformative then the posterior will be strongly influenced by the data (Wasserman, 2000). An informative prior will result in a posterior that is a mix of the prior and the data. As the prior becomes more informative, the amount of data needed to transform the distribution increases. If there is a large amount

with equal probability for all model sizes, and ones with varying probability for individual regressors. In this project, a uniform prior was chosen for two reasons. First, there was no knowledge of the model distribution so we aim to minimize the observer influence on the final distribution. Second, the MLR estimation (Section 2.4) assumes a uniform distribution so choosing a uniform distribution for BMA allows for consistency between methods. Using a uniform distribution is commonly done in BMA (Wasserman, 2000).

With a small number of regressors, the posterior of all possible models can be determined. However, with a large set of regressors, this computation becomes increasingly expensive. To reduce computation time for a large set of regressors, BMA can use Markov chain Monte Carlo (MCMC) model composition to directly approximate the posterior distribution (Wasserman, 2000). In our study, there are eight regressors so  $2^8 = 256$  models. It is possible to visit all models and to obtain an exact solution so an MCMC model composition was not employed.

BMA allows for the calculation of a metric called the posterior inclusion probability (PIP), which is used to evaluate the importance of a regressor in explaining the observed data. PIP is the sum of all posterior model probabilities (PMP) where the variable was included in the model (Zeugner and Feldkircher, 2015). A higher PIP indicates that the regressor is more important in the regression.

### 2.5.2 Methods

The BMA process was implemented in R (Appendix 4.3), using the Bayesian model statistics (BMS) package developed by Zeugner and Feldkircher (2015). The package computes the posterior distribution mean value of all  $\beta$  coefficients for topographic parameters as well as the percent variance explained by each parameter.

### 2.5.3 Results and Discussion

The regression coefficients generated using BMA and their relative importance is similar to those generated by MLR. The most important topographic parameter is Sx on Glacier 4 and elevation ( $z$ ) on Glaciers 2 and 13 (Tables 5 and 7). The variance explained by the BMA models is also comparable to that of the MLR models (Figures 33 and 34). The distribution of residuals is largest for Glacier 4 as a result of the poor fit of the regression (Figures 35). For a more detailed discussion of the regression of SWE and topographic parameters refer to Section 2.4.

The distribution of estimated glacier-wise SWE is different between glaciers but consistent with the distribution found using MLR (Figures 36). The estimated SWE range is greatest for Glacier 2 and there are a large number of outliers on Glacier 4. Mean estimated SWE decreases with distance from the topographic divide, which is consistent with the observed mean SWE (Figure ??). The range of regression coefficient values found using the different density options (Figure 37) and the semi-partial correlation range (Figure 38) are also similar to those found using MLR. The glacier wide estimates of SWE (Figure 39), difference between density option variation in SWE (Figure 41), and SWE difference as percent (Figure ??) resemble those of MLR. The mean percent difference is

Table 7: BMA coefficients for topographic regression between measured SWE and standardized topographic parameters. Topographic parameters are distance from centreline ( $d_C$ ), elevation ( $z$ ), aspect ( $\alpha$ ), slope ( $m$ ), northness ( $N$ ), curvature ( $\kappa$ ), and  $S_x$ . Since parameters are standardized, the magnitude of the coefficients is representative of their importance in predicting SWE. The root-mean-squared error (RMSE) between modelled SWE using those coefficients and observed SWE is also given. See Table 4 for description of density options.

	Parameter	<b>S1</b>	<b>F1</b>	<b>S2</b>	<b>F2</b>	<b>S3</b>	<b>F3</b>	<b>S4</b>	<b>F4</b>
<b>Glacier 4</b>	$z$	0.007	0.010	0.004	0.003	0.014	-0.001	0.005	0.004
	$d_C$	<0.001	<0.001	-0.001	-0.002	-0.001	-0.002	<0.001	-0.001
	$\alpha$	-0.022	-0.004	-0.008	-0.012	-0.015	-0.006	-0.007	-0.008
	$m$	-0.004	-0.003	-0.005	-0.004	-0.017	-0.007	-0.001	-0.013
	$N$	-0.006	-0.003	<0.001	-0.001	-0.005	-0.002	-0.003	<0.001
	$\kappa$	0.008	0.026	0.026	0.012	0.008	0.014	0.029	0.005
	$S_x$	-0.061	-0.044	-0.047	-0.047	-0.056	-0.051	-0.046	-0.052
<b>Glacier 2</b>	Intercept	0.618	0.578	0.623	0.631	0.618	0.635	0.614	0.637
	RMSE	0.104	0.099	0.104	0.098	0.110	0.110	0.108	0.112
	$z$	0.123	0.105	0.118	0.102	0.103	0.117	0.107	0.113
	$d_C$	0.005	0.015	0.010	0.006	0.011	0.003	0.009	0.004
<b>Glacier 13</b>	$\alpha$	-0.010	-0.022	-0.014	-0.006	-0.018	-0.004	-0.006	-0.007
	$m$	0.035	0.018	0.024	0.033	0.033	0.030	0.031	0.028
	$N$	0.013	0.025	0.016	0.005	0.018	0.006	0.003	0.011
	$\kappa$	0.001	<0.001	0.001	0.002	0.002	0.006	0.004	0.002
	$S_x$	0.037	0.034	0.032	0.033	0.039	0.037	0.042	0.036
	Intercept	0.280	0.263	0.273	0.235	0.275	0.241	0.276	0.243
	RMSE	0.082	0.075	0.078	0.068	0.077	0.070	0.081	0.070
<b>Glacier 13</b>	$z$	0.061	0.054	0.056	0.057	0.044	0.056	0.054	0.054
	$d_C$	<0.001	<0.001	0.001	<0.001	<0.001	<0.001	<0.001	0.001
	$\alpha$	-0.001	<0.001	-0.002	-0.002	<0.001	-0.002	<0.001	<0.001
	$m$	0.001	<0.001	0.002	-0.001	<0.001	<0.001	<0.001	0.001
	$N$	0.001	0.001	0.002	0.001	0.001	<0.001	0.001	<0.001
	$\kappa$	-0.021	-0.018	-0.016	-0.019	-0.021	-0.019	-0.023	-0.016
	$S_x$	0.003	0.003	0.007	0.001	0.006	0.005	0.001	0.002
Intercept	0.236	0.222	0.238	0.219	0.248	0.210	0.243	0.209	
	RMSE	0.063	0.056	0.063	0.058	0.068	0.056	0.067	0.052

highest for Glaciers 2 and 13 (22%) and lowest for Glacier 4 (15%). For a more detailed analysis of estimated SWE found using the regression coefficients, see Section 2.4. These consistent results indicate that the choice of regression method does not strongly affect the final regression of SWE and topographic parameters.

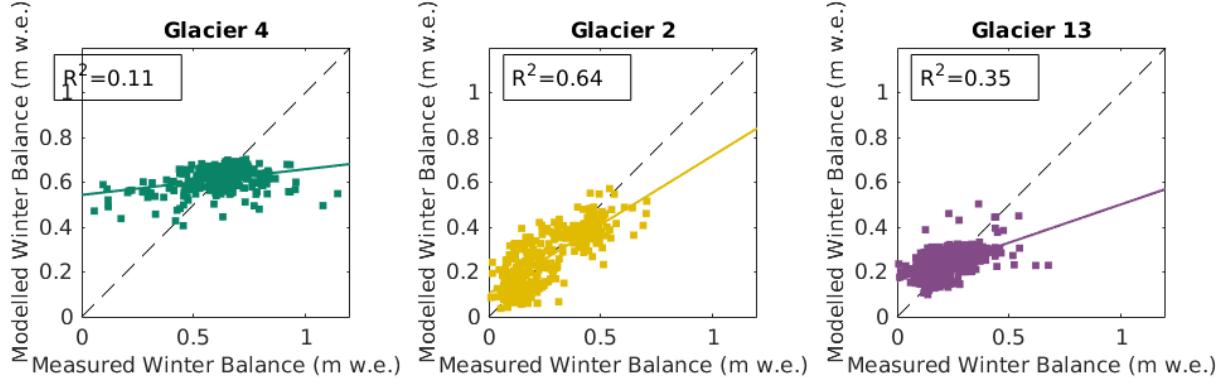


Figure 33: Comparison of predicted (BMA) and observed (original) snow water equivalent (SWE) for three study glaciers. The SWE values were calculated inverse distance weighted snowpit densities (S4).

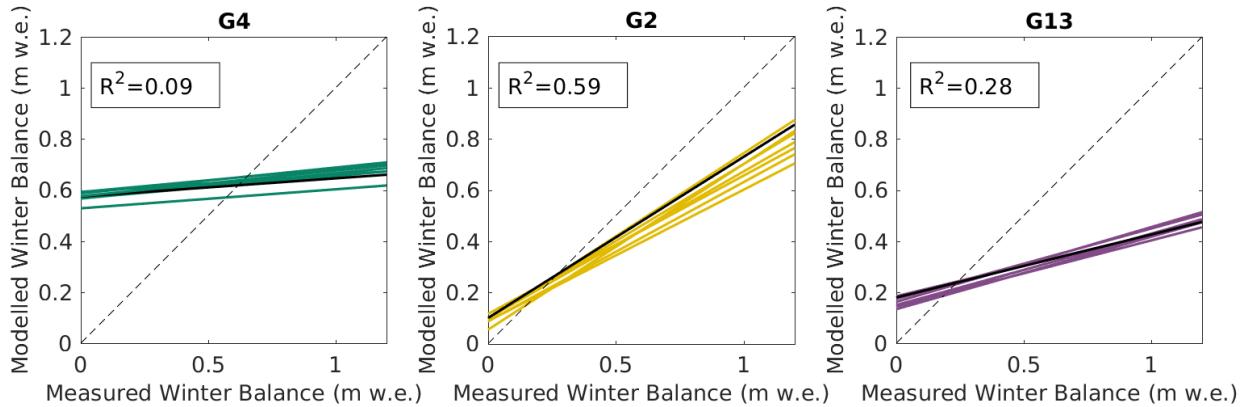


Figure 34: Plot of all linear fits between modelled (BMA) and observed SWE using eight options for calculating density. Mean  $R^2$  value is shown for each sub-plot and a reference 1:1 line is also provided. See Figure 33 for a plot of the data.

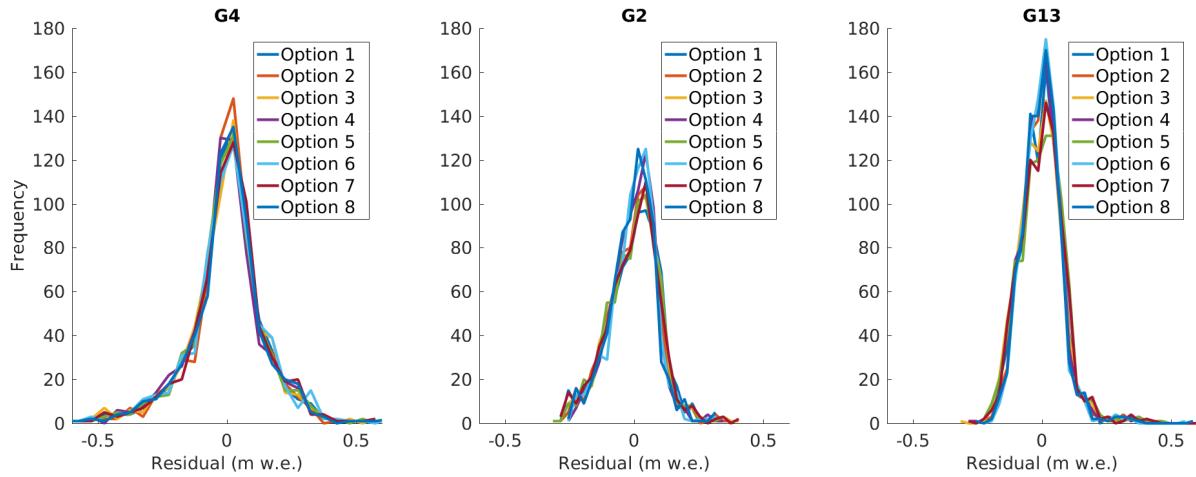


Figure 35: Residuals of SWE predicted using BMA for all options of estimating density.

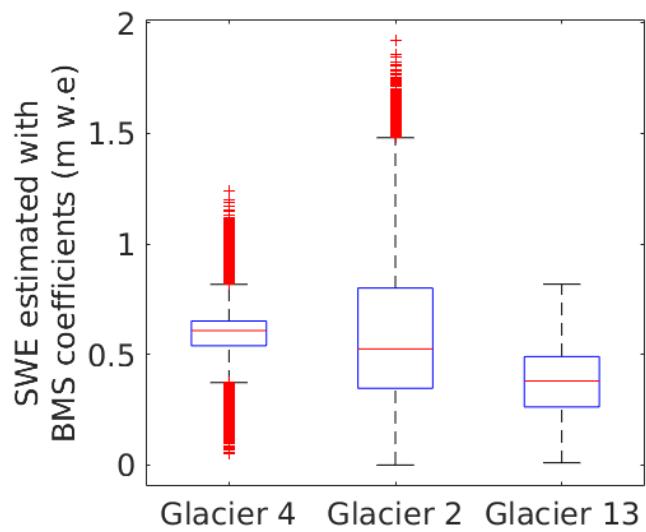


Figure 36: Summary of estimated SWE values found using BMA coefficients. Red line indicates median, blue box shows first quantiles, bars indicate minimum and maximum values (excluding outliers), and red crosses show outliers, which are defined as being outside of the range of 1.5 times the quartiles (approximately  $\pm 2.7\sigma$ ).

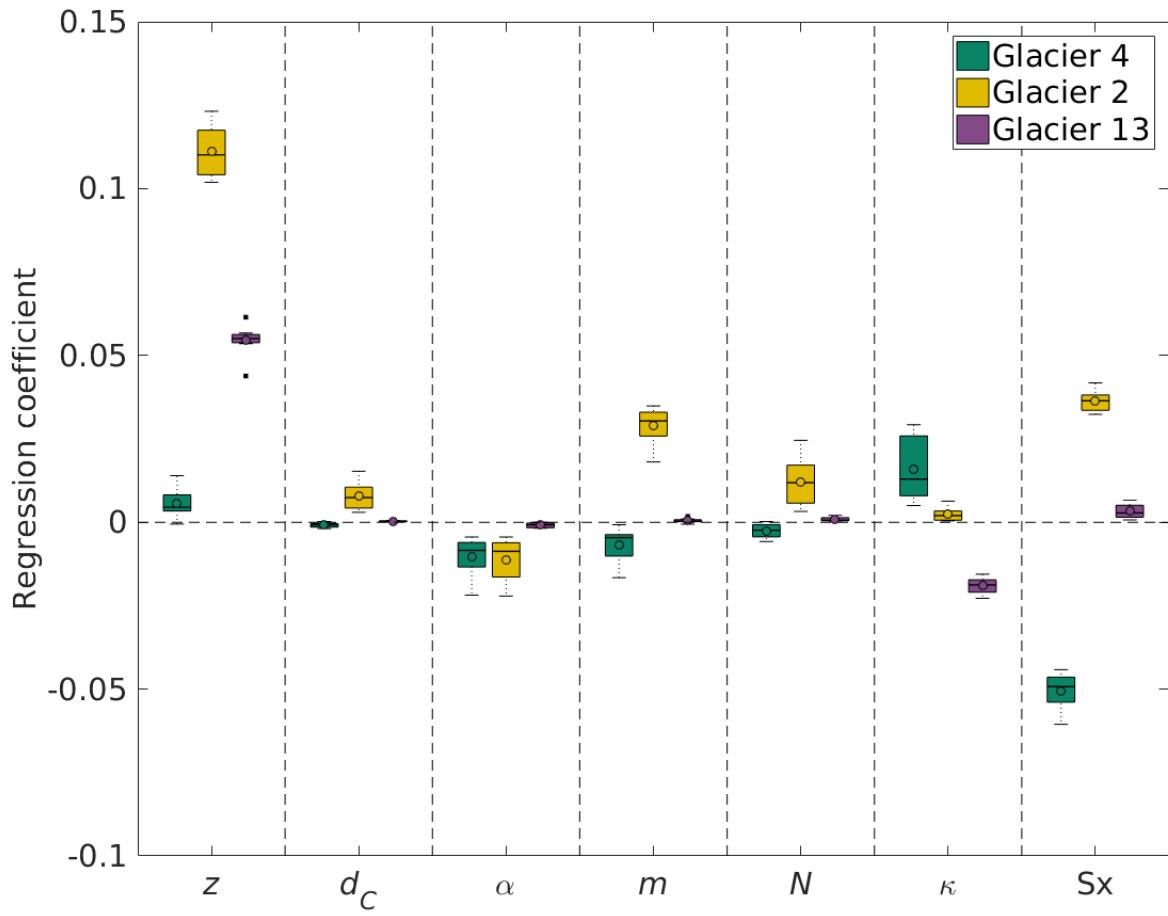


Figure 37: Boxplot showing the range of regressor coefficients explained by each topographic parameter for each option of estimating snow water equivalent (SWE) using BMA. Topographic parameters are distance from centreline ( $d_C$ ), elevation ( $z$ ), aspect ( $\alpha$ ), slope ( $m$ ), northness ( $N$ ), curvature ( $\kappa$ ), and  $Sx$ . Within each box, the mean is shown as a circle, the median as a horizontal line, the interquartile range (IQR) as a coloured box, two times the IQR as dashed lines beyond the box, and outliers as single points.

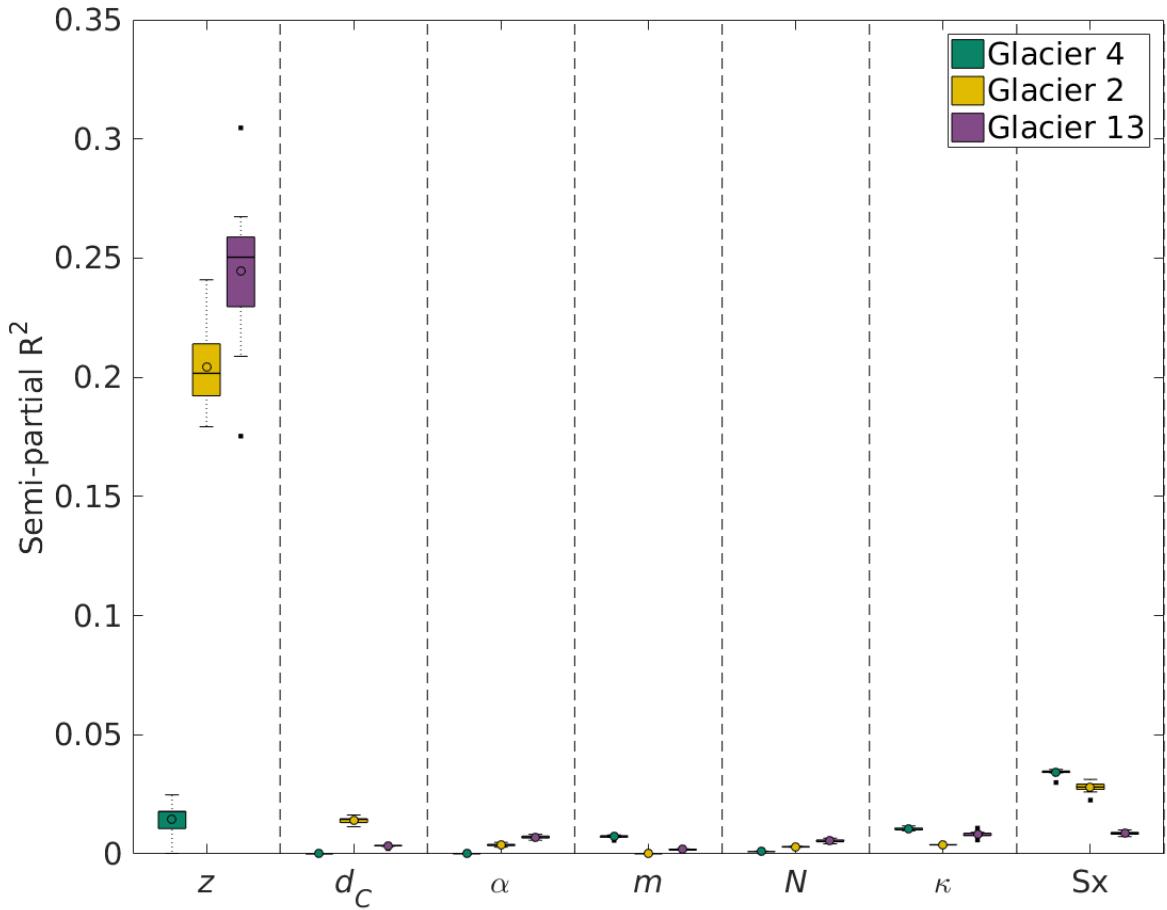


Figure 38: Boxplot showing the range of semi-partial correlation explained by each topographic parameter for each option of estimating snow water equivalent (SWE) using BMA. Topographic parameters are distance from centreline ( $d_C$ ), elevation ( $z$ ), aspect ( $\alpha$ ), slope ( $m$ ), northness ( $N$ ), curvature ( $\kappa$ ), and Sx. Within each box, the mean is shown as a circle, the median as a horizontal line, the interquartile range (IQR) as a coloured box, two times the IQR as dashed lines beyond the box, and outliers as single points.

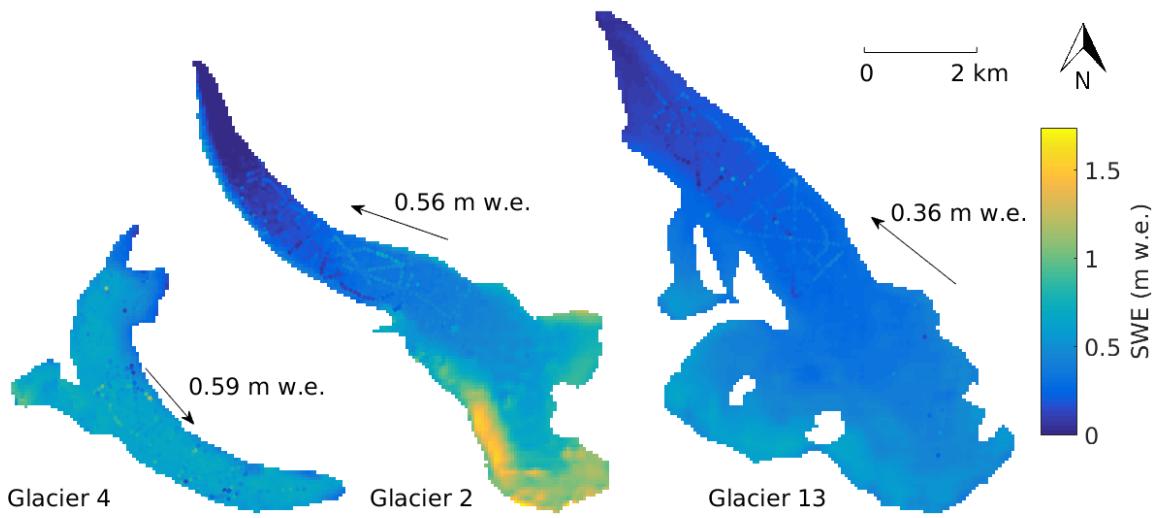


Figure 39: Modelled SWE using coefficients determined using BMA and density interpolated with inverse-distance weighting from snowpits (option 7). Observed SWE values are overlain on the maps.

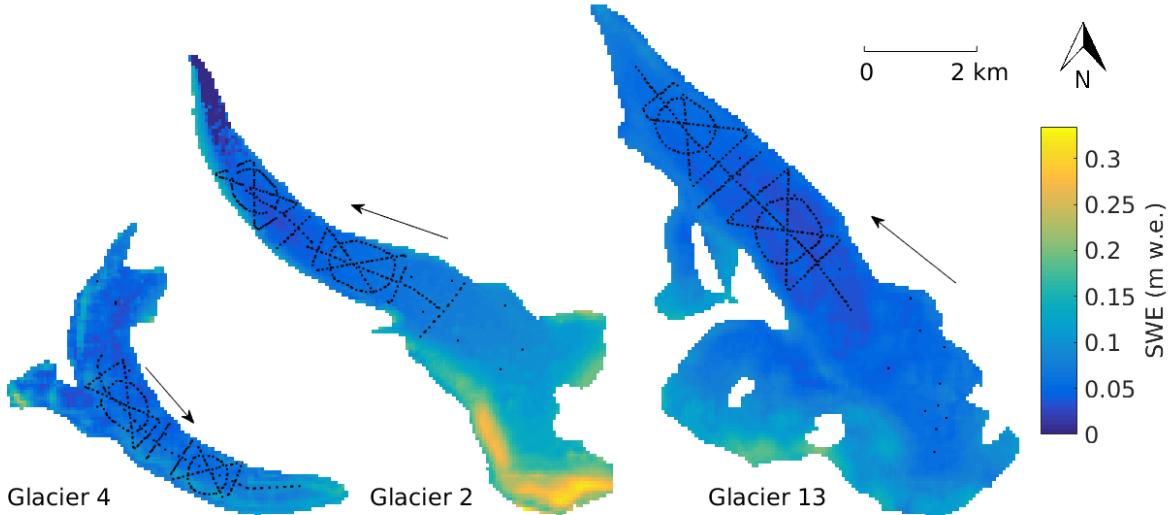


Figure 40: Map of the difference between maximum and minimum SWE values for each DEM cell between all density options using BMA coefficients.

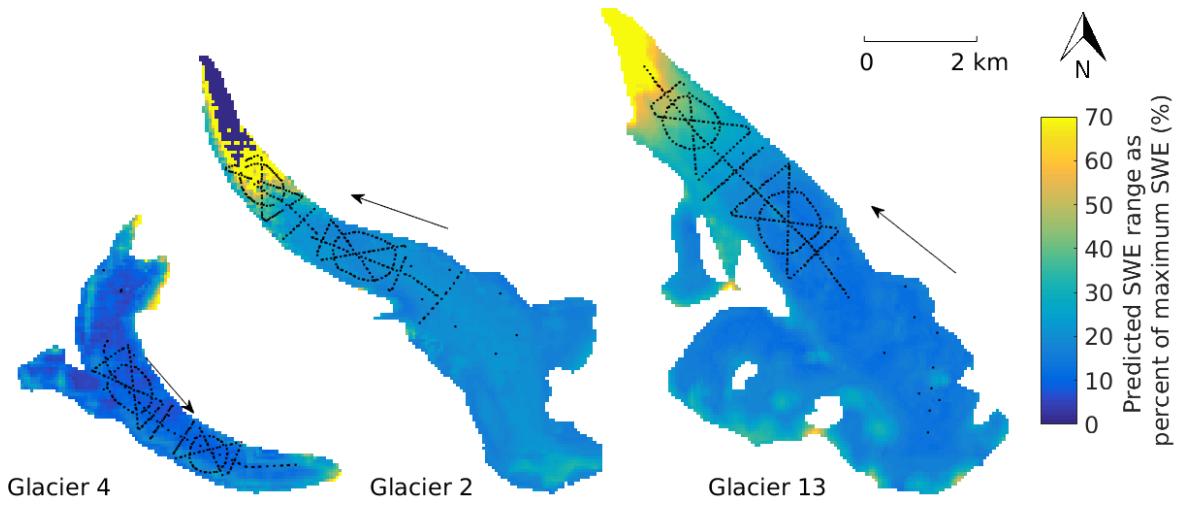


Figure 41: Map of the difference between maximum and minimum SWE values, expressed as a percent of the mean SWE, for each DEM cell between all density options using BMA coefficients. The colours have been scaled to highlight difference in the main part of the glaciers.

## 2.6 MLR and BMA comparison

The range of coefficient values resulting from different choices in density interpolation found using MLR and BMA is similar for all glaciers (Figure 43). The range of coefficient value for all glaciers are not significantly different between MLR and BMA ( $p>0.05$ ). Visually, these ranges always overlap, although the mean and median values can differ between these two methods. The range of unimportant variables is always small, indicating that for all density options these coefficients are small.

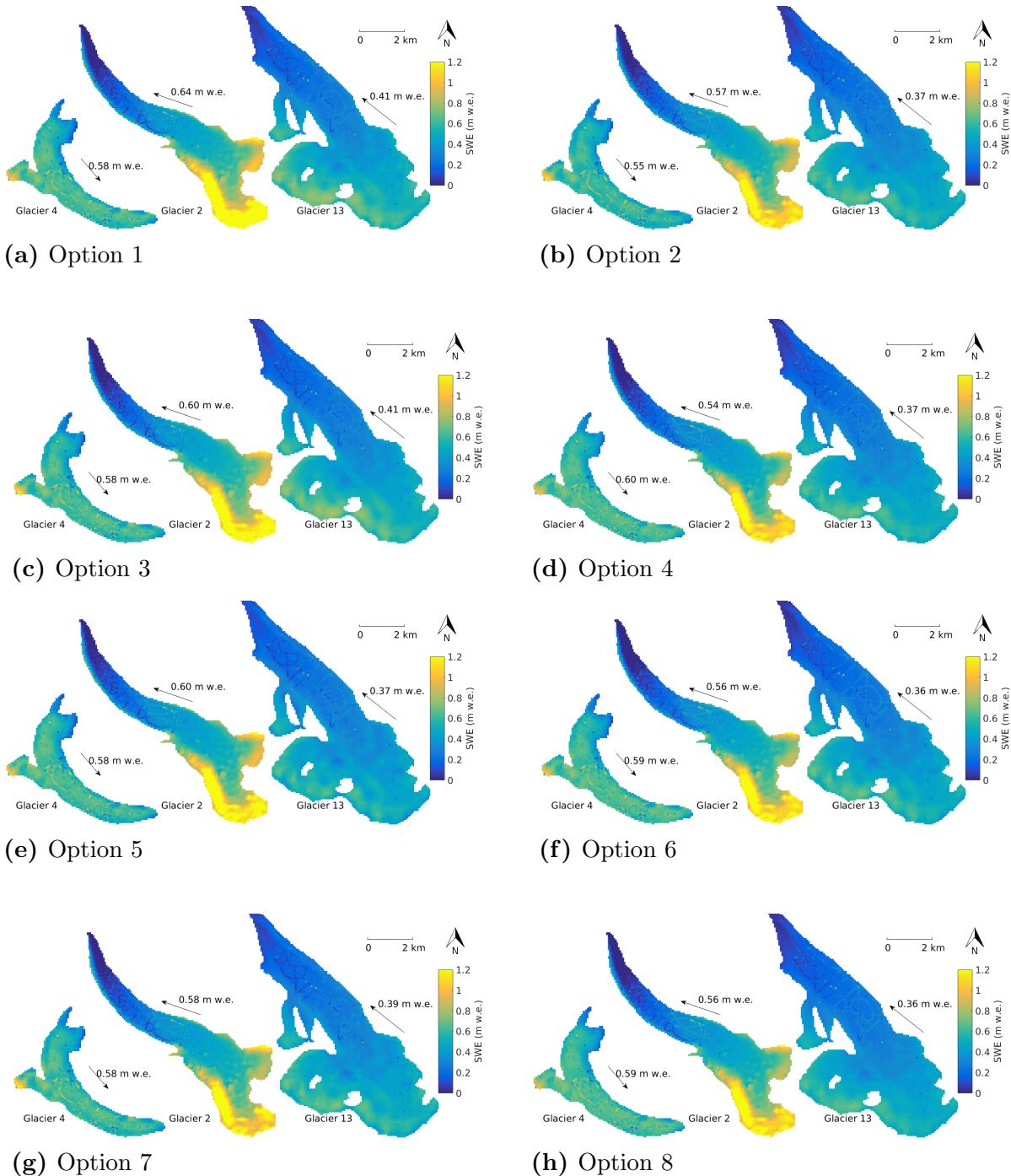


Figure 42: Map of modelled SWE using the BMS coefficient values for all density options. Measured SWE is plotted as overlaid filled circles.

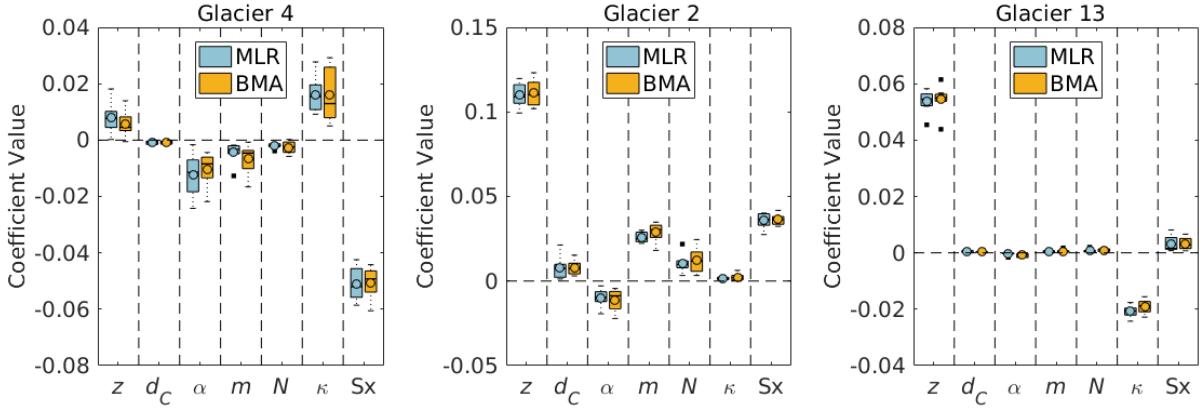


Figure 43: Boxplot showing the range of values of coefficients for each topographic parameter from both MLR and BMA analysis for Glacier 4 (left), Glacier 2 (middle), and Glacier 13 (right). Note the different y axes for the three glaciers. Topographic parameters are distance from centreline ( $d_C$ ), elevation ( $z$ ), aspect ( $\alpha$ ), slope ( $m$ ), northness ( $N$ ), curvature ( $\kappa$ ), and Sx. Within each box, the mean is shown as a circle, the median as a horizontal line, the interquartile range (IQR) as a coloured box, two times the IQR as dashed lines beyond the box, and outliers as single points.

The aspect, curvature, and Sx coefficients for Glacier 4 have the largest range of coefficients. Regardless of the choice of density option, the relative importance of these coefficients is the same (i.e. Sx will always be the most important). For Glacier 4, the choice of density option has a negligible impact on modelled SWE values - the regression fit is so poor that the fit does not change with the choice of density interpolation method (Figure 23 and 34). MLR and BMA appear to have similar ranges of coefficient values, although the mean and median within each regression method differs for Sx and slope, with MLR producing larger coefficients.

The range of coefficient values for Glacier 2 and 13 is generally small for all parameters indicating that the choice of density option does not have a large impact on modelled SWE. The range is especially small for Glacier 13 and most coefficients are close to zero, with the exception of elevation and curvature. This highlights the large semi-partial correlation of elevation on Glacier 13. Further, the relative importance of the coefficients does not change with choice of density option. The greatest range of coefficient values is for elevation, which is also the most significant coefficient. The type of regression model used does not appear to have a large impact on the coefficient values.

For all glaciers, the BMA regression produced a smaller RMSE and a slightly higher  $R^2$ . It is not clear why the BMA regression is better at minimizing RMSE. For some density options, estimated values of SWE can be significantly dif-

Table 8: ANOVA p-values between estimated SWE found using MLR and BMS regression coefficients for various density options. Significance is taken to be  $p < 0.05$ .

Density Option	Glacier 4	Glacier 2	Glacier 13
S1	0.03	0.14	<0.01
F1	<0.01	0.08	<0.01
S2 <sup>47</sup>	<0.01	0.37	<0.01
F2	0.97	<0.01	0.34
S3	<0.01	<0.01	<0.01
F3	0.02	0.84	<0.01

ferent when found using either MLR- or BMA-derived coefficients (Table 8). Although the range of coefficients is not different between BMA and MLR, the resulting SWE values are significantly different.

The range of coefficient values found using the eight density options is similar between MLR and BMA regression methods. Therefore, a single set of minimum and maximum (signed), and mean values for each regression coefficients are used to estimate SWE. These coefficient values are shown in Table 9.

The modelled SWE found using the mean coefficients (Figure 44) shows a relatively uniform SWE distribution on Glacier 4 and a strong elevation dependence on Glacier 2 and 13. This is also seen when using the minimum (Figure 45) and maximum (Figure 46) coefficient values. It is interesting to note that when using the minimum coefficients, there is almost no snow along the edge of the accumulation area on Glacier 4. The difference in SWE using the minimum and maximum coefficients shows that for all three glaciers, the accumulation area is most sensitive to choice of coefficient values.

Table 9: Final regression coefficients. Values are calculated from all coefficients found using MLR and BMA and using all density options. Topographic parameters are distance from centreline ( $d_C$ ), elevation ( $z$ ), aspect ( $\alpha$ ), slope ( $m$ ), northness ( $N$ ), curvature ( $\kappa$ ), and Sx.

	Topographic Parameter	Mean	Min	Max
Glacier 4	$d_C$	-0.0024	-0.0068	0.0000
	$z$	0.0118	0.0000	0.0280
	$\alpha$	0.0014	0.0000	0.0034
	$m$	-0.0162	-0.0290	-0.0047
	$N$	-0.0020	-0.0064	0.0013
	$\kappa_P$	-0.0022	-0.0113	0.0000
	$\kappa_T$	0.0001	-0.0003	0.0005
	Sx	-0.0435	-0.0534	-0.0271
	Intercept	0.6270	0.5749	0.6501
Glacier 2	$d_C$	0.0072	-0.0053	0.0203
	$z$	0.0578	0.0003	0.1171
	$\alpha$	-0.0019	-0.0132	0.0031
	$m$	-0.0002	-0.0247	0.0246
	$N$	0.0012	-0.0064	0.0114
	$\kappa_P$	-0.0005	-0.0040	0.0032
	$\kappa_T$	0.0012	-0.0003	0.0075
	Sx	-0.0039	-0.0534	0.0444
	Intercept	0.4377	0.2245	0.6483
Glacier 13	$d_C$	0.0034	-0.0053	0.0137
	$z$	0.0301	0.0003	0.0528
	$\alpha$	-0.0019	-0.0102	0.0031
	$m$	-0.0097	-0.0247	-0.0002
	$N$	-0.0005	-0.0064	0.0022
	$\kappa_P$	-0.0013	-0.0040	0.0000
	$\kappa_T$	-0.0007	-0.0030	0.0004
	Sx	-0.0187	-0.0534	0.0104
	Intercept	0.4251	0.2014	0.6483

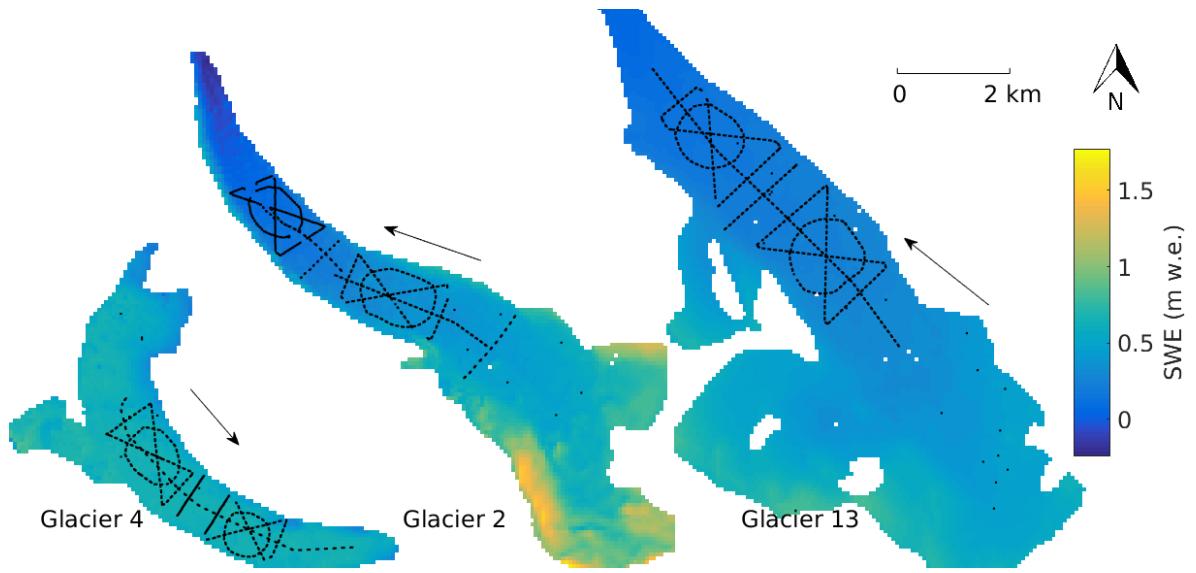


Figure 44: Map of modelled SWE using the mean coefficient values of MLR and BMA using all density options.

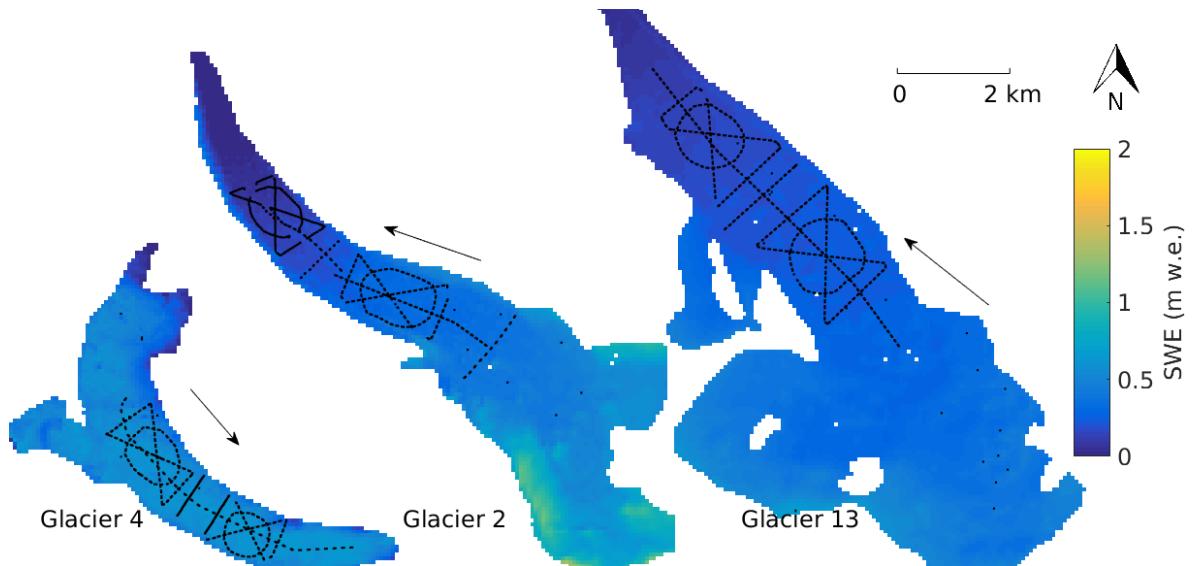


Figure 45: Map of modelled SWE using the minimum coefficient values of MLR and BMA using all density options.

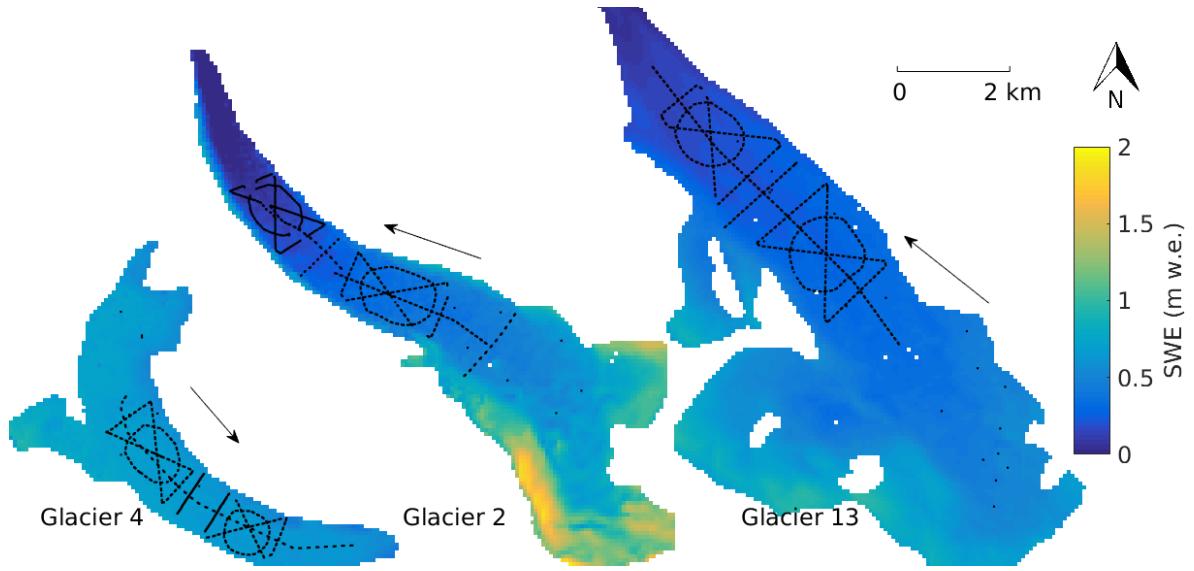


Figure 46: Map of modelled SWE using the maximum coefficient values of MLR and BMA using all density options.

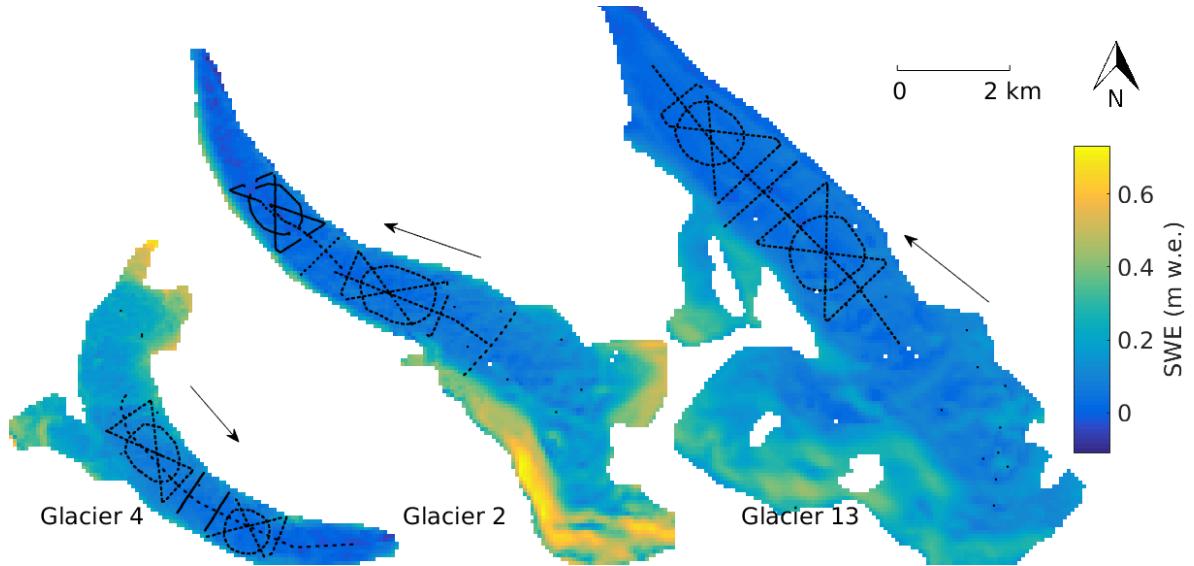


Figure 47: Map of the difference between maximum and minimum SWE values found using the maximum coefficient values of MLR and BMA using all density options.

### 3 Summary

In this portion of the project, the relation between topographic parameters and SWE at the basin scale was examined. First, a suitable DEM is produced by correcting and merging two SPOT-5 DEMs. Then, a series of topographic parameters are calculated for the study glaciers. The sampled topographic parameters are a poor representation

of the full range of parameters on the study glaciers. Major limitations include minimal sampling in the accumulation areas and a lack of sample locations with extreme values of topographic parameters (e.g. high elevation, steep slopes in the accumulation area). This is a major limitation of this study and there is likely a large error induced when extrapolating from relationships between SWE and topographic parameters.

A linear regression between observed SWE (using various density interpolation options) and topographic parameters is done using MLR and BMA. It is found that the choice of density interpolation and the regression method do not have a major impact on the values of topographic regressor coefficients. As a result, a set of mean, minimum, and maximum coefficient values is calculated.

The study glaciers showed varied relationships between SWE and topographic parameters. Glacier 4 has low coefficient values and the modelled SWE is a poor representation of observed SWE. The most significant parameter on Glacier 4 is Sx, which is negatively correlated with SWE. The regression for Glacier 2 indicates that elevation is the strongest predictor of SWE and is able to explain approximately 45% of the observed variability. A strong dependence on elevation is seen on Glacier 13 as well, although less ( $\sim 25\%$ ) of the observed variable could be explained by this parameter.

## 4 Appendix

### 4.1 Topographic parameters from QGIS to Matlab

The value of each topographic parameter at the sampling location is determined in QGIS. The sampling locations are imported to QGIS and the Point Sampling Tool is used to determine the value of the topographic parameter raster cell at each measurement location. The set of parameters that corresponds to each location is then exported to a .csv file and imported into Matlab with the script ‘Import\_Topo.m’. Note that selection of Sx values is completed first, as described in Section 1.2. The Sx map made with the combination of  $d_{max}$  and azimuth values that produces Sx values most strongly correlated with SWE is exported.

After importing values of sampled topographic parameters to Matlab, the sets of parameter values ( $x_p$ ) are standardized ( $x_s$ ) by  $x_s = \frac{x_p - \mu}{\sigma}$ . The resulting structure is called `topo_sampled` and is used in the regression to be able to compare the explanatory power of each parameter (Section 2.4 and 2.5). A non-standardized copy of the topographic parameters at the sampling locations is also kept for plotting purposes and is called `topo_sampled_ns`. Both structures are organized as a vector of values corresponding to the vectors in the `SWE` variable.

The topographic parameter rasters are also exported from QGIS as a .csv file and then imported into Matlab with the script ‘Import\_Topo.m’. The values are stored in the structure `topo_full`, where each cell corresponds to one DEM cell. Cells outside of the glacier outline have no value (NaN). The raster values were standardized using the mean and standard deviation of the sampled topographic parameters so that this set of full topographic parameters could be used for modelling SWE using the regression. A copy

of the non-standardized values was stored as `topo_full_ns`.

## 4.2 MLR software

The MLR was completed with the following steps (executed using the function '`MLRcalval.m`'):

1. The topographic parameters are imported to Matlab using the script 'Import\_Topo.m'.
2. The `topo_sampled` structure for one glacier as well as the `SWE` structure is passed into the function.
3. A set of initializations is completed. This includes 1) creating a logical matrix to choose all linear combinations of topographic parameters, 2) selecting the number of runs, 3) creating a matrix of random numbers for selecting data points in the cross validation procedure, 4) initializing matrices, and 5) converting the input structure to a table.
4. For each linear combination of topographic parameters, 1000 runs of a cross validation MLR are then executed. Two-thirds of the total data is randomly selected (Kohavi and others, 1995) to use for calculating the regression (using the function `regress()` which is a basic regression function with fast execution). The MLR equation is used to predict the SWE using the remaining one-third of the topographic parameters. The root-mean-squared error (RMSE) between the predicted and observed SWE values is then calculated and the set of regression coefficients that produce the lowest RMSE are then chosen for that combination of topographic parameters. The function `fitlm()` is then used to calculate the MLR from the set of data that gave the lowest RMSE. This function is slower but calculates a number of additional values that characterize the fit of the model. One of these values is the Bayesian information criterion (BIC), which allows for model selection among a finite set of models (Burnham and Anderson, 2004). The BIC from the best model for each combination of topographic parameters is saved.
5. A weighted sum of all models found using linear combinations of topographic parameters was then found. The BIC values for each model ( $BIC_i$ ) were used to determine the normalized weight of each model ( $w_i$ ) relative to the best model (lowest BIC value  $BIC_{min}$ ) according to Eq. 11.
6. The percent variance ( $var\%$ ) explained by each parameter was calculated using the equation  $var\% = \frac{SSr}{SSt} \times 100$ , where  $SSr$  is the sum of squares of the residual (fitted topographic parameters) and  $SSt$  is the total sum of squares (SWE observations). The final coefficients and the percent variance explained by each one can be found in the `coeffs_final` table within the function and in the `mlr` structure when run for all density options and glaciers in the main script 'TopoRegression.m'.
7. The residuals of the fit have also been calculated as a separate variable that can be returned when the function is called. The residual is calculated as the difference between the estimated SWE value and observed value.

### 4.3 BMA software

The BMA process was implemented in R, using the Bayesian model statistics (BMS) package developed by Zeugner and Feldkircher (2015). The function `BMS.R()` computes the posterior distribution mean value of all  $\beta$  coefficients for topographic parameters as well as the percent variance explained by each parameter using the following steps:

1. A portion of the data (2/3 as suggested by Kohavi and others (1995)) is randomly chosen as the calibration set and saved as a .mat file.
2. The Bayesian model statistics (BMS) package developed by Zeugner and Feldkircher (2015) is run in R (called through an operating system command in Matlab)
  - (a) The R script imports the .mat file with SWE and topographic parameter values. It then creates a data frame with the SWE values as the first column and the topographic parameters as the remaining columns.
  - (b) The BMS package is used to complete BMA for the imported values. A uniform model prior was chosen. The mean coefficient value, coefficient standard deviation, PIP, and the posterior probability of a positive coefficient (how probable it is for the sign of the coefficient to be positive) were computed.
  - (c) The coefficients are saved as a .mat file.
3. Regressor coefficient values calculated in R are loaded into Matlab and a data table is created with the coefficients.
4. The remaining portion of the data (1/3) are then used to calculate a modelled value of SWE at those locations. These are compared to the observed SWE values and a RMSE value is determined.
5. The above steps are completed 1000 times and the coefficients associated with the lowest RMSE are chosen.
6. Percent variance explained by each parameter is then calculated in the same way as for the MLR (see Section 4.2).
7. The final table of values includes the coefficients and percent variance explained for all topographic parameters associated with the lowest RMSE. It also includes the intercept and the actual RMSE value. This table is returned from the function.
8. The residuals of the best BMA fit are also calculated and can be return from the function.

## References

- Basist A, Bell GD and Meentemeyer V (1994) Statistical relationships between topography and precipitation patterns. *Journal of climate*, **7**(9), 1305–1315 (doi: 10.1175/1520-0442(1994)007<1305:SRBTAP>2.0.CO;2)
- Berthier E, Arnaud Y, Kumar R, Ahmad S, Wagnon P and Chevallier P (2007) Remote sensing estimates of glacier mass balances in the Himachal Pradesh (Western Himalaya, India). *Remote Sensing of Environment*, **108**(3), 327–338
- Berthier E, Schiefer E, Clarke GK, Menounos B and Rémy F (2010) Contribution of Alaskan glaciers to sea-level rise derived from satellite imagery. *Nature Geoscience*, **3**(2), 92–95
- Bring J (1996) A geometric approach to compare variables in a regression model. *The American Statistician*, **50**(1), 57–62 (doi: 10.1080/00031305.1996.10473543)
- 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)
- Cohen J, Cohen P, West SG and Aiken LS (2013) *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge
- Darlington RB (1968) Multiple regression in psychological research and practice. *Psychological bulletin*, **69**(3), 161
- Davis JC and Sampson RJ (1986) *Statistics and data analysis in geology*, volume 646. Wiley New York et al.
- Grömping U (2009) Variable Importance Assessment in Regression: Linear Regression versus Random Forest. *The American Statistician*, **63**(4), 308–319 (doi: 10.1198/tast.2009.08199)
- Grömping U (2015) Variable importance in regression models. *Wiley Interdisciplinary Reviews: Computational Statistics*, **7**(2), 137–152 (doi: 10.1002/wics.1346)
- 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

- 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)
- 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)
- Neteler M and Mitášová H (2008) *Open source GIS: a GRASS GIS approach*, volume 773. Springer Science & Business Media, 3 edition
- 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)
- Raftery AE, Gneiting T, Balabdaoui F and Polakowski M (2005) Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *Monthly Weather Review*, **133**(5), 1155–1174 (doi: 10.1175/MWR2906.1)
- 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)
- 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)
- 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
- Woo MK and Marsh P (1978) Analysis of Error in the Determination of Snow Storage for Small High Arctic Basins. *Journal of Applied Meteorology*, **17**(10), 1537–1541 (doi: 10.1175/1520-0450(1978)017<1537:AOEITD>2.0.CO;2)
- 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