Optimizing snow survey design for estimating winter balance of alpine glaciers

Alexandra PULWICKI,[1] Gwenn E. FLOWERS,[1]

[1] *Department of Earth Sciences, Faculty of Science, Simon Fraser University, Burnaby, BC, Canada*
Correspondence: Alexandra Pulwicki <apulwick@sfu.ca>

ABSTRACT.

INTRODUCTION

Estimates of basin-wide seasonal snow accumulation are critical for monitoring glacier mass balance and for predicting the availability and timing of surface runoff, especially in mountainous regions. The net accumulation and ablation of snow on a glacier over a winter season is known as the winter surface mass balance, or "winter balance" (WB) and is typically reported in meters of water equivalent (m w.e.) (Cogley and others, 2011). Winter balance accounts for half of the seasonally resolved mass balance, initializes ablation conditions and affects energy and mass exchange between the land and atmosphere (e.g. Hock, 2005; Réveillet and others, 2016).

Snow distribution is spatially variable so properties, such as snow depth, must be measured over an extensive area (e.g. Kinar and Pomeroy, 2015). In addition, the period of peak accumulation is short so snow measurement must be completed quickly and efficiently. As a result, representative measurements of snow depth are nearly impossible to obtain. Snow surveys must therefore be optimized in the extent and spacing of snow measurement locations, especially when labour-intensive methods like snow probing are used.

Optimal sampling schemes for snow probing are central to accurately estimating snow distribution and mass balance from *in situ* measurements. Measuring snow depth and travelling between measurement locations is both time consuming and can disturb the snow so care must be taken to choose a sampling scheme that avoids bias, allows for the greatest variability to be measured and minimizes distance travelled (Shea and Jamieson, 2010). There are a number of different designs that have been employed to obtain point measurements, including pure random (e.g. Elder and others, 1991), linear random (e.g. Shea and Jamieson, 2010), nested (e.g. Schweizer and others, 2008), gridded random (e.g. Bellaire and Schweizer, 2008; Elder and others, 2009; Bellaire and Schweizer, 2011) and gridded (e.g. Molotch and Bales, 2005; Kronholm and Birkeland, 2007; López-Moreno and others, 2011). Sampling designs that incorporate randomness are favourable because they limit sampling bias by varying sample spacing and direction. However, they are less efficient than sampling designs that incorporate grids. Grid-style sampling designs minimize travel distance but measurements are biased by regularly spaced intervals and linear orientations, which could result in an under representation of the snow variability (Kronholm and Birkeland, 2007).

Snow surveys on glaciers are conducted to estimate winter balance and multi-year sampling programs are often established to monitor changes in winter balance with time. Optimization of sampling design is rarely investigated because the locations of snow depth measurements are often dictated by field resources. An optimized sampling design requires (1) a sampling pattern that captures spatial variability and minimizes travel distance and (2) knowledge of the minimum number of measurement locations needed to accurately estimate WB. There are few studies that investigate the number of measurement locations needed to effectively sample WB distribution (c.f. Fountain and Vecchia, 1999; Walmsley, 2015). The sampling pattern used for most winter balance programs does not included randomness and measurements are typically collected along the glacier midline (e.g. **?**) to capture changes in snow depth due to orographic effects (e.g. Grünewald and others, 2014). However, midline transects are known to underestimate winter balance so transverse transects are often added to improve the reliability of the sampling scheme (e.g. Walmsley, 2015). An hourglass with inscribed circle (personal communication from C. Parr, 2016) is an alternative sampling design that is attractive because it is able to capture changes in WB with elevation but is not biased along the midline and is easy to travel. To our knowledge, no study has yet compared the ability of these sampling designs to capture spatial variability in WB.

The goal of our work is to provide insight into optimization of WB sampling design by investigating various sampling patterns and number of measurement locations. We use both synthetic- and real data to examine the effect of sampling design on estimates of glacier-wide WB for three alpine glaciers.

STUDY SITE

We investigate sampling design for winter balance surveys on three unnamed glaciers in the Donjek Range of the St. Elias Mountains, Yukon, Canada. The Donjek Range is located on the continental side of the St. Elias Mountains, which rise sharply from the Pacific Ocean and create a strong climatic gradient. Monitoring of snow distribution and glacier mass balance in the St. Elias Mountains began in the 1950s and 60s with a series of research programs, including Project "Snow Cornice" and the Icefield Ranges
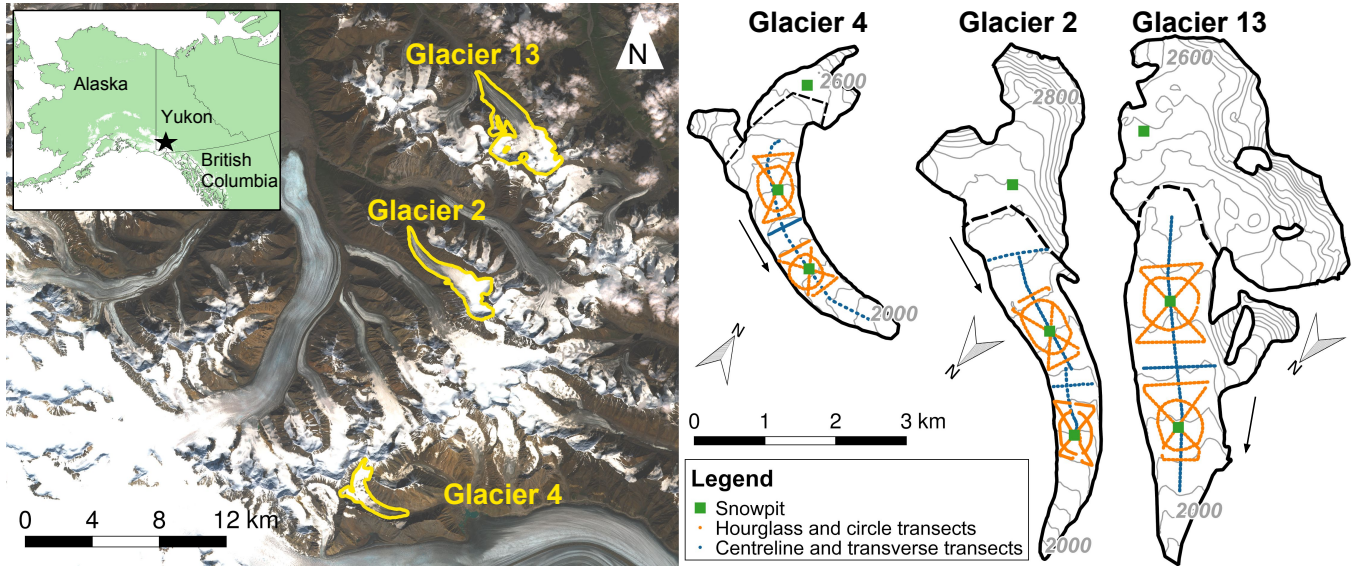
Fig. 1: Study area location and sampling design for Glaciers 4, 2 and 13. (Left) Study region in the Donjek Range of the St. Elias Mountains of Yukon, Canada (inset). Imagery from Landsat8 (29 August 2013, data available from the U.S. Geological Survey). (Right) Details of the snow-survey sampling design, with centreline and transverse transects (blue dots), hourglass and circle designs (orange dots) and locations of snow density measurements (green squares). Arrows indicate ice-flow directions. Approximate location of ELA on each glacier is shown as a black dashed line. Contour lines in increments of 50 m are shown in grey.

Research Project (Wood, 1948; Danby and others, 2003). More recent studies have focused on glaciological studies of selected alpine glaciers (e.g. Clarke, 2014) as well as estimates of regional glacier mass balance and dynamics (e.g. Arendt and others, 2008; Burgess and others, 2013; Waechter and others, 2015).

Glacier 4, Glacier 2 and Glacier 13 (labelling adopted from Crompton and Flowers (2016)) are small alpine glaciers (3.8–$12.6 \, \text{km}^2$) with simple geometries. Elevation of these glaciers ranges from 1900 to 3100 m a.s.l. and ELAs are located at $\sim$2500 m. The glaciers are generally oriented southeast-northwest in valleys with steep walls. We suspect that the glaciers are polythermal, based on a detailed study of Glacier 2 (Wilson and others, 2013) and related theoretical modelling (Wilson and Flowers, 2013). A detailed analysis of estimating winter balance on these three glaciers is presented by Pulwicki and others (2017).

METHODS

We aim to determine the optimal sampling pattern and number of measurement locations needed to obtain accurate estimates of WB for small alpine glaciers. We estimate distributed WB using both synthetic- and real values of gridcell WB. Real data are subsets of all gridcell values of WB, which is derived from direct measurements of snow depth and density, chosen based on different sampling designs. A linear regression is used to interpolate gridcell values resulting in an estimated WB. Synthetic data are generated by sampling a synthetic distribution of WB, which is based on field measurements, using various sampling designs. Uncertainty in gridcell WB

is incorporated by introducing noise into the synthetic distribution of WB. A linear regression of synthetic data on topographic parameters is then used to estimate WB. A set of estimated WB is generated by repeating this process with different levels of noise. We then directly compare the estimated WB and the synthetic distribution of WB. Three study glaciers, with differing spatial patterns of WB, are used to determine the applicability of our conclusions between glaciers.

Here we describe the process of collecting WB data in the field and interpolating/extrapolating to obtain the synthetic distribution of WB on the three study glaciers. Subsets of the real data are used to investigate the effect of various sampling designs on the error associated with estimating gridcell-scale values of WB. Then, we detail the methods used to obtain synthetic data and describe the various sampling patterns and number of measurement locations investigated in this study. A comparison of the synthetic distribution of WB and WB estimated using different sampling designs with synthetic data is presented.

Field data

Point-scale values of WB are found by obtaining direct measurements of snow depth and density (Figure 1). Snow depth was measured using a 3.2 m graduated aluminium avalanche probe. Measurement locations followed linear and curvilinear transects, which were similar between study glaciers, with a sample spacing of 10–60 m. Spacing was constrained by protocols for safe glacier travel. Each observer made 3–4 depth measurements within $\sim$1 m at each transect measurement location. We restricted snow-depth sampling to the ablation area, where the clear distinction
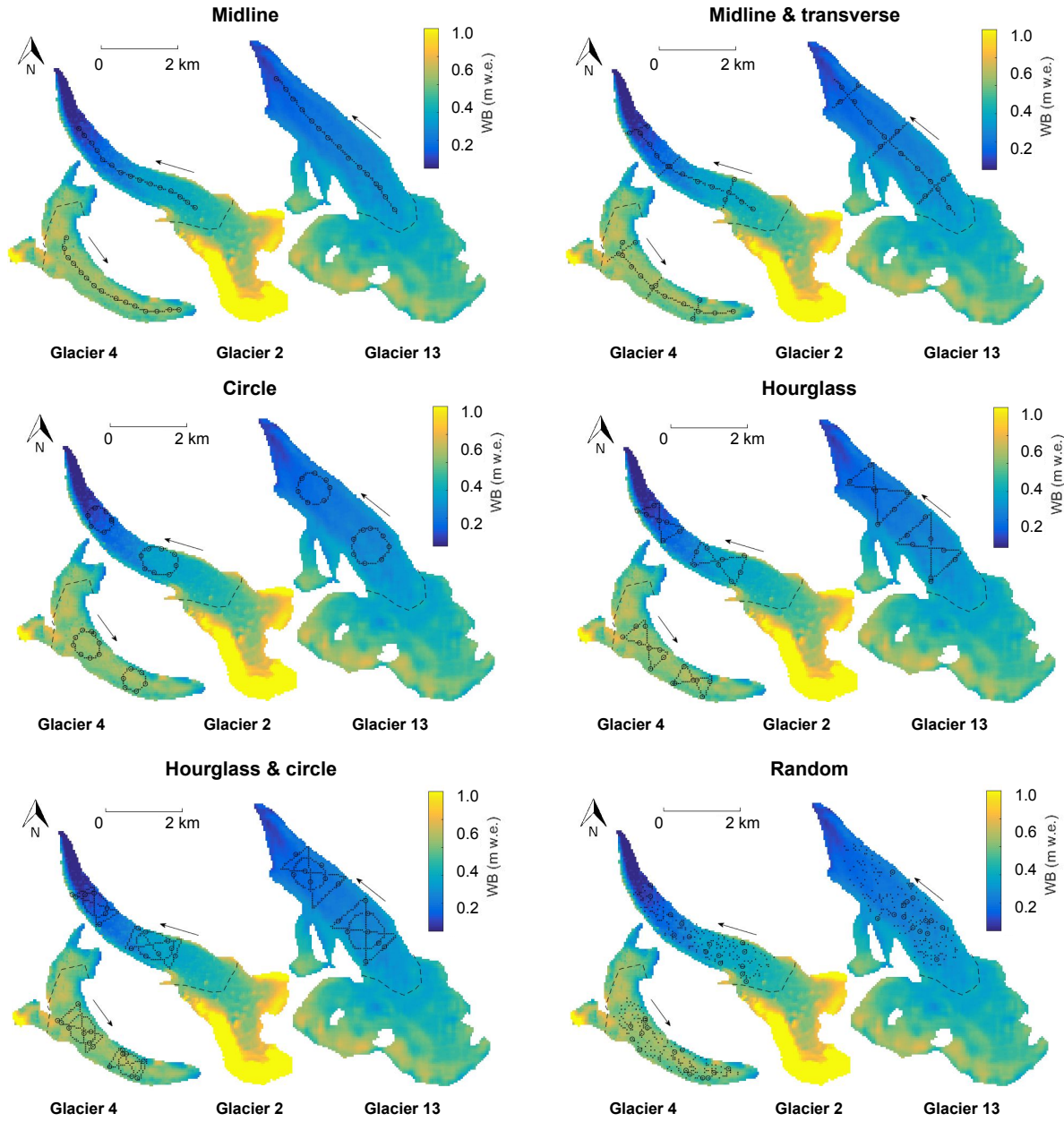
Fig. 2: Synthetic sampling designs on three study glacier. All synthetic observation measurement locations (small, black dots) and a subset including 15 measurement locations (open circles) is shown for each sampling pattern. The random sampling pattern shows a total of 200 sampling locations. Sampling designs are overlain on the prescribed winter balance (WB).

Table 1: Minimum travel distance (km) required to fulfil each sampling pattern. Distance for random design ($n = 40$) is the sum of distances between closest measurement locations and a range is provided for an array of different random locations.

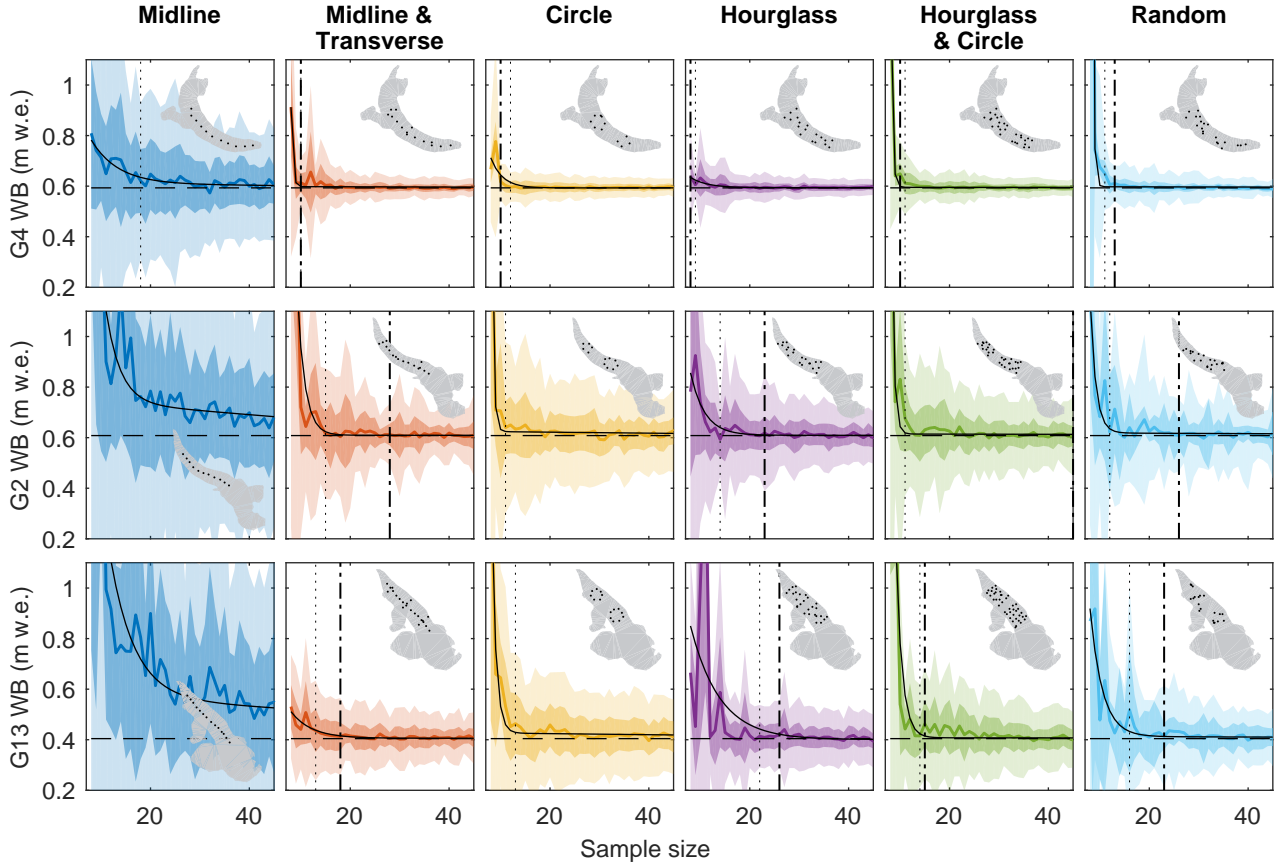|	Midline	Midline & Transect	Circle	Hourglass	Hourglass & Circle	Random
Glacier 4	6.3	8.3	4.8	6.9	11.1	5.5–10.2
Glacier 2	4.3	8.6	5.7	8.4	12.5	6.6–11.8
Glacier 13	8.0	11.0	7.0	10.6	16.8	8.3–14.2

Fig. 3: Glacier-wide winter balance (WB) estimated with various sampling designs for Glacier 4 (top row), Glacier 2 (middle row) and Glacier 13 (bottom row). Sampling patterns are shown in columns. Insets show an example of a sampling design corresponding to the panel. Mean WB from 100 runs of a linear regression of synthetic WB on topographic parameters for various patterns and sample sizes (solid coloured line) with second-order exponential fit (solid black line). Standard deviation in glacier-wide WB due to low noise (dark shading) and high noise (light shading). Glacier-wide WB of validation data (black dashed line). Sample size when fitted mean is within 0.05 m w.e. of true mean (black dotted line) and when standard deviation due to high noise is less than 0.15 m w.e. (black dotted-dashed line) are shown.

between snow and ice ensured that only snow from the current accumulation season was measured. In total, we collected more than 9000 snow-depth measurements throughout the study area. Snow density was measured using a wedge cutter in three snow pits that spanned a large portion of the elevation on each glacier. A mean density was then calculated for each glacier and this value was used to convert snow depth at all measurement locations to values of point-scale WB. Mean density was $348\pm13\,\mathrm{kg\,m^{-3}}$ on Glacier 4, $333\pm26\,\mathrm{kg\,m^{-3}}$ on Glacier 2 and $349\pm38\,\mathrm{kg\,m^{-3}}$ on Glacier 13. All point-scale values of WB located within a common DEM gridcell ($40\times40\,\mathrm{m}$) were averaged to obtain values of gridcell WB.

Real data

Sampling designs are derived from all measurement locations for each study glacier (Figure 1). We investigate numerous sampling designs that are unique combinations of six different sampling patterns and the number of measurement locations. Midline and midline & transverse

transects are the most common survey designs used in WB studies (e.g. Kaser and others, 2002; Machguth and others, 2006). The midline survey aims to capture changes in WB with elevation and transverse transects provide observation of lateral variations in WB. Hourglass and inscribed circle allow for sampling in multiple directions and are easy to travel (personal communication from C. Parr, 2016). We use hourglass and circle patterns separately and combined. Finally, a random pattern of measurement locations is obtained by selecting random gridcells from the full data set. For all sampling patterns we estimate WB using n measurement locations, where n ranges from a minimum of eight (constrained by using all seven topographic parameters for interpolation) to a maximum determined by the number of gridcells within a sampling pattern (ranges between 57–228 gridcells). The measurement locations are evenly distributed within the sampling pattern.

A linear regression is then used to interpolate the real values of WB for each glacier. Real data is regressed on
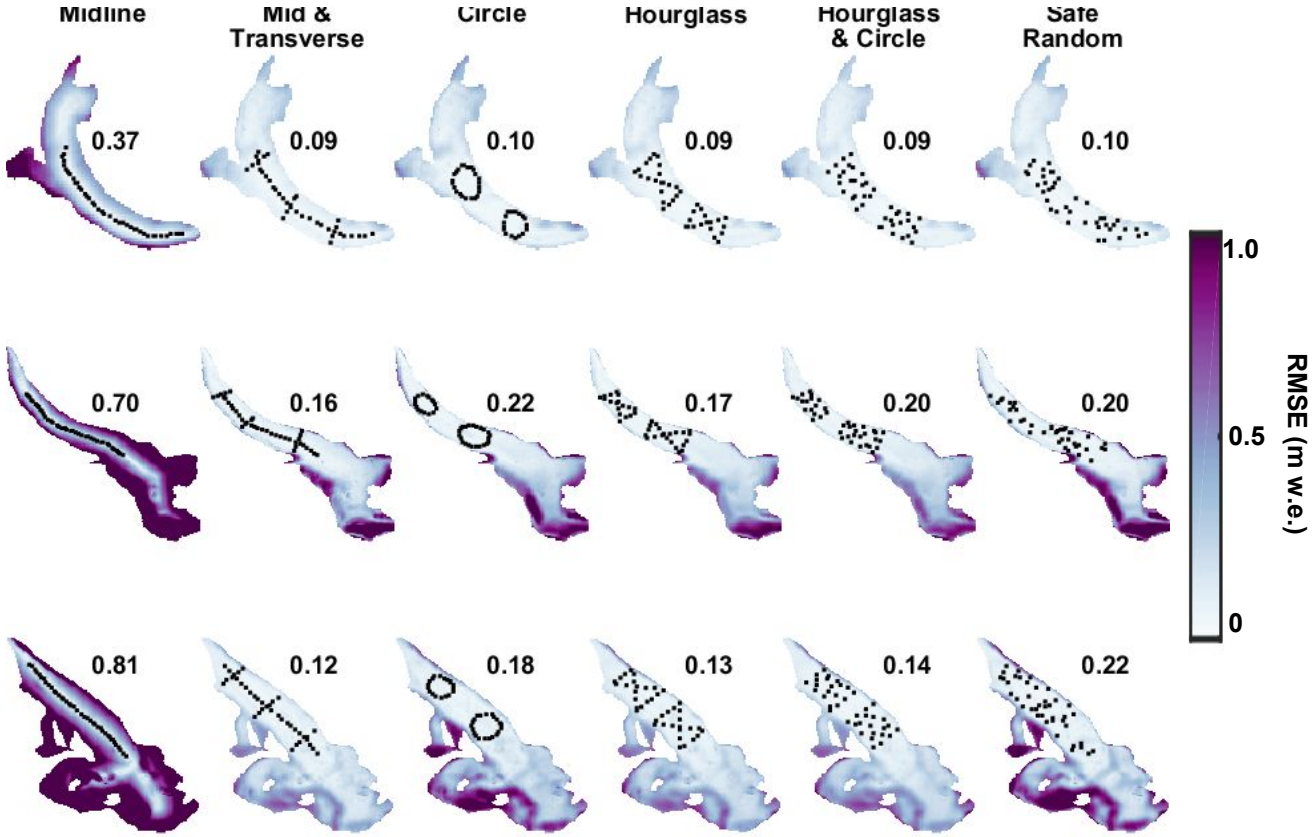
Fig. 4: Root-mean-squared-error (RMSE) between validation WB and WB estimated with various sample patterns ($n = 40$) for 100 runs of high noise linear regression. Sampling patterns are shown in columns for Glacier 4 (top row), Glacier 2 (middle row) and Glacier 13 (bottom row). RMSE for the entire glacier is displayed. Synthetic sampling locations are shown (black dots).

elevation, slope, aspect, distance from glacier centreline, "northness", curvature and a wind redistribution parameter. The linear regression calculates regression coefficients that minimize the sum of squares of the vertical deviations of each datum from the regression line (Davis and Sampson, 1986). The resulting regression coefficients are then applied to the topographic parameters associated with each gridcell to obtain an estimate of distributed WB. Results are presented as a RMSE value, which is calculated by taking the square root of the mean difference between WB estimated using various sampling designs and real data at all measurement locations.

Synthetic data

We obtain synthetic data using various sampling designs from the synthetic distribution of WB to test which sampling design best estimates WB. Gridcell values of WB that fall within each sampling design are extracted from the synthetic distribution of WB distribution and then used to estimate a new WB distribution. The WB distribution derived from various sampling designs is directly compared to the synthetic distribution of WB distribution to estimate spatially resolved error.

The synthetic distribution of WB is obtained by computing a linear regression of gridcell values of WB on topographic parameters for each glacier (Pulwicki and others, 2017). Topographic parameters are derived from a SPIRIT SPOT-5 DEM (Korona and others, 2009) and include commonly applied topographic parameters (e.g. McGrath and others, 2015) such as elevation, slope, aspect, distance from glacier centreline, "northness", curvature and a wind redistribution parameter. A linear regression, along with cross-validation and model averaging, is used to obtain a set of regression coefficients for the standardized topographic parameters on each glacier (Pulwicki and others, 2017). Distributed WB is found by multiplying fitted regression coefficients by corresponding topographic parameters for each gridcell. The distributed WB calculated using all available WB data is hereafter referred to as the synthetic distribution of WB.

Synthetic data are obtained for all of the sampling designs presented above. Midline, midline & transects, circle, hourglass, hourglass & circle and random sampling patterns are used (Figure 2) and values of n from eight to the maximum number of observation with each sampling pattern are investigated. All sampling patterns are
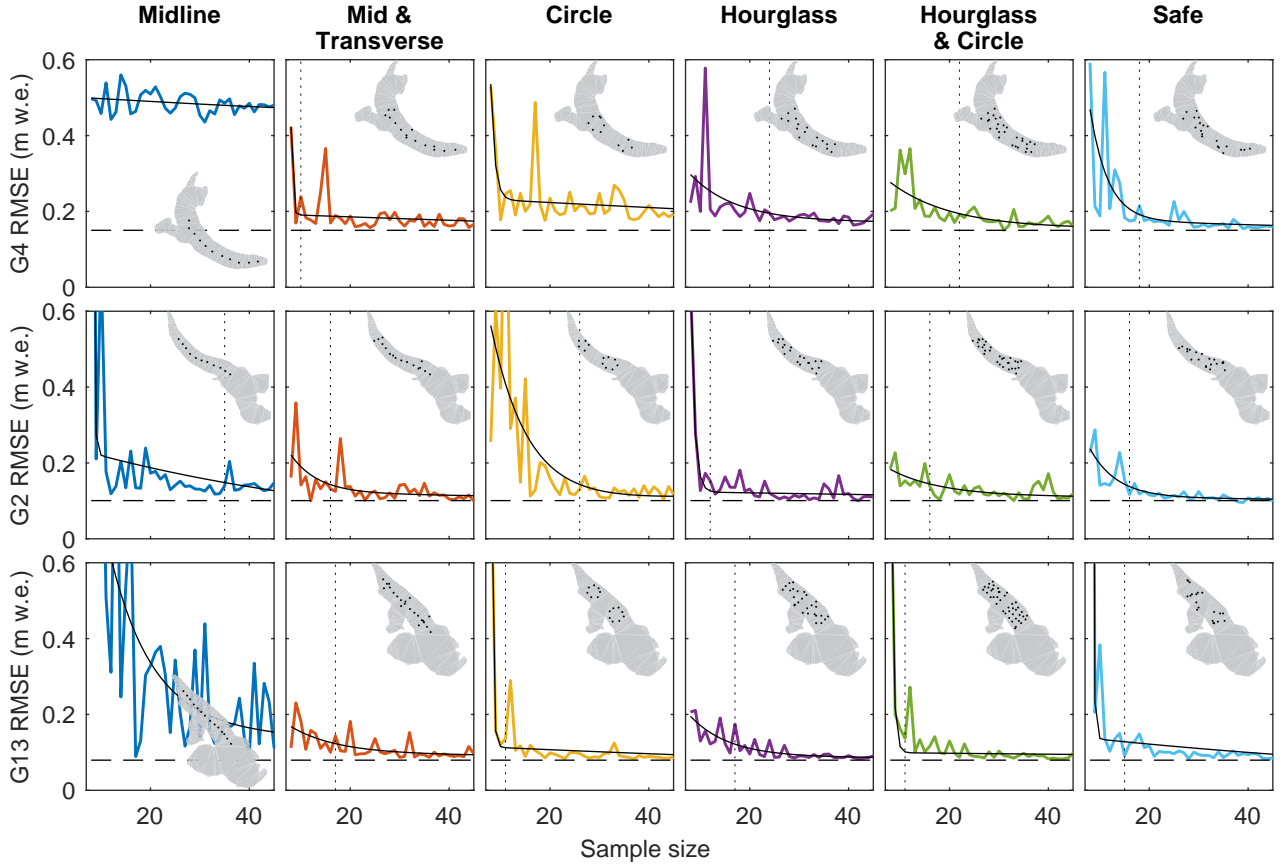
Fig. 5: Root-mean-squared-error (RMSE) between WB estimated with various sampling designs of real data and field values of WB at all measurement locations for Glacier 4 (top row), Glacier 2 (middle row) and Glacier 13 (bottom row). Sampling patterns are shown in columns. Insets show an example of a sampling design corresponding to the panel. Mean RMSE from 100 runs of a linear regression of real WB on topographic parameters for various patterns and sample sizes (solid coloured line) with second-order exponential fit (solid black line). Standard deviation in RMSE due to low noise (dark shading) and high noise (light shading). RMSE of estimated WB when all measurement locations are used in the linear regression (black dashed line). Sample size when fitted RMSE mean is within 0.05 m w.e. of true RMSE mean (black dotted line) and when standard deviation due to high noise is less than 0.01 m w.e. (black dotted-dashed line) are shown.

restricted to the ablation area, where terrain is accessible and direct measurements of snow depth are easy to obtain. The random pattern selects measurements locations within the area that encompasses the original sampling designs.

To simulate the process of measuring WB, we first obtain WB values from the synthetic distribution of WB distribution at selected measurement locations for each sampling pattern. Then, we add a low or high level of noise to the WB data. Low noise is defined by a normal distribution that is centred at zero and has a standard deviation equal to the mean standard deviation of WB data from a series of high-density gridcell-scale surveys on each glacier (Pulwicki and others, 2017). High noise is defined in the same way as low noise but the standard deviation of the normal distribution is three times larger, which is approximately equal to the standard deviation of the WB probability distribution that accounts for uncertainty due in grid-scale WB, density assignment, and interpolation for the three study glaciers (Pulwicki and others, 2017). For each

gridcell value of WB, a random number from the high or low noise distribution is added to obtain a synthetic observation of WB.

As above, a linear regression is calculated between synthetic data and topographic parameters and an estimated WB is then calculated. The process of adding random noise to gridcell-scale values of WB to obtain synthetic observations and fitting a regression to then estimate distributed WB is repeated 100 times to create a population of possible WB estimates. Each repetition uses a different set of random noise resulting in a range of WB values from which a mean WB and RMSE is then calculated. RMSE is found by taking the square root of the mean difference between all gridcells in the synthetic distribution of WB distribution and the WB distribution derived with the sampling design.

Performance metrics

To quantify the performance of each sampling pattern, we use two performance metrics. The first metric is 'convergence', which is defined as the number of measurement locations (n) needed to obtain a mean WB within 5% of the synthetic glacier-wide WB or the RMSE for WB estimated using all real data points. The relation between WB or RMSE and n is smoothed (...describe it here). Convergence is a metric that describes the minimum number of measurement locations required to have an accurate estimate of WB. The second metric is 'variability' (or 'sensitivity'?), which is defined as the number of measurement locations required to have the standard deviation of WB or RMSE due to high noise within 5% of the synthetic glacier-wide WB or the RMSE for WB estimated using all real data points. Again, the curve is smoothed. Variability is a metric of how sensitive the WB estimate is to uncertainty in grid-scale variability and interpolation. The total distance required to obtain measurements along a sampling pattern from a central location is also presented. An efficient sampling design would have lower n values for convergence and variability metrics as well as a short travel distance.

RESULTS

Synthetic data

Based on synthetic data, the most effective sampling patterns are midline & transect and hourglass for all three study glaciers. Midline & transect and hourglass quickly converge to the validation value of WB, have comparatively low standard deviations due to noise and involve the shortest travelling distances (Figure 3 and Table 1). To estimate a mean WB within 5% of the synthetic distribution of WB, ~5 fewer measurements are needed for the midline & transect design as compared to the hourglass design while travelling distance is slightly shorter for the hourglass design. The variability in WB estimates due to high noise becomes small ($< 0.15\,\text{m w.e.}$) when $n > 25$.

The remaining sampling patterns are less effective for estimating WB. Hourglass & circle and random sampling patterns also require small sample sizes but the standard deviations due to noise are larger and travelling distance is substantially greater. Further, the mean WB does not converge to the synthetic distribution of WB for $n < 50$ on Glaciers 2 and 13. Notably, hourglass & circle is not a substantial improvement on the hourglass sampling pattern given the considerable increase in travel distance required to sample the along the hourglass & circle. The circle pattern is not a reliable method of sampling because the standard deviation due to noise is comparatively high and the mean WB does not conver to the synthetic distribution of WB on Glaciers 2 and 13. By far the worst sampling pattern is the midline. WB values are highly sensitive to noise and even with a large number of measurement locations the validation values of WB are not achieved.

Although the convergence patterns are similar between glaciers, the standard deviation due to noise and the number of measurement locations needed to accurately estimate

WB are smallest for Glacier 4 and greatest for Glacier 13 for all sampling patterns. With all sampling patterns, WB is significantly overestimated at low n. If using midline & transect or hourglass sampling patterns, a sampling size of ~20 is sufficient to estimate glacier-wide WB with no marked improvement in accuracy or precision at larger sample sizes. Error in estimated WB is generally greatest in the accumulation area, especially for Glaciers 2 and 13 (Figure 4). When midline pattern is used, the errors are exceptionally high in areas that do not fall along the midline.

Real data

The RMSE rapidly decreases with an increased sample size for all sampling patterns on all three glaciers (Figure5). The sampling patterns trends are similar to those of the synthetic data. Midline & transverse, hourglass, hourglass & circle and random all perform well. However, the higher travelling distance of hourglass & circle and random make these patterns less efficient. Circle pattern appears to perform quite well although RMSE values are slightly higher. Again, the worst pattern is the midline, which results in exceptionally high values of RMSE.

An appropriate sample size is more difficult to determine based on sampling designs from real data. For all glaciers, $n > 30$ does not result in a marked decrease in RMSE. However, none of the sampling patterns result in the RMSE reaching that of the full data for $n < 50$.

DISCUSSION

The optimal sampling design for our study region is the hourglass pattern with a sample size of ~ 30. Based on both synthetic and real data, this design results in low error, is least sensitive to noise and involves relatively small travelling distance. The more broadly applied midline & transverse pattern also produces reliable WB estimates but is slightly less efficient for travelling. A sample size greater than 30 does not significantly improve the accuracy of WB estimated and does not decrease error. This surprisingly low number of measurement locations indicates that high-resolution sampling is not required to accurately estimate WB. Instead, taking measurements throughout the study basin should be prioritized when choosing a sampling design. If the goal is to estimate WB then field resources should be more strongly allocated to distributing measurement locations to capture basin-scale spatial trends in WB. Error is greatest in areas far from sampling locations and in areas with extreme values of topographic parameters.

The most effective sampling designs capture the dominant WB-elevation trend but also have measurement locations in gridcells that span the range of other topographic parameters. Despite the fact that the synthetic distribution of WB on Glaciers 2 and 13 is largely controlled by elevation, measurements locations that are not along the midline are needed to constrain the regression. When the midline pattern is used, the topographic parameters excluding elevation fall within a narrow range so the regression becomes sensitive to noise. Further, the accuracy of estimated WB does not improve at large n along a

midline pattern because the sampling locations do not capture relationships between WB and the remaining topographic parameters. However, if there are a large portion of measurement locations far from the midline then the elevation trend may be less strong resulting in larger errors. Random and hourglass & circle patterns may have larger errors on Glaciers 2 and 13 as a result of the higher proportion of measurement locations far from the midline.

On Glacier 4, the standard deviation of WB due to noise and the number of measurement locations needed to estimate WB is smaller than Glaciers 2 and 13. This trend is likely a result of the small influence of topographic parameters in the distribution of synthetic distribution of WB. The synthetic distribution of WB is derived from a linear regression between field measurements of WB and topographic parameters that explains little of the observed variance (R^2 =0.07). As a result, glacier-wide WB is approximately equal to the mean of WB field measurements. This mostly uniform distribution of snow can therefore be described with few measurement locations, regardless of where the measurements are obtained.

Further work into the optimization of snow survey sampling design is needed. The most obvious limitation of our study is the restriction of sampling design to accessible locations in the ablation area. Lack of measurements in the accumulation area is a sub-optimal sampling design. However, direct measurements of snow depth in the accumulation area are extremely costly because a large amount of time is needed to dig a snow pit, especially in regions with high accumulation. Given the importance of spatially distributed measurements though, there is likely an optimal balance between costly measurements in the accumulation area and less costly measurements in the ablation area. Use of more efficient snow-depth measurement methods, such a ground-penetrating radar (GPR) and lidar, in the accumulation area can also be investigated. Additional extensions to our work include an invetigation of the size and placement of hourglass patterns and to optimize measurement locations based on factors such as topographic parameters.

low n over estiamtes

CONCLUSION
ACKNOWLEDGEMENTS

REFERENCES

Arendt AA, Luthcke SB, Larsen CF, Abdalati W, Krabill WB and Beedle MJ (2008) Validation of high-resolution GRACE mascon estimates of glacier mass changes in the St Elias Mountains, Alaska, USA, using aircraft laser altimetry. *Journal of Glaciology*, **54**(188), 778–787 (doi: 10.3189/002214308787780067)

Bellaire S and Schweizer J (2008) Deriving spatial stability variations from penetration resistance measurements. In *Proceedings ISSW*, 188–194

Bellaire S and Schweizer J (2011) Measuring spatial variations of weak layer and slab properties with regard to snow slope stability. *Cold Regions Science and Technology*, **65**(2), 234–241 (doi: 10.1016/j.coldregions.2010.08.013)

Burgess EW, Forster RR and Larsen CF (2013) Flow velocities of Alaskan glaciers. *Nature communications*, **4**, 2146–2154 (doi: 10.1038/ncomms3146)

Clarke GK (2014) A short and somewhat personal history of Yukon glacier studies in the Twentieth Century. *Arctic*, **37**(1), 1–21

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

Crompton JW and Flowers GE (2016) Correlations of suspended sediment size with bedrock lithology and glacier dynamics. *Annals of Glaciology*, **57**(72), 1–9 (doi: 10.1017/aog.2016.6)

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

Davis JC and Sampson RJ (1986) *Statistics and data analysis in geology*. Wiley New York et al., 2nd edition

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

Elder K, Cline D, Liston GE and Armstrong R (2009) NASA Cold Land Processes Experiment (CLPX 2002/03): Field measurements of snowpack properties and soil moisture. *Journal of Hydrometeorology*, **10**(1), 320–329 (doi: 10.1175/2008JHM877.1)

Fountain AG and Vecchia A (1999) How many stakes are required to measure the mass balance of a glacier? *Geografiska Annaler: Series A, Physical Geography*, **81**(4), 563–573 (doi: 10.1111/1468-0459.00084)

Grünewald T, Bühler Y and Lehning M (2014) Elevation dependency of mountain snow depth. *The Cryosphere*, **8**(6), 2381–2394 (doi: 10.5194/tc-8-2381-2014)

Hock R (2005) Glacier melt: a review of processes and their modelling. *Progress in Physical Geography*, **29**(3), 362–391 (doi: 10.1191/0309133305pp453ra)

Kaser G, Fountain A, Jansson P and others (2002) *A manual for monitoring the mass balance of mountain glaciers*. IHP-VI- Technical documents in hydrology

Kinar N and Pomeroy J (2015) Measurement of the physical properties of the snowpack. *Reviews of Geophysics*, **53**(2), 481–544 (doi: 10.1002/2015RG000481)

Korona J, Berthier E, Bernard M, Rémy F and Thouvenot E (2009) SPIRIT SPOT 5 stereoscopic survey of Polar Ice: Reference images and topographies during the fourth International Polar Year (2007–2009). *ISPRS Journal of Photogrammetry and Remote Sensing*, **64**(2), 204–212 (doi: 10.1016/j.isprsjprs.2008.10.005)

Kronholm K and Birkeland KW (2007) Reliability of sampling designs for spatial snow surveys. *Computers & geosciences*, **33**(9), 1097–1110 (doi: 10.1016/j.cageo.2006.10.004)

López-Moreno JI, Fassnacht S, Beguería S and Latron J (2011) Variability of snow depth at the plot scale: implications for mean depth estimation and sampling strategies. *The Cryosphere*, **5**(3), 617–629 (doi: 10.5194/tc-5-617-2011)

Machguth H, Eisen O, Paul F and Hoelzle M (2006) Strong spatial variability of snow accumulation observed with helicopter-borne GPR on two adjacent alpine glaciers. *Geophysical Research Letters*, **33**(13), 1–5 (doi: 10.1029/2006GL026576)

McGrath D, Sass L, O'Neel S, Arendt A, Wolken G, Gusmeroli A, Kienholz C and McNeil C (2015) End-of-winter snow depth variability on glaciers in Alaska. *Journal of Geophysical Research: Earth Surface*, **120**(8), 1530–1550 (doi: 10.1002/2015JF003539)

Molotch NP and Bales RC (2005) Scaling snow observations from the point to the grid element: Implications for observation network design. *Water Resources Research*, **41**(11) (doi: 10.1029/2005WR004229)

Pulwicki A, Flowers G and Radić V (2017) Uncertainties in estimating winter balance from direct measurements of snow depth and density on alpine glaciers. *Journal of Glaciology*, (in submission)

Réveillet M, Vincent C, Six D and Rabatel A (2016) Which empirical model is best suited to simulate glacier mass balances? *Journal of Glaciology*, **63**(237), 1–16 (doi: 10.1017/jog.2016.110)

Schweizer J, Heilig A, Bellaire S and Fierz C (2008) Variations in snow surface properties at the snowpack-depth, the slope and the basin scale. *Journal of Glaciology*, **54**(188), 846–856 (doi: 10.3189/002214308787780058)

Shea C and Jamieson B (2010) Star: an efficient snow point-sampling method. *Annals of Glaciology*, **51**(54), 64–72 (doi: 10.3189/172756410791386463)

Waechter A, Copland L and Herdes E (2015) Modern glacier velocities across the Icefield Ranges, St Elias Mountains, and variability at selected glaciers from 1959 to 2012. *Journal of Glaciology*, **61**(228), 624–634 (doi: 10.3189/2015JoG14J147)

Walmsley APU (2015) *Long-term observations of snow spatial distributions at Hellstugubreen and Gråsubreen, Norway.* Master's thesis

Wilson N and Flowers G (2013) Environmental controls on the thermal structure of alpine glaciers. *The Cryosphere*, **7**(1), 167–182 (doi: 10.5194/tc-7-167-2013)

Wilson NJ, Flowers GE and Mingo L (2013) Comparison of thermal structure and evolution between neighboring subarctic glaciers. *Journal of Geophysical Research: Earth Surface*, **118**(3), 1443–1459 (doi: 10.1002/jgrf.20096)

Wood WA (1948) Project "Snow Cornice": the establishment of the Seward Glacial research station. *Arctic*, **1**(2), 107–112