Relating regional and point measurements of accumulation in southwest Greenland

Achim Heilig[1,2], Olaf Eisen[3,4], Martin Schneebeli[2], Michael MacFerrin[5], C. Max Stevens[6], Baptiste Vandecrux[7], and Konrad Steffen[8]

[1]Department of Earth and Environmental Sciences, LMU, Munich, Germany
[2]WSL Institute for Snow and Avalanche Research SLF, Davos Dorf, Switzerland
[3]Alfred Wegener Institute Helmholtz-Centre for Polar and Marine Research, Bremerhaven, Germany
[4]Department of Geosciences, University of Bremen, Bremen, Germany
[5]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO USA
[6]Department of Earth and Space Sciences, University of Washington, WA USA
[7]Department of Glaciology and Climate, Geological Survey of Denmark and Greenland, Copenhagen, Denmark
[8]Swiss Federal Research Institute WSL, Birmensdorf, Switzerland

Correspondence: Achim Heilig (heilig@r-hm.de)

Abstract. In recent decades, the Greenland ice sheet (GrIS) has frequently experienced record melt events, which significantly affected surface mass balance (SMB) and estimates thereof. SMB data are derived from remote sensing, regional climate models (RCMs), firn cores and automatic weather stations (AWSs). While remote sensing and RCMs cover regional scales with extents ranging from 1–10 km, AWS data and firn cores are point observations. To link regional scales with point measurements, we investigate the spatial variability of snow accumulation (b_s) within areas of approximately 1–4 km^2 and its temporal changes within two years of measurements. At three different sites of the southwestern GrIS (Swiss Camp, KAN-U, Dye-2), we performed extensive ground-penetrating radar (GPR) transects and recorded multiple snow pits. If the density is known and the snowpack dry, radar-measured two-way travel time can be converted to snow depth and b_s. We spatially filtered GPR transect data to remove small scale noise related to surface characteristics. The combined uncertainty of b_s from density variations and spatial filtering of radar transects is at 7–8% per regional scale of 1–4 km^2. Snow accumulation from a randomly selected snow pit is very likely representative of the regional scale of 1–4 km^2 (with probability $p = 0.8$ for a value within 10% of the regional mean for KAN-U, and $p > 0.95$ for Swiss Camp and Dye-2). However, to achieve such high representativeness of snow pits, it is required to determine the average snow depth within the vicinity of the pits. At Dye-2, the spatial pattern of snow accumulation was very similar for two consecutive years. Using target reflectors placed at respective end-of-summer-melt horizons, we additionally investigated the occurrences of lateral redistribution within one melt season. We found no evidence of lateral flow of meltwater in the current climate at Dye-2. Such studies of spatial representativeness and temporal changes in accumulation are necessary to assess uncertainties of the linkages of point measurements and regional scale data, which are used for validation and calibration of remote sensing data and RCM outputs.

1 Introduction

Numerous recent studies have documented a continuous mass loss from the Greenland ice sheet (GrIS) using remote sensing data and/or estimates from model simulations (e.g., Shepherd et al., 2012; Velicogna et al., 2014; Khan et al., 2015; van den Broeke et al., 2016; Sørensen et al., 2018; Mouginot et al., 2019). From 1980 to 2018, mass loss from the GrIS increased by a factor of six (Mouginot et al., 2019), and over the last two decades the major mass loss process has changed from solid ice discharge to surface mass balance (SMB) (Enderlin et al., 2014; van den Broeke et al., 2016; Mottram et al., 2019). SMB can be regarded as the sum of snow accumulation (b_s) and lateral redistribution by sublimation, wind and runoff. Depending on the location, lateral redistribution can increase SMB as well as decrease it. Over most of the GrIS, net accumulation is the dominating factor for SMB (Koenig et al., 2016), while recent negative trends in SMB are related to surface melt and runoff (Vaughan et al., 2013). Despite of all advances, SMB estimates remain a major source of uncertainty in ice-sheet mass-balance calculations (van den Broeke et al., 2009). This is because surface mass fluxes, such as snowfall and melt, cannot be measured by remote-sensing technology and derived estimates on snowfall can still have significant errors (Bennartz et al., 2019). Hence, predictions of SMB are usually obtained using scarce in situ measurements together with regional climate models (RCMs), which can introduce significant uncertainties (Vernon et al., 2013) as well. Different scales between in situ observations and simulations may also contribute to these uncertainties. The spatial resolution of RCMs and remote sensing data are limited to regional scales (on the order of one to tens of square kilometers), while in situ observations cover point data (on the order of a few square meters or less). Effects of wind redistribution, for instance, are leveled out for regional scales but can have significant influences at point scales. As a consequence, evaluation and validation of regional-scale data products using in situ data is difficult without knowledge of the spatial extent and representativeness of the point measurements. To date, only a few studies have investigated how representative point observations (e.g., snow pits, firn cores, mass-balance-stake readings, automatic weather station [AWS] measurements) are of the surrounding several square kilometers.

Within the last decade several studies have used radar systems to quantify accumulation variability in Greenland by tracking internal reflection horizons (IRHs) (e.g., Dunse et al., 2008; Miège et al., 2013; Hawley et al., 2014; Karlsson et al., 2016; Koenig et al., 2016; Lewis et al., 2017, 2019). While those studies aimed to track IRH variability using data from long ground transects of roughly 100 km (Miège et al., 2013) to more than 1000 km (Hawley et al., 2014) length or using airborne radar data (Karlsson et al., 2016; Koenig et al., 2016; Lewis et al., 2017), only Dunse et al. (2008) linked the point observations from snow pits and cores to the surrounding area. Koenig et al. (2016) used airborne radar data from NASA's Operation Ice Bridge to calculate accumulation rates with a stated uncertainty of 14%, and they compared their results to outputs from an RCM. They compare radar-derived accumulation to two sites with core data, but the locations of those sites are up to 8 km away from the radar track. Hence, it is not possible to identify whether mismatch between the core- and radar-derived accumulations is due to spatial variability or to assumptions in radar-data processing. Systematic offsets in b_s between radar data and RCM outputs, however, occur in northern Greenland with discrepancies between RCMs and radar up to 30% (Karlsson, personal communication). Other recent studies attempt to relate point observations of melt events within the percolation zone of the GrIS with annual atmospheric patterns (Graeter et al., 2018) or determine the mass of percolating liquid water and compare

percolation depths observed by upward-looking radar (upGPR) with temperature records in snow and firn (Heilig et al., 2018). In addition, several studies have quantified temporal accumulation variability using ice core records (e.g., Mosley-Thompson et al., 2001; Vandecrux et al., 2019). Since quantification of spatial representativeness of single point measurements for the surrounding square kilometers has only been conducted for one location in western Greenland so far (Dunse et al., 2008), there is a need to explore uncertainties at local and regional scales. The best means of resolving these uncertainties are to increase the spatial coverage of direct measurements (Farinotti et al., 2014) and to improve our understanding of how well point measurements represent a larger area.

Point observations, such as snow pits and ice cores are usually performed once a year at most. Such temporal snapshots limit the evaluation of spatial representativeness as they can be influenced by recent weather conditions. Hence, it is necessary to clarify whether regional accumulation patterns are consistent over more than one accumulation season to investigate if temporally continuous point measurements such as AWS data, upGPR and neutron probes remain representative.

Meltwater percolation can move mass from snow to the underlying firn (e.g., Charalampidis et al., 2016; Humphrey et al., 2012; Heilig et al., 2018) or even laterally along the surface slope (Humphrey et al., 2012). Hence, surface melt affects SMB (e.g., Sasgen et al., 2012) and accumulation (Heilig et al., 2018). However, it is unlikely that water percolation and mass redistribution are homogeneous over regional scales. Consequently, it is necessary to assess the impact of melt on temporal changes in accumulation distribution for the percolation zone of the GrIS.

The aim of this work is to relate point scales to regional scales of one to several square kilometers in area to improve our understanding of the representativeness of point measurements. For this purpose, we examine snow-pit and ground-penetrating radar (GPR) data from two sites within the percolation zone of the GrIS and one site at the equilibrium line gathered over several field seasons. For each site, we investigate density variability between measurements from up to six snow pits within an area of 4 km^2 made in a single season, process radar transects of up to 25 km recorded in close proximity to those snow pits, and spatially extrapolate the radar-derived accumulation to estimate area-wide accumulation variability. For temporal comparisons, we use continuous observations of accumulation and melt recorded by upGPR (Heilig et al., 2018). Our results show that spatial representativeness of snow accumulation for a point measurement (snow pit) is large but values can be affected by local wind-induced surface roughness. We recommend to apply multiple snow depth measurements at the vicinity of the pits to better assess accumulation on regional scales.

2 Methodology

2.1 Test site, instrumentation and data processing

We collected radar data along transects at three different locations on the southwestern GrIS over several years (Figure 1, Table 1). The sites were visited in spring of each year (see Table 1). At Swiss Camp a small transect was measured in May 2015 by towing a GPR trolley on foot. The measurements were triggered by an odometer wheel. Geolocation was only performed for starting and end points of some radar lines, and locations in between are interpolated. We used two different units for the recorded five radar transects. At Dye-2 and KAN-U in May 2017, we employed an IDS (Ingegneria dei Sistemi, Pisa, Italy)

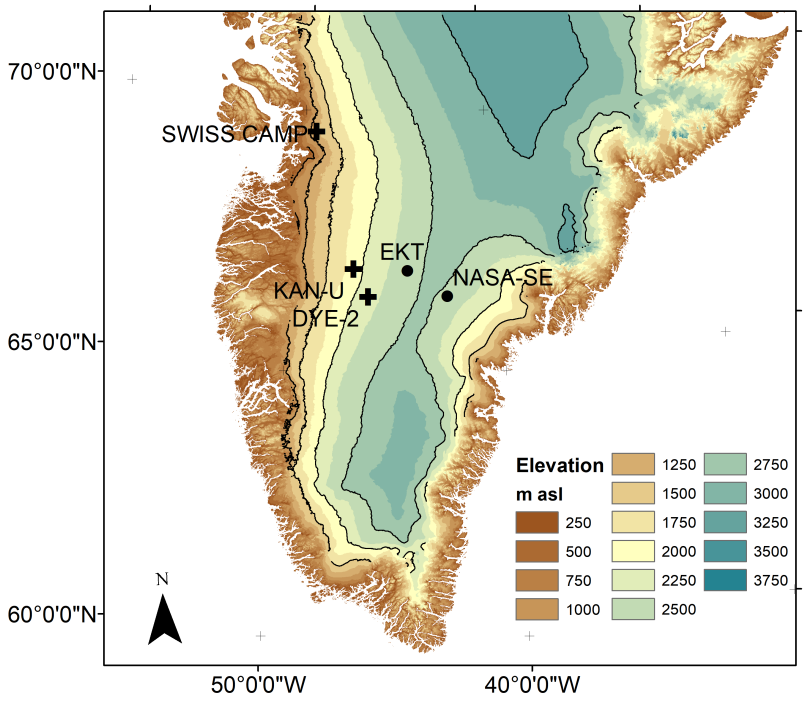

Figure 1. Map displaying locations of radar transects investigated for this study in Southern Greenland (black crosses). The black dots indicate additional locations, where snow pits were dug for snow density analysis. The colors are 250 m elevation bands with the maximum elevation per band indicated, the black contour lines are at 500 m intervals. The underlying digital elevation model was generated by Howat et al. (2014).

FastWave control unit with dual frequency antennas. The respective frequencies are listed in Table 1. Radar measurements at Swiss Camp in May 2015 and at KAN-U in May 2013 were conducted using a RAMAC system (MALA Geoscience, Sweden). The radar data from Swiss camp have 0.05 m trace distance along track. The transects at Dye-2 and KAN-U were recorded in time mode and dragging the antennas behind a snow machine. Because small variations in snow-machine speed cause recorded radar traces to be spaced unevenly, the traces are averaged to generate equidistant spacing. The resulting horizontal trace distance is 0.5 m for both Dye-2 transects and the 2017 KAN-U transect. The trace spacing along the 2013 KAN-U transect is 1.5 m because the snow-machine speed was faster. For the Dye-2 and KAN-U surveys, antennas were connected to a GPS receiver for geolocation of the GPR transect.

All recorded radar traces were processed in a very similar way. In case first arrivals were delayed by more than approximately 2 ns, we started with a correction for the DC shift. We corrected offsets in the zero line of each radar trace (wow) utilizing a dewow function and filtered low (approximately below 0.5 times the center frequency) and high frequency noise (approx. above 1.5 times the center frequency) applying bandpass filters. We further applied background removals to minimize disturbing

Table 1. Metadata for the five GPR transects analyzed in this study. Coordinates are presented in geographical coordinates with elevation in meters above sea level. DoA is date of acquisition.

Location	DoA	Coordinates	Trace distance [m]	Total length [km]	Antenna frequency [MHz]
KAN-U	May 2013	67.0011°N/ 47.0276°W/ 1860	1.5	15.3	800
Swiss Camp	May 2015	69.5552°N/ 49.3653°W/ 1170	0.05	0.35	1600
Dye-2	May 2016	66.4779°N/ 46.2856°W/ 2120	0.5	20.6	1600 and 600
KAN-U	April 2017	67.0011°N/ 47.0276°W/ 1860	0.5	10.9	200 and 600
Dye-2	May 2017	66.4779°N/ 46.2856°W/ 2120	0.5	24.9	200 and 600

effects from the direct wave and antenna ringing. For all radar transects, we corrected for divergence losses by gain functions and interpolated to equidistant traces. The zero-crossings of the snow surface reflections were corrected to be at time zero.

The measured quantity of radar transects is the two-way travel time (TWT with mathematical symbol τ) from the transmitter to the reflector and back to the antennas (e.g., Heilig et al., 2018). In dry snow and firn (with two contributing volume fractions $\theta_a + \theta_i = 1$), the wave propagation depends solely on the relation of air (θ_a) to ice volume fraction (θ_i) (e.g., Kovacs et al., 1995; Mätzler, 1996). Hence, with the bulk snow density (ρ_s, the average density of the entire snow column) measured in snow pits, we can convert from TWT to snow depth ($\overline{L_s}$) and the amounts of bulk accumulation (b_s with unit kg/m^2) using the equation

$$b_s = \overline{L_s}\rho_s \tag{1}$$

with

$$\overline{L_s} = \frac{\tau}{2}\frac{c}{\frac{\rho_s}{\rho_i}(\varepsilon_i^\beta - 1) + 1}. \tag{2}$$

The ice density ($\rho_i = 917$ kg/m^3), the exponent $\beta = 0.5$ (related to a medium with random orientation at the micro scale), the speed of light in vacuum (c) and the relative dielectric permittivity of ice ($\varepsilon_i = 3.18$) are constants taken from previous literature (e.g., Heilig et al., 2018). The reflections of the previous end-of-melt-season (EMS) horizons are clearly detectable in all radargrams. We relate internal reflecting horizons (IRHs) to depths at pit locations using the measured bulk snow density ρ_s within each pit. Accordingly, we choose the zero-crossing of the IRHs as the first break of the respective layer. To identify the EMS horizon of 2015 at Dye-2 in May 2017, we make use of target reflectors that were buried in May 2016 on the 2015 summer horizon. Hence, in May 2017, it was possible to revisit those locations with the radar and unambiguously distinguish between signal reflections arising from the 2015 and 2016 EMS horizons.

However, before applying a constant ρ_s over the entire length of the radar transects, one has to investigate the spatial heterogeneity in ρ_s over an area of comparable size. To accomplish this, we dug several snow pits at Dye-2 in May 2015 and 2016, at Swiss Camp in May 2015 and 2018 and at KAN-U in April 2016. In each pit, we measured the bulk density of the snow from the surface down to the previous season's melt surface (see Table 2 for details). The snow pits were dug at various

Table 2. Locations of snow density analyses with date of acquisition (DoA), number of snow pits (N), mean density ($\bar{\rho_s}$), density (ρ_s) range in percent of mean ρ_s and mean snow depth ($\overline{L_s}$).

Location	DoA	N	$\bar{\rho_s}$ [kg/m^3]	range in ρ_s [%]	$\overline{L_s}$ [m]
Dye-2	May 2015	6	355.5	-4 to +2	0.84
EKT	May 2015	5	341.3	-5 to +4	0.83
NASA SE	May 2015	2	364.5	-2 to +2	1.63
Swiss Camp	May 2015	4	358.4	-5 to +5	1.15
KAN-U	April 2016	4	346.0	-6 to +5	0.88
Dye-2	May 2016	6	320.1	-6 to +4	0.94
EKT	May 2016	3	339.2	-2 to +2	0.89
NASA SE	May 2016	2	369.7	-1 to +1	1.70
Swiss Camp	May 2018	3	351.3	-2 to +3	1.45

distances from each other, at maximum up to 1 km apart. In addition to locations where we collected radar data, we also investigated spatial variability in ρ_s at two more sites, EKT and NASA-SE (Figure 1). As these two sites are located within a distance of 45–60 km of the GrIS ice divide (W of the divide - EKT and E of the divide - NASA SE, see Figure 1), they extent our data analysis of spatial variability of ρ_s to the dry-snow zone. The recorded pits at NASA SE provide data for a high accumulation site as well. Table 2 displays the numbers of snow pits, the mean density of all pits for that site and year, and the ranges (minimum divided by mean and maximum divided by mean) in percent. To process the radar data collected at Dye-2 in May 2017, we use density data from firn cores to calculate radar wave speed between the summer 2016 and summer 2015 horizons. Snow temperature measurements ensured dry and subfreezing conditions.

For all three sites, long term meteorological observations exist. To discuss the meteorological conditions at each site, we use wind data from the GC-Net stations (Steffen and Box, 2001) at Dye-2 (September 2011 to May 2018, with gaps in between) and at Swiss Camp (May 2016 to May 2017) and the PROMICE station (van As et al., 2011) at KAN-U (April 2009 to September 2016).

2.2 Transect data analysis

The measured TWT of the GPR data are influenced by small-scale surface roughnesses and vertical time sampling. Wind-induced surface features, such as sastrugi, appear in 2-D radar transects as discontinuous, erratic noise. Ideally, we would have performed radar surveys on high-resolution grids (i.e. with spacing smaller than the characteristic length of the features) to spatially extrapolate such features to the non-surveyed areas. However, it was not possible to conduct such high-resolution surveys in the one to two days available at our sites. Instead, we apply spatial smoothing to minimize artifacts from vertical sampling and to remove wind-induced surface-feature noise.

The time sampling of the recorded GPR transects ranges from 0.05 ns per sample (Swiss Camp 2015) to 0.24 ns per sample (Dye-2 2017), corresponding to approximately 0.006 m and 0.028 m per sample respectively. For the longer transects at KAN-U and Dye-2 (Table 1), the vertical sampling is always coarser than 0.1 ns/sample. As displayed in Figure 2, the raw radar data for these transects are continuously fluctuating by ± 1 sample (corresponding to roughly ± 3 cm). Such effects are caused by amplitude clipping of the signal response and uncertainties of the zero-crossing as consequence of the vertical sampling. For each radar trace, we consistently picked the first strong positive half cycle and shifted the first break upwards to match the zero-crossing. However, due to a vertical sample intervals of 0.25 ns, it is likely that the strongest amplitudes shift by 1–2 samples for consecutive radar traces. To reduce effects caused by the amplitude shifts, in our (lower resolution) KAN-U and Dye-2 data, we applied a Savitzky-Golay filter (Savitzky and Golay, 1964) with frame length of 20 m and polynomial order of 3 (Figure 2, yellow line). At Swiss Camp with the much finer vertical sample interval, it is adequate to filter with 1 m frame length to reduce clipping and zero-crossing uncertainties.

At Swiss Camp, where we surveyed on a sub-meter grid, we are able to investigate small scale accumulation variability directly. For the other two sites, however, the transects were several kilometers in length and not in a regular grid. To enable quantitative geostatistical extrapolation over areas not surveyed with the radar, it is necessary to remove small-scale surface roughness from the data. With a horizontal sampling resolution of 0.5 to 1.5 m, the variability in the radar-derived snow depth is dominated by surface-wind features such as dunes and sastrugi. We determine the average wavelengths and amplitudes of all four longer transects (Dye-2 and KAN-U) by calculating the average distance (in meters) in between peaks in snow depth and the arithmetic mean of amplitude of those peaks. The average wavelengths are between 50–62 m and amplitudes range in between 6–8 cm. To filter such surface roughness, we again employ Savitzky-Golay filtering. We search numerically for filter frame lengths for which the average standard deviation within a 20 m radius (below half of the wavelength) around each radar trace is 1 cm or less. The resulting filter frame lengths range from 135 m (Dye-2, May 2016) to 210 m (KAN-U, May 2013), which allowed the removal of short wavelength variations with an amplitude of about ± 0.05 m and more (Figure 2 green line). A smoothing length of 20 m has been used by other recent studies dealing with large scale GPR transects as well (e.g., Lewis et al., 2019). We use the smoothed data for spatial extrapolation.

2.3 Spatial extrapolation

In order to analyze accumulation patterns over a larger area, it is necessary to extrapolate the data gathered along the radar transects. One radar trace provides a single depth estimate to a specific reflector. Combining GPR-derived snow accumulation transects with geostatistical techniques is a powerful method to model spatial occurrences of continuous subsurface features. Similar combinations of geophysical and stochastical techniques have been used in previous research (e.g., Rea and Knight, 1998; Tercier et al., 2000). The benefit of radar data is that numerous data pairs for a wide range of measurement distances are recorded enabling more constrained experimental variograms. Webster and Oliver (2007) state that sample size is directly related to the precision of variogram estimates, while variograms are used to estimate the variance of a parameter (here snow accumulation) at increasing intervals of distance in between measurements and in multiple directions. Before spatial extrapolation of a data parameter, the data must fulfill several prerequisites: data have to be spatially continuous and spatially correlated

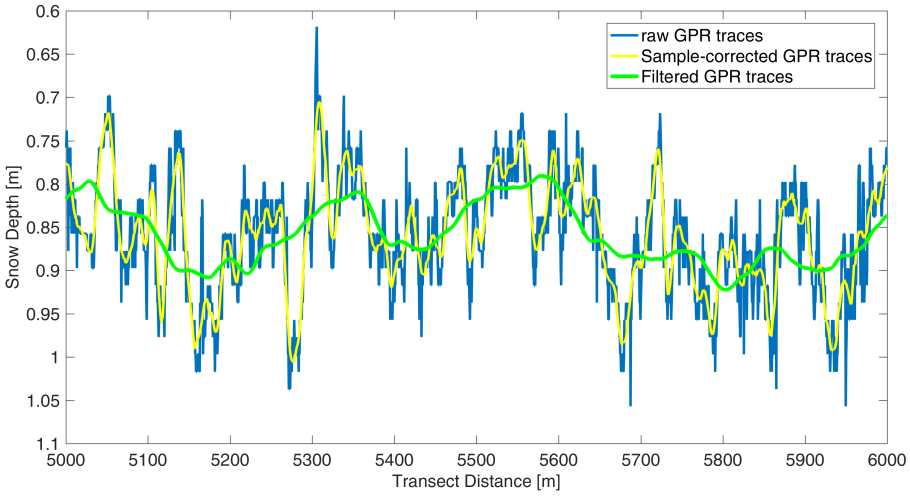

Figure 2. A 1 km section for radar-derived snow depths for the end-of-melt season 2016 horizon of the radar transect from May 2017 at Dye-2. Displayed are the raw first breaks (blue line), after being corrected for sample uncertainties (yellow line) and the final product being used to assess spatial variability for an area of several square kilometers (green line).

within a specific distance and the expected mean and variance of the data should be invariant in space (e.g., Rea and Knight, 1998). We used experimental variograms to investigate spatial correlation and snow accumulation at the surveyed sites is spatially continuous (accumulation occurred everywhere within the area of interest, governed by local weather conditions). To ensure that mean and variance are invariant, we investigated trends in X- and Y- direction separately and subtracted these trends before further analysis. At DYE-2 and KAN-U, we discovered accumulations trends in both, X- and Y- directions, over the distances surveyed, while at Swiss Camp, we found a simple one-dimensional trend.

For spatial extrapolation of the univariant parameter snow accumulation, we use ordinary kriging, which is the most robust and most commonly used method (Webster and Oliver, 2007). Ordinary kriging requires normal distribution of the data. Figure 3 displays the probability distributions of all five radar transects. If the distribution (plotted crosses) follows the straight line, the data are normally distributed. At least 10–80% of data match normality for all five GPR transects, and, consequently, no data transformation is applied. We used the Geostatistical Analyst toolbox in ArcGIS10.4.1 to perform the kriging.

After trend removal, the next step in ordinary kriging is to simulate variograms, which adequately mimic the calculated experimental variograms. Standard performance measures to assess the model accuracies are mean prediction errors (values should be at 0 kg/m^2) and root mean square (RMS) standardized prediction offsets (values should be 1). We present such accuracy assessments in Table 3. In addition, we found directional anisotropy of the covariance in all of the longer transects, which means that correlation ranges of accumulation vary with direction. Hence, we modeled variograms with different correlation ranges per direction. The correlation range marks the limit in distance of point pairs for being spatially dependent. Major and minor axis of the correlation range ellipsoid used for the variogram modeling are given in Table 3 as well. Swiss

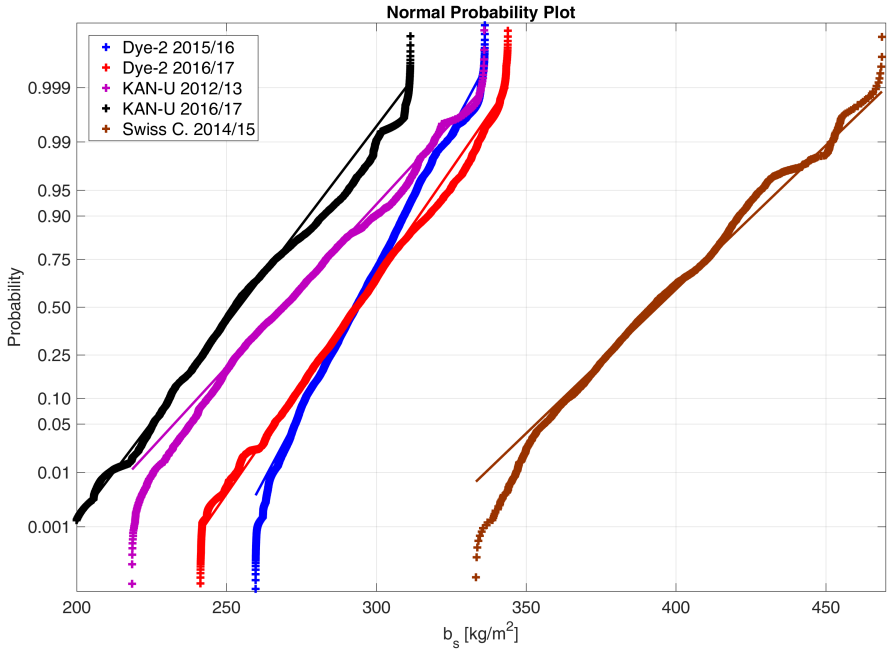

Figure 3. Normal probability plots displaying deviations from a standard normal distribution (straight line with corresponding color). The crosses plot the empirical probability versus the data value for each radar-determined b_s value. We display smoothed transects values for KAN-U and Dye-2 and solely vertical sampling corrected values for Swiss Camp.

Table 3. Kriging results with description of correlation ranges for the major and minor axis used in the variogram modeling, the resulting mean prediction error (pred. err.) and the root mean square (RMS) standardized prediction error.

Location	Anisotropy	Range major/minor [m]	Mean pred. err. [kg/m^2]	RMS standardized pred. err.
KAN-U 2012/13	y	274/ 91.5	0.01	1.02
Swiss Camp 2014/15	n	3.8	0	1.11
Dye-2 2015/16	y	126/ 80	0	0.83
KAN-U 2016/17	y	96.5/ 63	0	1.09
Dye-2 2016/17	y	156/ 52	0	1.33

Camp is an exception and can be modeled simply by an isotropic variogram. At Dye-2, a spherical variogram model provided highest prediction accuracies while at KAN-U and Swiss Camp, the usage of stable variogram modeling resulted in lowest mean prediction errors and best RMS standardized prediction offsets. The presented correlation ranges in Table 3 represent the direction-wise major extrapolation range. Nugget effects (description of the measurement errors) are small with values below 5 kg/m^2 for all transects. Our kriging outputs have a spatial resolution of 20 m by 20 m for the larger transects and of 0.1 m by 0.1 m for Swiss Camp.

To assess the distribution and spatial representativeness of the data, we calculate normalized accumulation values ($b_{s,N}$) and normalized cumulative probability distributions. Normalized accumulation is computed such that the individual kriged accumulation value (b_s) is divided by the mean kriged accumulation per site and campaign $\overline{b_s}$: $b_{s,N} = \frac{b_s}{\overline{b_s}}$. In Figures 4b, 6c and 7c, data distributions of $b_{s,N}$ are displayed as box plots with the whiskers set to the 5% and 95% percentiles respectively. Using the recorded radar traces, we determine whether any randomly located point measurement such as a snow pit is representative of the entire extrapolated area. We average all radar traces within a radius of 1 m around each radar trace (which represents a standard pit size) and scale this data point by the mean of the kriged output for the same campaign. Data distribution for each campaign including filtered and sampling-corrected data (see Section 2.2) are presented to describe offset dependencies. At KAN-U for the 2012/13 data, we increase the assumed pit size to an area with 2 m radius because of more sparse horizontal data resolution (1.5 m in between traces). Corner locations of radar transects with less than four (three for KAN-U 2012/13) neighboring traces within the respective search radius are excluded.

3 Results and Discussion

We first discuss errors associated with converting measured TWT to accumulation because understanding these errors is essential for assessing how representative a single point observation, such as a snow pit, is of a larger area; we present that assessment in Section 3.2. We then evaluate whether accumulation-patterns over two consecutive years at Dye-2 are different. Finally, we investigate how accumulation changes due to melt and liquid-water percolation. Such effects could be caused by strong lateral differences in melt or lateral flow of meltwater. In the following, to distinguish between offsets, deviations from mean and data distribution, we will describe offsets, deviations and uncertainties of b_s values in percentage (%) and data distribution as probability values of 0–1.

3.1 Error in travel time to accumulation conversion

We investigate the error that we introduce by assuming a single bulk density in the conversion from TWT to snow depth for an entire GPR transect. Hence, we determine the spatial variability in density within the respective area. Table 2 presents snow-pit data from our three study sites and two additional sites. The data were collected over three years, and the distances between pits ranged from a few meters up to 1 km, while snow depths ranged from 0.83 m to 1.70 m. The inclusion of two more sites close to the southern Greenland ice divide extents the data set to a low accumulation site west of the ice divide (EKT: $\overline{b_s} \sim 300$ kg/m^2) and a high accumulation sites east of the divide (NASA SE: $\overline{b_s} \sim 600$ kg/m^2). The range in density variation from $\bar{\rho}$ in Table 2—independent of distances in between pits—does not exceed -6 to $+5\%$ for nine snow pit campaigns in total, at five different locations for the southern GrIS. Calculated range averages for the last column in Table 2 are -3.7 to $+3.1\%$. We thus consider $\pm 5\%$ variation in average density to be a robust and conservative estimator of uncertainty within areas of several square kilometers for these regions. This corresponds well with observations by Proksch et al. (2016), who derived a mean measurement uncertainty for density of 2–5%.

Uncertainty in ρ_s results in only a small uncertainty in the derived $\overline{L_s}$: ρ_s factors into the conversion of τ to $\overline{L_s}$ as a fraction within the denominator (Equation 2). For our measured TWTs, a $\pm 5\%$ variation in ρ_s leads to a 0.7–1.4% uncertainty in $\overline{L_s}$ for bulk ρ_s values of 200–450 kg/m^3. Additional uncertainty in $\overline{L_s}$ is introduced by the smoothing applied to the larger transects. The average RMS deviation in snow depth of the smoothed transects from the sample-corrected transects at Dye-2 and KAN-U is 4.5 cm (5–6%). Combining the errors due to smoothing of radar traces and using a mean density for processing radar transects with observed ρ_s variations using Equation 1 leads to an average uncertainty in b_s of 7.0–7.9%. This uncertainty is significantly smaller than discrepancies between RCM simulations and Operation IceBridge airborne radar determinations (16%) (Koenig et al., 2016) and smaller than measured relative standard deviations in density observed within the same study (12%). However, to increase the robustness of accumulation estimates and to decrease effects of spatial extrapolation, we consider an estimated maximum uncertainty of 10% in b_s determined from radar data as a conservative estimate for regional catchments of size of 1–5 km^2.

3.2 Spatial representativeness of point accumulation

3.2.1 Swiss Camp

Figure 4a shows the measured radar grid and area-wide snow accumulation predicted by kriging for Swiss Camp. In May 2015, we measured a transect length of roughly 350 m with along-transect resolution of 5 cm and transect lines separated by 60 cm. Radar data were only filtered to remove sample-related noise (see Section 2.2), which allow us to identify small scale variabilities in b_s within 10 cm grid cells. The arithmetic mean of b_s within the surveyed area in Figure 4a is 393 kg/m^2 with a standard deviation of 28 kg/m^2 (7.1%). Within the northeasterly part of the presented accumulation distribution (Figure 4a), we find above average accumulation. Along the longer transect lines (from south to north), there are several spots with below average accumulation. Since the extrapolation was performed in accordance to the observed variogram range without boundary conditions being set (snow accumulation outside the measured grid existed, we just do not have information to quantify it), it is impossible to identify minimums and maximums as artifacts or actual variability patterns outside the grid lines. However, the observed minimums in b_s along the south-north transect lines are at regular distances between 8–10 m and are likely the result of wind-generated surface features. Prevailing wind direction is from the East with low variations (Figure A1a). Along the wind direction, the interpolated area range (East to West) does not exceed 21 m, which is less than the wavelength of the variability pattern observed at Dye-2 (Figure 2, Section 2.2). However, for the cross-wind direction, a wavelength of 8–10 m for dune dips seems to be apparent at Swiss Camp.

Figure 4b displays the normalized accumulation distribution ($b_{s,N}$) through box plots. The median (red horizontal line), interquartile range (IQR framed by the blue box), 5% and 95% percentiles (whiskers) and values outside a distribution of $p = 0.9$ (red crosses) are displayed. Similar to the recorded radar data (Figure 3), a large proportion of extrapolated b_s ($p > 0.9$) follow a normal distribution. In addition, arithmetic mean, median and mode for this data distribution at Swiss Camp are very similar ($\bar{b_s} = 392.5$ kg/m^2, $b_{s,med} = 391.4$ kg/m^2, $b_{s,mod} = 381.2$ kg/m^2) indicating symmetric data distribution as well (Fahrmeir et al., 2011). A normal distribution, hence, symmetric data distribution allows direct derivation of distribution probabilities.

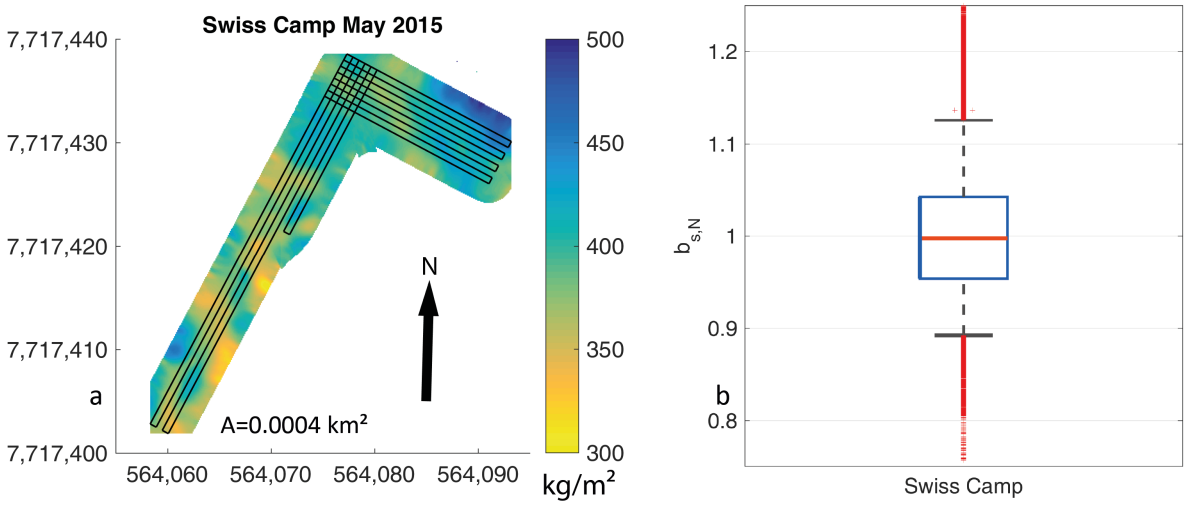

Figure 4. (a) Kriging results for the small radar transect at Swiss Camp for snow accumulation. The black lines display the recorded radar traces and the black arrow indicates geographic north. A total area of 400 m^2 can be covered by the applied spatial extrapolation. (b) Box plot displaying normalized data distribution ($b_{s,N}$) of kriged output with the red horizontal line showing the median, the box framing the interquartile range and the whiskers displaying the 5% and 95% percentiles. Outliers are shown as red crosses. The coordinates in (a) are given in UTM zone 22N with datum WGS1984.

For instance, the standard deviation for normalized b_s in Figure 4b is 0.07 which means that $\pm 7\%$ deviation comprise $p = 0.68$ of data. However, there is a slight difference in whisker lengths (5% percentile at $b_{s,N} = 0.89$, 95% percentile at $b_{s,N} = 1.13$) indicating a small shift towards higher values and asymmetry for the data distribution tales. Skewness of this data distribution equals 0.42. However, $p = 0.86$ of data are within the given $\pm 10\%$ uncertainty for the entire surveyed area. Consequently, the presented data distribution in Figure 4b indicates that with a probability of $p = 0.86$, the kriged 10 cm by 10 cm grid points are within 353–432 kg/m^2.

We use the recorded radar traces to numerically analyze how representative any pit location within the ~400 m^2 area would be. As described in Section 2.3, we define a search radius of 1 m around each radar trace. Radar-derived b_s values are averaged within the search radius. Results are plotted as normalized cumulative probability plot (Figure 5). Our analysis shows that with probability $p > 0.95$, any pit location would provide accumulation values within $\pm 10\%$ of the arithmetic mean for this 400 m^2 area at Swiss Camp. Only very few pit locations ($p < 0.05$) at Swiss Camp provide b_s values exceed a 10% deviation from arithmetic mean with an overestimation of about 15% at maximum.

3.2.2 Dye-2 and KAN-U

For the much longer radar transects at Dye-2 and KAN-U, we filtered out wind-induced surface variabilities of the radar traces to increase spatial extrapolation with enlarged variogram ranges from 10 – 30 m to 50 – 270 m (Table 3). Such filtering implies

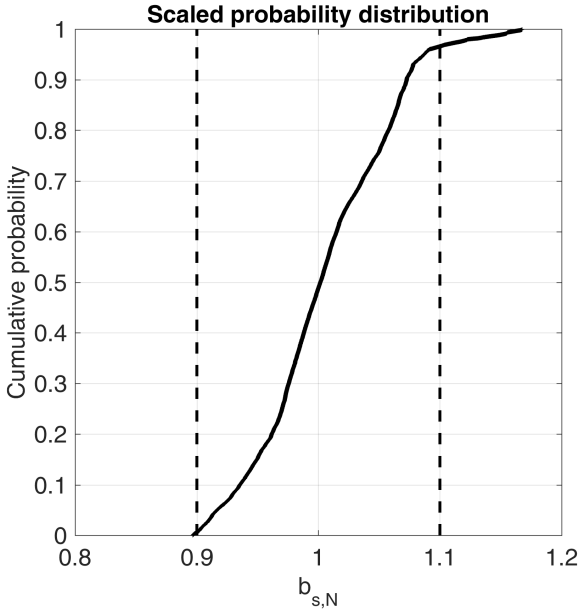

Figure 5. Normalized cumulative probability distribution of radar-derived accumulation (b_s) within any subset with 1 m radius of the GPR transect at Swiss Camp. The dashed vertical lines represent the determined uncertainty range of $\pm 10\%$.

spatial smoothing of surface roughnesses, which could be performed in the field by extensive snow-depth probings. Later in this section, we present comparisons for spatial representativeness of filtered and non-filtered GPR data. In 2016 and 2017, the radar transects were designed to follow the prevailing wind direction to better assess systematic inhomogeneities for Dye-2 and KAN-U in 2017 (see Figures 6, 7 and A1b and c).

At Dye-2, we recorded 21 km of continuous radar data in May 2016 and 25 km in May 2017. This results in geostatistical predictions of snow accumulation over an area of 2.4 km^2 for 2015/16 and almost 4 km^2 for 2016/17 (Figure 6a and b). The arithmetic mean for b_s in May 2016 is 293 kg/m^2 with standard deviation $\sigma = 11$ kg/m^2. 2016/17 results in almost identical values with a mean accumulation of 296 kg/m^2 and $\sigma = 15$ kg/m^2.

The box plots in Figure 6c represent the same quantiles as in Figure 4b. Data distribution for Dye-2 in 2015/16 is very homogeneous with an IQR of only $\pm 2.5\%$. The whiskers for the same year reach $\pm 6\%$. Hence, b_s in May 2015/16 varies only little with $p > 0.9$ of data within the uncertainty margins of $\pm 10\%$. Since already more than 95% of radar-derived b_s follow a normal distribution (Figure 3), values of extrapolated b_s have a high distribution symmetry as well. We observe slightly less homogeneity in the subsequent year at Dye-2. Here, the IQR increases to $\pm 3\%$, with 5% and 95% percentiles being slightly below the error margins of $\pm 10\%$. Transforming the named $b_{s,N}$ values to numbers: for b_s in 2015/16, we observe a likelihood of $p = 0.9$ that all extrapolated 20 m by 20 m pixels range from 275–311 kg/m^2. In May 2017, extrapolated b_s values for an area of 4 km^2 are at 266–326 kg/m^2 with a likelihood of $p > 0.9$.

The normalized cumulative probability distributions in Figure 6d demonstrate how representative a randomly located snow pit would be for the entire surveyed area. We analyzed both, the sample resolution corrected radar data (dotted lines) and the filtered data (solid lines, see Section 2.2). The filtered data in Figure 6d indicate that b_s measured in a snow pit anywhere within the radar survey (black lines Figures 6a, b) would be within the error margins of $\pm 10\%$ from the mean of the entire kriged area with a high probability ($p = 0.99$ for winter accumulation in 2015/16 and $p = 0.91$ for 2016/17). The unfiltered data, however, show a decreased representativeness with $p = 0.89$ in 2015/16 and $p = 0.77$ in 2016/17 for the same uncertainty range of $\pm 10\%$. Here, snow depth is solely derived from the snow pit. Such values demonstrate that b_s data derived simply from a snow pit without averaging snow depth for an area around the pit location will decrease the area-wide representativeness at Dye-2.

It is hard to explain the significantly low b_s variability in May 2016 at Dye-2. In theory, low wind speeds could lead to the absence of snow dunes and sastrugis and reduce the spatial heterogeneity of b_s. However, the recorded wind data do not confirm below average wind for this respective winter season. Determined statistics for wind speeds per winter season (01 Oct. – 01 May) at Dye-2 are very consistent over the last six years (2011/12 – 2016/17). We can only speculate that a snow fall event five days prior to the radar measurements in May 2016 caused the low spatial variability in b_s.

At KAN-U, the transect lengths and area coverage differed greatly between May 2013 and May 2017 (Figures 7a, b). The 2013 survey covered an area of 1 km^2 with a transect length of more than 15 km. In 2017, our radar surveys were approximately 11 km in length resulting in extrapolated area coverage of 1.8 km^2. The average winter accumulation for the entire area are at 272 kg/m^2 ($\sigma = 20$ kg/m^2) in 2013 and 253 kg/m^2 ($\sigma = 19$ kg/m^2) in 2017.

The box plots in Figure 7c demonstrate a more variable data distribution at KAN-U than at the other two sites. The IQR for extrapolated b_s in 2012/13 is at -6 to $+5\%$ around the arithmetic mean. In 2016/17, the IQR decreases to $\pm 4\%$ around the mean. For both years, the whiskers reach outside the error margins of $\pm 10\%$ and, consequently, indicate less than $p = 0.9$ of data being within the error margins at KAN-U ($p = 0.82$ for 2012/13 and 2016/17). Accumulation data of 2016/17 has a higher skewness of 0.37 in comparison to 2012/13 (skewness of 0.17). Similar to the recorded radar data (see Figure 3), the upper quartile in b_s is right-shifted towards higher values. This is due to the homogeneous peak in accumulation at the northeastern corner of the grid (Figure 7b). Here, we measured above-average b_s, which consequently led to above-average interpolated values. The larger spatial heterogeneity in accumulation at KAN-U than at Dye-2 and Swiss Camp results in snow pits being slightly less representative of the surrounding area; only 80% of the respective May pit locations would provide area-wide b_s values being within a 10% error (for both accumulation seasons). Again, if snow depth is not averaged around pit locations the likelihood of representing area-wide b_s decreases to $p = 0.68$ (2012/13) and $p = 0.64$ (2016/17).

Not all of the recorded radar transect grids are ideal for the applied geostatistical analyses. The distances between radar lines at Dye-2 and KAN-U in May 2017 are too large to allow interpolation between the lines. We had limited time available for radar surveys, and we chose to focus on surveying larger areas (up to 20 km^2) instead of only surveying dense grids. The results presented in Figures 6b and 7a give us confidence that the data gaps do not include major dips or peaks in snow accumulation because no such inhomogeneities exist within the areas of good spatial coverage.

The above results imply that a point measurement of b_s (snow pit, upGPR value, neutron probe, etc.) is representative for an area of roughly 4x4 km^2 at Dye-2 with a probability of $p \geq 0.9$ and an uncertainty of $\pm 10\%$ in case snow depth is averaged. For KAN-U, the spatial variability is slightly higher and, consequently, there is less certainty about how well a single measurement represents the surrounding area. However, we consider a probability of $p \geq 0.8$ with uncertainty of $\pm 10\%$ for both study sites as a resilient estimate.

To quantitatively assess the benefit of snow depth measurements in addition to a snow pit, we numerically assume a sinusoidal snow depth variation with wavelengths of 56 m (arithmetic mean of the previously presented range in wavelength for the GPR transects) and average amplitude of ± 6.8 cm (the fluctuations in snow depth from arithmetic mean). Averaging multiple snow depths (with a sampling distance of 1 m) from a 20 m long probing transect, result in a maximum possibly measured offset in snow depth of -20% (amplitude decreases to 5.4 cm). A 10 m long probing line reduces the maximum offset by -6% compared to single point measurements (6.4 cm amplitude). A 30 m long snow probing line, however, result in a decrease of maximum possible offsets by -44% (3.8 cm amplitude). An additional cross line of probings will further decrease offsets. Only if the surface features are aligned symmetrically in both probing directions, the maximum offset derived from both lines will theoretically remain stable. For a measured snow pit with $\rho_s = 350$ kg/m^2 and $\overline{L_s} = 1$ m, the combined regional uncertainty ($\pm 5\%$ density uncertainty, ± 6.8 cm snow depth variation) reduces from a single point measurement with $b_s = 350 \pm 42$ kg/m^2 to a maximum possible uncertainty of $b_s = 350 \pm 35$ kg/m^2 for just a single 20 m probing line. These numerical results confirm values for representativeness derived from geostatistical extrapolation. Hence, we recommend to combine a larger number of snow-depth probings within an area of at least 20 m by 20 m in the vicinity of the pits to increase the regional representativeness. Regional snow density variations of $\pm 5\%$ can be accepted if snow depth uncertainty is minimized. Snow probing lines can easily be performed with respectively low time consumption compared to multiple snow pits. In particular, the wind-induced surface roughness has to be accounted for to provide spatially-representative b_s values.

Averaging radar traces within 1 m radius results in a pit size of roughly 3 m^2. This is slightly too big for conventional pits with on average 1 m snow depth. However, the search radius is related to the horizontal data resolution of the radar traces and had to be further increased for the KAN-U site in 2012/13.

3.3 Interannual changes in accumulation patterns

At KAN-U only 0.16 km^2 were covered during both radar acquisitions and, consequently, we do not investigate changes in accumulation for spring 2013 and 2017. For Dye-2, we recorded radar transects for two consecutive winter accumulation seasons. However, multi-year intersecting radar transects and, hence, spatially-consistent area-wide b_s estimates are reduced. The intersecting area at Dye-2 comprises roughly 1.7 km^2. Here, we observe a slight trend in the north - south direction for both accumulation seasons (Figure 6a and b). While the most southerly parts of the transect show above area-wide average b_s values, the northern fringes are below the arithmetic mean of the area in b_s. However, for both years the trends (in north to south direction) are statistically non-significant and very low at 5 kg/m^2 per 1 km for 2015/16 and 8 kg/m^2 per 1 km for 2016/17. The respective coefficients of determination of accumulation with latitude are very low as well ($R^2 = 0.15$ for 2015/16 and $R^2 = 0.25$ for 2016/17). The parallel stripes, mainly visible in Figure 6b for the southern parts, are certainly artifacts provoked

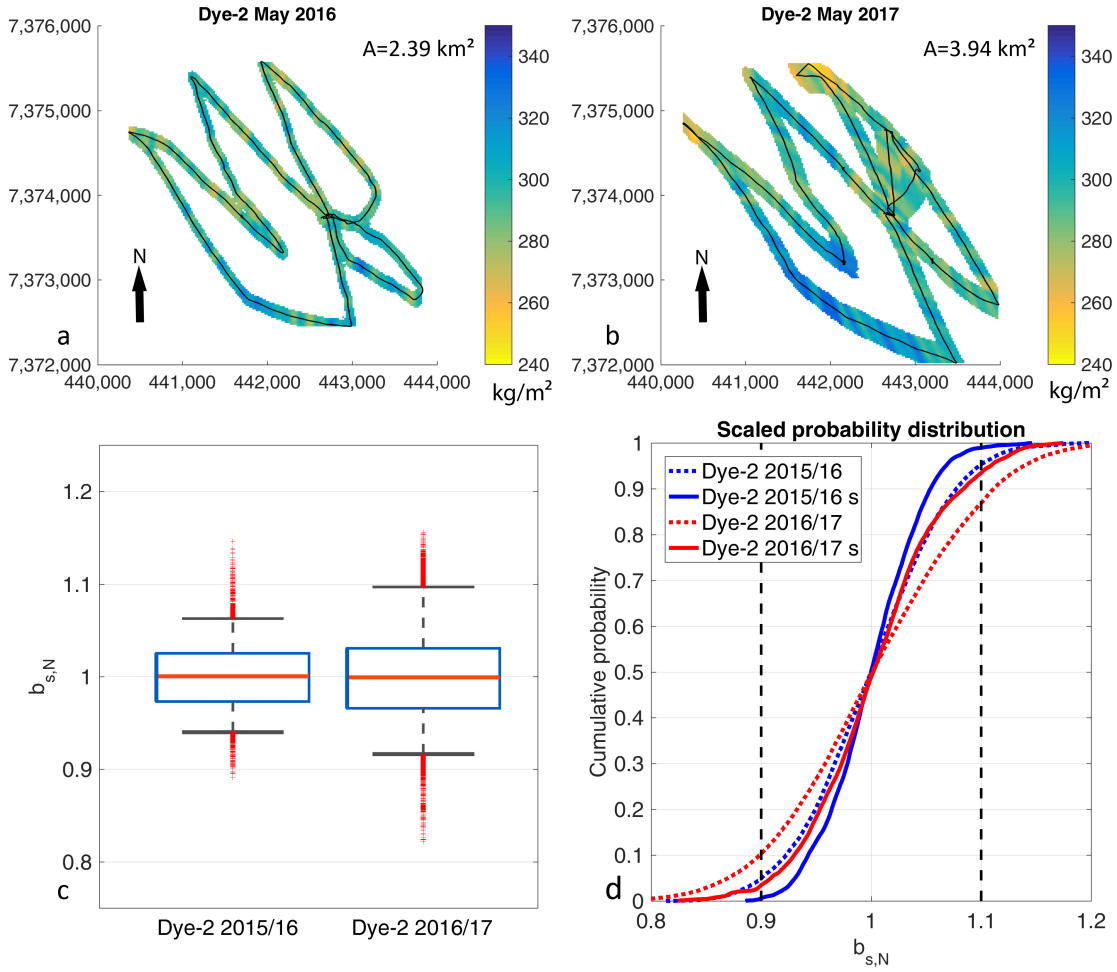

Figure 6. (a),(b) Kriging results for the radar transect at Dye-2 for snow accumulation. The black lines show the recorded radar traces and the black arrows indicate geographic north. A is the total area covered by extrapolation. (c) Box plots displaying normalized data distribution ($b_{s,N}$) of kriged outputs with the red horizontal line showing the median, the blue box framing the interquartile range and the whiskers displaying the 5% and 95% percentiles. Outliers are shown as red crosses. (d) Normalized cumulative probability distribution of radar-derived b_s within any subset with 1 m radius of the GPR transects. The dashed vertical lines represent the determined error margins of $\pm 10\%$. Compared are filtered (solid lines and letter s) with non filtered (dotted lines) radar transects. All map coordinates are given in UTM zone 23N with datum WGS1984.

by the grid design and the applied kriging. Local maximums in regular distances (150 – 220 m) occur along the transect line, however, the spatial extrapolation of these features is impossible due to the applied radar grid.

To quantitatively assess agreement in accumulation patterns, we used the respective normalized accumulation data and calculated the quotient. The cumulative data distribution of the quotients is presented in Figure 8. A constant area-wide quotient

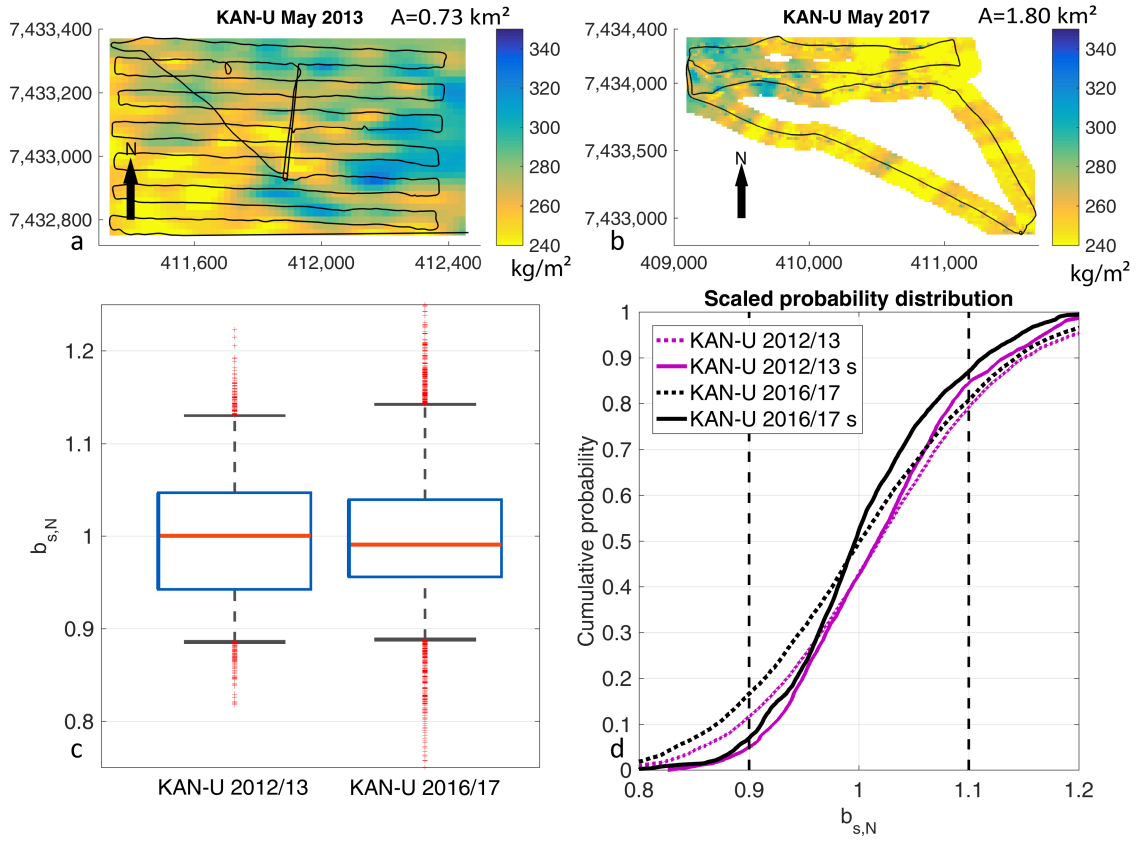

Figure 7. (a),(b) Kriging results for the radar transect at KAN-U for snow accumulation. The black lines show the recorded radar traces and the black arrows indicate geographic north. A is the total area covered by extrapolation. (c) Box plots displaying normalized data distribution ($b_{s,N}$) of kriged outputs with the red horizontal line showing the median, the blue box framing the interquartile range and the whiskers displaying the 5% and 95% percentiles. Outliers are shown as red crosses. (d) Normalized cumulative probability distribution of radar-derived b_s within any subset with 1 m radius (3 m radius for 2012/13) of the GPR transects. The dashed vertical lines represent the determined error margins of $\pm 10\%$. Compared are filtered (solid lines and letter s) with non filtered (dotted lines) radar transects. All map coordinates are given in UTM zone 23N with datum WGS1984.

of 1 would imply that the normalized accumulation patterns are exactly equal. For Dye-2, the probability of data being equally distributed in May 2016 and 2017 with a given uncertainty of $\pm 10\%$ is $p \geq 0.95$, meaning all intersecting locations of the accumulation pattern in two consecutive years at Dye-2 are similar.

3.4 Temporal changes in accumulation at Dye-2

370 During snow pit measurements in May 2016, we placed target reflectors at the EMS 2015 surface in each pit. These targets appear as hyperbolas in the radar data and make it possible to unambiguously identify that specific EMS for every subsequent

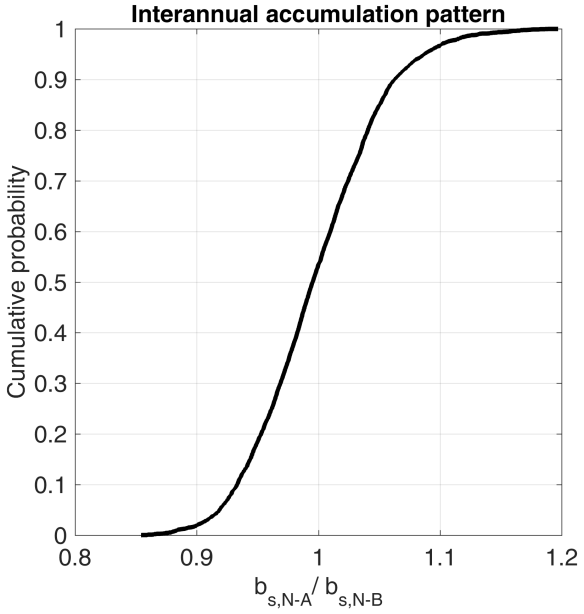

Figure 8. Interannual accumulation pattern comparison for intersecting transects at Dye-2. $b_{s,N-A}$ corresponds to normalized b_s for winter accumulation 2015/16 and $b_{s,N-B}$ to 2016/17.

radar campaign. We identified several targets in the May 2017 radar data. Hence, it is possible to detect changes in b_s that occurred between May 2016 (the last radar campaign) and the end of 2016 melt season (i.e. the start of the 2016/2017 accumulation season). However, such an analysis is only possible for intersecting areas of subsequent radar campaigns, which is 1.7 km^2 at Dye-2. The area for which both the summer 2015 and summer 2016 IRHs could be clearly identified decreases to 0.76 km^2. Ice movement contributes to uncertainties as well. Identical locations in May 2016 and May 2017 do not represent identical snow and firn layers, since we observed horizontal ice movement of 25 m at the upGPR location.

Instead of snow pit data, we used a firn core (drilled in May 2017) to determine the density of the layer between the 2015 and 2016 IRHs and to derive accumulation from TWT data as described in Equations 1 and 2. The firn between the 2015 IRH and the 2016 IRH is the net accumulation (accumulation minus meltwater percolation) between EMS 2015 and EMS 2016 ($b_{s,net}$), whereas the radar data collected in May 2016 is the winter accumulation, from EMS 2015 to May 2016. The changes that occurred over summer 2016 $\Delta b_s = b_{s,net} - b_s$ are the subtraction of the winter accumulation from the net accumulation for area intersections. The mean $\overline{\Delta b_s}$ for the intersecting transect areas (Figure 9a) for summer 2016 is 51 kg/m^2 with a standard deviation of 21 kg/m^2. Figure 9 documents the wide range in the data distribution. The negative values in Figure 9a occur only for six pixels and are likely artifacts.

Data distribution for $b_{s,net}$ is shown in Figure 9b as normalized values. Here, the distribution is less narrow than the winter accumulation in May 2016 (Figure 6d, blue line). Within the uncertainty margins, b_s decreases from $p = 0.99$ to $p = 0.88$ after one summer season. During summer 2016, melt caused a seasonal mass flux of 56 kg/m^2 into firn below EMS 2015 at

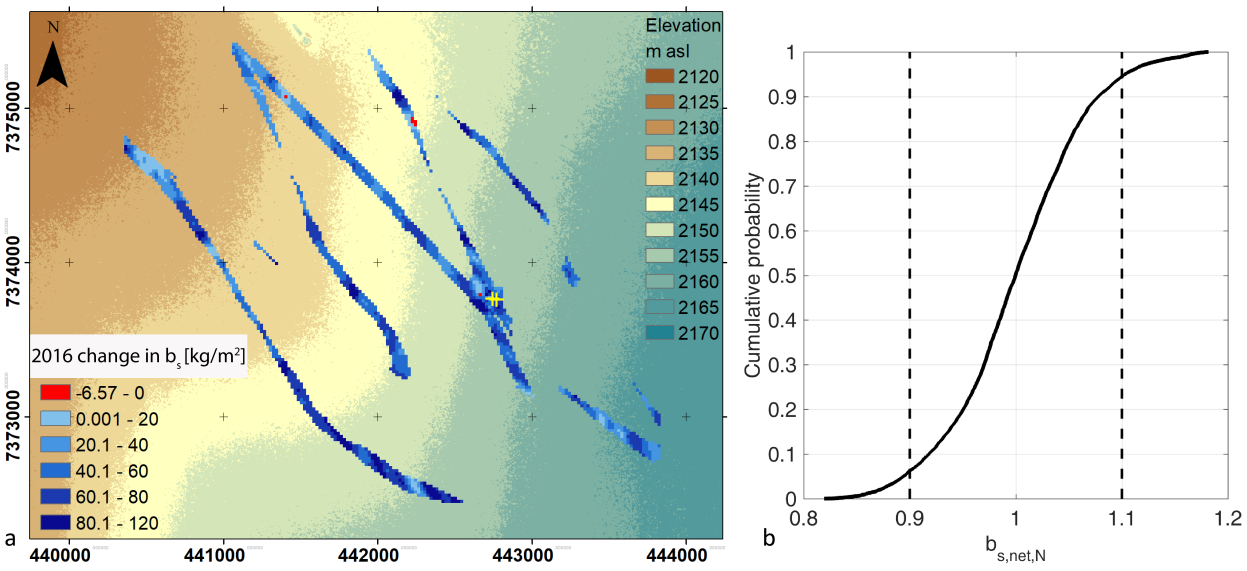

Figure 9. (a) Determined changes in accumulation (b_s) from May 2016 until September 2016. The yellow crosses show locations of the upGPR. The background displays a 2 m DEM (Porter et al., 2018) with 5 m contour color coding starting at 2120 m a.s.l. (brown color) and reaching 2165 m a.s.l. (blue color). The coordinates are given in UTM with datum WGS1984. (b) Cumulative probability distribution of the normalized net accumulation ($b_{s,net,N}$) of the layer between end-of-melt-season 2015 and end-of-melt-season 2016 layers.

the upGPR site (Heilig et al., 2018). It is likely that the seasonal mass flux is not homogeneous over the investigated area. In addition, the increased variability is in part due to mismatches in co-locating transects due to the ice movement. However, the mean change in b_s during summer 2016 corresponds almost exactly with observations derived from the upGPR (Heilig et al., 2018), which is 50.9 kg/m^2 from 01 May 2016 until the end of the melting period. This may be a coincidence or a confirmation of the benefits of upGPR, which averages a surface area of up to 10 m^2 compared to 1–3 m^2 area of a snow pit.

We cannot identify trends in b_s over the summer melt in 2016 associated with elevation; there are large differences within the same elevation band (Figure 9a). This implies that (i) no lateral redistribution of mass can be observed at Dye-2 during snow and firn melt and (ii) that melt and seasonal mass fluxes are much more inhomogeneous than accumulation distribution. These conclusions support the assumption made by Heilig et al. (2018) that in the current climate there is no systematic lateral mass redistribution during the melt season at Dye-2.

We also measured b_s in snow pits near the upGPR at Dye-2 in May 2018 and 2019. Although accumulation measured in May 2016 and May 2017 was very similar, the 2018 and 2019 data deviate strongly (Table 4). In 2018, b_s was more than 20% higher than in the previous two accumulation seasons. The accumulation measured in May 2019 was the lowest of the four years by a significant margin: 40% lower than the previous season and 23% lower than the next-lowest season (2017). This interannual accumulation variability is larger than the $\pm 10\%$ uncertainty in how well a b_s point measurement can be derived from radar data and usually represents the surrounding area. In agreement with Koenig et al. (2016), we conclude that annual or

Table 4. Winter snow accumulation (b_s) and snow density (ρ_s) measured in spring snow pits for Dye-2 and KAN-U compared with determined area wide arithmetic means.

Location date	b_s snow pit [kg/m^2]	$\overline{b_s}$ kriged area [kg/m^2]	$\overline{\rho_s}$ [kg/m^3]
Dye-2 May 2016	313	293	320
Dye-2 May 2017	294	296	334
Dye-2 May 2018	372	—	361
Dye-2 May 2019	225	—	364
KAN-U May 2013	319	271	358
KAN-U April 2017	246	252	316

more frequent density and b_s observations are necessary to estimate mean accumulation rates per region correctly. When snow depth is measured and averaged over an area of roughly 20×20 m^2, the value provides a reliable estimate of accumulation on regional scales of 1–20 km^2. Such data can be used for airborne radar campaigns and for validation of RCM simulations.

4 Conclusions

This study investigated how representative single point observations of b_s, such as snow pits, are for the surrounding 400 m^2 to 4 km^2 large areas. We used GPR to track IRHs created by summer melt surfaces along transects at three sites on the southwestern GrIS over the course of several field seasons. We derived maps of snow accumulation variability and compared them to snow pit and upGPR measurements. We found an uncertainty in radar-derived accumulation of 7–8%, which results from neglecting density variations along the radar transect and from applying a smoothing algorithm to minimize surface variability and layer-picking errors. In addition, we investigated the persistence of spatial patterns in accumulation over consecutive years and the influence of melt on an annual firn layer.

At all three sites, we found that point measurements such as snow pits represent the average b_s well over the study areas. A randomly selected snow pit location at any of the three sites would provide b_s values for the surrounding area (i.e. within $\pm 10\%$ of the areal mean) with a probability of $p = 0.8$ (KAN-U May 2013) to $p > 0.95$ (Swiss Camp May 2015 and Dye-2 May 2016). These likelihoods are independent of the size of investigated areas. However, not measuring and averaging snow depth over an area of at least 20×20 m decreases the probability of hitting arithmetic means by at least 10%. Snow-density variability is usually below $\pm 5\%$ on regional scales (1–4 km^2), while snow depth can vary significantly because of surface features such as dunes and sastrugi with various wavelengths ranging from submeters up to 60 m and more.

Our results suggest that there is only little change in accumulation patterns at Dye-2 for spring 2016 and 2017. However, the data only span two consecutive accumulation seasons that were almost identical in average density and accumulation. As such, we cannot confirm whether such persistence might be observed in seasons with significantly more or less accumulation or at different sites; this is a topic for future work.

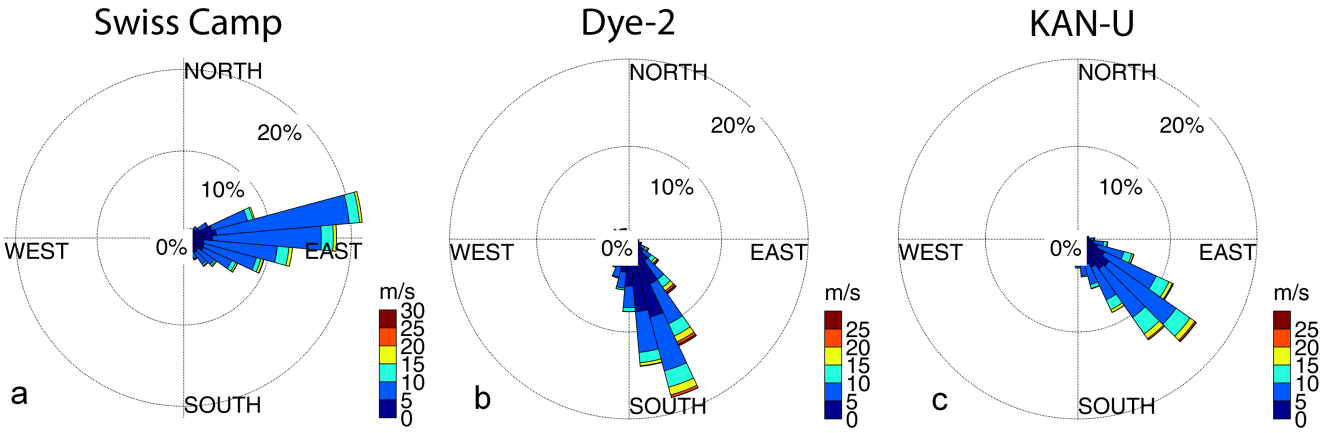

Figure A1. Prevailing wind distribution at Swiss Camp (a), Dye-2 (b) and KAN-U (c).

We also investigated the mass change that an accumulation layer (end of melt season to May) undergoes during the summer melt season using the GPR-transect data and continuous melt and accumulation observations from upGPR. We conclude that temporal changes in firn layer mass detected by the upGPR are representative of larger ($\sim$1 km^2) areas at Dye-2. We did not detect any patterns in summer melt along flowlines, suggesting that lateral meltwater flow at Dye-2 is not significantly redistributing mass. However, this could change with future warming in Greenland, which would influence data interpretation significantly of point measurements (AWS data, snow pits) and regional predictions by RCM and remote sensing.

This study aims to close the gap between point observations of b_s, which are meter scale, and remote-sensing data and RCMs, which have pixel sizes of $\sim 1 - 20$ km. We have shown that snow accumulation in the regions surrounding the three sites of the southwestern GrIS can be estimated well by point measurements as long as the snow depth is not influenced by surface roughness. To minimize such roughness effects, it is essential to determine the average snow depth over an area of several square meters. Ideally, snow-depth determinations— either directly via probings or derived from GPR transects— comprise spacings in between single points smaller than the characteristic length of the features and have an extent larger than the wavelengths of the features. Our analyses suggest that snow density does not vary greatly over km scales, and as such a single density measurement with numerous probed depths can suffice. Because interannual variability in accumulation can be significant, field measurements are essential for validating RCM predictions and remote sensing products.

Data availability. All GPR transects will become available on public data bases within the end of 2019. If needed earlier, the data are available from the lead author upon request. All other parameters are presented within this manuscript.

Competing interests. The authors declare that they have no conflict of interest

Acknowledgements. A. Heilig was supported by DFG grant (HE 7501/1-1). M. MacFerrin and C.M. Stevens were supported by the National Aeronautics and Space Administration (NASA) grant NNX15AC62G. We highly acknowledge support in logistics and preparation of the field campaigns from K. Young, T. Sheeley and staff from Polar Field Services. We appreciate the support by Christoph Mayer and the Geodesy and Glaciology group at the Bavarian Academy of Sciences and Humanities who provided the radar equipment in 2017 and had numerous suggestion and recommendations improving this work. We thank B. Gerling, L. Gambal, S. Samimi, T. Snow and S. Marshall for their assistance during field seasons.

References

Bennartz, R., Fell, F., Pettersen, C., Shupe, M. D., and Schuettemeyer, D.: Spatial and temporal variability of snowfall over Greenland from CloudSat observations, Atmospheric Chemistry and Physics, 19, 8101–8121, https://doi.org/10.5194/acp-19-8101-2019, 2019.

Charalampidis, C., van As, D., Colgan, W. T., Fausto, R. S., MacFerrin, M., and Machguth, H.: Thermal tracing of retained meltwater in the lower accumulation area of the Southwestern Greenland ice sheet, Annals of Glaciology, 57, 1–10, https://doi.org/10.1017/aog.2016.2, 2016.

Dunse, T., Eisen, O., Helm, V., Rack, W., Steinhage, D., and Parry, V.: Characteristics and small-scale variability of GPR signals and their relation to snow accumulation in Greenland's percolation zone, Journal of Glaciology, 54, 333–342, https://doi.org/10.3189/002214308784886207, 2008.

Enderlin, E. M., Howat, I. M., Jeong, S., Noh, M.-J., van Angelen, J. H., and van den Broeke, M. R.: An improved mass budget for the Greenland ice sheet, Geophysical Research Letters, 41, 866–872, https://doi.org/10.1002/2013GL059010, 2014.

Fahrmeir, L., Künstler, R., Pigeot, I., and Tutz, G.: Statistik: Der Weg zur Datenanalyse, Springer-Lehrbuch, Springer, Berlin and Heidelberg, 7. Auflage, korrigierter Nachdruck edn., 2011.

Farinotti, D., King, E. C., Albrecht, A., Huss, M., and Gudmundsson, G. H.: The bedrock topography of Starbuck Glacier, Antarctic Peninsula, as determined by radio-echo soundings and flow modeling, Annals of Glaciology, 55, 22–28, https://doi.org/10.3189/2014AoG67A025, 2014.

Graeter, K. A., Osterberg, E. C., Ferris, D. G., Hawley, R. L., Marshall, H. P., Lewis, G., Meehan, T., McCarthy, F., Overly, T., and Birkel, S. D.: Ice Core Records of West Greenland Melt and Climate Forcing, Geophysical Research Letters, 45, 3164–3172, https://doi.org/10.1002/2017GL076641, 2018.

Hawley, R. L., Courville, Z. R., Kehrl, L. M., Lutz, E. R., Osterberg, E. C., Overly, T. B., and Wong, G. J.: Recent accumulation variability in northwest Greenland from ground-penetrating radar and shallow cores along the Greenland Inland Traverse, Journal of Glaciology, 60, 375–382, https://doi.org/10.3189/2014JoG13J141, 2014.

Heilig, A., Eisen, O., MacFerrin, M., Tedesco, M., and Fettweis, X.: Seasonal monitoring of melt and accumulation within the deep percolation zone of the Greenland Ice Sheet and comparison with simulations of regional climate modeling, The Cryosphere, 12, 1851–1866, https://doi.org/10.5194/tc-12-1851-2018, 2018.

Howat, I. M., Negrete, A., and Smith, B. E.: The Greenland Ice Mapping Project (GIMP) land classification and surface elevation data sets, The Cryosphere, 8, 1509–1518, https://doi.org/10.5194/tc-8-1509-2014, 2014.

Humphrey, N. F., Harper, J. T., and Pfeffer, W. T.: Thermal tracking of meltwater retention in Greenland's accumulation area, Journal of Geophysical Research, 117, https://doi.org/10.1029/2011JF002083, 2012.

Karlsson, N. B., Eisen, O., Dahl-Jensen, D., Freitag, J., Kipfstuhl, S., Lewis, C., Nielsen, L. T., Paden, J. D., Winter, A., and Wilhelms, F.: Accumulation Rates during 1311–2011 CE in North-Central Greenland Derived from Air-Borne Radar Data, Frontiers in Earth Science, 4, D15 106, https://doi.org/10.3389/feart.2016.00097, 2016.

Khan, S. A., Aschwanden, A., Bjørk, A. A., Wahr, J., Kjeldsen, K. K., and Kjær, K. H.: Greenland ice sheet mass balance: A review, Reports on progress in physics. Physical Society (Great Britain), 78, 046 801, https://doi.org/10.1088/0034-4885/78/4/046801, 2015.

Koenig, L. S., Ivanoff, A., Alexander, P. M., MacGregor, J. A., Fettweis, X., Panzer, B., Paden, J. D., Forster, R. R., Das, I., McConnell, J. R., Tedesco, M., Leuschen, C., and Gogineni, P.: Annual Greenland accumulation rates (2009–2012) from airborne snow radar, The Cryosphere, 10, 1739–1752, https://doi.org/10.5194/tc-10-1739-2016, 2016.

Kovacs, A., Gow, A., and Morey, R.: The in-situ dielectric constant of polar firn revisited, Cold Regions Science and Technology, 23, 245–256, 1995.

Lewis, G., Osterberg, E., Hawley, R., Whitmore, B., Marshall, H. P., and Box, J.: Regional Greenland accumulation variability from Operation IceBridge airborne accumulation radar, The Cryosphere, 11, 773–788, https://doi.org/10.5194/tc-11-773-2017, 2017.

Lewis, G., Osterberg, E., Hawley, R., Marshall, H. P., Meehan, T., Graeter, K., McCarthy, F., Overly, T., Thundercloud, Z., and Ferris, D.: Recent precipitation decrease across the western Greenland ice sheet percolation zone, The Cryosphere, 13, 2797–2815, https://doi.org/10.5194/tc-13-2797-2019, 2019.

Miège, C., Forster, R. R., Box, J. E., Burgess, E. W., McConnell, J. R., Pasteris, D. R., and Spikes, V. B.: Southeast Greenland high accumulation rates derived from firn cores and ground-penetrating radar, Annals of Glaciology, 54, 322–332, https://doi.org/10.3189/2013AoG63A358, 2013.

Mosley-Thompson, E., McConnell, J. R., Bales, R. C., Li, Z., Lin, P.-N., Steffen, K., Thompson, L. G., Edwards, R., and Bathke, D.: Local to regional-scale variability of annual net accumulation on the Greenland ice sheet from PARCA cores, Journal of Geophysical Research: Atmospheres, 106, 33 839–33 851, https://doi.org/10.1029/2001JD900067, 2001.

Mottram, R., B. Simonsen, S., Høyer Svendsen, S., Barletta, V. R., Sandberg Sørensen, L., Nagler, T., Wuite, J., Groh, A., Horwath, M., Rosier, J., Solgaard, A., Hvidberg, C. S., and Forsberg, R.: An Integrated View of Greenland Ice Sheet Mass Changes Based on Models and Satellite Observations, Remote Sensing, 11, 1407, https://doi.org/10.3390/rs11121407, 2019.

Mouginot, J., Rignot, E., Bjørk, A. A., van den Broeke, M., Millan, R., Morlighem, M., Noël, B., Scheuchl, B., and Wood, M.: Forty-six years of Greenland Ice Sheet mass balance from 1972 to 2018, Proceedings of the National Academy of Sciences, 78, 201904 242, https://doi.org/10.1073/pnas.1904242116, 2019.

Mätzler, C.: Microwave permittivity of dry snow, IEEE Transactions on Geoscience and Remote Sensing, 34, 573–581, 1996.

Porter, C., Morin, P., Howat, I., Noh, M.-J., Bates, B., Peterman, K., Keesey, S., Schlenk, M., Gardiner, J., Tomko, K., Willis, M., Kelleher, C., Cloutier, M., Husby, E., Foga, S., Nakamura, H., Platson, M., Wethington, Michael, Jr., Williamson, C., Bauer, G., Enos, J., Arnold, G., Kramer, W., Becker, P., Doshi, A., D'Souza, C., Cummens, P., Laurier, F., and Bojesen, M.: ArcticDEM, https://doi.org/10.7910/DVN/OHHUKH [accessed on 14 March 2019], https://doi.org/10.7910/DVN/OHHUKH, 2018.

Proksch, M., Rutter, N., Fierz, C., and Schneebeli, M.: Intercomparison of snow density measurements: Bias, precision, and vertical resolution, The Cryosphere, 10, 371–384, https://doi.org/10.5194/tc-10-371-2016, 2016.

Rea, J. and Knight, R.: Geostatistical analysis of ground-penetrating radar data: A means of describing spatial variation in the subsurface, Water Resources Research, 34, 329–339, https://doi.org/10.1029/97WR03070, 1998.

Sasgen, I., van den Broeke, Michiel, Bamber, J. L., Rignot, E., Sørensen, L. S., Wouters, B., Martinec, Z., Velicogna, I., and Simonsen, S. B.: Timing and origin of recent regional ice-mass loss in Greenland, Earth and Planetary Science Letters, 333-334, 293–303, https://doi.org/10.1016/j.epsl.2012.03.033, 2012.

Savitzky, A. and Golay, M. J. E.: Smoothing and Differentiation of Data by Simplified Least Squares Procedures, Analytical Chemistry, 36, 1627–1639, https://doi.org/10.1021/ac60214a047, 1964.

Shepherd, A., Ivins, E. R., A, G., Barletta, V. R., Bentley, M. J., Bettadpur, S., Briggs, K. H., Bromwich, D. H., Forsberg, R., Galin, N., Horwath, M., Jacobs, S., Joughin, I., King, M. A., Lenaerts, Jan T M, Li, J., Ligtenberg, Stefan R M, Luckman, A., Luthcke, S. B., McMillan, M., Meister, R., Milne, G., Mouginot, J., Muir, A., Nicolas, J. P., Paden, J., Payne, A. J., Pritchard, H., Rignot, E., Rott, H., Sørensen, L. S., Scambos, T. A., Scheuchl, B., Schrama, Ernst J O, Smith, B., Sundal, A. V., van Angelen, Jan H, van de Berg, Willem J, van den Broeke, Michiel R, Vaughan, D. G., Velicogna, I., Wahr, J., Whitehouse, P. L., Wingham, D. J., Yi, D., Young, D., and Zwally, H. J.:

A reconciled estimate of ice-sheet mass balance, Science (New York, N.Y.), 338, 1183–1189, https://doi.org/10.1126/science.1228102, 2012.

Sørensen, L. S., Simonsen, S. B., Forsberg, R., Khvorostovsky, K., Meister, R., and Engdahl, M. E.: 25 years of elevation changes of the Greenland Ice Sheet from ERS, Envisat, and CryoSat-2 radar altimetry, Earth and Planetary Science Letters, 495, 234–241, https://doi.org/10.1016/j.epsl.2018.05.015, 2018.

Steffen, K. and Box, J.: Surface climatology of the Greenland Ice Sheet: Greenland Climate Network 1995-1999, Journal of Geophysical Research: Atmospheres, 106, 33 951–33 964, https://doi.org/10.1029/2001JD900161, 2001.

Tercier, P., Knight, R., and Jol, H.: A comparison of the correlation structure in GPR images of deltaic and barrier–spit depositional environments, GEOPHYSICS, 65, 1142–1153, https://doi.org/10.1190/1.1444807, 2000.

van As, D., Fausto, R. S., and the PROMICE project team: Programme for Monitoring of the Greenland Ice Sheet (PROMICE): first temperature and ablation records, GEUS, Geological Survey of Denmark and Greenland Bulletin, 23, 73–76, www.geus.dk/publications/bull, 2011.

van den Broeke, M., Bamber, J., Ettema, J., Rignot, E., Schrama, E., van de Berg, W. J., van Meijgaard, E., Velicogna, I., and Wouters, B.: Partitioning recent Greenland mass loss, Science (New York, N.Y.), 326, 984–986, https://doi.org/10.1126/science.1178176, 2009.

van den Broeke, M. R., Enderlin, E. M., Howat, I. M., Kuipers Munneke, P., Noël, B. P. Y., van de Berg, W. J., van Meijgaard, E., and Wouters, B.: On the recent contribution of the Greenland ice sheet to sea level change, The Cryosphere, 10, 1933–1946, https://doi.org/10.5194/tc-10-1933-2016, 2016.

Vandecrux, B., MacFerrin, M., Machguth, H., Colgan, W. T., van As, D., Heilig, A., Stevens, C. M., Charalampidis, C., Fausto, R. S., Morris, E. M., Mosley-Thompson, E., Koenig, L., Montgomery, L. N., Miège, C., Simonsen, S. B., Ingeman-Nielsen, T., and Box, J. E.: Firn data compilation reveals widespread decrease of firn air content in western Greenland, The Cryosphere, 13, 845–859, https://doi.org/10.5194/tc-13-845-2019, 2019.

Vaughan, D., Comiso, J., Allison, I., Carrasco, J., Kaser, G., Kwok, R., Mote, P., Murray, T., Paul, F., Ren, J., Rignot, E., Solomina, O., Steffen, K., and Zhang, T.: Observations: Cryosphere, chap. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, pp. 317–382, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Velicogna, I., Sutterley, T. C., and van den Broeke, M. R.: Regional acceleration in ice mass loss from Greenland and Antarctica using GRACE time-variable gravity data, Geophysical Research Letters, 41, 8130–8137, https://doi.org/10.1002/2014GL061052, 2014.

Vernon, C. L., Bamber, J. L., Box, J. E., Van den Broeke, M. R., Fettweis, X., Hanna, E., and Huybrechts, P.: Surface mass balance model intercomparison for the Greenland ice sheet, The Cryosphere, 7, 599–614, https://doi.org/10.5194/tc-7-599-2013, 2013.

Webster, R. and Oliver, M. A.: Geostatistics for environmental scientists, Statistics in practice, Wiley, Chichester, 2nd ed. edn., http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10257638, 2007.