

Water Resources Research



TECHNICAL REPORTS: DATA

10.1029/2019WR024907

Special Section:

Advances in remote sensing, measurement, and simulation of seasonal snow

Key Points:

- Ground-penetrating radar surveys during SnowEx 2017 provide a high-resolution, spatially extensive, and accurate validation data set
- Radar-derived snow depths exhibit high correlation with three independent methods: snow probes, airborne lidar, and satellite stereo imagery
- Median differences between GPR and other approaches are between -1 and -10 cm, which is within published accuracies of these techniques

Supporting Information:

- Supporting Information S1

Correspondence to:

D. McGrath,
daniel.mcgrath@colostate.edu

Citation:

McGrath, D., Webb, R., Shean, D., Bonnell, R., Marshall, H.-P., Painter, T. H., et al. (2019). Spatially extensive ground-penetrating radar snow depth observations during NASA's 2017 SnowEx campaign: Comparison with In situ, airborne, and satellite observations. *Water Resources Research*, 55. <https://doi.org/10.1029/2019WR024907>

Received 1 FEB 2019

Accepted 25 SEP 2019

Accepted article online 11 OCT 2019

©2019. The Authors.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Spatially Extensive Ground-Penetrating Radar Snow Depth Observations During NASA's 2017 SnowEx Campaign: Comparison With In Situ, Airborne, and Satellite Observations

Daniel McGrath¹ , Ryan Webb^{2,3,4} , David Shean⁵ , Randall Bonnell¹, Hans-Peter Marshall⁶, Thomas H. Painter⁷ , Noah P. Molotch^{3,8} , Kelly Elder⁹, Christopher Hiemstra¹⁰, and Ludovic Brucker^{11,12}
¹Department of Geosciences, Colorado State University, Fort Collins, CO, USA, ²Department of Civil, Construction, and Environmental Engineering, University of New Mexico, Albuquerque, NM, USA, ³Institute of Arctic and Alpine Research, University of Colorado Boulder, Boulder, CO, USA, ⁴Center for Water and the Environment, University of New Mexico, Albuquerque, NM, USA, ⁵Civil and Environmental Engineering, University of Washington, Seattle, WA, USA, ⁶Department of Geosciences, Boise State University, Boise, ID, USA, ⁷Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA, ⁸Department of Geography, University of Colorado Boulder, Boulder, CO, USA, ⁹Rocky Mountain Research Station, USDA Forest Service, Fort Collins, CO, USA, ¹⁰Cold Regions Research and Engineering Laboratory, U.S. Army Corps of Engineers, Fort Wainwright, AK, USA, ¹¹Cryospheric Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA, ¹²Goddard Earth Sciences Technology and Research Studies and Investigations, Universities Space Research Association, Columbia, MD, USA

Abstract Seasonal snow is an important component of Earth's hydrologic cycle and climate system, yet it remains challenging to consistently and accurately measure snow depth and snow water equivalent (SWE) across the range of diverse snowpack conditions that exist on Earth. The NASA SnowEx campaign is focused on addressing the primary gaps in snow remote sensing in order to gain an improved spatiotemporal understanding of this important resource and to further efforts toward a future satellite-based snow remote sensing mission. Ground-penetrating radar (GPR) is an efficient and mature approach for measuring snow depth and SWE. We collected ~ 1.3 million GPR snow depth observations during the NASA SnowEx 2017 campaign, yielding a spatially extensive (~ 133 -km total length) and high-resolution (~ 10 -cm lateral spacing) validation data set to assess various remote sensing and modeling approaches. We found high correlation between the GPR and manual snow probe derived snow depths ($r = 0.89$, $p < 0.0001$, root-mean-square error (RMSE) = 18 cm), but a median difference of -10 cm, which we attribute, in part, to probe penetration into the unfrozen subsurface. We also compared GPR-derived snow depths to two other independent estimates of snow depth, as an example of how this data set can be used for validation of remote sensing techniques: Airborne Snow Observatory lidar-derived snow depths ($r = 0.90$, $p < 0.0001$, median difference = -1 cm, RMSE = 14 cm) and preliminary DigitalGlobe WorldView-3 satellite-derived snow depths ($r = 0.70$, $p < 0.0001$, median difference = -3 cm, RMSE = 24 cm).

1. Introduction

Seasonal snow is a fundamental component of Earth's hydrologic cycle and climate system that also provides critical water resources to ~ 1.2 billion people globally (Barnett et al., 2005). Increasing atmospheric temperatures and changing circulation patterns have driven spatially variable reductions in total accumulated snow and the duration of the snow season (Clow, 2010; Luce et al., 2013; Mote et al., 2018; Stewart et al., 2005), thereby stressing ecological and human systems that rely on this resource. Despite the demonstrated importance and value of snow (Sturm et al., 2017), it remains challenging to consistently and accurately measure snow water equivalent (SWE) across the diverse range of snowpack conditions and landscapes that exist on Earth, at the spatial scales and temporal resolutions required for water management.

The NASA SnowEx program is focused on addressing the primary gaps in snow remote sensing to further efforts toward a future satellite-based snow remote sensing mission (Durand et al., 2018). The three primary techniques for snow depth or SWE retrieval being tested during the NASA SnowEx mission are (i) *depth*-

based, differencing “snow-off” and “snow-on” digital elevation models from lidar, stereo photogrammetry, and Ka-band interferometric synthetic aperture radar (InSAR) altimetry; (ii) *volume-based*, using active or passive signals in the microwave frequency range (2–89 GHz); and (iii) *InSAR phase-based*, which uses repeat differential L-band (1.26 GHz) InSAR to estimate changes in radar travel time through snow. Ground-based in situ measurements, such as ground-penetrating radar (GPR), play a critical role in calibrating and validating the performance of these different methods.

GPR is an efficient approach for quantifying spatiotemporal patterns in snow distribution across the landscape (e.g., Heilig et al., 2015; Holbrook et al., 2016; Lundberg et al., 2010; Marchand et al., 2003; Marshall & Koh, 2008; Mitterer et al., 2011; McGrath et al., 2018; Sand & Bruland, 1998; Schmid et al., 2014; Webb, 2018). This method measures the two-way travel time (*twt*) for an electromagnetic wave traveling from the antennas, through the snow, to various reflectors in the subsurface. The travel time depends on the distance traveled (layer thickness or snow depth), and the velocity of the electromagnetic wave. In a dry snowpack radar velocity only depends on density, and *twt* is typically converted to snow depth using a density estimate via an empirical relationship (e.g., Kovacs et al., 1995) or a calibration using coincident radar *twt* and snow depth measurements. Radar-estimated snow depth can subsequently be converted to SWE using observed or modeled snow densities (Sturm et al., 2010). In a snowpack containing liquid water, the conversion of *twt* to depth is more complicated, as even small volumes of liquid water content (LWC) significantly change the dielectric properties and thus the radar velocity (e.g., Bradford et al., 2009; Lundberg & Thunehed, 2000). The snow conditions during this study were primarily dry.

Here we present a spatially extensive GPR snow depth data set collected during the NASA SnowEx 2017 campaign at Grand Mesa, Colorado and compare GPR-derived snow depths to three independent snow depth observations: from manual probes, airborne lidar from the Airborne Snow Observatory (ASO), and stereo photogrammetry from DigitalGlobe WorldView-3 satellite imagery. As there are orders of magnitude more snow depth observations than SWE observations from the SnowEx 2017 campaign, and currently available remote sensing products measure depth rather than SWE, these comparisons focus on snow depth to minimize additional density assumptions in the snow depth-to-SWE conversion. However, it is important to note that GPR retrievals of SWE are less sensitive to the required estimate of density (due to the inverse relationship between radar velocity and density), than are SWE estimates from depth observations (e.g., Marshall et al., 2005).

2. Field Site

2.1. Overview

Airborne and ground-based field activities during the SnowEx 2017 campaign occurred at Grand Mesa and Senator Beck Basin, located in Western and Southwestern Colorado, respectively. Here we focus on observations collected at Grand Mesa, which has a total area of ~1,300 km² and ranges in elevation from 3,000 to 3,400 m above sea level (Figure 1). The western end of the mesa is relatively flat and is dominated by shrubland steppe. To the east, shrubland steppe transitions to tree islands embedded in a subalpine meadow matrix, eventually becoming continuous forest at higher elevations. Engelmann spruce is by far the dominant tree species on Grand Mesa. The Mesa Lakes SNOTEL station (NRCS 622, 3,048 m above sea level) is typically snow-covered between late October and late May and reaches a median maximum SWE of 46.5 cm in mid-April (1981–2010). Typical snow distribution patterns show a positive west-to-east gradient in snow depth on the Mesa, along with evidence for significant snow redistribution by strong westerly winds, particularly in the more open western sector (Figure 1).

2.2. Context for SnowEx 2017 Campaign

During the 2017 campaign, snow depth at the Mesa Lakes SNOTEL station ranged between 125 and 140 cm, while SWE increased from 40 to 43 cm between 7 and 17 February. The snowpack was above normal for this time of year, as median SWE on 7 February (1981–2010) was 26 cm. Snow surface conditions across the Mesa were variable, but tended to be wind-blown and compacted at the surface in the more exposed western sector and uncompacted in the more forested eastern sector. There is extensive recreational snowmobile traffic on the Mesa, such that many transects were crossed by snowmobile tracks.

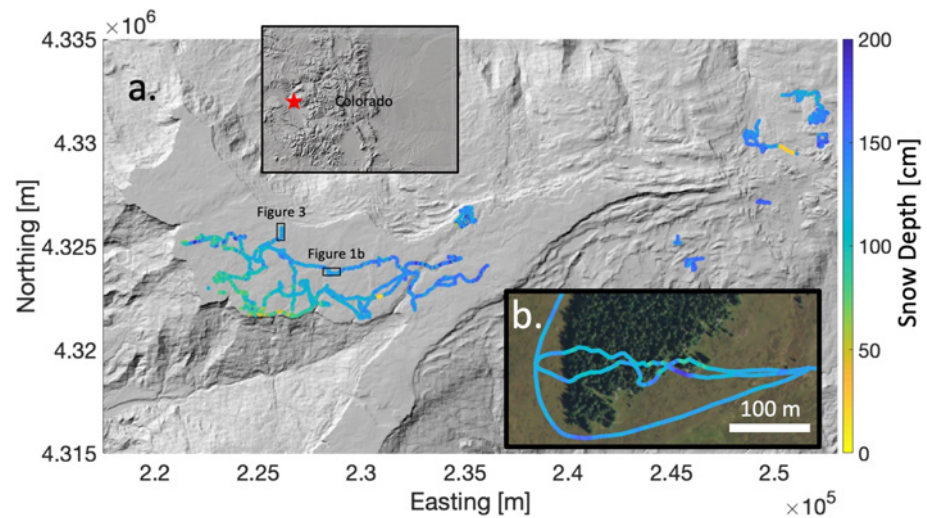


Figure 1. A snow depth derived from ground-penetrating radar surveys during SnowEx 2017 campaign overlaid on shaded relief map from 10-m National Elevation Dataset (NED) of Grand Mesa (EPSG:26913, NAD83 UTM 13 N). Inset: Shaded relief map of Colorado with Grand Mesa marked with a red star. Black outlines show the location of Figures 1(b) and 3. B detail of snow depth in the vicinity of a tree island on Grand Mesa.

The 2017 campaign was conducted in early February, in part to minimize the likelihood of melt and LWC in the snowpack. However, on 9 and 10 February, qualitative snow pit observations noted that the snowpack was damp or wet in the near surface layers, which was supported by temperature observations at or near 0 °C in the upper 10–20 cm (Figure S1). The impact of LWC on GPR velocity and derived snow depths is discussed in more detail in section 5.

3. Methods

3.1. Distributed in Situ Observations

SnowEx field observations at Grand Mesa occurred between 7 and 25 February 2017 and consisted of nearly 100 scientists collecting in situ observations (Brucker et al., 2017) and nine different sensors on five different aircraft (Kim et al., 2018). Here we focus on two core in situ observations: (i) snow depth measured roughly every 3 m using either a manual 3-m snow probe or a GPS equipped 1.2 m MagnaProbe (Sturm & Holmgren, 2018) along ~100 predefined ~300-m-long transects (Brucker et al., 2018) and (ii) snow density and snow temperature, measured at 10-cm vertical increments, in snow pits located ~30 m away from the majority of probing transects (Elder et al., 2018). We used a subset of these core observations collected between 7 and 17 February, resulting in >18,500 probed snow depths and 101 bulk (column-average) snow densities derived from snow pits (Figure 2(a)).

Geographic coordinates for each probe measurement (excluding MagnaProbe observations) were estimated by interpolating between georeferenced transect end points and any other midpoint poles placed along the transect (Hiemstra & Gelvin, 2018). Probe teams navigated between these poles using a combination of flagging installed in fall 2016 and consumer-grade GPS receivers. Geographic coordinates for observations close to the end points are more accurate, as these locations are known to better than 5 cm, and least accurate at locations farthest from poles/flagging.

3.2. Ground-Penetrating Radar

We collected GPR observations between 7 and 25 February 2017 at Grand Mesa using a Mala Geosciences GPR system with 1.6-GHz antennas. We mounted the GPR antenna in a plastic sled, which was either towed manually behind two individuals on skis or snowshoes (most common along forested transects) or a snowmobile (when transiting between transect lines or along transects out of forested areas). Here we present a subset of all GPR observations collected 7–17 February, as GPR surveys during this period focused on probed transects and continuous data collection between transects (Webb et al., 2019). Between 18 and 25 February,

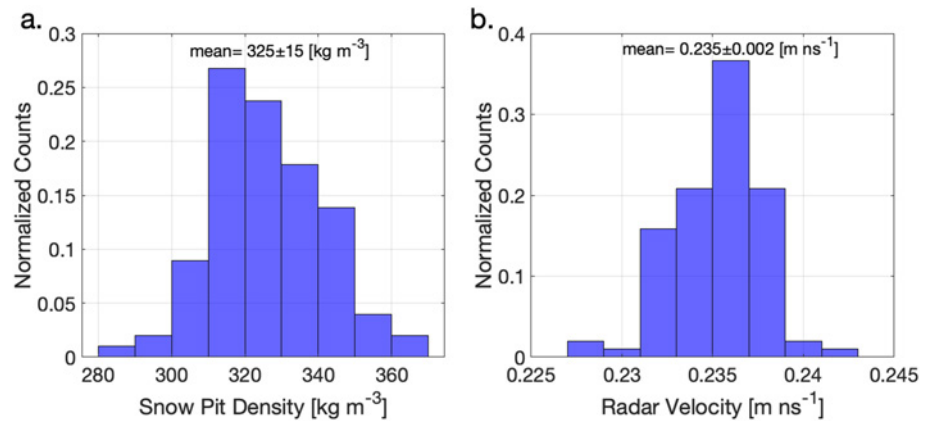


Figure 2. A histogram of snow pit densities collected between 7 and 17 February 2017. B histogram of empirically derived radar velocities derived from snow pit densities via estimates of dielectric constants using a standard empirical relationship (Kovacs et al., 1995).

GPR surveys primarily targeted terrestrial lidar survey locations and are not included here. GPR traces were geolocated with a consumer-grade (~3-m accuracy in open-sky conditions) GlobalSat BU-353-S4 GPS receiver.

Radargrams were processed by applying a time zero correction, dewow filter, linear gain, and a high-pass filter over a 30-trace window (Figure 3). Given the low surface slopes and consumer-grade GPS accuracy, the radargrams were not migrated. The GPR had an approximate ground footprint of 2 m² in typical conditions, a sample rate of 0.05 ns and an average trace spacing of ~0.15 m when pulled behind a snowmobile, and ~0.07 m when towed manually. The *twt* to the base of snowpack was semiautomatically picked using a phase-following algorithm, and converted to snow depth, *H_s*, as follows:

$$H_s = \left(\frac{twt}{2}\right) \cdot v_s, \quad (1)$$

where $v_s = c/\sqrt{\epsilon_s}$ is the velocity of electromagnetic waves in snow, $c = 3 \times 10^8$ m/s is the speed of light in a vacuum, and ϵ_s is the dielectric constant of snow. ϵ_s was calculated using an empirically derived relationship (Kovacs et al., 1995) based on the average observed snow pit densities (Elder et al., 2018). Radar velocity is not particularly sensitive to variations in density; for instance, an 18% difference in density (300 to 360 kg/m³) only results in a 4% difference in velocity. An alternative method for calculating v_s is described in section 5, where equation (1) is rearranged to solve for v_s directly using manual probe observed snow depths. However, given the circularity of this approach given subsequent comparisons to manual probes, it is only used illustratively.

For this analysis, we applied a temporal correction to align the GPR and remote sensing snow depth observations to a common date (corresponding to Airborne Snow Observatory (ASO) and WorldView data collections used in this analysis, respectively). This correction has not been applied to the archived data set (Webb et al., 2019). We derived this temporal correction from the mean net snow depth change over the elapsed interval between the observations (i.e., between the GPR surveys and ASO snow-on flight), as measured by arrays of ultrasonic snow depth sensors at two sites on Grand Mesa, which each contained 7–10 depth sensors spanning a range of canopy densities (Jennings et al., 2018). The median correction was –2 cm to make the GPR observations temporally consistent with the date of the ASO flight (max = 2 cm, min = –8 cm) and 2 cm for the date of the WorldView acquisition (max = 6 cm, min = –4 cm). We did not apply a correction to the GPR-derived snow depths prior to comparison with the snow probe data but we limited comparisons to those observations coincident ±two days. In order to compare GPR and probe derived snow depths, we first took the mean of all GPR snow depths within a 2-m radius around each probe location.

As noted previously, in situ probe observations were collected along 300-m-long transects, both perpendicular and parallel to the prevailing wind direction. To assess spatial variability in snow depths along these

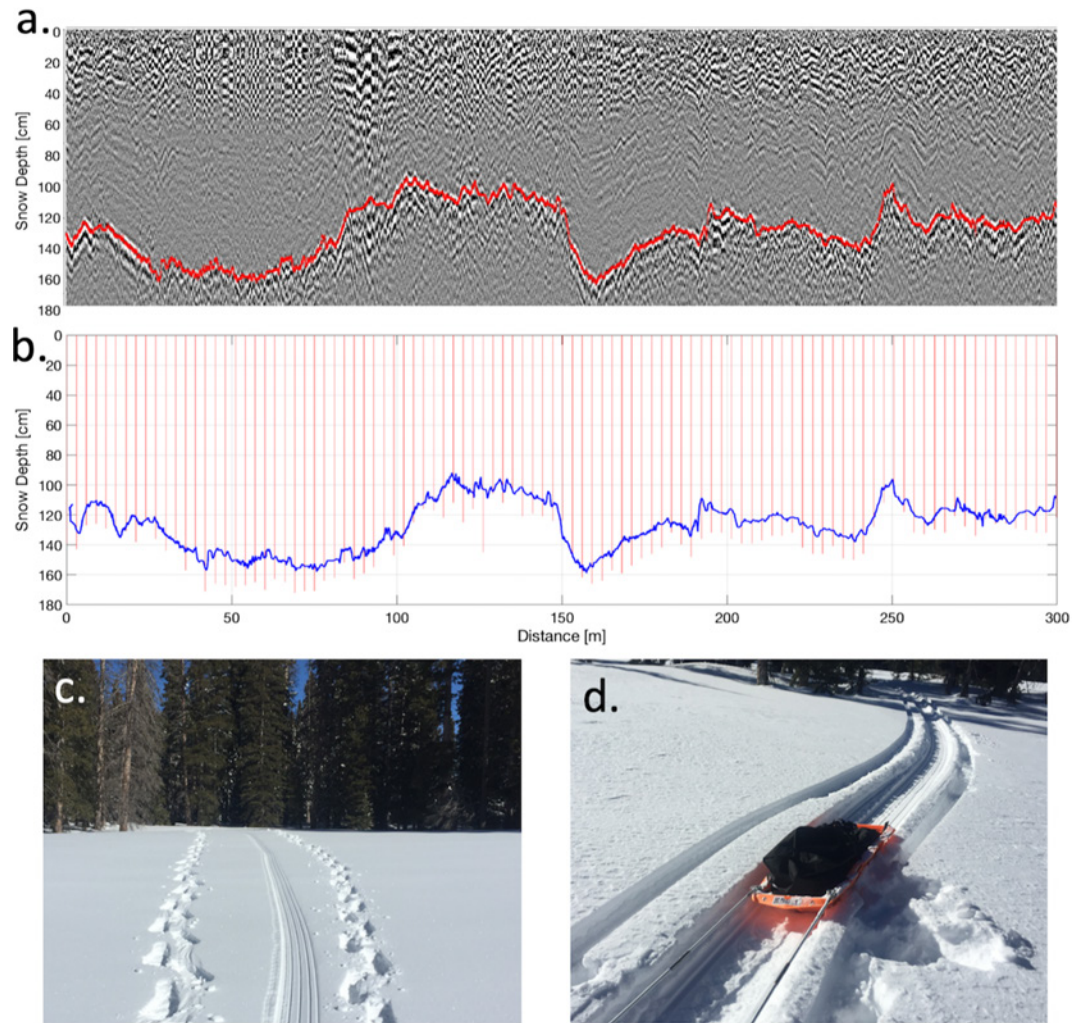


Figure 3. A processed radargram for transect 23. Red line shows semiautomated GPR picks. B GPR-derived snow depth (blue) and probe-derived snow depth (vertical red lines) for transect 23. C photo showing ~5-cm surface snow compaction when GPR sled was dragged manually. D photo showing ~10-cm surface snow compaction when GPR sled was dragged behind a snow machine.

transects, and potential differences in how this variability was quantified due to the lateral spacing of the observations (~0.1 m for GPR, 3 m for manual probes), we calculated variograms using

$$\gamma(h) = \left(\frac{1}{2N(h)} \right) \cdot \sum_{i=1}^{N(h)} (Hs_i - Hs_{i+h})^2, \quad (2)$$

where N is the number of pairs at lag spacing h and Hs are snow depths.

3.3. NASA's Airborne Snow Observatory

The Airborne Snow Observatory uses a Riegl Q1560 airborne laser scanner (1,064-nm wavelength) to measure surface elevations at different time periods. The lidar point clouds are filtered to preserve last returns, and “bare ground” digital terrain models (DTM) are produced to remove vegetation. Snow depths are calculated by differencing the snow-on DTM from the snow-off DTM (Deems et al., 2013; Hopkinson et al., 2004; Painter et al., 2016). For the analysis here, we utilize 3-m gridded snow-depth products (Painter, 2018) produced by differencing DTM products collected on 26 September 2016 (snow-off) and 8 February 2017 (snow-on). Previous work reported that ASO-derived snow depths had a mean absolute error of <8 cm and bias <1 cm (Painter et al., 2016); comparisons between ASO and ground-based lidar-derived snow depths at

Grand Mesa found a median difference of 5 cm and a mean absolute differences of 10 cm (Currier et al., 2019). To facilitate comparison, we gridded the GPR observations to match the ASO product using the median value of all GPR observations within a given cell. Only grid cells with 50 or more GPR observations were included in this comparison. To examine the influence of vegetation/forest canopy on ASO snow-depth retrievals, we parsed the analysis into two groups on the basis of ASO-derived vegetation height (<2 and > 2 m).

3.4. WorldView Stereo Digital Surface Models

Digital surface models (DSM) of Grand Mesa were generated from ~0.3-m DigitalGlobe WorldView-3 stereo satellite imagery using the NASA Ames Stereo Pipeline (ASP; Shean et al., 2016; Beyer et al., 2019). All individual WorldView DEMs were co-registered to the reference 2016-09-26 snow-off ASO lidar DSM using static control surfaces (e.g., bare ground identified in National Land Cover Database) using a robust iterative closest point (ICP) algorithm that solved for 3-D translations required to remove horizontal and vertical offsets (Shean et al., 2016, 2019). We then generated mosaics for WorldView DSMs acquired on the same date. A final ICP co-registration step removed any residual horizontal/vertical offsets between the reference snow-off WorldView DSM mosaic and the snow-on mosaic over static, snow-free, control surfaces. DSMs from 25 September 2016 (snow-off) and 1 February 2017 (snow-on) were posted at 8-m resolution and differenced to estimate snow depth (Shean, 2019). The extent of the 1 February imagery covered the western half of Grand Mesa. Previous work showed that WorldView-derived DSMs have relative accuracies of 10–50 cm over relatively shallow slopes (<10°; Shean et al., 2016). In contrast to the ASO DTM products, the Worldview products are DSMs (i.e., include vegetation heights), and thus, we limited comparisons to sectors of Grand Mesa without vegetation (i.e., where ASO-derived vegetation heights were equal to 0 m). Prior to comparing with the GPR observations, WorldView-derived snow depths were filtered to remove spurious outliers (snow depths <0 and >5 m). GPR observations were gridded as described for ASO, but given the coarser resolution of the preliminary WorldView DSM products, we used a minimum of 100 GPR observations per 8-m grid cell.

4. Results

4.1. Ground-Penetrating Radar Surveys and Comparison to in Situ Measurements

The mean ($\pm$ standard deviation) snow density measured in 101 snow pits between 7 and 17 February was 325 ± 15 kg/m³ (Figure 2(a)). From these bulk densities, we derived a mean ($\pm$ standard deviation) radar velocity of 0.235 ± 0.002 m/ns (Figure 2(b)), which we subsequently used to convert GPR-measured *twt* to snow depth. In total, the ~133 km of GPR profiles yielded more than 1.3 million point observations of snow depth across Grand Mesa (Figure 1(a)). GPR data collection is more efficient than manual snow probing, thus yielding greater spatial coverage and a more complete view of the overall west to east gradient of increasing snow depths. In addition, GPR-derived snow depths are at a nominally higher spatial resolution along the survey track (on the order of ~0.1 m for the GPR observations versus 3 m for manual probes in this campaign), thereby fully revealing meter-scale spatial variability in snow depths (e.g., effects of surface vegetation and wind redistribution; Figures 1(b) and 3) that could be potentially aliased at lower spatial sampling intervals. GPR-derived snow depths ranged from 16 to 330 cm across the mesa, with a mean and standard deviation of 128 ± 31 cm.

Errors in GPR-derived snow depths are due to uncertainty in *twt* (i.e., picking the air-snow interface and snow-ground interface) and the radar velocity, v_s . We assume that the uncertainties are independent and random, and thus, the uncertainties add in quadrature:

$$\sigma[H_s] = \sqrt{\left(\frac{\sigma[twt]}{twt}\right)^2 + \left(\frac{\sigma[v_s]}{v_s}\right)^2}, \quad (3)$$

where $\sigma[twt]$ is 0.2 ns (or four samples) and $\sigma[v_s]$ is 0.004 m/ns, which represents 1 standard deviation on either side of the mean (Figure 2(b)). This results in a fractional uncertainty of 0.03 or absolute uncertainty of ~4 cm, similar to previous studies (Marshall et al., 2005).

We used ~18,500 snow probe measurements made along ~100 transects (Brucker et al., 2018) during the subset of the SnowEx 2017 campaign data analyzed here (Figure 4(a)). Figure 3(b) illustrates the differences in

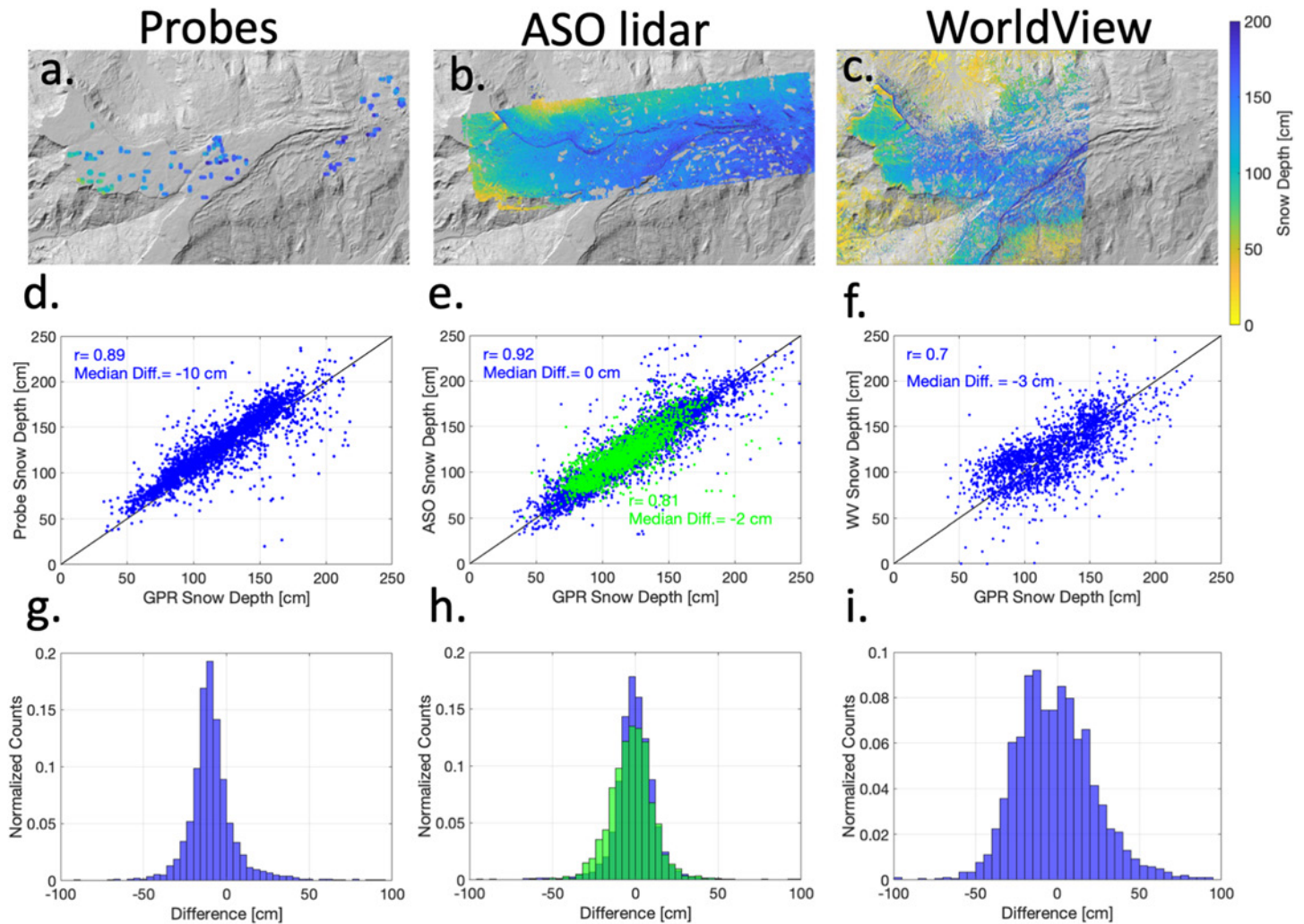


Figure 4. Snow depth derived from a manual probes, b ASO lidar DTMs, and c WorldView satellite DSMs. Scatterplots of snow depth from d manual probes, e ASO lidar DTMs, and f WorldView satellite DSMs compared to GPR-derived snow depths. Blue points in (e) and (f) have ASO-derived canopy heights <2 m, and green points have ASO-derived canopy heights >2 m. histograms of difference between (g) GPR-probes, (h) GPR-ASO, and (i) GPR-WorldView.

spatial sampling resolution of the GPR and probes along a 300-m transect. Along this profile, the probe-derived snow depths are deeper than the GPR-derived snow depths by ~15 cm. A comparison of all 2,823 temporally ($\pm$ two days) and spatially (within 2 m) coincident geolocated GPR snow depths and georeferenced probe depths (including both the manual probes and MagnaProbe observations) showed mean and median differences (GPR minus probe) of -9 and -10 cm, respectively (Figure 4(g)), with a correlation coefficient of $r = 0.89$, $p < 0.0001$, RMSE of 18 cm, and median absolute deviation (MAD) of 11 cm (Figure 4(d)).

As manual probing constitutes a primary field observation utilized by the SnowEx campaign and snow scientists more broadly, we sought to explore how differences in the lateral spacing between GPR (~0.10 m) and snow probes (~3 m) influenced assessments of spatial variability, as quantified with variograms. We computed variograms for 27 different transects with varying amounts of canopy cover (from 0 to 93% canopy, as determined using a 2-m threshold in the ASO-derived vegetation height product) using GPR and probe-derived snow depths independently. We subsequently took the mean of all variograms within two different canopy cover classes (0–25%, $n = 12$ and 25–100%, $n = 15$; Figure 5). Transects with <25% canopy cover had a significantly lower sill, or variance, than transects with >25% cover. Qualitatively, this is consistent with expectations, given that forest canopies can increase surface roughness due to downed trees, snow

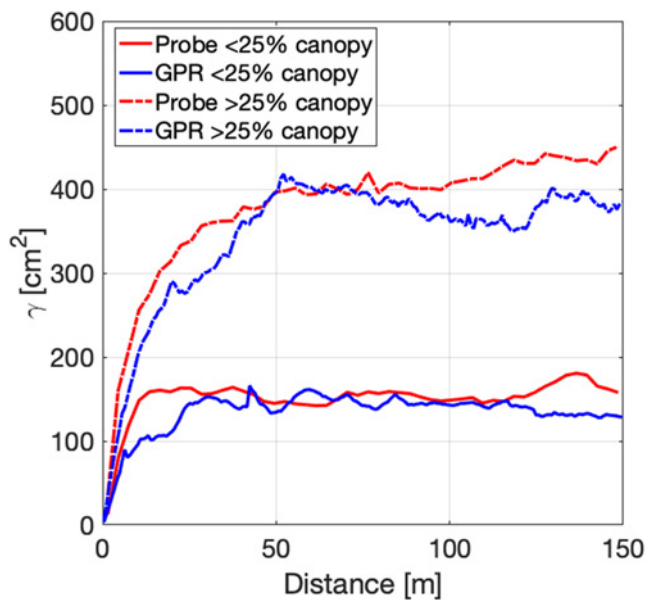


Figure 5. Mean variograms show spatial variability of observations along 27 transects ($n = 12 < 25\%$ canopy, $n = 15 > 25\%$ canopy).

canopy interception, and spatially varying wind redistribution patterns. Although the probe-based variograms have greater uncertainty (due to the relatively low number of pairs at a given lag distance; e.g., Webster & Oliver, 1992) than the GPR-based variograms, which have much greater sampling density, the two sets of variograms exhibit comparable features, with similar ranges (~ 10 m for $<25\%$ canopy and ~ 50 m for $>25\%$ canopy). On average, the probe-based variograms exhibited higher sills, or variances, likely reflecting the smaller footprint of these measurements (e.g., 1 cm^2 for an individual probe versus $\sim 2 \text{ m}^2$ for a single GPR trace).

4.2. Comparison to ASO and WorldView-Derived Snow Depths

As an example of the value of GPR for calibration and validation of snow remote sensing, we evaluated the accuracy of ASO and WorldView-derived snow depth products (Figures 4(b) and 4(c)) via a direct comparison with GPR-derived snow depths across Grand Mesa. There was strong agreement ($r = 0.92$, $p < 0.0001$, $\text{MAD} = 6 \text{ cm}$, $\text{RMSE} = 14 \text{ cm}$, mean and median differences of -1 and 0 cm) between ASO snow depths and GPR-derived snow depths (6,974 3-m grid cells) where canopy heights $< 2 \text{ m}$ (Figure 4(e)). The comparison for the 2,216 cells with $> 2 \text{ m}$ forest canopy showed similar agreement on average ($r = 0.81$, $\text{MAD} = 8 \text{ cm}$, $\text{RMSE} = 15 \text{ cm}$, mean and median differences of -2 cm), with lower correlation possibly caused by GPR geolocation errors under canopy. We

found moderate agreement ($n = 2,107$, $r = 0.7$, $p < 0.0001$, $\text{MAD} = 16 \text{ cm}$, $\text{RMSE} = 24 \text{ cm}$, mean and median differences of -1 and -3 cm) between WorldView snow depth estimates and GPR-derived snow depths (Figure 4(f)) for vegetation-free regions.

5. Discussion and Future Steps

GPR surveys during the SnowEx 2017 campaign yielded more than 1.3 million accurate individual snow depth observations, thereby providing a spatially extensive and high-resolution calibration/validation data set for airborne, satellite, and model-derived snow depths, at a fraction of the effort required for manual probing. This data set represents a more than 50-fold increase compared to the number of in situ depth observations collected as part of the extensive manual probing campaign, with a fraction of the field personnel, and is likely one of the most extensive validations of ASO-derived snow depths. Furthermore, although this analysis did not investigate InSAR phased-based observations (e.g., UAVSAR) due to the lack of changes in SWE between overflights during the 2017 campaign, GPR observations are very well suited to this task, as changes in phase from the L-band InSAR instruments should be equal to changes in ground-based GPR *twt* and thus GPR observations will be particularly valuable during the SnowEx 2020 campaign.

Across all comparisons, the GPR-derived snow depths showed the strongest agreement with the ASO snow depths where the canopy was $< 2 \text{ m}$ ($r = 0.92$, $\text{MAD} = 6 \text{ cm}$), followed by the manual probe snow depths ($r = 0.89$, $\text{MAD} = 11 \text{ cm}$). The agreement with ASO snow depths where the canopy was $> 2 \text{ m}$ was slightly reduced ($r = 0.81$, $\text{MAD} = 8 \text{ cm}$), which can be attributed to a number of difference sources, including a reduction in lidar point density on the ground due to canopy intercept (e.g., Deems et al., 2013), greater uncertainty in positional accuracy for the GPR observations, greater subpixel-scale spatial variability in snow depths (Figure 5), and the presence of subnivean void spaces beneath ground vegetation. The latter affects methods differently; the radar wave travels $\sim 25\%$ faster through the void space than through snow, thereby minimizing the impact, whereas lidar/WorldView would be unable to discriminate the void space from snow. Although the GPR-based assessment of the WorldView snow depths showed lower agreement (lower r value, larger RMSE) relative to the ASO comparison, it is important to note the major differences in platforms (aircraft versus satellite) and methodological maturity (operational versus development) between these approaches. Given these considerations, this preliminary analysis is quite encouraging and ongoing efforts to develop improved sensor corrections, stereo correlation/triangulation routines, and DSM co-registration methodology will yield higher-quality DSM products posted at $\sim 1\text{--}2 \text{ m}$ with improved spatial coverage (fewer data gaps) and horizontal/vertical accuracy.

The ASO and WorldView median snow depths differed from the median GPR-derived snow depths by -1 and -3 cm, although these differences are not statistically significant considering the respective uncertainties in each of these data sets. A larger median difference was observed between the GPR and probe-derived snow depths, where the probe depths exceeded the GPR snow depths by 10 cm. Likely contributing factors include probe penetration into an unfrozen ground surface and/or shallow vegetation (Sturm & Holmgren, 2018), uncertainty in the geolocation of the respective observations, differences in measurement footprints (square centimeter versus square meter, as noted in section 4.1), and the propensity for radar retrievals to be most sensitive to the shallowest depths within the radar footprint. Currier et al. (2019) found that ASO-derived snow depths during the SnowEx 2017 campaign were negatively biased by 6 cm compared to probe measurements, which they suggest may be due to probe penetration, thus potentially explaining a large portion of the observed difference in this study.

During field surveys, the radar sled/snowmobile compressed the surface snow on the order of ~ 5 to 10 cm, which presents an additional potential error source. Field photos (Figures 3(c) and 3(d)) suggest that the surface snow was compacted rather than displaced by the sled, and thus, neither the bulk SWE nor *twt* measured by the GPR would have been modified. In our analysis, we used a radar velocity based on the uncompressed snow densities (from snow pits), and thus, the reconstructed snow depths should not be affected by this compaction. If, however, some of the surface snow was displaced, this could be a contributing factor to the observed difference, as this typically low-density surface layer contributed proportionally more to the total snow depth (as measured by probes/ASO/WorldView) than to the total SWE (to which the measured radar travel time is sensitive). To examine the impact of sled compression further, we derived an additional radar velocity from a least squares regression between GPR-measured *twt* and coincident probe depths (Figure S2). This regression results in a slower radar velocity (0.228 m/ns) and *y* intercept of 14 cm. The slower velocity is in line with a higher mean snow density due to compaction and the *y* intercept likely reflects both the probe penetration into the subsurface discussed previously and surface snow compression by the radar sled. Snow depths derived from the two approaches (pit-derived velocity and regression-derived velocity) are nearly identical once the portion of the *y* intercept attributed to sled compression is accounted for. In summary, the observed median difference between probe and GPR-derived snow depths is likely the result of probe penetration into the subsurface and differences in measurement location/footprints/sensitivities. By using a radar velocity based on uncompressed snow densities, surface snow compression by the radar sled was not likely a major component of this difference. However, if using a radar velocity derived from direct comparisons to in situ snow depths, it is important to account for the effects of sled compression via the *y* intercept from the least squares solution.

As noted previously, both qualitative and quantitative (Figure S1) pit observations found the presence of LWC on 9 and 10 February 2017, which would potentially result in positively biased snow depths on these days, as the presence of liquid water lowers the radar velocity. Qualitative pit observations reported either moist or wet conditions (primarily in the upper 10–50 cm), which corresponds to 0–3 and 3–8% LWC by volume, respectively (Techel & Pielmeier, 2011). However, these qualitative observations carry considerable uncertainty, as many observers had limited experience with this type of subjective measurement, and many of the observations of wet snow coincided with snow temperatures significantly below 0 °C. Snow pit temperature measurements (Figure S1) did reach 0 °C at the surface and in the near surface on 9 and 10 February. Assuming a bulk LWC of 3% by volume and a density of 325 kg/m³, the corresponding radar velocity would be 0.196 m/ns (Roth et al., 1990). This would theoretically result in an $\sim 20\%$ overestimation in snow depth on these days. We evaluated the GPR-ASO snow depth differences by survey date and found GPR snow depths to be positively biased by ~ 3 cm relative to ASO depths on these days, suggesting that the impact of liquid water content on the overall assessment was minimal.

Future GPR field campaigns could be improved by a number of small modifications, including (i) multifrequency geodetic GNSS receivers to ensure that manual probes and GPR observations are spatially coincident; (ii) increased use of MagnaProbes, as the direct georeferencing facilitates comparison to other approaches; (iii) quantification of liquid water in the snowpack, given its strong impact on radar velocity (Bradford et al., 2009; Webb et al., 2018); and (iv) a sled mount that elevates the GPR antennas above the snow surface (Holbrook et al., 2016), thus eliminating surface snow compaction. However, such a design

is not without drawbacks as it requires two surfaces to be picked in the radargram, results in the loss of direct surface coupling of the instrument, and can make the surface more difficult to identify, especially after new snow, due to the smaller dielectric contrast at the air-snow interface.

6. Conclusions

The GPR data set presented here provides a spatially extensive and high-resolution data set that can be used to calibrate and validate other remote sensing observations from NASA's SnowEx 2017 campaign. In the comparisons shown here, we found strong agreement with ASO-derived snow depths and moderate agreement with preliminary WorldView-derived snow depths. GPR-derived snow depths agree well with traditional in situ manual probes, although probe depths were consistently greater than GPR depths, which we attribute to both penetration of the probes into the subsurface/soft vegetation and differences in measurement location/footprint/sensitivities. In summary, GPR is a powerful tool for efficiently measuring snow depths at high spatial resolution, thereby producing a valuable calibration/validation data set for other emerging methodologies (e.g., WorldView). In comparison to other approaches, GPR is significantly more efficient than manual probes, provides greater spatial coverage than terrestrial lidar, can be directly compared to InSAR phase-based approaches, and can be completed at a fraction of the cost of airborne lidar. GPR is not without limitations, including data collection solely along track and the dependency on density and wetness observations to constrain radar velocity, in addition to the time required for postprocessing. Despite these limitations, GPR surveys should be a priority for future snow field campaigns where cost-effective and spatially extensive coincident calibration/validation data sets are required, and especially for validation of techniques based on radar travel time, such as L-band InSAR.

Acknowledgments

D.M., R.W., R.B., and H.P.M. acknowledge NASA THP award 80NSSC18K0877 for the support. D.S. was supported by NASA THP award 80NSSC18K1405. Part of this work was performed at the Jet Propulsion Laboratory, California Institute of Technology under a contract with NASA. We thank the SnowEx leadership team and the ~100 scientists that volunteered during SnowEx 2017 campaign, without which this work would not have been possible. We also thank three anonymous reviewers and the Editor, Jessica Lundquist, for their thoughtful comments and suggestions that improved the manuscript. Data from the NASA SnowEx campaign, including the primary GPR data set described here, snow pit density profiles, and probe-derived snow depths, can be accessed at the National Snow and Ice Data Center: <https://nsidc.org/data/snowex>. ASO-derived snow depths from the SnowEx campaign are available here: <https://nsidc.org/data/aso/data-summaries>. The preliminary WorldView snow depth product is available at <http://doi.org/10.5281/zenodo.3381653>.

References

- Barnett, T., Adam, J., & Lettenmaier, D. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, 438, 303–309. <https://doi.org/10.1038/nature04141>
- Beyer, R., Alexandrov, O., & McMichael, S. (2019). NeoGeographyToolkit/StereoPipeline: ASP 2.6.2 (Version v2.6.2). Zenodo. <http://doi.org/10.5281/zenodo.3247734>
- Bradford, J. H., Harper, J. T., & Brown, J. (2009). Complex dielectric permittivity measurements from ground-penetrating radar data to estimate snow liquid water content in the pendular regime. *Water Resources Research*, 45, W08403. <https://doi.org/10.1029/2008WR007341>
- Brucker, L., Hiemstra, C., Marshall, H., & Elder, K. (2018). *SnowEx17 Community Snow Depth Probe Measurements, Version 1*. Boulder, Colorado USA: NASA National Snow and Ice Data Center Distributed Active Archive Center. <https://doi.org/10.5067/WKC6VFMT7JTF> January 30, 2019
- Brucker, L., Hiemstra, C. A., Marshall, H.-P., Elder, K., De Roo, R. D., Mousavi, M., et al. (2017). A first overview of SnowEx ground-based remote sensing activities during the winter 2016–2017. 2017 IEEE international geoscience and remote sensing symposium (IGARSS), Fort Worth, TX, 2017, pp. 1391–1394.
- Clow, D. W. (2010). Changes in the timing of snowmelt and streamflow in Colorado: A response to recent warming. *Journal of Climate*, 23, 2293–2306. <https://doi.org/10.1175/2009JCLI2951.1>
- Currier, W., Pflug, J., Mazzotti, G., Jonas, T., Deems, J., Bormann, K., et al. (2019). Comparing aerial lidar observations with terrestrial lidar and snow-probe transects from NASA's 2017 SnowEx campaign. *Water Resources Research*, 55, 6285–6294. <https://doi.org/10.1029/2018WR024533>
- Deems, J. S., Painter, T. H., & Finnegan, D. C. (2013). Lidar measurements of snow depth: A review. *Journal of Glaciology*, 59(215), 467–479. <https://doi.org/10.3189/2013JoG12J154>
- Durand, M., Gatebe, C., Kim, E., Molotch, N., Painter, T.H., Raleigh, M., et al. (2018). NASA SnowEx Science Plan: Assessing approaches for measuring water in Earth's Seasonal Snow, version 1.6.
- Elder, K., Brucker, L., Hiemstra, C., & Marshall, H. (2018). *SnowEx17 Community Snow Pit Measurements, Version 1*. Boulder, Colorado USA: NASA National Snow and Ice Data Center Distributed Active Archive Center. <https://doi.org/10.5067/Q0310G1XULZS> January 30, 2019
- Heilig, A., Mitterer, C., Schmid, L., Wever, N., Schweizer, J., Marshall, H. P., & Eisen, O. (2015). Seasonal and diurnal cycles of liquid water in snow—Measurements and modeling. *Journal of Geophysical Research: Earth Surface*, 120, 2139–2154. <https://doi.org/10.1002/2015JF003593>
- Hiemstra, C., & Gelvin, A. B. (2018). *SnowEx17 CRREL Differential GNSS Survey, Version 1*. [Snow Survey Stakes]. Boulder, Colorado USA: NASA National Snow and Ice Data Center Distributed Active Archive Center. <https://doi.org/10.5067/10HTB2UVNSEF> July 30, 2018
- Holbrook, W. S., Miller, S. N., & Provart, M. A. (2016). Estimating snow water equivalent over long mountain transects using snowmobile-mounted ground-penetrating radar. *Geophysics*, 81(1). <https://doi.org/10.1190/GEO2014-0121.1>
- Hopkinson, C., Sitar, M., Chasmer, L., & Treltz, P. (2004). Mapping snowpack depth beneath forest canopies using airborne Lidar. *Photogrammetric Engineering and Remote Sensing*, 70(3), 323–330.
- Jennings, K. S., Barnhart, T. B., & Molotch, N. P. (2018). *SnowEx17 Time Series Sonic Snow Depth Measurement Array, Version 1*. Boulder, Colorado USA: NASA National Snow and Ice Data Center Distributed Active Archive Center. <https://doi.org/10.5067/5YJEYNLS1YK4> December 10, 2018
- Kim, E., Gatebe, C., Hall, D., Newlin, J., Misakonis, A., Elder, K., et al. (2018). NASA's SnowEx campaign: Observing seasonal snow in a forested environment. 2017 IEEE international geoscience and remote sensing symposium (IGARSS), Fort Worth, TX, 2017, pp. 1388–1390. <https://doi.org/10.1109/IGARSS.2017.8127222>

- Kovacs, A., Gow, A. J., & Morey, R. M. (1995). The in-situ dielectric constant of polar firn revisited. *Cold Regions Science and Technology*, 23, 245–256.
- Luce, C. H., Abatzoglou, J. T., & Holden, Z. A. (2013). The missing mountain water: Slower westerlies decrease orographic enhancement in the Pacific Northwest USA. *Science*, 342(6164), 1360–1364. <https://doi.org/10.1126/science.1232335>
- Lundberg, A., Granlund, N., & Gustafsson, D. (2010). Towards automated 'ground truth' snow measurements – A review of operational and new measurement methods for Sweden, Norway and Finland. *Hydrological Processes*, 24, 1955–1970. <https://doi.org/10.1002/hyp.7658>
- Lundberg, A., & Thunehed, H. (2000). Snow wetness influence on impulse radar snow surveys theoretical and laboratory study. *Hydrology Research*, 31(2), 89–106.
- Marchand, W.-D., Killingtveit, Å., Wilén, P., & Wikström, P. (2003). Comparison of ground-based and airborne snow depth measurements with georadar systems, case study. *Nordic Hydrology*, 34(5), 427–448.
- Marshall, H., & Koh, G. (2008). FMCW radars for snow research. *Cold Regions Science and Technology*, 52(2), 118–131.
- Marshall, H. P., Koh, G., & Forster, R. (2005). Estimating alpine snowpack properties using FMCW radar. *Annals of Glaciology*, 40, 157–162.
- McGrath, D., Sass, L., O'Neil, S., McNeil, C., Candela, S. G., Baker, E. H., & Marshall, H. P. (2018). Interannual snow accumulation variability on glaciers derived from repeat, spatially extensive ground-penetrating radar surveys. *The Cryosphere*, 12, 3617–3633. <https://doi.org/10.5194/tc-12-3617-2018>
- Mitterer, C., Heilig, A., Schweizer, J., & Eisen, O. (2011). Upward-looking ground-penetrating radar for measuring wet-snow properties. *Cold Regions Science and Technology*, 69, 129–138.
- Mote, P. W., Li, S., Lettenmaier, D. P., Xiao, M., & Engel, R. (2018). Dramatic declines in snowpack in the western US. *NPJ Climate and Atmospheric Science*, 1(1), 1–6. <https://doi.org/10.1038/s41612-018-0012-1>
- Painter, T. H. (2018). *ASO L4 Lidar Snow Depth 3m UTM Grid, Version 1*. Boulder, Colorado USA: NASA National Snow and Ice Data Center Distributed Active Archive Center. <https://doi.org/10.5067/KIE9QNVG7HP0>
- Painter, T. H., Berisford, D. F., Boardman, J. W., Bormann, K. J., Deems, J. S., Gehrke, F., et al. (2016). The Airborne Snow Observatory: Fusion of scanning lidar, imaging spectrometer, and physically-based modeling for mapping snow water equivalent and snow albedo. *Remote Sensing of Environment*, 184, 139–152. <https://dx.doi.org/10.1016/j.rse.2016.06.018>
- Roth, K., Schulin, R., Fluhler, H., & Attinger, W. (1990). Calibration of time domain reflectometry for water-content measurement using a composite dielectric approach. *Water Resources Research*, 26(10), 2267–2273. <https://doi.org/10.1029/WR026i010p02267>
- Sand, K., & Bruland, O. (1998). Application of georadar for snow cover surveying. *Nordic Hydrology*, 29(4/5), 361–370.
- Schmid, L., Heilig, A., Mitterer, C., Schweizer, J., Maurer, H., Okorn, R., & Eisen, O. (2014). Continuous snowpack monitoring using upward-looking ground-penetrating radar technology. *Journal of Glaciology*, 60(221), 509–525. <https://doi.org/10.3189/2014JG13J084>
- Shean, D. E. (2019). Grand Mesa 2019-02-01 snow depth estimate (version 0.1) [data set]. Zenodo. <http://doi.org/10.5281/zenodo.3381653>
- Shean, D. E., Alexandrov, O., Moratto, Z. M., Smith, B. E., Joughin, I. R., Porter, C., & Morin, P. (2016). An automated, open-source pipeline for mass production of digital elevation models (DEMs) from very-high-resolution commercial stereo satellite imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 116, 101–117. <https://doi.org/10.1016/j.isprsjprs.2016.03.012>
- Shean, D. E., Bhushan, S., Lilien, D., & Meyer, J. (2019). dshean/demcoreg, Zenodo DOI release (Version v1.0.0). Zenodo. <http://doi.org/10.5281/zenodo.3243480>
- Stewart, I. T., Cayan, D. R., & Dettinger, M. D. (2005). Changes toward earlier streamflow timing across Western North America. *Journal of Climate*, 19, 1136–1155.
- Sturm, M., Goldstein, M. A., & Parr, C. (2017). Water and life from snow: A trillion dollar science question. *Water Resources Research*, 53, 3534–3544. <https://doi.org/10.1002/2017WR020840>
- Sturm, M., & Holmgren, J. (2018). An automatic snow depth probe for field validation campaigns. *Water Resources Research*, 54, 9695–9701. <https://doi.org/10.1029/2018WR023559>
- Sturm, M., Taras, B., Liston, G. E., Derksen, C., Jonas, T., & Lea, J. (2010). Estimating snow water equivalent using snow depth data and climate classes. *Journal of Hydrometeorology*, 11, 1380–1394. <https://doi.org/10.1175/2010JHM1202.1>
- Techel, F., & Pielmeier, C. (2011). Point observations of liquid water content in wet snow – investigating methodical, spatial and temporal aspects. *The Cryosphere*, 5, 405–418. <https://doi.org/10.5194/tc-5-405-2011>
- Webb, R. (2018). Using ground penetrating radar to assess the variability of snow water equivalent and melt in a mixed canopy forest, northern Colorado. *Frontiers in Earth Science*, 11(3), 482–495. <https://doi.org/10.1007/s11707-017-0645-0>
- Webb, R., McGrath, D., Hale, K., & Molotch, N. P. (2019). *SnowEx17 Ground Penetrating Radar, Version 2*. Boulder, Colorado USA: NASA National Snow and Ice Data Center Distributed Active Archive Center. <https://doi.org/10.5067/G21LGCNLFSC5> [August 28, 2019]
- Webb, R. W., Jennings, K. S., Fend, M., & Molotch, N. (2018). Combining ground-penetrating radar with terrestrial LiDAR to estimate the spatial distribution of liquid water content in seasonal snowpacks. *Water Resources Research*, 54, 10,339–10,349. <https://doi.org/10.1029/2018WR022680>
- Webster, R., & Oliver, M. A. (1992). Sample adequately to estimate variograms of soil properties. *Journal of Soil Science*, 43, 177–192.