https://doi.org/10.5194/egusphere-2023-2543
Preprint. Discussion started: 20 November 2023

Reanalyzing the spatial representativeness of snow depth at automated monitoring stations using airborne Lidar data

Jordan N. Herbert[1], Mark S. Raleigh[2], Eric E. Small[1]

[1]Department of Geological Sciences, University of Colorado Boulder, Boulder, 80309, USA
[2]College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, 97331, USA

Correspondence to: Jordan N. Herbert (jordan.herbert@colorado.edu)

https://doi.org/10.5194/egusphere-2023-2543
Preprint. Discussion started: 20 November 2023

Abstract. Automated snow station networks provide critical hydrologic data. Whether point observations represent snowpack at larger areas is an enduring question. Leveraging the recent proliferation of airborne Lidar snow depth data, we revisit the question of snow station representativeness at multiple scales surrounding 111 stations in Colorado and California (U.S.A.) from 2021–2023 (n= 476 total samples). In about 50% of cases, station depths were at least 10 cm higher than areal-mean snow depth (from Lidar) at 0.5 to 4 km scales. The nearest 50 m Lidar pixels had lower bias and were more often representative than coincident stations. The closest 3 m Lidar pixel often agreed (within 10 cm) with station snow depth, suggesting differences between station snow depth and the nearest 50 m Lidar pixel result from highly localized conditions, not the measurement method. Representativeness decreased as scale increased up to 6 km, mainly explained by the elevation of a site relative to the larger area. The bias direction at individual snow stations is temporally consistent, suggesting the relationship between station depth and that of the surrounding area may be predictable. Improving understanding of snow station representativeness could allow for more accurate validation of modelled and remotely sensed data.

1 Introduction

Mountain snowpack provides water to over a billion people worldwide (Dozier et al., 2016) and comprises the majority of water available in the western United States (Li et al., 2017). Snowmelt impacts agricultural activity (Qin et al., 2020), ecosystems (Blankinship et al., 2014; Dollery et al., 2006), and influences the magnitude and frequency of natural hazards such as wildfires, floods, and droughts (Dierauer et al., 2019; Musselman et al., 2018; Westerling et al., 2006). The amount and timing of water availability in snowmelt dominated watersheds is dependent on snowpack characteristics. Despite recent advances, existing remote sensing techniques do not allow for spatially and temporally continuous monitoring of snow water equivalent (SWE) in the complex terrain of mountain watersheds (Lettenmaier et al., 2015). Instead, assessments of water stored in mountain snowpack for hydrologic research and applications (e.g., streamflow forecasting) rely on a combination of ground-based snow sampling, remote sensing, and modelling (Pagano et al., 2009).

Automated stations (hereafter: snow stations), such as the Natural Resource Conservation Service's (NRCS) Snow-Telemetry (Snotel) network, provide temporally continuous, high-quality measurements of snow depth and SWE at over 900 locations throughout the western United States. At snow stations, the typical spatial support is 9 m^2 for SWE (snow pillow) and ~1 m^2 for depth (ultrasonic sensor). Snow courses provide approximately monthly measurements, typically using ~10 measurements across a 100 m transect. In this paper, we compare snow station and airborne light detection and ranging (Lidar) data to re-examine the representativeness of these point data for larger areas.

Snow stations are strategically located to maximize their utility for water supply forecasts (NRCS, 2011). Sites at high elevations and northern slopes are preferred, since locations with more persistent snow provide data for streamflow forecasts longer into the ablation season. Stations are built on flat surfaces, below tree line (between 2745–3350 m above sea level),

https://doi.org/10.5194/egusphere-2023-2543
Preprint. Discussion started: 20 November 2023

and in areas shielded from high winds (Molotch & Bales, 2006; NRCS, 2011; Woelders et al., 2020). The specific requirements for snow station locations, combined with their uneven distribution across the landscape, may increase the potential for bias when using station data to represent larger areas such as an entire watershed.

In addition to aiding water supply forecasts, snow station data have been applied to a wide array of applications in snow hydrology. Snow station data are frequently used to validate models (Pan et al., 2003; Schneider and Molotch, 2016) and as ground truth references for remotely sensed data (Klein and Barnett, 2003; Lievens et al., 2022; Painter et al., 2016). In these cases, station data are used as the "true" values against which the model and remotely sensed data are validated. However, the datasets being validated frequently represent areas on the hundred-meter to kilometer scale, much larger than the ~1-3 m sampling area of a snow station. Another common use for snow station data is input for data assimilation frameworks (Dechant and Moradkhani, 2011; Margulis et al., 2019; Slater and Clark, 2006; Smyth et al., 2020; Barrett, 2003). These applications also apply snow station data to represent the (usually much larger) scale of the model resolution. Data assimilation schemes could also be impacted by biases from combining datasets collected using different sampling methods, for example remotely sensed versus ground sampled. Finally, point station data have been spatially interpolated into gridded products (Broxton et al., 2019; Molotch et al., 2005; Lopez-Moreno et al., 2010). Even though the interpolation may include the influence of landscape factors such as elevation or aspect, the representativeness of the snow station data is typically unknown and is thus unaccounted for in the interpolation scheme.

Care is warranted when extrapolating snow station data to larger areas because the distribution of snow across a landscape can be highly variable, especially at meter-to-hundred-meter scales (Clark et al., 2011). As a result, many studies have assessed the utility of point data to represent larger areas. Evaluations of point measurement representativeness suggest single measurements are inadequate to represent areas as small as 10 m^2 (López-Moreno et al., 2011) or 30 m^2 (Fassnacht et al., 2018), and over 50 point measurements are required to represent an area of 300 m^2 (Watson et al., 2006). Molotch & Bales (2005) investigated snow station representativeness at six Snotel sites, estimating SWE at the 1 km^2 scale using manual snow depth and density measurements, binary regression trees, and remotely sensed snow-covered area data. They reported variable results in the two years examined. In 2001, four of the six sites overestimated mean snow depth, while two sites yielded values within 10% of the area-mean snow depth. In 2002, three sites were below the mean, two sites were well above, and one site was considered representative. Meromy et al. (2013) carried out a similar investigation at 15 snow stations, collecting over 200 snow depth measurements in the 1 km^2 area around each station. Upon conversion of snow depth to SWE using a uniform snow density, approximately half of snow stations had SWE values within 10% of the mean SWE of the surrounding area, with evidence of interannual consistency in the magnitude and direction of bias at individual sites.

Other studies have used high spatial resolution mapping of snow depth from airborne Lidar to assess snow station representativeness, though these efforts were limited in scope. Grünewald and Lehning, (2011) used data from five snow

https://doi.org/10.5194/egusphere-2023-2543
Preprint. Discussion started: 20 November 2023

stations and three Lidar surveys to assess if snow stations can accurately represent the change of snow depth with altitude. Grünewald and Lehning, (2015) used Lidar surveys from six different watersheds (one survey per watershed), finding sites that met the criteria for snow station locations (as opposed to using real station data) to assess snow station representativeness. After converting snow depth to SWE by assuming uniform snow density, these efforts found that snow stations typically overestimate SWE, possibly due to the sampling locations occurring on flat terrain compared to the more characteristically sloping mountainous terrain of the surrounding area. Of the sites that were deemed representative of the surrounding area (within 10% of the area-mean), there were no discernible similarities in topographic attributes that would serve as a predictor for "well-placed" sites.

The aforementioned studies were limited in the quantity and spatial-extent of study areas due to the labor intensity of manually collecting samples and the limited availability of high-resolution Lidar snow depth data. The recent proliferation of Lidar snow depth data in the western U.S. made possible by the Airborne Snow Observatory (ASO; Painter et al., 2016) provides an opportunity to assess the representativeness of snow monitoring stations using high confidence, spatially distributed Lidar data that are collocated with snow station locations. We utilize the Lidar snow depth data available in watersheds in Colorado and California to revisit the question of how representative the locations of snow monitoring stations are compared to the surrounding area, and whether snow depth differences of these locations and the mean of the surrounding area are consistent in time.

As in the studies described above, we assess spatial representativeness by comparing point snow depth to the mean snow depth of the surrounding area. For this analysis we utilize a combination of different scales and sources of point snow depth data (in meters) to calculate a metric: relative snow depth (RSD). RSD is calculated by subtracting the mean snow depth of the surrounding area (areal mean SD) from the point snow depth representing the snow station (point SD), following Eq. (1):

$$RSD = point\ SD - areal\ mean\ SD \tag{1}$$

We use RSD to determine if extrapolation of the point snow depth to the larger area would overestimate (if positive) or underestimate (if negative) the average depth of snow in the area. Using this metric, we address the following questions: (1) How variable is Lidar snow depth around operational snow stations? (2) What is the distribution of relative snow depth values, and how does RSD change when calculated using different spatial scales and point snow depths derived from different sensing techniques? (3) What impact do landcover and topography have on RSD? (4) Is the sign of RSD at a site consistent through time? While answering these questions we focus on snow depth, not SWE, because snow depth is the variable measured directly both by airborne Lidar and at snow stations.

https://doi.org/10.5194/egusphere-2023-2543
Preprint. Discussion started: 20 November 2023

2 Methods

2.1 Study sites and data

We selected locations in Colorado and California that have available airborne Lidar data from ASO and snow station data over the interval February 2021 through June 2023. California and Colorado have state sponsored programs to carry out airborne Lidar surveys to characterize snowpack during the snow season. Our analysis period coincides with the recent increased Lidar availability in both states, expanding the amount of study sites available for the analysis.

In Colorado, we utilized 40 Lidar surveys in 13 watersheds, containing 48 active snow stations, totaling 138 instances of coincident Lidar and snow station data. All Colorado Lidar surveys were carried out in April and May, typically with two surveys per year per basin. More data were available in California, where we utilized 108 Lidar surveys in 13 watersheds, containing 63 active snow stations, totaling 338 coincident Lidar-station comparisons. California surveys were conducted between January and June, with most surveys between March and May. Locations of the Lidar surveys and snow stations are summarized in Fig. 1 (and supplemental Tables 1 and 2). Between both states, we analyzed 476 instances of coincident Lidar-station data.

https://doi.org/10.5194/egusphere-2023-2543
Preprint. Discussion started: 20 November 2023

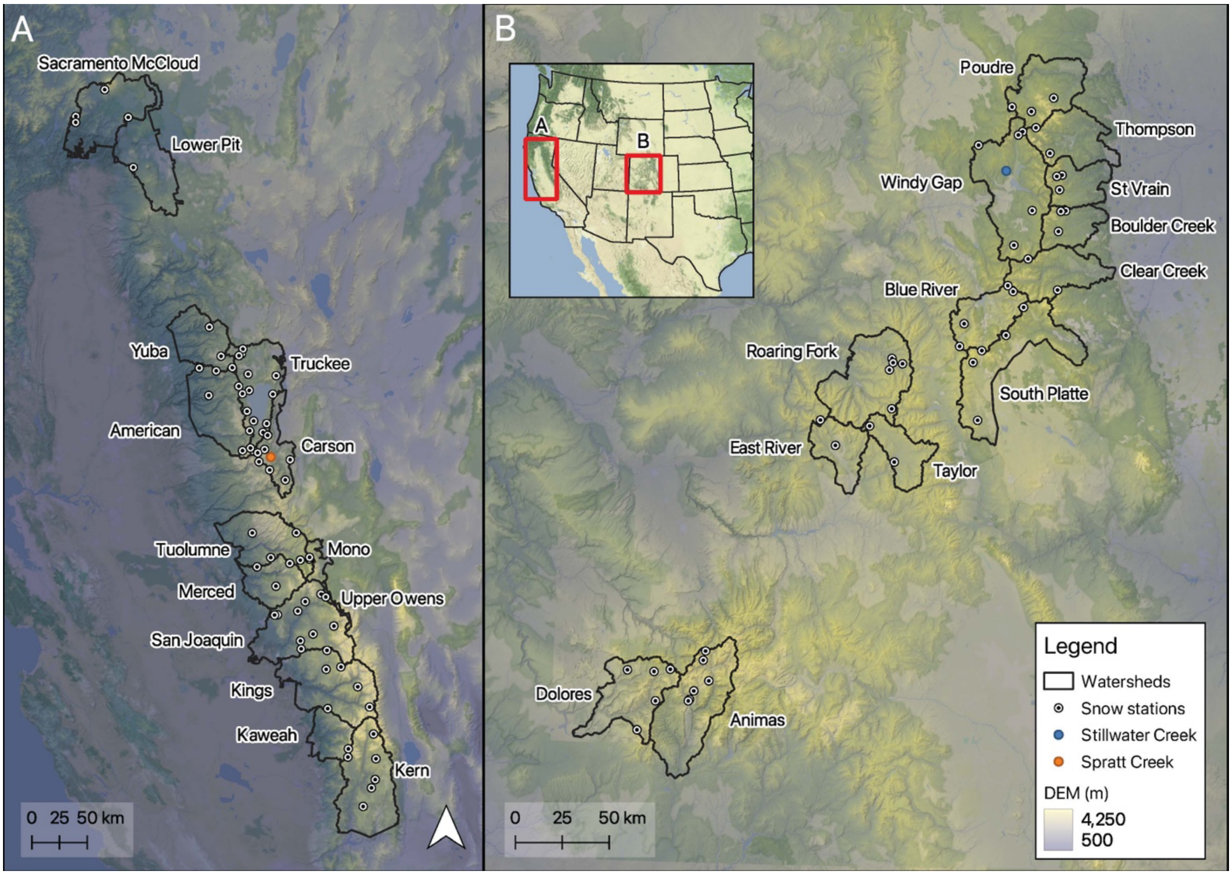


Figure 1: Locations of Lidar surveys and snow stations in (a) California and (b) Colorado, with watersheds labeled. Two stations (Stillwater Creek and Spratt Creek) are highlighted as these are used in subsequent examples. The base map was obtained from Stamen Design (© Stamen Design).

2.1.2 Snow Station Data

The NRCS and the California Department of Water Resources (CA-DWR) operate snow stations which monitor snow depth, SWE, and meteorological parameters at select locations in snow-dominated watersheds. These stations collect snow depth data using an ultrasonic sensor (precision: 13 mm) and SWE data by measuring the mass above a snow pillow (precision: 2.5 mm) (NRCS, 2011). Neither the meteorological data nor the SWE data are used here.

Although SWE is the critical variable for understanding water storage, we conduct our analyses using snow depth because it is the variable directly retrieved by Lidar surveys. Of the existing literature, one study (Molotch and Bales, 2005) directly measured SWE using a federal sampler to get distributed measurements of SWE, but was limited by the total amount of samples collected. Most other studies (e.g., Grünewald and Lehning, 2011, 2015; Meromy et al., 2013) converted snow depth to SWE by assuming a uniform snow density across the study site. Snow density is not uniform across the landscape and may

https://doi.org/10.5194/egusphere-2023-2543
Preprint. Discussion started: 20 November 2023

contribute considerable uncertainty in SWE estimations based on Lidar data (Meehan et al., 2023; Raleigh & Small, 2017; Wetlaufer et al., 2016). Converting values to SWE by assuming a uniform snow density provides no advantage over retaining the values as snow depth. Thus, we keep our analyses in terms of snow depth. Any results herein would be identical if we converted to SWE by multiplying snow depth with a chosen snow density (e.g., Grünewald and Lehning 2011, 2015; Meromy et al., 2013).

We downloaded hourly NRCS Snotel and CA-DWR snow depth data from all sites within the bounds of any watershed area surveyed by ASO with airborne Lidar in Colorado and California from 2021–2023. We acquired site coordinates (latitudes and longitudes) from the NRCS and CA-DWR websites. Due to the importance of accurate location data for this study, we verified the locations of each snow station using visual inspection of high resolution (~1 m) satellite imagery in Google Earth. We updated site coordinates in locations where the provided coordinates were visibly offset from an identifiable snow station. The coordinates were updated to the fifth decimal place in decimal degrees, providing ~1 m accuracy for the location of the center of the snow pillow. We assume that the depth sensor is located over the center of the pillow (which can be identified in the satellite images), although we recognize that this is not always true. The location of four CA-DWR site locations within the Lidar-surveyed watersheds could not be verified and were excluded from the analysis. Site coordinates are available in Supplemental Tables 1 and 2.

The CA-DWR snow depth data required additional quality control measures. In many cases, the snow depth sensor recorded meter-scale changes in hourly snow depth, often followed by an hourly change in the opposite direction of the same magnitude. We discarded clearly erroneous data that recorded hourly jumps of greater than 0.5 meters.

2.1.3 Lidar Data

We utilize all ASO Lidar snow depth data available in Colorado and California from 2021–2023. These datasets are available as gridded rasters at 50 m and 3 m resolutions in the Universal Transverse Mercator (UTM) coordinate system, WGS84. We use the 50 m datasets to analyze the distribution of snow depth surrounding a snow station and calculate the area-mean snow depth at a range of larger scales (see section 2.2). The 50 m scale is sufficient to capture snow depth distribution across the landscape at coarser analysis scales and requires much less storage and computational expense to manage. We also use a subset of 3 m gridded snow depth data, extracting the pixel coincident with the snow station.

Snow depth is retrieved from Lidar data by calculating the difference in surface elevation between snow-on and snow-off surveys. Lidar-derived digital terrain models typically have a root mean squared error (RMSE) on the order of 10 cm (Deems et al., 2013; Kraus & Pfeifer, 1998). Recent studies have reported improved accuracy of these estimates: 3 m snow depths record mean absolute errors of <8 cm, and 50 m snow depths record mean absolute errors of <2 cm (Painter et al., 2016).

https://doi.org/10.5194/egusphere-2023-2543
Preprint. Discussion started: 20 November 2023

It is worth noting that we do not exclude any Lidar data based on proximity to human activities (e.g., compacted snow in ski areas, deeper snow due to snow-making, snow removal on roads) which may impact areal mean snow depths. Snow stations are often built in secluded locations which we expect are minimally impacted by human activities, but this is limited to only the small (~30 m) area surrounding a snow station. Lidar surveys encompassing ski areas, towns, and roads have the potential to record snow depths that do not represent the "natural" snow depth that would have been measured in the absence of human impacts. We chose to not remove any Lidar surveys due to the difficulty of finding an objective, quantitative method to do so, and the changing degree of human impact at a site with scale. We found that at least eight snow stations are near ski areas, but did not find a consistent bias in the RSD across those sites.

2.2 Analyses

We conducted our analyses at three spatial scales typically employed in remote sensing and modeling applications: 0.5 km × 0.5 km, 1 km × 1 km, and 4 km × 4 km (hereafter: 0.5 km, 1 km, and 4 km scales) (Fig. 2, 3). The snow stations were centered within these squares (as in previous studies), though we acknowledge that snow stations will rarely be centered in remotely sensed gridded products. We separately repeated the same analyses using the 0.5 km MOD10A1F grid from the MODIS/Terra Snow Cover Daily L3 Global 500m SIN Grid data set (Riggs and Hall, 2020), and results (not shown) were not significantly changed as compared to the 0.5 km grid centered around a snow station. We evaluated the distribution of snow depth at each scale to gauge snow depth variability, calculating the 5^{th}–95^{th} percentile range of snow depth values using the 50 m resolution Lidar data (Fig. 2d, 3d). We then determined where point snow depth observations fall in the distribution of 50 m snow depths at each scale.

We use different data sources to represent the point snow depth value, allowing us to assess if the results are altered when using values obtained by different sampling methods or at a different Lidar pixel resolution. To represent the point snow depth, we use snow depth recorded by the snow station (station SD), the 50 m resolution Lidar pixel closest to the snow station (50 m SD) and the 3 m resolution Lidar pixel closest to the snow station (3 m SD). We utilize the 50 m resolution data because this resolution is used in a variety of streamflow forecasting schemes (Hedrick et al., 2018; Lahmers et al., 2022; NOAA, 2022). The station SD and high resolution 3 m SD data provide the most direct assessment at the location of a snow station. In addition, they allow us to assess if any biases are introduced by measurement technique (i.e., ground-based ultrasonic sensor versus Lidar).

A primary component of this investigation is the RSD (Eq. 1). We calculate the RSD for each spatial scale, using 50 m SD, 3 m SD, and station SD as "point" measurements. We use RSD as the primary indicator of whether a site is representative of the surrounding area. We deem a site to be representative if a point snow depth is within ±10 cm of the area-mean snow depth. We acknowledge that the range of "acceptable" RSD values varies based on the application and there is subjectivity in what constitutes a representative site. Our results could easily be adjusted using a different range of acceptable values. We do not use a percentage difference from the mean as an indicator of representativeness, as percentages can be overly influenced by the magnitude of snow depth, particularly with shallow snow depths (see Section 3.1 for further justification).

Snow stations are strategically placed on the landscape to maximize their utility for water supply forecasts (NRCS, 2011). We assess the impact of this strategic placement by calculating RSD for all possible snow station locations at each study site. Using Lidar data, we calculate the RSD value by sequentially setting each pixel in a study area as the snow station location. For example, we calculate 100 RSD values at the 0.5 km scale for the 100 pixels (each 50 m resolution) within the study area. We use these data to create a distribution of expected RSD values at a given scale. We then compare the distribution of the virtual RSD values to the distribution of real RSD values (across all 476 station-Lidar survey pairings) to see how strategic placement of snow stations compares to expected RSD values.

We assess variations in landcover and topography to test whether there are any discernable effects on RSD (Fig. 2b-c, 3b-c). To do so, we calculated relative elevation and relative fractional vegetation (FVEG). These metrics are similar to RSD; they are calculated by subtracting the area-mean value of elevation or FVEG from the pixel value closest to the snow station. We employ the National Land Cover Database (Dewitz, 2021) percent tree cover dataset (30 m resolution) for FVEG and the USGS National Elevation Dataset (10 m resolution) (Gesch et al., 2018) for elevation, both bilinearly resampled to the same 50 m grid of the Lidar data.

We expand on the three discrete scales that we employ for most of this investigation in order to more directly assess the effects of scale on snow distribution and on the representativeness of a station. Beginning at the point scale, we expand outward in 50-meter increments up to the 8 km scale. Doing so allows us to examine how the influence of landcover and topography on RSD changes with spatial scale. We also examine how the proportion of representative sites (using 50 m SD) changes with spatial scale. At some sites, expanding the analysis to scales greater than 4 km results in an analysis area that extends beyond the bounds of the Lidar scan. For the increased scale analysis, we only included sites in which 90% or more of the grid cells contain snow depth values at the 8 km scale. This reduced the number of snow stations in our scale analysis (i.e., section 3.3 only) to n=56 (from 111), but ensured that we did not conflate changes in mean areal snow depth and the number of sites analyzed at different scales.

Is the sign and magnitude of RSD at a site consistent through time? We address this question by calculating RSD at each snow station over all available Lidar surveys in the three-year period. For this temporal consistency assessment, we include all sites that have data points spanning at least two years (n = 71 sites). To assess temporal consistency at snow stations, we partition the data into three groups: those where RSD tends to be negative (i.e., mean RSD < -10 cm), where RSD tends to be near zero, or where RSD tends to be positive (i.e., mean RSD > +10 cm). We then analyze the distribution of all points within these three groups.

https://doi.org/10.5194/egusphere-2023-2543
Preprint. Discussion started: 20 November 2023

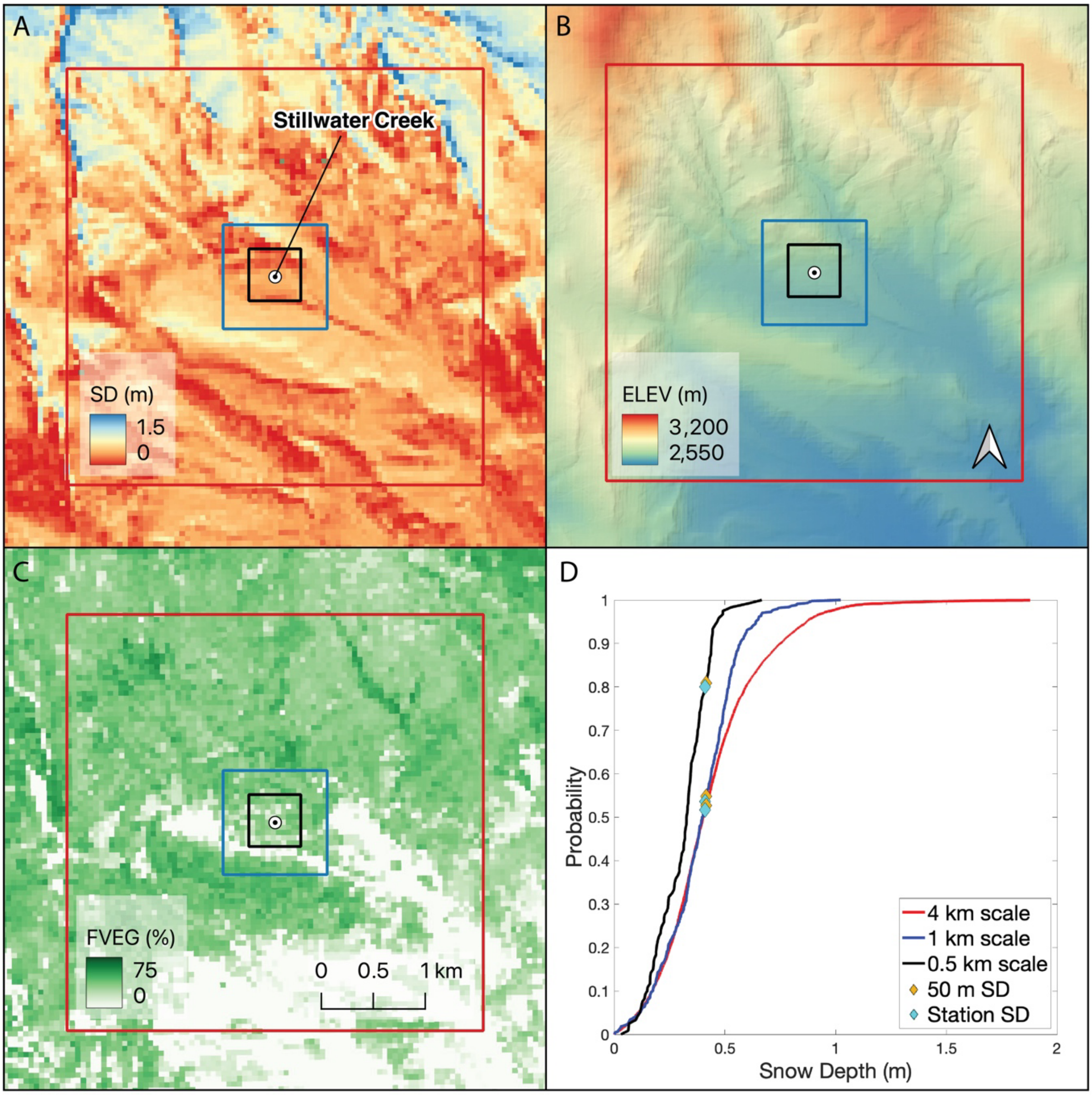


Figure 2: The spatial distribution at 50 m resolution of (a) Lidar snow depth, (b), elevation, and (c) fractional vegetation. The squares represent spatial scales of 0.5 km (black), 1 km (blue), and 4 km (red). (Dd Cumulative density functions of snow depth at each of the three scales with 50 m SD and station SD plotted on the distribution for the Stillwater Creek snow station in Colorado, April 16, 2023.

https://doi.org/10.5194/egusphere-2023-2543
Preprint. Discussion started: 20 November 2023

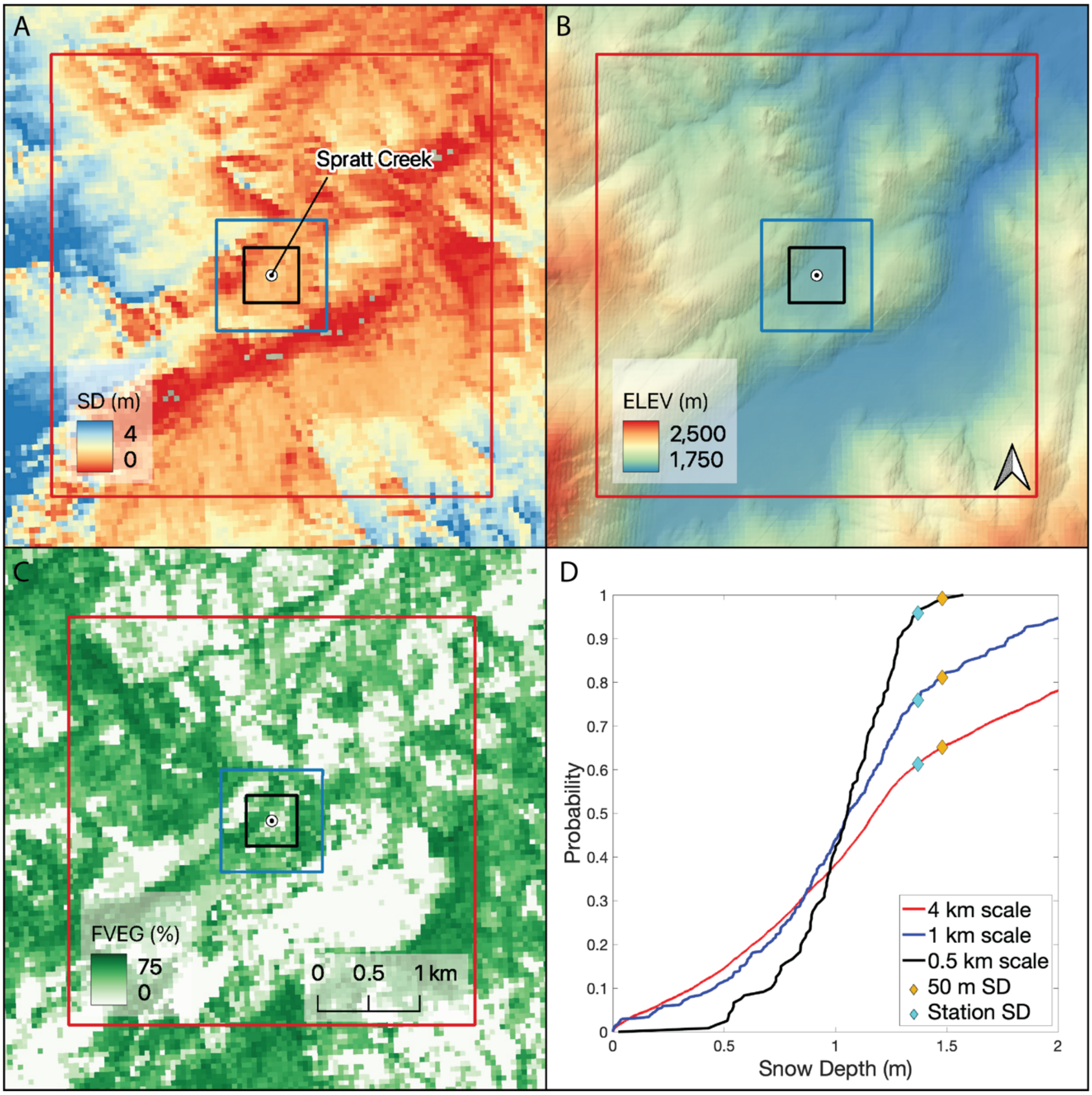


Figure 3: Same as Fig. 2 but for the Spratt Creek snow station in California, March 31, 2023. Note that the CDF in (d) is truncated and that there are snow depth values exceeding 2 m at the 1 km and 4 km scales.

235 **3 Results**

3.1 Snow depth variability

The spatial variability of snow depth influences the likelihood that a snow station is representative of the surrounding area. A higher range of snow depths increases the maximum magnitude of the RSD, whereas a limited snow depth range has a smaller

https://doi.org/10.5194/egusphere-2023-2543
Preprint. Discussion started: 20 November 2023

maximum RSD. For example, a site with a 20 cm range of snow depths would have a maximum RSD value of 10 cm (assuming a normal distribution), guaranteeing the station to be representative. Recall that we define a representative site as being within 10 cm of the areal mean. Here, we examine the statistical distribution of snow depths surrounding snow stations and its role on site representativeness, with a focus on the 0.5 km scale.

The 5–95th percentile range varies greatly between sites and between study region (Colorado vs. California, Fig. 4a, f). The mode for the 5–95th percentile range is 0.4–0.5 m in Colorado and between 0–0.1 m in California; the latter a result of Lidar surveys occurring when more study sites were mostly snow-free. Aside from these low values, most sites have a range of snow depths between 0.3–0.6 at the 0.5 km scale in both Colorado and California. The maximum 5–95th percentile range is about 1 meter in Colorado and 2.4 meters in California, likely due to deeper snowpacks in California. The median range in Colorado is 0.46 m and 0.61 m in California.

The CDF plots demonstrate a range of possible scenarios created from different snow depth distributions. Sites characterized by lower snow depth variability (Fig. 4b, c) are less likely to have point snow depths far from the median due to the limited range of snow depths. For example, at the Michigan Creek Snotel site (Fig. 4b) the Lidar and Snotel point values correspond to the 7th and 95th percentiles, yet both values are within 10 cm depth of the median value. Point snow depths at Ward Creek (Fig. 4g) are similarly marked by percentiles at the upper end of the distribution, while the values themselves are within ~15 cm of the median. Rubicon #2, a site with more moderate spatial variability (Fig. 4h) is characterized by point snow depths which are close to the median in percentile and magnitude. These example sites with low-to-moderate snow depth variability demonstrate that 1) using the percentile proximity to the median is not always an effective indicator of representativeness, and 2) low or moderate snow depth variability generally increases the likelihood of a point snow depth being representative of the surrounding area.

Sites with higher snow depth variability (Fig. 4d, e, i, j) allow for greater differences between the median and point snow depth. For example, at Scotch Creek (Fig. 4e) the closest 50 m SD (yellow marker) is a good representation of the area-wide snow depth, falling 10 cm from the median, at the 60th percentile. Conversely, the station SD (blue marker) is not representative as it is 0.46 m greater than the median snow depth value, near the 100th percentile of the areal distribution.

These results highlight that snow depth variability differs from site to site and that percentile from the median is influenced by the range of snow depth values. Identifying snow depth variability at sites is one important factor that controls the likelihood that a site will be representative of the surrounding area.

https://doi.org/10.5194/egusphere-2023-2543
Preprint. Discussion started: 20 November 2023

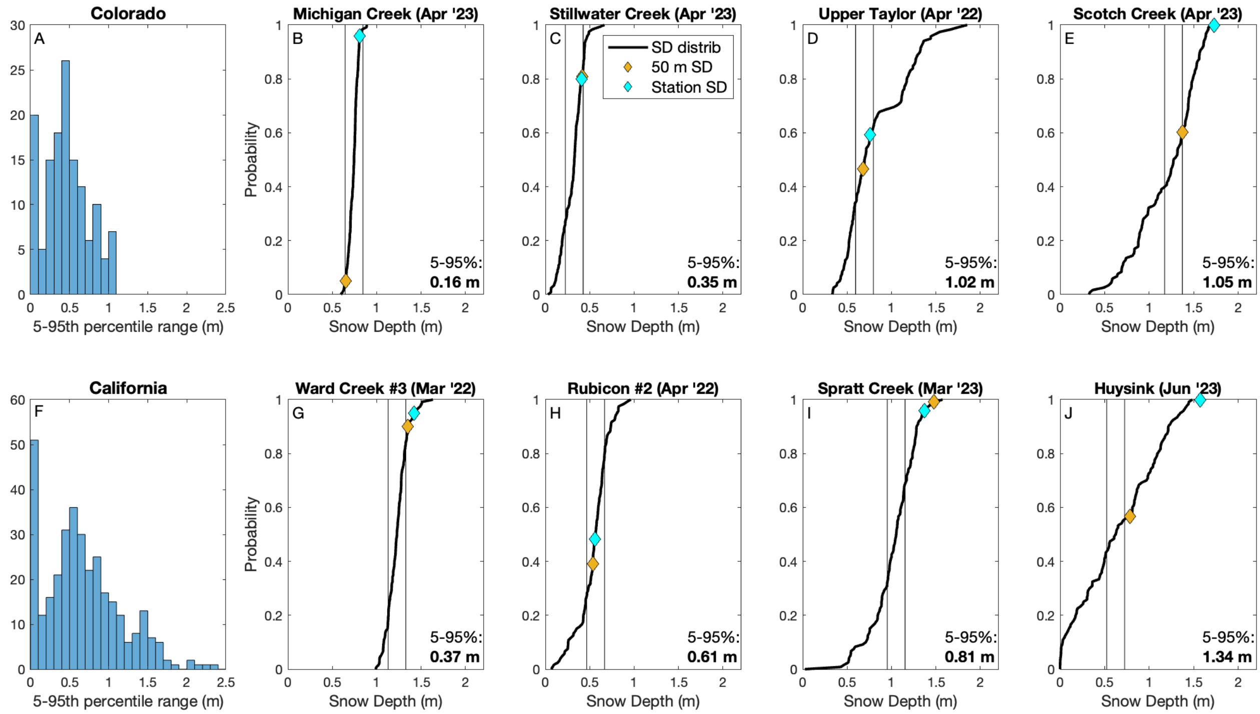


Figure 4: **a) Histogram plot of the range of the 5–95th percentile Lidar snow depth values around snow stations (n=138) in Colorado. b-e) Cumulative density function plots at select sites in Colorado spanning low to high snow depth variability at the 0.5 km scale. f-j) Same as the first row with sites from California (n=338). Point snow depths are plotted with their corresponding probabilities within the snow depth distribution, with blue for station observed and yellow for the 50 m Lidar pixel containing the snow station. Vertical black lines represent the range of snow depth values which are within ±10 cm of the median snow depth.**

3.2 Site representativeness

We now examine the distribution of RSD values, and how the distribution changes when the RSD is calculated using different scales and point snow depths. This is compared to virtual stations that represent the distribution of RSDs from all points within the study area. When using station SD as the point measurement, 35%, 33%, and 28% of the snow stations are representative (RSD within ±10 cm) at the 0.5 km, 1 km, and 4 km scales, respectively (Fig. 5, Table 1). As representativeness decreases with scale, RMSE also increases, with values of 0.46, 0.48, and 0.54 m from small to large scales. There is a clear high bias of RSD values, with ~50% of stations yielding RSD values greater than 10 cm at all three scales.

The proportion of representative sites increases when using 50 m SD as the point value. ~50% of points are representative at the 0.5 km and 1 km scales. Representativeness again decreases with scale, with 38% of points being representative at the 4 km scale (Fig. 5, Table 1). Relative to the station SD case, RMSE values are decreased when using the 50 m SD, yielding

https://doi.org/10.5194/egusphere-2023-2543
Preprint. Discussion started: 20 November 2023

values of 0.20, 0.24, 0.35 m for the 0.5 km, 1 km, and 4 km scales (Table 1). At all three scales there is a slight high bias, with ~15% more sites being "high" (RSD > 10 cm) as opposed to "low" (RSD < -10 cm).

The virtual snow station analysis suggests that 50 m SD locations more effectively represent the surrounding area than if they were placed randomly (Fig. 5 and Table 1). Compared to virtual locations, real site placement (using 50 m SD) increases the frequency of representative sites and lowers the frequency of low sites at all three scales. The frequency of high sites is approximately equal between the 50 m SD and virtual site placement values at all three scales.

These results highlight that 1) point snow depths are more likely to be representative of the surrounding area at finer scales (0.5 km) rather than at coarser scales (4 km), 2) high biases in snow depth occur more frequently for station SD compared to the 50 m lidar SD at the same location, and 3) non-representative sites are more likely to be biased high than biased low at all three scales and for all point data sources.

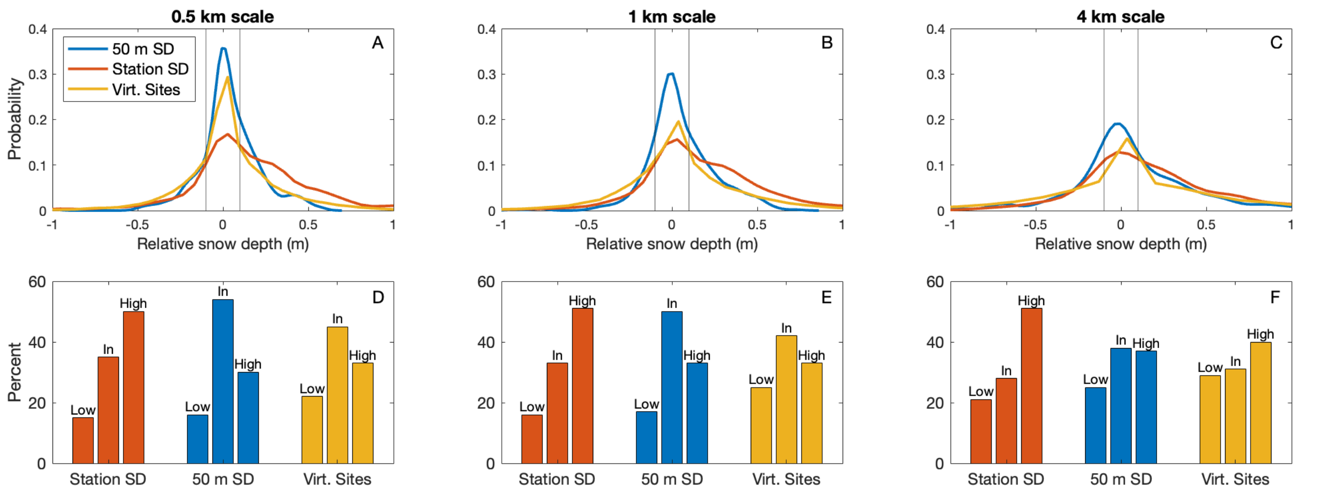


Figure 5: **a-c) Probability density functions of RSD at the 0.5, 1, and 4 km scales, using 50 m SD, station SD, and the virtual station locations as point values. d-f) The relative distribution of RSD values which are less than -10 cm (low), within 10 cm (in), or above 10 cm (high) for each of the point values at each scale.**

https://doi.org/10.5194/egusphere-2023-2543
Preprint. Discussion started: 20 November 2023

Table 1: The percentage of coincident Lidar-snow station data points where RSD is less than -10 cm (Low), within ±10 cm (In), or above 10 cm (High) for each scale, using the 50 m Lidar, Station SD, and the virtually placed snow stations. Median, mean, and RMSE of the RSD values are also presented.

Scale	Point Data	Low (%)	In (%)	High (%)	Median RSD (m)	Mean RSD (m)	RMSE (m)
0.5 km	Station SD	15	35	50	0.10	0.15	0.46
	50 m SD	16	54	30	0.01	0.03	0.20
	Sim. Site	22	45	33	0.00	0.03	0.32
1 km	Station SD	16	33	51	0.11	0.15	0.48
	50 m SD	17	50	33	0.00	0.04	0.24
	Sim. Site	25	42	33	0.00	0.04	0.40
4 km	Station SD	21	28	51	0.11	0.18	0.54
	50 m SD	25	38	37	0.00	0.06	0.35
	Sim. Site	29	31	40	0.00	0.06	0.67

3.3 Topography and fractional forest cover

In this section, we examine question 3: What impact do landcover and topography have on RSD? We found significant correlations between relative elevation and RSD, but no significant correlations between relative fractional vegetation and RSD. Therefore, we only focus on relationships between RSD and relative elevation hereafter, and do not show results related to relative fractional vegetation.

Analysis of the three primary scales demonstrates that the correlation (as indicated with R^2) between RSD and relative elevation increases with scale, coincident with the increased range of relative elevation values (Fig. 6). At the 4 km scale, the slope of the linear regression indicates that RSD increases by 16 cm for every 100 m of relative elevation. The positive slope is consistent with our expectation of temperature and precipitation lapse rates resulting in deeper snow at higher elevations. The y-intercept of the linear regression is 8 cm, suggesting that if relative elevation is zero, RSD is still slightly high biased. This is consistent with the high bias observed earlier when using 50 m SD as the point value (Fig. 5 and Table 1).

We expanded our analysis to all scales from 50 m to 8 km at 50-meter increments to better understand the interplay of scale and topographic effects on RSD. As discussed in section 2.2, we only include sites in which 90% or more of the grid cells contain valid snow depth values at the 8 km scale. The correlation between RSD and relative elevation (as indicated with R^2)

https://doi.org/10.5194/egusphere-2023-2543
Preprint. Discussion started: 20 November 2023

steadily increases with scale until ~7 km, where it levels off at a value of ~0.52 (Fig 7a). The relationship between RSD and relative elevation is significant ($p<0.05$) at scales greater than 0.5 km (Fig. 7a). The increasing correlation with scale is likely linked to a growing range of elevation values; this is demonstrated by the consistently increasing standard deviation of elevation with scale (Fig. 7c). The proportion of representative sites decreases with scale (Fig. 7d). As scale increases, sites are more likely to have higher magnitude relative elevation values, leading to higher magnitude RSD values (and fewer representative sites). The decrease in representativeness levels off at the ~6 km scale (Fig. 7d). At the 6–8 km scales, approximately 25% of sites are representative, 47% of sites are high, and 28% are low. The change in RSD with elevation falls between ~15–18 cm per 100-meters elevation for scales greater than 1 km (Fig. 7b). At scales less than 1 km, the slope falls between 13–16.5 cm per 100-meters, and exhibits much less consistency with scale, likely due to low correlations at these small scales ($R^2 < 0.1$).

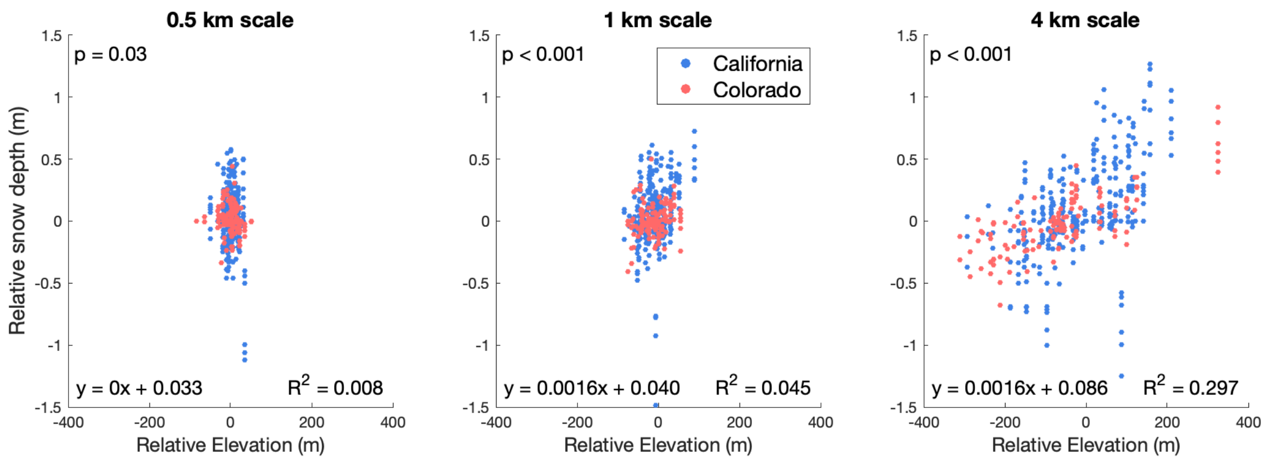
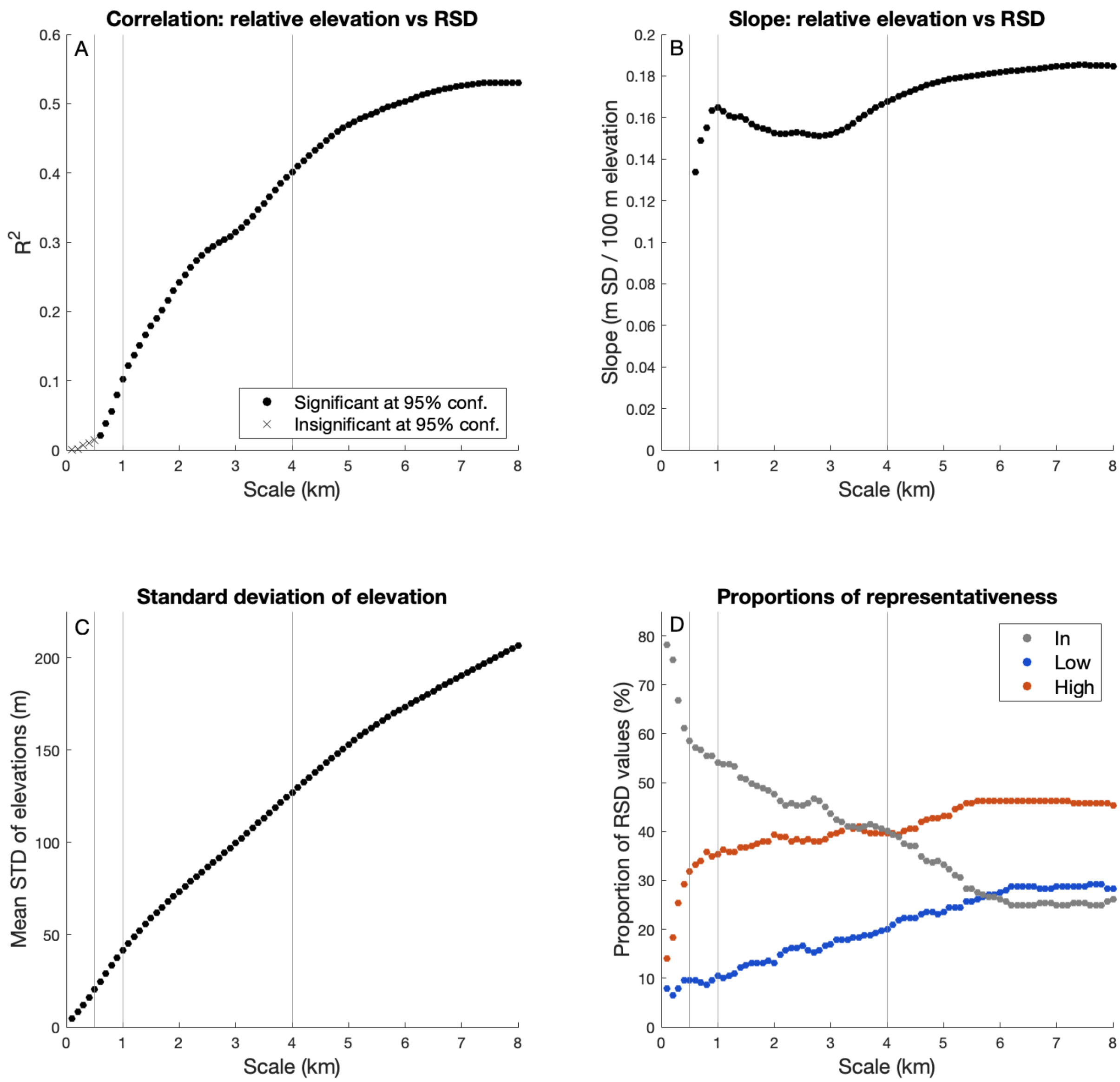


Figure 6: Scatter plots showing the relationship between relative elevation and relative snow depth at the three spatial scales using 50 m SD data to represent the point value.

https://doi.org/10.5194/egusphere-2023-2543
Preprint. Discussion started: 20 November 2023



335 **Figure 7: a) Spatial scale versus R^2 correlation between RSD and relative elevation. Points with p-values less than 0.05 are marked with a filled circle and sites with p-values greater than 0.05 are marked with an "x" marker. b) scale versus regression slope for the relationship between RSD and relative elevation. Only points with significant relationships (from plot a) are shown. c) Scale versus the mean standard deviation of elevation calculated from all sites. d) The proportion of low, representative (in), or high sites with scale when using 50 m SD as the point value. Vertical lines mark the three spatial scales primarily used in this investigation.**

https://doi.org/10.5194/egusphere-2023-2543
Preprint. Discussion started: 20 November 2023

3.4 Point Snow Depth Comparisons

The source and sensing scale of snow depth observations influences whether a site is deemed representative. RSD based on station SD yields a higher bias than 50 m SD (Fig. 5). There are two possible explanations for this bias: 1) snow stations tend to be installed in locations with relatively deep snow compared to the surrounding 50 m area, or 2) there is a systematic bias caused by the difference between remotely sensed Lidar and in-situ station ultrasonic measurements of snow depth. To assess the cause of these differences we now compare the 50 m SD, the 3 m SD, and the station SD with each other (Fig. 8).

Station SDs are systematically higher than the 50 m SDs, with 48% of station SDs being over 10 cm greater than their 50 m SD counterparts and only 9% being at least 10 cm lower than the 50 m SD (Fig. 8a). The station SD and 3 m SD match each other more closely (Fig. 8b); 64% of points are within ±10 cm of each other, with minimal bias. The 3 m SD to 50 m SD comparison (Fig. 8c) yields similar results as the snow station SD to 50 m SD comparison (Fig. 8a), with a similar high bias. The greater similarity between the 3 m SD and station SD values suggest that the high bias in RSD at stations is caused by higher snow depth at the station location, not due to differences in measurement technique (i.e., airborne Lidar vs. a ground-based ultrasonic sensor).

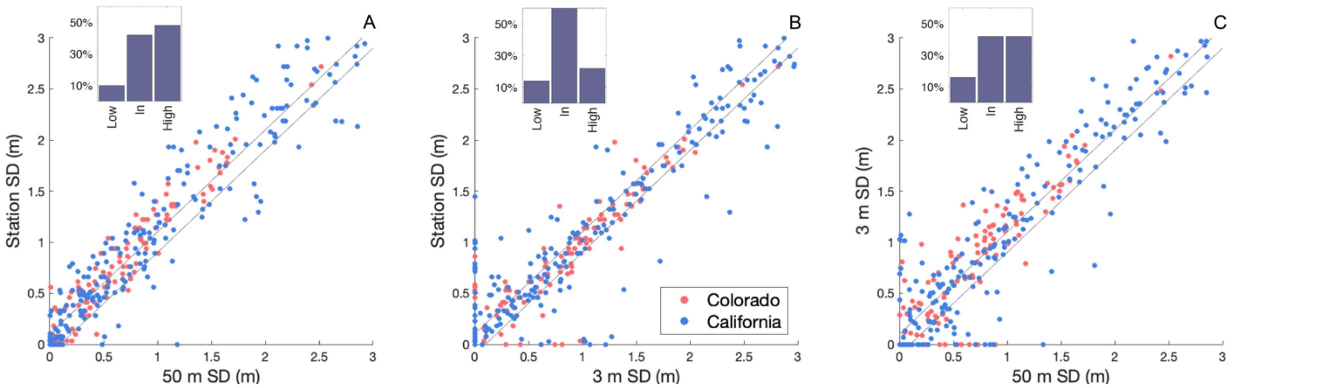


Figure 8: Scatter plots comparing the three different options for point snow depth: (a) station versus 50 m Lidar pixel, (b) station versus 3 m Lidar pixel, and (c) 3 m Lidar pixel versus 50 m Lidar pixel. Points inside the black lines are within ±10 cm of each other. Histogram insets represent percentage of points that are below, within, or above the ±10 cm threshold represented by the black lines.

3.5 Temporal consistency of RSD at snow stations

RSD values at individual sites demonstrate temporal consistency from survey to survey at all three scales (Fig. 9). For this analysis, we used sites with three or more lidar surveys. We grouped the sites into three categories: those with RSD values that are typically negative, near zero, or positive at the 0.5 km, 1 km, and 4 km scales (Fig 9 a-c). Each group contains the same number of sites at each of the three scales. Violin plots of the three categories (Fig. 9d-f) illustrate a divide between the three

https://doi.org/10.5194/egusphere-2023-2543
Preprint. Discussion started: 20 November 2023

groups. By definition, the sites in the low group are classified by almost exclusively negative RSD values whereas sites in the high group are classified by almost exclusively positive RSD values. For example, at the 0.5 km scale, 94 of 101 RSD values in the biased low group are less than or equal to zero. Similarly, 140 out of 160 RSD values are greater than or equal to zero at the high sites. The proportions of low and high sites are similar at the 1 km and 4 km scales. These results demonstrate the consistency of the sign of RSD values at a site through time.

The temporal consistency of RSD at a site must be influenced by more than just relative elevation. As demonstrated in section 3.3, the magnitude of RSD values increases in tandem with the increased magnitude of relative elevation values. However, there is still a clear temporal consistency in the sign of RSD at the smaller (0.5 km and 1 km) scales, where relative elevation has less of an influence (Fig. 9a-c). The 0.5 km scale is particularly striking; relative elevation magnitudes are generally less than 25 m (Fig. 9a), but the proportion of points in the low and high groups displaying unidirectional bias that is similar to the 4 km scale (Fig. 9d, f).

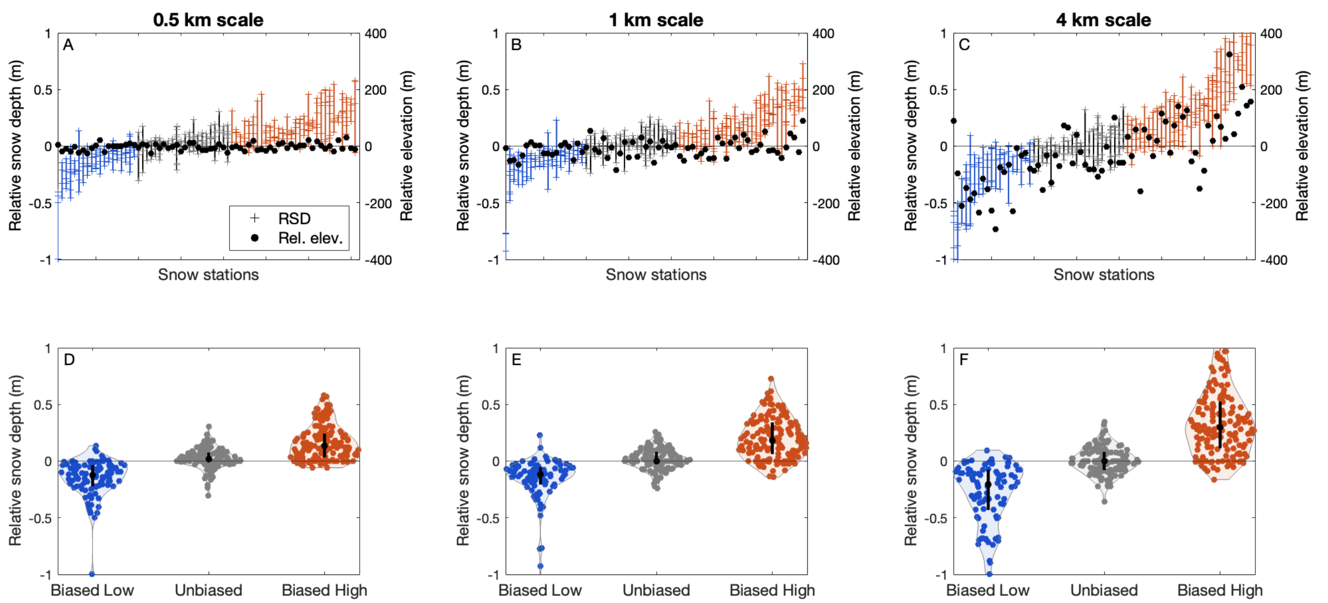


Figure 9: The temporal consistency of relative snow depth at snow stations with three or more Lidar surveys. a-c) Each snotel site (x-axis) plotted against the RSD calculated from the 50 m SD for each Lidar survey at 0.5 km, 1 km, and 4 km scales. Crosses represent individual RSD values and the lines represent the range of RSD values at a given site. Stations are ordered from lowest to highest mean RSD for each scale (thus snow stations are in different orders for each scale). Relative elevation values are also plotted as black circles on the right y-axis. d-f) Distribution plots of qualitatively grouped snow stations which are typically biased low, unbiased, or biased high for the three scales. The black bars with circles represent the median and interquartile range of the RSD values.

4 Discussion

4.1 High bias tendency at operational snow stations

We found that in ~50% of cases, station SDs were not spatially representative, exceeding the area-mean snow depth by at least 10 cm at all three spatial scales (Fig. 5). The tendency of snow stations to overrepresent the surrounding area is not unprecedented. Grünewald et al. (2011; 2015) found that snow stations typically overestimate the mean snow depth of both the corresponding elevation band as well as the entire catchment when analyzing snow depth surrounding areas that fit the qualifications for a snow station location. Meromy et al. (2013) analyzed 53 samples, designating a site representative if the station SD was within ±10% of the area-mean SD. Using that definition, 51% of their station SDs were representative, 30% were high, and 19% were low at the 1 km scale. This distribution more closely matches the distribution we observed when using 50 m SD as the point snow depth, but still demonstrates a slight high bias. It is important to note that the use of percentage from the area-mean snow depth is different than our use of magnitude from the area-mean snow depth, which could affect the results (see the analysis in section 3.1 for justification on how we defined representativeness).

Comparing point snow depths to each other demonstrates that the station SD values are consistently higher than the 50 m SD values at the coincident Lidar pixels (Fig. 8). The general agreement between the 3 m SD and station SD values, two independent data sources, suggests that the elevated values are not a result of sensor bias (i.e., Lidar vs. ultrasonic depth sensor), but rather fine scale (several m) location. A higher proportion of sites are representative of larger areas when using 50 m SD as opposed to station SD (Fig. 5). This suggests that the high bias at the fine-scale station location lowers representativeness. Removing this bias from snow station snow depths would mitigate this problem at some sites, but risks deteriorating representativeness in sites that do not exhibit oversampling.

Why are station SDs higher than the corresponding 50 m SD values? There are multiple possibilities: (1) there is a persistent bias caused by snow station location or (2) the bias is caused by the snow station infrastructure. Grünewald et al. (2011) suggested that elevated station snow depths were a result of flat terrain at a snow station compared to the sloping terrain characteristic of a mountain watershed. Persistent shielding effects or placement within forest gaps could provide another location-based explanation for the high bias. The snow pillow, a flat, vegetation-free structure with different thermal properties than the surrounding forest floor could also cause this bias. Further work is required to ascertain the exact cause of higher snow depths recorded at snow stations compared to the surrounding 50 m area. A final explanation could be that snow density is systematically lower at the snow station, so the increased SD would not actually result in differences in SWE. Density could be lower due to altered thermal exchange at the snow-ground interface due to the snow pillow (i.e., hence changing metamorphism) or due to wind sheltering (e.g., reduced rates of settlement and compaction of newer snow). This final issue highlights the limitations of working in terms of snow depth, since spatial variations in density can influence snow depth variations (e.g., Bonnell et al., 2023; Meehan et al., 2023). Knowledge of both depth and density are needed to accurately resolve spatial distributions of SWE.

https://doi.org/10.5194/egusphere-2023-2543
Preprint. Discussion started: 20 November 2023

4.2 Influence of landcover and topography

Local vegetation and topography influence the distribution of snow across the landscape (Clark et al., 2011). Previous efforts which used statistical approaches (e.g., binary regression trees) to identify the physiographic controls on snow depth surrounding a snow station determined both elevation and fractional vegetation to be major controls on snow depth variability (Meromy et al., 2013; Molotch & Bales, 2006). Our investigation included a simple correlation analysis on the effects of relative elevation and relative fractional vegetation on RSD at the 50 m scale.

Snow depth generally increases with elevation, likely due to increased precipitation and colder temperatures, except at the highest altitudes where wind redistribution is more significant (Grünewald et al., 2014). We found that relative elevation and RSD have significant correlations at scales greater than 0.5 km (Fig. 7a). The correlation increases with scale as the magnitudes of relative elevation values increase, allowing lapse rate effects to outweigh other local controls on snow distribution.

The results show that the proportion of representative sites decreases with scale until plateauing at the 6 km scale (Fig. 7d). The close matching of the representativeness curve (Fig. 7d) to the R^2 curve (Fig. 7a) suggests that these relationships are closely linked. Within the range of scales we assessed in the available data, the larger the scale, the less likely an individual site is to be representative, until the 6–8 km scale. It is unclear why the proportion of representative sites stabilizes at the 6 km scale, but one possible explanation is that other local factors controlling mean areal snow depth keep the impact of relative elevation on RSD from increasing further. It is important to note that high magnitude relative elevation values are the primary cause for deteriorating representativeness at larger scales, not the scale itself. The linear regressions at the 1 and 4 km scales suggests that if relative elevation was zero, RSD would be 4 cm and 9 cm (depending on scale) due to a slight high bias in RSD (Fig. 6). In comparison, relative elevation alters RSD by ~15–18 cm per 100-meters (Fig. 7b). Thus, sites with high magnitude relative elevation values should be adjusted accordingly. It is important to note that the slope (change in RSD per change in relative elevation) calculated here is a mean slope of all sites used in this study. Local factors impact the rate of snow depth change with elevation, so calculating a slope of relative elevation versus RSD at an individual site would be a more accurate way to adjust RSD.

Despite previous studies' identification of fractional vegetation as a major control on snow depth distribution (Meromy et al., 2013; Molotch and Bales, 2006), we found no significant relationships between relative fractional vegetation and relative snow depth. There were also no noticeable differences in the distribution of relative fractional vegetation with scale. The relationship between vegetation and snowpack is complex and non-linear, and (depending on climate) may shift within a single snow season (e.g., less deep snow in the forest in mid-winter but deeper snow in the forest in the spring melt season) (Lundquist et al., 2013; Mazzotti et al., 2020). Other metrics may be more effective at describing the effect of vegetation on snow depth. Additionally, there may be different snow depth regimes within subcanopy zones and gaps in a forest (e.g., Currier & Lundquist, 2018). Vegetation intercepts snowfall, reduces incoming shortwave radiation, decreases wind speeds, and emits

https://doi.org/10.5194/egusphere-2023-2543
Preprint. Discussion started: 20 November 2023

longwave radiation. As such, the relationship between fractional vegetation and snow depth is much more complex than the comparatively simple lapse rate effects of elevation on temperature and precipitation. Accurately simulating forest effects on snow cover also requires extremely high spatial resolutions (<5 m) (Clark et al., 2011; Mazzotti et al., 2021). Our use of a 50 m fractional vegetation dataset may be too coarse to accurately capture forest effects.

4.3 Temporal consistency of station biases

Snow stations exhibit within season and interannual consistency in the directional bias of RSD. Two thirds of sites with three or more lidar surveys have data demonstrating almost exclusively unidirectional bias (Fig. 9). Similarly, Meromy et al. (2013) found consistent bias direction and magnitude at many sites in their investigation. Another study analyzing basin-wide snowpack using Lidar data found consistent patterns of snowpack in years with similar meteorological characteristics (Pflug and Lundquist, 2020). Topography, landcover, and typical storm tracks are relatively static on annual timescales (e.g., Liston, 1999). If these are the factors that control snow depth distribution it is not unexpected that RSD biases would also be similar from year to year at a given site.

Given this consistency, it may only take a few Lidar surveys at a site to determine the relationship of a snow station to the surrounding area at a certain scale. Previously, this would require labor intensive manual sampling of snow depth surrounding a snow station. Thus, we could increase the utility of the temporally continuous snow station data with just a few Lidar surveys. The consistency we observe provides the opportunity to adjust snow station data based on the typical RSD bias at a site for other applications. Doing so would cause the adjusted value to be more in line with the area-mean snow depth, improving its utility for remote sensing ground-truthing, data assimilation, or model validation efforts.

5 Conclusions

We analyzed snow depth distributions surrounding snow stations at three scales using coincident Lidar-snow station data in Colorado and California from 2021–2023. Snow stations record snow depths within ±10 cm of the area-mean snow depth in about 1/3 of cases at all three scales, while overrepresenting the areal-mean snow depth in ~50% of cases. When using the closest 50 m Lidar point to represent the snow station SD, the frequency of site representation is increased to ~50% at the 0.5 and 1 km scales. Representativeness increases when using the 50 m Lidar point because snow station locations record snow depths which are on average ~10 cm greater than the surrounding 50 m area. Representativeness decreases with scale because relative elevation magnitudes increase, causing lapse rates to impact relative snow depth via changes in mean areal snow depth. The directional bias of RSD at a snow station is consistent from survey to survey. Together, these results suggest there is an opportunity to increase the utility of snow stations for model validation and ground truthing. Future work should focus on determining the underlying influences that cause site bias, potentially allowing for *a priori* identification of a site's relationship to the surrounding area. Adjusting snow station data based on the consistent high bias compared to the surrounding 50 m area, or based on the typical trend of RSD would increase the ability of a snow station to better represent the surrounding area, particularly at scales of 1 km or less.

Code Availability

Data Availability

The data used herein are all publicly available.

Author contribution

Herbert carried out the analyses, created the figures, and wrote the manuscript. Small and Raleigh helped design the experiments as well as edit the figures and manuscript text.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

This work was supported by the National Aeronautics and Space Administration (NASA) under three grants: Water Resources Program Grant No. 80NSSC22K0928 and Terrestrial Hydrology Program Grants No. 80NSSC22K0685 and NNX17AL41G. We would like to thank the institutions which provided invaluable data for this investigation. Landcover and topography data were provided by the United States Geological Survey. Snow station data were provided by the USDA NRCS and the CA-DWR (https://www.wcc.nrcs.usda.gov/snow/, https://cdec.water.ca.gov/snow/current/snow/). Lidar data were provided by ASO, Inc. (https://www.airbornesnowobservatories.com/).

References

Barrett, A. P.: National Operational Hydrologic Remote Sensing Center SNOw Data Assimilation System (SNODAS) Products at NSIDC., Boulder, CO, 2003.

Blankinship, J. C., Meadows, M. W., Lucas, R. G., and Hart, S. C.: Snowmelt timing alters shallow but not deep soil moisture in the Sierra Nevada, Water Resour Res, 50, 1448–1456, https://doi.org/10.1002/2013WR014541, 2014.

https://doi.org/10.5194/egusphere-2023-2543
Preprint. Discussion started: 20 November 2023

Bonnell, R., McGrath, D., Hedrick, A. R., Trujillo, E., Meehan, T. G., Williams, K., Marshall, H., Sexstone, G., Fulton, J., Ronayne, M. J., Fassnacht, S. R., Webb, R. W., and Hale, K. E.: Snowpack relative permittivity and density derived from near-coincident lidar and ground-penetrating radar, Hydrol Process, 37, https://doi.org/10.1002/hyp.14996, 2023.

Broxton, P. D., van Leeuwen, W. J. D., and Biederman, J. A.: Improving Snow Water Equivalent Maps With Machine Learning of Snow Survey and Lidar Measurements, Water Resour Res, 55, 3739–3757, https://doi.org/10.1029/2018WR024146, 2019.

Clark, M. P., Hendrikx, J., Slater, A. G., Kavetski, D., Anderson, B., Cullen, N. J., Kerr, T., Örn Hreinsson, E., and Woods, R. A.: Representing spatial variability of snow water equivalent in hydrologic and land-surface models: A review, https://doi.org/10.1029/2011WR010745, 2011.

Currier, W. R. and Lundquist, J. D.: Snow Depth Variability at the Forest Edge in Multiple Climates in the Western United States, Water Resour Res, 54, 8756–8773, https://doi.org/10.1029/2018WR022553, 2018.

Dechant, C. M. and Moradkhani, H.: Improving the characterization of initial condition for ensemble streamflow prediction using data assimilation, Hydrol Earth Syst Sci, 15, 3399–3410, https://doi.org/10.5194/hess-15-3399-2011, 2011.

Deems, J. S., Painter, T. H., and Finnegan, D. C.: Lidar measurement of snow depth: A review, https://doi.org/10.3189/2013JoG12J154, July 2013.

Dewitz, J.: National Land Cover Database, United States Geological Survey, 2021.

Dierauer, J. R., Allen, D. M., and Whitfield, P. H.: Snow Drought Risk and Susceptibility in the Western United States and Southwestern Canada, Water Resour Res, 55, 3076–3091, https://doi.org/10.1029/2018WR023229, 2019.

Dollery, R., Hodkinson, I. D., and Jónsdóttir, I. S.: Impact of warming and timing of snow melt on soil microarthropod assemblages associated with Dryas-dominated plant communities on Svalbard, Ecography, 29, 111–119, https://doi.org/10.1111/j.2006.0906-7590.04366.x, 2006.

Dozier, J., Bair, E. H., and Davis, R. E.: Estimating the spatial distribution of snow water equivalent in the world's mountains, https://doi.org/10.1002/wat2.1140, 1 May 2016.

Fassnacht, S. R., Brown, K. S. J., Blumberg, E. J., López Moreno, J. I., Covino, T. P., Kappas, M., Huang, Y., Leone, V., and Kashipazha, A. H.: Distribution of snow depth variability, Front Earth Sci, 12, 683–692, https://doi.org/10.1007/s11707-018-0714-z, 2018.

Gesch, D. B., Evans, G. . A., Oimoen, M. J., and Arundel, S.: The National Elevation Dataset, American Society for Photogrammetry and Remote Sensing, 2018.

Grünewald, T. and Lehning, M.: Altitudinal dependency of snow amounts in two small alpine catchments: can catchment-wide snow amounts be estimated via single snow or precipitation stations?, 2011.

Grünewald, T. and Lehning, M.: Are flat-field snow depth measurements representative? A comparison of selected index sites with areal snow depth measurements at the small catchment scale, Hydrol Process, 29, 1717–1728, https://doi.org/10.1002/hyp.10295, 2015.

Grünewald, T., Bühler, Y., and Lehning, M.: Elevation dependency of mountain snow depth, Cryosphere, 8, 2381–2394, https://doi.org/10.5194/tc-8-2381-2014, 2014.

Hedrick, A. R., Marks, D., Havens, S., Robertson, M., Johnson, M., Sandusky, M., Marshall, H. P., Kormos, P. R., Bormann, K. J., and Painter, T. H.: Direct Insertion of NASA Airborne Snow Observatory-Derived Snow Depth Time Series Into the iSnobal Energy Balance Snow Model, Water Resour Res, 54, 8045–8063, https://doi.org/10.1029/2018WR023190, 2018.

Klein, A. G. and Barnett, A. C.: Validation of daily MODIS snow cover maps of the Upper Rio Grande River Basin for the 2000-2001 snow year, Remote Sens Environ, 86, 162–176, https://doi.org/10.1016/S0034-4257(03)00097-X, 2003.

Kraus Ł, K. and Pfeifer, N.: Determination of terrain models in wooded areas with airborne laser scanner data, ISPRS Journal of Photogrammetry & Remote Sensing, 193–203 pp., 1998.

Lahmers, T. M., Kumar, S. V., Rosen, D., Dugger, A., Gochis, D. J., Santanello, J. A., Gangodagamage, C., and Dunlap, R.: Assimilation of NASA's Airborne Snow Observatory Snow Measurements for Improved Hydrological Modeling: A Case Study Enabled by the Coupled LIS/WRF-Hydro System, Water Resour Res, 58, https://doi.org/10.1029/2021WR029867, 2022.

Lettenmaier, D. P., Alsdorf, D., Dozier, J., Huffman, G. J., Pan, M., and Wood, E. F.: Inroads of remote sensing into hydrologic science during the WRR era, Water Resour Res, 51, 7309–7342, https://doi.org/10.1002/2015WR017616, 2015.

Li, D., Wrzesien, M. L., Durand, M., Adam, J., and Lettenmaier, D. P.: How much runoff originates as snow in the western United States, and how will that change in the future?, Geophys Res Lett, 44, 6163–6172, https://doi.org/10.1002/2017GL073551, 2017.

Lievens, H., Brangers, I., Marshall, H. P., Jonas, T., Olefs, M., and De Lannoy, G.: Sentinel-1 snow depth retrieval at sub-kilometer resolution over the European Alps, Cryosphere, 16, 159–177, https://doi.org/10.5194/TC-16-159-2022, 2022.

Liston, G. E.: Interrelationships among Snow Distribution, Snowmelt, and Snow Cover Depletion: Implications for Atmospheric, Hydrologic, and Ecologic Modeling, 1999.

López-Moreno, J. I., Fassnacht, S. R., Beguería, S., and Latron, J. B. P.: Variability of snow depth at the plot scale: Implications for mean depth estimation and sampling strategies, Cryosphere, 5, 617–629, https://doi.org/10.5194/tc-5-617-2011, 2011.

Lundquist, J. D., Dickerson-Lange, S. E., Lutz, J. A., and Cristea, N. C.: Lower forest density enhances snow retention in regions with warmer winters: A global framework developed from plot-scale observations and modeling, Water Resour Res, 49, 6356–6370, https://doi.org/10.1002/wrcr.20504, 2013.

Margulis, S. A., Fang, Y., Li, D., Lettenmaier, D. P., and Andreadis, K.: The Utility of Infrequent Snow Depth Images for Deriving Continuous Space-Time Estimates of Seasonal Snow Water Equivalent, Geophys Res Lett, 46, 5331–5340, https://doi.org/10.1029/2019GL082507, 2019.

Mazzotti, G., Essery, R., Moeser, C. D., and Jonas, T.: Resolving Small-Scale Forest Snow Patterns Using an Energy Balance Snow Model With a One-Layer Canopy, Water Resour Res, 56, https://doi.org/10.1029/2019WR026129, 2020.

Mazzotti, G., Webster, C., Essery, R., and Jonas, T.: Increasing the Physical Representation of Forest-Snow Processes in Coarse-Resolution Models: Lessons Learned From Upscaling Hyper-Resolution Simulations, Water Resour Res, 57, https://doi.org/10.1029/2020WR029064, 2021.

Meehan, T. G., Hojatimalekshah, A., Marshall, H.-P., Deeb, E. J., O'neel, S., Mcgrath, D., Webb, R. W., Bonnell, R., Raleigh, M. S., Hiemstra, C., and Elder, K.: Spatially distributed snow depth, bulk density, and snow water equivalent from ground-based and airborne sensor integration at Grand Mesa, Colorado, USA, https://doi.org/10.5194/tc-2023-141, n.d.

Meromy, L., Molotch, N. P., Link, T. E., Fassnacht, S. R., and Rice, R.: Subgrid variability of snow water equivalent at operational snow stations in the western USA, Hydrol Process, 27, 2383–2400, https://doi.org/10.1002/hyp.9355, 2013.

Molotch, N. P. and Bales, R. C.: Scaling snow observations from the point to the grid element: Implications for observation network design, Water Resour Res, 41, 1–16, https://doi.org/10.1029/2005WR004229, 2005.

Molotch, N. P. and Bales, R. C.: SNOTEL representativeness in the Rio Grande headwaters on the basis of physiographics and remotely sensed snow cover persistence, in: Hydrological Processes, 723–739, https://doi.org/10.1002/hyp.6128, 2006.

Molotch, N. P., Colee, M. T., Bales, R. C., and Dozier, J.: Estimating the spatial distribution of snow water equivalent in an alpine basin using binary regression tree models: The impact of digital elevation data and independent variable selection, Hydrol Process, 19, 1459–1479, https://doi.org/10.1002/hyp.5586, 2005.

Moreno, J. I. L., Latron, J., and Lehmann, A.: Effects of sample and grid size on the accuracy and stability of regression-based snow interpolation methods, Hydrol Process, 24, 1914–1928, https://doi.org/10.1002/hyp.7564, 2010.

Musselman, K. N., Lehner, F., Ikeda, K., Clark, M. P., Prein, A. F., Liu, C., Barlage, M., and Rasmussen, R.: Projected increases and shifts in rain-on-snow flood risk over western North America, https://doi.org/10.1038/s41558-018-0236-4, 1 September 2018.

NOAA: CBRFC Water Year in Review, National Oceanic and Atmospheric Administration, 2022.

NRCS: Part 622 Snow Survey and Water Supply Forecasting National Engineering Handbook, National Resource Conservation Service, 2011.

Pagano, T. C., Garen, D. C., Perkins, T. R., and Pasteris, P. A.: Daily updating of operational statistical seasonal water supply forecasts for the Western U.S., J Am Water Resour Assoc, 45, 767–778, https://doi.org/10.1111/j.1752-1688.2009.00321.x, 2009.

Painter, T. H., Berisford, D. F., Boardman, J. W., Bormann, K. J., Deems, J. S., Gehrke, F., Hedrick, A., Joyce, M., Laidlaw, R., Marks, D., Mattmann, C., McGurk, B., Ramirez, P., Richardson, M., Skiles, S. M. K., Seidel, F. C., and Winstral, A.: The Airborne Snow Observatory: Fusion of scanning lidar, imaging spectrometer, and physically-based modeling for mapping snow water equivalent and snow albedo, Remote Sens Environ, 184, 139–152, https://doi.org/10.1016/j.rse.2016.06.018, 2016.

Pan, M., Sheffield, J., Wood, E. F., Mitchell, K. E., Houser, P. R., Schaake, J. C., Robock, A., Lohmann, D., Cosgrove, B., Duan, Q., Luo, L., Higgins, R. W., Pinker, R. T., and Tarpley, J. D.: Snow process modeling in the North American Land

https://doi.org/10.5194/egusphere-2023-2543
Preprint. Discussion started: 20 November 2023

Data Assimilation System (NLDAS): 1. Evaluation of model simulated snow water equivalent, Journal of Geophysical Research D: Atmospheres, 108, https://doi.org/10.1029/2003jd003994, 2003.

Pflug, J. M. and Lundquist, J. D.: Inferring Distributed Snow Depth by Leveraging Snow Pattern Repeatability: Investigation Using 47 Lidar Observations in the Tuolumne Watershed, Sierra Nevada, California, Water Resour Res, 56, https://doi.org/10.1029/2020WR027243, 2020.

Qin, Y., Abatzoglou, J. T., Siebert, S., Huning, L. S., AghaKouchak, A., Mankin, J. S., Hong, C., Tong, D., Davis, S. J., and Mueller, N. D.: Agricultural risks from changing snowmelt, Nat Clim Chang, 10, 459–465, https://doi.org/10.1038/s41558-020-0746-8, 2020.

Raleigh, M. S. and Small, E. E.: Snowpack density modeling is the primary source of uncertainty when mapping basin-wide SWE with lidar, Geophys Res Lett, 44, 3700–3709, https://doi.org/10.1002/2016GL071999, 2017.

Riggs, G. and Hall, D.: Continuity of modis and viirs snow cover extent data products for development of an earth science data record, Remote Sens (Basel), 12, 1–20, https://doi.org/10.3390/rs12223781, 2020.

Schneider, D. and Molotch, N. P.: Real-time estimation of snow water equivalent in the Upper Colorado River Basin using MODIS-based SWE Reconstructions and SNOTEL data, Water Resour Res, 52, 7892–7910, https://doi.org/10.1002/2016WR019067, 2016.

Slater, A. G. and Clark, M. P.: Snow Data Assimilation via an Ensemble Kalman Filter, 2006.

Smyth, E. J., Raleigh, M. S., and Small, E. E.: Improving SWE Estimation With Data Assimilation: The Influence of Snow Depth Observation Timing and Uncertainty, Water Resour Res, 56, https://doi.org/10.1029/2019WR026853, 2020.

Watson, F. G. R., Anderson, T. N., Newman, W. B., Alexander, S. E., and Garrott, R. A.: Optimal sampling schemes for estimating mean snow water equivalents in stratified heterogeneous landscapes, J Hydrol (Amst), 328, 432–452, https://doi.org/10.1016/j.jhydrol.2005.12.032, 2006.

Westerling, A. L., Hidalgo, H. G., Cayan, D. R., and Swetnam, T. W.: Warming and earlier spring increase Western U.S. forest wildfire activity, Science (1979), 313, 940–943, https://doi.org/10.1126/SCIENCE.1128834/SUPPL_FILE/WESTERLING.SOM.PDF, 2006.

Wetlaufer, K., Hendrikx, J., and Marshall, L.: Spatial heterogeneity of snow density and its influence on snow water equivalence estimates in a large mountainous basin, Hydrology, 3, https://doi.org/10.3390/hydrology3010003, 2016.

Woelders, L., Lukas, J., Payton, L., and Duncan, B.: Snowpack monitoring in the Rocky Mountain West: A user guide, 2020.