

Subgrid variability of snow water equivalent at operational snow stations in the western USA

Leah Meromy,^{1*} Noah P. Molotch,^{1,2} Timothy E. Link,³ Steven R. Fassnacht⁴ and Robert Rice⁵

¹ Department of Geography, Institute of Arctic and Alpine Research, University of Colorado at Boulder, Boulder, CO, USA

² Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

³ Department of Forest Ecology and Biogeosciences, University of Idaho, Moscow, ID, USA

⁴ ESS–Watershed Science, Colorado State University, Fort Collins, CO, USA

⁵ Sierra Nevada Research Institute, University of California at Merced, Merced, CA, USA

Abstract:

The spatial distribution of snow water equivalent (SWE) is a key variable in many regional-scale land surface models. Currently, the assimilation of point-scale snow sensor data into these models is commonly performed without consideration of the spatial representativeness of the point data with respect to the model grid-scale SWE. To improve the understanding of the relationship between point-scale snow measurements and surrounding areas, we characterized the spatial distribution of snow depth and SWE within 1-, 4- and 16-km² grids surrounding 15 snow stations (snowpack telemetry and California snow sensors) in California, Colorado, Wyoming, Idaho and Oregon during the 2008 and 2009 snow seasons. More than 30 000 field observations of snowpack properties were used with binary regression tree models to relate SWE at the sensor site to the surrounding area SWE to evaluate the sensor representativeness of larger-scale conditions. Unlike previous research, we did not find consistent high biases in snow sensor depth values as biases over all sites ranged from 74% overestimates to 77% underestimates. Of the 53 assessments, 27 surveys indicated snow station biases of less than 10% of the surrounding mean observed snow depth. Depth biases were largely dictated by the physiographic relationship between the snow sensor locations and the mean characteristics of the surrounding grid, in particular, elevation, solar radiation index and vegetation density. These scaling relationships may improve snow sensor data assimilation; an example application is illustrated for the National Operational Hydrologic Remote Sensing Center National Snow Analysis SWE product. The snow sensor bias information indicated that the assimilation of point data into the National Operational Hydrologic Remote Sensing Center model was often unnecessary and reduced model accuracy. Copyright © 2012 John Wiley & Sons, Ltd.

KEY WORDS snow; modelling; snow water equivalent; water resources; SNOTEL

Received 5 December 2011; Accepted 20 April 2012

INTRODUCTION

The western USA largely depends on mountain snowmelt for water supply (Barnett *et al.*, 2005; Bales *et al.*, 2006). Quantifying this source of water is particularly important for water resource management to support municipal, industrial and agricultural sectors. In mountainous regions, complex interactions of climatological and physiographic variables have limited our understanding of the processes controlling water availability from snowmelt. Such complicating factors include substantial temperature and precipitation gradients with elevation, distinct wet–dry seasonal transitions and great interannual climate variability (Bales *et al.*, 2006). Enhanced knowledge of subalpine hydrological processes, particularly those governing snow accumulation and melt, will enable a more precise hydrologic modelling and a clearer understanding of the potential effects of climate change on water availability.

Hydrologic modelling in mountainous regions with seasonal snow cover has historically been limited by the paucity of ground-based hydrometric measurements – especially with regard to snow observations. Remote sensing has been successfully used to estimate the spatial distribution of snow water equivalent (SWE) at larger regional (e.g. >100 000 km²) and global scales across relatively uniform terrain (Lee *et al.*, 2005; Derksen, 2008; Tedesco and Narvekar, 2010). However, the limitations in accuracy and the relatively coarse spatial resolution make remote sensing techniques insufficient for characterizing snowpack at watershed scales (e.g. ~1000 km²) (Cline *et al.*, 1998). As a result, a breadth of research has resolved SWE distribution via interpolation of ground-based measurements (Carroll and Cressie, 1996; Balk and Elder, 2000; Winstral *et al.*, 2002; Fassnacht *et al.*, 2003; Anderton *et al.*, 2004; Molotch *et al.*, 2005; Lopez-Moreno *et al.*, 2010). Those approaches have provided important information for updating, evaluating and calibrating distributed hydrologic models and remote sensing retrieval algorithms. An important aspect of using ground observations in the context of broader-scale SWE estimation is the issue of the spatial representativeness of point measurements and

*Correspondence to: Leah Meromy, Department of Geography, Institute of Arctic and Alpine Research, University of Colorado, Boulder, Campus Box 450 UCB, Boulder, CO 80309, USA.
E-mail: leah.meromy@colorado.edu

the model grid cell scale. Indeed, numerous studies indicate that individual point observations of SWE are not necessarily representative of the surrounding grid element (Molotch and Bales, 2005, 2006; Neumann *et al.*, 2006; Rice and Bales, 2010; Lopez-Moreno *et al.*, 2011).

Meteorological data, SWE and other snowpack attributes in the western USA are primarily collected at more than 730 monitoring sites such as snowpack telemetry (SNOTEL) stations operated by the National Resources Conservation Service (NRCS, 2010). In California, a similar network of more than 110 stations is maintained independently by the California Cooperative Snow Survey (CA-DWR, 2010). Because these data are used for regression-based water supply prediction models, the snow station locations were selected to represent the hydrologic and general environmental conditions of a given watershed and typically in specific locations where snow accumulates earlier and ablates later than surrounding areas (Daly *et al.*, 2000). However, many sites in mountain watersheds, which may be more representative than current monitoring stations, are not easily accessible; thus, the sites selected for snow stations may not be representative of SWE in their respective watersheds (Molotch and Bales, 2005). Moreover, because these data are frequently used to update and evaluate hydrologic models, the relationship between the point SWE and the mean SWE of the surrounding grid cell must be determined to assess how representative a given snow station site is relative to the surrounding grid cell. Identifying consistent relationships between the point measurement and the surrounding grid is very useful because it can facilitate upscaling point snow station SWE measurements to the surrounding area, thus improving snowpack SWE estimation.

The primary objective of this study was to improve the understanding of the relationship between the measured SWE at snow stations and the SWE of the surrounding grid element. The specific objectives were (i) to evaluate how representative data from selected snow stations are relative to the mean observed snow depth and SWE of the surrounding grid, (ii) to determine how snow station representativeness changes from accumulation to melt and at varying spatial scales and (iii) to evaluate the model assimilations of SWE by the National Operational Hydrologic Remote Sensing Center (NOHRSC) in the context of the snow station sampling biases.

Study area

This study focused on 15 sites in the western USA in three different geographic regions, including Sierra Nevada (SN), Southern Rockies (SR) and Pacific Northwest (PNW) (Figure 1). Within each region, sites were selected to include as much intraregional variability in climatic regime within the constraints of resources and field team travel times. Five sites in the SN were selected, including two sites in Yosemite National Park and three sites in the Eastern Sierra. The elevation of these sites ranged from 2149 m at Gin Flat to 2957 m at Rock Creek (Table I). As a result of the significant elevation and climate gradients

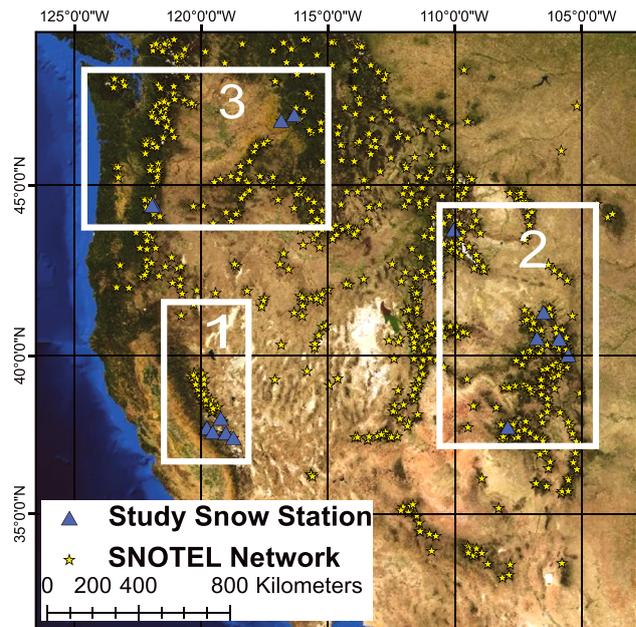


Figure 1. Western USA showing the areas of the three study regions: (1) SN, (2) US SR and (3) PNW. Snow station locations used in this study are marked with a blue triangle, the SNOTEL network is marked with yellow stars

associated with orographic precipitation, the variability in average annual maximum SWE accumulation at the SN sites is considerable at 800 mm (Table I). Within the USA SR region, six sites cover a broad range of climate conditions with a low elevation of 2573 m at South Brush Creek, Wyoming, and a high elevation of 3109 m at Lizard Head, Colorado. All sites were chosen for having long-term data – at least 25 years – and based on their variation in snow climatology built on work from Fassnacht and Derry (2010). A large latitudinal range is represented in this region from approximately 38°N to 43°N. The average SWE accumulation at these sites is highly variable, with a low of 335 mm at South Brush Creek, Wyoming, and a high of 640 mm at Togwotee Pass, Wyoming. Four sites in the PNW were selected; two sites in the interior PNW in Idaho and two sites in central Oregon in the Cascade Range were selected. The PNW sites cover a range of elevations from 975 to 1460 m, and average annual maximum SWE accumulation is highly variable with a low of 257 mm at Sherwin, Idaho, and a high of 991 mm at Hogg Pass, Oregon.

Study period

The study was carried out around the approximate time of maximum snow accumulation and near the middle of the 2008 and 2009 snowmelt seasons. In general, the time of maximum SWE accumulation was mid-March through mid-April, and the middle of the snowmelt season generally was in mid-April to mid-May depending on location (for exact survey dates, see Appendix A). Relative to the 30-year average 1 April snow course SWE values (anomalies from 1971 to 2000 average) at each site, the SR and SN sites generally had below

Table I. Attributes of the 15 snow stations used in the analysis

Site no.	Site	Location (dd)	Elevation (m) ^{a,b}	Maximum SWE (mm) ^c	Average temperature (°C) ^{a,b}	
					2008	2009
1	Gin Flat, California	37.768, -119.773	2149	930	8.6	7
2	Ostrander, California	37.638, -119.553	2499	958	7.7	6.8
3	Mammoth Pass, California	37.611, -119.034	2835	1176	2.7	3.7
4	Rock Creek, California	37.459, -118.736	2957	333	1.8	2
5	Virginia Ridge Lakes, California	38.072, -119.238	2879	483	4.9	4.9
6	Niwot Ridge, Colorado	40.037, -105.546	3021	318	2.2	3.1
7	Togwotee Pass, Wyoming	43.75, -110.059	2920	640	-0.6	0.1
8	Joe Wright, Colorado	40.532, -105.888	3085	546	1.1	2
9	Lizard Head, Colorado	37.802, -107.924	3109	409	1.4	2.3
10	Dry Lake, Colorado	40.534, -106.782	2560	577	3.4	4.4
11	South Brush Creek, Wyoming	41.328, -106.504	2573	335	3	3.8
12	Sherwin, Idaho	46.951, -116.34	975	257	6	6.7
13	Moscow, Idaho	46.807, -116.855	1433	533	5.7	6.7
14	Santiam Junction, Oregon	44.434, -121.945	1140	406	5.9	7.1
15	Hogg Pass, Oregon	44.42, -121.857	1460	991	4.9	6.1

^a Source: NRCS National Water and Climate Center.

^b Source: CA-DWR California Data Exchange Center.

^c Values are average values of 1 April SWE from 1971 to 2000, measurements from nearby snow courses and snow stations

average peak SWE values in both 2008 and 2009, whereas PNW sites had close to average or above average peak SWE values (Figure 2). Snow courses very close to snow stations in this study were used for this comparison because of their longer record at most sites.

METHODS

To assess snow station point-to-area biases of each site, we conducted distributed snow surveys around each station at peak snow accumulation and during snowmelt in both 2008 and 2009. At all sites, snow depths were interpolated to the 1, 4 and 16 km² area surrounding the

stations using binary regression tree models. Snow station biases were then examined (i) by evaluating the difference between observed snow depth at each station and the mean of the snow survey, (ii) by evaluating the difference between observed snow depth at each station and binary regression tree estimates of snow depth from the surrounding areas at 1, 4 and 16 km² scales and (iii) by evaluating SWE biases at 1, 4 and 16 km² scales using our snow density observations combined with binary regression tree results. We take this approach so that we can evaluate biases using both the direct snow survey observations and a statistical model that allows us to extend the analysis to scales larger than the sampled area.

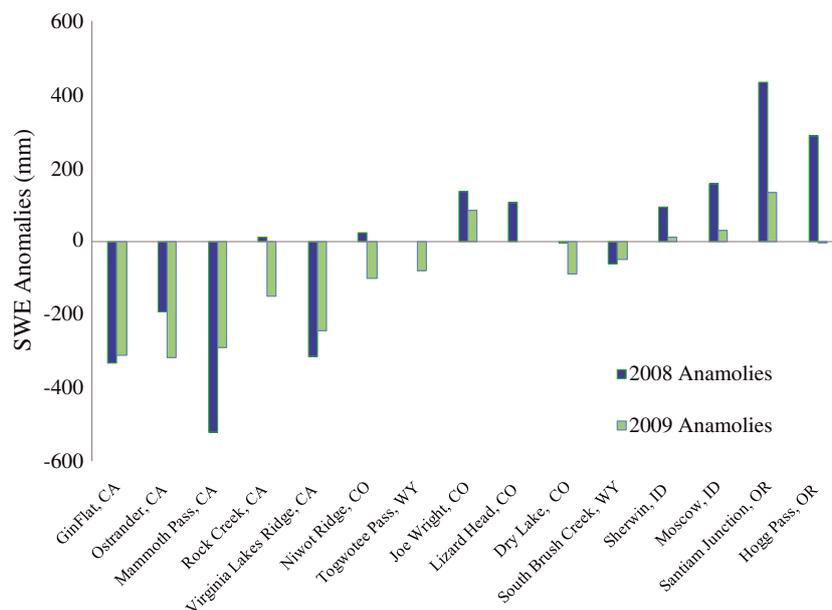


Figure 2. SWE anomalies of the two study years, 2008 and 2009, from the 1 April 1971–2000 average of snow courses and snow stations. Sites 1–5 are SN sites, sites 6–11 are SR sites and sites 12–15 are PNW sites

Moreover, by extending our bias estimation to SWE, we place these assessments in the context of the most hydrologically relevant variable. Detailed descriptions of these methods are described in the following sections.

Field methods

During each of the snow surveys, depth measurements were collected around each snow station using a 1-cm diameter aluminium probe along 11 transects spaced 100 m apart (Figure 3). Along each transect, points were sampled with a spacing of 50 m forming a 1-km² rectangle surrounding the snow sensor site for a total of approximately 231 point locations (Figure 3). Triplicate depth measurements were taken at each survey point for the SR and SN sites, whereas five measurements were taken at each point for PNW sites (centre, north, south, east and west). The number of depth samples for each survey and site can be found in the appendix with the number of observations made in 2008 totalling 15 379 and 19 761 in 2009.

Snow density, grain size and shape and snow temperature were measured with a minimum of one snow pit per site near the snow pillow, as well as one to four other pits in open areas and under canopy locations. Snow density was measured with a 1000-ml stainless steel snow cutter at 10-cm intervals; a 250-ml stainless steel snow cutter was used at some PNW sites at 10-cm intervals. At some surveys in the PNW, a standard Federal Sampler was used in the interest of time, given the deep snowpacks that developed at some of the higher elevation sites. Federal Sampler measurements were adjusted based on known biases (Work *et al.*, 1965; Goodison *et al.*, 1981). Density values from each pit were averaged to calculate an average site snow density value. Differences in snow density measurement protocols were considered when interpreting results.

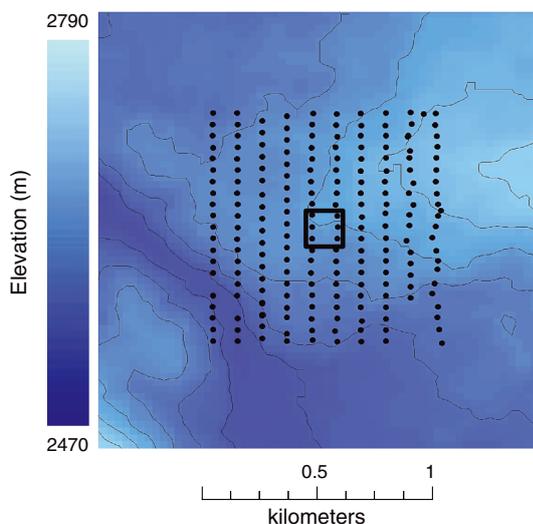


Figure 3. DEM (30-m resolution) and location of snow depth measurements around the South Brush Creek SNOTEL site. Contour lines are at 30-m intervals. Black centre square marks SNOTEL location. Black dots represent locations of snow depth measurements. At each dot, three or five depths were sampled

Snow station networks

Data from the SNOTEL and the California snow sensor networks were compared with our field measurements. In both sensor networks, each station generally consists of a snow pillow for SWE, an air temperature sensor and a storage precipitation gauge. Some enhanced SNOTEL stations also collect additional data including soil moisture and temperature, solar radiation, wind speed/direction, humidity, barometric pressure and precipitation (tipping bucket rain gauge).

Binary regression tree models

Measured snow depth point values were interpolated to 1, 4 and 16 km² scales at a 30-m resolution using binary regression tree models (Chambers and Hastie, 1993). In the regression tree model formulation, we used different combinations of the independent variables described in the next section to predict snow depth. The combinations of variables used in tree model development were determined through cross validation in which 100 iterations of model runs were performed (Molotch *et al.*, 2005). Variables and regression tree model sizes that minimized model deviance were selected as optimal predictors following commonly applied procedures (Balk and Elder, 2000; Erxleben *et al.*, 2002; Molotch *et al.*, 2005). The methods used to perform the cross validation and to construct the regression trees are thoroughly explained by Molotch *et al.* (2005).

At the watershed scale, snow depth and SWE distribution studies have shown that nonlinear relationships exist between snow accumulation and topographic variables (Elder *et al.*, 1991; Elder *et al.*, 1998; Molotch *et al.*, 2005). Regression tree snow depth models account for these nonlinearities and have been used in a variety of studies to model snow depth distribution (Elder *et al.*, 1998; Balk and Elder, 2000; Erxleben *et al.*, 2002; Winstral *et al.*, 2002; Molotch *et al.*, 2005). These empirically derived models are built through a process known as binary recursive partitioning, splitting the data into progressively more homogenous subsets until reaching a user-specified number of terminal nodes. The split at each branch is chosen such that the sum of the squared deviations from the mean is minimized (Chambers and Hastie, 1993). These regression models have limited utility over larger areas with relatively sparse observations because they may not adequately capture the range of physical parameters controlling snow dynamics. In situations where spatially dense snow sampling has been conducted, regression tree models provide a powerful means to resolve relationships between physiography and snow distribution to estimate the spatial variability of snow depth and SWE surrounding each snow station site in the study areas.

Regression tree model output consists of spatially continuous estimates of snow depth around individual snow stations. Biases using these snow depth surfaces were determined by comparing snow depth values at SNOTEL and California snow sensor locations to the average regression-tree estimated snow depth over the 1-, 4- and 16-km² areas surrounding each site. Similarly,

SWE biases were estimated by comparing snow station SWE observations to estimates of surrounding grid-element SWE. SWE distribution was estimated by multiplying the average snow pit density value at each site by the regression-tree modelled snow depth in each 30-m pixel and dividing by the density of water (Erxleben *et al.*, 2002; Molotch *et al.*, 2005). Far fewer density measurements (compared with depth measurements) can be used to characterize an area's SWE because density varies much less in a given area than depth (Goodison *et al.*, 1981).

Physiographic variables

We used the following independent variables to construct the regression tree models and to interpolate snow depth (dependent variable) at each site: elevation, index of clear sky, potential incoming solar radiation, percent canopy cover, slope, aspect and maximum upwind slope. The selection of these variables was based on results from related studies (Elder *et al.*, 1998; Balk and Elder, 2000; Erxleben *et al.*, 2002; Winstral *et al.*, 2002; Molotch *et al.*, 2005), which demonstrated that these variables explain snow distribution relatively well. Elevation data were obtained from the Shuttle Radar Topography Mission finished 1 arc second digital elevation model (DEM) for each site (Farr *et al.*, 2007). Slope and aspect were derived from the DEM using a geographic information system.

Shortwave radiation. A shortwave radiation index was calculated from the DEM for each 30-m pixel of each site using a variation of the methods of Molotch *et al.* (2005). Daily potential incoming shortwave radiation was calculated per pixel for the 15th day of each month using the ESRI ArcGIS Solar Analyst tool (Fu and Rich, 1999), starting from 15 November 2007 to the 15th day of the given survey month in 2008. For 2009 surveys, the daily potential incoming shortwave radiation was calculated from 15 November 2008 until the 15th day of the respective 2009 survey month. For each year, the average of these individual days was then calculated and used as an index of total snow season solar radiation (Molotch and Bales, 2005; Molotch *et al.*, 2005).

Maximum upwind slope. The mean maximum upwind slope (S_x) is a parameter that helps explain the variability in snow deposition and scour due to wind redistribution and specific terrain features (Winstral *et al.*, 2002; Molotch *et al.*, 2005). The meteorological data used to estimate prevailing wind direction were obtained from the following sources for each site: California Data Exchange Center (California sites), MesoWest (Oregon sites), Center for Snow and Avalanche Studies (Lizard Head Pass, Colorado), Desert Research Institute's Storm Peak Laboratory (Dry Lake, Colorado), NOAA's Quality Controlled Local Climatological Data from the National Climatic Data Center (Joe Wright, Colorado, and South Brush Creek, Wyoming), Idaho National Laboratory (Idaho sites), NRCS (Togwotee Pass, Wyoming)

and University of Colorado Ameriflux towers on Niwot (Niwot Ridge, Colorado).

Vegetation. Vegetation density can have a large effect on snow distribution by changing the energy balance at the snow-atmosphere interface, by intercepting precipitation and by affecting the surface roughness and winds that transport snow (McKay and Gray, 1981; Musselman *et al.*, 2008; Molotch *et al.*, 2009; Veatch *et al.*, 2009). To account for these processes in our regression tree model, we obtained percent canopy cover data as a proxy for vegetation density for each site from the National Land Cover Database (NLCD) (<http://www.mrlc.gov/>). The 30-m resolution NLCD values range from 0 to 1, where 0 indicates no canopy and 1 indicates a completely closed canopy.

NOHRSC National Snow Analysis data assimilation

The NOHRSC National Snow Analysis (NSA) system is a distributed model primarily used to estimate SWE distribution across the USA and parts of Canada. For a system of its scale, the NSA contains a relatively detailed physical representation of the processes that control snow accumulation and melt at a 1-km² spatial resolution. The model contains an explicit characterization of snowpack properties such as snow temperature, snow density and snow depth. The model energy balance includes an explicit representation of incoming and outgoing solar and long-wave radiation and turbulent fluxes. The model estimates of SWE are routinely compared with the observed SWE at snow stations, and if model errors exceed a given threshold, the ground-based snow station SWE observations are assimilated into the model (Carroll *et al.*, 2001). Because a snow station's depth or SWE do not necessarily represent the surrounding 1-km² grid cell in topographically complex regions, such assimilation procedures could benefit greatly from detailed assessments of snow station representativeness.

Considering NSA's modelled SWE is updated or assimilated to the observed value at a given snow station once a residual exceeds a threshold value, we examined the difference between the NSA-modelled SWE and the snow station observed SWE at the time of station SWE assimilation. NSA SWE and assimilation information were obtained from NOHRSC (<http://nohrsc.noaa.gov/>, accessed 10/2011). We then compared those differences to the point-to-area biases we found from field observations to assess NSA data assimilation within the context of snow station representativeness.

RESULTS

Field surveys

The mean snow depth of all Colorado surveys was 127.7 cm and ranged from 0 to 330 cm. The mean density was 340 kg m⁻³, and thus mean SWE was 434 mm (Table II). The mean SWE at the California sites was

Table II. Summary of 2008 and 2009 snow depth and snow density measurements made during the peak accumulation and mid-melt field campaigns^a

	2008		2009		2008		2009	
	Peak	Mid-melt	Peak	Mid-melt	Peak	Mid-melt	Peak	Mid-melt
	Colorado				California			
Minimum	0	0	0	7	0	0	0	0
Maximum	305	330	283	266	316	277	348	311
Mean	148.9	128.5	118.7	114.5	150.5	65.7	136.8	75.2
SD	31.7	54	48.8	40.7	63.5	59.6	77.5	78.9
Mean density	320	370	310	340	430	440	370	410
Mean station depth	148.3	114.3	113.8	105.5	156	66.8	148.4	81
	Idaho				Oregon			
Minimum	0	0	0	0	10	0	30	0
Maximum	290	227	325	214	441	356	425	306
Mean	134.1	76.5	127.6	37.5	269.5	184.4	203.8	104.5
SD	52.9	51.6	62.9	41.4	65.2	64.2	64.7	67.6
Mean density	383	446	330	440	400	505	380	440
Mean station depth	161.5	90.5	154.4	85.5	245.9	149.95	194.5	105

^a Snow depth values are in centimetres; snow density values are in kilograms per cubic metre.

439 mm, the average snow depth was 107.1 cm and the average snow density was 410 kg m⁻³; snow depths ranged from 0 to 348 cm. The average snow depth at Idaho sites was 93.9 cm, with depths ranging from 0 to 325 cm. The average density was 400 kg m⁻³, giving an average SWE of 376 mm. Similar to the California surveys, the more maritime snowpack in Oregon rendered higher average densities (430 kg m⁻³) and a much deeper snowpack (average depth, 190.6 cm; range, 0–441 cm). Thus, the greatest SWE was observed at the Oregon sites with a mean of 819 mm (for greater regional detail and survey-specific data, see Table II and Appendix A).

The determination of snow station representativeness is a fairly subjective process. We assess representativeness using 10% bounds around the mean and median observed values. We recognize the 10% threshold as being fairly arbitrary, but it provides a means to summarize our results across all sites. In this regard, 6 of 13 sites in the 2008 accumulation season had snow depth values at the snow station that were within 10% of the mean of the surrounding area. During the 2008 melt season, 5 of 11 sites were within the 10% threshold. During the 2009 accumulation period, 9 of 15 sites were representative of the surrounding grid at the 10% threshold, and 7 of 14 sites were representative during melt season assessments. Overall, 27 of the 53 surveys had SNOTEL or snow station biases within 10% of the surrounding mean observed depth, and 40 of 53 were within 20% of the surrounding mean observed depth (Figures 4a–4d).

Cumulative probability plots of snow depth near peak accumulation reveal the spatial variability around snow stations. Snow station representativeness based on the $\pm 10\%$ about the median threshold is indicated by the regions bounded by the dotted lines in Figure 5. These cumulative probability plots show the distribution of snow depth around SNOTEL sites with respect to the median snow depth value (intersection at the 0.5 probability value)

as well as areas above and below the 40th and 60th percentiles (intersection of the 0.4 and 0.6 probability values). Similarly, the shapes of these distribution functions reveal the nature of the variability in snow depth across the continuum of measured values; steeper slopes indicate relatively uniform distributions and shallow slopes indicate relatively heterogeneous distributions.

SR sites show little consistency between years, although most indicate that the SNOTEL underestimates snow depth (Figure 5a). Of the 53 surveys, 14 indicated that the SNOTEL provided a representative value for peak snow depth. The slopes of the cumulative probability lines are very similar, indicating similar spatial variability in snow depth at the SR sites except two Joe Wright surveys and the Togwotee Pass survey, which had more variable distribution as indicated by the lower slopes in the cumulative probability plots (Figure 5a). SR surveys from mid-melt also show only two sites (Dry Lake and South Brush Creek both in 2008) to be within 10% of the median observed snow depth value, demonstrating that most sites during melt were not representative of the surrounding 1-km² grid. Interestingly, most cumulative probability plots at the SR sites (Figures 5a and 5b) show similar levels of spatial variability in snow depth as indicated by the relatively consistent slopes of the lines. Thus, sites such as Joe Wright consistently deviated from the other sites.

At the SN sites, depth biases were relatively consistent from year to year during the accumulation surveys. In particular, Ostrander and Gin Flat sites had positive biases in both years, and the relative magnitudes of these biases were consistent from year to year, falling in the 70th to 85th percentiles in both years. Similarly, both the Mammoth Pass and the Rock Creek sites reported negative biases during accumulation in both years with biases mostly falling in the 30th to 50th percentile. Only one site in the SN region during accumulation fell within 10% of the median snow depth (Figure 5c). During mid-melt,

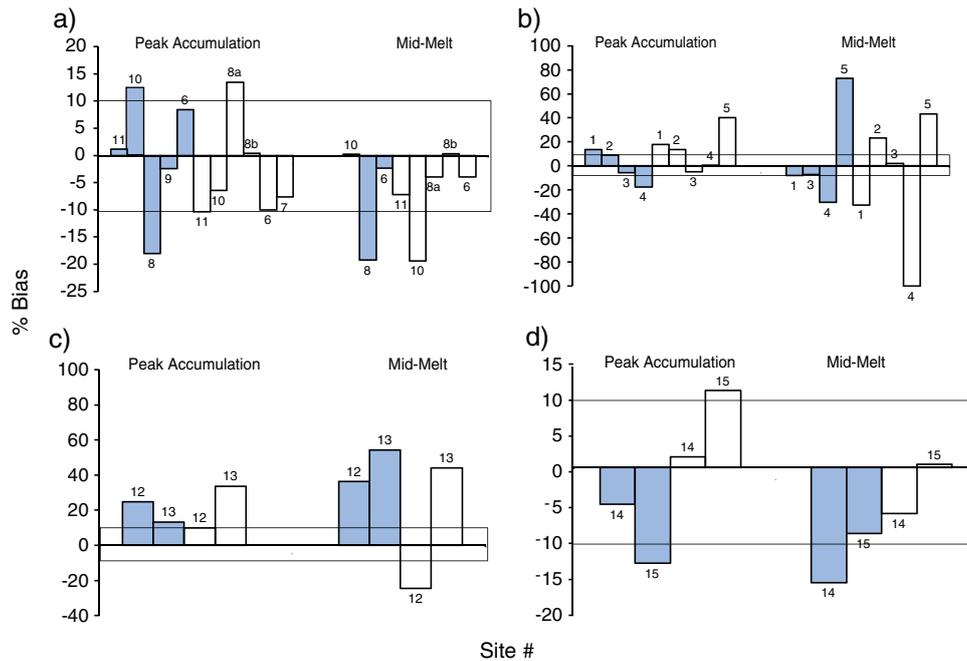


Figure 4. Percent bias of each snow station compared with the surrounding mean of observed depths; for example, 13% positive bias means the snow station depth was 13% greater than the surrounding observed mean snow depth. (a) SR sites, (b) SN sites, (c) Idaho sites and (d) Oregon sites. Sites are identified as they are numbered in Table I; for example, site 11 is South Brush Creek, Wyoming. Shaded bars are percent biases from 2008 surveys, and white bars are percent biases from 2009 surveys. Boxes mark 10% above and below mean observed snow depth. Note that different scales were used in each plot. Joe Wright a and b in 2009 are due to the additional surveys at during peak accumulation and mid-melt

however, three SN sites were within 10% of the median snow depth of the surrounding areas (Figure 5d). Overall, observed snow depths exhibited a larger range in the SN versus the SR.

At the PNW sites, depth biases were somewhat consistent from year to year during accumulation. The Moscow Mountain site overestimated snow depth relative to the surrounding area with depth values falling in the 70th to 90th percentile in both years. Conversely, the Sherwin site was more representative in both years with SNOTEL depth values falling around the 60th percentile relative to the surrounding area. The Hogg Pass and the Santiam Junction sites exhibited neutral to negative biases in both years falling between the 10th and the 50th percentile (Figure 5e). PNW sites show similar slopes on the cumulative probability plots with more even distributions during accumulation than during mid-melt. Mid-melt slopes on the cumulative probability plots for Hogg Pass in both 2008 and 2009 show a tail in snow depth distribution more skewed toward shallower depths. However, like California sites, there is a large amount of variability in snow depth from site to site; for example, Hogg Pass had consistently greater depths than the other sites during both survey seasons. During both accumulation and mid-melt surveys, three SNOTEL sites fell within 10% of the median snow depth of the surrounding areas (Figures 5e and 5f). Thus, both mean and median-based assessments of 'representative' snow stations show that less than half of all surveys indicated representative snow stations. Furthermore, the station biases were frequently consistent in sign and magnitude from year to year – an important discovery that is revisited in the discussion section.

Binary regression trees

The regression tree models explained 47% of the spatial snow depth variability on average across all sites and years. R^2 values varied greatly but were greater than 0.4 for 23 of the 39 stable models, and 10 were greater than 0.6, indicating relatively strong explanatory capability. Idaho sites had the highest average R^2 values averaging 0.57 during the accumulation surveys and 0.61 during mid-melt surveys. Averaged over all sites in all regions, R^2 values during the accumulation season (0.44) were 14% lower than during mid-melt (0.50), indicating strong statistical relationships between independent variables and snow distribution during the ablation season (Table III).

Regression tree model size was constrained based on the cross validation of model deviance and tree size (Molotch *et al.*, 2005). The number of terminal nodes, which minimized model deviance while maximizing the returned tree size, was selected. By maximizing tree size near the minimum model deviance, an increased R^2 is achieved. For example, model deviance at Dry Lake, Colorado, in May 2008 reached a minimum of approximately 12 terminal nodes with the model using the independent variables solar radiation, elevation, maximum upwind slope, vegetation and slope. Although R^2 values are not criteria for tree size selection, it is important to note that R^2 values began to plateau at more than 12 terminal nodes. Hence, the model deviance criterion avoids overfitting the model (Figure 6). The regression tree was constructed for this site using the aforementioned variables and specifying 12 terminal nodes (Figure 7). This procedure was repeated for all regression tree models.

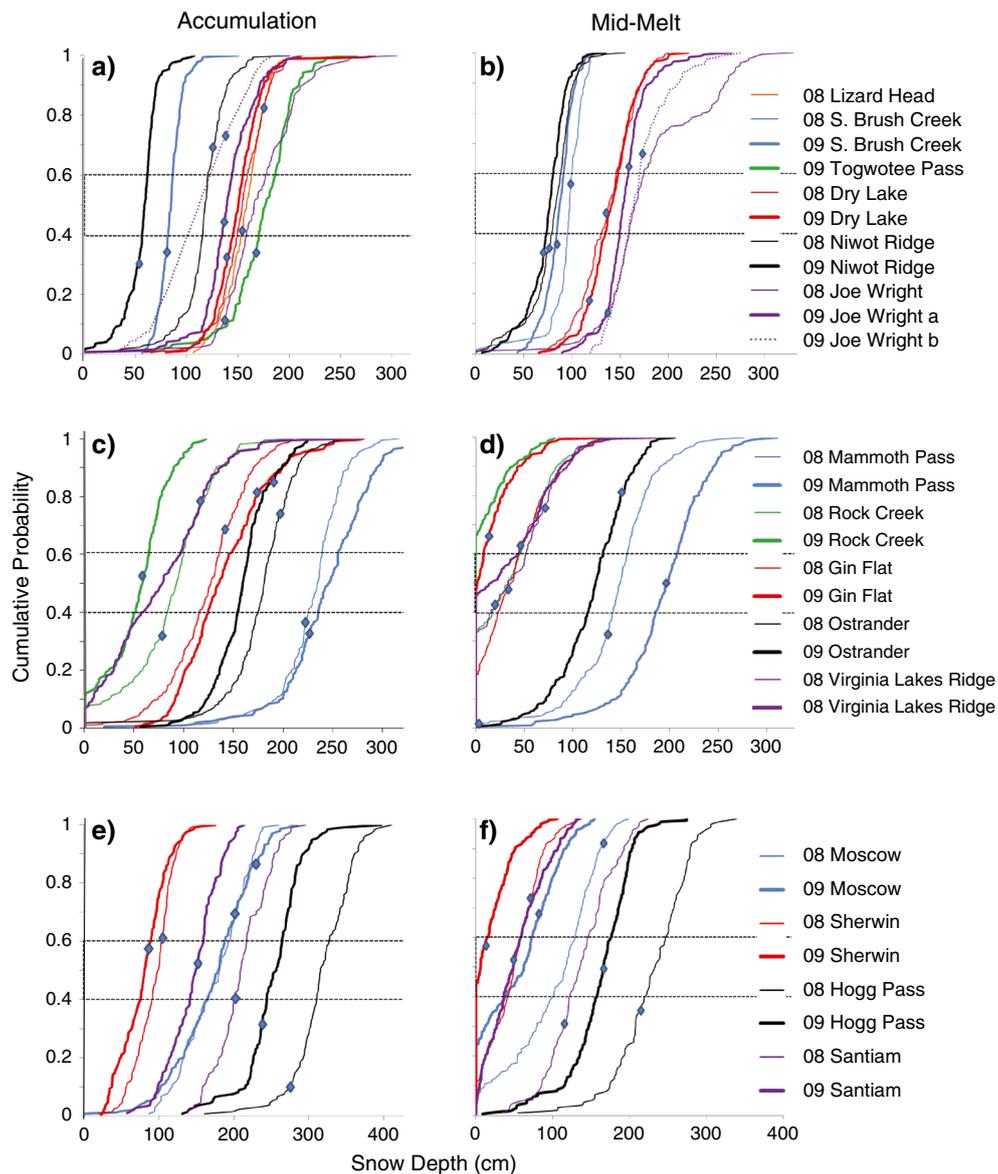


Figure 5. Statistical distribution of snow depth. Snow station depth values are marked with diamonds, and the dotted line region highlights $\pm 10\%$ of the median snow depth values. Plots of snow accumulation in the SR, SN and PNW are shown in plots a, c and e, respectively. Mid-melt plots for the SR, SN and PNW are shown in panels b, d and f, respectively

Regression tree snow depth models produced from data at some sites were unstable. This was indicated by a lack of minima in graphs of deviance versus the number of terminal nodes (Chambers and Hastie, 1993). Unstable models were produced from data at five sites during peak accumulation surveys and four sites during mid-melt surveys (Table III). Model instability at these sites was likely due to in part the relatively low topographic variability that diminishes relationships between snow depth and physiography. Less varied topography will generally require a much larger sample size to produce a stable model. In total, 53 surveys were conducted. However, the four surveys at Hogg Pass and the one at Togwotee Pass were not used to create regression trees because not all independent variables could be obtained. There were 48 regression assessments that in total produced 39 stable regression tree models. Thus, snow station

representativeness assessments based on regression tree models were limited to these 39 cases with stable models.

The size of the regression tree models also varied from site to site, from year to year and for accumulation versus ablation season. Overall regression tree model sizes six and seven terminal nodes for the 2008 accumulation and ablation seasons, respectively. In 2009, tree model size was seven terminal nodes for the accumulation and 8 for the melt season on average. Tree model sizes ranged from 4 to 12 terminal nodes in 2008 and from 4 to 13 terminal nodes in 2009. Regional differences show that California and Idaho tree models had the largest average number of terminal nodes (Table III).

The identification of the physiographic variables controlling snow distribution may provide an explanation of snow station biases relative to the surrounding areas. In this regard, the highest-ranking variables (level 1 in the

Table III. Cross validation of R^2 results

Site	Peak 2008	No. T nodes	Mid-melt 2008	No. T nodes	Peak 2009	No. T nodes	Mid-melt 2009	No. T nodes	
1	Gin Flat, California	0.44	9	0.41	5	0.46	10	0.42	5
2	Ostrander, California	N/M	–	N/A	–	0.44	10	0.50	10
3	Mammoth Pass, California	0.30	6	N/M	–	0.35	7	0.24	5
4	Rock Creek, California	0.45	6	0.42	5	0.38	7	N/M	–
5	Virginia Lakes Ridge, California	N/A	–	0.60	10	0.64	11	0.65	8
6	Niwot Ridge, Colorado	0.38	–	0.35	5	0.28	5	N/M	5
7	Togwotee Pass, Wyoming	N/A	–	N/A	–	N/A	–	N/A	–
8	Joe Wright, Colorado	0.26	4	0.60	8	0.28	4	0.20	4
9	Lizard Head, Colorado	0.30	5	N/A	–	N/A	–	N/A	–
10	Dry Lake, Colorado	0.48	5	0.70	12	0.42	5	0.59	8
11	South Brush Creek, Wyoming	0.56	–	N/A	5	N/M	–	0.28	5
12	Sherwin, Idaho	0.39	6	0.29	5	0.39	6	N/M	–
13	Moscow Mountain, Idaho	0.81	12	0.76	10	0.67	9	0.79	13
14	Santiam Junction, Oregon	0.15	–	0.30	4	0.58	7	0.81	11
15	Hogg Pass, Oregon	N/A	–	N/A	–	N/A	–	N/A	–

N/A, no survey was performed at that time or that the data were otherwise unable to be processed; N/M, the model was unstable with no minimum deviance; No. T nodes, the number of terminal nodes specified in the regression tree construction.

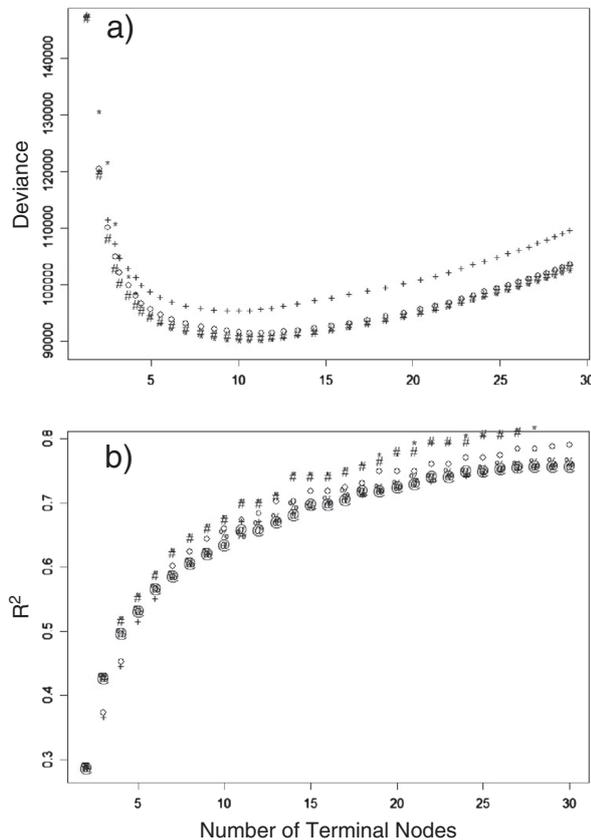


Figure 6. Plots of (a) model deviance versus number of terminal nodes and (b) R^2 versus number of terminal nodes for Dry Lake, Colorado, May 2008. These indicate optimal tree size in this case is 12 terminal nodes

regression tree; e.g. Figure 7) controlling snow depth were solar radiation and maximum upwind slope during the 2008 and 2009 accumulation season; aspect was also significant during accumulation in 2009 (Table IV). Solar radiation was a consistent correlate of snow distribution as indicated by its persistent appearance in upper tree levels in both years. Similarly, elevation was consistently

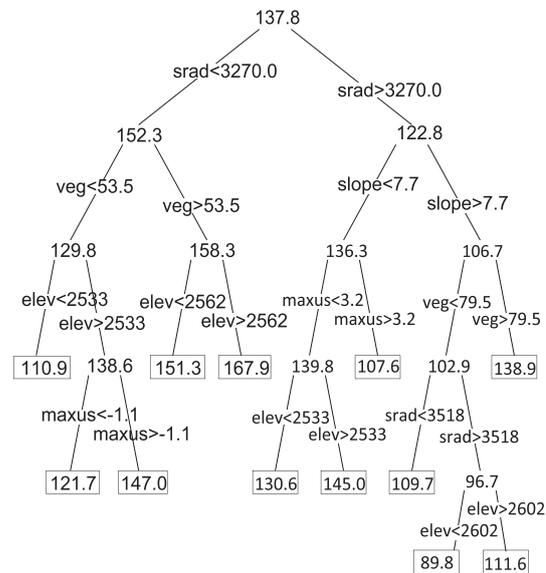


Figure 7. Regression tree snow depth model, Dry Lake, Colorado (May 2008), with 12 terminal nodes. Each node represents a snow depth; regular nodes have no surrounding shape, and terminal nodes are rectangles. Srad, potential shortwave solar radiation; elev, elevation; veg, 2001 NLCD percentage of canopy cover data; maxus, maximum upwind slope

important as indicated by its frequent presence in the second and third tree levels. Vegetation was also important for many sites, with the percentage of canopy cover frequently appearing in the second regression tree level (Table IV). Slope was typically ranked lower than other variables, although there is a significant correlation between the slope and aspect and the solar radiation index. Solar radiation likewise dominated the highest levels for the 2008 and 2009 melt season models, with the NLCD percentage of canopy cover as the second most frequently occurring variable in the first tree level (Table IV).

Maps of SWE distribution (Figure 8) help illuminate snow station biases when examined in tandem with bias

Table IV. Frequency of appearance of physiographic variables at different levels in regression tree snow depth models for peak accumulation and ablation surveys^a

Accumulation	Potential SW radiation		Elevation (m)		MAXUS		NLCD		Slope		Aspect	
	2008	2009	2008	2009	2008	2009	2008	2009	2008	2009	2008	2009
Tree level												
1	3	4	1	1	3	2	1	1	0	0	1	3
2	5	4	5	5	1	1	5	4	1	1	0	2
3	1	2	5	6	4	6	2	1	2	3	0	3
4	0	6	3	1	2	3	1	2	0	1	0	0
5	0	2	1	2	1	1	0	1	0	0	0	1
6	1	0	0	1	0	1	0	1	0	0	0	1
Ablation												
1	5	4	0	0	2	1	3	3	0	0	0	1
2	2	3	3	3	6	1	5	1	2	4	1	6
3	3	6	8	10	5	2	2	1	0	0	0	0
4	7	3	1	2	2	3	0	2	0	0	0	2
5	0	1	7	4	1	0	0	1	0	0	0	1
6	0	2	0	0	1	0	1	0	1	0	0	0

MAXUS, maximum upwind slope.

^a For example, a value of six for solar radiation at tree level four indicates that solar radiation appeared in the fourth level on six of the models that year.

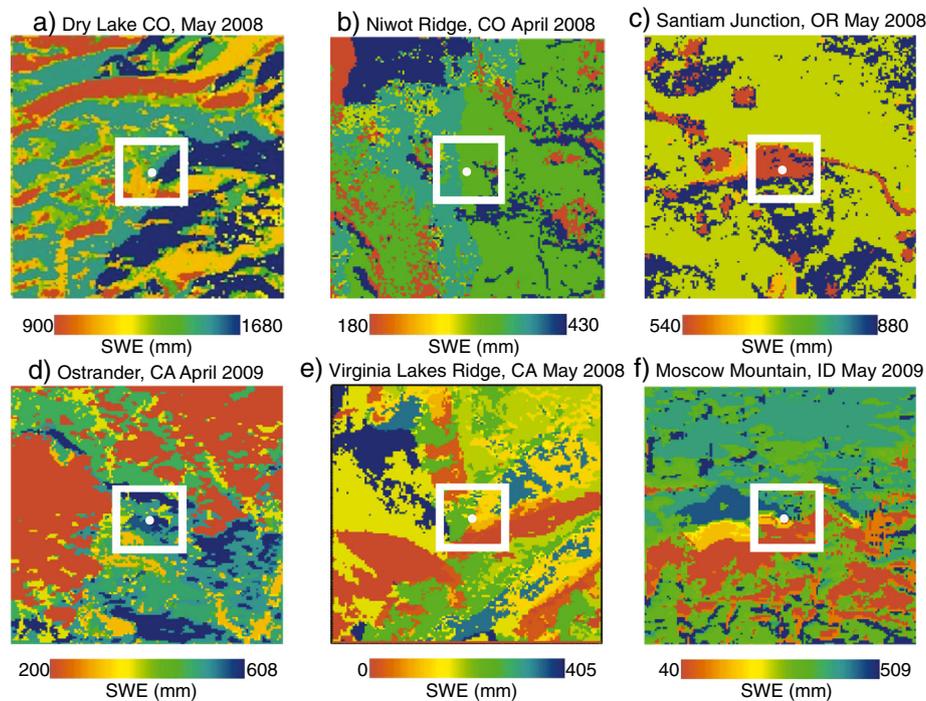


Figure 8. Spatial distribution of modelled SWE during mid-ablation periods. Each map is 4×4 km; the centre white box (1 km^2) highlights the area surrounding the SNOTEL or snow station survey location (white dot) and the terrain immediately around it. Note that different scales are used to illustrate spatial variability in SWE at each site

and cumulative probability plots (Figures 4 and 5). The first panel in Figure 8 shows a very high degree of modelled spatial variability in SWE at Dry Lake, Colorado, in May 2008. The most highly ranked variable on the regression tree for this site and time was solar radiation followed by vegetation and slope, the patterns of which can be seen in the ribbon-like features from east to west. Note that despite this considerable spatial variability, the snow station SWE value was quite representative (Figure 8a, white box; Figure 4a, Site #10; Figure 5b, red line). At the Niwot Ridge, Colorado, site, the SNOTEL SWE value was also quite representative whereas the spatial variability of SWE around the SNOTEL was considerably lower (Figure 8b). In this regard, we see

that Niwot Ridge consistently had representative SWE values (Figure 4a) because of the uniform distribution and representative placement of the station. At the Santiam Junction, Oregon, site, we also note a fairly uniform spatial SWE pattern but a low bias in the snow station location (note the light green of surrounding area vs the orange area at SNOTEL in Figure 8c); these low biases occurred in three of four assessments (Figure 4d). At the Ostrander, California, site, we note a spatial SWE pattern that varies at a relatively larger scale with the snow station location in an area of relatively high accumulation (observe the northwest to southeast pattern of increased SWE shaded blue-green in Figure 8d). This site consistently had high (>10%) biases in

all three assessments (Figure 4b). The Virginia Lakes Ridge, California, site is placed in an area of high SWE accumulation (note the green area in Figure 8e); SNOTEL depth values consistently overestimated the surrounding area by greater than 40% (Figure 4b). Similarly, the Moscow Mountain, Idaho, site was placed in an area of relatively high SWE accumulation (blue-green area in Figure 8f) with snow depth biases exceeding 30% in three of four assessments (Figure 4c).

Because of the spatial heterogeneity of snow depth and SWE, the modelled snow depth and SWE also vary at different spatial scales. On average and across all scales, SWE biases (the snow station SWE point observation compared with the mean surrounding modelled SWE) were of greater magnitude during mid-melt compared with accumulation. Of the 15 pairs of accumulation versus mid-melt assessments, 9 assessments show increases in bias during mid-melt (Figure 9). Whether the biases were positive or negative tended to follow different patterns by region. For models at peak accumulation, 1- and 4-km² scales had a lower magnitude bias, thus indicating that SWE is more accurately estimated at smaller spatial scales. The total magnitude of biases was greater in 2009 versus 2008 across all three spatial scales and regardless of time in the snow season.

The one exception to this is the peak accumulation bias at 16 km²; the average bias at this scale and seasonal timing was 9.97 cm for 2008 and slightly less at 9.57 cm for 2009 (Table V). These trends are further discussed in the following section.

An opposite trend is evident in the model biases at the three spatial scales during accumulation compared with melt. At peak accumulation, the magnitude of the model biases increases with increasing spatial scale from 1 to 16 km². However, the model biases generally decrease as spatial scale increases during snowmelt (Table V). Of the 39 models, 21 had at least a 15% difference in SWE bias across all three spatial scales (Figure 9). Of these sites, just over half were from models during snowmelt assessments.

The percent bias plots modelled at the three spatial scales reveal few sites with consistent biases from peak accumulation to mid-melt or from year to year. Only two sites during the same seasonal timing in 2008 and 2009 (site 1, Gin Flat peak accumulation; site 2, Mammoth Pass peak accumulation; Figures 9c and 9d) show a similar sign and magnitude percent bias. From peak accumulation to mid-melt within the same year, only three sites (site 8, Joe Wright in 2008 and 2009; and site 13, Moscow in 2009) show similar sign and magnitude of model percent bias (Figures 9a, 9b and 9f).

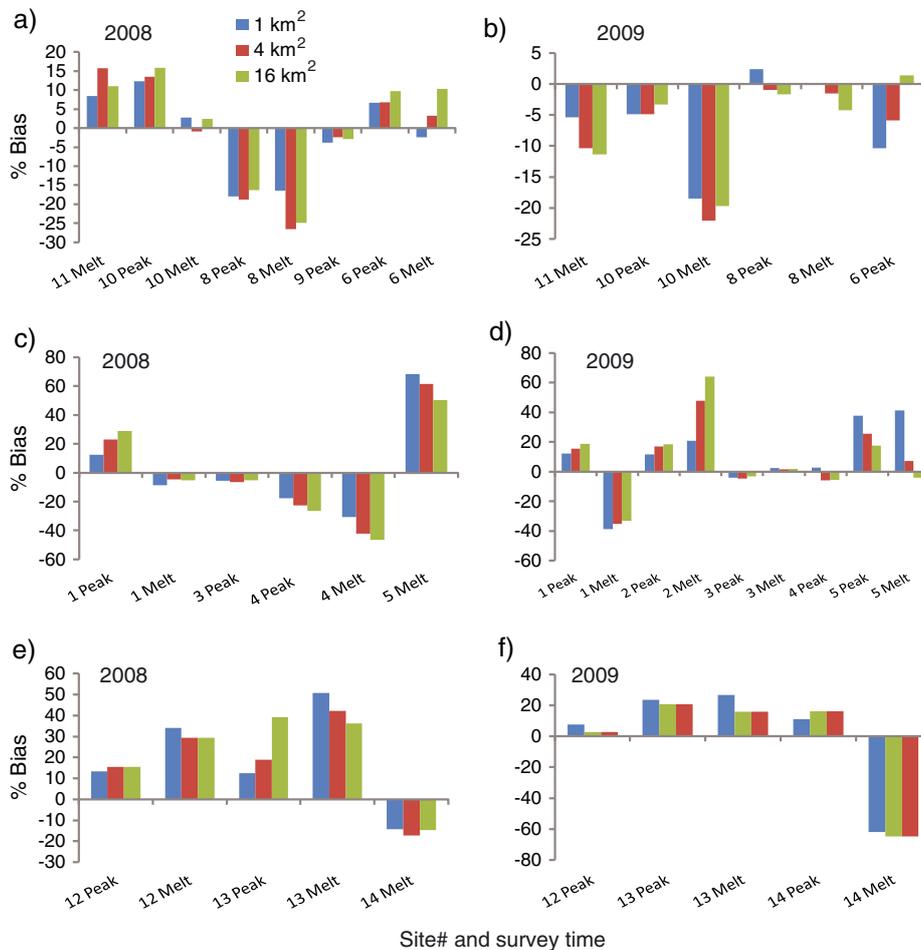


Figure 9. Modelled SWE percent bias (snow station SWE – average modelled SWE) at three spatial scales. Plots are for 2008 in the (a) SR, (c) SN and (e) PNW. The 2009 plots are (b) SR, (d) SN and (f) PNW. Sites are numbered as in Table I

Table V. Statistics of SWE percent bias at three scales

	1-km ² peak	1-km ² melt	4-km ² peak	4-km ² melt	16-km ² peak	16-km ² melt
Colorado 2008	-0.70	-1.89	-0.24	-2.11	1.57	-0.31
Colorado 2009	-4.27	-7.98	-3.90	-11.31	-1.19	-11.76
California 2008	-3.60	9.56	-2.12	4.88	-1.01	-0.51
California 2009	12.07	6.40	9.53	5.36	9.21	7.21
Idaho 2008	12.91	42.41	17.20	35.83	27.33	32.81
Idaho 2009	19.29	40.43	15.68	26.76	11.76	16.00
Oregon 2008		-14.32		-17.22		-14.63
Oregon 2009	-0.20	-53.15	11.01	-61.80	16.13	-64.65
Average 2008 ^a	5.74	17.05	6.52	15.01	9.97	12.06
Average 2009 ^a	8.96	26.99	10.03	26.31	9.57	24.91
Total average ^a	7.58	22.02	8.52	20.66	9.74	18.48

^a Averages are all averaged absolute values of percent bias.

NSA NOHRSC application

As an example application of how the snow station biases reported earlier can affect hydrologic model results, we focus on the NOHRSC NSA system (Carroll *et al.*, 2001; Rutter *et al.*, 2008). Following the work of Molotch and Bales (2005), we examined the difference between the NOHRSC-modelled SWE and the snow station observed SWE. NOHRSC evaluates residuals between the observed SWE at snow stations and their modelled 1-km² SWE estimates on a daily basis. On the basis of the magnitude of these residuals, a decision is made as to whether modelled SWE will be updated (assimilated) to the observed value. We compared the NOHRSC-modelled SWE, observed SWE and our observed biases just before an 'assimilation' period and found that, in many circumstances, the bias in NOHRSC's model is consistent but inversely proportional to the bias of the snow station measurement. This indicates that the model is actually estimating grid-element SWE better than the point observation at the snow station.

For the 50 of our surveys for which NOHRSC has modelled SWE, 31 surveys displayed this relationship. Specifically, 18 surveys showed a positive snow station bias – the snow station overestimating SWE with regard to the surrounding 1 km² survey – coinciding with the NOHRSC model underestimating SWE and assimilating the model to the snow station observation. Virginia Lakes Ridge in 2008 demonstrates this case well. The NOHRSC model underestimates the SNOTEL observed SWE (Figure 10a); however, according to our snow survey data, the observed SWE at the SNOTEL site is much higher than the observed SWE over the surrounding 1-km² area. Specifically, average observed SWE was 17 cm less than the SNOTEL SWE, resulting in a 74% overestimate at the SNOTEL. The consideration of this bias when evaluating NOHRSC model residuals in this example would indicate that updating the model to the observed SWE might be unwarranted.

Thirteen surveys showed that snow stations were underestimating SWE with regard to the surrounding 1 km², whereas the NOHRSC model overestimated SWE. The NOHRSC model bias and the surveyed percent bias at Santiam Junction in May 2008 show that the model

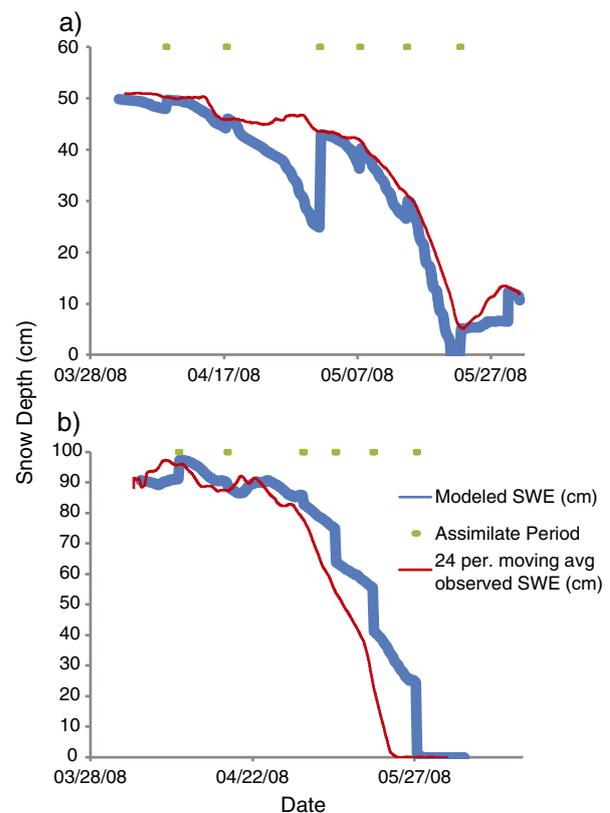


Figure 10. NOHRSC models of SWE over the 1-km² grid surrounding various snow station sites. Observed SWE is from the snow pillow, not manual measurements. (a) Virginia Lakes Ridge during May 2008; (b) Santiam Junction in May 2008

overestimated the observed SNOTEL SWE, but the point SNOTEL SWE significantly undersampled the SWE relative to the surrounding grid element. The average observed SWE across the grid was 12 cm greater than the SNOTEL SWE, resulting in a 16% underestimate at the SNOTEL (Figure 10b). In this example, the NOHRSC model is again estimating the true grid-element SWE better than the SNOTEL value, and thus assimilations of the SNOTEL value would deteriorate model SWE estimates and may therefore be unwarranted.

If we take into account these point-to-area biases of the snow stations, we can derive an adjusted NOHRSC model

bias that better represents the true residual. We were able to show this improvement by adjusting the bias for 22 assessments (Figure 11). For example, site 3 (Mammoth Pass, California) in Figure 11a shows that the snow station underestimates SWE with regard to the surrounding grid by approximately 5 cm, but the NOHRSC model overestimates SWE by approximately 4 cm, and thus the adjusted model bias is 1 cm of SWE. We note that at the three SN sites for which NOHRSC does not take point assimilations (Ostrander Lake, Mammoth Pass and Rock Creek), we used modelled SWE values for the grid encompassing each snow station location.

DISCUSSION

An unprecedented amount of distributed SWE data enabled a geographically comprehensive analysis of snow station representativeness and terrain controls on small-scale snow depth variability in the western USA. Just under half (11 of 24) of the surveys conducted during the accumulation season had snow station depth values that were within 10% of the mean observed depths over the respective surrounding areas. Similarly, just over half (15 of 29) of the surveys conducted during the melt season had snow station depth values within 10% of the mean observed depth (Figure 4). This demonstrates snow depth and SWE biases from the point-to-area scale and illustrates the large amount of subgrid variability that is not typically accounted for in distributed land surface and hydrologic models, even on the order of a 1-km² resolution. One way to resolve this issue may be to

redesign the sensor network such that four or five sensors are placed at optimal locations (Rice and Bales, 2010) instead of using one sensor to represent a larger, spatially variable area. This could be sufficient to better estimate the 1-km² spatial mean (Neumann *et al.*, 2006).

Snow depth during the ablation period exhibited greater subgrid variability. More variability during melt is intuitive, given the high degree of spatial heterogeneity associated with topographic controls on incident solar radiation, which largely controls snowmelt distribution (Marks and Dozier, 1992; Cline, 1997), and solar radiation increases through the melt season. The high tree level ranking of maximum upwind slope as well as the frequency of elevation in the second and third levels of the regression trees highlights the strength of the topographic control on snow accumulation associated with orographic effects and redistribution of snow by wind as found by Anderton *et al.* (2004). Not all sites had variables appearing consistently at certain levels. However, given the overall consistencies from year to year in the variables that control snow distribution, there are inherent consistencies in the biases at the snow stations as noted previously. Hence, the results from this 2-year study provide a means whereby biases evaluated during this study can be applied to other years and potentially in real time. In addition, we note that although our results indicate consistent biases from year to year in many cases, these results must be viewed with caution as they are based on only 2 years of data. Further study is needed to identify consistencies in snow station biases over several years.

This higher observed subgrid variability during snowmelt is also reflected in the modelled SWE (Figure 9 and Table V). Differences in vegetation density will also affect the net radiation and therefore influence melt timing and magnitude as well (Faria *et al.*, 2000; Veatch *et al.*, 2009), hence, it is expected that differences will increase between any single point and the spatial mean. This variability hampers ablation season modelling efforts that are critical in numerous hydrological and ecological studies focusing on the spring melt pulse in snowmelt-dominated water systems (DeBeer and Pomeroy, 2010). Thus, the consideration of variability during both the accumulation and ablation season in sensor network design is needed (Rice and Bales, 2010).

Additional complications in hydrologic modelling include the greater variability of snow distribution in lower accumulation years. As a result, significant nonlinearities exist with regard to snowmelt model performance (e.g. the NOHRSC model). The different results observed here across sites and years are partially attributable to these nonlinearities.

Differences in SWE distribution and modelled SWE during snowmelt may be attributed to a few different factors. Using mean snow density to interpolate SWE assumes uniform density, which is a tenuous assumption. However, Erxleben *et al.* (2002) indicated that snow density is not well correlated to physiography, and therefore complex interpolation models are unwarranted. Differences in SWE distribution and modelling during snowmelt may also be attributed to the differences in

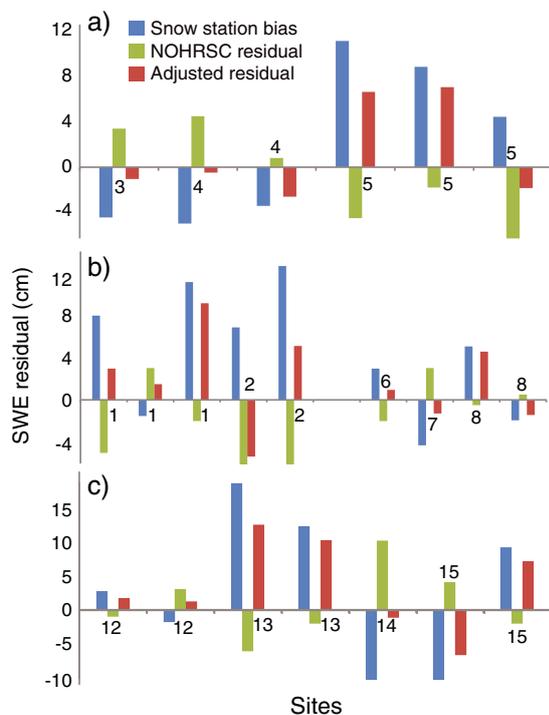


Figure 11. Plots of snow station bias (snow station - mean surrounding), NOHRSC residuals and residuals adjusted if NOHRSC model residual is added. Plots show (a) Eastern SN sites, (b) Yosemite, California, and SR sites and (c) PNW sites. Plots are numbered as in Table I

snow density observation methods in this study. As found by Fassnacht *et al.* (2010), core samplers tend to underestimate snow density compared with the more labour-intensive traditional snow pits, particularly when a snowpack is melting (Fassnacht *et al.*, 2010). Snow pits were dug at most surveys in this study; however, some PNW surveys used a Federal Sampler (snow tube) or both snow pits and a Federal Sampler to measure snow density. We recognize that this may lead to potential inaccuracies in SWE interpolation.

Examining modelled SWE at different spatial scales, we found that the highest SWE biases at peak accumulation occurred at the coarser scale of 16 km² with the 4-km² model, also showing higher biases than at the 1-km² scale. This has relevance to other multiresolution snowpack modelling studies (Pavelsky *et al.*, 2011), which found better agreement between model and point SWE when models were run at smaller spatial scales (9 and 3 km²). Hence, our study can help inform bias assessments for high resolution snowpack modelling (Anderton *et al.*, 2004; Caldwell *et al.*, 2009).

Many in the hydrologic community assume point data to be representative of an area for model evaluation, data assimilation and refining remote sensing algorithms for retrieval of snowpack characteristics (Daly *et al.*, 2000; Fassnacht *et al.*, 2003; Pan *et al.*, 2003; Andreadis and Lettenmaier, 2006; Durand *et al.*, 2008). However, considering many biases found in this study, the NOHRSC SWE model is actually estimating grid-element SWE better than the point observation at the snow station, explicitly accounting for the snow station biases that might improve the NOHRSC SWE assimilation procedure and better inform other modelling and data assimilation efforts. Given the widespread use of the NOHRSC SWE product, improvements in the SWE estimates will have dramatic effects in a variety of scientific and management disciplines (Azar *et al.*, 2008).

We acknowledge that using the empirically derived regression tree models to extrapolate snowpack information to a 16-km² grid may be problematic if the larger-scale landscape encompasses a wider range of variables than were used to create the model. However, for the purposes of this study, we feel regression trees are justified because the extrapolated area is relatively small and close to the area where data were collected.

Of the previous studies using binary regression tree snow distribution models (Elder, 1995; Elder *et al.*, 1998; Balk and Elder, 2000; Erxleben *et al.*, 2002; Winstral *et al.*, 2002; Anderton *et al.*, 2004; Molotch and Bales, 2005; Molotch *et al.*, 2005; Musselman *et al.*, 2008), only those of Erxleben *et al.* (2002) and Molotch and Bales (2005) were conducted at similar spatial scales and in areas with relatively dense canopy cover. The explanatory power of the regression tree models developed in the present study (average $R^2=0.47$) was good relative to these two previous works (average R^2 values of 0.25 and 0.56, respectively), particularly because the average R^2 value of the top ten models in this study was 0.7. The variability in regression model structure and explanatory

power at the different sites is consistent with these previous studies. This variability affirms the lack of transferability of regression trees over larger scales (Elder, 1995).

A major issue with snow station location physiography is that they tend to be located in small forest gaps on flat ground. The logistical difficulties of setting up instruments on slopes and in dense vegetation can largely account for snow station placement. This signifies a methodological need to track SWE on different slopes with more varying canopy coverage. One suggestion is that the SNOTEL and California snow sensor networks could be expanded to include a small cluster of depth sensors and snow pillows so that SWE could be estimated from the snow station-derived snow density. With the advancement of wireless sensor networks (Kerkez *et al.*, 2010), the opportunity exists to establish a network of instrument clusters to enhance the current SWE measurement network. In addition, future snow measurement networks could use the methods performed here to identify locations that can consistently provide robust estimates of grid cell mean SWE.

We have made the case that the SNOTEL bias should be considered when doing model assimilations, yet the biases are not always consistent between years and at different times during the snow season. This indicates that intensive surveys performed either manually or by remote sensing techniques will not necessarily reveal general SNOTEL site biases; they may just show biases at the particular moment of the survey. There are several factors contributing to the relationships between SNOTEL point biases and NOHRSC model residuals. Further study is needed for an exhaustive analysis of these factors. One noteworthy point is that several snow stations are located in small forest clearings, which favour the oversampling of SWE – particularly around the time of peak accumulation – compared with the densely forested immediate surroundings (Gin Flat and Virginia Lakes Ridge in California and Niwot Ridge in Colorado). Most surveys at these sites did, in fact, show that the snow station oversampled SWE relative to the surrounding grid but the NOHRSC model undersampled SWE (Figure 11a, b).

These bias relationships suggest that snow monitoring strategies should be altered and that there is a need to find another way to more accurately quantify snow distribution over scales of 1 km² or larger. Regression tree analysis provides valuable information on key variables affecting snow depth and could hence be used to design generalized monitoring strategies as well as specific strategies if resources exist for some form of intensive manual surveys as were performed in this study. In addition, we only study 15 sites and what is greatly needed is a generalized approach to estimate biases at all of the more than 700 SNOTEL sites in the west based solely on physiography. The shaped solution approach of Melloh *et al.* (2008) is one potential direction; modes of snow distribution variability can be described in three-dimensional space based on a variety of terrain-related explanatory variables.

The data collected here could be used to derive shaped solutions of snow depth (Melloh *et al.*, 2008). Shaped solutions that are consistent between sites of similar physiography can be aggregated to develop ensembles of generalized models that could be used to characterize biases across the SNOTEL network.

CONCLUSION

Overall, we found that 27 of the 53 surveys had SNOTEL or snow station biases within 10% of the surrounding mean observed depth, and 40 of the 53 surveys were within 20% of the surrounding mean observed depth. This indicates that approximately half of the sampled sites had representative snow station measurements relative to the surrounding area. In addition, we found that modelled snow depth and SWE also vary at different spatial scales; 21 of 39 scaling assessments had at least a 15% difference in SWE bias between 1 and 16 km². On average and across all spatial scales, SWE biases (snow station SWE point observation compared with the mean surrounding modelled SWE) were of greater magnitude during mid-melt compared with the accumulation season, highlighting the challenges associated with using these point data to evaluate spatially explicit models. In this regard, the comparisons between the snow station biases observed in this study and the model assimilations of the NOHRSC SWE product indicate that the assimilations may degrade model performance because the NOHRSC model is, in fact, estimating SWE more accurately on a 1-km² scale than the point snow station observations. An improved understanding of point-to-area scaling relationships will help improve snow depth and SWE estimates for hydrologic modelling efforts in snowmelt-dominated mountain catchments. More accurate SWE estimates will become increasingly critical as water resources in the western USA become progressively difficult to manage amid rapidly changing climate conditions.

ACKNOWLEDGEMENTS

This research was made possible by the NSF Hydrological Sciences (grant nos. EAR1032295 and EAR1032308) and the NOAA Office of Hydrologic Development (grant no. NA07NWS4620016). The authors thank their partners at NOAA, including Tom Carroll, Don Cline, Carrie Olheiser, Andy Rost and Pedro Restrepo. They also thank everyone who helped in the field.

REFERENCES

- Anderton SP, White SM, B Alvera. 2004. Evaluation of spatial variability in snow water equivalent for a high mountain catchment. *Hydrological Processes* **18**: 435–453.
- Andreadis KM, Lettenmaier DP. 2006. Assimilating remotely sensed snow observations into a macroscale hydrology model. *Advances in Water Resources* **29**: 872–886.
- Azar AE, Ghedira H, Romanov P, Mahani S, Tedesco M, Khanbilvardi R. 2008. Application of Satellite Microwave Images in Estimating Snow Water Equivalent. *Journal of the American Water Resources Association* **44**: 1347–1362.
- Bales RC, Molotch NP, Painter TH, Dettinger MD, Rice R, Dozier J. 2006. Mountain hydrology of the western United States. *Water Resources Research* **42**: W08432.
- Balk B, Elder K. 2000. Combining binary decision tree and geostatistical methods to estimate snow distribution in a mountain watershed. *Water Resources Research* **36**: 13–26.
- Barnett TP, Adam JC, Lettenmaier DP. 2005. Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature* **438**: 303–309.
- CA-DWR. 2010. California Department of Water Resources, Snow reports.
- Caldwell P, Chin HNS, Bader DC, Bala G. 2009. Evaluation of a WRF dynamical downscaling simulation over California. *Climatic Change* **95**: 499–521.
- Carroll SS, Cressie N. 1996. A comparison of geostatistical methodologies used to estimate snow water equivalent. *Water Resources Bulletin* **32**: 267–278.
- Carroll T, Cline D, Fall G, Nilsson A, Li L, Rost A. 2001. NOHRSC operations and the simulation of snow cover properties for the coterminous U.S. 69th Annual Meeting of the Western Snow Conference, Sun Valley, ID.
- Chambers J, Hastie T. 1993. *Statistical Models in S*. CRC Press: Boca Raton, FL.
- Cline DW. 1997. Snow surface energy exchanges and snowmelt at a continental, midlatitude Alpine site. *Water Resources Research* **33**: 689–701.
- Cline DW, Bales RC, Dozier J. 1998. Estimating the spatial distribution of snow in mountain basins using remote sensing and energy balance modeling. *Water Resources Research* **34**: 1275–1285.
- Daly SF, Davis R, Ochs E, Pangburn T. 2000. An approach to spatially distributed snow modelling of the Sacramento and San Joaquin basins, California. *Hydrological Processes* **14**: 3257–3271.
- DeBeer CM, Pomeroy JW. 2010. Simulation of the snowmelt runoff contributing area in a small alpine basin. *Hydrology and Earth System Sciences* **14**: 1205–1219.
- Derksen C. 2008. The contribution of AMSR-E 18.7 and 10.7 GHz measurements to improved boreal forest snow water equivalent retrievals. *Remote Sensing of Environment* **112**: 2701–2710.
- Durand M, Kim EJ, Margulis SA. 2008. Quantifying uncertainty in modeling snow microwave radiance for a mountain snowpack at the point-scale, including stratigraphic effects. *IEEE Transactions on Geoscience and Remote Sensing* **46**: 1753–1767.
- Elder K. 1995. *Snow distribution in alpine watersheds*. University of California: Santa Barbara.
- Elder K, Dozier J, Michaelsen J. 1991. SNOW ACCUMULATION AND DISTRIBUTION IN AN ALPINE WATERSHED. *Water Resources Research* **27**: 1541–1552.
- Elder K, Rosenthal W, Davis RE. 1998. Estimating the spatial distribution of snow water equivalence in a montane watershed. *Hydrological Processes* **12**: 1793–1808.
- Erxleben J, Elder K, Davis R. 2002. Comparison of spatial interpolation methods for estimating snow distribution in the Colorado Rocky Mountains. *Hydrological Processes* **16**: 3627–3649.
- Faria DA, Pomeroy JW, Essery RLH. 2000. Effect of covariance between ablation and snow water equivalent on depletion of snow-covered area in a forest. *Hydrological Processes* **14**: 2683–2695.
- Farr TG, Rosen PA, Caro E, Crippen R, Duren R, Hensley S, Kobrick M, Paller M, Rodriguez E, Roth L, Seal D, Shaffer S, Shimada J, Umland J, Werner M, Oskin M, Burbank D, Alsdorf D. 2007. The shuttle radar topography mission. *Reviews of Geophysics* **45**: RG2004.
- Fassnacht SR, Derry JE. 2010. Defining similar regions of snow in the Colorado River Basin using self-organizing maps. *Water Resources Research* **46**: W04507.
- Fassnacht SR, Dressler KA, Bales RC. 2003. Snow water equivalent interpolation for the Colorado River Basin from snow telemetry (SNOTEL) data. *Water Resources Research* **39**(8): 1208.
- Fassnacht SR, Heun CM, Lopez-Moreno JJ, Latron J. 2010. Variability of Snow Density Measurements in the Rio Esera Valley, Pyrenees Mountains, Spain. *Cuadernos de Investigación Geográfica (Journal of Geographical Research)* **36**: 59–72.
- Fu P, Rich PM. 1999. The Solar Analyst 1.0 User Manual. Helios Environmental Modeling Institute, LLC.
- Goodison BE, Ferguson WH, McKay GA. 1981. Measurement and data analysis. In *Handbook of Snow*, Gray DM, Male DH (eds). The Blackburn Press: Caldwell, NJ; 191–274.
- Kerker B, Glaser SD, Bales RC, Saksa PC. 2010. Design and development of a wireless sensor network to monitor snow depth in multiple catchments in the American River basin, California: hardware

- selection and sensor placement techniques. in American Geophysical Union Fall Meeting 2010, San Francisco, CO.
- Lee SW, Klein AG, Over TM. 2005. A comparison of MODIS and NOHRSC snow-cover products for simulating streamflow using the Snowmelt Runoff Model. *Hydrological Processes* **19**: 2951–2972.
- Lopez-Moreno JIL, Latron J, Lehmann A. 2010. Effects of sample and grid size on the accuracy and stability of regression-based snow interpolation methods. *Hydrological Processes* **24**: 1914–1928.
- Lopez-Moreno JI, Fassnacht SR, Begueria S, Latron J. 2011. Variability of snow depth at the plot scale: implications for mean depth estimation and sampling strategies. *The Cryosphere* **5**: 617–629.
- Marks D, Dozier J. 1992. Climate and energy exchange at the snow surface in the alpine region of the Sierra Nevada 2. Snow cover energy balance. *Water Resources Research* **28**: 3043–3054.
- McKay GA, Gray DM. 1981. The Distribution of Snowcover. In *Handbook of Snow*, Gray DM, Male DH (eds). The Blackburn Press: Caldwell, NJ; 153–187.
- Melloh RA, Richmond P, Shoop SA, Affleck RT, Coutermarsh BA. 2008. Continuous mapping of distributed snow depth for mobility models using shaped solutions. *Cold Regions Science and Technology* **52**: 155–165.
- Molotch NP, Bales RC. 2005. Scaling snow observations from the point to the grid element: Implications for observation network design. *Water Resources Research* **41**: W11421.
- Molotch NP, Bales RC. 2006. SNOTEL representativeness in the Rio Grande headwaters on the basis of physiographics and remotely sensed snow cover persistence. *Hydrological Processes* **20**: 723–739.
- Molotch NP, Colee MT, Bales RC, Dozier J. 2005. Estimating the spatial distribution of snow water equivalent in an alpine basin using binary regression tree models: the impact of digital elevation data and independent variable selection. *Hydrological Processes* **19**: 1459–1479.
- Molotch NP, Brooks PD, Burns SP, Litvak M, Monson RK, McConnell JR, Musselman K. 2009. Ecohydrological controls on snowmelt partitioning in mixed-conifer sub-alpine forests. *Ecohydrology* **2**: 129–142.
- Musselman KN, Molotch NP, Brooks PD. 2008. Effects of vegetation on snow accumulation and ablation in a mid-latitude sub-alpine forest. *Hydrological Processes* **22**: 2767–2776.
- Neumann NN, Derksen C, Smith C, Goodison B. 2006. Characterizing local scale snow cover using point measurements during the winter season. *Atmosphere-Ocean* **44**: 257–269.
- NRCS. 2010. NRCS webpage.
- Pan M, Sheffield J, Wood EF, Mitchell KE, Houser PR, Schaake JC, Robock A, Lohmann D, Cosgrove B, Duan QY, Luo L, Higgins RW, Pinker RT, Tarpley JD. 2003. Snow process modeling in the North American Land Data Assimilation System (NLDAS): 2. Evaluation of model simulated snow water equivalent. *Journal of Geophysical Research-Atmospheres* **108**(D22): 8850.
- Pavelsky TM, Kapnick S, Hall A. 2011. Accumulation and melt dynamics of snowpack from a multiresolution regional climate model in the central Sierra Nevada, California. *Journal of Geophysical Research* **116**.
- Rice R, Bales RC. 2010. Embedded-sensor network design for snow cover measurements around snow pillow and snow course sites in the Sierra Nevada of California. *Water Resources Research* **46**: W03537.
- Rutter N, Cline D, Li L. 2008. Evaluation of the NOHRSC snow model (NSM) in a one-dimensional mode. *Journal of Hydrometeorology* **9**: 695–711.
- Tedesco M, Narvekar PS. 2010. Assessment of the NASA AMSR-E SWE Product. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **3**: 141–159.
- Veatch W, Brooks PD, Gustafson JR, Molotch NP. 2009. Quantifying the effects of forest canopy cover on net snow accumulation at a continental, mid-latitude site. *Ecohydrology* **2**: 115–128.
- Winstral A, Elder K, Davis RE. 2002. Spatial snow modeling of wind-redistributed snow using terrain-based parameters. *Journal of Hydrometeorology* **3**: 524–538.
- Work RA, Stockwell HJ, Freeman TG, Beaumont RT. 1965. Accuracy of field snow surveys in western United States, including Alaska. Technical Report 163. US Army Cold Regions Research and Engineering Laboratory, Army Material Command, Hanover, NH.

APPENDIX A

Summary of snow depth and density measurements made during the peak accumulation and mid-melt 2008 and 2009 field campaigns. Dates of surveys are beneath site names. Depth values are reported in centimetres, and density values are in kilograms per cubic metre.

	South Brush Creek, Wyoming, 4/5	Dry Lakes, Colorado, 4/4, 5/2	Joe Wright, Colorado, 4/3, 5/1	Niwot Ridge, Colorado, 4/7, 5/5	Lizard Head, Colorado, 3/16
2008					
Minimum	Mid-melt 0	Peak 59	Peak 0	Peak 17	Peak 106
Maximum	125	212	305	201	212
Mean	95	154	171	117	157
SD	20	23	37	23	20
<i>n</i>	66	173	134	183	241
Average density	288	370	287	299	328
Snow station depth	97	173	140	127	153
	Virginia Lakes Ridge, California, 5/9	Mammoth Pass, California, 3/25, 5/6	Rock Creek, California, 4/4, 5/2	Gin Flat, California, 4/2, 5/2	Ostrander Lake, California, 3/19
Minimum	Mid-Melt 0	Peak 82	Peak 0	Peak 0	Peak 0
Maximum	179	316	210	237	258
Mean	40	230	88	125	177
SD	38	34	41	39	38
<i>n</i>	232	160	207	231	231
Average density	420	452	392	478	433
Snow station depth	69	217	73	141	193
	Moscow Mountain, Idaho, 4/11, 5/5	Sherwin, Idaho, 3/23, 4/30	Hogg Pass, Oregon, 3/15, 5/15	Santiam Junction, Oregon, 3/16, 5/13	
Minimum	Peak 24	Peak 0	Peak 117	Peak 10	
Maximum	290	194	441	375	
Mean	175	94	318	211	
SD	45	31	41	43	
<i>n</i>	131	134	134	111	
Average density	394	372	402	398	
Snow station depth	228	95	298.8	193	
2009	South Brush Creek, Wyoming, 3/1, 3/29	Dry Lakes, Colorado, 2/28, 3/28	Joe Wright, Colorado, 1/31, 4/2	Niwot Ridge, Colorado, 3/6, 4/3	Togwootee, Wyoming, 3/17
Minimum	Peak 56.3	Peak 79	Peak 0.5	Peak 0	Peak 65
Maximum	150	283	182.7	107	266
Mean	84.1	150	126.7	56	175
SD	11.9	24	36.5	17	31
<i>n</i>	223	189	170	175	158
Average density	289	325	295	310	320
Snow station depth	75	140	143.8	51	162

(Continues)

(Continued)

	Virginia Lakes Ridge, California, 4/2, 5/4	Mammoth Pass, California, 3/27, 4/29	Rock Creek, California, 3/29, 5/1	Gin Flat, California, 3/11, 5/7	Ostrander, California, 3/18, 4/22
	Peak	Peak	Peak	Peak	Peak
	Mid-melt	Mid-melt	Mid-melt	Mid-melt	Mid-melt
Minimum	0	20	0	51	57
Maximum	276	348	123	281	226
Mean	77	244	54	141	159
SD	50	45	30	43	29
<i>n</i>	231	226	201	231	231
Average density	310	363	342	439	404
Snow station depth	108	232	54	167	181
	Moscow, Idaho, 4/7, 5/8	Sherwin, Idaho, 3/10, 4/24	Hogg Pass, Oregon, 3/19, 5/1	Santiam Junction, Oregon, 3/16, 4/28	Joe Wright b, Colorado, 2/27, 5/2
	Peak	Peak	Peak	Peak	Peak
	Mid-melt	Mid-melt	Mid-melt	Mid-melt	Mid-melt
Minimum	0	8	110	30	0
Maximum	325	192	425	250	209
Mean	172.8	79.6	253	146	140
SD	53.5	33.4	42.1	38.1	27
<i>n</i>	236	222	136	115	175
Average density	326	337	390	37	287
Snow station depth	221.5	87.5	239	150	141