RESEARCH ARTICLE

Distribution of snow depth variability

S.R. FASSNACHT (✉)[1,2,3,4], K.S.J. BROWN[1], E.J. BLUMBERG[5], J.I. LÓPEZ MORENO[6], T.P. COVINO[1,3], M. KAPPAS[4], Y. HUANG[7], V. LEONE[8], A.H. KASHIPAZHA[1]

1 ESS-Watershed Science, Colorado State University, Fort Collins, CO 80523-1476, USA
2 Cooperative Institute for Research in the Atmosphere, Fort Collins, CO 80523-1375, USA
3 Natural Resources Ecology Laboratory, Fort Collins, CO 80523-1499, USA
4 Geographisches Institut, Georg-August-Universität Göttingen, 37077 Göttingen, Deutschland
5 Geosciences, Colorado State University, Fort Collins, CO 80523-1482, USA
6 Instituto Pirenaico de Ecología, CSIC, Campus de Aula Dei, 50080 Zaragoza, España
7 State Key Laboratory of Hydroscience and Engineering, Tsinghua University, Beijing 100084, China
8 Seconda Università degli Studi di Napoli, Viale Abramo Lincoln, 5, 81100 Caserta CE, Italia

Abstract Snow depth is the easiest snowpack property to measure in the field and is used to estimate the distribution of snow for quantifying snow storage. Often the mean of three snow depth measurements is used to represent snow depth at a location. This location is used as a proxy for an area, typically a digital elevation model (DEM) or remotely sensed pixel. Here, 11, 17, or 21 snow depth measurements were used to represent the mean snow depth of a 30-m DEM pixel. Using the center snow depth measurement for each sampling set was not adequate to represent the pixel mean, and while the use of three snow depth measurements improved the estimate of mean, there is still large error for some pixels. These measurements were then used to determine the variability of snow depth across a pixel. Estimating variability from few points rather than all in a measurement was not sufficient. The sampling size was increased from one to the total per pixel (11, 17, or 21) to determine how many point samples were necessary to approximate the mean snow depth per pixel within five percent. Binary regression trees were constructed to determine which terrain and canopy variables dictated the spatial distribution of the snow depth, the standard deviation of snow depth, and the sample size to within 5% of the mean per pixel. One location was measured in two years just prior to peak accumulation, and it is shown that there was little to no inter-annual consistency in the mean or standard deviation.

Keywords uncertainty, sampling, binary regression trees, snow telemetry

1 Introduction

Snowmelt derived from mountain headwaters provides water resources to over half of the world's population (Debarbieux and Price, 2012). Accordingly, the ability to measure or estimate the snowpack in mountain landscapes is critical to water resource management (Bales et al., 2006). While the importance of information on mountain snowpack is clear, obtaining these measures remains challenging. This is in part because snow depth and snow water equivalent (SWE) are both spatially and temporally variable, and mountain regions are generally difficult to access. As a step toward addressing this challenge we evaluated methodologies to increase the efficiency of snow surveys and to enhance remotely derived estimates.

The correlation between snow depth and terrain/canopy variables has been well established (Meiman, 1968; Elder et al., 1991; Balk and Elder, 2000; Erxleben et al., 2002; Erickson et al., 2005; Molotch and Bales, 2005; Hultstrand et al., 2006; López-Moreno and Nogués Bravo, 2006; López-Moreno et al., 2011; Sexstone and Fassnacht, 2014). Existing methods to estimate snow depths integrate digital elevation models (DEMs) and land cover (LC) data within a geographic information system (GIS) analysis framework. With these approaches, only a few snow depth measurements, often from one (Elder et al., 2009) to three (Molotch and Bales, 2005) points, are used to represent the area of one DEM or LC pixel (e.g., 30 m).

Clearly, this raises the question: how much variability in snow depth exists within a pixel (López-Moreno et al., 2011)? Various assumptions (Sturm and Wagner, 2010) and analyses (Erickson et al., 2005; Deems et al., 2008; López-Moreno et al., 2015) have demonstrated that snow

accumulation and distribution patterns are consistent over time for mean snow depth within a given pixel. Thus, a second question follows: is the variability of snow depth per pixel also consistent over time? To address these two questions, we pose the following more specific questions for our research objectives:

1) How different are the estimates of snow depth per pixel using one or three points compared to using more measurements per pixel? Specifically, do more measurements per pixel provide more information?

2) How many measurements are needed per pixel to estimate the mean value of a pixel within an acceptable limit?

3) How do the pixel mean and pixel standard deviations vary inter-annually for the same location at the same time of the year?

4) How do meteorological driving forces, represented by terrain and canopy variables, change in relative effect on snow depth distribution between sites and inter-annually at the same site?

2   Background on snow depth sampling

From a water resource management perspective, information on water storage or SWE is preferred to estimates of snow depth (Bales et al., 2006). However, snow depth (d_s) information is the easiest snowpack property to measure. SWE is the product of d_s and the depth averaged snowpack density, and density has been shown to be less spatially variable than SWE or depth (Logan, 1973; Fassnacht et al., 2010; López-Moreno et al., 2013). Operationally, the Natural Resource Conservation Service (NRCS) has measured snowpack properties (d_s and SWE) for more than 80 years with the snow course sites, and over the past 40 years has used data from the automated snowpack telemetry (SNOTEL) network to forecast spring and summer runoff volumes (U.S. Department of Agriculture, 2011, 2016) across the Western United States where snowmelt contributes water supply to more than 60 million people. SNOTEL data have a daily temporal resolution that represents the dynamic evolution of the snowpack (Fassnacht et al., 2014), but only represent a small area (~10 m²). Thus, individual measurements or stations tend not to be representative of the area(s) surrounding them (Molotch and Bales, 2005; Kashipazha, 2012; Meromy et al., 2013), and such a single fixed point measurement is not a statistically useful tool to represent the average snow depth of an area, even with a relatively uniform snow pack (Neumann et al., 2006). This issue arises for manual field surveys used to evaluate the distribution of snow (e.g., Meiman, 1968; Mould, 1985; Elder et al., 1991; Balk and Elder, 2000; Erxleben et al., 2002; Erickson et al., 2005; Molotch and Bales, 2005; Hultstrand et al., 2006; López-Moreno and Nogués Bravo, 2006; Sexstone and Fassnacht, 2014). Most of these surveys use one (Elder et al., 2009) or three (Elder et al., 1991) snow depth measurements to represent one pixel (e.g., 30 m) of a DEM.

Rice and Bales (2010) suggested that a network of sensors could be used to provide a better estimate of snow depth, especially for areas that cannot be readily or easily surveyed manually. On average, 5 sensors out of a network up to 44 yielded a snow depth estimate within 25% of the basin mean, depending on the snow variability (Rice and Bales, 2010). More sensors are likely needed to achieve this level of accuracy in terrain with strong topographic variability (Neumann et al., 2006). López-Moreno et al. (2011) sampled 121 points at 15 relatively homogenous 100 m² plots in the Spanish Pyrenees and found that five to seven points produced an average snow depth within 5% of the mean value based on the 121 points. In this paper we used 11, 17, or 21 snow depth measurements per location or pixel size of 900 m² for over 150 pixels to evaluate snow depth variability within a given location and to determine the optimal number of points.

3   Data and methods

3.1   Study sites and sampling

Snow depth was measured manually about two SNOTEL stations using evolving sampling strategies. The first survey was about the Togwotee Pass SNOTEL station in Northwestern Wyoming (Fig. 1(a)) on 17 March 2009 (TP09), while the second and third surveys were about the Joe Wright SNOTEL station near Cameron Pass in Northern Colorado (Fig. 1(a)) on 1 May 2009 (JW09) and 2 May 2010 (JW10). The Togwotee Pass site has a mixed canopy density throughout (Fig. 1(b)-i), while the Joe Wright site is moderately to densely forested, except along the road through the middle of the study area (Fig. 1 (b)-ii). Both sites are located in the Spruce-Fir forest composed mainly of *Picea engelmanii* and *Abies lasiocarpa* trees.

Snow depth measurements were taken to cover a 1 km×1 km area about the SNOTEL station in 9 or 10 transects each separated by 100 m (Meromy et al., 2013; Fig. 1(b)). A 1 cm diameter aluminum probe (extendable in 1 m lengths) was used to measure snow depth to the nearest one centimeter. Measurements were taken at 50 m intervals along each transect. During snow sampling, a hand-held Garmin or Trimble global positioning system (GPS) unit was used to navigate to each sampling location. Each person was instructed to navigate to within 10 m for each coordinate, and record the GPS coordinates at the center point to the nearest one-meter.

The simplest and most basic sampling design was three snow depth measurements per pixel taken 5 m apart, denoted -5, 0, and $+5$ m (Meromy et al., 2013; Fig. 1(c)). At the first site (TP09), two sets of four points were added between each of the basic three measurements (-5, 0, and
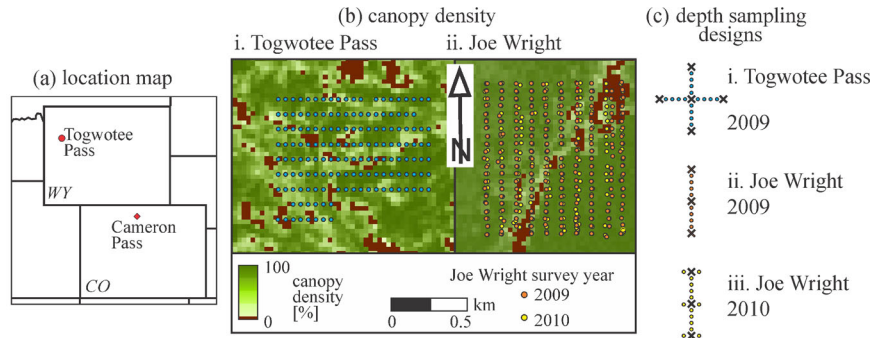
Fig. 1 (a) Location map for the i. Togwotee Pass and ii. Joe Wright SNOTEL stations in Wyoming and Colorado, respectively. (b) Canopy density in percent for the area surveyed: i. Togwotee Pass was surveyed in 2009, and ii. Joe Wright was surveyed in 2009 (red symbols) and iii. 2010 (black symbols). (c) The snow depth sampling design at each sampling plot is i) 21 points for Togwotee Pass (2009), ii. 11 points for Joe Wright 2009, and iii. 17 points for Joe Wright 2010.

+ 5 m) plus an additional five points to the left and right of center point yielding 21 points per pixel (Fig. 1(c)-i). The 10 additional points to the left and right added a substantial amount of sampling time, and were only used for TP09. Eleven points were used per pixel for the Joe Wright 2009 survey (Fig. 1(c)-ii). To increase the number of points per pixel, while only slightly increasing the sampling time, one point was added each to the left and right of the start (-5), center (0) and end ($+5$) resulting in 17 points per pixel at Joe Wright in 2010 (Fig. 1(c)-iii). These surveys produced 157, 203, and 206 pixels for TP09, JW09, and JW10, respectively. Any measurements that were by the road or in the anthropogenically produced snow bank beside the road were noted and not used in the analysis.

3.2   Analysis and statistics

We examined the impact that the number of sampling points per pixel had on computed mean and standard deviations (Tables 1 and 2; Fig. 2). Firstly, we compared the center snow depth measurement to the depths collected along the 11 points in a row (Fig. 1(c)-ii) for all three datasets, and then all points for TP09 (21 points, Fig. 1(c)-i) and JW10 (17 points, Fig. 1(c)-iii). Secondly, we used the basic three measurements in a row (-5, 0, $+5$ m; Meromy et al., 2013), and then the five measurements in a plus from TP09 (Fig. 1(c)-i). For TP09 and JW10 we also compared the 11 in a row versus all measurements. Lastly, we computed the difference from the pixel mean snow depth for all samples per pixel as a function of the number of points per pixel (Fig. 3). These comparisons were evaluated using the Nash Sutcliffe coefficient of efficiency (NSCE) (Nash and Sutcliffe, 1970), the coefficient of determination (R^2), and the slope of the best-fit regression.

To evaluate repeatability of sampling (Erickson et al., 2005; Sturm and Wagner, 2010), comparisons of the two years of snow survey data for the Joe Wright site were conducted by overlapping both datasets to ensure that each

corresponding data point from both years was within the same 30-m pixel (Fig. 4). Any 2009 measurement location that did not have measurements in 2010 but was within one pixel of a 2010 measurement was also compared; this increased the number of comparable points between the two survey years (Fig. 4).

Various statistical techniques (e.g., Erxleben et al., 2002; López-Moreno and Nogués Bravo, 2006) have been used to determine which terrain and canopy variables explain the spatial distribution of snow. A comparison of methods for the same extent used here (~ 1 km^2 shown in Fig. 1(b)) was applied across several Rocky Mountain sites in Colorado for the NASA Cold Land Process Experiment, and binary regression trees were determined to be the most successful at explaining the variability of snow depth (Erxleben et al., 2002). Accordingly, we created binary regression trees (Fig. 5) following Elder et al. (1991) using the Matlab data processing language. The trees were developed to identify the sequence of terrain and canopy variables that best distributed mean and standard deviations of snow depths (Table 3), as well as to determine the number of measurements needed to produce a value within 5% of the pixel mean (López-Moreno et al., 2011) (hereinafter called measurements to 5%, Table 3). The binary regression trees were built with a series of branching decisions ending in a group of measurements or "leaf". A minimum leaf size equal to 5% of each site's measurements was used so that each terminal node contained a reasonable number of measurements. Each binary regression tree was analyzed across all possible mergers of decision points and leaves to find the tree with the lowest standard error. Standard error here is a function of the relative standard deviation within each leaf. A tree with one leaf for each measurement will have an error of zero, and error generally increases as number of leaves decrease. There are exceptions where reducing leaves and decision points produces slightly more homogenous terminal nodes (Meromy et al., 2013).
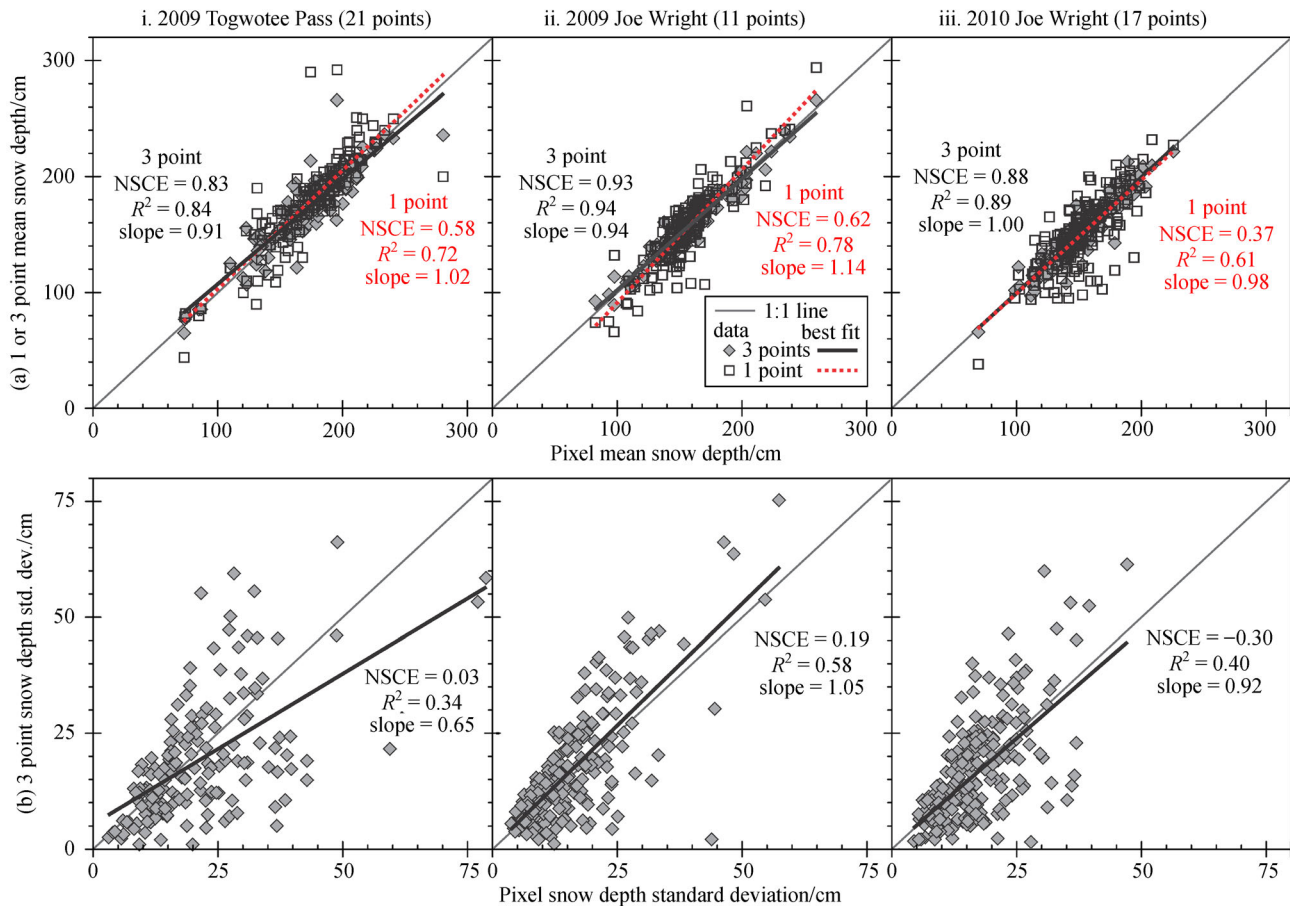
Fig. 2 Comparison of (a) mean and (b) standard deviation comparison between 3 point statistics (grey diamonds) and plot statistics for three snow depth surveys about SNOTEL stations: i. Togwotee Pass in March 2009, ii. Joe Wright in May 2009, and iii. Joe Wright in May 2010. The center or 1 point snow depth is also compared to the pixel mean (open box). The maximum number of points per station-date are in parentheses.

Table 1 Statistics mean summarizing the comparison of fewer sampling points (1, 3, or 5) to all sampling points (11, 17, or 21)

| Comparison	Togwotee Pass 2009			Joe Wright 2009			Joe Wright 2010		
	NCSE	R^2	slope	NCSE	R^2	slope	NCSE	R^2	slope
1 to 11	0.61	0.73	1.02	0.62	0.78	1.14	0.44	0.62	0.96
1 to 17		N/A			N/A		0.37	0.61	0.98
1 to 21	0.58	0.72	1.02		N/A			N/A	
3 to 11	0.90	0.90	0.93	0.93	0.94	0.94	0.89	0.89	0.96
3 to 17		N/A			N/A		0.88	0.89	1.0
3 to 21	0.83	0.84	0.91		N/A			N/A	
5 to 21	0.90	0.91	0.92		N/A			N/A	
11 to 17		N/A			N/A		0.98	0.94	0.98
11 to 21	0.94	0.94	0.98		N/A			N/A	

3.3 Geospatial data

Manual snow depth sampling was carried out at a resolution of 50 m to 100 m while a 30-m DEM was used for topographic data. The statistical analyses identified the terrain and canopy variables that describe the distribution and variability of snow depth about the two SNOTEL stations. These terrain variables included eleva-
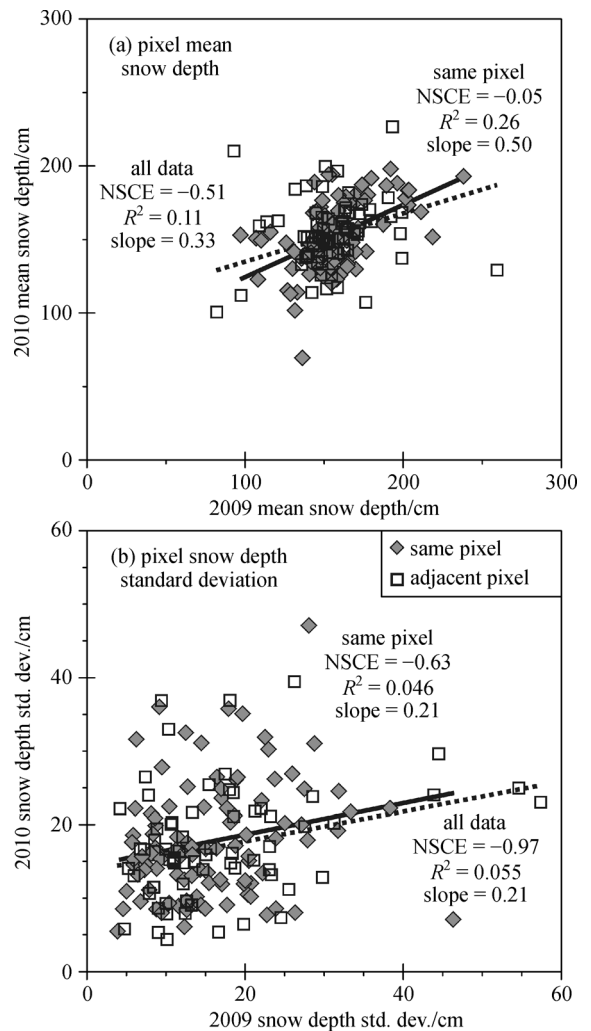
Table 2 Statistics standard deviation summarizing the comparison of fewer sampling points (1, 3, or 5) to all sampling points (11, 17, or 21)

| Comparison	Togwotee Pass 2009			Joe Wright 2009			Joe Wright 2010		
	NCSE	R^2	slope	NCSE	R^2	slope	NCSE	R^2	slope
3 to 11	0.22	0.55	0.96	0.19	0.58	1.05	−0.20	0.39	0.87
3 to 17		N/A			N/A		−0.30	0.40	0.92
3 to 21	0.03	0.34	0.65		N/A			N/A	
5 to 21	0.46	0.64	0.95		N/A			N/A	
11 to 17		N/A			N/A		0.76	0.79	0.94
11 to 21	0.35	0.46	0.58		N/A			N/A	
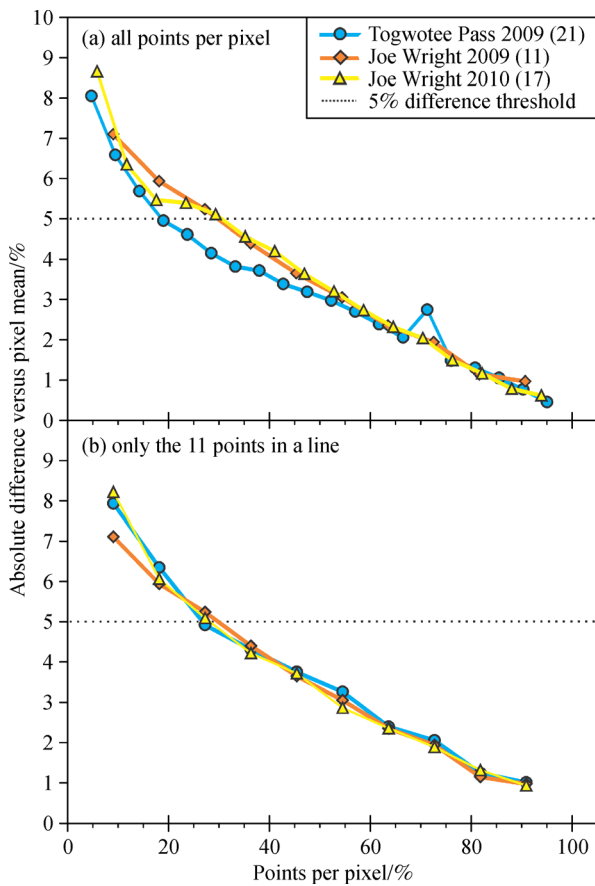
Fig. 3 Mean absolute difference from the pixel mean snow depth for all samples per pixel as a function of the number of points per pixel for the three sampling dates. The number in the (a) parentheses is the number of samples per pixel, while (b) is for only the 11 points in a row (as per Fig. 1(c)-ii). The 5% difference threshold is shown as the dotted line.

Fig. 4 Joe Wright SNOTEL inter-annual (May 2009 and 2010) comparison of (a) mean snow depth, and (b) standard deviation for data at the 99 pixels within the same pixel (grey diamonds) and 70 sets of pixels that were within one pixel. A pixel is derived from the 30-meter U.S. Geological Survey (2018) digital elevation model < nationalmap.usgs.gov > . The Nash-Sutcliffe coefficient of efficiency statistic is not included as it is less than zero in all cases.

tion (Dingman et al., 1988; Fassnacht et al., 2003), slope as the sine of slope (Sexstone and Fassnacht, 2014), northness (Fassnacht et al., 2012), eastness (Wallace and Gass, 2008), cumulative monthly clear sky solar radiation (Dozier and Frew, 1990), and maximum upwind slope
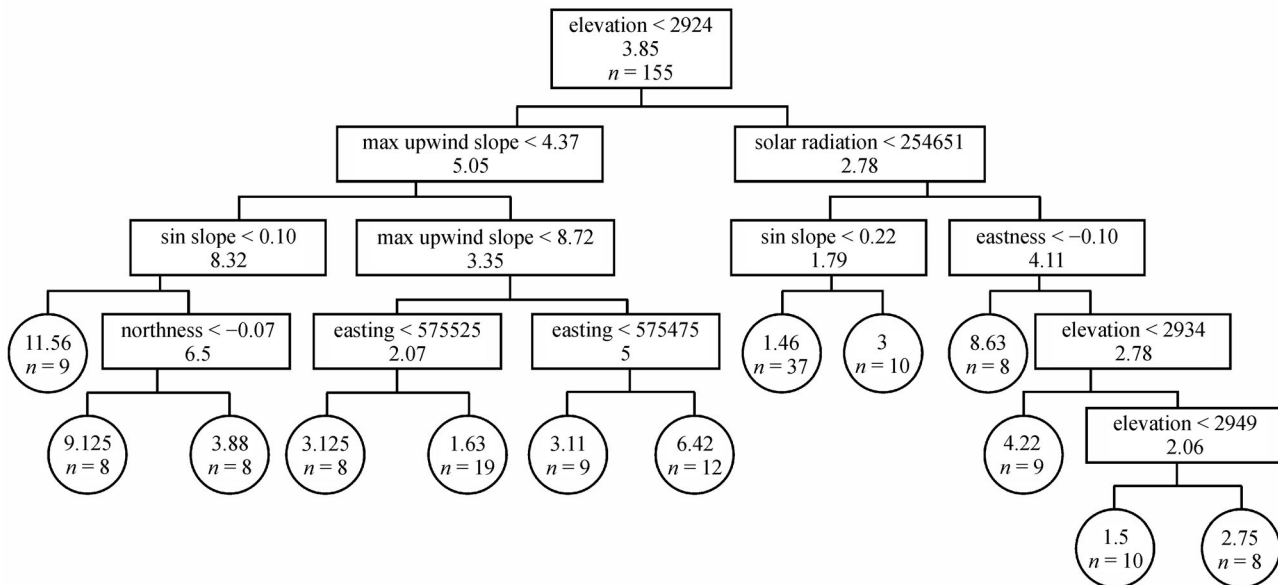
Fig. 5 Togwotee Pass 2009 binary regression trees for depth to 5% of the mean per pixel. Boxes represent decision nodes with the true being the node to the left below. The value below the variable represents the mean value of all pixels remaining. Circles represent terminal nodes with the mean of the number of points to within 5% of snow depth, for the number of points (n) in that category.

(Winstral et al., 2002). Northness is the product of sine of slope and cosine of aspect to represent an integration of wind and sun influences with a steep north facing slope approaching a value of 1 and a steep south facing slope approaching a value of -1. Eastness is the product of sine of slope and sine of aspect to represent wind processes. Maximum upwind slope describes the topographic shelter or exposure relative to a specific wind direction (Winstral and Marks, 2002), and can explain the variability in snow deposition due to wind transport and redistribution (Winstral et al., 2002).

Elevation data were obtained from the U.S. Geological Survey (2018) 30-m DEM, and used to derive the terrain variables. Canopy cover offers snow protection from the wind, and provides shade on sunny aspects, and was thus also used. Canopy cover data were also obtained from the U.S. Geological Survey (2018). Data points were overlain onto the DEM and the other variables in the ArcMap Geographical Information System (GIS) software. The value for the variables at each data point was then extracted.

4 Results

When center points were compared to overall pixel means, only Joe Wright 2010 was not considered good or very good (Fig. 2(a)), as per the classification of Moriasi et al. (2007). The mean of the three main measurement points (-5, center, $+5$; Fig. 1(c)) is better than the center point measurement alone for representing the overall pixel mean, as shown by the improved NSCE and R^2 values (Tables 1

and 2, Fig. 2). For Togwotee Pass 2009 and Joe Wright 2010, measuring 11 points was nearly the same as measuring all 21 (TP09) or 17 (JW10) points, in terms of estimating the pixel mean (Table 1; Fig. 2(a)). The five extreme points arranged in a plus at Togwotee Pass did produce a more accurate mean than three points in a row (Fig. 1(c)-i versus Fig. 1(c)-ii).

The standard deviation computed from just three measurements is not adequate to represent the variation across a pixel from all measured points (Table 2; Fig. 2(b)). The five extreme points at Togwotee Pass (Fig. 1(c)-i) is better than the 11 points in a row at representing the variability within each pixel (Table 2). At Joe Wright 2010, the standard deviation from 11 points in a row is similar to the pixel standard deviation computed with the extra left and right points at -5, 0, and $+5$, i.e., 17 points in total (Fig. 1(c)-iii).

The average number of points per pixel required to represent the mean of the pixel depends on the acceptable error (Fig. 3), and we used the 5% threshold suggested by López-Moreno et al. (2011). Only 20 (TP) to 30% (JW) of the actual points measured, i.e., 3 to 5 points, are required to obtain the mean within 5% (Fig. 3(a)). Since there are more measurements per pixel at TP than JW (Fig. 1(c)), the deviation from the mean is smaller even with a lower proportion of the total points (Fig. 3). When using only the 11 points in a row configuration (Fig. 1(c)-ii), the relation between difference from mean and the number of points are almost identical (Fig. 3(b)). However, there is considerable variability among individual pixels (Tables 1 and 2; Fig. 2).

There is inter-annual consistency in the average number

of points necessary to estimate the pixel mean (Fig. 4(a)), but much less consistency among the mean of individual pixels within the same year (Fig. 4(a)). This decreases further from the fit of R^2 = 0.26 to when nearby measurements (within one pixel) are added (R^2 = 0.11). The snow depth was almost the same for the early May sampling at Joe Wright; the measurement means were 156 and 151 cm, and the SNOTEL measured depths were 177 and 180 cm for 2009 and 2010, respectively. At the SNOTEL station, melt started after the sampling (6 May 2009 and 16 May 2010). There is almost no correlation of pixel standard deviation between years (Fig. 3(b)).

The relative influence of terrain and canopy variables was somewhat more consistent. Eastness appears as an explanatory variable in all nine binary regression trees (BRTs), with canopy density and elevation appearing in eight of the nine BRTs (Table 3). Canopy density appears early in the binary regression trees for mean snow depth and standard deviation (Table 3(b) and 3(c)). Elevation appears at the top of the binary regression trees for measurements to 5% at all three sites (Table 3(a)), and in the second (JW09 and JW10) or third (TP09) branch for mean, but later (or not at all JW09) for standard deviation. Sine of slope appears in the first or second branch and then again in the third branch for the standard deviation. There is more consistency between the order of appearance of terrain and canopy variables for the two years at JW than for the same year at TP and JW (Table 3).

5 Discussion

Previous surveys have used only three measurement points in the sub alpine (e.g., Molotch and Bales, 2005; Meromy et al. (2013) around SNOTEL stations) or up to five measurement points in the alpine (e.g., Elder et al. (1991) at Emerald Lake, Sierra Nevada CA or Hultstrand et al. (2006) at GLEES, Snowy Range WY) to evaluate snowpack depth and variability per sample location. The NASA Cold Land Process (CLPX) used only one snow depth measurement, but with a random stratified sampling design which yielded 550 snow depth measurements over a 1 km² area (Erxleben et al., 2002). In a study in the Pyrenees, five to seven samples were adequate to represent one location at a 5% threshold (López-Moreno et al., 2011). However, the sites used by López-Moreno et al. (2011) were chosen as they appeared to have homogenous snow distribution on the surface, and only 15 plots of 121 points each were surveyed twice. A reliable sampling strategy is five measurement points taken at 5-m intervals (Fig. 1(c)-i), but this is more difficult in forested areas and complex terrain.

The 2009 and 2010 Joe Wright SNOTEL snow surveys used a sampling strategy of 11 and 17 points, respectively. In these two surveys only 99 sampling locations were within the same pixel and 70 were in an adjacent pixel (Fig. 4) due to the complexity of the terrain, the dense canopy, and safety/logistics of sampling. It should be noted that in this dense forest, the GPS accuracy was occasionally reported to be as poor as 10 m, implying that a specific pixel measurement may actually be within a different pixel. The 10-m transect (Fig. 1(c)) is not necessarily all within one 30-m pixel, and could even be sampled across three pixels (Fig. 1(c)-i). Snow surveyors were told to target each sampling pixel within 10 m of a specified center point, which could lead to up to maximum potential error of 40 m for points between the 2009 and 2010 survey. Furthermore, since snow depth, terrain, and

Table 3 Summary of the order of appearance in the binary regression tree for the (a) points to within 5% of the pixel mean from top 1 to last node (e.g., Fig. 2), (b) mean, and (c) standard deviation, and. The same number in the order list identifies that variable occurring more than once in the particular level. The fit statistic is the relative standard errors using a minimum leaf size of 5% of total number of samples

Variable	(a) Measurements to five 5% of mean per pixel			(b) Mean snow depth per pixel			(c) Standard deviation per pixel		
	Togwotee Pass 2009	Joe Wright 2009	Joe Wright 2010	Togwotee Pass 2009	Joe Wright 2009	Joe Wright 2010	Togwotee Pass 2009	Joe Wright 2009	Joe Wright 2010
Canopy density	N/A	2,4	4	3	1	3,5	1	2,5	1
Latitude	N/A	2,3	4,5	N/A	6	N/A	N/A	4,5,5	4,5,5
Longitude	4,4	N/A	6,7	4	N/A	1,3,4,5	4,4	N/A	5
Elevation	1,4,5	1,5,6	1,2,8	3,5,9	2,3,6	2,4,6	3,4	N/A	4
Sine of slope	3,3	N/A	5	5,6,7,9,10	3	N/A	2,3,5	1,3	2,3.4
Max upwind slope	2,3	4	7,8	8	N/A	N/A	3,6	6	3,5
Eastness	3	3,7,8	3	1	4,5,6	5,5	2,5	5,7	3
Northness	4	3,4	N/A	4	N/A	N/A	N/A	3,8	2
Solar radiation	2	N/A	N/A	2	N/A	4	4	4	N/A
Relative standard error	0.0321	0.0121	0.0229	0.0354	0.0283	0.0184	0.0708	0.0409	0.0182

canopy all vary at scales finer than the 30-m resolution data used here (Fassnacht and Deems, 2006), measuring more than three points per pixel (e.g., 11 to 21; Fig. 1(c)) could reduce sampling error.

Much of the snow hydrology literature suggests that there is a temporal consistency in the spatial patterns of the distribution of snow (e.g., Erickson et al., 2005; Sturm and Wagner, 2010; Webb, 2017). In open areas or areas with low density canopy, the snow distribution should be consistent through each snow year, as the areas are affected by the same physical features each season, while in more closed canopy areas, the snow distribution may vary through seasons due to changes in vegetative structure and density, especially over time (Rice and Bales, 2010). The patterns can be less consistent intra-annually than inter-annually over small domains (Kashipazha, 2012); this justifies further investigations into the patterns at the seasonal scale (Webb, 2017). For the Joe Wright data, the difference between 2009 to 2010 (Figs. 4(a) and 4(b)) could not be explained by land cover type, as 93% of the pixels were classified as evergreen, nor by canopy density, as most of the pixels had a dense canopy (91% had a canopy density greater than 50%, 70% of the pixels has a canopy density greater than 75%).

Binary regression trees (e.g., Fig. 5) illustrate the systematic relation of snow properties per pixel to terrain and canopy variables (Table 3; Elder et al., 1991; Erxleben et al., 2002; López-Moreno and Nogués Bravo, 2006). Elevation influences the distribution of snow (Fassnacht et al., 2003), but more so at Joe Wright (Table 3) as the elevation range was 50% greater (3055 to 3295 m at Joe Wright versus 2840 to 3000 m at Togwotee Pass). Slope also affected the distribution of snow, appearing often in the mean regression tree at Joe Wright. Sine of slope and maximum upwind slope were only related when the local slope faces into the direction of the wind from the west at TP and south at JW (Meromy et al., 2013).

Canopy density was an important factor due to its effects on wind and solar radiation (Winstral et al., 2002; Veatch et al., 2009), and it appeared in the binary regression trees for all snow surveys for both mean snow depth (Table 3(b); Fassnacht et al., 2017) and standard deviation (Table 3(c)) per pixel, but fewer for measurements to 5% (Table 3(a)). Both study sites are forested with an average of 63% and 74% at Togwotee Pass and Joe Wright, respectively (Fig. 1 (b)). Similarly, modeled clear sky cumulative solar radiation (plus eastness and northness) appeared in the regression trees and had some correlation to standard deviation at Joe Wright. Since all the measurements were taken prior to the start of melt, canopy would further be relevant for the distribution of snow during ablation (Molotch et al., 2009).

How much snow varies within a pixel (Table 3(c); Fig. 2 (b)) is not directly a function of the same terrain/canopy variables that drive the distribution of snow across the landscape (Table 3(b)), nor how many measurements are required to reach 5% (Table 3(a); Fig. 5). This difference is related to in part to scale (Blöschl, 1999); snow depth and vegetation are correlated at shorter scales (Deems et al., 2006) when trees are present (Trujillo et al., 2007). At longer distances, snow depth is more related to elevation (Fassnacht et al., 2003, 2009; Deems et al., 2006).

Since each of the Joe Wright surveys composes the same area and uses 99 of the same sample points, it is expected that the key variables influencing the average snow depths for each survey should be the same or similar. Variables such as canopy cover, slope, and aspect are expected to be constant between years, and other factors, such as the driving local meteorology (given by solar radiation and maximum upwind slope), are consistent from year to year (Meromy et al., 2013). Here, we showed limited inter-annual consistency in mean snow depth (Fig. 4(a)). There is no previous research examining the distribution of variability; no consistency was found in the variability of snow depth for a specific pixel between years (Fig. 4(b)). However, the correlation to the terrain and canopy variables is consistent (Table 3). More data should be collected to draw more solid conclusions between variability in snow depth at the same location over different years (e.g., Erickson et al., 2005).

With 11 to 21 samples per pixel over more than 150 pixels (Fig. 1(b)), our measurements have higher accuracy than the single SNOTEL measurement (Kashipazha, 2012; Meromy et al., 2013). The 11 measurement points in a row (Joe Wright 2009; Fassnacht et al., 2017) was the easiest to perform in the field, while 21 measurement points in a plus (Togwotee Pass) was the most difficult, especially in dense canopy and steep areas. The 17 measurement point per pixel survey from Joe Wright in 2010 seemed to be the most reasonable compromise to ensure representative samples (Kashipazha, 2012), and efficiency in time and energy spent to sample. On average, sampling three of 11 points in a row yields a mean snow depth within 5% (López-Moreno et al., 2011) of the pixel mean (Fig. 3(b)). The additional measurements for TP09 (21 points) and JW10 (17 points) improve estimation of snow depth in a pixel because these additional 6 to 10 points are off the row of 11 samples. However, for TP09, the 5 meters measurements taken off the main sampling line were difficult due to the dense canopy (Fig. 1(b)) and slope of the terrain at some of the pixels. The 17 point sampling strategy (Fig. 1(c)-iii) was more efficient. To determine the variability across a pixel using the standard deviation, the five extreme points in the 21 point sampling strategy (Fig. 1(c)-i) were best strategy (Tables 1 and 2). The different needs depend on the final purpose of the research and must be considered when designing a sampling strategy; sampling to estimate the mean would be used for an overall estimation of the snow resource while sampling to estimate the standard deviation would be used to assess sub-grid variability.

6 Conclusions

Snow depth sampling strategies, the distribution of snow, and its sub-pixel variability were investigated at two sites. Using one snow depth measurement is not adequate to estimate snow depth within a sampling pixel, when compared to 11, 17, or 21 measurements. While increasing the sampling to three points is an improvement when compared to using 11 measurement, but more measurements provide a better estimate of the mean snow depth across a sampling pixel. Using an acceptable error of sampling within 5% of the overall mean, measuring between 3 and 6 points was optimal. The results were similar for each snow survey. Adding sampling off of the main transect improved the estimation of pixel mean snow depth. However, sampling five meters to the left and right of the main transect increased the sampling time, so sampling one meter on each side at three locations in each sampling transect was a good compromise between sampling time and quantity of data collected.

On a pixel by pixel basis, the mean snow depth was not consistent for two similar sampling years; there was no correlation between years for the standard deviation of snow depth within the same pixels. The same key variables appeared more often in the binary regression trees across years and sites, but the hierarchy was inconsistent. The most relevant variables were elevation, canopy density, and eastness (the product of sine of slope and sine of aspect).

References

Bales R C, Molotch N P, Painter T H, Dettinger M, Rice R, Dozier J (2006). Mountain hydrology of the western United States. Water Resour Res, 42(8): W08432

Balk B, Elder K (2000). Combining binary decision tree and geostatistical methods to estimate snow distribution in a mountain watershed. Water Resour Res, 36(1): 13–26

Blöschl G (1999). Scaling issues in snow hydrology. Hydrol Processes, 13(14–15): 2149–2175

Debarbieux B, Price M F (2012). Mountain regions: a global common good? Mt Res Dev, 32(S1): S7–S11

Deems J S, Fassnacht S R, Elder K J (2006). Fractal distribution of snow depth from LiDAR data. J Hydrometeorol, 7(2): 285–297

Deems J S, Fassnacht S R, Elder K J (2008). Interannual consistency in fractal snow depth patterns at two Colorado mountain sites. J Hydrometeorol, 9(5): 977–988

Dingman S, Seely-Reynolds D, Reynold D (1988). Application of Kriging to estimating mean annual precipitation in a region of orographic influence. J Am Water Resour Assoc, 24(2): 329–339

Dozier J, Frew J (1990). Rapid calculation of terrain parameters for radiation modeling from digital elevation data. IEEE Trans Geosci Remote Sens, 28(5): 963–969

Elder K, Dozier J, Michaelsen J (1991). Snow accumulation and distribution in an alpine watershed. Water Resour Res, 27(7): 1541–1552

Elder K J, Cline D W, Liston G E, Armstrong R (2009). NASA cold land processes experiment (CLPX 2002–2003): field measurements of snowpack properties and soil moisture. J Hydrometeorol, 10(1): 320–329

Erickson T, Williams M, Winstral A (2005). Persistence of topographic controls on the spatial distribution of snow in rugged mountain terrain, Colorado, United States. Water Resour Res, 41(4): W04014

Erxleben J, Elder K, Davis R (2002). Comparison of spatial interpolation methods for estimating snow distribution in the Colorado Rocky Mountains. Hydrol Processes, 16(18): 3627–3649

Fassnacht S R, Deems J S (2006). Measurement sampling and scaling for deep montane snow depth data. Hydrol Processes, 20(4): 829–838

Fassnacht S R, Deitemeyer D C, Venable N B H (2014). Capitalizing on the daily time step of snow telemetry data to model the snowmelt components of the hydrograph for small watersheds. Hydrological Processes, 28(16): 4654–4668

Fassnacht S R, Dressler K A, Bales R C (2003). Snow water equivalent interpolation for the Colorado River Basin from snow telemetry (SNOTEL) data. Water Resour Res, 39(8): 1208

Fassnacht S R, Dressler K A, Hultstrand D M, Bales R C, Patterson G G (2012). Temporal inconsistencies in coarse-scale snow water equivalent patterns: Colorado River Basin snow telemetry-topography regressions. Pirineos, 167(0): 165–185

Fassnacht S R, Heun C, López Moreno J I, Latron J (2010). Variability of snow density measurements in the Esera valley, Pyrenees mountains, Spain. Cuadernos de Investigación Geográfica, 36(1): 59–72

Fassnacht S R, Williams M W, Corrao M V (2009). Changes in the surface roughness of snow from millimetre to metre scales. Ecol Complex, 6(3): 221–229

Fassnacht S R, Wyss D, Heering S M (2017). Räumliches Denken – ein Beispiel im Schnee / A Spatial Thinking Research-Didactic Example in Snow. Geoöko, 37(3-4)

Hultstrand D, Fassnacht S R, Stednick J (2006). Geostatistical methods for estimating snowmelt contribution to the alpine water balance. Proceedings of the Annual Western Snow Conference, (Las Cruces NM), 74: 149–154

Kashipazha A H (2012). Practical snow depth sampling around six snow telemetry (SNOTEL) stations in Colorado and Wyoming, United States. Dissertation for M.S. degree. Watershed Science Program, Colorado State University, Fort Collins, Colorado USA

Logan L (1973). Basin-wide water equivalent estimation from snowpack depth measurements. In: Role of Snow and Ice in Hydrology. IAHS AIHS Publ. 107: 864–884

López-Moreno J I, Fassnacht S R, Begueria S, Latron J (2011).

Variability of snow depth at the plot scale: implications for mean depth estimation and sampling strategies. Cryosphere, 5(3): 617–629

López-Moreno J I, Fassnacht S R, Heath J T, Musselman K, Revuelto J, Latron J, Morán-Tejeda E, Jonas T (2013). Small scale spatial variability of snow density and depth over complex alpine terrain: implications for estimating snow water equivalent. Adv Water Resour, 55: 40–52

López-Moreno J I, Nogués Bravo D (2006). Interpolating local snow depth data: an evaluation of methods. Hydrol Processes, 20(10): 2217–2232

López-Moreno J I, Revuelto J, Fassnacht S R, Azorín-Molina C, Vicente-Serrano S M, Morán-Tejeda E, Sexstone G A (2015). Snowpack variability across various spatio-temporal resolutions. Hydrol Processes, 29(6): 1213–1224

Meiman J R (1968). Snow accumulation related to elevation, aspect, and forest canopy. In Proceedings of Snow Hydrology Workshop Seminar, Canadian National Committee of the International Hydrological Decade, Fredericton NB, pp. 35–47

Meromy L, Molotch N, Link T, Fassnacht S R, Rice R (2013). Subgrid variability of snow water equivalent at operational snow stations in the western USA. Hydrol Processes, 27(17): 2383–2400

Molotch N, Brooks P, Burns S, Litvak M, Monson R, McConnell J, Musselman K (2009). Ecohydrological controls on snowmelt partitioning in mixed conifer sub alpine forests. Ecohydrology, 2 (2): 129–142

Molotch N P, Bales R C (2005). Scaling snow observations from the point to the grid-element: implications for observation network design. Water Resour Res, 41: W11421

Moriasi D N, Arnold J G, Van Liew M W, Bingner R L, Harmel R D, Veith T L (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans ASABE, 50(3): 885–900

Mould B (1985). New Day Rising. Chapter 1 in New Day Rising, Solid State Tuners, Taylor, Texas

Nash J E, Sutcliffe J V (1970). River flow forecasting through conceptual models part I–A discussion of principles. J Hydrol (Amst), 10(3): 282–290

Neumann N, Derksen C, Smith C, Goodison B (2006). Characterizing local scale snow cover using point measurements during the winter season. Atmos-ocean, 44(3): 257–269

Rice R, Bales R (2010). Embedded sensor network design for snow cover measurements around snow pillow and snow course sites in the Sierra Nevada of California. Water Resour Res, 46(3): W03537

Sexstone G A, Fassnacht S R (2014). What drives basin scale spatial variability of snowpack properties in the Front Range of northern Colorado? Cryosphere, 8(2): 329–344

Sturm M, Wagner A (2010). Using repeated patterns in snow distribution modeling: an Arctic example. Water Resour Res, 46(12): W12549

Trujillo E, Ramírez J A, Elder K J (2007). Topographic, meteorologic, and canopy controls on the scaling characteristics of the spatial distribution of snow depth fields. Water Resour Res, 43(7): W07409

U.S. Department of Agriculture (2011). Snow Survey and Water Supply Forecasting. National Engineering Handbook Part 622, Water and Climate Center, Natural Resources Conservation Service

U.S. Department of Agriculture (2016). Snow Survey and Water Supply Forecasting Program. National Water and Climate Center, Natural Resources Conservation Service

U.S. Geological Survey (2018). The National Map. U.S. Department of Interior

Veatch W, Brooks P, Gustafson J, Molotch N (2009). Quantifying the effects of forest canopy cover on net snow accumulation at a continental, mid latitude site. Ecohydrology, 2(2): 115–128

Wallace C, Gass L (2008). Elevation derivatives for Mojave Desert tortoise habitat models. US Geological Survey Open File Report 2008-1283, 7pp

Webb R (2017). Using ground penetrating radar to assess the variability of snow water equivalent and melt in a mixed canopy forest, Northern Colorado. Front Earth Sci, 11(3): 482–495

Winstral A, Elder K, Davis R (2002). Spatial snow modeling of wind redistributed snow using terrain based parameters. J Hydrometeorol, 3(5): 524–538

Winstral A, Marks D (2002). Simulating wind fields and snow redistribution using terrain based parameters to model snow accumulation and melt over a semi arid mountain catchment. Hydrol Processes, 16(18): 3585–3603