Title

Optimizing embedded sensor network design for catchment-scale snow-depth estimation using LiDAR and machine learning

Author

Oroza, Carlos A

Publication Date

2016-10-22

Peer reviewed

AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2016WR018896

Key Points:
- A machine-learning algorithm for optimizing snow sensor placements is compared to expert placements in an existing sensor network
- The spatial and temporal transferability of the algorithm is then assessed in 14 total LiDAR surveys in two uninstrumented catchments
- The accuracy of the snow depth estimated from the sensor measurements is higher than expert and randomized placements in each of the surveys

Correspondence to:
C. A. Oroza,
coroza@berkeley.edu

Citation:
Oroza, C. A., Z. Zheng, S. D. Glaser, and D. Tuia (2016), Optimizing embedded sensor network design for catchment-scale snow-depth estimation using LiDAR and machine learning, *Water Resour. Res.*, *52*, 8174–8189, doi:10.1002/2016WR018896.

Received 11 MAR 2016
Accepted 23 SEP 2016
Accepted article online 27 SEP 2016
Published online 22 OCT 2016

Optimizing embedded sensor network design for catchment-scale snow-depth estimation using LiDAR and machine learning

Carlos A. Oroza[1], Zeshi Zheng[1], Steven D. Glaser[1], Devis Tuia[2], and Roger C. Bales[1,3]

[1]Department of Civil and Environmental Engineering, University of California, Berkeley, California, USA, [2]Department of Geography, University of Zurich, Zurich, Switzerland, [3]Sierra Nevada Research Institute, University of California, Merced, California, USA

Abstract We evaluate the accuracy of a machine-learning algorithm that uses LiDAR data to optimize ground-based sensor placements for catchment-scale snow measurements. Sampling locations that best represent catchment physiographic variables are identified with the Expectation Maximization algorithm for a Gaussian mixture model. A Gaussian process is then used to model the snow depth in a 1 km^2 area surrounding the network, and additional sensors are placed to minimize the model uncertainty. The aim of the study is to determine the distribution of sensors that minimizes the bias and RMSE of the model. We compare the accuracy of the snow-depth model using the proposed placements to an existing sensor network at the Southern Sierra Critical Zone Observatory. Each model is validated with a 1 m^2 LiDAR-derived snow-depth raster from 14 March 2010. The proposed algorithm exhibits higher accuracy with fewer sensors (8 sensors, RMSE 38.3 cm, bias = 3.49 cm) than the existing network (23 sensors, RMSE 53.0 cm, bias = 15.5 cm) and randomized placements (8 sensors, RMSE 63.7 cm, bias = 24.7 cm). We then evaluate the spatial and temporal transferability of the method using 14 LiDAR scenes from two catchments within the JPL Airborne Snow Observatory. In each region, the optimized sensor placements are determined using the first available snow raster for the year. The accuracy in the remaining LiDAR surveys is then compared to 100 configurations of sensors selected at random. We find the error statistics (bias and RMSE) to be more consistent across the additional surveys than the average random configuration.

1. Introduction

A major challenge of hydrologic science in montane regions relates to estimating the spatial variability of snow cover [*Bales et al.*, 2006; *Trujillo and Lehning*, 2015; *McCreight et al.*, 2014; *Guan et al.*, 2013]. Multiple independent variables govern the distribution of snow cover, including elevation, slope, aspect, and the distribution of canopy [*Faria et al.*, 2000; *Musselman et al.*, 2008; *Lehning et al.*, 2011; *Helfricht et al.*, 2014]. Nonstationary effects such as climate warming and changes in vegetation structure may significantly alter the timing and magnitude of storage and runoff in these watersheds [*Goulden and Bales*, 2014; *Flanner et al.*, 2009]. Existing regression-based hydrologic models, which use statistical relations from historical hydrographs to predict runoff and inform allocation decisions [*Perkins et al.*, 2009; *Rosenberg et al.*, 2011; *Rango and Martinec*, 1995] will have limited accuracy as conditions deviate from historical norms and thus may prove to be inadequate for predictions in water management.

Recent research has focused on improving hydrologic models by assimilating remote-sensing and in situ measurements with distributed energy-balance models to better estimate storage and runoff [*Guan et al.*, 2013]. These methods use well-developed remote-sensing [*Painter et al.*, 2003; *Rosenthal and Dozier*, 1996; *Dozier*, 1989; *Pälli et al.*, 2002; *Egli et al.*, 2012] and energy-balance models [*Marks et al.*, 1992; *Link and Marks*, 1999; *Brubaker et al.*, 1996] to estimate snow and snowmelt processes across basins. In situ measurements for these methods are presently limited to snow pillows and snow courses, which largely sample flat, open terrain [*Molotch and Bales*, 2006], yet the distribution of snow cover can vary considerably as a function of topographic features. To address this, in situ sensor measurements can be deployed to capture the mean and variance of the snow depth, which can be used to inform models that use these statistics as inputs [e.g., *Essery and Pomeroy*, 2004]. Alternatively, individual sensor measurements can be used together with

distributed models to estimate snow distribution across uninstrumented regions [*Balk and Elder*, 2000; *Erickson et al*., 2005; *Erxleben et al*., 2002; *Fassnacht et al*., 2003; *Harshburger et al*., 2010].

Prior studies have investigated the feasibility of using wireless-sensor networks to distribute representative snow measurements over a broader landscape [*Kerkez et al*., 2012; *Rice and Bales*, 2010]. These studies demonstrated that wireless-sensor networks can be configured to provide simultaneous measurements of snow-depth distribution, solar forcing, and subsurface exchange across a 1 km^2 region. Based on these findings, 14 additional 1 km^2 area networks were deployed across the American River Basin to develop a real-time water information system [*Kerkez et al*., 2010]. Existing deployment methods for wireless-sensor networks require extensive field surveys in order to identify sampling regions for a limited budget of sensor nodes. These practices are resource intensive, provide no guarantees on accuracy, and are unsustainable if wireless-sensor networks are going to see larger-scale adoption for real-time monitoring.

Although snow-depth statistics are site specific [*Grünewald et al*., 2013], and are often unknown prior to deployment, the distribution of independent physiographic variables is often known at high resolution [e.g., from LiDAR surveys, *Deems et al*., 2013]. Sensors could be deployed in representative locations based on the remotely sensed data. Then, site-specific parameters for a snow-depth estimator could be inferred from the initial deployment in order to determine whether additional sensors are needed. This process would be well-suited to the field of machine learning, which includes "unsupervised" algorithms (which identify patterns in the independent variables without observing the dependent variable), and "supervised" or "active" algorithms (which employ limited observations of the independent variable in order to improve the estimator). A number of recent studies have applied machine-learning algorithms to topics in hydrology, such as runoff and streamflow estimation [*Solomatine and Shrestha*, 2009; *Londhe and Charhate*, 2010], evapotranspiration modeling [*Torres et al*., 2011], streamflow forecasting [*Rasouli et al*., 2012], assessment of the contamination of groundwater [*Khalil et al*., 2005a], and estimation of needs for reservoir releases [*Khalil et al*., 2005b; *Ticlavilca and McKee*, 2010].

In the present study, we evaluated a two-step machine-learning method to optimize sensor locations for catchment-scale snow-depth observatories. First, we used a Gaussian mixture model (an unsupervised algorithm) to identify representative sampling locations in a LiDAR-derived feature space (i.e., the multivariate distribution of independent variables that exist within the catchment). Second, we used a Gaussian process to estimate the catchment-scale distribution of snow-depth and model uncertainty, then placed additional sensors in high-uncertainty regions (i.e., supervised updates). The aim of this process is to determine the distribution of sensors that will minimize the RMSE and bias of the Gaussian process estimate throughout the catchment. The aims of the present study are to (i) determine how many sensors are needed in the unsupervised step in order to estimate the catchment-scale parameters for the estimator (i.e., the regression weights and autocorrelation for each independent variable), (ii) assess how many additional sensors are needed in the supervised step to optimally instrument the catchment, (iii) compare the accuracy of the snow-depth model using placements determined with the algorithm to an existing network and to randomized placements, and (iv) determine the spatial and temporal transferability of the method by evaluating the accuracy of the algorithm across multiple regions and over multiple LiDAR flights.

2. Methods

2.1. Study Areas and Data Collection

We evaluated the accuracy of the algorithm in three 1 km^2 catchments. The first study site is located in the Southern Sierra Critical Zone Observatory (SSCZO) (37°4′N, 119°11′W), within the rain-snow transition of the Sierra Nevada near Fresno, California. In water year (WY), 2010 (1 October 2009 to 30 September 2010), a 23-node wireless-sensor network was installed (Figure 1), spanning a 1.5 km transect in a forested headwater catchment where approximately 50–60% of annual precipitation falls as snow [*Bales et al*., 2011, Figure 9]. The catchment covers elevations between 1900 and 2100 m a.s.l., which are centrally distributed at 2000 m a.s.l. The mean and median slopes in the catchment are 11.6° and 10.2°, respectively. Though a range of aspects exist within the catchment, the distribution of "northness" is weighted toward 180°, as the catchment faces SW, toward the Central Valley. The canopy distribution is bimodal: most of the site is either open (penetration fraction = 1) or dense (penetration fraction = 0). For the present study, sensor locations in the existing wireless network were measured with a Trimble GPS (10 cm horizontal accuracy).

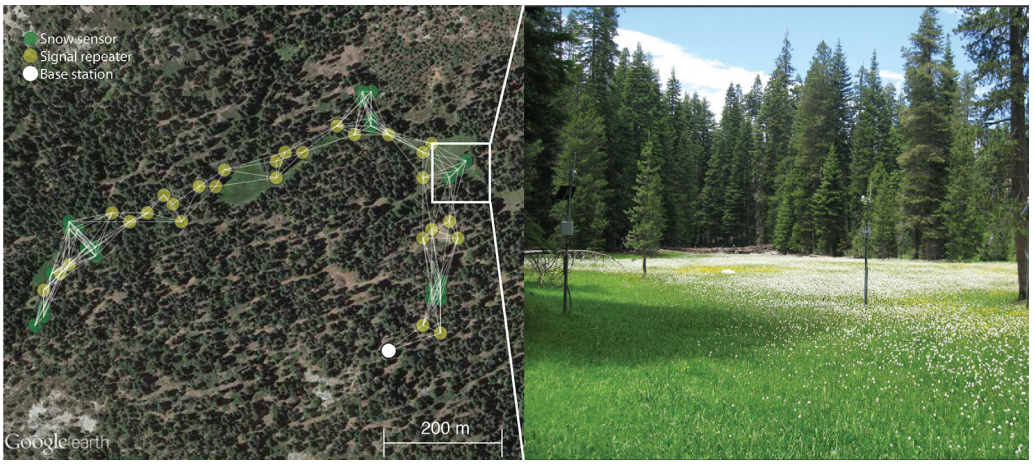

Figure 1. An existing wireless-sensor network at the Southern Sierra Critical Zone Observatory. Clusters of ultrasonic point measurements of snow depth are distributed across a 1.5 km transect in a 1 km^2 catchment. Data from each cluster are relayed every 15 min through a network of wireless elements to a base station with a real-time data uplink.

Node locations in the existing network are clustered such that the distribution of sensor measurements can be used to represent the catchment-scale mean and variance. Particular attention was paid to the effect of canopy, with clusters of sensors measuring the drip-edge, under-canopy, and open regions at multiple locations in the catchment. Although these measurements are spatially proximate, they represent 23 distinct measurements at dissimilar points in the feature space when used in the Gaussian-process estimator. Figure 1 shows the overall sensor distribution and typical network structure. The right-hand plot shows a cluster of sensors in the network that were placed to capture drip-edge-to-open gradient in the NE region of the catchment. Other clusters in the network are designed to capture under-canopy effects as well as gradients of slope, aspect, and elevation. The existing network provides a representative sampling of canopy and aspect, but significantly under-represents high and low elevations (elevations between 1850–1950 and 2020–2100 m are not covered), as well as high values of slope (slopes above 15° are not covered).

Physiographic variables (Figure 2) and a snow on/snow off raster (Figure 3) were gathered from the NSF Open Topography database, opentopography.org [accessed 1 February 2016]. Elevation, slope, and aspect extracted from LiDAR data were processed in ArcMap 10.2. On average, 9.21 points per square meter were used to generate the 1 m^2 DEM, canopy, and snow-depth rasters. Elevation information was stored as a

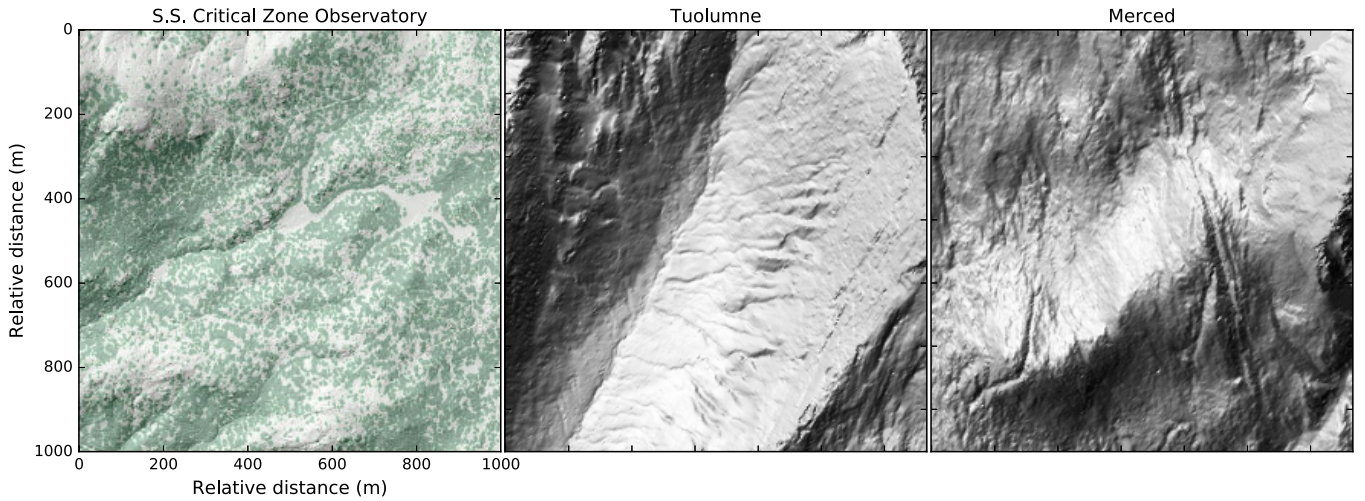

Figure 2. Hillshade maps showing the LiDAR-derived digital elevation models at each of the three sites used in the present study. At the Southern Sierra Critical Zone Observatory (left-hand plot), the distribution of canopy is overlain in green. The raster at the Southern Sierra Critical Zone Observatory was derived from LiDAR data in *Anderson et al.* [2012]. The rasters at the Tuolumne and Merced locations were derived from the JPL Airborne Snow Observatory [*Painter et al.*, 2016].

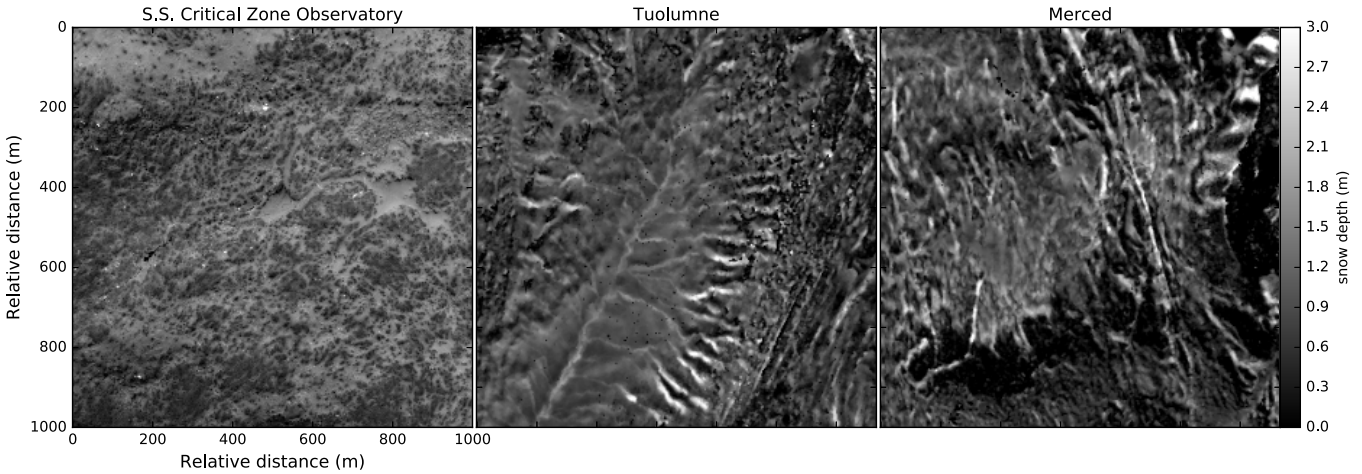

Figure 3. LiDAR-derived snow depth in each of the 1 km^2 catchments was used to determine the optimized sensor locations for each site. At Tuolumne and Merced, these are the first rasters available for WY2014. The accuracy of the placements in the remaining LiDAR surveys is shown in Figures 9 and 10.

point cloud in raw LiDAR data and the points of ground returns were gridded, averaged, and smoothed in order to create a high-resolution digital elevation model (DEM). Slope and aspect were calculated at 1 m^2 resolution from the gradient of the DEM in the longitudinal and latitudinal directions in the nine adjacent grid cells. The "northness" component of aspect was used in the present study (0–180°), as N/S differences affect snow cover due to differences in solar forcing.

The snow-depth raster was calculated by differencing the snow on/snow off surveys from 14 March 2010 [*Anderson et al*., 2012]. The snow depth is Gaussian-distributed between 0 and 3 m (mean = 1.42, standard deviation = 0.50). A comparison against ground-truth surveys conducted during the LiDAR surveys showed 0.1 m of vertical error in the snow-depth raster. There were no returns in 28.5% region due to dense canopy. In these regions, we first segmented trees from the LiDAR canopy height model using the variable-area local maxima algorithm and a typical pine tree height-diameter ratio [*Swetnam and Falk*, 2014; *Wonn and O'Hara*, 2001] and found that 2.3% of regions with no returns were tree trunks. We then set the snow depth at these points to zero and gap-filled the remaining unobserved regions using bi-linear interpolation. The interpolation scheme was based on a prior analysis of snow-depth variability at a nearby instrument cluster [*Bales et al*., 2011], which determined that under-canopy measurements were typically 30 cm lower than drip-edge. Finally, we calculated the canopy penetration fraction from the LiDAR point cloud based on the methods presented in *Zheng et al*. [2016]. The raw and gap-filled LiDAR data are shown in Figure 4.

Sites for the long-term analysis were selected from within the Airborne Snow Observatory (ASO) [*Painter et al*., 2016]. The ASO mission runs LiDAR surveys every other week (beginning the last week of March) in the Merced and Tuolumne river basins of the Sierra Nevada. For the present study we selected a 1 km^2 catchment from within each basin. Both sites were selected from above the tree line. This provided two advantages for the present study: it allowed us to evaluate the transferability of the algorithm in a new environment (i.e., not densely forested), and meant that the snow rasters did not have to be interpolated under dense canopy. The accuracy of the algorithm could therefore be evaluated based solely on measured values. The Tuolumne site is located at (38.044591, −119.427359°) at 3100 m a.s.l. and is due East from the Hetch Hetchey reservoir. The Merced site is located at (37.737945, −119.302983°), 3450 m a.s.l. The 3 m resolution LiDAR rasters of site characteristics and snow depth are shown in Figures 2 and 3, respectively. At the Tuolumne catchment, the elevation range is 3085–3340 m, the aspect is predominately west-facing and east-facing, and the mean and median slopes are 19.3 and 18.2°, respectively. The Merced catchment spans 3270–3680 m, the aspect is predominately south-facing, and the mean and median slopes are 20.0 and 17.5°, respectively. Dates for the LiDAR surveys at each site are listed in Table 2. Based on the in situ measurements from the Gin Flat snow pillow (located between the Merced and Tuolumne basins at 2150 m), the differences in snow depth between the seven LiDAR scenes were primarily attributable to: accumulation (between scenes 1 and 2), melt (2 and 3), melt (3 and 4), accumulation (4 and 5), melt (5 and 6), accumulation, and then melt (6 and 7).

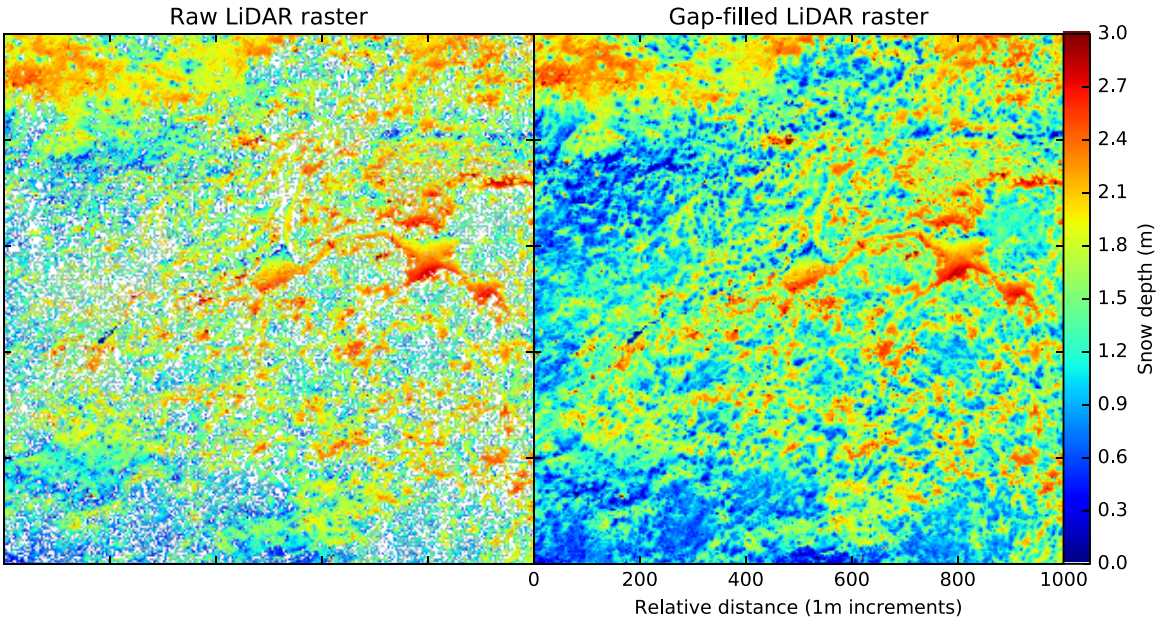

Figure 4. (left) Raw and (right) gap-filled LiDAR snow rasters at the Southern Sierra Critical Zone Observatory. Dense canopy resulted in no returns for under-canopy measurements in 26.2% of the catchment. These points were bilinearly interpolated from adjacent measurements (described in section 2.1). The snow rasters in Tuolumne and Merced did not require gap-filling.

2.2. Identification of Representative Sampling Locations

In the first step of the proposed method, we determined the distribution of sensors that is most-representative of the LiDAR-derived feature space. This is accomplished using a Gaussian mixture model, which assumes that a feature space (i.e., the combined $\mathbb{R}^6$, $\mathbf{x}=[x^{\text{lat}}, x^{\text{lon}}, x^{\text{slope}}, x^{\text{aspect}}, x^{\text{elevation}}, x^{\text{canopy}}]$ LiDAR data from section 2.1) is a product of a finite number of latent (unobserved) components (i.e., sensors). The sensor's ability to observe each point in the feature space is represented using a multivariate normal distribution (equation (1)). This is the parametric expression of each component of the mixture. The expected value of such a normal distribution is the sensor's location in the feature space. Multiple Gaussian distributions (multiple sensor locations) were combined and weighted with mixing parameters from an ensemble of M mixture elements in equations (2) and (3). The combined ability of all sensors to observe all points in the feature space is represented using a likelihood function (equation (4)).

Formally, the Gaussian mixture model is a linear superposition of D-dimensional multivariate normal distributions, $\mathcal{N}$, with expected value μ and covariance Σ applied to data $\mathbf{x}$, which are vectors including the four LiDAR derived variables. Each component is defined as:

$$\mathcal{N}(\mathbf{x}|\mu,\Sigma)=\frac{1}{(2\pi)^{D/2}}\frac{1}{|\Sigma|^{1/2}}\exp\{-\frac{1}{2}(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)\} \tag{1}$$

and their superposition, with mixing parameters, π_m, with m being an index denoting a single mixture component, is defined as:

$$p(x)=\sum_{m}=1^{M}\pi_m\mathcal{N}(\mathbf{x}|\mu_m,\Sigma_m) \tag{2}$$

subject to:

$$\sum_{m=1}^{M}\pi_m=1 \tag{3}$$

The complete log-likelihood function, evaluated over all the LiDAR-derived measurements, N is given by:

$$\ln\ p(\mathbf{x}_n|\pi,\mu,\Sigma)=\sum_{n=1}^{N}\ln\left\{\sum_{m=1}^{M}\pi_m\mathcal{N}(\mathbf{x_n}|\mu_m,\Sigma_m)\right\} \tag{4}$$

We used the Expectation Maximization (EM) algorithm [*McLachlan and Peel*, 2004; *Pedregosa et al*., 2011] to retrieve the sensor locations. This is an iterative process in which the algorithm tries to recover the most-likely parameter estimates for the mixture of multivariate $\mathcal{N}_m$ to explain the data. We used a spherical covariance function and updated the model weights, covariance, and means with each iteration. Once the maximization step no longer increases the log-likelihood, the process terminates and the optimized sensor locations have been found. Like many gradient-based optimization methods, EM converges to local minima. Therefore we used 100 initializations from randomized starting points and selected the result with the maximum likelihood. We parallelized this process onto four computational cores and subsampled the feature space (using a 1-in-16 point subsampling) to make the evaluation of many starting points computationally tractable. We then performed a nearest neighbor search through the full feature space (i.e., not subsampled) in order to find the physical location that most closely matches the features of each Gaussian mixture model mean.

2.3. Snow-Depth Model

In the second step, the LiDAR-derived snow-depth measurements at the locations proposed by the Gaussian mixture model were used to estimate the distribution of snow depth across the catchment using a Gaussian process [*Rasmussen and Williams*, 2006]. A Gaussian process is a regression technique that predicts a dependent variable (in our case, the snow depth, *Y*) using a set of dependent variables, *X*, which are expected to be informative for its prediction. It establishes a covariance-based model that, using known input-output relationships (the training data at the measuring stations locations, $\{\mathbf{x}_n, y_n\}_{n=1}^{N}$), is able to predict the snow depth on new, unseen locations where the inputs can be obtained. In our specific application, the physiographic variables (slope, aspect, elevation, and canopy) were used as dependent variables in the estimation (i.e., for a single point in $\mathbb{R}^4$, $\mathbf{x}=[x^{\text{slope}}, x^{\text{aspect}}, x^{\text{elevation}}, x^{\text{canopy}}]$). The Gaussian process (equation (5)) combines the *N* point measurements $\mathbf{X}^N$ from each sensor station and the four LiDAR-derived physiographic variables to estimate the mean snow depth, $\bar{Y}$, using a covariance function *K*,

$$Y \sim GP(\bar{Y}, K), \tag{5}$$

where the $\sim$ symbol means "is distributed following." An illustration of the Gaussian process is shown in Figure 5: given the observed data and uncertainty (sensor stations, in red) and a prior over the shape and parametrization of the functions which are likely to be observed (e.g., all functions are Gaussians with mean $\bar{Y}$ and covariance *K*), the Gaussian process estimates a posterior distribution with a predictive average (dashed line) and variance (the uncertainty of the prediction—the gray envelope). Note that close to the observed data, the variance approaches zero. Readers interested in the mathematical details may refer to Chapter 2 in the book by *Rasmussen and Williams* [2006]. In the present study, we used a squared exponential covariance function (K_{se} in equation (6), which depends on four variables: $\mathbf{x}$ and $\mathbf{x}'$ are two points in the domain, *d* is the distance between them, and *l* is the characteristic length scale (computed individually for each independent variable in the normalized feature space). We used this covariance, since we assume that the snow depth varies smoothly with respect to the input variables. The autocorrelation in the model is controlled by the *l* parameter. We assume the algorithm has no prior information about this parameter; it must be estimated only from the point measurements at each sensor. Regression weights and autocorrelation are estimated using a constrained optimization by linear approximation algorithm in the software package SciPy (fmin cobyla) [*Pedregosa et al*., 2011]. The error at each measurement (0.1 m based on the LiDAR error) was quantified using the regularization in equation (7), where y_i is the measured snow depth at point *i*, and σ_i is the measurement variance at y_i.

$$K_{se}(x,x')=exp\left(-\frac{||x-x'||^2}{2l^2}\right)=exp\left(-\frac{||d||^2}{2l^2}\right) \tag{6}$$

$$r_i=\left[\frac{\sigma_i}{y_i}\right]^2 \tag{7}$$

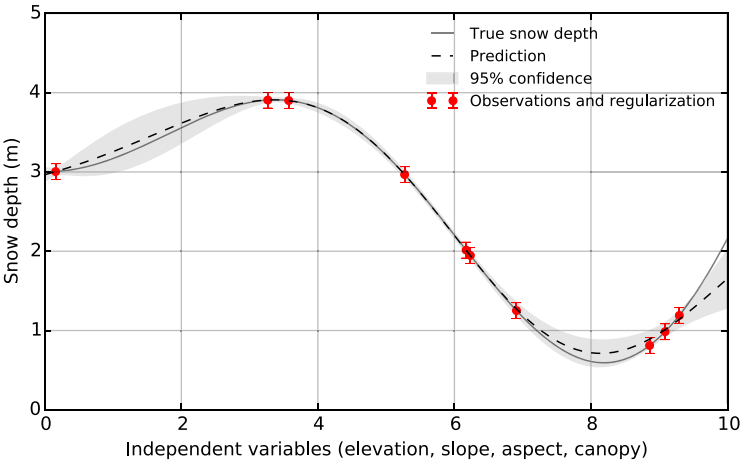

Figure 5. Illustration of a Gaussian process shown in one of the four dimensions. Uncertainty in the observations (0.1 m in LiDAR data) is quantified with a regularization at each point (equation (7)). The confidence of the prediction at intermediate points is a function of the covariance parameters estimated from measurements at the sensor locations. The estimated uncertainty is employed by the supervised updating process.

2.4. Supervised Updates

In addition to estimating the spatial mean of the snow depth, the Gaussian process estimates the distribution of model uncertainty from the covariance matrix of the posterior distribution, which is a function of the independent variable weights and estimated autocorrelation. This provides a basis for placing additional sensors with the aim of minimizing uncertainty throughout the catchment. Using the measurements from the sensors placed in section 2.2, we modeled the snow depth using the Gaussian process and extracted a new feature space corresponding to the highest 1% and 10% of model uncertainty. We then performed clustering with the Gaussian mixture model in each feature space to select new sensor locations that both reduce the uncertainty of the model and are unique with respect to the combination of topographic variables they sample. The additional placements are considered supervised updates because they rely on observations of the dependent variable (i.e., the estimated autocorrelation and regression weights in the snow-depth model) to determine optimal locations. In the machine-learning community, the procedure is known as "active learning" [*Settles*, 2012]. They are distinct from the placements in section 2.2, which are unsupervised in the sense that no observations of the dependent variable (snow depth) are used in their determination. Sensors placed in regions of high uncertainty will reduce the uncertainty throughout the domain at points with similar combinations of physiographic features.

2.5. Model Evaluation

The error in the snow-depth model was determined by differencing the predicted snow depth under each scenario from the LiDAR-derived snow-depth raster (Figure 3). We quantified the accuracy of the snow-depth model under each placement scenario using two standard metrics: RMSE (equation (8)) and Bias (equation (9)). In each equation, n is the number of points in the model, $\hat{y}_i$ is the model prediction at point i, and y_i is the true snow depth.

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i-y_i)^2} \tag{8}$$

$$\mathrm{Bias}=\sum_{i=1}^{n}(\hat{y}_i-y_i) \tag{9}$$

2.6. Optimal Number of Sensors

We investigated how many sensors are needed in the unsupervised step by examining the error in the snow-depth model in a range of placement scenarios. In the first set of scenarios, we began the unsupervised placements with two sensors, increasing to 23 sensors. We took the optimal number of sensors to be when the marginal improvement in RMSE was less than 10%. Given the placements from the unsupervised process, we added sensors using the process described in section 2.4, up to the budget of 23 sensors. Finally, with the optimal number of sensors determined, we compared the snow-depth estimate resulting from the proposed placements to an equivalent number of randomly chosen, but spatially distributed sensors. We evaluated the snow-depth model under 100 randomized configurations to determine the expected accuracy of the randomized placements.

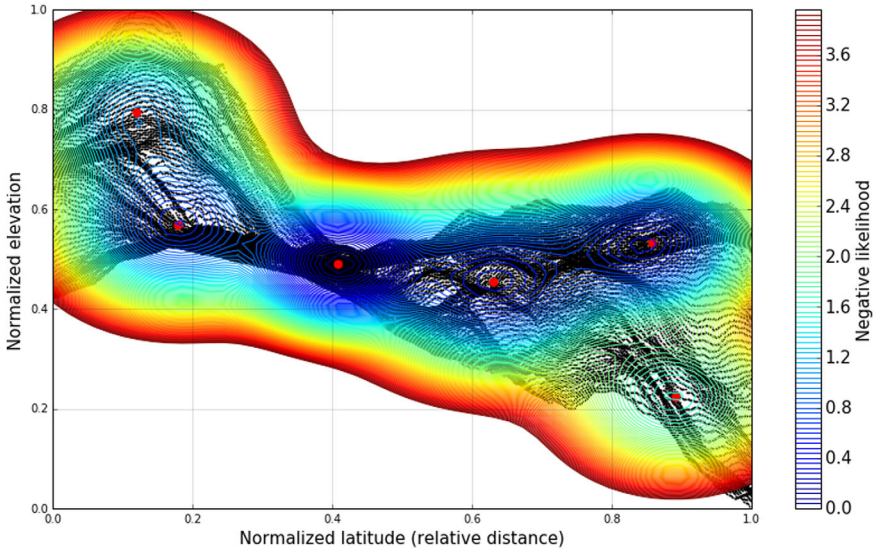

Figure 6. Representative sensor locations were determined using a Gaussian mixture model in the space of independent variables from Figure 2. Black points represent a two-dimensional projection of the $\mathbb{R}^6$ LiDAR data onto the latitude-elevation plane. The optimized sensor locations (indicated by red points) are the expected values of the latent Gaussians. The likelihood function (contour lines) quantifies how well each point in the LiDAR-derived feature space is observed given the locations and covariance of each sensor (see equation (4)). The optimal parameters (expected value and covariance) for the model were estimated using the EM algorithm [*McLachlan and Peel*, 2004].

2.7. Long-Term Accuracy

Only one LiDAR survey was available for the region covered by the existing wireless-sensor network, which is insufficient to assess the long-term accuracy of the algorithm. To address this, we chose two regions within the JPL Airborne Snow Observatory (described in section 2.1) that had multiple LiDAR surveys during WY2014. First, we performed the unsupervised placements based on the independent variables at each site (described in section 2.2). Second, we performed the supervised update step (described in section 2.4) in the first snow-depth raster of WY2014. We then evaluated the accuracy of the model estimated from the sensors in seven additional LiDAR surveys for each catchment, and quantified the bias and RMSE in each LiDAR survey. Results were then compared to the average of 100 random configurations of sensors. Finally, we evaluated the ability of the sensors to represent the catchment-scale mean and standard deviation of snow depth in each of the LiDAR surveys.

3. Results

Using the methods described in section 2.6, we selected six sensors in the unsupervised step for the SSCZO catchment. An illustration of the converged Gaussian mixture model for this configuration is shown in Figure 6. Optimal sensor placements (the expected values of the latent Gaussians) are shown as red markers. The likelihood function (quantifying how well the space of independent variables is observed under the current sensor configuration), is shown as contour lines. Points that are well-observed are shown in blue; poorly observed points are shown in red. As it is not possible to show the full six-dimensional feature space, Figure 6 illustrates a two-dimensional projection of the output. Adding more than six sensors in the unsupervised step did not significantly improve the accuracy of the algorithm (see Table 1). With 6 sensors, the accuracy of the algorithm was 43.23 cm; at 23 sensors, this was reduced only to 42.90 cm. Intermediate values were variable, and in one case as high as 73.88 cm (at 14 sensors). Two sensors were then placed in the supervised step (using the 10% uncertainty threshold), reducing the RMSE to 38.29 cm. The effect of the update is shown in Figure 7: the error becomes more tightly distributed around zero, but the bias remains essentially unchanged (it increases by slightly less than one centimeter).

In the 100 sample evaluation of randomized placements of eight sensors, the average RMSE was 63.7 cm. The distribution was heavy-tailed: most of the results were clustered between 40 and 75 cm RMSE, but there were eight outcomes with greater than 100 cm RMSE. A configuration corresponding to one of the average runs is shown in the middle plot of Figure 8. The spatial distribution is similar to the proposed

Table 1. RMSE, Bias With Increasing Number of Sensors in the Unsupervised Step at SSCZO

Sensors	Bias (cm)	RMSE (cm)
5	35.58	277.5
6	2.57	43.2
7	7.47	52.8
8	9.00	63.8
9	0.16	59.3
10	−0.23	62.7
11	−9.73	52.6
12	−14.15	46.0
13	−9.11	56.0
14	−3.52	73.9
15	−4.67	63.4
16	−3.02	47.1
17	−4.32	66.3
18	3.58	52.4
19	−16.79	65.2
20	4.67	43.3
21	−8.25	43.9
22	−9.74	46.6
23	5.41	42.9

placements, however the output of snow-depth model reveals a slight over-estimation of the snow-depth, particularly in the SW regions of the catchment (right-hand plot in Figure 8), resulting in a positive bias for the estimator (blue line in Figure 7).

The snow-depth model estimated from the existing 23 placements has higher error (RMSE 53.0 cm, bias 15.49 cm) than the estimates from the machine-learning method. The accuracy of the snow-depth model is high (less than 15 cm error) near the sensor clusters. However, the error is very high (greater than 1 m error) in the NW and SW region, where the snow-depth model over-estimates the true snow depth. This produces a more heavy-tailed error distribution than the proposed or random placements, and an overall overestimate of the true snow depth within the catchment (green line in Figure 7). Despite this, the true mean (1.42 m) of the snow-depth raster are better estimated by the sensors in the existing network (i.e. without using interpolation, but simply averaging the measurements to estimate a mean value for the catchment). The mean and standard deviation of the sensor measurements is 1.76/0.49, 1.54/0.52, and 1.48/0.57 m for the proposed, random, and existing network respectively.

In the Tuolumne ASO site, 11 sensors were placed in the unsupervised step, and 4 were added in the supervised step (using the 1% uncertainty threshold). The bias and RMSE of the 15 locations proposed by the algorithm in the first LiDAR scene are higher than at SSCZO (left-hand plot, Figure 9). This negative bias persists and decreases in magnitude in the remaining LiDAR surveys (top line in Table 2) and the RMSE is bounded between 64.2 and 70.6 cm in the remaining surveys. The bias and RMSE in the average of the 100 of the random trials are worse throughout the remaining LiDAR surveys (second number in top two lines of Table 2) and there is greater variability in both quantities. Neither the proposed nor the random placements capture the small-scale variability along the drainages in the center of the catchment (see spatial distribution of error in the right-hand plots, Figure 9). Although the machine-learning algorithm consistently outperforms the average of the random placements in estimating location-specific error, the correspondence between the measured snow mean and the snow variance is inconsistent (bottom two lines in Table 2). The

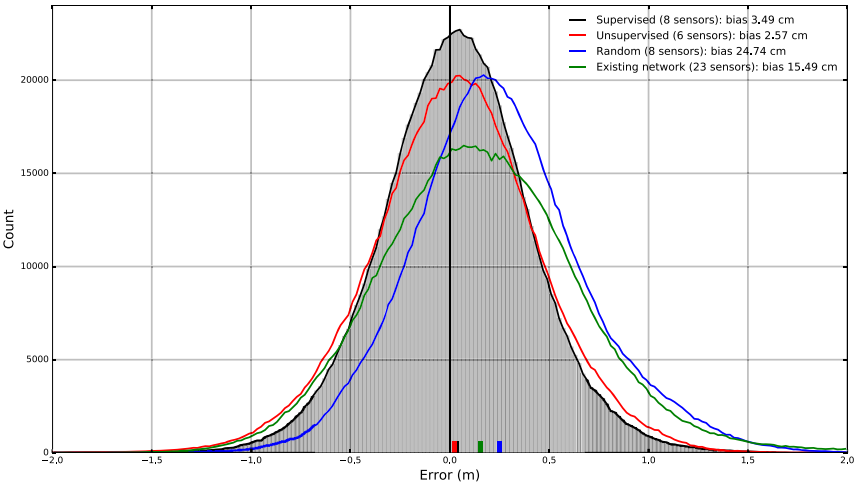

Figure 7. Histograms showing the distribution of error in the snow-depth model under each sensor configuration from Figure 8. The placements proposed by the machine-learning algorithms have lower bias and RMSE than the eight random sensors and the existing network. The RMSE is lowered in the supervised update, but the bias slightly increases compared to the unsupervised placements.

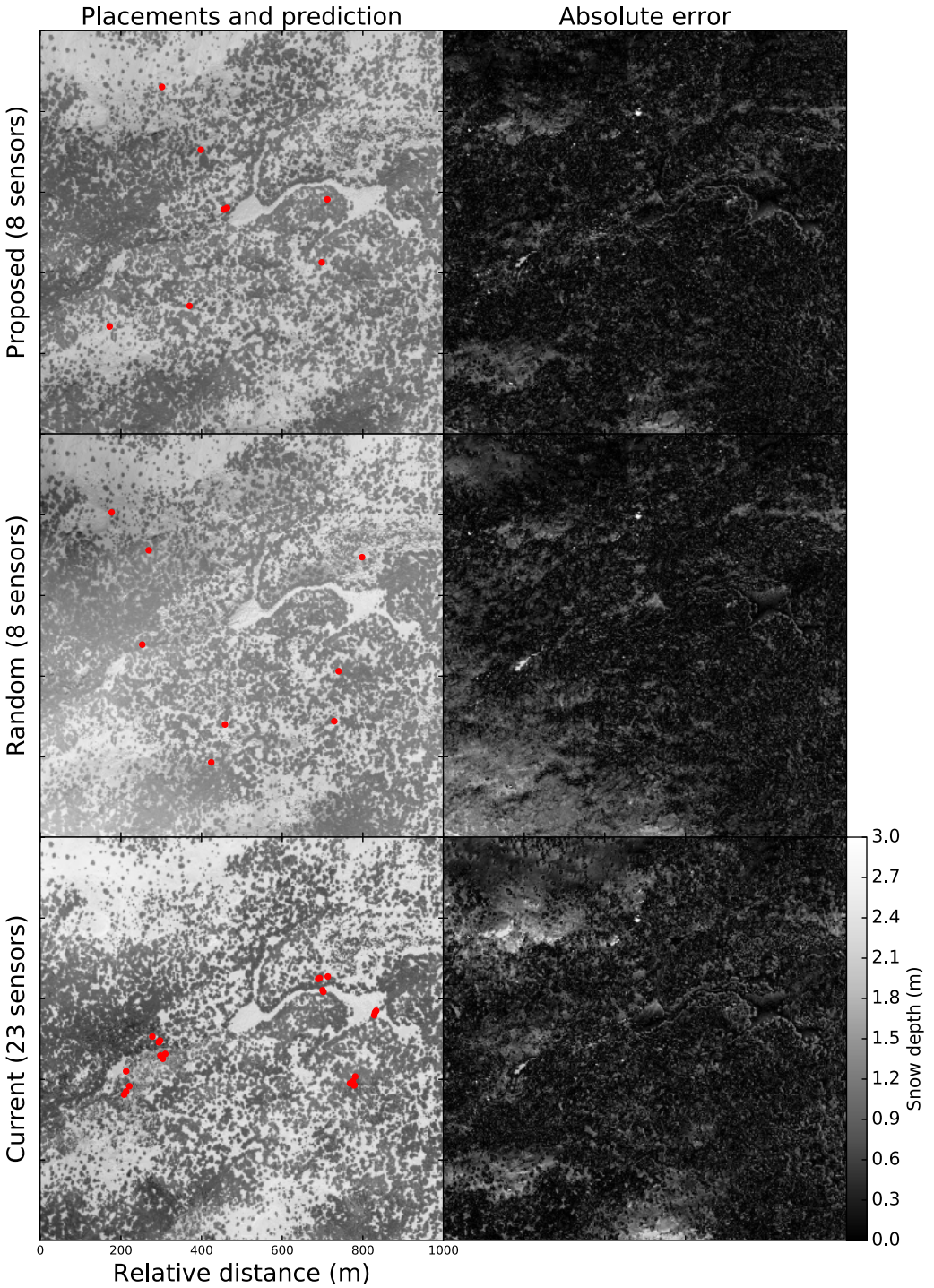

Figure 8. Predicted snow depth (left-hand plots) and error (right-hand plots) using three sensor configurations: (top) proposed, (middle) random, and (bottom) existing. The resolution of the snow-depth model is 1 m^2, the scale in the left-hand plot is 0–3.0 m. The error of each model is computed relative to the LiDAR snow-depth raster in Figure 3. The distribution of errors is compared in Figure 7.

mean and standard deviation are generally underestimated by the proposed placements and overestimated by the random placements.

In the Merced ASO site (Figure 10), 13 sensors were placed in the unsupervised step, and 1 was added in the supervised step (using the 10% uncertainty threshold). The RMSE and bias of the 14 locations proposed by the algorithm are lower than at Tuolumne (top lines in the "Merced" subsection of Table 2). The bias

Table 2. Long-Term Error and Snow-Depth Statistics in ASO LiDAR Data

Tuolumne	3-23-14	4-07-14	4-13-14	4-20-14	4-28-14	5-02-14	5-11-14
Bias[a]	−12.2/44.4	−15.3/47.3	−9.9/43.4	−13.9/35.4	−3.5/50.4	−5.6/27.3	−10.3/37.8
RMSE[b]	60.9/89.0	70.6/108.2	64.5/97.5	66.4/80.3	65.5/105.6	64.2/74.9	64.2/76.4
Mean[c]	115/101/136	153/130/172	127/105/145	110/75/125	124/97/142	94/69/115	85/59/107
Std[d]	58/30/88	65/35/109	62/35/98	64/44/105	65/41/102	61/37/96	61/40/95
Merced	3-24-14	4-06-14	4-14-14	4-23-14	4-29-14	5-03-14	5-12-14
Bias	0.4/−1.5	−2.4/−6.8	6.1/9.0	0.8/17.1	−5.5/12.8	0.6/6.4	5.2/12.6
RMSE	58.0/70.1	73.1/82.5	65.1/84.0	62.5/87.7	66.6/74.5	61.6/74.0	59.4/81.1
Mean	76/84/67	114/117/100	73/80/68	67/75/65	77/77/68	71/75/66	54/61/48
Std	58/49/48	69/64/63	61/52/49	61/54/52	66/60/58	63/59/52	57/49/46

[a]Bias of the Gaussian process estimator (proposed/random, cm).
[b]RMSE of the Gaussian process estimator (proposed/random, cm).
[c]Mean snow depth (measured/proposed/random, cm).
[d]Standard deviation of the snow depth (measured/proposed/random, cm).

starts near zero and does not deviate above 6.1 cm. By contrast, the bias and RMSE resulting from the average of the random trials are high and variable throughout the remaining LiDAR surveys (second number in the top lines the "Merced" subsection of Table 2). Again, much of the small-scale variability in the catchment is not captured by either the random or proposed placements. The relationship between the mean and standard deviation of the sensor measurements compared to the catchment-scale statistics is again inconsistent (bottom two lines the "Merced" subsection of Table 2).

4. Discussion

Although individual point measurements of snow depth are poor estimators of the local mean, our results indicate that an automated process can be used to find a limited combination of representative placements that can be used to estimate the catchment-scale snow cover. Eight placements in strategic locations produced a better catchment-scale estimate than 23 placements in the existing network, likely due to the under-representation of high values of slope and elevation and the closer spatial distribution of the existing sensor clusters. The long-term analysis indicated that the method is transferable to other types of environments, and that the accuracy of the sensors placement determined from the first LiDAR survey persists in the remaining surveys. This was not true of the random placements, which showed greater variability in the bias and RMSE in the remaining surveys.

At the ASO catchments (Tuolumne and Merced), the RMSE of the snow-depth model was higher than at the SSCZO, despite the additional sensors placed by the algorithm. This is likely due to the greater exposure of the terrain above the treeline in the ASO sites. At the SSCZO, the highest errors are in exposed (i.e., unvegetated) areas in the NW, SW, and meadows (right-hand plots, Figure 8). At the ASO sites, the highest errors are distributed throughout the catchments (right-hand plots in Figures 9 and 10). The dense canopy closures at SSCZO likely constrained the wind redistribution effects to smaller regions within the catchment, thereby reducing the overall error and required number of sensors. It should also be noted that many of the under-canopy measurements in the SSCZO raster were interpolated. This may have affected the relative error and required number of sensors between the ASO and SSCZO sites by producing a more homogeneous distribution of under-canopy snow depth.

The present study is conducted in small catchments sampled at very high spatial resolution, whereas typical hydro-metrological studies are conducted at lower resolution (e.g., spatial resolutions on the order of 50–100 m). Running the algorithm at lower resolution will likely affect the accuracy of the spatial snow cover model and placements determined by the algorithm. If the above analysis is repeated with the LiDAR data averaged to 50 m resolution, the accuracy of the model increases and fewer sensors are placed by the algorithm. The variance explained by the low-resolution model in the first snow rasters for SSCZO, Merced, and Tuolumne is 56%, 59%, and 36% with 8, 10, and 8 sensors, respectively. Using the high-resolution data, the variance explained 45%, 6%, and 18% with 8, 14, and 15 sensors, respectively. This is likely due to small-scale variability being averaged out at lower resolutions. The temporal persistence of the accuracy is also observed in the 50 m data: all snow cover models chosen by the algorithm had lower RMSE than the

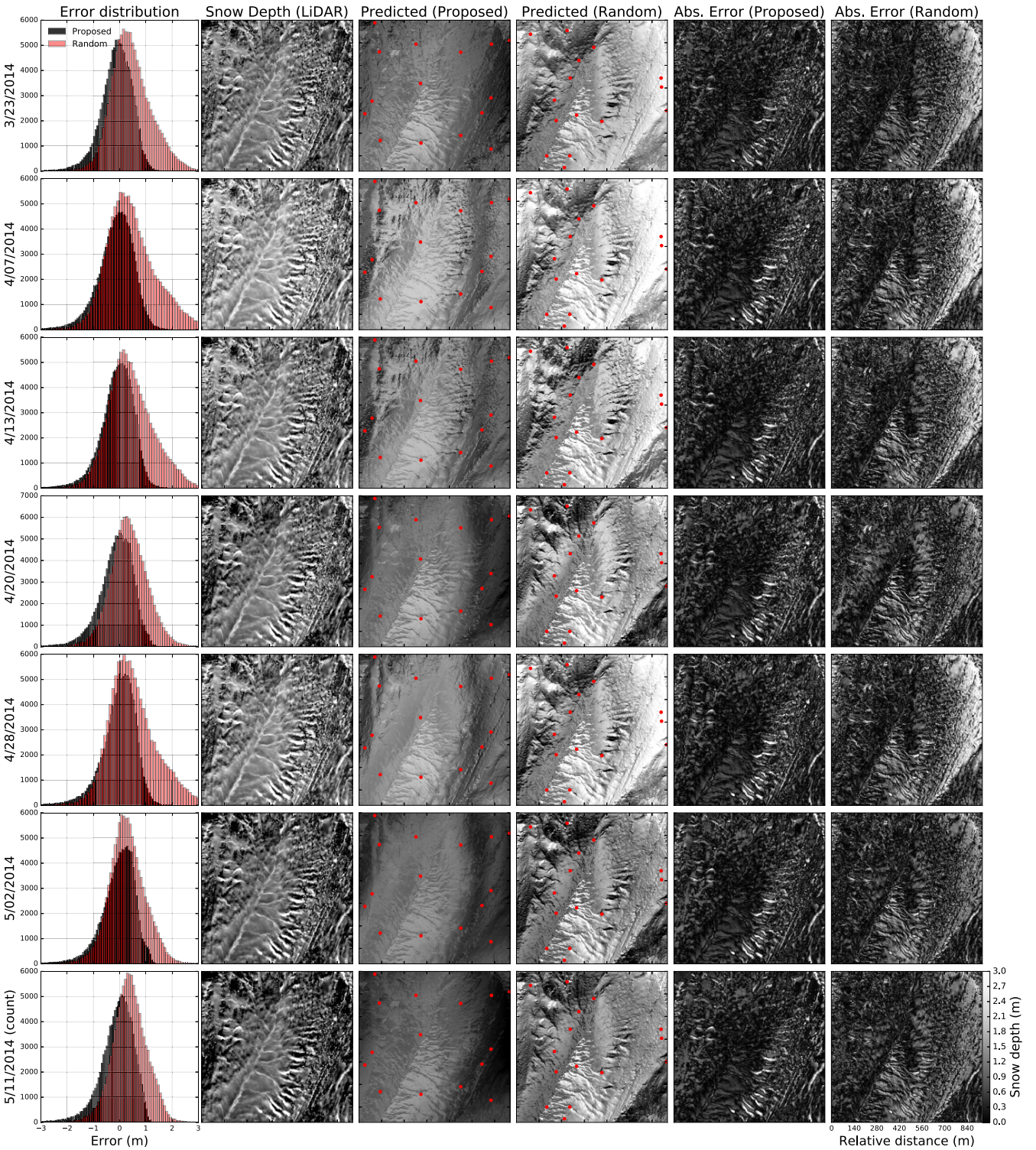

Figure 9. Snow-depth model accuracy (comparing proposed placements to the average of 100 random placements) in the remaining 2014 LiDAR surveys at Tuolumne. Sensor placements are overlain as red points.

average random configuration. The range of model accuracies in the present study is consistent with the findings in *Erxleben et al*. [2002], which analyzes the accuracy of multiple snow cover models in three 1 km^2 catchments. The study finds the models explain between 6.8% and 31% of the snow-depth variance, depending on the specific catchment and statistical model.

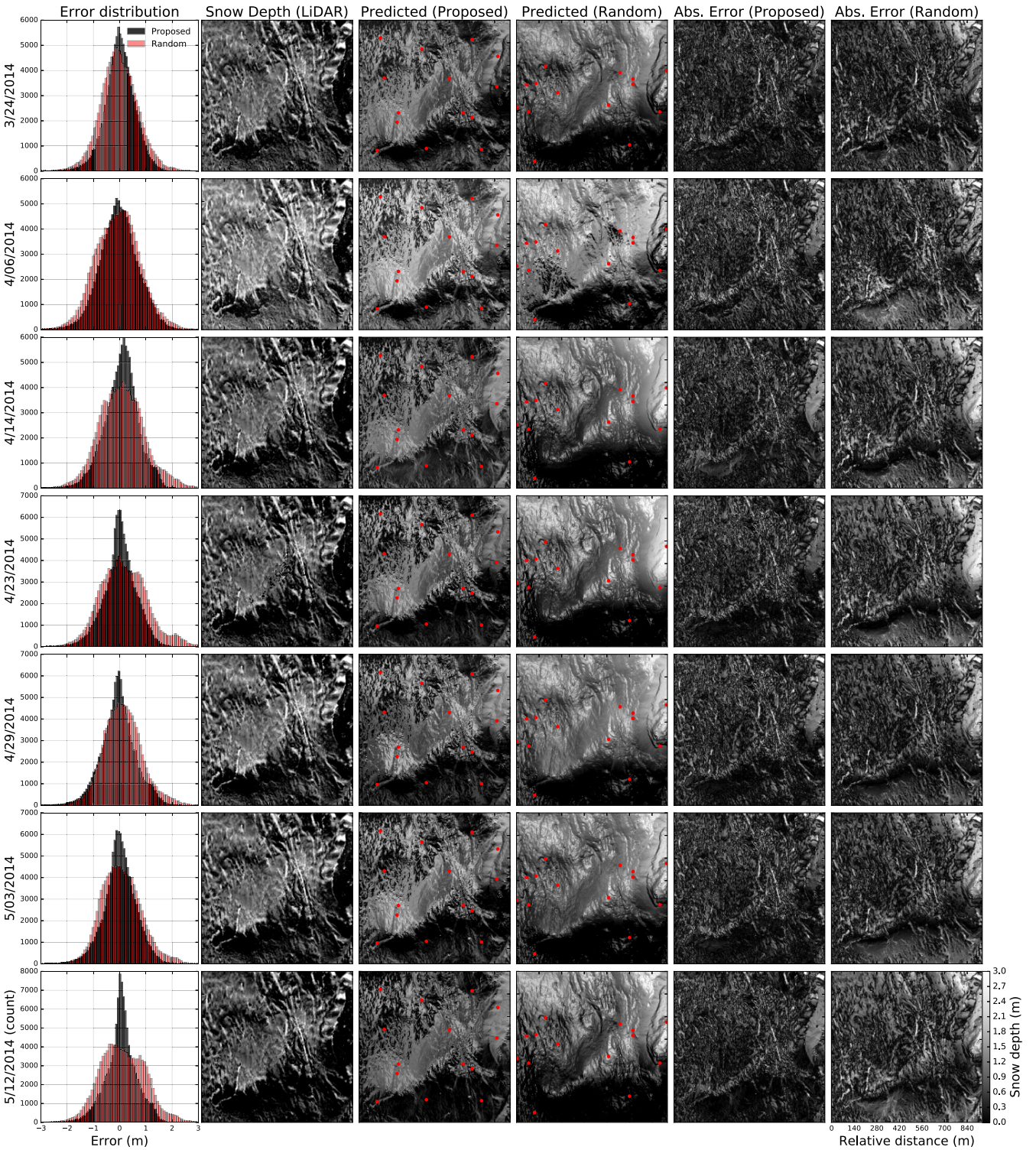

Figure 10. Snow-depth model accuracy (comparing proposed placements to the average of 100 random placements) in the remaining 2014 LiDAR surveys at Merced. Sensor placements are overlain as red points.

The accuracy of the high-resolution model in the present study may be improved by adding independent variables that capture the effects of small-scale terrain features and wind redistribution. None of the algorithms in the ASO catchments accurately modeled the accumulation of snow depth near sharp transitions

in the DEM, (e.g., along the drainages in Figure 9, and along the ridges of Figure 10). Similarly, none of the algorithms captured the increasing, S-facing gradient of snow depth across the meadow in the Southern Sierra Critical Zone Observatory, which was likely a result of wind redistribution. This effect has been observed in a number of prior studies [*Winstral and Marks*, 2002; *Molotch and Bales*, 2005], and directional variables have been suggested to account for directional redistribution of snow (e.g., in *Molotch and Bales* [2005]). Given the dense canopy throughout the catchment in the present study, we did not include a directional variable in the feature space. In catchments with less dense distributions of canopy, it would likely be necessary. Care should be taken in mixed open/forested regions that the directional bias measured in open regions is not translated to the regions with dense canopy.

It would be beneficial to evaluate the proposed method in a vegetated region over long time spans in which there are regular under-canopy measurements (from an in situ network or synoptic survey) and multiple LiDAR flights. The accuracy of the algorithm may exhibit greater temporal variability in such environments, since the influence of canopy is different throughout the year (affecting snow interception during the accumulation phase, and longwave radiation enhancement during the ablation phase). The method should also be evaluated in alpine catchments that exhibit lower temporal persistence in the spatial snow cover. The coefficients of determination between the first raster and successive rasters at Tuolumne and Merced are: 0.82, 0.81, 0.74, 0.80, 0.76, 0.73 and 0.93, 0.93, 0.86, 0.91, 0.90, 0.87, respectively. The long-term bias and RMSE may exhibit greater variability in catchments with lower temporal persistence.

Recent studies have demonstrated that mutual-information-based sensor placement strategies can outperform placements based on Gaussian Process uncertainty [*Krause et al*., 2008]. This approach was not used in the present study as the mutual information algorithm is NP-complete with complexity $O(kn^4)$ (where k is the number of sensors and n is the number of possible locations). The fourth-order dependence on the number of potential locations can yield computationally intractable problems when using LiDAR data. The number of candidate placements in the numerical studies in *Krause et al*. [2008] is on the order of 100–200, compared to over 1,000,000 in the present study. The mutual information-based strategy may be feasible to implement at lower spatial resolutions or by reducing the complexity using submodularity (discussed in *Krause et al*. [2008]). It therefore warrants further research for this application.

There are practical considerations in observatory design that are not considered in the present study. Regions of the catchment may be inaccessible due to terrain attributes and other access constrains. This can be addressed by defining a set of inaccessible placements in the feature space from Figure 6. If the optimal sensor location is determined to be on an inaccessible grid element, a search to the nearest viable point in the feature space will be output as the optimal point. The proposed algorithm also requires a greater spatial distribution than the clustering approach. Recent field deployments have indicated that 1 km scale wireless-sensor networks can be deployed using existing hardware in a variety of terrains. If the spatial extent is limited, the spatial coordinates could be removed from the unsupervised step, and an algorithm could determine the most-representative, spatially proximate distribution of nodes.

It should be noted that the "supervised updating" step in the present study would require two field deployments: one to gather data to estimate the distribution of uncertainty throughout the catchment (which also requires knowledge of the catchment-scale independent-variable distributions), and a second to add sensors in high-uncertainty regions. In practice, the marginal gain from the supervised updates may not be worth the marginal cost, though in our analysis, the supervised placements outperform unsupervised placements. In the present study, the first LiDAR flight was used to estimate the uncertainty and perform supervised placements. The selected locations may change depending on which raster is used, as the autocorrelation and regression weights will change throughout the year. Given additional years of LiDAR data, it may be better to perform supervised placements based on a long-term analysis of the uncertainty within the catchment.

The true optimal number of nodes per site should be determined by considering the marginal value of the improved information. Combined with an understanding of the marginal cost of each additional placement, this would determine an optimal number of sensor nodes by comparing marginal cost and value functions. This approach would capture site-specific cost/value tradeoffs for each sensor network.

5. Conclusion

The research reported here suggests that a machine-learning algorithm can be used to identify snow sensor locations in catchment-scale observatories prior to field deployment. In the three regions considered in the present study, the placements determined from the algorithm exhibited higher accuracy and less bias than an existing sensor network and an equivalent number of randomly selected locations. The accuracy of the algorithm was found to be consistent when the temporal transferability was evaluated in 14 LiDAR scenes within the Airborne Snow Observatory.

The aim of the present study was to determine sensor locations to be used together with a snow-depth regression model to estimate the catchment-scale distribution. Without the Gaussian process estimate, we found that the mean and standard deviation of the sensor measurements alone was not consistently more accurate than random or existing placements. Therefore, the proposed algorithm is likely better suited to situations in which the snow depth is to be modeled at the catchment scale, as opposed to distributed models which take the mean and variance as an input.

While the research reported here indicates that structuring catchment-scale observatories prior to field deployment using remotely sensed data may be a feasible alternative to conducting field surveys, further research is needed to quantify the accuracy of the algorithm over longer time spans, particularly in forested environments. It would also be useful to investigate how increased information from these networks translates to value for downstream users. This would provide a more objective metric for the marginal value of information, which could be compared to the marginal cost of network establishment to determine the best locations for new catchment-scale observatories.

Acknowledgments
This research was supported by the National Science Foundation, through the Southern Sierra Critical Zone Observatory (EAR-1331939); a Major Research Instrumentation grant (EAR-1126887); NSF-IGERT award DGE-0903711; UC-ANR grant 11-986; California Department of Water Resources; the U.C. Water Security and Sustainability Research Initiative; and the University of California Center for Information Technology Research in the Interest of Society. D. Tuia acknowledges the Swiss National Science Foundation, under grant PP00P2-150593. We also acknowledge the cooperation of M. Meadows, and E. Stacey at the SSCZO. LiDAR data used in this study are available via opentopography.org (SSCZO catchment) and by contacting the JPL Airborne Snow Observatory: aso.jpl.nasa.gov (Tuolumne and Merced catchments).

References

Anderson, S., G. Qinghua, and E. Parrish (2012), *Snow-on and Snow-off Lidar Point Cloud Data and Digital Elevation Models for Study of Topography, Snow, Ecosystems and Environmental Change at Southern Sierra Critical Zone Observatory*, Southern Sierra CZO, Univ. of California at Merced, (digital media), Calif., doi:10.5069/G9GF0RFP.

Bales, R. C., N. P. Molotch, T. H. Painter, M. D. Dettinger, R. Rice, and J. Dozier (2006), Mountain hydrology of the western United States, *Water Resour. Res.*, *42*, W08432, doi:10.1029/2005WR004387.

Bales, R. C., J. W. Hopmans, A. T. O'Geen, M. Meadows, P. C. Hartsough, P. Kirchner, C. T. Hunsaker, and D. Beaudette (2011), Soil moisture response to snowmelt and rainfall in a Sierra Nevada mixed-conifer forest, *Vadose Zone J.*, *10*(3), 786–799, doi:10.2136/vzj2011.0001.

Balk, B., and K. Elder (2000), Combining binary decision tree and geostatistical methods to estimate snow distribution in a mountain watershed, *Water Resour. Res.*, *36*(1), 13–26, doi:10.1029/1999WR900251.

Brubaker, K., A. Rango, and W. Kustas (1996), Incorporating radiation inputs into the snowmelt runoff model, *Hydrol. Processes*, *10*(10), 1329–1343, doi:10.1002/(SICI)1099-1085(199610)10:10 < 1329::AID-HYP464 > 3.0.CO;2-W.

Deems, J. S., T. H. Painter, and D. C. Finnegan (2013), Lidar measurement of snow depth: A review, *J. Glaciol.*, *59*(215), 467–479, doi:10.3189/2013JoG12J154.

Dozier, J. (1989), Spectral signature of alpine snow cover from the landsat thematic mapper, *Remote Sens. Environ.*, *28*, 9–22, doi:10.1016/0034-4257(89)90101-6.

Egli, L., T. Jonas, T. Grünewald, M. Schirmer, and P. Burlando (2012), Dynamics of snow ablation in a small alpine catchment observed by repeated terrestrial laser scans, *Hydrol. Processes*, *26*(10), 1574–1585.

Erickson, T. A., M. W. Williams, and A. Winstral (2005), Persistence of topographic controls on the spatial distribution of snow in rugged mountain terrain, Colorado, United States, *Water Resour. Res.*, *41*, W04014, doi:10.1029/2003WR002973.

Erxleben, J., K. Elder, and R. Davis (2002), Comparison of spatial interpolation methods for estimating snow distribution in the Colorado Rocky Mountains, *Hydrol. Processes*, *16*(18), 3627–3649, doi:10.1002/hyp.1239.

Essery, R., and J. Pomeroy (2004), Implications of spatial distributions of snow mass and melt rate for snow-cover depletion: Theoretical considerations, *Ann. Glaciol.*, *38*(1), 261–265.

Faria, D., J. Pomeroy, and R. Essery (2000), Effect of covariance between ablation and snow water equivalent on depletion of snow-covered area in a forest, *Hydrol. Processes*, *14*, 2683–2695, doi:10.1002/1099-1085(20001030)14:15 < 2683::AID-HYP86 > 3.0.CO;2-N.

Fassnacht, S. R., K. A. Dressler, and R. C. Bales (2003), Snow water equivalent interpolation for the Colorado River Basin from snow telemetry (snotel) data, *Water Resour. Res.*, *39*(8), 1208, doi:10.1029/2002WR001512.

Flanner, M. G., C. S. Zender, P. G. Hess, N. M. Mahowald, T. H. Painter, V. Ramanathan, and P. J. Rasch (2009), Springtime warming and reduced snow cover from carbonaceous particles, *Atmos. Chem. Phys.*, *9*(7), 2481–2497, doi:10.5194/acp-9-2481-2009.

Goulden, M. L., and R. C. Bales (2014), Mountain runoff vulnerability to increased evapotranspiration with vegetation expansion, *Proc. Natl. Acad. Sci. U. S. A.*, *111*(39), 14,071–14,075, doi:10.1073/pnas.1319316111.

Grünewald, T., J. Stötter, J. W. Pomeroy, R. Dadic, I. Moreno Baños, J. Marturià, M. Spross, C. Hopkinson, P. Burlando, and M. Lehning (2013), Statistical modelling of the snow depth distribution in open alpine terrain, *Hydrol. Earth Syst. Sci.*, *17*(8), 3005–3021, doi:10.5194/hess-17-3005-2013.

Guan, B., N. P. Molotch, D. E. Waliser, S. M. Jepsen, T. H. Painter, and J. Dozier (2013), Snow water equivalent in the Sierra Nevada: Blending snow sensor observations with snowmelt model simulations, *Water Resour. Res.*, *49*, 5029–5046, doi:10.1002/wrcr.20387.

Harshburger, B. J., K. S. Humes, V. P. Walden, T. R. Blandford, B. C. Moore, and R. J. Dezzani (2010), Spatial interpolation of snow water equivalency using surface observations and remotely sensed images of snow-covered area, *Hydrol. Processes*, *24*(10), 1285–1295, doi:10.1002/hyp.7590.

Helfricht, K., J. Schöber, K. Schneider, R. Sailer, and M. Kuhn (2014), Interannual persistence of the seasonal snow cover in a glacierized catchment, *J. Glaciol.*, *60*(223), 889–904, doi:10.3189/2014JoG13J197.

Kerkez, B., R. Rice, S. Glaser, R. Bales, and P. Saksa (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, Abstract #IN34A-07 presented at 2010 Fall Meeting, vol. 1, p. 07, AGU, San Francisco, Calif.

Kerkez, B., S. D. Glaser, R. C. Bales, and M. W. Meadows (2012), Design and performance of a wireless sensor network for catchment-scale snow and soil moisture measurements, *Water Resour. Res.*, *48*, W09515, doi:10.1029/2011WR011214.

Khalil, A., M. N. Almasri, M. McKee, and J. J. Kaluarachchi (2005a), Applicability of statistical learning algorithms in groundwater quality modeling, *Water Resour. Res.*, *41*, W05010, doi:10.1029/2004WR003608.

Khalil, A., M. McKee, M. Kemblowski, and T. Asefa (2005b), Sparse Bayesian learning machine for real-time management of reservoir releases, *Water Resour. Res.*, *41*, W11401, doi:10.1029/2004WR003891.

Krause, A., A. Singh, and C. Guestrin (2008), Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies, *J. Mach. Learning Res.*, *9*, 235–284.

Lehning, M., T. Grünewald, and M. Schirmer (2011), Mountain snow distribution governed by an altitudinal gradient and terrain roughness, *Geophys. Res. Lett.*, *38*, L19504, doi:10.1029/2011GL048927.

Link, T., and D. Marks (1999), Distributed simulation of snowcover mass- and energy-balance in the boreal forest, *Hydrol. Processes*, *13*(14), 2439–2452, doi:10.1002/(SICI)1099-1085(199910).

Londhe, S., and S. Charhate (2010), Comparison of data-driven modelling techniques for river flow forecasting, *Hydrol. Sci. J.*, *55*(7), 1163–1174, doi:10.1080/02626667.2010.512867.

Marks, D., J. Dozier, and R. E. Davis (1992), Climate and energy exchange at the snow surface in the alpine region of the Sierra Nevada: Meteorological measurements and monitoring, *Water Resour. Res.*, *28*(11), 3029–3042, doi:10.1029/92WR01483.

McCreight, J., A. Slater, H. Marshall, and B. Rajagopalan (2014), Inference and uncertainty of snow depth spatial distribution at the kilometre scale in the Colorado Rocky Mountains: The effects of sample size, random sampling, predictor quality, and validation procedures, *Hydrol. Processes*, *28*(3), 933–957, doi:10.1002/hyp.9618.

McLachlan, G., and D. Peel (2004), *Finite Mixture Models*, John Wiley, N. Y.

Molotch, N. P., and R. C. Bales (2005), Scaling snow observations from the point to the grid element: Implications for observation network design, *Water Resour. Res.*, *41*, W11421, doi:10.1029/2005WR004229.

Molotch, N. P., and R. C. Bales (2006), SNOTEL representativeness in the Rio Grande headwaters on the basis of physiographics and remotely sensed snow cover persistence, *Hydrol. Processes*, *20*(4), 723–739, doi:10.1002/hyp.6128.

Musselman, K., N. Molotch, and P. Brooks (2008), Effects of vegetation on snow accumulation and ablation in a mid-latitude sub-alpine forest, *Hydrol. Processes*, *22*(15), 2767–2776, doi:10.1002/hyp.7050.

Painter, T. H., J. Dozier, D. A. Roberts, R. E. Davis, and R. O. Green (2003), Retrieval of subpixel snow-covered area and grain size from imaging spectrometer data, *Remote Sens. Environ.*, *85*(1), 64–77, doi:10.1016/S0034-4257(02)00187-6.

Painter, T. H., et al. (2016), The airborne snow observatory: Fusion of scanning lidar, imaging spectrometer, and physically-based modeling for mapping snow water equivalent and snow albedo, *Remote Sens. Environ.*, *184*, 139–152.

Pälli, A., J. C. Kohler, E. Isaksson, J. C. Moore, J. F. Pinglot, V. A. Pohjola, and H. Samuelsson (2002), Spatial and temporal variability of snow accumulation using ground-penetrating radar and ice cores on a svalbard glacier, *J. Glaciol.*, *48*(162), 417–424.

Pedregosa, F., et al. (2011), Scikit-learn: Machine learning in Python, *J. Mach. Learning Res.*, *12*, 2825–2830.

Perkins, T. R., T. C. Pagano, and D. C. Garen (2009), Innovative operational seasonal water supply forecasting technologies, *J. Soil Water Conserv.*, *64*(1), 15A–17A, doi:10.2489/jswc.64.1.15A.

Rango, A., and J. Martinec (1995), Revisting the degree-day method for snowmelt computations, *JAWRA J. Am. Water Resour. Assoc.*, *31*(4), 657–669, doi:10.1111/j.1752-1688.1995.tb03392.x.

Rasmussen, C. E., and C. K. I. Williams (2006), *Gaussian Processes for Machine Learning*, MIT Press, Cambridge, Mass.

Rasouli, K., W. W. Hseih, and A. J. Cannon (2012), Daily streamflow forecasting by machine learning methods with weather and climate inputs, *J. Hydrol.*, *414–415*, 284–293, doi:10.1016/j.jhydrol.2011.10.039.

Rice, R., and R. C. Bales (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*, W03537, doi:10.1029/2008WR007318.

Rosenberg, E. A., A. W. Wood, and A. C. Steinemann (2011), Statistical applications of physically based hydrologic models to seasonal streamflow forecasts, *Water Resour. Res.*, *47*, W00H14, doi:10.1029/2010WR010101.

Rosenthal, W., and J. Dozier (1996), Automated mapping of montane snow cover at subpixel resolution from the Landsat Thematic Mapper, *Water Resour. Res.*, *32*(1), 115–130, doi:10.1029/95WR02718.

Settles, B. (2012), *Active Learning*, Morgan & Claypool, Long Island, N. Y., doi:10.2200/S00429ED1V01Y201207AIM018.

Solomatine, D. P., and D. L. Shrestha (2009), A novel method to estimate model uncertainty using machine learning techniques, *Water Resour. Res.*, *45*, W00B11, doi:10.1029/2008WR006839.

Swetnam, T. L., and D. A. Falk (2014), Application of metabolic scaling theory to reduce error in local maxima tree segmentation from aerial LiDAR, *For. Ecol. Manage.*, *323*, 158–167, doi:10.1016/j.foreco.2014.03.016.

Ticlavilca, A. M., and M. McKee (2010), Multivariate Bayesian regression approach to forecast releases from a system of multiple reservoirs, *Water Resour. Manage.*, *25*(2), 523–543, doi:10.1007/s11269-010-9712-y.

Torres, A. F., W. R. Walker, and M. McKee (2011), Forecasting daily potential evapotranspiration using machine learning and limited climatic data, *Agric. Water Manage.*, *98*(4), 553–562, doi:10.1016/j.agwat.2010.10.012.

Trujillo, E., and M. Lehning (2015), Theoretical analysis of errors when estimating snow distribution through point measurements, *Cryosphere*, *9*(3), 1249–1264, doi:10.5194/tc-9-1249-2015.

Winstral, A., and D. Marks (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, doi:10.1002/hyp.1238.

Wonn, H., and K. O'Hara (2001), Height:diameter ratios and stability relationships for four northern rocky mountain tree species, *Western J. Appl. For.*, *16*(2), 87–94.

Zheng, Z., P. B. Kirchner, and R. C. Bales (2016), Topographic and vegetation effects on snow accumulation in the southern Sierra Nevada: A statistical summary from lidar data, *Cryosphere*, *10*(1), 257–269, doi:10.5194/tc-10-257-2016.