Sensor Placement Strategies for SWE Estimation in the American River Basin

Stephen C. Welch, Department of Civil and Environmental Engineering, UC Berkeley

Branko Kerkez, Department of Civil and Environmental Engineering, UC Berkeley

Roger C. Bales, Sierra Nevada Research Institute, UC Merced

Steven D. Glaser, Department of Civil and Environmental Engineering, UC Berkeley

Karl Rittger, Bren School of Environmental Science & Management, UC Santa Barbara

Robert R. Rice, Sierra Nevada Research Institute, UC Merced

Contact information for corresponding author:

Stephen Welch
stephencwelch@gmail.com
404-310-8361
Abstract

An 11-year dataset of spatially distributed snow water equivalent (SWE) was used to inform a quantitative, near-optimal sensor placement methodology for real-time SWE estimation in the American River basin of California. Rank-based clustering was compared to geographically based clustering (sub-basin delineation) to determine the existence of stationary covariance structures within the overall SWE dataset. The historical SWE data, at 500 x 500 m resolution, were split into eight years of training and three years of validation data. Within each cluster, a quantitative sensor-placement algorithm, based on maximizing the metric of Mutual Information, was implemented and compared to random placement. Gaussian Process models were then built from validation data points selected by the algorithm to evaluate the efficacy of each placement approach. Rank based clusters remained stable inter-annually, suggesting that rankings of pixel-by-pixel SWE exhibit stationary features that can be exploited by a sensor-placement algorithm, yielding a 200 mm average root mean square error (RMSE) for twenty randomly selected sensing locations. This outperformed geographic and basin-wide placement approaches, which generated 460 mm and 290 mm RMSE, respectively. Mutual Information-based sampling provided the best placement strategy, improving RMSE between 0 and 100 mm compared to random placements. Increasing the number of rank-based clusters consistently lowered average RMSE from 400 mm for one cluster to 175 mm for eight clusters, for twenty total sensors placed. To optimize sensor placement we recommend a strategy that couples rank-based clustering with Mutual Information-based sampling design.
**Index terms:** [4894] Instruments, sensors, and techniques; [1972] Sensor web;

**Keywords:** sensor networks, snowpack and hydrologic monitoring, snow water equivalent, water information systems, optimal sensor placement
1. Introduction

Seasonal snowpack and glaciers provide water to more than one sixth of the world’s population and are resources at risk in a changing climate [Barnett et al., 2005]. The mountains of the western United States offer no exception, with snowmelt thought to account for up to 80% of annual streamflow [Daly et al., 2001]. Water managers must rely on a few point estimates of snow water equivalent (SWE), the amount of water contained in the snowpack, in a given mountain basin to inform streamflow forecasting. Accurate estimation of SWE across a basin for forecasting snowmelt runoff, particularly in a changing climate, is impeded by both limited on-ground measurements and high inter-annual variability in snow accumulation and ablation. Snowpack variability is driven by the intersection of climate variability with complex physiographic spatial patterns [Bales et al., 2006].

Though SWE is measured at more than 1700 points across the western U.S. through a combination of automated snow sensors and manual snow courses, these measurements are insufficient to resolve SWE variability at the basin scale [Bales et al., 2006; Rice and Bales, 2010]. Deployed with the intent of informing statistical streamflow predictions, the majority of current snow courses and snow sensors are placed in easily accessible, flat, open, and relatively mid-elevation locations where snow cover is persistent. These physiographically homogeneous locations result in redundant observations that may not reflect basin behavior as a whole. Moreover, it has been shown these sensors are not representative of immediately surrounding areas [Molotch and Bales, 2006], and that the bias between the surrounding terrain and snow sensors are of greater magnitude during the ablation period [Meromy et al., in press]. This strategy
undersamples high elevations where snow can linger into the summer months, and low elevations where melting can occur throughout the winter [Bales et al., 2006].

Improving real-time SWE estimates is a two-part effort: expanding and improving ground-based snow-measurement methods to more accurately capture SWE variability (sensor placement), and distributing these ground measurements to infer SWE at the basin scale (estimation). These two problems are equally important, inherently coupled, and must be solved in tandem.

Ideally, SWE can be estimated at a resolution that reduces the subgrid heterogeneity to a level where the majority of variability can be modeled explicitly [Blöschl, 1999]. Smaller-scale (<100 km²) SWE-distribution approaches, using various geostatistical techniques such as binary regression trees and kriging, have been implemented successfully, but require dense measurements that are typically confined in space and time [Balk and Elder, 2000; Erikson et al., 2005; Erxleben et al., 2002].

Satellite-derived snow-covered area (SCA) measurements aid in distributing ground measurements at the basin scale, and are appealing from a real-time operational hydrology perspective [Painter et al., 2009; Rice et al., 2011]. Remotely sensed SCA products have been used in conjunction with ground-based SWE measurements through various interpolation strategies to build spatial SWE estimates [Fassnacht et al., 2003; Molotch et al., 2004; Harshburger et al., 2009]. These estimates have proven reliable for predicting existing snow stations, but do not create representative estimates of basin-wide SWE distribution due to the inherent bias in ground measurements.

Energy-balance and precipitation-based models have been implemented on the basin scale and shown to compare well to snow-sensor measurements [Shamir and
Reconstruction models provide an energy-balance based retrospective tool for estimating SWE distribution without using ground observations [Cline et al., 1998; Molotch and Margulis, 2009; Rice et al., 2011]. These models use satellite derived melt-out dates and SCA time series to back calculate SWE using an energy balance for each grid element. While this method does provide a retrospective, unbiased, and independent estimate of SWE historical distribution, it inherently cannot be used to make real-time SWE estimates.

Sensor placement falls under the larger umbrella of sampling design, a robust and growing field within statistics and computer science that has seen limited application to snow hydrology. Expanding the current snowpack sensing infrastructure requires the determination of how many sensors (or sampling locations) are required and where they should be placed within the basin to maximize the amount of captured information. Molotch and Bales [2005] and Rice and Bales [2010] searched for optimal sensor placements that capture the mean and variability of SWE across a larger grid element, but did not evaluate potential resulting placement error or the amount of information captured by would-be placements. While it is generally acknowledged that improved strategic sampling is needed to build better large-scale sensor deployments and SWE estimates [Bales et al., 2011], no quantitative methods or performance metrics have been proposed.

The aim of this research is to use a multi-year historical reconstruction of SWE for the American River basin in California's Sierra Nevada to identify sensing locations for accurate, real-time estimation of basin-wide SWE. This research addresses three questions. First, what properties of basin-wide SWE distribution are temporally stationary? Second, how can we use these properties to inform sensor placement? Finally,
can we develop a quantitative sampling method to generate sensor placements based on historical SWE observations? A more general aim is to develop a ground-based sampling strategy that will become a core element of a larger water-information system for the Sierra Nevada and other mountain basins.

2. Background

2.1 Sampling strategies. Given a large set of observable locations, selection of the “most informative” subset is a problem faced within numerous disciplines. Selecting ideal sensing locations for observation of spatial phenomena, in our case SWE, is a direct instance of this problem. One sensor-placement approach is to assume that sensors exhibit disk-shaped sensing regions, and to search for a set of sampling locations that provides ideal spatial coverage; this approach is formalized as the “art gallery problem” in the computer science literature [Gonzalez-Banos and Latombe, 2001]. Within snow hydrology, the concept of circular impact areas has been applied to estimate unobserved locations through the use of inverse distance weighting (IDW), optimal distance averaging (ODA), and kriging [Fassnacht et al., 2003; Carroll, 1995]. These techniques are often combined with methods such as linear regression against elevation to derive more robust estimates of SWE [Carroll and Cressie, 1996]. While these studies have validated the efficacy of this approach for estimation purposes, these principles prove difficult to apply in sampling design, and risk over or under sampling snowpack behavior.

It is well documented that SWE and snow depth vary with physiographic variables, such as elevation, aspect, slope, and canopy cover [e.g., Fassnacht et al., 2003]. A reasonable sampling strategy would be to “evenly sample” across each predictor.
variable. Such strategies have been described in the computer science literature (see \textit{Latin hypercube sampling} [Mackay et al., 1979]). Unfortunately, it is difficult to quantify the effectiveness of such a placement scheme since evenly sampling physiographic features may not adequately capture the relative importance of any given feature.

An alternative approach is to couple sensor placement with modeling. This involves developing a statistical model of a phenomenon based on previous observations, and then designing a placement scheme that improves model performance. The statistics community has addressed the question of maximizing the quality of parameter estimates in linear models through \textit{sampling design} [e.g., Atkinson, 1988]. Algorithms have also been developed to select sampling locations to yield the best possible “error reduction” in linear models [Das and Kempe, 2008].

A specific modeled-based approach, developed by Cressie [1991], involves developing a \textit{Gaussian Process} (GP) model for an underlying spatial phenomena, either from expert knowledge or a pilot deployment. Use of this inference framework assumes underlying Gaussian distributed data. The power of this Bayesian technique stems from its ability to provide variance estimates for predictions made at un-instrumented locations, allowing us to frame a sensor placement approach by selecting those locations that most reduce model uncertainty across the estimated field.

Given a set of sparse spatial observations, a GP model can be used to make predictions at un-instrumented locations. Assuming that our sampling space is comprised of a finite set $\mathcal{V}$ of random variables $x$, and we have made observations at a subset of points $\mathcal{A} \subset \mathcal{V}$, it is possible to make predictions at the remaining set at the un-instrumented locations $\mathcal{Y} \subset \mathcal{V}$ through:
\[
\mu_{y|\mathcal{A}} = \mu_y + \sum_{y\in\mathcal{A}} \sum_{\mathcal{A}\setminus\mathcal{A}}^{-1} (x_{\mathcal{A}} - \mu_y),
\]

\[
\sigma_{y|\mathcal{A}}^2 = \sum_{yy} - \sum_{y\in\mathcal{A}} \sum_{\mathcal{A}\setminus\mathcal{A}}^{-1} \sum_{\mathcal{A}y},
\]

where \(\mu_{y|\mathcal{A}}\) and \(\sigma_{y|\mathcal{A}}^2\) are the conditional mean and variance of the estimates at the un-instrumented locations, and \(\sum_{mn}\) is the covariance matrix of two sets of vectors (e.g. \(\sum_{y\in\mathcal{A}}\) is the covariance of the un-instrumented locations \(y\) and instrumented locations \(\mathcal{A}\), and \(\sum_{\mathcal{A}\setminus\mathcal{A}}^{-1}\) is the inverse of covariance matrix of the instrumented locations \(\mathcal{A}\)) [Rasmussen and Williams, 2006].

Quantitatively, the GP framework allows for a number of placement strategies, and has been used in conjunction with Information-theoretic measures [Kemppainen et al., 2008; Guestrin et al., 2005]. Information Theory is a mathematical framework that can be used to quantify the amount of uncertainty that can be informed within a process [Shannon, 1948]. A direct measure of problem uncertainty is given by Entropy [Pierce, 1980]. In our sensor-placement approach we seek to find a means by which to quantify and inform the uncertainty of making SWE predictions at the un-instrumented locations \(y \subset \mathcal{V}\) given observations of SWE at locations \(\mathcal{A} \subset \mathcal{V}\).

The conditional entropy of a Gaussian variable \(y\) given another Gaussian variable \(\mathcal{A}\) is [Guestrin et al., 2005]:

\[
H(y|\mathcal{A}) = \frac{1}{2} \log(2\pi e \sigma_{y|\mathcal{A}}^2)
\]

The above expression estimates the entropy, or uncertainty, \(H\), at a set of un-instrumented locations \(y\), given a set of selected sensing locations \(\mathcal{A}\). Entropy-based placement
methods typically seek to maximize the entropy, $H(\mathcal{A})$, providing more opportunities for new information, i.e. adding sensors in regions of the highest model uncertainty. This approach results in the set of sensor locations that are most uninformed about each other [Kemppainen et al., 2008; Lee and Queyranne 1995]. A direct extension of entropy-based methods use Mutual Information ($MI$), another information-theoretic quantity, which when maximized seeks to reduce the error of estimates for unobserved locations by selecting the placement $\mathcal{A}$ that is most informative about unobserved locations $y$ [Guestrin et al., 2005; Krause et al., 2007]:

$$\arg\max_{\mathcal{A} \subseteq V} MI(y) = \arg\max_{\mathcal{A} \subseteq V} H(y) - H(y|\mathcal{A})$$ (4)

In general, model-based sampling design presents a difficult computational problem, leading to the use of greedy sensor-placement algorithms. Rather than finding the optimal set of sensor locations, greedy algorithms place sensors one at a time, with each step incrementally maximizing the above expression. Although greedy algorithms cannot guarantee a truly optimal solution, they are considered near-optimal, and provide performance guarantees, ensuring that algorithm performance in no worse than some fraction of the optimal solution [Guestrin et al., 2005; Krause et al., 2007].

2.2 Regionalization (clustering) of data. In certain conditions the parameterization of a single statistical model may not be sufficient to explain the variability of an entire data set. In the case of SWE, it may be necessary to use multiple models, each of which describes a subset of the observed data. Effective grouping of data into homogenous subsets has been shown to improve model performance [Chipman et al., 2002; Hannah and Dunson, 2011], and is a common pre-analysis step within hydrologic investigations.
[Serreze et al., 1999; Clark et al., 2001]. Observations are often grouped by geographic proximity [Moore and McKendry, 1996], but more-complex statistical methods have also been applied to snow data such as principal component analysis [McGinnis, 1997; Cayan, 1996], and self-organizing maps [Fassnacht et al., 2010].

Wu and Lui [2012] searched for statistically similar regions within soil-moisture data using a coarse-grained ordering approach. This method assumes that although overall moisture levels will vary significantly throughout space and time, certain locations will typically be drier or wetter than others. This phenomenon is measured by sorting from least to greatest, by wetness, all locations at each time step and analyzing the stability of this ordering over time. Such ordering is ultimately used to divide the study area in homogeneous regions, or clusters, roughly corresponding to wetter and drier areas that can be explained by separate statistical models.

3. Methods

An eleven-year spatially distributed SWE data set for the American River basin, with a grid size of 500 x 500 m, was used to determine fixed measurement locations for estimating SWE during snowmelt under a range of inter-annual climate conditions. While SWE variability at the sub-500m-scale is an important part of an overall SWE estimation framework – the work presented here assumes homogenous behavior at the 500m scale. Scaling small-scale snowpack variability to the larger pixel scale is discussed in Kerkez et al. [2012]. The 11 years of data were then divided into training and testing sets, the first eight years (2001-2008) used to train the sensor-placement algorithm and to develop a historical SWE covariance matrix, and the next three years (2009 -2011) used for evaluation of the placement algorithms.
3.1 Study area. The American River flows westward from the crest of the Sierra Nevada, with its three forks emptying into the Folsom Reservoir. Basin elevations range from 200 m at Folsom reservoir to over 3000 m in the Desolation Wilderness of the Eldorado National Forest. Snowfall typically is greatest and persists the longest at higher elevations. Of the 4741 km² above Folsom, 2155 km² is above 1500 m, and 242 km² is above 2400 m. Pixels with mean elevations below 1500 m were not used in the analysis, as these elevations do not significantly contribute to overall SWE in the American River basin (Figure 1). The snow-covered portion of the American River Basin extends from oak-pine forest at the lower snowline, through mixed conifer forest, through red fir forest to subalpine above treeline. The mixed-conifer forest extends over a broad elevation band, and includes a heterogeneous mix of tree heights and canopy densities. Species in the mixed conifer are white fir, ponderosa pine, sugar pine and incense cedar, plus an array of shrubs under the canopy and in disturbed areas. There are 12 snow pillows and 26 snow courses in the basin. While the methods presented in this paper can be applied across many regions, the American River basin was chosen for the purposes of this study because it is representative of a number of main rivers draining the west slope of the Sierra Nevada. Further, multiple ecosystem services in the American River basin depend on accurate and timely hydrologic information and forecasts.

3.2 Data. The spatial SWE dataset used in this analysis was developed using a pixel-by-pixel energy-balance reconstruction of snowmelt, summed back in time from when snowcover disappeared to the beginning of seasonal snowmelt. Methods are summarized below, with details given in Rittger et al. [2011] and Rittger [2012]. The reconstruction method combines the satellite-derived snow-cover-depletion
record of the melt to retrospectively estimate how much snow had existed at every pixel. The technique has been validated in the Sierra Nevada [Cline et al., 1998; Rittger et al., 2011] and applied to large basins at multiple scales [Molotch and Margulis, 2008; Rice et al., 2011]. SWE was estimated using a snowmelt model that combines both energy-balance and temperature-index methods [Brubaker et al., 1996]. Air temperature and incoming solar and longwave radiation were downscaled from 1/8° NLDAS2 reanalysis data [Cosgrove et al., 2003] to adjust for elevation, topography and vegetation [Dubayah, 1992; Dozier and Frew, 1990; Garen and Marks, 2005]. Reflected solar radiation was estimated using MODSCAG albedo [Painter et al., 2009] and outgoing longwave radiation was estimated using the Stefan-Boltzmann equation assuming the snow temperature was the minimum of air temperature and 0° C.

Fractional SCA used to scale the potential melt estimate for each pixel was derived from the MODSCAG model [Rittger et al., 2012; Dozier et al., 2008]. Validation of maximum SWE was performed in a 3x3 grid cell region around each snow pillow and snow course by calculating the mean error in the nine cells and the best cell within that region. For snow pillows the error RMSE was 242 mm and 186 mm with a mean difference of 19 mm and 3 mm. For snow courses the RMSE was 296 mm and 230 mm with a mean difference of 77 mm and 18 mm. Snow pillows are considered a better validation because they measure snow daily or hourly, while snow courses only measure SWE on the first of each month. The MODIS satellite became operational in the year 2000, limiting the present work to 11 years of reconstructed data. The reconstruction product is shown in Figure 2 for the American River Basin through the 2006 melt season. It should be noted that although the year 2006 was used here for a number of
visualizations, all modeling and sensor placement efforts utilized the complete
reconstructed SWE data set. 2006, known to be particularly wet year, was simply chosen
for visualization purposes to show the variability of the melting snowpack throughout the
melt season. Prior to modeling, SWE training data were standardized each day by
subtracting the spatial mean and dividing by the standard deviation. Since mean SWE
varied significantly each year, this standardization step was necessary to extract
stationary features of SWE variability. Historical monthly snow-course and daily snow-
sensor data were obtained from the California Data Exchange Center (http://cdec.water.ca.gov).

3.3 Clustering analysis. Multiple clustering methods were evaluated to investigate the
existence of stationary subsets that could be used to inform sensor placement within the
American River SWE data. The first and simplest method, called global clustering,
treated the entire American River basin as one cluster on which to conduct sensor
placement. The second method, geographic clustering, intended to capture smaller-scale
SWE variability by delineating the larger basin into six third-order sub basins. The third
method, rank clustering, clustered similar regions in space through the analysis of
historical behavior of each SWE pixel. Daily SWE data were ordered from least to
greatest, resulting in a “rank” for each pixel for every day studied. The sum of ranks was
taken for all days studied, and pixels were then clustered based on sum rank values using
a K-means algorithm as the number of clusters used was varied between two and eight.
K-means is a simple clustering algorithm that minimizes the Euclidean distance between
each cluster’s centroid in K-dimensional space and the data within that cluster [Mackay,
2003]. The K-means algorithm was used here in a single dimension as a simple tool to
ensure that the most-similarly ranked pixels were assigned to the same clusters. Rank stability was evaluated by plotting histograms of pixel rank [Wu and Lui, 2012].

Although ranking were not necessarily stable throughout the year, rankings during melt seasons were relatively stable (Figure 3). It is difficult to establish a consistent “melt-season” definition year to year; during our analysis simply using April 15- June 1 produced the stability needed to carry out our analyses.

### 3.5 Dimensionality reduction

Both GP modeling, and a number of other placement schemes, rely heavily on the covariance matrix of the input data. Performing the needed operations with covariance matrices requires the matrix to be well conditioned and close to full rank. Poorly conditioned or rank-deficient covariance matrices can result from two or more pixels having SWE time series that are too similar. If all the information known about one pixel is very close to that of another, it becomes computationally impossible to store the difference within a covariance matrix that must describe the entire space. This results in covariance matrices that require computationally impossible numerical precision to be operated on. A solution to this problem is to reduce the dimension of the input data – group highly statistically similar pixels together, and use representative pixels from each independent group in modeling and placement algorithms and then projected back onto the full space for final predictions. Dimensionally reduced data are at the core of our placement and modeling algorithms, and then projected back onto the full space for final predictions. A K-means algorithm was applied to all the clusters resulting from the previous section, isolating sufficiently independent time series, from which a historical covariance matrix could then be constructed. Unlike the K-means approach that was used above to build rank based clusters, K-means was used here in a higher-
dimensional space to find and group pixels with the most similar time series. While we are analyzing approximately 7000 pixels, the nature of the input data allows us to effectively operate on 50-300 representative pixels per cluster. An important implication here is that the nature of the underlying dataset ultimately determines the precision with which sensors can be placed.

3.6 Sensor Placement. Sensor-placement algorithms were applied within clusters, and constructed under the premise of maximizing information drawn from historical data. The approach was evaluated by placing sensors using historical training data for 2001-2008, and evaluating the placements on data sets in years 2009-2011. Years 2009-2011 were used strictly for validation of the placement approach, and were not in any way used to inform the placements or estimations. Two sensor-placement schemes were compared: a Mutual Information (MI) based placement and random placements chosen from the dimensionally reduced subset described in the previous section. A MATLAB toolbox for Submodular Function Optimization was used to carry out the optimization problem for the MI-based placement [Krause et al., 2010].

In the current analysis, “sensor” refers to an instrumented 500 × 500 m pixel, and it is assumed that the average SWE in that area is accurately measured by the sensor. In practice, a “sensor” of this scale consists of 10-20 individual, fixed snow-measurement nodes that strategically sample within the pixel and are the resulting data combined to give a pixel-scale SWE value [Rice et al., 2011; Bales et al., 2011; Kerkez et al., 2012]. When comparing global placement to cluster-based placements, one must decide in what order to add sensors to clusters. Here, sensors were simply added numerically by cluster number (e.g. for 5 clusters and 12 total sensors placed, clusters 1 and 2 each have three
sensors, and clusters 3 through 5 each have two sensors).

3.7 Modeling. To evaluate potential placements, Gaussian Process models were used to make predictions of SWE based on historical observations as well as observed locations selected for each placement scheme. The 2001-2008 data were used to build covariance matrices, effectively “training” the GP model. The covariance matrix was then used in conjunction with the locations selected by each sensor placement to derive predictions for years 2009-2011 at un-instrumented locations using equation (1). Predictions were then back-standardized using the sample mean and variance of observed locations from the training set. This constraint meant that it was necessary to place at least two sensors within each cluster in order to calculate a sample standard deviation.

3.8 Evaluation. The root mean square error (RMSE) between the original SWE product (years 2009-2011) and model output was used to evaluate the efficacy of a given placement. RMSE was used to compare placement approaches, the effect of adding more sensors, as well as the effect of varying the number of clusters. To evaluate placements across multiple time steps, average RMSE was calculated as:

\[
AvgRMSE = \frac{1}{J} \sum_{j=1}^{J} \frac{1}{T} \sum_{t=1}^{T} \sqrt{\frac{\sum_{i=1}^{n} (\hat{x}_i - x_i)^2}{n}}
\]  

(5)

Where \( J \) is the number of trails conducted, \( T \) is the number of time slices used, \( n \) is the number of pixels analyzed, \( \hat{x} \) is estimated SWE, and \( x \) is actual SWE. This formulation of RMSE permitted cumulative evaluation of RMSE across three years. RMSE was only calculated for pixels with actual SWE above 10 mm, using the assumption that nearly snow-free pixels can be identified in real time using satellite SCA products. For random
placements, ten trials were conducted and mean RMSE was computed ($J = 10$), while MI-based placements required RMSE only to be computed once ($J = 1$).

4. Results

4.1 Rank-cluster analysis. Rank-based clustering behavior was consistent throughout the eight-year analysis period. Results from the 2006 melt season are shown in Figure 4, which were comparable to those seen in other years. Rank-based clustering resulted in clusters that roughly followed basin elevation bands (Figure 4a). Higher variability was typically observed in higher-elevation clusters (Figure 4b). On April 1st, 2006 (shown here to be representative of the beginning of melt) standard deviations varied between approximately 400 mm for the highest SWE cluster (cluster 5, 2100 mm peak SWE) to 150 mm for the lowest SWE cluster (cluster 1, 700 mm peak SWE). In the 2006 melt season, each cluster exhibited a noticeably unique melt-out date ranging over a two-month span from late April to late June. Rank-based clusters divided the bimodal SWE distribution of the whole basin (April 15, 2006) into near-Gaussian distributions of various spreads (Figure 4c). Rank-based clusters divided physiographic predictor variables primarily by elevation (Figure 5), generating near-Gaussian distributions of elevation, with the exception of the lowest elevation-cluster, which was cut off at 1500 m as described in the methods section. Vegetation cover is normalized on a zero to ten scale, with zero indicating no vegetation and ten indicating full vegetation cover. No notable correlation was seen between elevation and aspect within clusters. When the tails of the elevation distributions for each cluster were examined for relationships to physiographic variables (possibly explaining why two pixels of the same elevation were grouped in different clusters), only latitude and longitude demonstrated notable
Inter-annual differences between rank-based clusters ranged from 22 to 41%, with a median value of 31% (Table 1). This is a measure of stationarity that evaluates if a pixel retained its cluster assignment between these two periods. Differences were larger when comparing results for 15-day periods within one year, with values of 0.13-0.67 and a median near 0.5 (Table 2). Cluster assignments varied significantly when comparing different periods during a single year, but agreed when comparing multiple years to each other.

**4.2 Modeling Results.** Analysis of sensor-placement methods showed that average RMSE over the three years studied (2009-2011) decreased with the number of sensors placed (Figure 6). Diminishing returns occurred between 20 and 30 sensors placed, after which placing further sensors, or taking additional samples, yielded little improvement in RMSE. The global MI-based placement algorithm failed to converge for more than 24 sensors placed.

Sensor placement within a rank-based clustering approach outperformed geographic clustering and global placement. Within each clustering method, MI based sensor placements generally offered improvements in RMSE when compared to random placements within the same clusters. MI-based placements within geographic clusters offered an RMSE improvement between 0 and 150 mm compared to a random approach, while MI-based placements within rank-based clusters yielded between 30 mm greater and 30 mm lower RMSE than random placements.

To investigate temporal RMSE behavior, and taking the 2009 melt season as a
typical year, daily RMSE for rank-based and geographic clustered placements showed a
gen-er-al decrease approaching melt-out of the snowpack. Global placement showed
unstable and high RMSE values (Figure 7). Rank-based cluster RMSE was consistently
below 200 mm, while geo clustering had RMSE values below 500 mm. Rank-based
clustering, along with MI-based sensor placement showed the best RMSE performance
throughout the melt season.

Given the performance of a rank-based clustering approach, an analysis was
carried out to determine the effect on RMSE when varying the number of clusters while
keeping the number of sensors per cluster steady. RMSE generally decreased as more
rank-based clusters were used in the placement (Figure 8), where the use of one cluster is
equivalent to a global clustering approach. RMSE decreased from a range of 400-600
mm in the case of global placement, to less than 200 mm when using eight clusters.
Diminishing returns in RMSE became apparent in the three to six cluster range, after
which further rank-based clustering offered no significant improvement in RMSE. MI-
based placement offered slight improvements over random placement.

The final proposed sensing regions resulting from coupling rank-based clustering
with MI-based sensor placement are shown in Figure 9. As discussed, the nature of the
input data limits the precision of the covariance structure. Once optimal locations were
selected using the placement algorithm, they were back projected from the dimensionally
reduced data set into the original, higher-dimensional data set. Rather than calculating
single optimal locations, the algorithm produces optimal sensing areas where each pixel
is considered equivalent. In figure 9 each round corresponds to placing one sensor within
each cluster; all locations indicated in a single round are equally valuable.
5. Discussion

5.1 Temporal stability through rank based clustering. A major goal of this analysis was to seek out stationary metrics to inform sensor placement. As seen through the reconstructed data, SWE is non-stationary with regard to annual mean and variance, making it difficult to infer one statistical model to cover the full range of annual spatiotemporal fluctuations. Some level of stationarity does however need to be extracted to make use of historical observations for sensor placement. It is well known that SWE is correlated to a number of stationary physiographic variables such as elevation, slope, aspect, and vegetation [Fassnacht et al., 2003]. However, this does not mean that the correlation between SWE and its physiographic predictors is stationary. Perhaps the most common correlation metric used when studying SWE, the slope of the regression line taken between SWE and elevation, is shown in Figure 10. While this relationship may be robust over a single time step, it is temporally non-stationary, both inter- and intra-melt season. This non-stationarity impedes the use of these model parameters for purposes of sensor placement.

Rather than finding absolute trends in the data, rank-based clustering isolates subsets of that data that exhibit relatively similar behavior through time. Tables 1 and 2 show that rank-based clustering consistently isolated common clusters between years, with worst-case inter-annual cluster differences not exceeding 41%, while other measures of basin-wide SWE, such as regression against elevation (Figure 10), vary more than one order of magnitude inter-annually. Although SWE itself is quite variable, certain locations have consistently higher or lower SWE than others – this phenomenon is well exploited by rank-based clustering. This provided a strong starting point for sensor-
placement efforts. While Table 2 does show higher intra-annual differences, these can largely be attributed to the melt out of pixels toward the end of a melt season: 48% of pixels with snow during the April 1-15 period had no snow by June 1. Implementing cluster analysis on a multi-year data set maximizes the use of all historical data while providing a balance between long-term trends and inter-seasonal variability. In an effort to measure the representativeness of a multi-year cluster analysis, a sensitivity analysis was conducted to determine the impact of clusters on final optimal placements. Ideal placements were computed for each year and compared to the eight-year aggregate. Results ranged from 28% to 38% overlap in ideal sensor-placement regions between the eight-year aggregate placements and individual yearly placements.

The results shown in Figure 5 indicate that rank-based clusters divide primarily on elevation, a reasonable result considering the strong correlation broadly seen within the literature [Fassnacht et al., 2003, Bales et al., 2006]. Of interest here is that elevation is not simply binned into bands, but divided into overlapping near-Gaussian distributions. A similar result is observed when examining SWE histograms for each cluster, which also exhibit near-Gaussian distributions (Figure 4). This trend may have strong implications for sensor placement, and could partially explain why a random sensor-placement approach performed well. Examining the tails of each cluster’s elevation distribution showed dependence on latitude and longitude, but no other physiographic variables. Dependence of cluster elevation tails on latitude and longitude was expected however, as latitude and longitude are correlated to elevation in the American River basin. The lack of correlation with aspect, slope and vegetation does not rule out combinations of physiographic variables influencing the clustering below the 500-m pixel scale however.
It has been established from snow surveys that these physiographic variables influence SWE at the 500-m pixel scale, and are important determinants of the number and location of samples required within a pixel [Molotch and Bales, 2005; Rice and Bales, 2011].

5.2 Placement strategy within clusters. Dividing the sampling space using rank-based clustering is the first step of the placement approach described here. Within the derived clusters, further steps were taken to identify sensors placements that best inform future predictions. Dimensionality reduction, a computationally necessary step, is inherently part of the placement process. Typical clusters (1000-7000 pixels) were reduced to approximately 50 representative pixels to form computationally workable covariance matrices. It is within this reduced space that MI and random placement methods were carried out. Allowing selection exclusively from this subset means that we are already guaranteed reasonably independent observations, regardless of how placements are chosen in the reduced space. In practice, working in this reduced (and standardized) space limits the potential benefits of a MI placement approach, explaining the inconsistent RMSE improvements generated by MI placement over random placement seen in figures 6, 7, and 8. Similar results were observed by Wu and Lui [2012] on soil-moisture data.

While MI-based placement did not always necessarily outperform random placement, the approach does offer other benefits. Perhaps the most significant is a performance guarantee. The RMSE results of random-sampling approaches were the mean of many trials, spanning a range of good and poor placements. It may be entirely feasible that a real-world random-sampling approach could select one such poor placement. On the other hand, mutual-information-based placements are unique for a
given covariance matrix, and will always select the same sensor locations in the same order, consistently guaranteeing the indicated performance. A second benefit of the MI-based sensor-placement approach is the ability to start with existing snow-pillow and snow-course sites, and then select additional sites through the placement algorithm. This should further improve estimation performance by maximizing the utility of data obtained from existing sites. The validation of the proposed algorithms used a historical covariance structure to predict SWE at future years based on observations made through relatively sparse sensing locations. Given the statistical nature of the modeling approach, it is feasible to update the covariance structure through real-time observations, and to add more sensors as needed. Given reported RMSE performance and consistent near-optimal placement guarantees, we recommend an MI-based approach.

In some cases the RMSE slightly increased though the addition of more sensors (Figure 6). These perturbations can be explained through the real-world nature of the data set. RMSE is not a monotonic function, and the validation data set may contain noise or outliers. Existence of outliers or data-set noise can adversely affect the RMSE behavior.

5.4 Number of Clusters. As shown in Figure 8, three to eight clusters should be sufficient to model SWE and conduct sensor placement within the American River basin. Modeling with more clusters does not improve RMSE performance significantly. Furthermore, from the perspective of good statistical practice, it is desirable to reduced model complexity to avoid the possibility of over fitting. With the current historical data sets we thus propose that three to five rank-based clusters are an optimal choice for the American River. The number of available sensor should then be divided equally throughout each cluster, or added proportionally to the variance of clusters, as done by
5.5 Overall Sampling Approach. Based on the RMSE performance of the placement approaches implemented here, we are confident that one pixel in each of the regions shown in figure 9 should be instrumented to derive better spatial estimates of SWE. In practice, specific points can be chosen within each proposed area based on accessibility or site preference. One such ideal placement resulting from our proposed method is shown in Figure 1.

The method proposed in this paper is designed to select sensing locations across a large area (over 1000 km²). Each sensor location was treated as a 500 × 500 m pixel. It was assumed that instrumenting a pixel provides a noise-free representative measurement for that area. The locations selected should be thought of as strategically placed sets of individual nodes, with the set of nodes together forming the spatial sensor. Realistically, SWE is known to vary significantly at small scales and it is unlikely that a single measurement within a 500-m cell will accurately represent mean SWE of the area [Molotch and Bales, 2005; Rice and Bales, 2010]. Multiple sensors (low cost snow-depth sensors coupled with pillow measurements) will be required to reliably instrument one pixel. Sensor-placement strategies for smaller-scale areas thus need to be considered as part of a basin-scale design. Thus the overall strategy involves a hierarchical placement: i) optimal locations at the large scale (more than one square kilometer) should be selected based on the methods presented in this paper, and ii) estimates of mean SWE at each one of these locations should be derived though a secondary placement scheme at the smaller scale to capture local variability in aspect, vegetation, and energy balance [Kerkez et al., 2012]. Each of the locations selected in this paper can then be thought of as a spatial
sensor made up of a set of individual sensor nodes. We are confident that this presents the most realistically manageable sensor placement approach when considering the future instrumentation of large basins such as that of the American River. The blending of these strategically placed on-the-ground measurements with broad-coverage satellite and aircraft measurements will offer unprecedented estimates of snowpack, soil moisture, vegetation state and energy balance, and snowmelt [Rice et al., 2009].

5.6. Implications for operational hydrology. The design approach described above can be used to both evaluate the information potential of existing operational snow-measurement sites as well as serve as a tool for upgrading an existing system to provide more spatially representative snowcover information. That is, this approach can indicate where, across a larger watershed, new measurements should be taken given an already existing infrastructure, and which existing snow-course, snow-pillow or meteorological-station sites are should be considered for upgrading. Note that upgrading a local site involves adding sensor nodes to obtain a spatial estimate at the 1-4 km² scale [Rice and Bales, 2011; Kerkez et al. 2012]. This two-part strategy is currently being implemented in the American River Basin, and will provide a full-scale test of the design approach. This enhanced ground-based system is also an essential component of operational systems under consideration that blend ground data with satellite snowcover information [Bales et al., 2008; Rice et al., 2011] and with aircraft LiDAR snow-depth information [Kerkez et al., 2012; Kirchner et al., 2012].

6. Conclusions

The approach in this paper provides a sampling strategy that uses historical data to select sensing locations that will be most informative for future real-time estimation of
SWE in the American River basin. While the spatio-temporal SWE patterns are a non-stationary process, a rank-based clustering approach derives a set of regions that remain relatively stable over time. Following standardization of the annual SWE data, the variability of SWE within rank-based clusters was explained through a stationary covariance structure. This covariance structure was used to inform a quantitative placement scheme, by greedily placing sensors to maximize Mutual Information within each cluster. By splitting the existing data set into a training and test subsets, our approach reduces the RMSE of the estimation as a function of the number of sensors placed and outperforms non-rank-based clustering approaches. It is shown that there is a point of marginal returns, after which placing more sensors does not improve estimation performance significantly. We recommend the use of rank-based clustering, coupled with placement based on MI, as the near-optimal choice to determine the number of required sensors, and their respective locations. The sampling strategy presented here is generalizable to other basins, given the availability of reconstructed SWE datasets. Further, the non-stationary behavior of SWE is indicative of hydrologic variables, suggesting the value of a similar approach applied elsewhere, such in Wu and Lui [2012]. Our future work will focus on deploying sensor-network clusters in the locations proposed in this paper, and will evaluate the efficacy of these networks to estimate SWE in the American River basin.

Acknowledgements

This research was supported by the National Science Foundation, through the Southern Sierra Critical Zone Observatory (EAR-0725097) and a Major Research Instrumentation grant (EAR-1126887), California Department of Water Resources (Agreement
460008548), and the University of California Center for Information Technology Research in the Interest of Society. We also acknowledge the cooperation of M. Anderson and F. Gehrke, California Department of Water Resources; J. Dozier, UC Santa Barbara; and X. Meng, UC Merced.
References


Kirchner, P.B., R. C. Bales, K. N. Musselman, N. P. Molotch (2012), Under-canopy snow accumulation and ablation measured with airborne scanning LiDAR altimetry and in-situ instrumental measurements, southern Sierra Nevada, California, *Fall meeting abstracts (C14B-08)*, American Geophysical Union, San Francisco.


doi:10.1029/95WR03640.


### Tables

**Table 1. Inter-annual cluster difference, fraction of pixels in a different cluster between years**

<table>
<thead>
<tr>
<th></th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>0.302</td>
<td>0.315</td>
<td>0.397</td>
<td>0.340</td>
<td>0.342</td>
<td>0.336</td>
<td>0.312</td>
<td>0.316</td>
<td>0.363</td>
</tr>
<tr>
<td>2002</td>
<td>0.253</td>
<td>0.269</td>
<td>0.300</td>
<td>0.242</td>
<td>0.313</td>
<td>0.296</td>
<td>0.328</td>
<td>0.278</td>
<td>0.278</td>
</tr>
<tr>
<td>2003</td>
<td>0.290</td>
<td>0.223</td>
<td>0.250</td>
<td>0.287</td>
<td>0.316</td>
<td>0.304</td>
<td>0.291</td>
<td>0.408</td>
<td>0.408</td>
</tr>
<tr>
<td>2004</td>
<td>0.308</td>
<td>0.344</td>
<td>0.339</td>
<td>0.366</td>
<td>0.329</td>
<td>0.408</td>
<td>0.366</td>
<td>0.329</td>
<td>0.408</td>
</tr>
<tr>
<td>Dry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wet</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 2. Intra-melt-season cluster difference, fraction of pixels in a different cluster between periods**

<table>
<thead>
<tr>
<th></th>
<th>Apr 15-30</th>
<th>May 1-15</th>
<th>May 16-31</th>
<th>June 1-15</th>
<th>Apr 15</th>
<th>Apr 16-30</th>
<th>May 1-15</th>
<th>May 16-31</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>0.322</td>
<td>0.548</td>
<td>0.634</td>
<td>0.671</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>0.443</td>
<td>0.551</td>
<td>0.592</td>
<td>0.281</td>
<td>0.196</td>
<td>0.592</td>
<td>0.281</td>
<td>0.196</td>
</tr>
<tr>
<td>2003</td>
<td>0.132</td>
<td>0.132</td>
<td>0.132</td>
<td>0.132</td>
<td>0.132</td>
<td>0.132</td>
<td>0.132</td>
<td>0.132</td>
</tr>
</tbody>
</table>
List of Figures

**Figure 1.** The American River Basin above 1500 m elevation, divided into approximately 7000 500-m square pixels. Snow-pillow and snow-course locations from California Cooperative Snow Survey. Optimal sensing locations selected through the method proposed in this paper are shown as green squares.

**Figure 2.** (a) Snow course, snow pillow, and interpolation product over 2006 melt season (b) Time slices of interpolation product shown for May 1, May 15, and June 1.

**Figure 3.** (a) One pixel chosen randomly from each 300 meter elevation band shown in Figure 1 (b) time series SWE for each pixel (c) Ranks across melt season.

**Figure 4.** (a) Five rank based clusters (b) Cluster mean and standard deviation over 2006 season (c) Histograms of SWE by cluster for April 15.

**Figure 5.** Histograms of physiographic variables divided by rank-based cluster.

**Figure 6.** Average RMSE vs. sensors placed, averaged over 2009-2011 melt seasons. Rank-based results shown for 5 clusters.

**Figure 7.** (a) 5, 50, and 95th percentiles of SWE across whole basin (b) RMSE over 2009 melt season. Rank-based cluster results shown for five clusters.

**Figure 8.** Average RMSE vs. number of rank based clusters. At least two sensors are required per cluster.

**Figure 9.** Recommended regions for placements in three rank-based clusters using Mutual Information based placement algorithm. Round one is the most informative placement, round two the second most and so on. One sensor should be placed within a given region each round. All proposed locations within a given region are equally valuable.

**Figure 10.** Boxplots of regression parameters for relationship between elevation and SWE computed daily across each melt season for American River Basin.
Figure 1. The American River Basin above 1500 m elevation, divided into approximately 7000 500-m square pixels. Snow-pillow and snow-course locations from California Cooperative Snow Survey. Optimal sensing locations selected through the method proposed in this paper are shown as green squares.
Figure 2. (a) Snow course, snow pillow, and interpolation product over 2006 melt season (b) Time slices of interpolation product shown for May 1, May 15, and June 1.
Figure 3. (a) One pixel chosen randomly from each 300 meter elevation band shown in Figure 1 (b) time series SWE for each pixel (c) Ranks across melt season.
Figure 4. (a) Five rank based clusters (b) Cluster mean and standard deviation over 2006 season (c) Histograms of SWE by cluster for April 15.
Figure 5. Histograms of physiographic variables divided by rank-based cluster.
Figure 6. Average RMSE vs. sensors placed, averaged over 2009-2011 melt seasons. Rank-based results shown for 5 clusters.
Figure 7. (a) 5, 50, and 95th percentiles of SWE across whole basin (b) RMSE over 2009 melt season. Rank-based cluster results shown for five clusters.
Figure 8. Average RMSE vs. number of rank based clusters. At least two sensors are required per cluster.
Figure 9. Recommended regions for placements in three rank-based clusters using *Mutual Information* based placement algorithm. Round one is the most informative placement, round two the second most and so on. One sensor should be placed within a given region each round. All proposed locations within a given region are equally valuable.
Figure 10. Boxplots of regression parameters for relationship between elevation and SWE computed daily across each melt season for American River Basin.