Title

Spatial snow water equivalent estimation for mountainous areas using wireless-sensor networks and remote-sensing products

Authors

Zheng, Zeshi
Molotch, Noah P
Oroza, Carlos A
et al.

Publication Date

2018-09-01

DOI

10.1016/j.rse.2018.05.029

Peer reviewed

Remote Sensing of Environment, Vol. 215, 44-56, 2018
https://doi.org/10.1016/j.rse.2018.05.029

Spatial Snow Water Equivalent Estimation for Mountainous Areas Using Wireless-Sensor Networks and Remote-Sensing Products

Zeshi Zheng[1]*, Noah P. Molotch[2], Carlos A. Oroza[1], Martha Conklin[3], Roger C. Bales[1, 3]

Abstract

We developed an approach to estimate snow water equivalent (SWE) through interpolation of spatially representative point measurements using a *k*-nearest neighbors (*k*-NN) algorithm and historical spatial SWE data. It accurately reproduced measured SWE, using different data sources for training and evaluation. In the central-Sierra American River basin, we used a *k*-NN algorithm to interpolate data from continuous snow-depth measurements in 10 sensor clusters by fusing them with 14 years of daily 500-m resolution SWE-reconstruction maps. Accurate SWE estimation over the melt season shows the potential for providing daily, near real-time distributed snowmelt estimates. Further south, in the Merced-Tuolumne basins, we evaluated the potential of *k*-NN approach to improve real-time SWE estimates. Lacking dense ground-measurement networks, we simulated *k*-NN interpolation of sensor data using selected pixels of a bi-weekly Lidar-derived snow water equivalent product. *k*-NN extrapolations underestimate the Lidar-derived SWE, with a maximum bias of -10 cm at elevations below 3000 m and +15 cm above 3000 m. This bias was reduced by using a Gaussian-process regression model to spatially distribute residuals. Using as few as 10 scenes of Lidar-derived SWE from 2014 as training data in the *k*-NN to estimate the 2016 spatial SWE, both RMSEs and MAEs were reduced from around 20-25 cm to 10-15 cm comparing to using SWE reconstructions as training data. We found that the spatial accuracy of the historical data is more important for learning the spatial distribution of SWE than the number of historical scenes available. Blending continuous spatially representative ground-based sensors with a historical library of SWE reconstructions over the same basin can provide real-time spatial SWE maps that accurately represents Lidar-measured snow depth; and the estimates can be improved by using historical Lidar scans instead of SWE reconstructions.

Keywords

Mountain hydrology
Snow water equivalent
Wireless-sensor networks
Sierra Nevada
k-Nearest neighbors
Reconstructions
MODIS
Lidar

Citation

Zeshi Zheng, Noah P. Molotch, Carlos A. Oroza, Martha H. Conklin, Roger C. Bales, Spatial snow water equivalent estimation for mountainous areas using wireless-sensor networks and remote-sensing products, Remote Sensing of Environment, Volume 215, 15 September 2018, Pages 44-56, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2018.05.029.

[1] *Department of Civil and Environmental Engineering, University of California, Berkeley, California, USA*
[2] *Institute of Arctic and Alpine Research, Department of Geography, University of Colorado, Boulder, Colorado, USA*
[3] *Sierra Nevada Research Institute, University of California, Merced, California, USA*
***Corresponding author**: zeshi.z@berkeley.edu

1. Introduction

In the state of California, ecosystem processes and water supplied for agricultural and urban uses depend on the snowpack in the Sierra Nevada as the primary source of spring and summer streamflow (Bales et al. (2006)). As the prediction of water availability and flood peaks depend in part on snowpack conditions, accurate knowledge of the snowpack can assist decision making for water resources management (California Department of Water Resources (2013)).

Current decision making for water management in California during the snowmelt season relies on ground measurements in the Sierra Nevada, which include continuous snow-pillow and snow-depth sensor measurements, and monthly manual snow surveys (Molotch and Bales (2005)). Ground stations are sparsely placed in the mountains compared to the spatial scale of each watershed. Therefore, the measurements may not be representative of physiographic features required to capture spatial variability of snow depth and snow water equivalent (SWE), either at the site or basin scale. Satellite-based remote sensing, such as MODIS and Landsat, has been used to map snow coverage at regional to global scales. However, they provide only pixel-wise fractional snow-coverage information, with no direct information on snow depth or SWE (Dozier et al. (2008); Molotch and Margulis (2008); Painter et al. (2009); Raleigh et al. (2013); Rittger et al. (2013); Rosenthal and Dozier (1996)). A modeled snow-data product that is commonly used in the Continental United States is the Snow Data Assimilation System (SNODAS), which integrates snow information from both satellite and ground stations, providing daily snow depth and snow water equivalent information at 1-km^2

resolution (Barrett (2003)). Recent work validating the SNODAS spatial product with Lidar suggested that the performance of SNODAS in the Tuolumne River basin is less accurate than 3-D ($x, y,$ elevation) bilinear interpolation of ground stations (Bair et al. (2016)).

Snow-coverage information and modeled spatial land-surface meteorological data can be used to back-integrate SWE from the snow melt-out date to the date of maximum SWE at the beginning of the snowmelt season. This technique has been applied across several mountain ranges and is referred to as the SWE-reconstruction technique (Bair et al. (2016); Guan et al. (2013); Margulis et al. (2016); Rittger (2012)). Although SWE reconstruction captures both temporal variability and spatial variability, it can only be done at the end of the season when the daily energy inputs and snow covered area are known (Cline et al. (1998)).

As a complement to satellite-based estimates of snow distribution, numerous statistical models have been developed to interpolate point-based snow information. Multivariate linear regression, commonly used in previous studies, can relate physiographic variables, historical SWE data, and snow covered area imagery with the observed SWE; and the accuracy is reasonably better than techniques such as inverse-distance weighting and simple kriging (Schneider and Molotch (2016); Fassnacht et al. (2003)). However, the linear-regression-based methods do not provide spatially smooth maps and the independent variables do not necessarily have a linear relationship with SWE (Zheng et al. (2016)). Other than regression, one category of methods that have shown promise are nearest-neighbor-based algorithms. These algorithms are attractive because they are easy to implement, nonparametric, learning based, and can learn linear and nonlinear trends in observations (Ni and Nguyen (2009)). Simulations and estimations at either fine temporal or spatial resolutions using parametric models can be computationally intensive. Nearest-neighbor approaches have therefore become an alternative solution to many problems in spatio-temporal modeling, not only for their advantage in time complexity, but also for their superior accuracy and ability to preserve patterns from observations. The k-NN algorithm has been used for multivariate time-series simulation for weather forecasting (Rajagopalan and Lall (1999)), disaggregating meteorological time-series data to finer time scales (Prairie et al. (2007); Kalra and Ahmad (2011)), and downscaling spatial climate-model data (Gangopadhyay and Clark (2005)). The k-NN algorithm was found to be superior in preserving the spatio-temporal covariability of the observation than multivariate autoregressive approaches.

To address the issues in presently available basin-scale water-balance data, a prototype real-time observation network that includes monitoring the snow conditions is being developed for the headwater areas of the American River basin in the Sierra Nevada (Zhang et al. (2017)). The system enables combining ground measurements of snow depth and historical SWE reconstruction using a k-nearest neighbors (k-NN) algorithm for real-time spatial SWE estimation (Larose (2005)).

This work documents the k-nearest neighbors spatial-SWE-estimation method and evaluates the estimates against a spatial SWE product that is derived from Lidar-measured snow depth. Three questions that motivated the present study are:

1. Does a k-NN approach for spatial SWE interpolation in mountainous regions provide accurate SWE estimates relative to other products?

2. How is the error of the k-NN estimation distributed with regard to topographic variables?

3. Is it possible to further decrease the error of the k-NN estimates by distributing the residuals spatially?

2. Methods

We applied the k-nearest neighbors (k-NN) algorithm to estimate spatial snow water equivalent (SWE) in three basins in the Sierra Nevada, California, USA (Figure 1a, Table 1). The experiment for the American River basin focused on estimating the 2014 spatial SWE using 10 clusters of snow-depth measurements for 2014 from wireless-sensor networks, and historical SWE reconstructions based on MODIS from 2001 to 2013, aiming to evaluate the k-NN estimates temporally over the melt season. The SWE reconstructions were used by the algorithm for learning the SWE spatial distribution embedded in the data set. We did similar experiments in the Merced (2014) and Tuolumne (2014, 2016) basins using Lidar-based SWE estimates to evaluate the k-NN results spatially. For these two basins, since we have fewer sensor networks deployed, we instead selected representative pixels as hypothetical sensor-network locations based on physiographic variables using a Gaussian-mixture model; and used these Lidar-based SWE values for the k-NN experiments. In this setup we used historical SWE reconstructions, historical Lidar-derived SWE, and historical SNODAS SWE as spatial training data to explore if different data sources matter. The spatial results over the two basins were evaluated using the Lidar-derived SWE as a ground-truth data set.

Table 1. Topography summary of the three studied basins (above 1500-m elevation only)

Basin	Elevation range, m	Area, km^2
American	142−3070	2116
Tuolumne	408−3870	1136
Merced	1021−3927	853

2.1 American River basin analysis using wireless-sensor network data

The 10 wireless-sensor networks (Table 2) were deployed in the seasonally snow-covered region of the 5570 km^2 American basin (Figure 1b). Each network has ten or eleven sensor nodes (Figure 1c) that measure snow depth, temperature, relative humidity, soil moisture, and short-wave solar radiation (Zhang et al. (2017); Brun-Laguna et al. (2016)). The placements were strategically selected, aiming to capture snow depth and meteorological variability from elevation gradients, south versus north-facing slopes, steep versus flat areas, and various vegetation densities. All sensors take measurements at a 15-minute intervals, and the network manager of each sensor cluster forwards the data to a central webserver (Zhang et al. (2017)). Daily data averaged over each cluster were used in the current analysis.

2.2 Snow water equivalent reconstruction data

Snow water equivalent reconstruction is an existing data set that was produced by estimating historical spatial SWE for past snowmelt seasons (Guan et al. (2013); Molotch et al. (2017)). The time extent of the SWE reconstructions is from 2000 to 2014 and the spatial extent covers the entire Sierra Nevada. The SWE-reconstruction method uses a snow-surface energy and mass-balance model:

$$M_p \rho L = S\downarrow(1-\alpha) + LW\downarrow + LW\uparrow + SH + LH \tag{1}$$

where M_p(m) is the potential snowmelt (assuming full snow coverage), ρ(kg/m^3) is the liquid-water density, L(kJ/kg) is the latent heat of fusion, $S\downarrow$ (J/m^2) is the subcanopy insolation, α(unitless) is snow albedo, $LW\downarrow$ (J/m^2) is the downwelling longwave radiation, $LW\uparrow$ (J/m^2) is the longwave radiation emitted from the snowpack. SH(J/m^2) and LH(J/m^2) are sensible heat exchange and latent heat exchange accordingly. We need to note that the SWE-reconstruction model did not

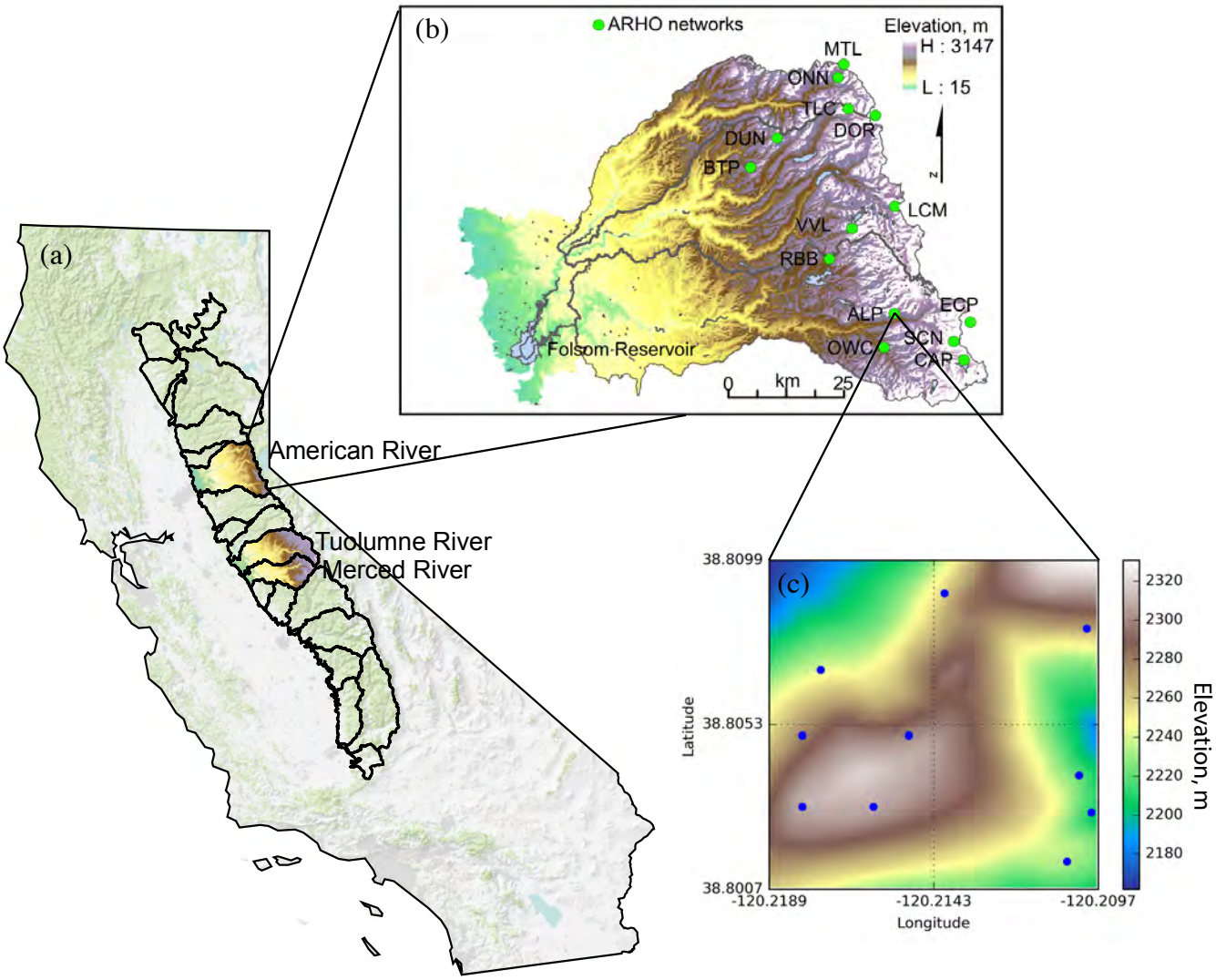

Figure 1. (a) The study areas of this research work are American River, Merced River and Tuolumne River basins that lie on the west side of the Sierra Nevada. (b) American River basin is instrumented with 14 wireless-sensor networks, which forms the American River Hydrologic Observatory (ARHO) (Zhang et al. (2017)). Please note the analysis in this study used 10 clusters' data. (c) Sensor locations of site Alpha in the Observatory, showing that the sensor locations of the network capture elevation gradient at the site.

Table 2. Topographic Statistics of Sensor Clusters in the American River basin.

Cluster	Latitude	Longitude	Elevation, m	Slope, deg	Aspect, deg	Canopy Coverage, %
Alpha	38.80	-120.21	2319	2.7	181.1	55
Bear Trap	39.09	-120.58	1538	3.5	328.7	61
Caples Lake	38.71	-120.04	2430	4.1	220.6	20
Rice Creek*	39.15	-120.37	2024	7.0	246.6	30
Duncan Peak	39.15	-120.51	2111	10.4	264.3	42
Echo Peak	38.85	-120.07	2510	16.8	197.1	21
Lost Corner*	39.02	-120.21	2310	13.6	266.6	15
Mount Lincoln	39.29	-120.33	2405	8.1	194.4	2
Onion Creek	39.28	-120.36	1888	1.5	247.7	24
Owens Camp*	38.74	-120.24	1552	3.5	223.3	25
Robbs Saddle	38.91	-120.38	1828	3.5	107.9	32
Schneiders	38.75	-120.07	2694	9.9	293.3	42
Talbot Camp*	39.19	-120.38	1756	4.8	151.2	60
Vanvleck	38.94	-120.31	2023	2.7	202.8	48

* These sites were operational after the summer in 2014 so they were not included in the k-NN analysis.

account for precipitation that occurs during the melt period, which may introduce bias in the estimates. The potential snowmelt M_p is scaled by the fractional snow-covered area (f_{SCA}) derived from MODIS to estimate the actual daily snowmelt,

$$M = M_p \times f_{SCA} \tag{2}$$

The time-series SWE for the season is calculated by back integrating the daily snowmelt since snow meltout:

$$\mathrm{SWE}_0 = \mathrm{SWE}_n + \sum_{j=1}^{n} M_j \tag{3}$$

where SWE_n is SWE at time step n, SWE_0 is the initial SWE, and M_j is the actual snowmelt during time step j. The initial SWE at each model pixel can be reconstructed at the time when snow disappearance observed from the satellite ($f_{SCA} = 0$):

$$\mathrm{SWE}_0 = \sum_{j=1}^{n} M_j \text{ when } \mathrm{SWE}_n = 0 \tag{4}$$

2.3 *k*-nearest-neighbor-based SWE estimation

A k-nearest neighbors algorithm was used to find the closest snowpack conditions in historical SWE reconstructions to real-time ground observations from the sensor networks data that are introduced in Section 2.1. The daily historical SWE reconstructions that were estimated for April 1st to August 31st for each year from 2001 to 2013, resulted in a dataset containing 1988 d-dimensional data points, where d is the number of independent variables, with each variable being a SWE value estimated from a spatially scaled ground measurement of snow depth. Before the execution of the k-NN algorithm, preprocessing was involved:

1. The ground measurements are of snow depth, whereas the SWE reconstructions are snow water equivalent. The conversion factor between the two is snow density. The spatial variation of snow water equivalent in the Sierra Nevada caused by snow density is small compared to the snow depth (Painter et al. (2016)); and no spatial patterns in snow density were apparent in the 2014 data (Zhang et al. (2017)). In our model, the density value was estimated as the mean ratio of the co-located snow pillows and snow-depth sensors' measurements in the basin. We tested the relationships between the snow densities and terrain variables but did not find any. Therefore in further calculations we assumed that the snow density is uniform across the model spatial extent at each time step.

2. The spatial resolution (i.e. 500 m) of the SWE reconstructions is consistent with our model, however the sensor-network measurements are not. To make the sensor-network measurements consistent, we grouped the sensors bounded by the same SWE-reconstruction pixel and took their mean as the ground observation that is representative of the SWE-reconstruction pixel; these values were used in the k-NN algorithm. The grouped sensors were mapped to 49 SWE-reconstruction pixels. In this manner the multiple sensors within each 500-m pixel account for some of the sub-pixel variability found within the pixel. That said, the degree to which we sample the sub-grid variability cannot be explicitly determined. A detailed analysis of sub-pixel snow depth variability is beyond the scope of the current study.

3. Certain pixels in the SWE reconstructions that correspond to sensor-network measurements can be highly biased and inaccurate based on the comparison between 2014 sensor-network measurements and the SWE reconstructed for the same year. The bias has been found to affect the distance metric for the k-NN algorithm (Li et al. (2016); Liu et al. (2014)). In order to filter out the biased and inaccurate SWE-reconstruction pixels, we calculated the correlation coefficient in time between each pair of ground measurements and SWE reconstructions during the 2014 snowmelt. A threshold (p-value ≤ 0.001) was used to decide if the pixel was retained: if the p-value exceeded the threshold then the ground measurements and SWE reconstructions for that pixel were excluded. Of the 49 pixels, 21 were retained after filtering; and the retained pixels' correlation matrices are shown in Figure 2. All pixel pairs show that the R^2s are above 0.5. A high temporal correlation between pixels suggests that the spatial patterns of SWE are consistent in time but with the basin-wide mean SWE changing over the season. This filtering step is not needed if the historical spatial data is from measurements like Lidar.

4. The default setting of a k-NN algorithm uses the Euclidean distance as a metric for comparing pair-wise distance between data points. The unique distribution and collinearity of a mountain snowpack's spatial distribution implies that using Euclidean distance might unevenly weight each dimension when calculating the distance (Figure 2). Therefore, we subtracted the mean from each variable and then used the Mahalanobis-distance metric to normalize the data by its covariance matrix,

$$d(\vec{x}_1, \vec{x}_2) = \sqrt[\mathrm{fn}]{(\vec{x}_1 - \vec{x}_2)^T S^{-1} (\vec{x}_1 - \vec{x}_2)} \tag{5}$$

where $d(\vec{x}_1, \vec{x}_2)$ is the Mahalanobis distance between two vectors $\vec{x}_1$ and $\vec{x}_2$, S is the empirical covariance matrix. The ill-conditioned empirical covariance matrix was adjusted by adding a small number to all diagonal elements such that the covariance matrix will become full rank without losing much information and its inverse is computable after the adjustment.

5. In the k-nearest neighbors algorithm, we stored the filtered 21 pixel values from SWE reconstructions into a Balltree data structure (Cayton (2009)). The number of nearest neighbors to be searched was determined by a transformed leave-one-out cross-validation tuning scheme. A typical cross-validation scheme leaves out part of the samples as a testing set for evaluating the predictive model trained by the remaining samples (training set). In our task, we need to test if the method is adequate to estimate SWE spatially and holding out SWE reconstruction samples does not help evaluating the spatial predictive performance. Therefore, we left out one pixel in the training set and evaluated on it in validation; and by doing so we validated whether the method developed can estimate SWE at unobserved locations.

The 21 selected pixels' time series from the 2001-2013 SWE reconstructions were formated to the training data set and the corresponding snow-depth ground measurements from 2014 were converted to SWE and formated to the testing data set. We ran the k-NN algorithm using these formated data. The number of nearest neighbors (k) was chosen to be 30 since the error calculated in the leave-one-out cross-validation did not significantly decrease when k became larger than 30.

2.4 Analysis using Lidar data

The Merced-Tuolumne basins were surveyed by airborne Lidar at sub-monthly temporal frequency during snowmelt by NASA's Airborne Snow Observatory (ASO). The surveyed areas have elevation ranges of 1021 to 3927 m in the Merced basin and 408 to 3871 m in the Tuolumne basin. ASO produces a 50-m resolution SWE product that is derived from Lidar-measured snow-depth and modeled snow density (Painter et al. (2016)). We used the ASO products from 2014 and 2016 (Table 3) to evaluate the k-NN estimation method spatially.

ASO collected and processed Lidar under the following standards: The raw Lidar data were collected using a Riegl Q1560 airborne laser scanner to measure surface elevations. Snow-depth maps were calculated by subtracting snow-off gridded surfaces from snow-on gridded surfaces. The spatial resolution of the Lidar snow-depth raster product is 3-m resolution and the snow-depth estimates have mean absolute errors of < 8 cm,

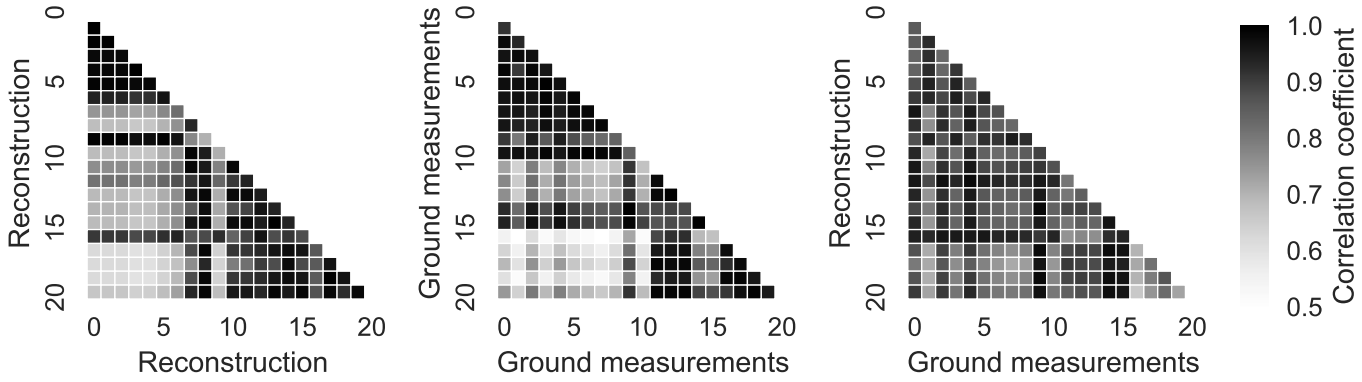

Figure 2. Lower triangle of correlation matrices of SWE reconstruction, ground measurements and correlation between the two for the American River basin, with autocorrelations excluded.

Table 3. Lidar survey dates in 2014 and 2016 at Merced and Tuolumne basins.

	2014	2016
Merced	03-24, 04-06, 04-14, 04-23, 04-29 05-03, 05-12	N/A
Tuolumne	03-23, 04-07, 04-13, 04-20, 04-28 05-02, 05-11, 05-17, 05-27, 05-31	04-01, 04-07, 04-16, 04-26, 05-27 06-07, 06-13, 06-20

with bias < 1 cm when compared with manually measured depths at the 15×15m scale (Painter et al. (2016)). This 3-m snow-depth raster product was resampled to 50-m resolution to match the spatially distributed version of the snowcover energy and mass balance model resolution (ISNOBAL) (Marks et al. (1999)). The resampled 50-m snow-depth data were then converted to SWE by multiplying by the spatial snow density estimated from the ISNOBAL model. Our k-NN simulations were based on the 50-m resolution SWE products. To enable simulating k-NN using Lidar data, the ASO 50-m resolution SWE product was resampled again to 500-m resolution by averaging all 50-m grids inside the corresponding 500-m grid. The SWE estimates from the k-NN simulation were evaluated against the 500-m Lidar product. The evaluation of the k-NN simulation was compared to an evaluation of SNODAS against the Lidar data. SNODAS is an operational, spatial snow product. It produces a spatially distributed SWE estimates at 1-km resolution in near real time since 2004. SNODAS assimilates a physically based model with SNOTEL observations and remotely sensed snow-covered-area images (Barrett (2003)). Previous work suggests that SNODAS works well in environments that are geographically closer to the SNOTEL stations, and the accuracy decays as the distance increases (Clow et al. (2012)). The snow-depth comparison results between SNODAS and Lidar in Colorado's Rocky Mountains show that the accuracy is reasonable in general (Hedrick et al. (2015); Schneider and Molotch (2016)). Therefore we used it as a SWE-estimation performance benchmark.

Since there are only 4 small sensor clusters deployed in the Merced and Tuolumne basins, and in order to conduct similar experiments as what we did in the American River basin, we used a Gaussian-mixture model to select physiographically representative locations and treated them as sensor-network instrumented locations. These locations' corresponding pixel values were extracted from the Lidar-derived SWE maps and they were used as inputs in the k-NN algorithm for both basins. After running the k-NN algorithm, Gaussian-process regression was conducted to distribute residuals spatially across these two basins. In order to determine if the spatial SWE estimates perform better than a typical SWE product, the k-NN SWE estimates were compared to SNODAS.

As the spatial distribution of SWE is highly dependent on physiographic variables such as elevation, slope, aspect and vegetation (Zheng et al. (2016)), the 3-m resolution digital elevation models (DEM) of both basins retrieved from the snow-off Lidar measurements were resampled to 500-m resolution DEMs and were further processed to derive slope and aspect at 500-m resolution. A 30-m resolution National Land Cover Database 2011 (NLCD 2011) USFS Tree Canopy Analytical dataset was clipped, resampled, and georegistered to the same resolution and extent as the 500-m resolution DEMs. These 4 variables were combined with latitude and longitude to form a 6-dimensional vector (i.e., the combined $\mathbb{R}^6$, $\mathbf{x} = [x^{\text{latitude}}, x^{\text{longitude}}, x^{\text{elevation}}, x^{\text{slope}}, x^{\text{aspect}}, x^{\text{canopy}}]$). Selection of the most-representative locations for siting sensor networks was accomplished using a Gaussian mixture model (Oroza et al. (2016)). The Gaussian mixture model is a parametric probability density function represented as a weighted sum of Gaussian component densities as given by the equation,

$$p(\mathbf{x}|\lambda) = \sum_{i=1}^{M} w_i \mathcal{N}(\mathbf{x}|\mu_i, \Sigma_i) \tag{6}$$

where w_i is the mixture weight and $g(\mathbf{x}|\mu_i, \Sigma_i)$ is the component Gaussian density for each Gaussian component. Each component density is a D-variate Gaussian function of the form,

$$\mathcal{N}(\mathbf{x}|\mu_i, \Sigma_i) = \frac{1}{(2\pi)^{D/2}|\Sigma_i|^{1/2}} \exp\{-\frac{1}{2}(\mathbf{x}-\mu_i)'\Sigma_i^{-1}(\mathbf{x}-\mu_i)\} \tag{7}$$

with mean vector μ_i and covariance matrix Σ_i. The mixture weights satisfy the constraint that $\sum_{i=1}^{M} w_i = 1$. The objective of the Gaussian

mixture model is to maximize the likelihood function (8) by estimating the parameters using the Expectation Maximization (EM) algorithm.

$$p(X|\lambda) = \prod_{t=1}^{T} p(\mathbf{x}_t|\lambda) \tag{8}$$

where λ is the collection of parameters (i.e., $\lambda = \{w_i, \mu_i, \Sigma_i\} \quad i = 1,...,M$).

The estimated mean vector of each Gaussian component is the expected 6-dimensional vector. Since the given set of data does not necessarily have a data point equal to this mean, the data point that has the smallest Euclidean distance to the estimated mean vector is selected as the sampling location.

The number of sampling locations needed is decided by the Akaike information criterion (AIC). AIC is a measure of the relative quality between statistical models that are fitted with the same given set of data. It offers a relative estimate of the information lost when a given model is used to represent a process that generates the given data. In doing so it addresses the trade-off between the model's goodness of fit and model complexity. For the Gaussian-mixture model, the parameter that differentiates between models is the number of Gaussian components (i.e., sampling locations in this case). By using the AIC equation,

$$\mathrm{AIC}(\lambda) = -2\ln(p(X|\lambda)) + 2d \tag{9}$$

where $p(X|\lambda)$ is the maximized likelihood and d is the number of Gaussian components, the integer value of d was selected as the number of sampling locations for which the corresponding model had the lowest AIC score.

The sensor-network locations selected were used to extract 500-m resolution Lidar-derived SWE pixels. These pixels were used to simulate k-NN algorithm with historical SWE reconstructions (described in Section 2.3). The preprocessing steps were not applied because they are specific to the snow-depth sensor data. The k-NN estimated SWE maps were evaluated against the Lidar-derived SWE map on a pixel-wise basis and the error statistics calculated. The error statistics used were root-mean-square error (RMSE),

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{10}$$

and mean-absolute error (MAE),

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{11}$$

where n is the number of pixels, $\hat{y}$ is the k-NN estimated pixel value and y is the Lidar-derived pixel value.

2.5 Bias-correction on *k*-NN simulation using Gaussian process regression

The SWE-reconstruction product is inherently embedded with systematic bias because of (i) the inaccuracy to infer the final date of the seasonal snow cover from remote sensing; (ii) errors and coarse spatial scales of meteorological data; and (iii) weaknesses in the snow model used for SWE reconstructions (Slater et al. (2013)). The k-NN interpolated SWE could be biased because the spatial distribution of SWE depends on SWE reconstructions. To correct the bias inherited from SWE reconstructions, we calculated the k-NN estimation residuals for all sensor observed locations and the residuals were modeled using a Gaussian-process regression. The entire process of applying k-NN algorithm and Gaussian-process regression is called k-NN+GP in the following text.

Gaussian-process regression, the same as simple kriging in geostatistics, is a technique that has been extensively used in spatial inference. The regression relates the observation y to an underlying function $f(x)$ through a Gaussian noise model:

$$y = f(x) + \mathcal{N}(0, \sigma_n^2) \tag{12}$$

The essence of regression is the search for $f(x)$; in Gaussian process regression, what relates one observation to another is the *covariance function*, $k(x,x')$. For example, a popular choice of covariance function is the *squared exponential function*, and by folding the Gaussian noise into the covariance function, we have

$$k(x,x') = \sigma_f^2 \exp\left[\frac{-(x-x')^2}{2l^2}\right] + \sigma_n^2\delta(x,x') \tag{13}$$

where σ_f^2 is the maximum allowable covariance, l is the length scale parameter of the covariance function, and $\delta(x,x')$ is the Kronecker delta function.

Since the key assumption in Gaussian-process modeling is that our data can be represented as a sample from a multivariate Gaussian distribution, we have

$$\begin{bmatrix} \mathbf{y} \\ y_* \end{bmatrix} = \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K & K_*^{\mathrm{T}} \\ K_* & K_{**} \end{bmatrix}\right) \tag{14}$$

where K, K_* and K_{**} can be summarized in the equations below.

$$K = \begin{bmatrix} k(x_1,x_1) & k(x_1,x_2) & \cdots & k(x_1,x_n) \\ k(x_2,x_1) & k(x_2,x_2) & \cdots & k(x_2,x_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(x_n,x_1) & k(x_n,x_2) & \cdots & k(x_n,x_n) \end{bmatrix} \tag{15}$$

$$K_* = [k(x_*,x_1) \quad k(x_*,x_2) \quad \cdots k(x_*,x_n)] \tag{16}$$

$$K_{**} = k(x_*,x_*) \tag{17}$$

From the above equations,

$$y_*|\mathbf{y} \sim \mathcal{N}(K_*K^{-1}\mathbf{y}, K_{**} - K_*K^{-1}K_*^{\mathrm{T}}) \tag{18}$$

In doing the SWE residual interpolation, the independent variables $\mathbf{x}$ are the same as those being used in the Gaussian mixture model. The Matérn covariance function was used for its better prediction accuracy than other functions. The parameters in the covariance function were optimized with regards to SWE residuals at observed locations.

2.6 Using historical Lidar scans and SNODAS as training data

Considering that the k-NN interpolation method for estimating spatial SWE can be used with any historical spatial SWE product, not just SWE reconstructions, we repeated the k-NN and k-NN+GP analysis for 2016 over the Tuolumne basin, but with all SWE reconstructions replaced with two different historical data sets, the 2014 Lidar-derived SWE and the SWE estimated from SNODAS for the years of 2004−2015. In the setup with 2014 Lidar data, only 10 scenes are included in the training set, which is only 0.5% of the number of historical SWE reconstructions.

When running the k-NN algorithm we used $k = 1$ because we had a limited number of historical Lidar scenes. The setup with SNODAS has 1461 scenes so we kept the same configuration as in the analysis using the SWE reconstructions. We compared these estimates based on the mean bias error (MBE), systematic root-mean-square error ($RMSE_s$), unsystematic root-mean-square error ($RMSE_{us}$) that are introduced in Willmott (1982).

3. Results

3.1 *k*-nearest neighbors plus Gaussian-process regression (*k*-NN+GP) using SWE reconstructions

Spatial estimates of American River SWE of three days around peak season in 2014 are shown in Figure 3. Snow accumulated across the high-elevation region of the basin during March and early April and the peak SWE is observed on April 5th. The snowpack melted quickly within two weeks. The SWE for most high-elevation regions are below 0.25 m and only the peak regions of the mountains have about 0.5 m of SWE, as is shown in the spatial SWE map estimated for the May 3rd. A comparison between ground observations and the k-NN estimates (Figure 4) suggests that the k-NN algorithm has low bias on a basin-wide scale, but can be highly biased for certain pixels. The bias is likely attributable to both error induced by upscaling ground measurements to 500-m resolution and the inherent bias in the SWE reconstructions as sensor-network measurements are more accurate than SWE reconstructions.

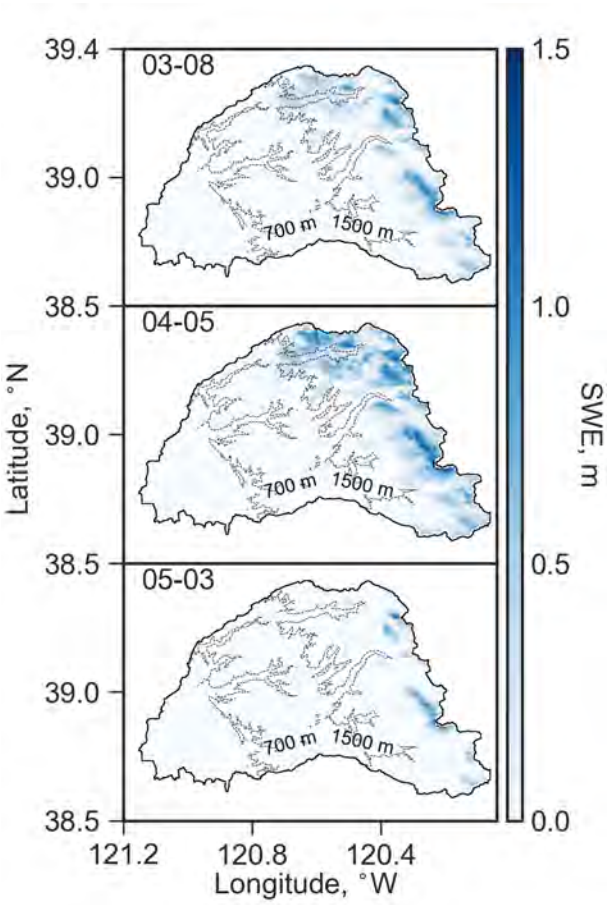

Figure 3. Snow water equivalent maps in the American River basin created using ground measurements and historical SWE-reconstruction data through *k*-NN algorithm. Bi-weekly maps are shown in Figure S1.

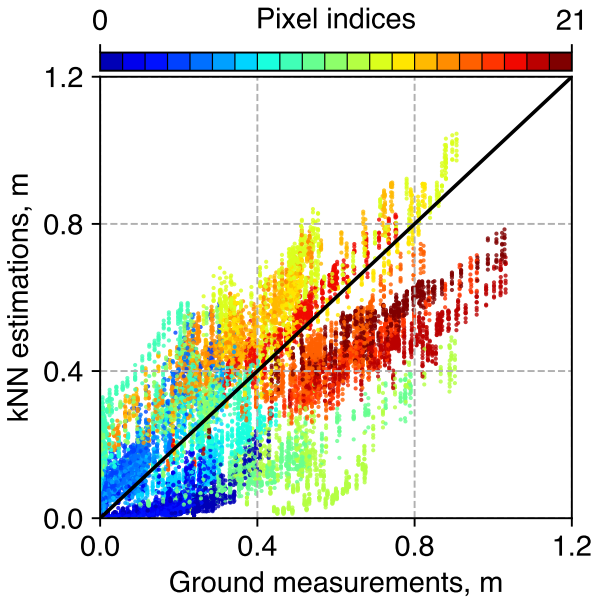

Figure 4. Scatter plot showing ground measurements and searched *k*-nearest neighbors at 21 locations (each location is a different color, referred as pixel index) in the American River basin. The estimations vs. measurements shows 1:1 trend overall; however, the relationship for each individual sensor is not necessarily 1:1 and the intercept for each may not be zero.

Although SWE reconstructions are not accurate estimates of SWE, it is still interesting to determine whether k-NN could be a technique for extrapolating the ground measurements across the basin. Figure 5 shows a time series comparison between ground observations, k-NN estimates, and SWE reconstructions at 3 selected 500-m-resolution pixels. The k-NN estimation for each pixel is evaluated using leave-one-out cross-validation. Comparing the k-NN estimates with ground observations, some pixels have low bias at all times, while most of the pixels differ somewhat throughout time series. We found that the bias does not depend on elevation (Figure S2). The magnitude of error could be as large as 0.4 m for certain pixels. The SWE-reconstruction curve suggests that SWE reconstructions underestimate SWE at all locations during the entire snowmelt season and the magnitude of error can be as large as 0.8 m for certain pixels.

To evaluate the k-NN estimation in the Merced and Tuolumne basins, 24 locations in the Merced and 33 locations in the Tuolumne were selected using a Gaussian mixture model (Figure 6). Figure 7 shows a pixel-wise comparison of three estimation methods with the Lidar-derived SWE. Point colors on the scatter plots represent the estimated Gaussian distribution kernel density of these data points over $\mathbb{R}^2$. The k-NN+GP estimation exhibits minimum bias relative to Lidar (fitted linear regression line is closest to the 1:1 line in Figure 7 on the majority of days in both basins). Slopes and coefficients of determination presented in Figure 8 suggest the same. The RMSEs and MAEs of these estimation methods for each Lidar date are shown as a time series, also in Figure 8. From these we observe that RMSEs and MAEs of the k-NN and the k-NN+GP estimates are smaller than SNODAS for the majority of days in 2016 over the Tuolumne basin. Both error statistics can be reduced as much as 10 cm during a normal snow season (2016). Similar results were observed from the analysis over the Merced basin in 2014. The estimated bias in reconstructed SWE starts to increase with elevation. The elevation-band-mean SWE estimated from k-NN has a maximum bias of -15 cm at elevations below 3000 m and +15 cm above 3000

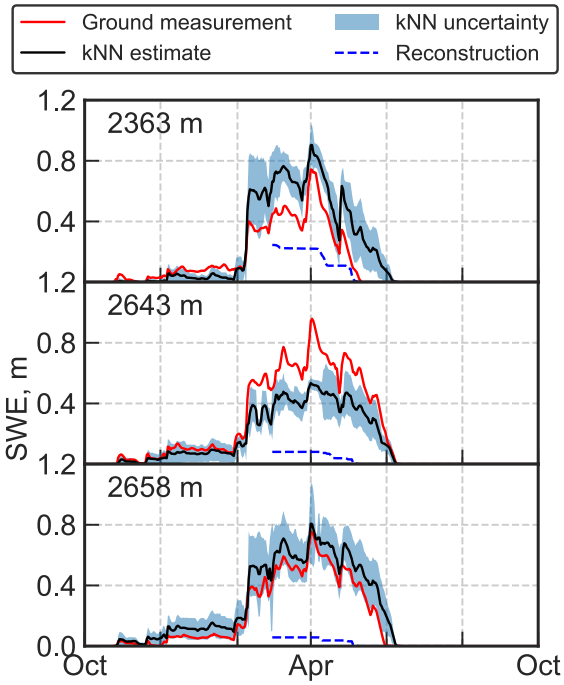

Figure 5. Daily snow water equivalent time series from the ground measurements, *k*-NN estimations, and SWE reconstructions at 3 selected pixels, in the American River basin for the 2014 water year. See Figure S2 for 20 locations sorted by elevations.

m. *k*-NN+GP can correct part of the bias over a few elevation bands. However, SNODAS has large underestimation bias for regions above 3000 m (Figure 9).

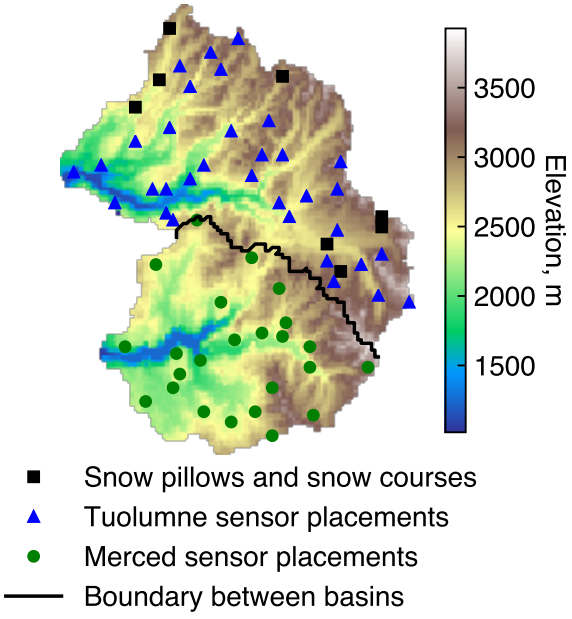

Figure 6. Snow pillows, snow courses, and selected simulation sensor locations at Merced and Tuolumne basins.

3.2 Estimation using historical Lidar-derived SWE and SNODAS as training set

Although only 10 Lidar-derived SWE maps produced in 2014 were included in the training set, the prediction SWE for 2016 is highly correlated with the SWE derived from Lidar (correlation inferred from scatterplots and kernel densities in Figure 10). The detailed statistics (Table 4) reinforce that by showing that R^2s are in the range of 0.82−0.87 for these four days. Also, both RMSEs and MAEs estimated using Lidar-derived SWE are at least 10 cm less than that using SWE reconstructions. Also, the systematic error decreases drastically when GP is used. The only drawback we found from this experiment is that the slopes of the scatter plots decrease as the season approaches to the end. The MBEs suggest that the basin-wide mean SWE is unbiased if using SNODAS SWE as training data but the R^2s suggest it can only explain 40−50% of the variability (Table 4).

3.3 *k*-NN estimation accuracy versus number of SWE reconstruction scenes

The performance of a statistical prediction model depends in part on the size of the training data set, as verified in computer vision and machine learning research (Zhu et al. (2016)). This finding agrees with our intuition to include as many SWE-reconstruction scenes as possible in the data set used to train the nearest neighbors algorithm. Since each SWE-reconstruction scene is embedded with a unique spatial distribution of snow water equivalent, the estimation accuracy should be improved as long as more scenes are stored. In other words, the *k*-NN algorithm will have a higher probability of finding closer neighbors when more scenes are included in the training set.

We verified this hypothesis by randomly selecting SWE-reconstruction scenes and increasing the number of scenes incrementally in our training set from 10% to 100% of the total number of scenes available (rounding up the percent scene numbers to the closest integer if needed). We compared the *k*-NN estimation results with the 500-m Lidar-derived SWE from the Tuolumne basin, using RMSEs as error statistics representing estimation performance. RMSEs decrease when more SWE-reconstruction scenes are included in the training set for most days simulated for the Tuolumne basin (Figure 11). Similar patterns were observed when running the same experiment on the American-River-basin wireless-sensor-network data (Figure 12), in which the cross-validation RMSEs are shown because a Lidar evaluation set is not available. Both figures imply that in general the decrease of the RMSEs tends to become saturated when the training set has more than 50% of all SWE reconstructions, suggesting 1000 SWE reconstructions are sufficient for capturing spatial SWE variability. There is also case when RMSEs are not sensitive to the number of SWE reconstructions (April 23rd, 2014, Figure S4), when Lidar data was collected immediately following a storm. The SWE-reconstruction model did not include the melting season precipitation events in the surface energy and mass balance calculation, which could explain why the *k*-NN results estimated for post-precipitation days were insensitive to the number of SWE-reconstruction scenes.

4. Discussion

4.1 Importance of ground measurements

The American River basin wireless-sensor networks provided unprecedented spatially distributed snowpack observations. The spatio-temporal data allow us to evaluate existing spatial SWE products and enable real-time spatial SWE estimation. The comparison between the sensor-network observations and SWE reconstructions over time (Figure 5, S2) suggests that the SWE reconstructions we are using are biased over the sensor locations. However, the bias was reduced in the *k*-NN estimates that used SWE reconstructions as training data. This emphasizes that

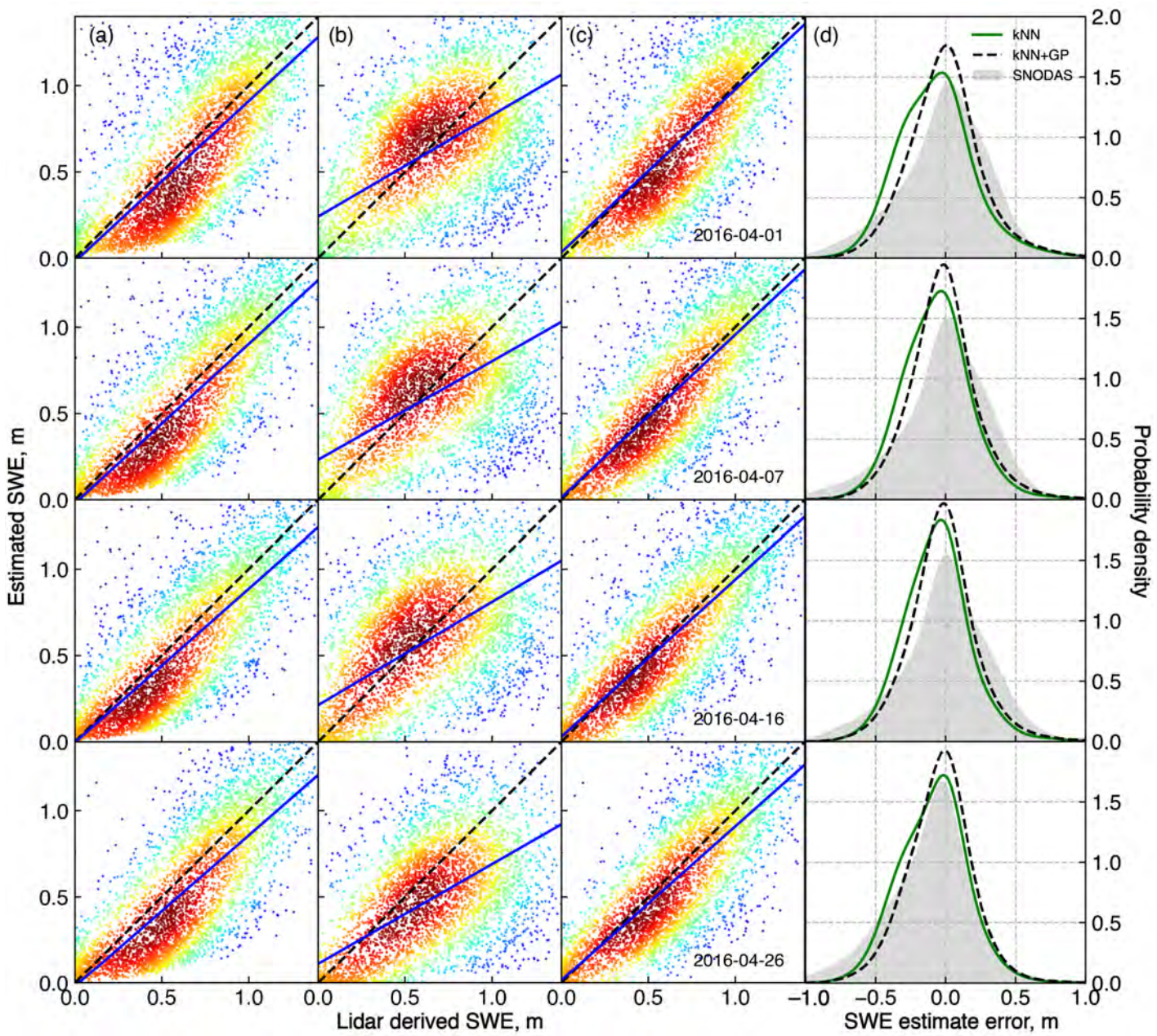

Figure 7. (a) *k*-NN (b) SNODAS and (c) *k*-NN+GP estimated snow water equivalent versus Lidar derived snow water equivalent on four dates during the peak season in 2016 at the Tuolumne River basin, with color showing the point density (red is high density and blue is low density). (d) Snow water equivalent estimation error distribution shows that *k*-NN+GP is the best spatial estimation method.

accurate and spatially distributed ground measurements are important for correcting the error when the historical training data are biased. The more ground measurement locations we have, the less error that the *k*-NN estimates would have, even though it is used with a biased training data set (Figure S5).

The ground measurements are also important for residual distribution. Although the ground measurements used in the Tuolumne-Merced experiment were simulated from the Lidar products, the results that are shown in Table 4 suggest that spatially dense and representative ground measurements are important when systematic bias exist in the training data and non-removable from the *k*-NN method. The bias in the *k*-NN estimates can be corrected systematically with a Gaussian process regression model trained from the residuals over the ground measurement locations.

The importance of the ground measurements is also relative to the quality of the historical training data. Comparing the analysis from using SWE reconstructions versus Lidar, we believe there is a trade-off between the quality of the historical training data and the amount of ground measurements that are needed. With more Lidar scans or more accurate historical SWE products available, we may not need to have that many sensor locations for accurate *k*-NN estimates and the residual distribution procedure may not be needed.

4.2 Bias and uncertainties in the *k*-NN SWE estimates

In general, the *k*-NN approach intrinsically preserves the autocorrelation structure of the spatial SWE, which was also found and reported in other applications of the *k*-NN method (Gangopadhyay and Clark (2005); Prairie et al. (2007)). These studies also found that the *k*-NN method

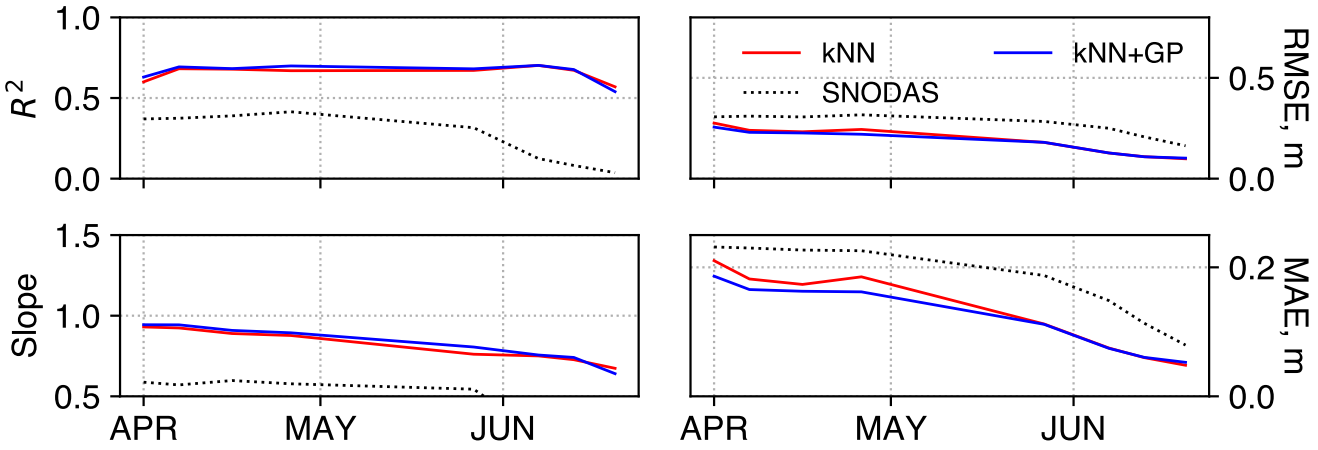

Figure 8. Slopes and coefficients of determination estimated from linear regressions between SWE estimations vs. Lidar, RMSE and MAE of SWE estimations and Lidar at Tuolumne basin in 2016. Statistics for other years and basins are shown in Figure S3.

Table 4. Statistics estimated from predicted SWE and Lidar-derived SWE using historical Lidar and SNODAS as training set.

Date	Training Data	+GP	MAE, m	MBE, m	RMSE, m	$RMSE_s$, m	$RMSE_{us}$, m	Slope	R^2
2016-04-01	SNODAS		0.24	-0.08	0.31	0.15	0.27	0.61	0.37
		✓	0.20	**-0.02**	0.26	0.09	0.24	0.73	0.51
	Lidar		0.43	0.43	0.48	0.48	0.05	0.34	**0.82**
		✓	**0.11**	-0.05	**0.15**	0.05	0.14	**0.93**	**0.82**
2016-04-07	SNODAS		0.23	0.03	0.30	0.17	0.24	0.51	0.36
		✓	0.22	**0.02**	0.29	0.16	0.24	0.54	0.39
	Lidar		0.38	0.38	0.45	0.45	0.04	0.33	**0.89**
		✓	**0.10**	-0.03	**0.14**	0.05	0.13	**0.88**	0.85
2016-04-16	SNODAS		0.23	0.07	0.31	0.19	0.24	0.49	0.34
		✓	0.20	0.04	0.27	0.14	0.23	0.62	0.48
	Lidar		0.34	0.34	0.42	0.42	0.04	0.34	**0.89**
		✓	**0.11**	**-0.02**	**0.15**	0.05	0.14	**0.88**	0.84
2016-04-26	SNODAS		0.22	0.09	0.30	0.19	0.23	0.52	0.39
		✓	0.21	0.08	0.28	0.17	0.22	0.56	0.43
	Lidar		0.37	0.37	0.44	0.44	0.03	0.34	**0.92**
		✓	**0.09**	**-0.02**	**0.13**	0.05	0.12	**0.87**	0.87

MBE: mean-bias error
$RMSE_s$: systematic root-mean-square error
$RMSE_{us}$: unsystematic root-mean-square error

does not fully preserve the temporal autocorrelation structure, which was not observed in our results (not shown) and the finding can be specific to particular applications. The unexplained spatial variabilities are mostly due to the bias and inaccuracies of the historical training data, which can be decomposed into the following aspects:

1. As was discussed in Section 2.5, multiple error sources can contribute to the bias in SWE reconstructions; and thus affect the estimation accuracy of the *k*-NN interpolation method. Since we can use a Gaussian-process regression model to correct the bias in the *k*-NN SWE estimates, the bias is likely due to physiographic features affecting the SWE reconstructions.

2. The MODIS f_{SCA}, which was the basic of the SWE-reconstruction products that we used, was derived using the method in Painter et al. (2009). The local zenith for f_{SCA} can be significantly affected by the subpixel topographic variability, which can cause errors in the estimate of both fractional snow cover and derived snow properties. In densely vegetated area, quantifying fractional snow cover is beyond our capacity of using MODIS and we may only detect snow's presence or absence, which contributes uncertainties in SWE reconstructions over forested areas, with the bias accumulating throughout the entire modeling season (Raleigh et al. (2013)). Also, MODIS has the detection capacity down to 10−15% of fractional snow covered area; and thus the inferred final date of seasonal from MODIS snow coverage can be earlier than the actual melt-out date (Painter et al. (2009); Liu et al. (2008); Rice et al. (2011)).

3. Other than the snow-covered-area images, errors can also be caused by the spatial meteorological data that are used for estimating the heat flux that drives the snowmelt. These data have a much larger spatial scale than the SWE-reconstruction model, and downscaling is often performed by predicting values at the smaller scale using a statistical model that takes topographic variables as predicting features. Since the model is trained using the coarser-scale data, applying the model at a different scale, especially a smaller scale can cause errors in the estimates (Hostetler (1994); Fowler and Wilby (2007)).

4. In addition to these snowmelt estimation errors in most SWE reconstructions, the SWE-reconstruction model used in this study does not capture precipitation events during the melt season (Guan et al. (2013)).

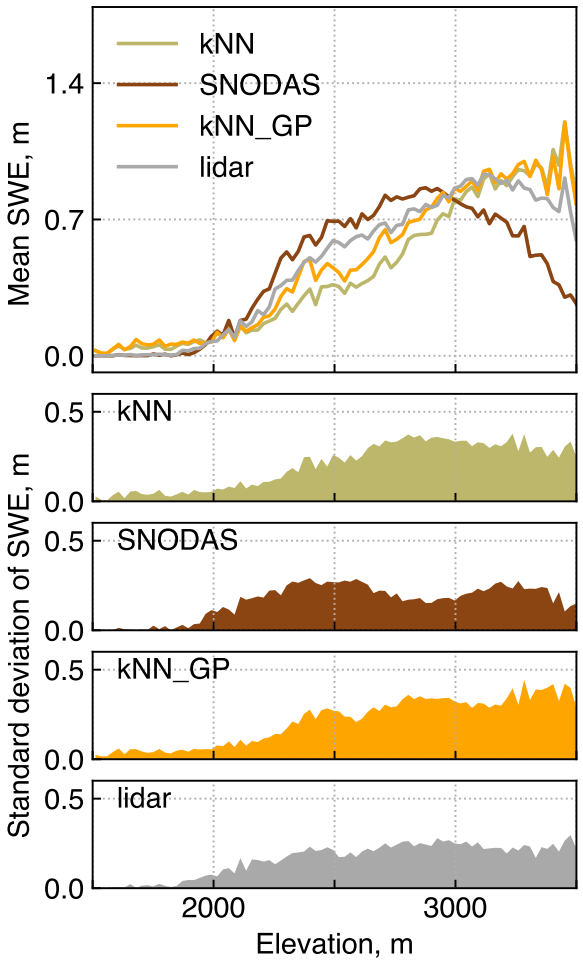

Figure 9. The Tuolumne basin's snow water equivalent mean and standard deviation along the elevation gradient.

Warm winds usually come with rain-on-snow events and they can drive snowpack to melt faster by altering the energy balance (Garvelmann et al. (2014)) and it is more likely to happen in the rain-snow transition zones (1500−2000 m in elevation) in the Sierra Nevada (Lundquist et al. (2008)). Not considering precipitation events can greatly affect the accuracy of the estimated snow distributions at different elevation ranges.

Comparing to the analysis using the SWE reconstructions, the accuracy of the estimates using historical Lidar as training set improved significantly (Figure 10). Most reported error statistics (Table 4) were greatly reduced in the k-NN estimates if the GP was applied. This suggests that the GP model is effective on correcting the systematic errors in the k-NN estimation. Erratically, the coefficients of determination (R^2s) are generally greater in the k-NN estimates that are without the GP corrections comparing to these with the GP corrections, suggesting that the spatial distribution of SWE is stationary, such that the real-time SWE distribution can be accurately quantified by blending the real-time ground measurements and historical Lidar scans.

The uncertainties of the estimates that use SNODAS as training data are much greater than that using either Lidar or SWE reconstructions. The statistics of the k-NN analysis suggest that the historical SNODAS SWE estimates cannot capture the spatial pattern of the SWE distribution.The GP does not boost the performance, as observed in the analysis using historical Lidar as training data, which suggests that the residuals cannot be related with the physiographic variables and most of the errors are unsystematic and are difficult to be corrected.

Comparing to linear-regression-based interpolation methods, the RMSEs of the SWE estimates from k-NN+GP that uses historical Lidar as training set are all below 15 cm, which are similar to the errors reported in Fassnacht et al. (2003). We need to notice that the errors reported in our study could be more credible because the results are evaluated with a spatial data set, while previous studies only reported cross-validation errors. Also, most previous interpolation studies use SNOTEL data in the Rocky Mountains as ground observations, which can be biased training samples of the studied area because the SNOTEL sites in the Rocky Mountains can overestimate surrounding SWE by up to 200% (Molotch and Bales (2005)). Considering that the total snow accumulation in the 2016 water year is about the average of the long-term historical records, an acceptable spatio-temporal SWE estimation should have RMSEs below 20 cm.

4.3 Potential of the *k*-NN SWE estimation approach

Considering that SWE reconstructions can capture the spatial patterns of SWE, plus the high density of scenes available for these daily data, it can be an adequate data set to use with the k-NN method even though SWE reconstructions underperform in estimating the absolute magnitude of SWE. However, the experiment introduced in Section 2.6 and 3.2 suggests that the k-NN approach does not have to be limited to using SWE reconstructions as the training set, and any improvements that lead to better SWE distribution estimates can help in advancing the k-NN estimates. Other time series of historical spatial SWE data can be used as the training set, these include:

1. A more-accurate SWE-reconstruction data set produced by finer-resolution meteorological data and better f_{SCA} product processed by a more-advanced snow-coverage-mapping algorithm.

2. A new archive of Lidar-derived SWE maps. We observed a performance boost with using only 10 scenes of historical Lidar instead of historical SWE reconstructions. Although those scenes were collected during the recent drought, we can see they are able to help predict spatial SWE in a normal SWE year. It will be also interesting to see how much more the k-NN approach can be improved when we have a longer series of Lidar-derived SWE maps.

3. A new satellite-based SWE mapping product. We envision that in the future there will be Lidar systems helping to map the land surface from the space, in real time. Considering they can be easily affected by cloud cover, we can apply the k-NN method with their historical record to fill in the gap and better serve real-time water-resources decision making.

We summarized the existing spatial SWE products for the Sierra Nevada that were introduced in previous studies and can be used for studying the spatial distribution of SWE in mountainous area (Table 5). The information provided in these data sets are worth studying for their rich content of spatio-temporal distribution of SWE, which can potentially motivate new techniques being developed for real-time spatial SWE estimation and water-resources decision making. Also, with similar techniques as k-NN, the number of real-time Lidar flights can be greatly reduced during the winter season and thus reduce the financial cost to water agencies. With changes in global climate and increased variability of regional climate in the mountains, it requires having techniques that are data driven and model assisted, like the k-NN approach, for providing more-accurate and timely spatial SWE estimates. Thus more-precise and

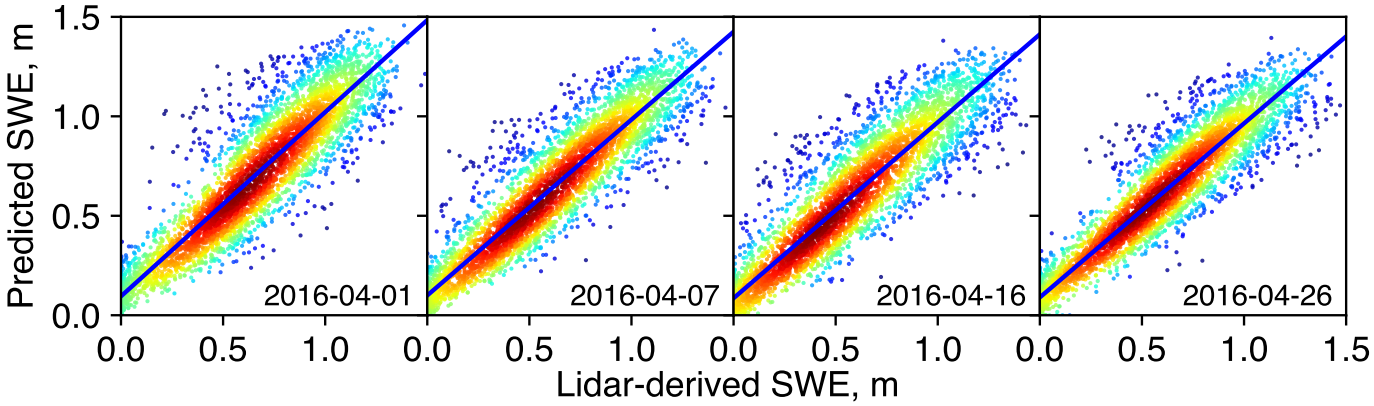

Figure 10. ***k*-NN+GP estimated snow water equivalent versus Lidar derived snow water equivalent on four dates during the peak season in 2016 at Tuolumne River basin, using the 2014 Lidar-derived SWE as training set, with color showing the point density (red is high density and blue is low density).**

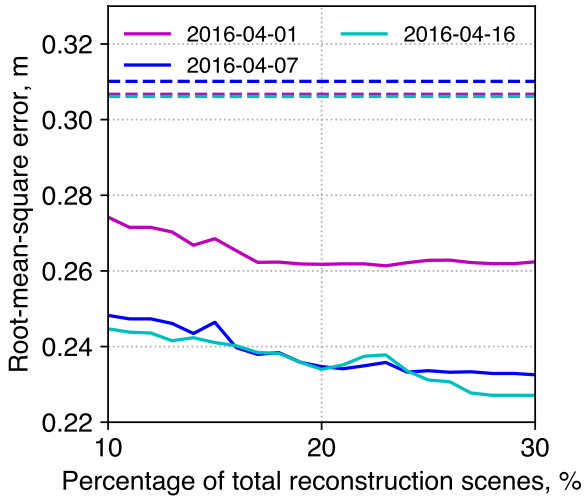

Figure 11. Solid lines are root-mean-square error versus the percentage of total number of SWE-reconstruction scenes (1988 scenes) that were included in the *k*-NN method. The errors are estimated from the results of the analysis for the 2016 Tuolumne basin data. The dashed lines are the errors of SNODAS at the same dates, which are used as baselines of prediction performance. The RMSEs decrease saturates around 30% of total number of SWE-reconstruction scenes. RMSEs and MAEs for more days in 2014 for the Merced basin are shown in Figure S4.

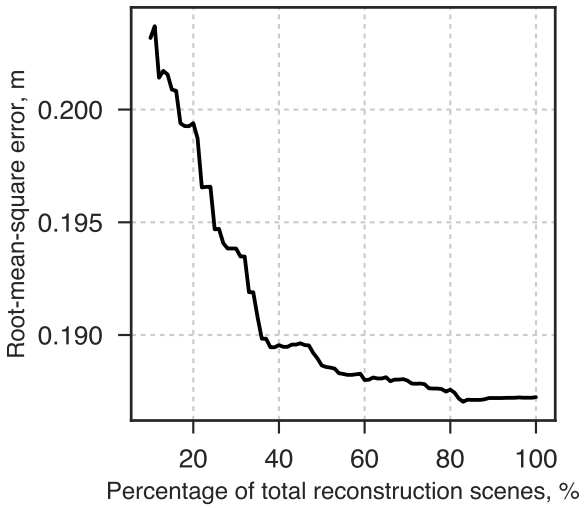

Figure 12. Root-mean-square error versus the percentage of total number of SWE-reconstruction scenes (1988 scenes) that were included in the *k*-NN method. The errors are estimated from the results of the analysis for the 2014 American River basin data.

efficient water-resources management can be achieved for coping with the future challenges.

5. Conclusions

k-NN and *k*-NN+GP are effective approaches to blend ground-measured snowpack and remotely sensed snow-coverage information to estimate spatial snow water equivalent. The *k*-NN approach, using spatially representative snow-depth sensor data, built on historical spatial SWE estimates, can provide a near-real-time spatial estimate of SWE. The residual-distribution analysis using the Gaussian-process model indicated that the estimation errors from the *k*-NN method depend on the topography of each modeled pixel. By using the Gaussian-process model to distribute the errors spatially, estimation errors can be reduced by about 5−10% during the peak season. The *k*-NN and *k*-NN+GP simulation results using Lidar data from the Merced and Tuolumne basins reached $R^2 = 0.7$ for most modeled days during a normal snow season in 2016, with RMSEs $<$ 15 cm ($<$ 25 cm) and MAEs $<$ 10 cm ($<$ 20 cm) in 2014 (2016). The comparison with SNODAS indicates that these two approaches are able to provide a more-accurate estimation, regardless of climate conditions (2014 is a drought year; 2016 is a normal year). The decreasing RMSE with increasing number of SWE reconstructions suggests that the more historical information in a training set, the better quality the *k*-NN can achieve. The *k*-NN method can also be significantly improved as more-accurate historical spatial-SWE products, such as Lidar snow-depth products, become available.

Acknowledgments

The work presented in this paper is supported by the UC Office of the President's Multi-Campus Research Programs and Initiatives (MR-15-328473) through the UC Water Security and Sustainability Research

Table 5. Spatial SWE products information

Type	Resolution	Year range	Remote-sensing data source	Literature
Reanalysis	100 m	1985-2015	Landsat	Margulis et al. (2016)
Reconstruction	500 m	2000-2016	MODIS	Dozier et al. (2016)
Reconstruction	500 m	2000-2016	MODIS	Rittger et al. (2016)
Regression	5 km	1950-2003 (1 April only)	N/A	Howat and Tulaczyk (2005)
Reanalysis	1/16 deg	1915-2011	N/A	Livneh et al. (2013)

Initiative (grant no. 13941). We acknowledge the National Science Foundation (NSF) through a Major Research Instrumentation Grant (EAR-1126887) and the Southern Sierra Critical Zone Observatory (NSF Award Numbers 1331939 and 1239521). We also acknowledge Ziran Zhang and Sami Malek's assists on data acquisition and Tessa Maurer's help on literature refinement, and NASA Airborne Snow Observatory for providing the Lidar data.

References

Bair, E. H., Rittger, K., Davis, R. E., Painter, T. H., and Dozier, J. (2016). Validating reconstruction of snow water equivalent in California's Sierra Nevada using measurements from the NASA Airborne Snow Observatory. *Water Resources Research*, pages 1943–1959.

Bales, R. C., Molotch, N. P., Painter, T. H., Dettinger, M. D., Rice, R., and Dozier, J. (2006). Mountain hydrology of the western United States. *Water Resources Research*, 42(8).

Barrett, A. P. (2003). National Operational Hydrologic Remote Sensing Center SNOw Data Assimilation System (SNODAS) Products at NSIDC. *NSIDC Special Report 11*, (Natl. Snow and Ice Data Cent.: Boulder, CO):19.

Brun-Laguna, K., Oroza, C., Zhang, Z., Malek, S., Watteyne, T., and Glaser, S. (2016). Sierranet: Monitoring the snowpack in the sierra nevada: Demo. In *Proceedings of the Eleventh ACM Workshop on Challenged Networks*, CHANTS '16, pages 33–34, New York, NY, USA. ACM.

California Department of Water Resources (2013). California' s Flood Future: Recommendations for Managing the State's Flood Risk. Technical Report November.

Cayton, L. (2009). Efficient bregman range search. In Bengio, Y., Schuurmans, D., Lafferty, J. D., Williams, C. K. I., and Culotta, A., editors, *Advances in Neural Information Processing Systems 22*, pages 243–251. Curran Associates, Inc.

Cline, D. W., Bales, R. C., and Dozier, J. (1998). Estimating the spatial distribution of snow in mountain basins using remote sensing and energy balance modeling. *Water Resources Research*, 34(5):1275.

Clow, D. W., Nanus, L., Verdin, K. L., and Schmidt, J. (2012). Evaluation of SNODAS snow depth and snow water equivalent estimates for the Colorado Rocky Mountains, USA. *Hydrological Processes*, 26(17):2583–2591.

Dozier, J., Bair, E. H., and Davis, R. E. (2016). Estimating the spatial distribution of snow water equivalent in the world's mountains. *Wiley Interdisciplinary Reviews: Water*, 3(3):461–474.

Dozier, J., Painter, T. H., Rittger, K., and Frew, J. E. (2008). Time-space continuity of daily maps of fractional snow cover and albedo from MODIS. *Advances in Water Resources*, 31(11):1515–1526.

Fassnacht, S. R., Dressler, K. A., and Bales, R. C. (2003). Snow water equivalent interpolation for the Colorado River Basin from snow telemetry (SNOTEL) data. *Water Resources Research*, 39(8).

Fowler, H. J. and Wilby, R. L. (2007). Beyond the downscaling comparison study. *International Journal of Climatology*, 27(12):1543–1545.

Gangopadhyay, S. and Clark, M. (2005). Statistical downscaling using K-nearest neighbors. *Water Resources Research*, 41(W02024).

Garvelmann, J., Pohl, S., and Weiler, M. (2014). Variability of Observed Energy Fluxes during Rain-on-Snow and Clear Sky Snowmelt in a Midlatitude Mountain Environment. *Journal of Hydrometeorology*, 15(3):1220–1237.

Guan, B., Molotch, N. P., Waliser, D. E., Jepsen, S. M., Painter, T. H., and Dozier, J. (2013). Snow water equivalent in the Sierra Nevada: Blending snow sensor observations with snowmelt model simulations. *Water Resources Research*, 49(8):5029–5046.

Hedrick, A., Marshall, H.-P., Winstral, A., Elder, K., Yueh, S., and Cline, D. (2015). Independent evaluation of the SNODAS snow depth product using regional-scale lidar-derived measurements. *The Cryosphere*, 9(1):13–23.

Hostetler, S. W. (1994). Hydrologic and atmospheric models: The (continuing) problem of discordant scales. *Climatic Change*, 27(4):345–350.

Howat, I. M. and Tulaczyk, S. (2005). Climate sensitivity of spring snowpack in the Sierra Nevada. *Journal of Geophysical Research: Earth Surface*, 110(F4):n/a–n/a.

Kalra, A. and Ahmad, S. (2011). Evaluating changes and estimating seasonal precipitation for the Colorado River Basin using a stochastic nonparametric disaggregation technique. *Water Resources Research*, 47(W05555).

Larose, D. T. (2005). k -Nearest Neighbor Algorithm. In *Discovering Knowledge in Data*, pages 90–106. John Wiley & Sons, Inc., Hoboken, NJ, USA.

Li, X.-J., Ma, X.-J., and Zhang, J.-X. (2016). Robust feature screening for varying coefficient models via quantile partial correlation. *Metrika*, pages 1–33.

Liu, J., Li, R., and Wu, R. (2014). Feature Selection for Varying Coefficient Models With Ultrahigh-Dimensional Covariates. *Journal of the American Statistical Association*, 109(505):266–274.

Liu, J., Woodcock, C. E., Melloh, R. A., Davis, R. E., McKenzie, C., and Painter, T. H. (2008). Modeling the view angle dependence of gap fractions in forest canopies: Implications for mapping fractional snow cover using optical remote sensing. *Journal of Hydrometeorology*, 9(5):1005–1019.

Livneh, B., Rosenberg, E. A., Lin, C., Nijssen, B., Mishra, V., Andreadis, K. M., Maurer, E. P., and Lettenmaier, D. P. (2013). A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States: Update and Extensions*. *Journal of Climate*, 26(23):9384–9392.

Lundquist, J. D., Neiman, P. J., Martner, B., White, A. B., Gottas, D. J., and Ralph, F. M. (2008). Rain versus Snow in the Sierra Nevada, California: Comparing Doppler Profiling Radar and Surface Observations of Melting Level. *Journal of Hydrometeorology*, 9(2):194–211.

Margulis, S. A., Cortés, G., Girotto, M., and Durand, M. (2016). A Landsat-Era Sierra Nevada Snow Reanalysis (1985–2015). *Journal of Hydrometeorology*, 17(4):1203–1221.

Marks, D., Domingo, J., Susong, D., Link, T., and Garen, D. (1999). A spatially distributed energy balance snowmelt model for application in mountain basins. *Hydrological Processes*, 13(12-13):1935–1959.

Molotch, N. P. and Bales, R. C. (2005). Scaling snow observations from the point to the grid element: Implications for observation network design. *Water Resources Research*, 41(11):1–16.

Molotch, N. P., Guan, B., and Lestak, L. R. (2017). Snow Water Equivalent (SWE) for Water Supply and Management in California. INSTAAR and NASA's Jet Propulsion Laboratory, University of Colorado at Boulder, Colorado, USA. Digital media available by request.

Molotch, N. P. and Margulis, S. a. (2008). Estimating the distribution of snow water equivalent using remotely sensed snow cover data and a spatially distributed snowmelt model: A multi-resolution, multi-sensor comparison. *Advances in Water Resources*, 31(11):1503–1514.

Ni, K. and Nguyen, T. (2009). An Adaptable k-Nearest Neighbors Algorithm for MMSE Image Interpolation. *IEEE Transactions on Image Processing*, 18(9):1976–1987.

Oroza, C. A., Zheng, Z., Glaser, S. D., Tuia, D., and Bales, R. C. (2016). Optimizing embedded sensor network design for catchment-scale snow-depth estimation using LiDAR and machine learning. *Water Resources Research*, 52(10):8174–8189.

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

Painter, T. H., Rittger, K., McKenzie, C., Slaughter, P., Davis, R. E., and Dozier, J. (2009). Retrieval of subpixel snow covered area, grain size, and albedo from MODIS. *Remote Sensing of Environment*, 113(4):868–879.

Prairie, J., Rajagopalan, B., Lall, U., and Fulp, T. (2007). A stochastic nonparametric technique for space-time disaggregation of streamflows. *Water Resources Research*, 43(W03432).

Rajagopalan, B. and Lall, U. (1999). A k -nearest-neighbor simulator for daily precipitation and other weather variables. *Water Resources Research*, 35(10):3089–3101.

Raleigh, M. S., Rittger, K., Moore, C. E., Henn, B., Lutz, J. a., and Lundquist, J. D. (2013). Ground-based testing of MODIS fractional snow cover in subalpine meadows and forests of the Sierra Nevada. *Remote Sensing of Environment*, 128:44–57.

Rice, R., Bales, R. C., Painter, T. H., and Dozier, J. (2011). Snow water equivalent along elevation gradients in the merced and tuolumne river basins of the sierra nevada. *Water Resources Research*, 47(8). W08515.

Rittger, K. (2012). *Spatial estimates of snow water equivalent in the Sierra Nevada*. PhD thesis, University of California, Santa Barbara.

Rittger, K., Bair, E. H., Kahl, A., and Dozier, J. (2016). Spatial estimates of snow water equivalent from reconstruction. *Advances in Water Resources*, 94:345–363.

Rittger, K., Painter, T. H., and Dozier, J. (2013). Assessment of methods for mapping snow cover from MODIS. *Advances in Water Resources*, 51:367–380.

Rosenthal, W. and Dozier, J. (1996). Automated mapping of montane snow cover at subpixel resolution from the Landsat Thematic Mapper. *Water Resources Research*, 32(1):115–130.

Schneider, D. and Molotch, N. P. (2016). Real-time estimation of snow water equivalent in the Upper Colorado River Basin using MODIS-based SWE Reconstructions and SNOTEL data. *Water Resources Research*, 52(10):7892–7910.

Slater, A. G., Barrett, A. P., Clark, M. P., Lundquist, J. D., and Raleigh, M. S. (2013). Uncertainty in seasonal snow reconstruction: Relative impacts of model forcing and image availability. *Advances in Water Resources*, 55:165–177.

Willmott, C. J. (1982). Some Comments on the Evaluation of Model Performance. *Bulletin of the American Meteorological Society*, 63(11):1309–1313.

Zhang, Z., Glaser, S. D., Bales, R. C., Conklin, M., Rice, R., and Marks, D. G. (2017). Technical report: The design and evaluation of a basin-scale wireless sensor network for mountain hydrology. *Water Resources Research*, 53(5):4487–4498.

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

Zhu, X., Vondrick, C., Fowlkes, C. C., and Ramanan, D. (2016). Do We Need More Training Data? *International Journal of Computer Vision*, 119(1):76–92.