

1 **Estimating relationships between snow water equivalent, snow**
2 **covered area, and topography to extend the Airborne Snow**
3 **Observatory dataset**

4 Dominik Schneider^{1,2}, Noah P. Molotch^{1,2,3,*}, Jeffrey S. Deems⁴

5 Correspondence to: noah.molotch@colorado.edu

6 ¹Dept. of Geography, Guggenheim 110, 260 UCB, Boulder Colorado 80309, USA

7 ²Institute of Arctic and Alpine Research, University of Colorado Campus Box 450, Boulder, CO 80309, USA

8 ³Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA

9 ⁴National Snow and Ice Data Center, CIRES, 449 UCB, University of Colorado, Boulder, CO, 80309, USA

10
11 **Abstract**

12 A new spatio-temporal dataset from the ongoing Airborne Snow Observatory (ASO) provides an unprecedented look
13 at the spatial and temporal patterns of snow water equivalent (SWE) over multiple years in the Tuolumne Basin,
14 California, USA. We found that fractional snow covered area (fSCA) significantly improves our ability to model the
15 distribution of SWE based on relationships between SWE, fSCA, and topography. Further, the broad availability of
16 satellite images of fSCA facilitates the transfer of these relationship to different years with minimal degradation in
17 performance compared with models fit on the same day, by considering variations in SWE depth as expressed by
18 differences in fSCA between years. The crux of this proposition is in selecting the model to transfer. We offer a
19 method with which to select a model from another year based on the similarity in SWE distribution at existing snow
20 pillows in the area. the selected predictions exhibit a mild decrease in r^2 (0.02) and moderate increases in %MAE
21 (15%) and %Bias (10%) from the best transferred predictions ($r^2=0.83$, %MAE=33%, %Bias=1%). The results
22 motivate further refinement in the technique used to select the *best model* because if these dates can be identified then
23 SWE can be modeled at accuracy levels equivalent to models generated from ASO data collected on the day of interest.
24 Lastly, we found that models from ASO observations of anomalously low snowpacks in 2015 still transferred to other
25 years, although the same cannot be said for the reverse. This research provides a first attempt at extending the value
26 of ASO beyond the observations and we expect ASO will continue to provide insights for improving water resource
27 management for years to come.
28

Deleted: ($r^2=0.85$, %MAE=33%, %Bias=1%)

Deleted: . Comparison of the best transferred predictions and

Deleted: results in

32 1 Introduction

33 The spatial distribution of snow water equivalent (SWE) largely controls the timing and magnitude of streamflow
34 (Stewart et al., 2004) and is important for ecology (Giersch et al., 2016; Litaor et al., 2008) in snow-dominated
35 catchments around the world. In these catchments, accurate assessments of snowpack water storage are critical for
36 ensuring robust estimates of seasonal water supply. Nevertheless, SWE is poorly measured operationally with only
37 sparse point measurements on the order of one measurement location per thousand square kilometers of snow covered
38 terrain in the western United States (Meromy et al., 2012; Rice and Bales, 2010). Additional manual measurements in
39 the form of snow courses add little information about the spatial variability and typically occur on the order of only
40 three times a season (López-Moreno et al., 2013).

41 Satellite remote sensing provides spatially explicit information about the snowpack but current satellites are
42 unable to measure SWE directly at the scales relevant for water resources management in mountainous terrain (e.g.
43 the western United States) (Dozier, 2011). The need for improved information regarding the quantity and distribution
44 of SWE has led to the development of new measuring techniques including the application of ground penetrating radar
45 (GPR) (Marshall and Koh, 2008), Global Positioning Systems (GPS) (Gutmann et al., 2012; Koch et al., 2014), Light
46 Detection and Ranging (lidar) (Deems et al., 2013; Prokop, 2008; Schirmer et al., 2011), and photogrammetry (Bühler
47 et al., 2015; Nolan et al., 2015). Although these techniques are typically limited to snow depth, snow depth can be
48 converted to SWE using modeled snow density or in situ snow pit observations (Elder et al., 1991; Jonas et al., 2009;
49 Painter et al., 2016; Sturm et al., 2010).

50 Airborne systems have significantly improved the ability to measure SWE distribution at high spatial
51 resolution and at extents relevant to water resource management (i.e. > 100 km²). However, most previous studies of
52 snow distribution using airborne data have been limited to snap-shots in time, limiting the ability to empirically
53 transfer observations to time periods outside of those directly sampled. Since 2013, the National Aeronautics and
54 Space Administration (NASA), Jet Propulsion Laboratory, Airborne Snow Observatory (ASO) has acquired weekly
55 observations of snow properties from approximately the time of annual peak SWE to the end of the snowmelt season
56 in the Tuolumne Basin, California totaling 28 flights from 2013-2016 (Painter et al., 2016). ASO measures snow depth
57 by differencing the lidar-derived elevations from snow-on and snow-off flights and infers albedo and snow extent
58 based on spectroradiometric measurements. Density is computed using the iSNOBAL energy-balance model (Marks
59 et al., 1999) based on inferred albedo, local meteorological measurements, and constrained by in situ snow pillow
60 observations and manual snow course observations (Painter et al., 2016). Estimated snowpack density is multiplied
61 by the lidar-derived snow depth to estimate SWE. These weekly observations of SWE distribution over multiple years
62 represent a new opportunity for understanding the spatial and temporal dynamics of snow distribution. Moreover, the
63 intensive repeat sampling of ASO may provide an opportunity to extend observed SWE patterns beyond the time
64 periods directly observed by ASO – a goal of the work presented here. In this context, the ability to extend expensive
65 ASO data in time could dramatically expand the applicability of ASO data to future time periods without incurring
66 the costs associated with future airborne acquisitions.

67 A potential application of ASO measurements relates to the possibility of developing statistical relationships
68 between ASO data and other snow and terrain data that are more routinely available. Statistical models of this type

Deleted: plant

Deleted: .

Deleted: they can still capture the majority of the variability in SWE because snow depth varies an order of magnitude greater than density. Snow depth distributions

Deleted: .

Deleted: surfaces

Deleted: a

Deleted: flight

Deleted: An

Deleted: is used to estimate

Deleted: , and subsequently convert

Deleted: In this context,

82 have been extensively used to estimate relationships between snow point measurements and physiography, though
 83 previous efforts were applied to manual SWE data sets of limited spatial coverage (Balk and Elder, 2000; Elder et al.,
 84 1998; Erickson et al., 2005; Fassnacht et al., 2003; López-Moreno and Nogués-Bravo, 2006; Molotch et al., 2005;
 85 Schneider and Molotch, 2016). At small scales (i.e. < 10 km²), dedicated sampling of headwater catchments has led
 86 to models that explain between 20% and 65% of the variability in snow depth based on physiographic variables at
 87 30m resolution (Balk and Elder, 2000; Elder et al., 1998; Erxleben et al., 2002). Importantly, Erickson et al. (2005)
 88 found persistence in the topographic controls on snow depth distribution and successfully parameterized a multi-year
 89 model relating the physiographic variables of a small headwater catchment to annual peak snow depth by scaling the
 90 mean based on in situ measurements. Later studies have since confirmed an inter-annual consistency in snow depth
 91 distribution based on high resolution lidar measurements (Deems et al., 2008; Schirmer et al., 2011; Trujillo et al.,
 92 2009) and the inter-annual persistence of topographic controls on snow depth distribution (Grünewald et al., 2013;
 93 Revuelto et al., 2014). These previous works suggest that relationships between lidar snow depth measurements and
 94 topography may be useful for extending lidar measurements of snow in time.

95 The lidar measurements provide snow depth, and coincident density measurements or estimates are required
 96 to estimate SWE. However, snow depth measurements still capture the majority of the variability in SWE because
 97 snow depth varies an order of magnitude greater than density (Jonas et al., 2009; López-Moreno et al., 2013; Mizukami
 98 and Perica, 2008). Hence, we expect the results from the previous studies be relevant here, i.e. inter-annual consistency
 99 in SWE distribution will be similar to that of snow depth. Given the difficulty of extensively measuring density and
 100 SWE, operational SWE observation networks that use snow pillows to measure SWE have been used to relate
 101 physiography and SWE. In this regard, multiple studies have explained up to 82% of the variability in SWE based on
 102 physiographic variables (Fassnacht et al., 2003; Harshburger et al., 2010; Schneider and Molotch, 2016). These studies
 103 aimed to understand the processes controlling snow distribution and to apply this knowledge to interpolate point
 104 observations of SWE to >1000 km² for a single point in time. The snow pillow networks, however, are constrained to
 105 flat sites in a relatively narrow elevation band below treeline. The spatio-temporal dataset of SWE from ASO, which
 106 covers the full range of physiographic variables present in a 1,175 km² watershed, provides an unprecedented
 107 opportunity to develop relationships between SWE and topography and test their persistence across several years.

108 Given that topographic variables are largely static in time, additional time-variant variables should be useful
 109 in the context of explaining the spatio-temporal distribution of SWE. Remotely sensed snow covered area (SCA) has
 110 long been recognized to provide information relevant to snowpack water storage and consequently expected summer
 111 streamflow (Cline et al., 1998; Good and Martinec, 1987; Martinec and Rango, 1981; Molotch and Margulis, 2008;
 112 Potts, 1937; 1944). Currently, SCA is commonly estimated from SWE in hydrologic models through a depletion curve
 113 parameterization in order to constrain melt production to the areal extent of snow cover (Anderson, 1973; Clark et al.,
 114 2011; Lawrence et al., 2011; Livneh et al., 2010; Luce and Tarboton, 2004; Niu et al., 2011). The utility of depletion
 115 curves to provide sub-grid information in physically-based modeling suggests that fSCA should provide additional
 116 information in statistical models of SWE distribution. The assumption of a consistent relationship between fSCA and
 117 SWE is predicated on the fact that SWE distribution is extremely heterogeneous over complex terrain, and terrain
 118 features are progressively uncovered during the melt season. This process varies only slightly each year because of

Deleted: Lidar is only capable of measuring

Deleted: scarce, despite

Deleted: being the more important hydrologic variable.

Deleted: more

Deleted: and therefore largely controls the variability of SWE

Deleted: In this context,

Deleted: data

Deleted: with regard

Deleted: model-scale

Deleted: Upon melt out,

similarities in the meteorology, e.g. wind direction, that drive accumulation patterns and solar exposure that drives melt processes (Luce and Tarboton, 2004; Revuelto et al., 2016; Sturm and Wagner, 2010). Snow covered area is also relatively easy to measure due to the distinctive spectral signature of snow compared to soil, rock, and vegetation. In fact, photographic estimates of fSCA have been utilized for seven decades within hydrologic applications (Parsons and Castle, 1959; Potts, 1937; 1944) and today robust observations of fSCA can be obtained from variety of ground-based, aerial and satellite optical imagers (Bloschl et al., 1991; Dozier et al., 1981; Hall et al., 2001; Kimbaurer and Bloschl, 1994; König and Sturm, 1998; Painter et al., 2009; Rittger et al., 2013; Rosenthal and Dozier, 1996).

The repeat observations from ASO provide a unique dataset of concurrent SWE and fSCA over multiple years with which to develop robust statistical relationships between SWE, fSCA, and topography. Given that fSCA is widely observable from a variety of satellites, these relationships could then be used to estimate SWE for any date on which fSCA observations are available. Hence, the objective of this research is to use ASO-derived relationships between SWE (dependent variable) and fSCA and physiography (independent variables) to estimate SWE distribution for time periods when ASO data are not available. We aim to test how well statistical models of the relationship between SWE, fSCA and physiography transfer in time. We ask (1) Does fSCA improve statistical models of SWE distribution? (2) Can statistical models of SWE distribution be transferred directly from one year to another? (3) How can we determine which SWE distribution from the ASO record best represents a date of interest?

We present our SWE distribution modelling framework and show the utility of including the time-variant variable fSCA for improving the SWE distribution estimates. Next, we evaluate the impact of transferring models from one year to another. Lastly, we present a methodology for identifying which models of SWE distribution, from the ensemble of historical ASO acquisitions, best represents the SWE distribution for hypothetically unsampled dates of interest. We then discuss the results in the context of extending ASO to unsampled dates.

2 Site Description

We used a SWE dataset from the Tuolumne River basin in the Sierra Nevada mountains in California, USA (Fig. 1). The basin is 1,175 km² in area, consisting of 48% vegetation, 50% rock, 2% water, and small isolated areas with permanent snow/ice. The elevation range is 1127 m to 3965 m, encompassing 4 distinct ecological zones ranging from lower montane forest to alpine (NPS, 2016). The lower montane forest ranges from 1127 m to 1800 m elevation and consists of a diverse mix of coniferous and deciduous trees. The upper montane forest ranges from 1800 m to 2450 m elevation and primarily consists of coniferous species such as red fir (*Abies magnifica*) and lodgepole pine (*Pinus contorta*). Elevations from 2450 m to 2900 m are considered subalpine and consist of a mix of meadows and coniferous forest. The highest elevation band above 2900 m is an alpine zone that is devoid of tree cover and contains limited herbaceous vegetation. This alpine zone contains areas with large granitic features, talus slopes and boulder fields. Snowmelt from the basin runs off into the Hetch Hetchy reservoir, which is the main water supply for the City of San Francisco.

Analysis of SWE measurements from snow pillows show that historical mean peak SWE ranges from 0.3 m to 1.5 m with a mean of 0.8 m. Each of the study years was characterized by below average snowpack with 2014 and 2015 characterized as a severe dry snow drought (Harpold et al., 2017), experiencing only 36% and 35% of the

Deleted: ,

Deleted: out

Deleted:),

Deleted: Further

Deleted: .

climatological mean peak SWE, respectively. The years 2013 and 2016 also experienced below average snowpack conditions, but less severely, with the data from the snow pillows reporting 62% and 71% of climatological peak SWE, respectively. Typically, minimum temperatures range from -12°C in winter to 3°C in summer and maximum temperatures range from 4°C in winter to 22°C in summer (Cristea et al., 2017).

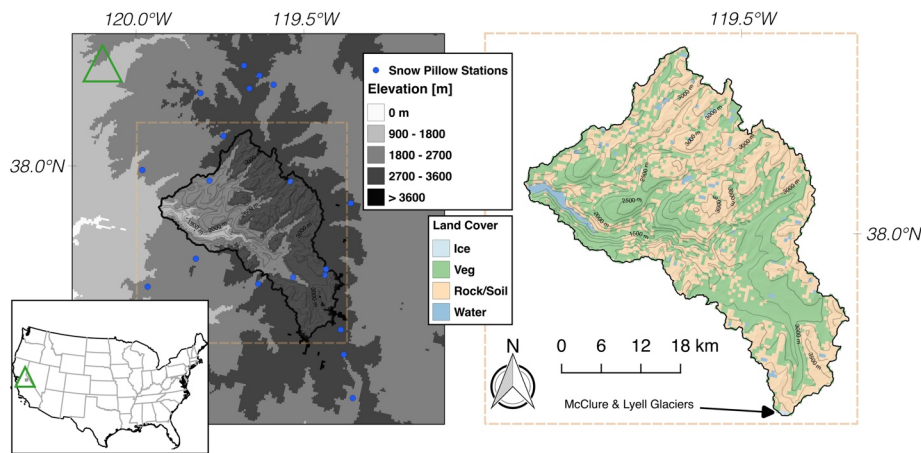


Figure 1. Elevation with snow pillows shown as blue dots (left) and land cover map of the Tuolumne Basin (right). The land cover data was provided by ASO and was derived from spectral information from summer snow-off flight. The snow pillows locations were obtained from the California Department of Water Resources. 250 m contour lines are shown on both maps.

3 Data Sources

We utilize the ASO dataset in the Tuolumne River Basin in the Sierra Nevada mountains of California. The ASO mission consists of airborne lidar in conjunction with a hyperspectral spectroradiometer (Painter et al., 2016). The dataset consists of 3 m snow depth maps for which snow-free areas are masked using spectral information from the spectroradiometer, 3 m vegetation height map and 50 m SWE maps. The dataset is distributed in the UTM zone 11 and WGS84 datum map projection system.

In order to conduct our analysis at a spatial scale that will ensure transferability to more widely available data, we mean-aggregated the 50 m SWE maps to 500 m. We chose the 500 m resolution because it is the scale for which commonly-used daily satellite fSCA products are available from the Moderate Resolution Imaging Spectroradiometer (MODIS) (Painter et al., 2009; Salomonson and Appel, 2004). Subsequently, we converted the 3 m snow depth maps to 501 m using mean aggregation and then bilinearly resampled to the 500 m SWE grid. We also used a binary aggregation of the 3 m snow depth maps to obtain 501 m fSCA maps, which were then resampled to 500 m by nearest neighbor to preserve snow free pixels. We used the Global Multi-resolution Terrain Elevation Data

Deleted: with snow pillows shown as blue dots.

Deleted: a 3 m resolution snow-free digital elevation model (DEM),

Deleted: DEM and

Deleted: We computed the physiographic variables used in the modelling framework [Table 1] from the 500 m DEM using open source GIS software, including GDAL, SAGA GIS, and R. We chose the 500 m resolution because it is relevant to water resource management and is the scale for which daily satellite fSCA images are available from the Moderate Resolution Imaging Spectroradiometer, which is a commonly used fSCA data product.

204 [2010 digital elevation model \(data available from the U.S. Geological Survey at <https://lta.cr.usgs.gov/GMTED2010>\)](https://lta.cr.usgs.gov/GMTED2010)
205 [reprojected to the 500 m SWE grid to compute the physiographic predictor variables \(Table 1\).](#)

Deleted: Lastly, we aggregated and bilinearly resampled the vegetation height map to 500 m.

207 **Table 1. List of physiographic variables considered in the multiple linear regression to model SWE distribution. Source**
208 **includes studies in which these variables have been used and the source of the algorithm, if applicable. Citations in the**
209 **source column are by no means exclusive.**

Variable	Units/Derivation Specifics	Source
UTM Northing	meters	Fassnacht et al. (2003)
UTM Easting	meters	Fassnacht et al. (2003)
elevation	meters	Elder et al., (1991, 1998); Balk and Elder (2000)
northness	$\sin(\text{slope}) \cdot \cos(\text{aspect})$; ranges 0-1; dimensionless	Fassnacht et al. (2003); Molotch et al. (2005)
eastness	$\sin(\text{slope}) \cdot \sin(\text{aspect})$; ranges 0-1; dimensionless	Fassnacht et al. (2003); Molotch et al. (2005)
topographic position index (TPI)	elevation difference of a pixel from the mean of the surrounding pixels using a 9x9 pixel window; meters	Reuelto et al. (2014); GDAL (2015)
vector ruggedness measure (VRM)	3-dimensional measure of the variation of slope and aspect in a 9x9 pixel window; not correlated with slope or aspect; ranges 0-1; dimensionless	Veitinger et al. (2015); Sappington et al. (2007); Conrad et al. (2015)
standard deviation of slope	standard deviation of slope in 7x7 window around each pixel; shown to detect changes in slope at multiple scales; radians	Marchand and Killingveit (2005); Lopez-Moreno et al. (2014); Grohman et al. (2007)
vegetation height	measured by ASO; used in place of forest canopy density from previous studies; meters	Molotch and Bales (2005, 2006); Painter et al. (2016)

211 The ASO flies approximately weekly from near peak SWE to the end of the melt season. As such, there were
212 six flights in 2013, ten in 2014, eight in 2015, and eight in 2016. Only the final four flights of the 2016 season were
213 available for this study resulting in a total of 28 SWE maps. [Painter et al. \(2016\)](#) report a mean absolute vertical snow
214 depth error of ≤ 8 cm and a bias of ≤ 1 cm when compared with manually measured snow depths at the 15×15 m scale.
215 Further details about the mission and processing can be found in Painter et al. (2016).

Deleted: The 2014 and 2015 seasons were characterized by a severe dry snow drought and 2013 and 2016 also experienced below average snowpack conditions, but less severely. Painter et al.

216 We also obtained daily SWE measurements from [the 17](#) snow pillows operated by the California Department
217 of Water Resources that are within 20 km of the Tuolumne watershed boundary. [We chose 20 km from a range of 10](#)
218 [km to 50 km because it produced the best results.](#) The stations range from 2000 m to 3250 m. We downloaded the
219 adjusted SWE records, which have been manually quality controlled, for the 2013-2016 water years. No further
220 adjustments were performed. The data can be downloaded from <http://cdec.water.ca.gov/>.

Deleted: 54

Deleted: <http://cdec.water.ca.gov/>.

221 4 Methods

222 We use linear regression to model the distribution of SWE for every ASO flight. The explanatory variables we consider
223 are ASO-observed fSCA and [topographic](#) variables previously used in the literature (Table 1). [The topographic](#)
224 [variables were directly computed from the 500 m digital elevation model with the exception of vegetation height](#)
225 [which was derived from the 3 m ASO vegetation height by mean aggregating to 501 m and bilinearly resampling to](#)
226 [500 m. Of note, we use three different windowed variables \(topographic position index \(TPI\), vector ruggedness index](#)

Deleted: physiographic

(VRM), and standard deviation of slope); to the best of our knowledge these have not previously been used for resolutions larger than 25 m. Comparison with ASO SWE at 500 m shows mean correlations of 0.3, -0.26, and -0.31, respectively, and we consider these variables indicative of preferential deposition from orographic updraft (Dadic et al., 2010; Garvert et al., 2007; Lehning et al., 2008). The window sizes were chosen based on the maximum correlation between ASO SWE and the variable, where the marginal increase in correlation from the next smaller window size was greater than 10%. We tested each odd window size from 3x3 to 15x15 pixels which equate to 1.5 km to 7.5 km scale, respectively.

We present results from two models: 1. “PHV” is a multiple linear regression that consists only of physiographic variables as independent variables. This represents the traditional approach for estimating SWE distribution based on relationships with physiography; 2. “PHV-FSCA” is a multiple linear regression that includes both physiographic variables and fSCA as independent variables. We make these distinctions to demonstrate the utility of fSCA as a time-variant variable, which has not yet been explicitly explored for its utility in directly estimating the distribution of SWE. fSCA values including zero are used to fit the PHV-FSCA models. Regression estimates from both PHV and PHV-FSCA are masked to snow covered areas as observed by ASO. Once we illustrate the utility of the statistical models for characterizing ASO SWE distribution patterns for a given day, we then explore how these statistical models can be transferred to time periods without ASO observations. With regard to all statistical models, we report the squared Pearson correlation coefficient (r^2) as a measure of the relative spatial pattern between the modeled SWE distribution and ASO observed SWE distribution. We also report the mean absolute error as a percent of mean observed SWE (%MAE) as a measure of the accuracy of the modeled SWE distribution. Lastly, we report bias as a percent of mean observed SWE (%Bias) as a measure of the systematic over or under-prediction by the model.

4.1 SWE Models: Same Day

To examine the utility of fSCA as a predictor, we compared the SWE distributions modeled with PHV and PHV-FSCA using a split sampling strategy; these are known as same-day models. We randomly split the data for each date into a training (80%) and test (20%) dataset to evaluate overall model performance on this date. This insures that we are not evaluating the model with the same data used to create the model. Furthermore, this procedure is replicated 20 times to provide 20 different subsets with which to evaluate model performance; this is more robust than a single replication. More replications were computationally prohibitive. This split sample strategy is an important initial step in transferring ASO data in time as it is necessary to first show that fSCA and physiographic variables can be used to adequately model ASO-observed SWE on the date of acquisition. Once this is established, the transferability of the models in time can then be explored – as described in the next section.

4.2 SWE Models: Transferred in Time

We evaluated a second set of predictions whereby each date is modeled using all the data, i.e. not split, and then this model is used to predict SWE on dates that ASO flew in different years. In this manner, we simulate SWE on the date ASO flew using models from other years and then we use the ASO data on the date of interest strictly to evaluate the

Deleted: . No one

Deleted: the

Deleted: of fSCA for

Deleted: greater than and equal to

Deleted: at discrete points in time

Deleted: show

Deleted: Discrete Time

Deleted: We split

278 model estimates of SWE. Hereafter, we refer to the date of the model (i.e. the date of the ASO observation for which
 279 the model is developed) as the *model date* and the date being predicted as the *transfer date*. This results in 28 models
 280 of SWE distribution for PHV and PHV-FSCA each because there are 28 ASO flights. Given our primary goal of
 281 estimating the SWE distribution for unsampled dates, we apply models developed for each ASO acquisition to all
 282 other dates except for dates within the same year as that in which the model was developed. For example, in 2013
 283 there were a total of 6 ASO flights out of the total of 28 flights during the four-year study period. This leaves 22 flights
 284 from other years that can be used to develop statistical models of SWE that can be transferred to the dates in 2013. By
 285 conducting our model tests in this manner, we are robustly testing the transferability of models from a given *model*
 286 *date* with regard to simulating SWE distribution on a given *transfer date*, in a hypothetical future year. Due to a
 287 different number of ASO observations in each year, the prediction ensemble size differs for each year. As noted above,
 288 for each date in 2013 there are 22 potential models. For each date in 2014 there are 18 potential models; for each date
 289 in 2015 there are 20 potential models; and for each date in 2016 there are 24 potential models. These models are
 290 referred to as *transferred models*, and for each *transfer date* we identify the *best model* from another year based on
 291 error statistics generated from the ASO data acquired on the *transfer date*. We refer to this model as the *best model*.

292 The ability of a model to transfer from the one date to another will vary based on how well the relationships
 293 between the dependent variable (SWE) and explanatory variables are captured and the similarity of the SWE
 294 distributions. Here we quantify the SWE similarity between dates using the mean absolute error (MAE) of SWE
 295 recorded at snow pillows on both dates within 20 km of the basin as shown in Figure 1; station data was not weighted
 296 by distance to the basin or elevation. For each *transfer date*, there is an ensemble of predictions from the *model dates*
 297 from the other years. Each of the *model dates* exhibit a similarity with the SWE distribution of the *transfer date*. In
 298 order to pick which model date exhibits the greatest similarity with the *transfer date* without having an ASO
 299 observation, we calculate the MAE of SWE at the snow pillows between each pair of model-transfer dates and select
 300 the *model date* with the lowest MAE. We compare the prediction performances from this model selection procedure
 301 (denoted *selected model*) with those of the *best models*. We recognize that a given SWE pattern from the snow pillows
 302 may correspond to multiple basin-wide SWE patterns 1, but given our objective of extending ASO observations to
 303 real-time applications, we must use other data that is available. Some discussion of other options is presented in the
 304 Discussion.

305 4.3 Statistical Model

306 The multiple linear regression models described above, i.e. PHV and PHV-FSCA, are based on a regularized
 307 regression model applied in an elastic net framework (Friedman et al., 2010). The benefit of a regularized regression
 308 over standard regression is that it reduces overfitting while permitting all theoretically relevant variables to be
 309 included, rather than removing potentially useful variables due to multicollinearity. Regularized regression increases
 310 the predictive ability of a model with multiple predictor variables by penalizing the objective function used to estimate
 311 the parameter set. The elastic net is an extension of ordinary least squares, which estimates parameter coefficients by
 312 minimizing the residual sum of squares (RSS) as the objective function (Eqn 1):
 313

Deleted: a total of

Deleted: more

Deleted: .

Deleted: nearby snow pillows.

Deleted: as implemented in the glmnet package in R .

Deleted: conceptual

$$RSS = \sum_{i=1}^n \left(y_i - \sum_{j=1}^p \beta_j x_{ij} \right)^2 \quad \text{Eqn (1)}$$

where y_i is the response variable at the i^{th} observation, β_j is the coefficient for predictor variable j , and x_{ij} is predictor variable j at each observation i . The elastic net penalizes RSS by two different types of regularization techniques, known as L1 and L2, that have opposing properties (more on this below). In this regard, the elastic net estimates the regression parameters β by minimizing RSS in Eqn 2:

$$RSS = \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda P_\alpha(\beta) \quad \text{Eqn (2)}$$

where

$$P_\alpha(\beta) = \sum_{j=1}^p \left[\frac{1}{2} (1 - \alpha) \beta_j^2 + \alpha |\beta_j| \right] \quad \text{Eqn (3)}$$

β_0 is the intercept, λ controls the magnitude of the penalty, and α changes the relative influences of the L1 and L2 regularizations. In practice, this shrinks the coefficient values towards zero for variables exhibiting multicollinearity and for predictors with low explanatory power. Penalized coefficients have lower variance and therefore variables can be selected without resorting to a discrete selection procedure, e.g. a p-value threshold such as in step-wise regression. Consequently, resulting parameter sets are more robust predictors for independent data (Zou and Hastie, 2005).

The elastic net has two tuning parameters, λ and α , which are determined through cross validation. When $\alpha=1$, the penalty is composed completely of the L1 penalty and commonly known as Lasso regression. When $\alpha=0$, the penalty is composed completely of the L2 penalty and is commonly known as Ridge regression. The advantage of the elastic net is that α can range between 0 and 1 and therefore inherits the properties of both L1 and L2 regularization. L1 regularization is commonly used for model selection because predictor coefficients can be shrunk to zero and effectively removed from the model. However, in the presence of correlated predictor variables, one predictor variable would be randomly selected while the others are removed. This can result in decreased predictive performance since variables with some explanatory power are no longer in the model. With L2 regularization, regression coefficients will shrink towards zero but with an asymptote at zero. This is the preferred type of regularization in the presence of multicollinearity because all variables would remain in the model but with smaller coefficients. For further details, we direct the reader to Zou and Hastie (2005) and Hastie et al. (2009).

We do not directly treat spatial correlation in our models due to the large computational demands of fitting the covariance function for ~4000 pixels for 28 dates. Neglecting spatial correlation is another potential source for regression coefficients to be overfit and consequently we do not interpret them for physical meaning (Cressie, 1993;

Deleted: to account

Deleted: less

Deleted: select variables

Deleted: alpha

Deleted: The elastic net provides a framework to choose the best compromise between the L1 and L2 penalties.

Erickson et al., 2005). Nonetheless, we show utility with our methods without addressing spatial correlation and expect the results presented herein would improve if spatial correlation were explicitly treated (Carroll and Cressie, 1997).

5 Results

5.1 SWE Models: Same Day

Table 2 shows that PHV-FSCA outperforms PHV in all metrics in all years except %Bias (where both models exhibit close to 0 %Bias, as expected from a regression). The model PHV-FSCA explains on average between 80% and 85% of the variance in SWE distribution in any given year whereas PHV only explains between 56% and 67%. We similarly see improvement in %MAE where PHV-FSCA exhibits mean annual %MAE between 26% and 40% compared with PHV which yields mean annual %MAE between 47% and 66%. In summary, we note a substantial improvement in r^2 and %MAE for distributing SWE when fSCA is included as an additional variable.

Table 2. Mean prediction performance for PHV and PHV-FSCA from each observation date using a split-sampling approach.

	PHV			PHV-FSCA		
	r^2	%MAE	%Bias	r^2	%MAE	%Bias
2013	0.56	66	1	0.83	33	0
2014	0.67	47	2	0.85	26	1
2015	0.58	59	-1	0.82	32	1
2016	0.61	65	4	0.80	40	2
Mean	0.61	57	1	0.83	31	1

5.2 SWE Models: Transferred in Time

Figure 2 shows March 23, 2014 as an example transfer date that was predicted using a PHV-FSCA model from May 25, 2013. Overall, we see similar spatial trends between the observed and modeled SWE distributions, but we see lighter colors in the observed map indicating higher SWE. The mean observed SWE is 0.21 m compared to 0.22 m modeled. The range of observed SWE is 0-0.67 m while the modeled SWE ranges from 0-0.38 m. The standard deviations of observed and modeled SWE are 0.14 m and 0.12 m, respectively. The difference map shows large areas of agreement to within 0.05 m SWE and a qualitative comparison with Figure 1 suggests these are mainly forested areas. We see areas of under prediction (red) mostly above tree line in the north and southern tip, and areas of over prediction (blue) below tree line in the south. These errors correspond to the steepest slopes, which may be an effect of coarsening the data to 500 m. The snow extent is very similar between the modeled SWE and observed SWE because only areas observed to have fSCA greater than 0 were predicted.

Deleted: Discrete Time

Deleted: 78

Deleted: 86

Deleted: 55

Deleted: 27

Deleted: 41

Deleted: 50

Deleted: 71

Deleted: in 2014

Deleted: darker purples

Deleted: 23

Deleted: 21

Deleted: 75

Deleted: 39

Deleted: 13

Deleted: 11

Deleted: above

Deleted: A comparison with Google Earth® aerial imagery confirms that

Deleted: pixels that exhibit very large negative differences (bright red pixels) are areas with persistent snow for much

Deleted: year

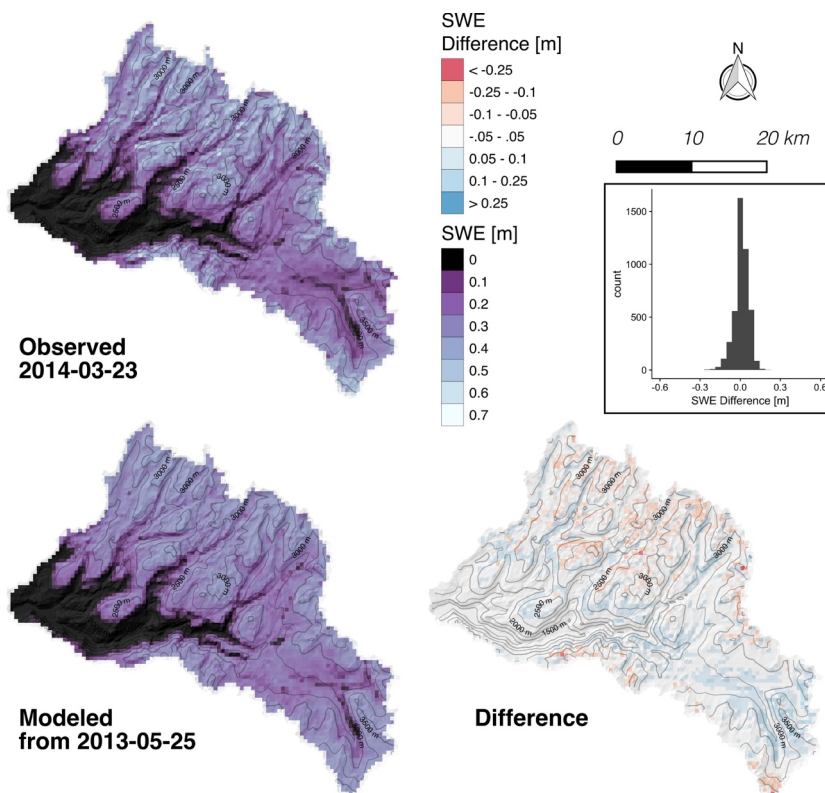


Figure 2. An example of SWE distribution from 2014-03-23. Observed SWE (top left); modeled SWE using PHV-FSCA (bottom left); the difference, modeled SWE – observed SWE (bottom right). A shaded relief and 250 m contours are used to highlight the topography. A histogram of the differences (top right).

Figure 3 shows the range of r^2 for transferred models for PHV and PHV-FSCA on each date based on models created in other years. Additionally, we highlight the best model performance for each transfer date by a diamond. We observe unanimous improvement across all dates with PHV-FSCA compared to PHV, except for April 9, 2015 where PHV yielded 0.01 higher r^2 . PHV-FSCA yields the highest mean best r^2 of 0.83 (mean of diamonds) compared to PHV with a mean best r^2 of 0.61. We also observe a notable decrease in r^2 for PHV towards the end of the season while PHV-FSCA exhibit a consistent r^2 . The standard deviation of r^2 for PHV is 0.13 and for PHV-FSCA it is 0.05.

Deleted: , draped over

Deleted: map (bottom right).

Deleted: indicate

Deleted: In this regard,

Deleted: .

Deleted: 84

Deleted: 6

Deleted: 14

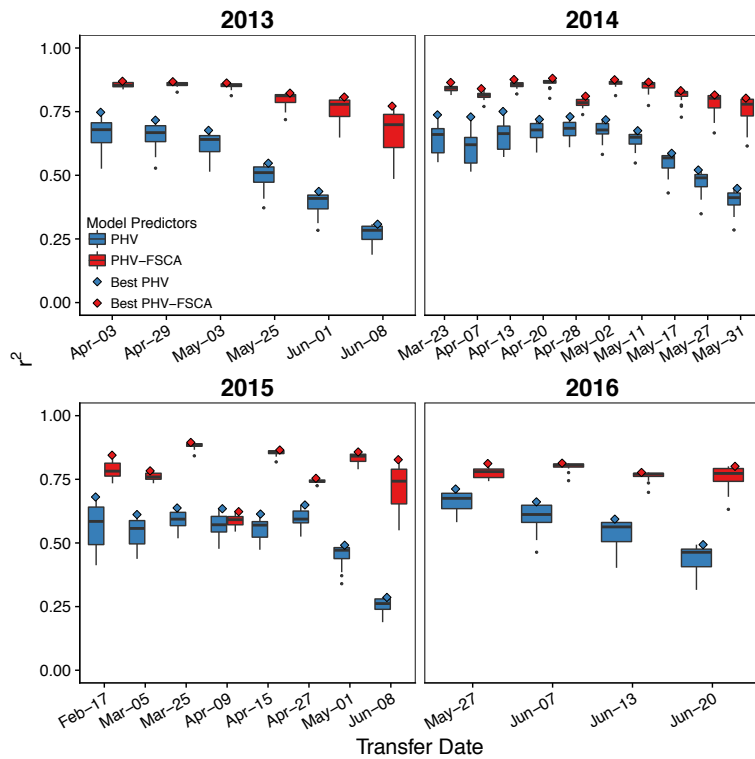


Figure 3. The range of r^2 for each simulation date from the transferred models. The diamond highlights the best model for each transfer date. The boxplots represent the interquartile range with vertical lines to denote the 5th and 95th percentiles. Black dots are outliers.

Figure 4 also clearly shows PHV-FSCA transferred models to exhibit better, i.e. lower, %MAE compared to PHV. April 9, 2015 is a notable exception where PHV exhibits the best transferred model. The mean best %MAE (mean of diamonds) for PHV-FSCA is 33% while the mean best %MAE for PHV is 60%. Particularly obvious in these panels is the upward distribution shift and larger range for PHV later in the season compared to only a minimal increase in %MAE for PHV-FSCA. The standard deviations of the best %MAE are 9% for PHV-FSCA and 25% for PHV.

Deleted: represents

Deleted: the best

Deleted: lowest

Deleted: with transferred models

Deleted: 63

Deleted: 10

Deleted: 26

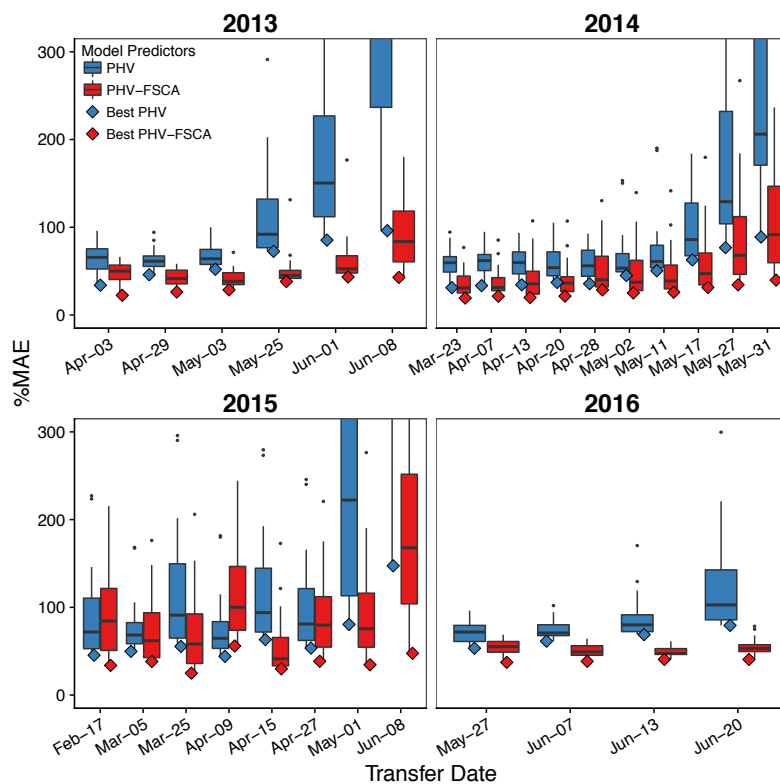


Figure 4. The range of %MAE prediction errors for each transfer date from the transferred models. The diamond highlights the best model for each transfer date. The y-axis is limited for clarity. The boxplots represent the interquartile range with vertical lines to denote the 5th and 95th percentiles. Black dots are outliers.

Figure 5 shows that the *best transferred models* (diamonds) for all models exhibit close to zero bias. The mean best %Bias for PHV is 1% and for PHV-FSCA it is 1%. However, we note that the variability in %Bias increases more dramatically at the end of the season, especially for PHV. The standard deviations of the best (i.e. lowest) %Bias for PHV and PHV-FSCA are 15% and 7%, respectively.

Deleted: represents

Deleted: both

Deleted: 2

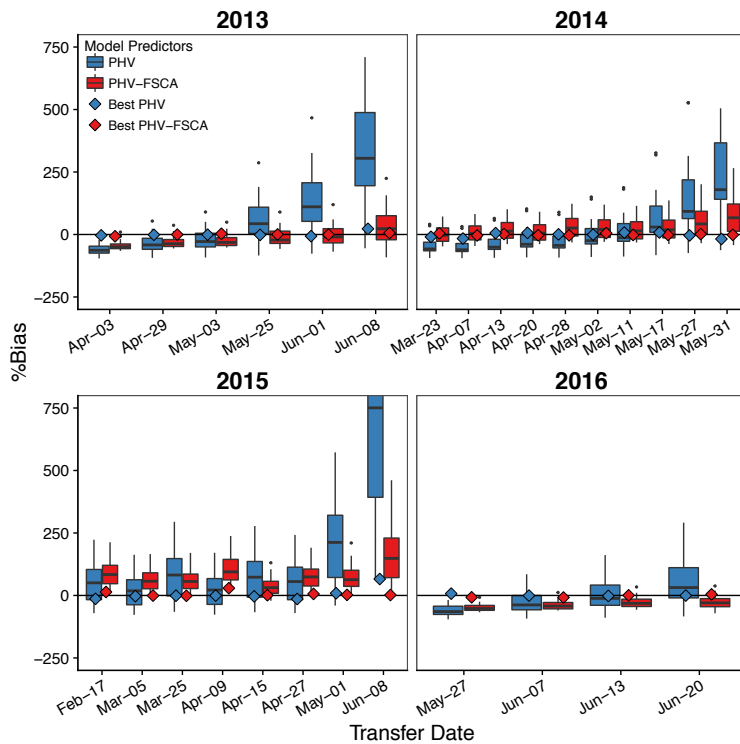


Figure 5. The range of %Bias prediction errors for each transfer date from the transferred models. The diamond highlights the best model for each transfer date. The y-axis is limited for clarity. The boxplots represent the interquartile range with vertical lines to denote the 5th and 95th percentiles. Black dots are outliers.

Following the demonstration of [near](#) unanimous improvement with fSCA as a predictor variable, we compare just the *best transferred models* of PHV-FSCA presented in Figs. 3, 4, 5 with the split-sample models of PHV-FSCA from the previous section, to assess the performance degradation one would expect due to transferring a model between years. We observe that PHV-FSCA models can be transferred to another year with little degradation in performance (Table 3). The yearly r^2 of the *best transferred model* are always within 1% of the *split sample model* and, on average, explains the same amount of variance in SWE distribution. The yearly mean %MAE of the *transferred model* is always within [.5%](#) of the *same-day model* with the mean [.2%](#) higher. The yearly mean magnitude of %Bias is actually lower, i.e. better, in [2014](#) with the *best transferred model* compared to the *same-day model*, and the overall mean is the same.

Table 3. Mean model performance comparison with PHV-FSCA for the split sample model and the best transferred model for each simulation date.

Deleted: represents

Deleted: 6

Deleted: split sample

Deleted: 1

Deleted: two years

Deleted: split sample

	Same Day Model			Best Transferred Model		
	r^2	%MAE	%Bias	r^2	%MAE	%Bias
2013	0.83	33	0	0.83	34	0
2014	0.85	27	1	0.85	27	0
2015	0.82	33	1	0.81	38	6
2016	0.80	40	2	0.8	39	-3
Mean	0.83	31	1	0.83	33	1

5.3 Evaluating the Selected Model Performance

Table 4 summarizes and compares the yearly mean statistics from PHV-FSCA for the *best models* and the *selected models* (i.e. those chosen based on snow pillow similarity) transferred from another year. The model selection procedure results in similar yearly r^2 to the *best models* but increases in both %MAE and %Bias are apparent with the *selected models*. Compared to the *best models*, yearly mean %MAE for *selected models* increases between 4% and 17% and yearly absolute %Bias increases between 2% and 34%. The years 2013, 2014, and 2015 yielded increases in %MAE of 17%, 16%, and 16%, respectively for *selected* versus *best models*. In 2016 selected model %MAE exhibited an increase of only 4%. The absolute %Bias increases 2% in 2013, 7% in 2014, 34% in 2015, and 24% in 2016 for *selected* versus *best models*.

Table 4. Yearly and overall prediction errors of PHV-FSCA for best transferred models and the selected models. Best transferred model errors involved fitting a model to all the data on each date and using these models to predict SWE on dates in other years. Only the best model date-simulation date pair is considered. The selected model errors are derived from the same ensemble of model date-simulation date pairs, but the model is selected based on the pillow SWE similarity described in the text.

	Best Transferred Model			Selected Model		
	r^2	%MAE	%Bias	r^2	%MAE	%Bias
2013	0.83	34	0	0.82	51	2
2014	0.85	27	0	0.82	43	7
2015	0.81	38	6	0.79	54	40
2016	0.8	39	-3	0.79	43	-27
Mean	0.83	33	1	0.81	48	11

Figure 6 shows the difference for each *transfer date* between the errors of the *best models* (diamonds in Figs. 3, 4, 5) and the errors of the *selected models* (these are shown in Fig. 7). We focus on %MAE and %Bias from PHV-FSCA only because r^2 showed generally consistent performance for a given transfer date (Table 4; comparatively small vertical range of the red boxplots in Fig. 3 compared with Figs. 4 and 5).

The mean difference in %MAE between the *selected* and *best models* is 15% and the *selected models* exhibited the same %MAE as the *best model* on two dates (Fig. 6a). The range in performance difference is between 0% and 43% with a mean of 15% and median of 12%. The differences in %MAE were generally lowest in 2016 with

Deleted: 3

Deleted: 18

Deleted: 3

Deleted: 37

Deleted: 18%,

Deleted: 3

Deleted: 3

Deleted: 6

Deleted: 37

Deleted: s

Deleted: as open circles

Deleted: 6

Deleted: purple

Deleted: model

Deleted: only one date

Deleted: 45

a mean of 4% and standard deviation of 5%. In contrast, 2013, 2014, and 2015 exhibited both higher mean differences (18%, 17%, and 16%, respectively) and higher standard deviations in %MAE (13%, 7%, and 13%, respectively).

The *best model* had a lower absolute magnitude %Bias of between 0% and 67% with a mean of 30% and median of 29%. The selected model was the same as the *best model* on one date (Fig. 6b). The yearly mean difference in error was consistently higher for %Bias than %MAE, with means of 27% in 2013, 32% in 2014, 33% in 2015, and 21% in 2016. The standard deviation in the error difference was lowest in 2014 (8%) compared to 14% in 2013, 20% in 2015, and 11% in 2016.

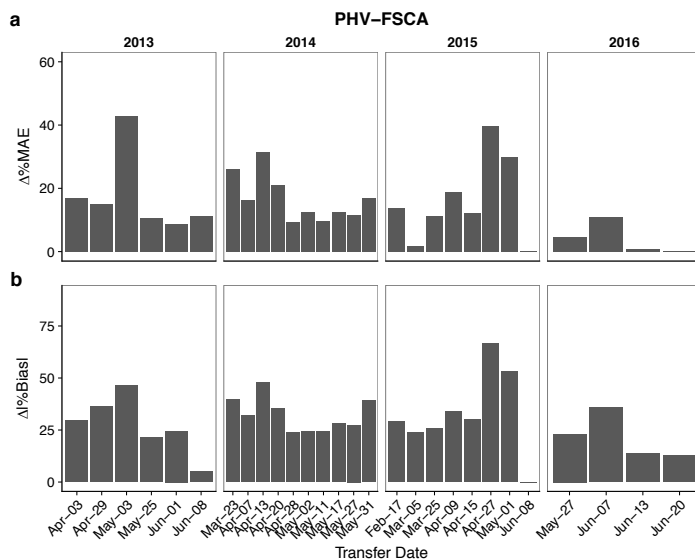


Figure 6. The increase in prediction error for %MAE (a) and %Bias (b) between the *best model* and the *selected model*.

We evaluate the prediction errors from different model years to see if there are any systematic differences in the predictions generated; i.e. do some years yield better predictions of other years? This allows us to determine how sensitive predictions for transfer dates will be given the existing ensemble of observations. In this context, we note distinct differences in the predictive ability of models from different years, which has important implications for the ability of a model to transfer in time.

Figure 7a shows that, on average, models from 2015 produced the lowest %MAE and models from 2016 produce predictions with the largest %MAE. The mean %MAE in 2016 is 113% compared to 75% with 2013 models, 59% with 2014 models and 47% with 2015 models. We observe consistency in a year's ability to predict another year relative to the overall distribution, i.e. the colored dots are typically clustered within the range of the boxplots. We also note in 2016 that the *best models* always came from 2013, but in 2013 the *best models* only came from 2016 for

Deleted: 3

Deleted: 4

Deleted: 16

Deleted: 18

Deleted: 14%, 8

Deleted: 14

Deleted: 73

Deleted: 29

Deleted: 26

Deleted: only

Deleted: 25

Deleted: 30

Deleted: 34

Deleted: 20

Deleted: 9

Deleted: 16

Deleted: 22

Deleted: 118

Deleted: 79

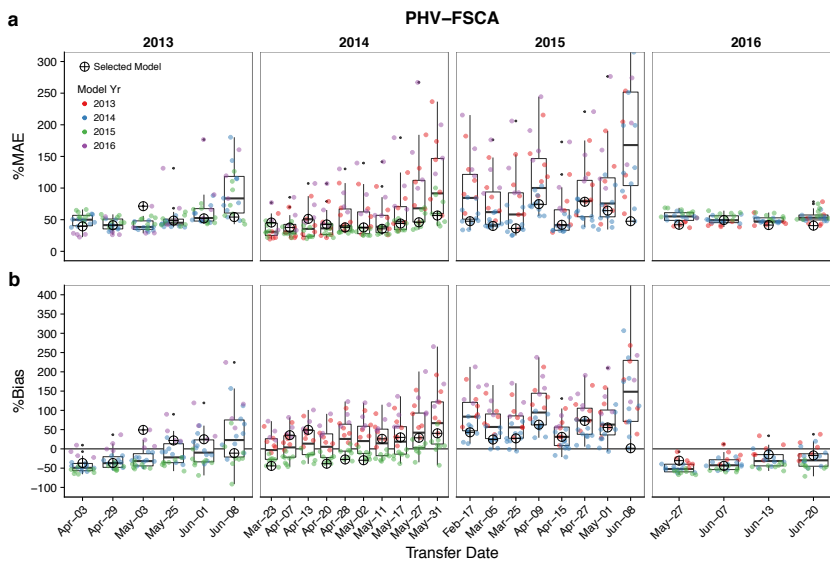
Deleted: 69

Deleted: 50

561 the first three flights. In 2014, the bulk of the models around the median performance were from 2015 and vice-versa
562 in 2015. In these two years, the poorest predictions were from 2013 and 2016.

563 Figure 7b show that models from 2016 also produce the largest %Bias with a mean of 94% compared with
564 56% from 2013, 11% from 2014, and 30% from 2015. We again see consistency in the location of the colored dots
565 relative to the boxplots, thus suggesting models from certain years will consistently estimate the SWE distribution of
566 certain other years better. Similar to the %MAE results, 2013 produces the lowest %Bias in 2016, but in 2013 the
567 inverse is only true for the first three flights as the distribution shifts up.

568



569

570 Figure 7. The prediction errors for PHV-FSCA for each date; %MAE (a) and %Bias (b). The colored dots are the prediction
571 errors coded by model year and jittered to prevent over plotting. The symbol 'O' represents the selected model. The
572 boxplots represent the entire distribution of prediction errors for each date. The boxes represent the interquartile range
573 with vertical bars for the 5th and 95th percentiles. Small black dots are statistical outliers.

574 6 Discussion

575 6.1 SWE Distribution Modeling

576 The relationship between snow covered area and SWE is well established as evident by the common use of depletion
577 curves in hydrologic modeling (Anderson, 1973; Clark et al., 2011; Liston, 1999; Luce et al., 1999). We invert the
578 idea of the depletion curve in this study by using the spatial distribution of fSCA to predict the spatial distribution of
579 SWE. The basis for this approach is relatively well established given that fSCA is sensitive to topographic complexity
580 and the spatial distribution of SWE (Cristea et al., 2017; Donald et al., 1995; Fassnacht et al., 2016; Niu and Yang,
581 2007; Walters et al., 2014). We extend the concept of linking fSCA with topography from these previous works to

Deleted: 97

Deleted: 21

Deleted: 26

Deleted: model

Deleted: In this regard,

Deleted: Open circles represent

Deleted: .

estimating SWE distribution. In this regard, model performance consistently improved when fSCA was used as a predictor, with an average %MAE decrease from 60% with PHV to 33% with PHV-FSCA. The utility of fSCA as an explainer of SWE distribution is not an unprecedented result. König and Sturm (1998) showed that in situ measurements of SWE were correlated with fSCA from aerial photographs. In addition, Marchand et al. (2005) suggested that the sub-grid standard deviation of physiography could improve regression models of SWE distribution because it accounts for sub-grid snow depth variability. Remotely sensed fSCA provides a means of capturing this sub-grid variability in SWE without requiring higher resolution data to compute the variance of physiography for each pixel.

The statistics for the selected PHV-FSCA model reported in this study compare favorably to SWE distribution statistics reported previously. Headwater catchment scale studies, based on intensive field data, have been able to achieve r^2 values from as low as 0.18 to as high as 0.65 (Balk and Elder, 2000; Elder et al., 1991; 1998; Erxleben et al., 2002; López-Moreno and Nogués-Bravo, 2006; Molotch and Bales, 2005) compared to the mean r^2 of 0.81 for the selected models in this study. The results presented herein also compare favorably with larger scale studies (i.e. > 1000 km²) of snow distribution. Fassnacht et al. (2003) reported average yearly RMSE between 0.12 and 0.16 m and 0 m bias when cross-validated with snow pillows. Harshburger et al. (2010) reported an average r^2 of 0.82 and RMSE of 0.05 m when cross-validated with snow pillows. Schneider and Molotch (2016) reported a mean RMSE of 0.08 m and mean %Bias of 11% for the selected model in this study. The favorable error statistics reported here are even more encouraging when considering the differences in evaluation methods of these previous studies. In this regard, the error values reported here are quite robust in that we compare against spatially explicit observations over relatively large spatial scales (i.e. > 1000 km²). In contrast, the aforementioned works were evaluated against relatively sparse observations in relatively homogenous, flat, forested environments (Fassnacht et al., 2003; Harshburger et al., 2010; Schneider and Molotch, 2016).

Bair et al. (2016) compared a retrospective SWE reconstruction to the same ASO observations (2013-2015) and reported yearly mean %MAE between 20% and 31% and yearly mean %Bias between -11% and 10%. These yearly statistics are better than those reported in this study (Tables 3, 4), but are the result of a much more complicated energy balance model that can only be run after the snow has disappeared. The selected model in this study is a simple linear regression that can be applied in real-time, thus we consider our results valuable for applications where real-time estimates of SWE distribution are needed. In this regard, we compare our selected model results to SWE estimates from the U.S. National Weather Service's operational Snow Data Assimilation System (SNODAS). SNODAS produces spatially distributed SWE estimates for the coterminous United States at 1 km by assimilating a physically based model with snow pillow observations and remotely sensed snow covered area. The SWE product from SNODAS is the only high-resolution, gridded SWE product available at a daily time step for the continental United States and is available from <http://nsidc.org> (Barrett, 2003). Previous work has shown the physically based model to perform well at the point scale (Rutter et al., 2008) but suffer in alpine zones because it does not consider wind redistribution (Clow et al., 2012). In our own analysis, the yearly mean r^2 between ASO and SNODAS ranges from 0.04 in 2016 to 0.36 in 2015 with a mean of 0.17. The yearly mean %MAE ranges from 120% in 2014 to 274% in 2013 with a mean

Deleted: 61

Deleted: 32

Deleted: 83

Deleted: These papers, which cover only a few square kilometers, represent a far more simplistic problem with regard to characterizing relationships between snow accumulation and physiographic variables.

Deleted: 07

Deleted: Furthermore

Deleted: SNOTEL

Deleted:

of 199%. The yearly mean %Bias ranges from -10% in 2016 to 236% in 2015. In comparison to the PHV-FSCA selected model in Table 4, SNODAS exhibits a mean %MAE 4 times greater than that of PHV-FSCA and a mean %Bias 8 times higher than PHV-FSCA. Both models poorly predicted the anomalous conditions in 2015, which experienced high rain fraction and ephemeral snowpacks at low to mid elevations. While it is clear that the errors with PHV-FSCA are considerably lower than with SNODAS overall, SNODAS is a complex system that attempts to capture the snow dynamics across the entire United States compared with PHV-FSCA which was trained using a very specialized data set in the study region.

We also show that SWE distributions can be related to fSCA and physiography in one year and applied to another year. The performance of PHV-FSCA was quite similar when applied to the date at which the model was trained (i.e. same-day models with a split sample) versus applying the model to other years (i.e. transferred models) as evidenced by minimal decrease in prediction skill and minimal increase in prediction error when we compare same-day models with the best transferred models. Recall that the same-day model was trained and tested on the same day and the transferred model was trained in a different year from which the model was applied, yet the mean r^2 in the same-day is equivalent to that of the best transferred model (Table 3). Moreover, the average %MAE of the best transferred model exhibits only a 2% difference from the same-day model and the mean %Bias exhibits no difference. Thus, if we are able to identify the best model for a given date of interest we would see minimal degradation in predictive ability relative to a model derived from data acquired on the date of interest. However, we see significant differences in predictive ability from models of certain years (Fig. 7), which suggests that relationships between SWE and physiography are only similar between specific years, not as uniformly as put forth by previous studies (Erickson et al., 2005; Grünwald et al., 2013). Also, it is unclear as to the impact of climate non-stationarity with respect to the ability to transfer models to future years. Even so, for each year in this dataset there exists a corresponding year from which accurate predictions can be made.

The benefits of using fSCA as a predictor variable in the transferred models are particularly large at the end of the season when the errors are highest (Figs. 3, 4, 5). For the last 2 dates of each year, the average difference in %MAE between the best PHV model and best PHV-FSCA model was 50% compared to an average difference of 18% for the other dates. The improvements seen by including fSCA as a predictor are noteworthy because the ensemble of models trained using ASO observations could then be applied using remotely sensed fSCA from satellites, and recent advances in downscaling fSCA may further improve SWE estimates at the end of the melt season when only small patches of snow persist (Cristea et al., 2017; Li et al., 2015; Walters et al., 2014). Nonetheless, the degree to which satellite-based fSCA will improve model performance toward the end of the snowmelt season will be partially dictated by the accuracy of the fSCA data, which is subject to increasing uncertainty at low fSCA values and with downscaling (Painter et al., 2009; Rittger et al., 2013). Optical fSCA products such as MODSCAG also suffer in forested areas since snow cover is occluded by the canopy (Raleigh et al., 2013; Rittger et al., 2013). A viewable gap fraction correction is typically used to extrapolate fSCA to the occluded portions of a pixel by assuming that fSCA is the same under the canopy as it is in canopy gaps (Molotch and Margulis, 2008), and future studies should evaluate the sensitivity of PHV-FSCA to this assumption. Cloud cover can also obscure a satellite's view of the snow and therefore making it difficult to estimate the snow extent. However, Slater et al. (2013) showed that gaps of 5 or more consecutive

Deleted: We refer the reader

Deleted: "Selected Model" of Table 4 for

Deleted: error summary. In this regard

Deleted: .

Deleted: discrete time

Deleted: . In this regard, we see a

Deleted: split sample

Deleted: split sample

Deleted: . Table 3 shows

Deleted: split sample model to be

Deleted: .

Deleted: 1

Deleted: split sample

Deleted: different

Deleted:) and conclude

Deleted: 54

Deleted: 20

Deleted: .

Deleted: .

693 days are rare with MODIS, thus suggesting that weekly estimates of SWE would be feasible. Lastly, omission and
 694 commission errors of cloud identification can provide erroneous estimates of fSCA although Parajka and Blöschl
 695 (2008) report an overall filtering accuracy of 96% in the Alps. Nonetheless, this is still an active area of research
 696 (Dozier et al., 2008; Parajka and Bloeschl, 2008; Rittger et al., 2013; Xia et al., 2012).

697 6.2 Considerations for Extending the ASO Record

698 The value of ASO during the year flown is significant for water management because it provides high resolution SWE
 699 information with low uncertainty compared to traditional estimation methods and therefore facilitates more accurate
 700 water supply forecasts. The work presented here provides a first step in realizing the value of ASO subsequent to
 701 active operations. We find that the number of flights within a year affects the mean and variance of the predictions in
 702 other years by <1%; this was quantified by iteratively selecting between 1 and the number of observations in a given
 703 year 100 times and assessing the change in error. In other words, a greater breadth of value can be derived by flying
 704 ASO once per year for 10 years as opposed to 10 times in a single year.

705 One might also consider the value of a single flight each year for estimating SWE distribution on other dates in
 706 the same year and in this case we found the mean r^2 to be 0.82, %MAE to be 39%, and %Bias to be 7% when
 707 considering the best model for each date. These values are similar to the yearly values from the best transferred model
 708 from other years (Tables 3, 4), which supports the prior assertion that the good or equivalent performance of models
 709 from other years suggests a strong degree of persistence in snow patterns. We note the positive %Bias and suggest
 710 that this perhaps should not come as a surprise since we only used prior flights for the same year, which measured
 711 deeper snowpacks and greater fSCA than those modeled; ASO flights typically start around peak SWE. Future work
 712 should identify the importance of the fSCA state for transferring a model, and whether an ASO flight in the
 713 accumulation season might yield better results.

714 It is clear from our results that flights in one year do not necessarily transfer well to another year, e.g. 2016 to
 715 2014 and 2016 to 2015 (Fig. 7). The model selection procedure relies on operational snow pillows to identify similar
 716 patterns of SWE between historical dates (i.e. model dates) and the date of interest (i.e. transfer dates). The relative
 717 homogeneity of these stations may have contributed to the general inability of these stations to select the best historical
 718 date for a given date of interest, because distinct basin-wide SWE patterns may correspond to similar patterns at the
 719 snow pillows. It is well established that these snow pillows may not adequately represent the SWE of the surrounding
 720 terrain (Meromy et al., 2012; Molotch and Bales, 2006; Rice et al., 2011), but they provide the only source of real-
 721 time SWE information. However, the minimal degradation in prediction performance when considering the best
 722 transferred model motivates improvements for identifying the dates from the past with the most similar SWE
 723 distribution. As shown here, if these dates can be identified, the SWE can be modeled at accuracy levels that are
 724 equivalent to models generated from ASO data collected on the date of interest. We also tested the similarity in
 725 remotely sensed fSCA as a method for identifying the historical date with the most similar SWE distribution to the
 726 date of interest. We found that anomalous SWE and fSCA distributions make this method less robust than using snow
 727 pillow data. However, a completely remote-sensing based approach would be useful in data sparse regions where
 728 ground stations do not exist. A potentially robust remote-sensing based method could be to track fSCA through time

Deleted: The downside to ASO is the relatively high cost of operation compared to traditional measurement campaigns.

Deleted: it is better to perform

Deleted: than

Deleted: one

Deleted: spatial representativeness

Deleted: . In this context,

Deleted: .

Deleted: should motivate

and select a similar SWE distribution based on the trajectory of fSCA rather than a single snapshot of fSCA. It is also important to note that flying ASO during a year with an anomalously low snowpack such as in 2015 in California does not necessarily reduce the predictive capacity of models for future years. In our case, 2015 provided useful estimates of SWE distribution in all the other years, including 2016, even though the models from the relatively wet year of 2016 did not transfer well to the very dry year of 2015 (Fig. 7).

Intuitively, it is not surprising that the models from 2016 did not transfer to other years because there was much more snow than in 2014 and 2015, which were very low snow years. Less intuitive, however, is why the inverse worked well, i.e. the models from 2014 and 2015 did transfer relatively well to 2016. In this regard, we see lower mean fSCA in 2016 for the same mean SWE in 2015 in the ASO data. This means that for the same mean SWE, there are deeper pockets of snow covering a smaller area remaining at the end of a higher snow year (i.e. 2016) than a lower snow year (i.e. 2015). Therefore, we would expect the relationships between topographic variables and SWE to become increasingly disparate from the underlying terrain with a deeper snow accumulation. Thus, the regression coefficients derived from years with deeper snowpacks do not adequately represent relationships found during years with shallow snowpacks. However, this does not mean that the model from the last date of the season (which is also a shallow snowpack) transfers well because this SWE distribution still largely represents the dominant spatial patterns from the peak SWE distribution (Egli and Jonas, 2009; Liston, 1999; Luce et al., 1999).

7 Conclusion

We estimated the relationships between SWE, physiography, and fSCA from ASO data. Our results show that fSCA information strongly improves statistical SWE models, and these models can be reliably transferred directly from one year to another due to strong persistence in snow accumulation and melt patterns. We estimate SWE in years beyond the ASO observation record with a mean r^2 of .83, mean %MAE of 33%, and mean %Bias of 1%. The relationships transfer robustly in time with no degradation in r^2 or %Bias and only 2% in %MAE when comparing predictions between models fit on the same day and the best models from a different year. Models with fSCA as a predictor transfer better than those without, and we suggest that the inclusion of fSCA provides information with regard to the variability of the SWE resulting from different accumulation and melt dynamics due to differences in terrain roughness. The availability of satellite images of fSCA thus can facilitate the transfer of modeled relationships based on ASO observations to dates when no airborne snow depth measurements exist. The crux of this approach is in selecting the model to transfer. We offer a method with which to select a model from another year based on the similarity in SWE distribution at existing snow pillows in the area. Comparison of the best predictions and the selected predictions results in a mild decrease in r^2 (0.02) and moderate increases in %MAE (15%) and %Bias (10%). The results motivate further refinement in the technique used to select the best model because if these dates can be identified then SWE can be modeled at accuracy levels equivalent to models generated from ASO data collected on the day of interest. Lastly, although SWE distributions simulated in years with anomalous SWE distributions (2014, 2015) had the highest errors, models from these years still yielded good performance in 2013 and 2016. Thus, the benefit of ASO in anomalously dry years is two-fold: water managers receive accurate information during a year that is difficult to model, but also these observed SWE distributions can be used to simulate SWE distributions in future,

Deleted: and

Deleted: the temporal consistency in

Deleted: relationships

Deleted: used to

Deleted: ,

Deleted: 85

Deleted: 1

Deleted: In this regard,

Deleted: proposition

Deleted: presented above

less anomalous years. Overall, ASO provides an unprecedented observation of the relationships between SWE, fSCA, and physiography. The ASO dataset facilitates improved understanding of these relationships in both time and space and should lead to better information for water managers.

8 Data Availability

All data and code is available at <https://osf.io/5643w/>

Deleted: from the corresponding author

Deleted: noah.molotch@colorado.edu.

9 Acknowledgements

D. Schneider thanks his PhD committee for their insightful comments. The authors would like to acknowledge NASA for supporting D. Schneider with an Earth and Space Science Fellowship (NNX14AL27H) and funding the time of N.P. Molotch with grant numbers NNX17AF50G and 80NSSC17K0071.

10 References

- Anderson, E.: Snow Accumulation and Ablation Model, US Department of Commerce, Silver Spring, MD. 1973.
- Bair, E. H., Rittger, K. E., Davis, R. E., Painter, T. H. and Dozier, J.: Validating reconstruction of snow water equivalent in California's Sierra Nevada using measurements from the NASA Airborne Snow Observatory, *Water Resour. Res.*, n/a–n/a, doi:10.1002/2016WR018704, 2016.
- Balk, B. and Elder, E.: Combining binary decision tree and geostatistical methods to estimate snow distribution in a mountain watershed, *Water Resources Research*, 36(1), 13–26, doi:10.1029/1999WR900251, 2000.
- Barrett, A. P.: National Operational Hydrologic Remote Sensing Center SNOW Data Assimilation System (SNODAS) Products at NSIDC, National Snow and Ice Data Center, Boulder, CO, USA. 2003.
- Bloschl, G., Kirnbauer, R. and Gutknecht, D.: Distributed Snowmelt Simulations in an Alpine Catchment 1. Model Evaluation on the Basis of Snow Cover Patterns, *Water Resources Research*, 27(12), 3171–3179, doi:10.1029/91WR02250, 1991.
- Bühler, Y., Marty, M., Egli, L., Veitinger, J., Jonas, T., Thee, P. and Ginzler, C.: Snow depth mapping in high-alpine catchments using digital photogrammetry, *The Cryosphere*, 9(1), 229–243, doi:10.5194/tc-9-229-2015, 2015.
- Carroll, S. S. and Cressie, N.: A comparison of geostatistical methodologies used to estimate snow water equivalent - Reply, *Journal of the American Water Resources Association*, 33(1), 221–222, 1997.
- Clark, M. P., Hendrikx, J., Slater, A. G., Kavetski, D., Anderson, B., Cullen, N. J., Kerr, T., Örn Hreinsson, E. and Woods, R. A.: Representing spatial variability of snow water equivalent in hydrologic and land-surface models: A review, *Water Resources Research*, 47(7), doi:10.1029/2011WR010745, 2011.
- Cline, D. W., Bales, R. C. and Dozier, J.: Estimating the spatial distribution of snow in mountain basins using remote sensing and energy balance modeling, *Water Resources Research*, 34(5), 1275–1285, doi:10.1029/97WR03755, 1998.
- Clow, D. W., Nanus, L., Verdin, K. L. and Schmidt, J.: Evaluation of SNODAS snow depth and snow water equivalent estimates for the Colorado Rocky Mountains, USA, edited by M. Pelto and K. RRichard, *Hydrological Processes*, 26(17), 2583–2591, doi:10.1002/hyp.9385, 2012.

820 Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V. and Boehner, J.:
821 System for Automated Geoscientific Analyses (SAGA) v. 2.1.4, Geoscientific Model Development, 8(7), 1991–
822 2007, doi:10.5194/gmd-8-1991-2015, 2015.

823 Cressie, N. A. C.: Statistics for spatial data, J. Wiley. 1993.

824 Cristea, N. C., Breckheimer, I., Raleigh, M. S., HilleRisLambers, J. and Lundquist, J. D.: An evaluation of terrain-
825 based downscaling of fractional snow covered area data sets based on LiDAR-derived snow data and orthoimagery,
826 Water Resour. Res., 53(8), 6802–6820, doi:10.1002/2017WR020799, 2017.

827 Dadic, R., Mott, R., Lehning, M. and Burlando, P.: Wind influence on snow depth distribution and accumulation
828 over glaciers, J. Geophys. Res., 115(F1), n/a–n/a, doi:10.1029/2009JF001261, 2010.

829 Deems, J. S., Fassnacht, S. R. and Elder, K. J.: Interannual Consistency in Fractal Snow Depth Patterns at Two
830 Colorado Mountain Sites, J. Hydrometeorol., 9(5), 977–988, doi:10.1175/2008JHM901.1, 2008.

831 Deems, J. S., Painter, T. H. and Finnegan, D. C.: Lidar measurement of snow depth: a review, Journal of Glaciology,
832 59(215), 467–479, doi:10.3189/2013JoG12J154, 2013.

833 Donald, J. R., Soulis, E. D., Kouwen, N. and Pietroniro, A.: A Land Cover-Based Snow Cover Representation for
834 Distributed Hydrologic-Models, Water Resour. Res., 31(4), 995–1009, doi:10.1029/94WR02973, 1995.

835 Dozier, J.: Mountain hydrology, snow color, and the fourth paradigm, Eos Trans. AGU, 92(43), 373–374,
836 doi:10.1029/2011EO430001, 2011.

837 Dozier, J., Painter, T. H., Rittger, K. E. and Frew, J. E.: Time–space continuity of daily maps of fractional snow
838 cover and albedo from MODIS, Advances in Water Resources, 31(11), 1515–1526,
839 doi:10.1016/j.advwatres.2008.08.011, 2008.

840 Dozier, J., Schneider, S. R. and McGinnis, D. F.: Effect of grain size and snowpack water equivalence on visible and
841 near-infrared satellite observations of snow, Water Resources Research, 17(4), 1213–1221,
842 doi:10.1029/WR017i004p01213, 1981.

843 Egli, L. and Jonas, T.: Hysteretic dynamics of seasonal snow depth distribution in the Swiss Alps, Geophys. Res.
844 Lett., 36(2), n/a–n/a, doi:10.1029/2008GL035545, 2009.

845 Elder, K., Dozier, J. and Michaelsen, J.: Snow accumulation and distribution in an Alpine Watershed, Water
846 Resources Research, 27(7), 1541–1552, doi:10.1029/91WR00506, 1991.

847 Elder, K., Rosenthal, W. and Davis, R. E.: Estimating the spatial distribution of snow water equivalence in a
848 montane watershed, Hydrological Processes, 12(10-11), 1793–1808, doi:10.1002/(SICI)1099-
849 1085(199808/09)12:10/11<1793::AID-HYP695>3.3.CO;2-B, 1998.

850 Erickson, T. A., Williams, M. W. and Winstral, A.: Persistence of topographic controls on the spatial distribution of
851 snow in rugged mountain terrain, Colorado, United States, Water Resour. Res., 41(4), W04014,
852 doi:10.1029/2003WR002973, 2005.

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

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

857 Fassnacht, S. R., Sexstone, G. A., Kashipazha, A. H., Jasinski, M. F., Kampf, S. K., Thaden, Von, B. C. and López
858 Moreno, J. I.: Deriving snow-cover depletion curves for different spatial scales from remote sensing and snow
859 telemetry data, *Hydrological Processes*, 30(11), 1708–1717, doi:10.1002/hyp.10730, 2016.

860 Friedman, J., Hastie, T. and Tibshirani, R.: Regularization Paths for Generalized Linear Models via Coordinate
861 Descent, *Journal of Statistical Software*, 33(1), 1–22, 2010.

862 Garvert, M. F., Smull, B. and Mass, C.: Multiscale Mountain Waves Influencing a Major Orographic Precipitation
863 Event, <http://dx.doi.org/10.1175/JAS3876.1>, 64(3), 711–737, doi:10.1175/JAS3876.1, 2007.

864 GDAL Development Team: GDAL - Geospatial Data Abstraction Library, Version 1.11.3. 2015.

865 Giersch, J. J., Hotaling, S., Kovach, R. P., Jones, L. A. and Muhlfeld, C. C.: Climate-induced glacier and snow loss
866 imperils alpine stream insects, *Glob Chang Biol*, 23(7), 2577–2589, doi:10.1111/gcb.13565, 2016.

867 Good, W. and Martinec, J.: Pattern recognition of air photographs for estimation of snow reserves, *Annals of*
868 *Glaciology*, 1987.

869 Grohmann, C. H., Smith, M. J. and Riccomini, C.: Multiscale Analysis of Topographic Surface Roughness in the
870 Midland Valley, Scotland, *Geoscience and Remote Sensing, IEEE Transactions on*, 49(1), 1200–1213,
871 doi:10.1109/TGRS.2010.2053546, 2011.

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

875 Gutmann, E. D., Larson, K. M., Williams, M. W., Nievinski, F. G. and Zavorotny, V.: Snow measurement by GPS
876 interferometric reflectometry: an evaluation at Niwot Ridge, Colorado, *Hydrological Processes*, 26(19), 2951–2961,
877 2012.

878 Hall, D. K., Foster, J. L., Salomonson, V. V., Klein, A. G. and Chien, J. Y. L.: Development of a technique to assess
879 snow-cover mapping errors from space, *Geoscience and Remote Sensing, IEEE Transactions on*, 39(2), 432–438,
880 doi:10.1109/36.905251, 2001.

881 Harpold, A., Dettinger, M. and Rajagopal, S.: Defining snow drought and why it matters, *EOS - Earth & Space*
882 *Science News*, 98, doi:10.1029/2017EO068775, 2017.

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

886 Hastie, T., Tibshirani, R. and Friedman, J.: *The Elements of Statistical Learning*, Springer New York, New York,
887 NY. 2009.

888 Jonas, T., Marty, C. and Magnusson, J.: *Journal of Hydrology*, *Journal of Hydrology*, 378(1-2), 161–167,
889 doi:10.1016/j.jhydrol.2009.09.021, 2009.

890 Kimbauer, R. and Blochl, G.: Wie ähnlich sind Ausaperungsmuster von Jahr zu Jahr (How similar are snow cover
891 patterns from year to year)? *Deutsche Wässerkundliche Mitteilungen*, 37(5/6), 113–121, 1994.

892 Koch, F., Prasch, M., Schmid, L., Schweizer, J. and Mauser, W.: Measuring Snow Liquid Water Content with Low-
893 Cost GPS Receivers, *Sensors* 2014, Vol. 14, Pages 20975–20999, 14(11), 20975–20999, doi:10.3390/s141120975,
894 2014.

895 König, M. and Sturm, M.: Mapping snow distribution in the Alaskan Arctic using aerial photography and
896 topographic relationships, *Water Resources Research*, 34(12), 3471–3483, doi:10.1029/98WR02514, 1998.

897 Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., Zeng, X., Yang,
898 Z.-L., Levis, S., Sakaguchi, K., Bonan, G. B. and Slater, A. G.: Parameterization improvements and functional and
899 structural advances in Version 4 of the Community Land Model, *J. Adv. Model. Earth Syst.*, 3(3), M03001–n/a,
900 doi:10.1029/2011MS000045, 2011.

901 Lehning, M., Löwe, H., Ryser, M. and Raderschall, N.: Inhomogeneous precipitation distribution and snow transport
902 in steep terrain, *Water Resour. Res.*, 44(7), 123, doi:10.1029/2007WR006545, 2008.

903 Li, H., He, Y., Hao, X., Che, T., Wang, J. and Huang, X.: Downscaling Snow Cover Fraction Data in Mountainous
904 Regions Based on Simulated Inhomogeneous Snow Ablation, *Remote Sensing*, 7(7), 8995–9019,
905 doi:10.3390/rs70708995, 2015.

906 Liston, G. E.: Interrelationships among Snow Distribution, Snowmelt, and Snow Cover Depletion: Implications for
907 Atmospheric, Hydrologic, and Ecologic Modeling, *Journal of Applied Meteorology*, 38(10), 1474–1487,
908 doi:10.1175/1520-0450(1999)038<1474:IASDSA>2.0.CO;2, 1999.

909 Litaor, M. I., Williams, M. and Seastedt, T. R.: Topographic controls on snow distribution, soil moisture, and
910 species diversity of herbaceous alpine vegetation, Niwot Ridge, Colorado, *J. Geophys. Res.*, 113(G2), n/a–n/a,
911 doi:10.1029/2007JG000419, 2008.

912 Livneh, B., Xia, Y., Mitchell, K. E., Ek, M. B. and Lettenmaier, D. P.: Noah LSM Snow Model Diagnostics and
913 Enhancements, *J. Hydrometeorol.*, 11(3), 721–738, doi:10.1175/2009JHM1174.1, 2010.

914 López-Moreno, J. I. and Nogués-Bravo, D.: Interpolating local snow depth data: an evaluation of methods,
915 *Hydrological Processes*, 20(10), 2217–2232, doi:10.1002/hyp.6199, 2006.

916 López-Moreno, J. I., Fassnacht, S. R., Heath, J. T., Musselman, K. N., Revuelto, J., Latron, J. B. P., Morán-Tejeda,
917 E. and Jonas, T.: Small scale spatial variability of snow density and depth over complex alpine terrain: Implications
918 for estimating snow water equivalent, *Advances in Water Resources*, 55, 40–52,
919 doi:10.1016/j.advwatres.2012.08.010, 2013.

920 López-Moreno, J. I., Revuelto, J., Fassnacht, S. R., Azorín-Molina, C., Vicente-Serrano, S. M., Morán-Tejeda, E.
921 and Sextone, G. A.: Snowpack variability across various spatio-temporal resolutions, *Hydrological Processes*,
922 29(6), 1213–1224, doi:10.1002/hyp.10245, 2014.

923 Luce, C. H. and Tarboton, D. G.: The application of depletion curves for parameterization of subgrid variability of
924 snow, *Hydrological Processes*, 18(8), 1409–1422, doi:10.1002/hyp.1420, 2004.

925 Luce, C. H., Tarboton, D. G. and Cooley, K. R.: Sub-grid parameterization of snow distribution for an energy and
926 mass balance snow cover model, *Hydrological Processes*, 13(12-13), 1921–1933, doi:10.1002/(SICI)1099-
927 1085(199909)13:12/13<1921::AID-HYP867>3.0.CO;2-S, 1999.

928 Marchand, W. D. and Killingtveit, A.: Statistical probability distribution of snow depth at the model sub-grid cell
929 spatial scale, *Hydrological Processes*, 19(2), 355–369, doi:10.1002/hyp.5543, 2005.

930 Marks, D., Domingo, J., Susong, D., Link, T. and Garen, D.: A spatially distributed energy balance snowmelt model
931 for application in mountain basins, *Hydrological Processes*, 13(12-13), 1935–1959, doi:10.1002/(SICI)1099-
932 1085(199909)13:12/13<1935::AID-HYP868>3.0.CO;2-C, 1999.

933 Marshall, H.-P. and Koh, G.: FMCW radars for snow research, *Cold Regions Science and Technology*, 52(2), 118–
934 131, 2008.

935 Martinec, J. and Rango, A.: Areal distribution of snow water equivalent evaluated by snow cover monitoring, *Water*
936 *Resour. Res.*, 17(5), 1480–1488, 1981.

937 Meromy, L., Molotch, N. P., Link, T. E., Fassnacht, S. R. and Rice, R.: Subgrid variability of snow water equivalent
938 at operational snow stations in the western United States, *Hydrological Processes*, 27(17), 2383–2400,
939 doi:10.1002/hyp.9355, 2012.

940 Mizukami, N. and Perica, S.: Spatiotemporal Characteristics of Snowpack Density in the Mountainous Regions of
941 the Western United States, *Journal of Hydrometeorology*, 9(6), 1416–1426, doi:10.1175/2008JHM981.1, 2008.

942 Molotch, N. P. and Bales, R. C.: Scaling snow observations from the point to the grid element: Implications for
943 observation network design, *Water Resources Research*, 41(11), doi:10.1029/2005WR004229, 2005.

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

947 Molotch, N. P. and Margulis, S. A.: Estimating the distribution of snow water equivalent using remotely sensed
948 snow cover data and a spatially distributed snowmelt model: A multi-resolution, multi-sensor comparison, *Advances*
949 *in Water Resources*, 31(11), 1503–1514, doi:10.1016/j.advwatres.2008.07.017, 2008.

950 Molotch, N. P., Colee, M. T., Bales, R. C. and Dozier, J.: Estimating the spatial distribution of snow water
951 equivalent in an alpine basin using binary regression tree models: the impact of digital elevation data and
952 independent variable selection, *Hydrological Processes*, 19(7), 1459–1479, doi:10.1002/hyp.5586, 2005.

953 Niu, G.-Y. and Yang, Z.-L.: An observation-based formulation of snow cover fraction and its evaluation over large
954 North American river basins, *J. Geophys. Res.*, 112(D21), doi:10.1029/2007JD008674, 2007.

955 Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D.,
956 Rosero, E., Tewari, M. and Xia, Y.: The community Noah land surface model with multiparameterization options
957 (Noah-MP): 1. Model description and evaluation with local-scale measurements, *J. Geophys. Res.*, 116(D12), n/a–
958 n/a, doi:10.1029/2010JD015139, 2011.

959 Nolan, M., Larsen, C. F. and Sturm, M.: Mapping snow-depth from manned-aircraft on landscape scales at
960 centimeter resolution using Structure-from-Motion photogrammetry, *The Cryosphere Discuss.*, 9(1), 333–381,
961 doi:10.5194/tcd-9-333-2015, 2015.

962 NPS: Plants - Yosemite National Park, [online] Available from: <https://www.nps.gov/yose/learn/nature/plants.htm>
963 (Accessed 5 October 2016), 2016.

964 Painter, T. H., Berisford, D. F., Boardman, J. W., Bormann, K. J., Deems, J. S., Gehrke, F., Hedrick, A., Joyce, M.,
965 Laidlaw, R., Marks, D., Mattmann, C., McGurk, B., Ramirez, P., Richardson, M., Skiles, S. M., Seidel, F. C. and
966 Winstral, A.: The Airborne Snow Observatory: Fusion of scanning lidar, imaging spectrometer, and physically-
967 based modeling for mapping snow water equivalent and snow albedo, *Remote Sensing of Environment*, 184, 139–
968 152, doi:10.1016/j.rse.2016.06.018, 2016.

969 Painter, T. H., Rittger, K. E., McKenzie, C., Slaughter, P., Davis, R. E. and Dozier, J.: Retrieval of subpixel snow
970 covered area, grain size, and albedo from MODIS, *Remote Sensing of Environment*, 113(4), 868–879,
971 doi:10.1016/j.rse.2009.01.001, 2009.

972 Parajka, J. and Bloeschl, G.: Spatio-temporal combination of MODIS images - potential for snow cover mapping,
973 *Water Resources Research*, 44(3), doi:10.1029/2007WR006204, 2008.

974 Parsons, W. J. and Castle, G. H.: Aerial reconnaissance of mountain snow fields for maintaining up-to-date forecasts
975 of snow melt runoff during the melt season, *Western Snow Conference*, Reno, Nevada. 1959.

976 Potts, H. L.: Snow surveys and runoff forecasting from photographs, Proceedings of the 5th Annual Western
977 Interstate Snow Survey Conference, 1937.

978 Potts, H. L.: A photographic snow survey method of forecasting runoff, Proceedings of the 12th Annual Western
979 Interstate Snow Survey Conference, 1944.

980 Prokop, A.: Assessing the applicability of terrestrial laser scanning for spatial snow depth measurements, Cold
981 Regions Science and Technology, 54(3), 155–163, doi:10.1016/j.coldregions.2008.07.002, 2008.

982 Raleigh, M. S., Rittger, K. E., Moore, C. E., Henn, B., Lutz, J. A. and Lundquist, J. D.: Ground-based testing of
983 MODIS fractional snow cover in subalpine meadows and forests of the Sierra Nevada, Mon. Wea. Rev., 128, 44–57,
984 doi:10.1016/j.rse.2012.09.016, 2013.

985 Revuelto, J., Jonas, T. and López Moreno, J. I.: Backward snow depth reconstruction at high spatial resolution based
986 on time-lapse photography, Hydrological Processes, 30(17), 2976–2990, doi:10.1002/hyp.10823, 2016.

987 Revuelto, J., López-Moreno, J. I., Azorín-Molina, C. and Vicente-Serrano, S. M.: Topographic control of snowpack
988 distribution in a small catchment in the central Spanish Pyrenees: intra- and inter-annual persistence, The
989 Cryosphere, 8(5), 1989–2006, doi:10.5194/tc-8-1989-2014, 2014.

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

992 Rice, R., Bales, R. C., Painter, T. H. and Dozier, J.: Snow water equivalent along elevation gradients in the Merced
993 and Tuolumne River basins of the Sierra Nevada, Water Resour. Res., 47, doi:10.1029/2010WR009278, 2011.

994 Rittger, K. E., Painter, T. H. and Dozier, J.: Assessment of methods for mapping snow cover from MODIS,
995 Advances in Water Resources, 51, 367–380, doi:10.1016/j.advwatres.2012.03.002, 2013.

996 Rosenthal, W. and Dozier, J.: Automated Mapping of Montane Snow Cover at Subpixel Resolution from the
997 Landsat Thematic Mapper, Water Resources Research, 32(1), 115–130, doi:10.1029/95WR02718, 1996.

998 Rutter, N., Cline, D. and Li, L.: Evaluation of the NOHRSC snow model (NSM) in a one-dimensional mode, Journal
999 of Hydrometeorology, 9(4), 695–711, doi:10.1175/2008JHM861.1, 2008.

1000 Salomonson, V. V. and Appel, I.: Estimating fractional snow cover from MODIS using the normalized difference
1001 snow index, Remote Sensing of Environment, 89(3), 351–360, doi:10.1016/j.rse.2003.10.016, 2004.

1002 Sappington, J. M., Longshore, K. M. and Thompson, D. B.: Quantifying Landscape Ruggedness for Animal Habitat
1003 Analysis: A Case Study Using Bighorn Sheep in the Mojave Desert, Journal of Wildlife Management, 71(5), 1419–
1004 1426, doi:10.2193/2005-723, 2007.

1005 Schirmer, M., Wirz, V., Clifton, A. and Lehning, M.: Persistence in intra-annual snow depth distribution:
1006 1. Measurements and topographic control, Water Resources Research, 47(9), W09516–n/a,
1007 doi:10.1029/2010WR009426, 2011.

1008 Schneider, D. and Molotch, N. P.: Real-Time Estimation of Snow Water Equivalent in the Upper Colorado River
1009 Basin using MODIS-based SWE Reconstructions and SNOTEL data, Water Resources Research,
1010 doi:10.1002/2016WR019067, 2016.

1011 Slater, A. G., Barrett, A. P., Clark, M. P., Lundquist, J. D. and Raleigh, M. S.: Uncertainty in seasonal snow
1012 reconstruction: Relative impacts of model forcing and image availability, Advances in Water Resources, 55(0), 165–
1013 177, doi:10.1016/j.advwatres.2012.07.006, 2013.

1014 Stewart, I. T., Cayan, D. R. and Dettinger, M. D.: Changes toward earlier streamflow timing across western North
1015 America, *Journal of Climate*, 18(8), 1136–1155, 2004.

1016 Sturm, M. and Wagner, A. M.: Using repeated patterns in snow distribution modeling: An Arctic example, *Water*
1017 *Resour. Res.*, 46(12), n/a–n/a, doi:10.1029/2010WR009434, 2010.

1018 Sturm, M., Taras, B., Liston, G. E., Derksen, C., Jonas, T. and Lea, J.: Estimating Snow Water Equivalent Using
1019 Snow Depth Data and Climate Classes, *Journal of Hydrometeorology*, 11(6), 1380–1394,
1020 doi:10.1175/2010JHM1202.1, 2010.

1021 Trujillo, E., Ramirez, J. A. and Elder, K. J.: Scaling properties and spatial organization of snow depth fields in sub-
1022 alpine forest and alpine tundra, *Hydrological Processes*, 23(11), 1575–1590, doi:10.1002/hyp.7270, 2009.

1023 Veitinger, J., Sovilla, B. and Purves, R. S.: Influence of snow depth distribution on surface roughness in alpine
1024 terrain: a multi-scale approach, *The Cryosphere*, 8(2), 547–569, doi:10.5194/tc-8-547-2014, 2014.

1025 Walters, R. D., Watson, K. A., Marshall, H.-P., McNamara, J. P. and Flores, A. N.: A physiographic approach to
1026 downscaling fractional snow cover data in mountainous regions, *Remote Sensing of Environment*, 152, 413–425,
1027 doi:10.1016/j.rse.2014.07.001, 2014.

1028 Xia, Q., Gao, X., Chu, W. and Sorooshian, S.: Estimation of daily cloud-free, snow-covered areas from MODIS
1029 based on variational interpolation, *Water Resour. Res.*, 48(9), 1480–1488, doi:10.1029/WR017i005p01480, 2012.

1030 Zou, H. and Hastie, T.: Regularization and variable selection via the elastic net, *Journal of the Royal Statistical*
1031 *Society Series B-Statistical Methodology*, 67(2), 301–320, doi:10.1111/j.1467-9868.2005.00503.x, 2005.

1032

1033