



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 ($r^2=0.85$, %MAE=33%, %Bias=1%) compared with models fit on the same day, by considering
18 variations in SWE depth as expressed by differences in fSCA between years. The crux of this proposition is in selecting
19 the model to transfer. We offer a method with which to select a model from another year based on the similarity in
20 SWE distribution at existing snow pillows in the area. Comparison of the best transferred predictions and the selected
21 predictions results in a mild decrease in r^2 (0.02) and moderate increases in %MAE (15%) and %Bias (10%). The
22 results motivate further refinement in the technique used to select the best model because if these dates can be
23 identified then SWE can be modeled at accuracy levels equivalent to models generated from ASO data collected on
24 the day of interest. Lastly, we found that models from ASO observations of anomalously low snowpacks in 2015 still
25 transferred to other years, although the same cannot be said for the reverse. This research provides a first attempt at
26 extending the value of ASO beyond the observations and we expect ASO will continue to provide insights for
27 improving water resource management for years to come.

28



29 1 Introduction

30 The spatial distribution of snow water equivalent (SWE) largely controls the timing and magnitude of streamflow
31 (Stewart et al., 2004) and is important for plant ecology (Litaor et al., 2008) in snow-dominated catchments around
32 the world. In these catchments, accurate assessments of snowpack water storage are critical for ensuring robust
33 estimates of seasonal water supply. Nevertheless, SWE is poorly measured operationally with only sparse point
34 measurements on the order of one measurement location per thousand square kilometers of snow covered terrain.
35 Additional manual measurements in the form of snow courses add little information about the spatial variability and
36 typically occur on the order of only three times a season.

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

47 Airborne systems have significantly improved the ability to measure SWE distribution at high spatial
48 resolution and at extents relevant to water resource management (i.e. $> 100 \text{ km}^2$). However, most previous studies of
49 snow distribution using airborne data have been limited to snap-shots in time, limiting the ability to empirically
50 transfer observations to time periods outside of those directly sampled. Since 2013, the National Aeronautics and
51 Space Administration (NASA), Jet Propulsion Laboratory, Airborne Snow Observatory (ASO) has acquired weekly
52 observations of snow properties from approximately the time of annual peak SWE to the end of the snowmelt season
53 in the Tuolumne Basin, California (Painter et al., 2016). ASO measures snow depth by differencing the lidar-derived
54 surfaces from a snow-on and snow-off flight and infers albedo and snow extent based on spectroradiometric
55 measurements. An energy-balance model is used to estimate snowpack density, and subsequently convert snow depth
56 to SWE. These weekly observations of SWE distribution over multiple years represent a new opportunity for
57 understanding the spatial and temporal dynamics of snow distribution. Moreover, the intensive repeat sampling of
58 ASO may provide an opportunity to extend observed SWE patterns beyond the time periods directly observed by ASO
59 – a goal of the work presented here. In this context, the ability to extend expensive ASO data in time could dramatically
60 expand the applicability of ASO data to future time periods without incurring the costs associated with future airborne
61 acquisitions.

62 A potential application of ASO measurements relates to the possibility of developing statistical relationships
63 between ASO data and other snow and terrain data that are more routinely available. In this context, statistical models
64 have been extensively used to estimate relationships between snow point measurements and physiography (Balk and
65 Elder, 2000; Elder et al., 1998; Erickson et al., 2005; Fassnacht et al., 2003; López-Moreno and Nogués-Bravo, 2006;



66 Molotch et al., 2005; Schneider and Molotch, 2016). At small scales (i.e. $< 10 \text{ km}^2$), dedicated sampling of headwater
67 catchments has led to models that explain between 20% and 65% of the variability in snow depth based on
68 physiographic variables at 30m resolution (Balk and Elder, 2000; Elder et al., 1998; Erxleben et al., 2002).
69 Importantly, Erickson et al. (2005) found persistence in the topographic controls on snow depth distribution and
70 successfully parameterized a multi-year model relating the physiographic variables of a small headwater catchment to
71 annual peak snow depth by scaling the mean based on in situ measurements. Later studies have since confirmed an
72 inter-annual consistency in snow depth distribution based on high resolution lidar measurements (Deems et al., 2008;
73 Schirmer et al., 2011; Trujillo et al., 2009) and the inter-annual persistence of topographic controls on snow depth
74 distribution (Grünewald et al., 2013). These previous works suggest that relationships between lidar snow depth
75 measurements and topography may be useful for extending lidar measurements of snow in time.

76 Lidar is only capable of measuring snow depth and coincident density measurements are scarce, despite SWE
77 being the more important hydrologic variable. However, snow depth varies an order of magnitude more than density
78 and therefore largely controls the variability of SWE (Mizukami and Perica, 2008). Hence, we expect the results from
79 the previous studies be relevant here, i.e. inter-annual consistency in SWE distribution will be similar to that of snow
80 depth. Given the difficulty of extensively measuring density and SWE, operational SWE observation networks that
81 use snow pillows to measure SWE have been used to relate physiography and SWE. In this regard, multiple studies
82 have explained up to 82% of the variability in SWE based on physiographic variables (Fassnacht et al., 2003;
83 Harshburger et al., 2010; Schneider and Molotch, 2016). These studies aimed to understand the processes controlling
84 snow distribution and to apply this knowledge to interpolate point observations of SWE to $>1000 \text{ km}^2$ for a single
85 point in time. The spatio-temporal dataset of SWE from ASO provides an unprecedented opportunity to develop
86 relationships between SWE and topography and test their persistence across several years.

87 Given that topographic variables are largely static in time, additional time-variant variables should be useful
88 in the context of explaining the spatio-temporal distribution of SWE. In this context, remotely sensed snow covered
89 area data has long been recognized to provide information with regard to snowpack water storage and consequently
90 expected summer streamflow (Good and Martinec, 1987; Martinec and Rango, 1981; Potts, 1937; 1944). Currently,
91 SCA is commonly estimated from SWE in hydrologic models through a depletion curve parameterization in order to
92 constrain melt production to the areal extent of snow cover (Anderson, 1973; Clark et al., 2011; Lawrence et al., 2011;
93 Livneh et al., 2010; Luce and Tarboton, 2004; Niu et al., 2011). The utility of depletion curves to provide sub-model-
94 scale information in physically-based modeling suggests that fSCA should provide additional information in statistical
95 models of SWE distribution. The consistent relationship between fSCA and SWE is predicated on the fact that SWE
96 distribution is extremely heterogeneous over complex terrain. Upon melt out, terrain features are progressively
97 uncovered. This process varies only slightly each year because of similarities in the meteorology, e.g. wind direction,
98 that drive accumulation patterns and solar exposure that drives melt out (Luce and Tarboton, 2004). Snow covered
99 area is also relatively easy to measure due to the distinctive spectral signature of snow compared to soil, rock, and
100 vegetation. In fact, photographic estimates of fSCA have been utilized for seven decades within hydrologic
101 applications (Parsons and Castle, 1959; Potts, 1937; 1944) and today robust observations of fSCA can be obtained
102 from variety of ground-based, aerial and satellite optical imagers (Bloschl et al., 1991; Dozier et al., 1981; Hall et al.,



103 2001; Kirnbauer and Blöschl, 1994; König and Sturm, 1998; Painter et al., 2009; Rittger et al., 2013; Rosenthal and
104 Dozier, 1996).

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

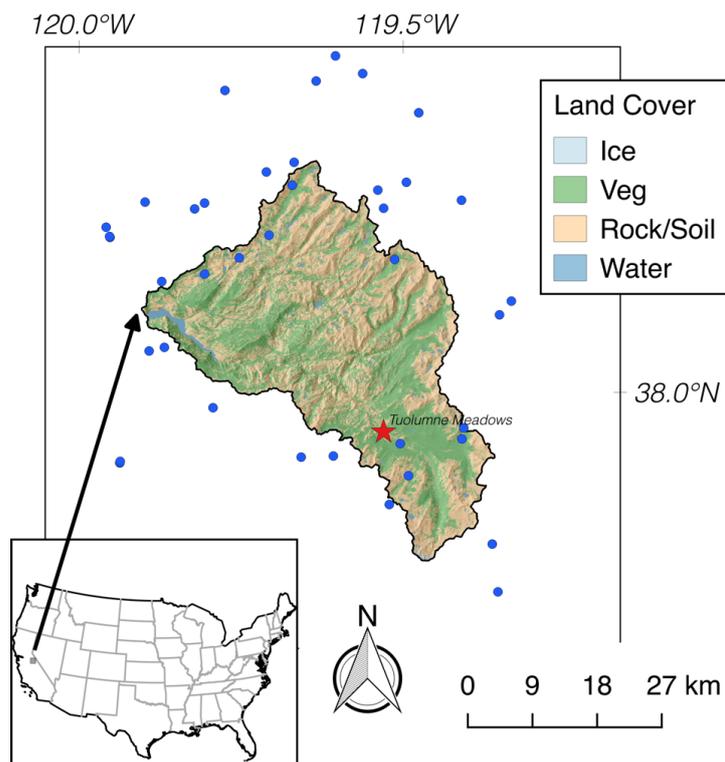
114 We present our SWE distribution modelling framework and show the utility of including the time-variant
115 variable fSCA for improving the SWE distribution estimates. Further, we evaluate the impact of transferring models
116 from one year to another. Lastly, we present a methodology for identifying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

119 **2 Site Description**

120 We used a SWE dataset from the Tuolumne River basin in the Sierra Nevada mountains in California, USA (Fig. 1).
121 The basin is 1,175 km² in area, consisting of 48% vegetation, 50% rock, 2% water, and small isolated areas with
122 permanent snow/ice. The elevation range is 1127 m to 3965 m, encompassing 4 distinct ecological zones ranging from
123 lower montane forest to alpine (NPS, 2016). The lower montane forest ranges from 1127 m to 1800 m elevation and
124 consists of a diverse mix of coniferous and deciduous trees. The upper montane forest ranges from 1800 m to 2450 m
125 elevation and primarily consists of coniferous species such as red fir and lodgepole pine. Elevations from 2450 m to
126 2900 m are considered subalpine and consist of a mix of meadows and coniferous forest. The highest elevation band
127 above 2900 m is an alpine zone that is devoid of tree cover and contains limited herbaceous vegetation. This alpine
128 zone contains areas with large granitic features, talus slopes and boulder fields. Snowmelt from the basin runs off into
129 the Hetch Hetchy reservoir, which is the main water supply for the City of San Francisco.

130

131



132

133 **Figure 1. Land cover map of the Tuolumne Basin with snow pillows shown as blue dots. The land cover data was provided**
134 **by ASO and was derived from spectral information from summer snow-off flight. The snow pillows locations were obtained**
135 **from the California Department of Water Resources.**

136 3 Data Sources

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

142 In order to conduct our analysis at a spatial scale that will ensure transferability to more widely available
143 data, we mean-aggregated the 50 m SWE maps to 500 m. Subsequently, we converted the 3 m DEM and snow depth
144 maps to 501 m using mean aggregation and then bilinearly resampled to the 500 m SWE grid. We computed the
145 physiographic variables used in the modelling framework [Table 1] from the 500 m DEM using open source GIS
146 software, including GDAL (GDAL Development Team, 2015), SAGA GIS (Conrad et al., 2015), and R (R Core
147 Team, 2015). We chose the 500 m resolution because it is relevant to water resource management and is the scale for
148 which daily satellite fSCA images are available from the Moderate Resolution Imaging Spectroradiometer, which is
149 a commonly used fSCA data product (Painter et al., 2009; Salomonson and Appel, 2004). We also used a binary



150 aggregation of the 3 m snow depth maps to obtain 501 m fSCA maps, which were then resampled to 500 m by nearest
 151 neighbor to preserve snow free pixels. Lastly, we aggregated and bilinearly resampled the vegetation height map to
 152 500 m.

153

154 **Table 1. List of physiographic variables considered in the multiple linear regression to model SWE distribution. Source**
 155 **includes studies in which these variables have been used and the source of the algorithm, if applicable. Citations in the**
 156 **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)
zness	sine(slope); ranges 0-1; dimensionless	Balk and Elder (2000); Erxleben et al. (2002); Fassnacht et al. (2003)
northness	cosine(aspect); ranges 0-1; dimensionless	Balk and Elder (2000); Erxleben et al. (2002); Fassnacht et al. (2003)
eastness	sine(aspect); ranges 0-1; dimensionless	Balk and Elder (2000); Erxleben et al. (2002); Fassnacht et al. (2003)
topographic position index (TPI)	elevation difference of a pixel from the mean of the surrounding pixels; meters	Revuelto et al. (2014); GDAL (2015)
vector ruggedness measure (VRM)	3-dimensional measure of the variation of slope and aspect; 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 3x3 window around each pixel; shown to detect changes in slope at multiple scales; radians	Marchand and Killingveit (2005); Lopez 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)

157

158

159 The ASO flies approximately weekly from near peak SWE to the end of the melt season. As such, there were
 160 six flights in 2013, ten in 2014, eight in 2015, and eight in 2016. Only the final four flights of the 2016 season were
 161 available for this study resulting in a total of 28 SWE maps. The 2014 and 2015 seasons were characterized by a severe
 162 dry snow drought (Harpold et al., 2017) and 2013 and 2016 also experienced below average snowpack conditions, but
 163 less severely. Painter et al. (2016) report a mean absolute vertical snow depth error of 8 cm and a bias of 1 cm when
 164 compared with manually measured snow depths at the 15 × 15 m scale. Further details about the mission and
 165 processing can be found in Painter et al. (2016).

166

167

168

169

We also obtained daily SWE measurements from 54 snow pillows operated by the California Department of
 Water Resources that are within 20 km of the Tuolumne watershed boundary. The stations range from 2000 m to 3250
 m. We downloaded the adjusted SWE records, which have been manually quality controlled, for the 2013-2016 water
 years. No further adjustments were performed. The data can be downloaded from <http://cdec.water.ca.gov/>.

170 4 Methods

171

172

We use linear regression to model the distribution of SWE for every ASO flight. The explanatory variables we consider
 are ASO-observed fSCA and physiographic variables previously used in the literature (Table 1). We present results



173 from two models: 1. “PHV” is a multiple linear regression that consists only of physiographic variables as independent
174 variables. This represents the traditional approach for estimating SWE distribution based on relationships with
175 physiography; 2. “PHV-FSCA” is a multiple linear regression that includes both physiographic variables and fSCA
176 as independent variables. We make these distinctions to demonstrate the utility of fSCA as a time-variant variable. No
177 one has explicitly explored the utility of fSCA for directly estimating the distribution of SWE. fSCA values greater
178 than and equal to zero are used to fit the PHV-FSCA models. Regression estimates from both PHV and PHV-FSCA
179 are masked to snow covered areas as observed by ASO. Once we illustrate the utility of the statistical models for
180 characterizing ASO SWE distribution patterns at discrete points in time, we then show how these statistical models
181 can be transferred to time periods without ASO observations. With regard to all statistical models, we report the
182 squared Pearson correlation coefficient (r^2) as a measure of the relative spatial pattern between the modeled SWE
183 distribution and ASO observed SWE distribution. We also report the mean absolute error as a percent of mean
184 observed SWE (%MAE) as a measure of the accuracy of the modeled SWE distribution. Lastly, we report bias as a
185 percent of mean observed SWE (%Bias) as a measure of the systematic over or under-prediction by the model.

186 4.1 SWE Models: Discrete Time

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

196 4.2 SWE Models: Transferred in Time

197 We evaluated a second set of predictions whereby each date is modeled using all the data, i.e. not split, and then this
198 model is used to predict SWE on dates that ASO flew in different years. In this manner, we simulate SWE on the date
199 ASO flew using models from other years and then we use the ASO data on the date of interest strictly to evaluate the
200 model estimates of SWE. Hereafter, we refer to the date of the model (i.e. the date of the ASO observation for which
201 the model is developed) as the *model date* and the date being predicted as the *transfer date*. This results in 28 models
202 of SWE distribution for PHV and PHV-FSCA each because there are a total of 28 ASO flights. Given our primary
203 goal of estimating the SWE distribution for unsampled dates, we apply models developed for each ASO acquisition
204 to all other dates except for dates within the same year as that in which the model was developed. For example, in
205 2013 there were a total of 6 ASO flights out of the total of 28 flights during the four-year study period. This leaves 22
206 flights from other years that can be used to develop statistical models of SWE that can be transferred to the dates in
207 2013. By conducting our model tests in this manner we are more robustly testing the transferability of models from a



208 given *model date* with regard to simulating SWE distribution on a given *transfer date*. Due to a different number of
209 ASO observations in each year, the prediction ensemble size differs for each year. As noted above, for each date in
210 2013 there are 22 potential models. For each date in 2014 there are 18 potential models; for each date in 2015 there
211 are 20 potential models; and for each date in 2016 there are 24 potential models. These models are referred to as
212 *transferred models*, and for each transfer date we identify the best model from another year based on error statistics
213 generated from the ASO data acquired on the transfer date. We refer to this model as the *best model*.

214 The ability of a model to transfer from the one date to another will vary based on how well the relationships
215 between the dependent variable (SWE) and explanatory variables are captured and the similarity of the SWE
216 distributions. Here we quantify the SWE similarity between dates using the mean absolute error (MAE) of SWE
217 recorded at nearby snow pillows. For each transfer date, there is an ensemble of predictions from the model dates from
218 the other years. Each of the model dates exhibit a similarity with the SWE distribution of the transfer date. In order to
219 pick which model date exhibits the greatest similarity with the transfer date without having an ASO observation, we
220 calculate the MAE of SWE at the snow pillows between each pair of model-transfer dates and select the model date
221 with the lowest MAE. We compare the prediction performances from this model selection procedure (denoted *selected*
222 *model*) with those of the best models.

223 4.3 Statistical Model

224 The multiple linear regression models described above, i.e. PHV and PHV-FSCA, are based on a regularized
225 regression model applied in an elastic net framework as implemented in the glmnet package in R (Friedman et al.,
226 2010). The benefit of a regularized regression over standard regression is that it reduces overfitting while permitting
227 all conceptual variables to be included, rather than removing potentially useful variables due to multicollinearity.
228 Regularized regression increases the predictive ability of a model with multiple predictor variables by penalizing the
229 objective function used to estimate the parameter set. The elastic net is an extension of ordinary least squares, which
230 estimates parameter coefficients by minimizing the residual sum of squares (RSS) as the objective function (Eqn 1):
231

$$232 \quad \text{RSS} = \sum_{i=1}^n \left(y_i - \sum_{j=1}^p \beta_j x_{ij} \right)^2 \quad \text{Eqn (1)}$$

233
234 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
235 variable j at each observation i . The elastic net penalizes RSS by two different types of regularization techniques,
236 known as L1 and L2, that have opposing properties (more on this below). In this regard, the elastic net estimates the
237 regression parameters β by minimizing RSS in Eqn 2:
238

$$239 \quad \text{RSS} = \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda P_{\alpha}(\beta) \quad \text{Eqn (2)}$$



240

241 where

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

243

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

249 The elastic net has two tuning parameters, λ and α , which are determined through cross validation. When
250 $\alpha=1$, the penalty is composed completely of the L1 penalty and commonly known as Lasso regression. When $\alpha=0$,
251 the penalty is composed completely of the L2 penalty and is commonly known as Ridge regression. The advantage of
252 the elastic net is that alpha can range between 0 and 1 and therefore inherits the properties of both L1 and L2
253 regularization. L1 regularization is commonly used for model selection because predictor coefficients can be shrunk
254 to zero and effectively removed from the model. However, in the presence of correlated predictor variables, one
255 predictor variable would be randomly selected while the others are removed. This can result in decreased predictive
256 performance since variables with some explanatory power are no longer in the model. With L2 regularization,
257 regression coefficients will shrink towards zero but with an asymptote at zero. This is the preferred type of
258 regularization in the presence of multicollinearity because all variables would remain in the model but with smaller
259 coefficients. The elastic net provides a framework to choose the best compromise between the L1 and L2 penalties.
260 For further details, we direct the reader to Zou and Hastie (2005b) and Hastie et al. (2009).

261 We do not directly treat spatial correlation in our models due to the large computational demands of fitting
262 the covariance function for ~4000 pixels for 28 dates. Neglecting spatial correlation is another potential source for
263 regression coefficients to be overfit and consequently we do not interpret them for physical meaning (Cressie, 1993;
264 Erickson et al., 2005). Nonetheless, we show utility with our methods without addressing spatial correlation and expect
265 the results presented herein would improve if spatial correlation were explicitly treated (Carroll and Cressie, 1997).

266 5 Results

267 5.1 SWE Models: Discrete Time

268 Table 2 shows that PHV-FSCA outperforms PHV in all metrics in all years except %Bias (where both models exhibit
269 close to 0 %Bias, as expected from a regression). The model PHV-FSCA explains on average between 78% and 86%
270 of the variance in SWE distribution in any given year whereas PHV only explains between 55% and 67%. We similarly
271 see improvement in %MAE where PHV-FSCA exhibits mean annual %MAE between 27% and 41% compared with



272 PHV which yields mean annual %MAE between 50% and 71%. In summary, we note a substantial improvement in r^2
 273 and %MAE for distributing SWE when fSCA is included as an additional variable.

274

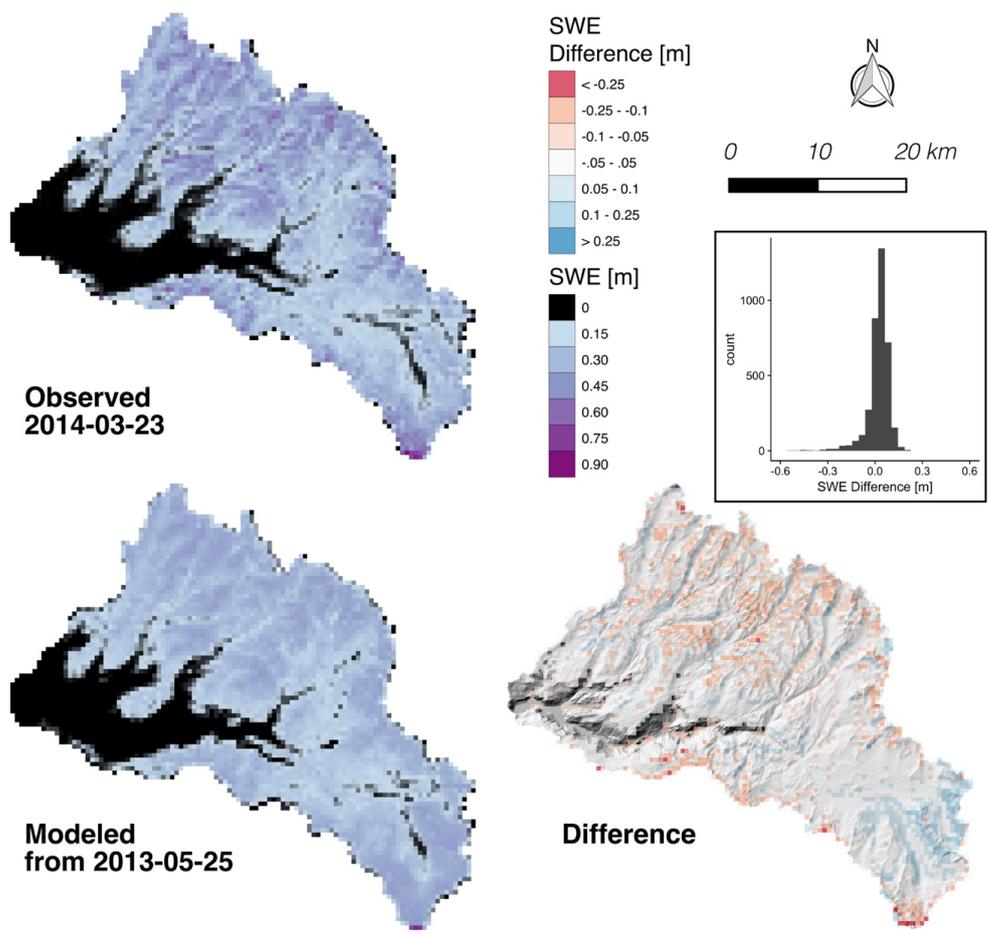
275 **Table 2. Mean prediction performance for PHV and PHV-FSCA from each observation date using a split-sampling**
 276 **approach.**

	PHV			PHV-FSCA		
	r^2	%MAE	%Bias	r^2	%MAE	%Bias
2013	0.55	71	0	0.86	33	1
2014	0.67	50	1	0.86	27	1
2015	0.57	64	-3	0.83	33	1
2016	0.59	71	2	0.82	41	2
Mean	0.6	61	0	0.85	32	1

277
 278

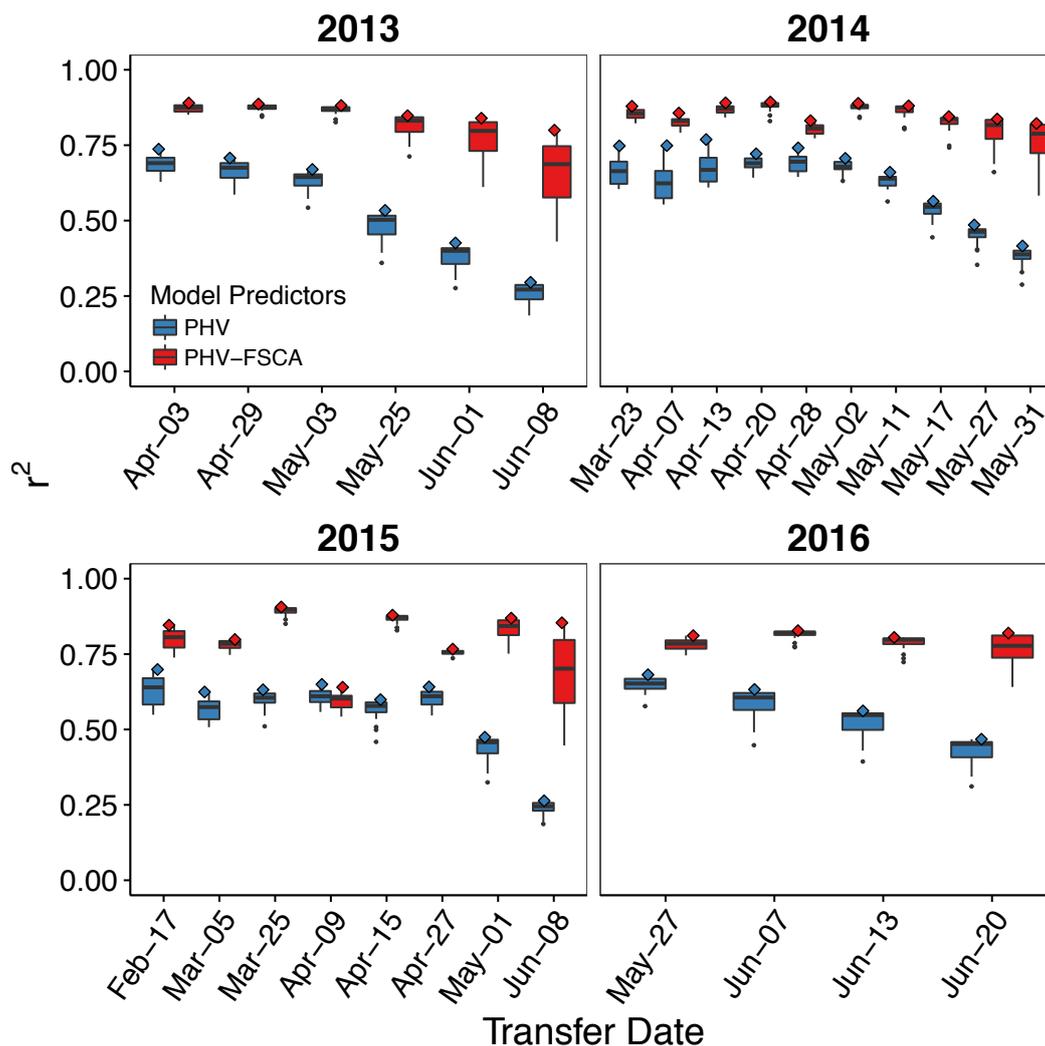
279 5.2 SWE Models: Transferred in Time

280 Figure 2 shows March 23, 2014 as an example transfer date in 2014 that was predicted using a PHV-FSCA model
 281 from May 25, 2013. Overall, we see similar spatial trends between the observed and modeled SWE distributions, but
 282 we see darker purples in the observed map indicating higher SWE. The mean observed SWE is 0.23 m compared to
 283 0.21 m modeled. The range of observed SWE is 0-0.75 m while the modeled SWE ranges from 0-0.39 m. The standard
 284 deviations of observed and modeled SWE are 0.13 m and 0.11 m, respectively. The difference map shows large areas
 285 of agreement to within 0.05 m SWE and a qualitative comparison with Figure 1 suggests these are mainly forested
 286 areas. We see areas of under prediction (red) mostly above tree line in the north and areas of over prediction (blue)
 287 above tree line in the south. A comparison with Google Earth® aerial imagery confirms that the pixels that exhibit
 288 very large negative differences (bright red pixels) are areas with persistent snow for much of the year. The snow extent
 289 is very similar between the modeled SWE and observed SWE because only areas observed to have fSCA greater than
 290 0 were predicted.



291
 292 **Figure 2.** An example of SWE distribution from 2014-03-23. Observed SWE (top left); modeled SWE using PHV-FSCA
 293 (bottom left); the difference, modeled SWE – observed SWE, draped over a shaded relief map (bottom right). A histogram
 294 of the differences (top right).

295
 296 Figure 3 shows the range of r^2 for transferred models for PHV and PHV-FSCA on each date based on models
 297 created in other years. Additionally, we indicate the best model performance for each transfer date by a diamond. In
 298 this regard, we observe unanimous improvement across all dates with PHV-FSCA compared to PHV. PHV-FSCA
 299 yields the highest mean best r^2 of 0.84 (mean of diamonds) compared to PHV with a mean best r^2 of 0.6. We also
 300 observe a notable decrease in r^2 for PHV towards the end of the season while PHV-FSCA exhibit a consistent r^2 . The
 301 standard deviation of r^2 for PHV is 0.14 and for PHV-FSCA it is 0.05.
 302



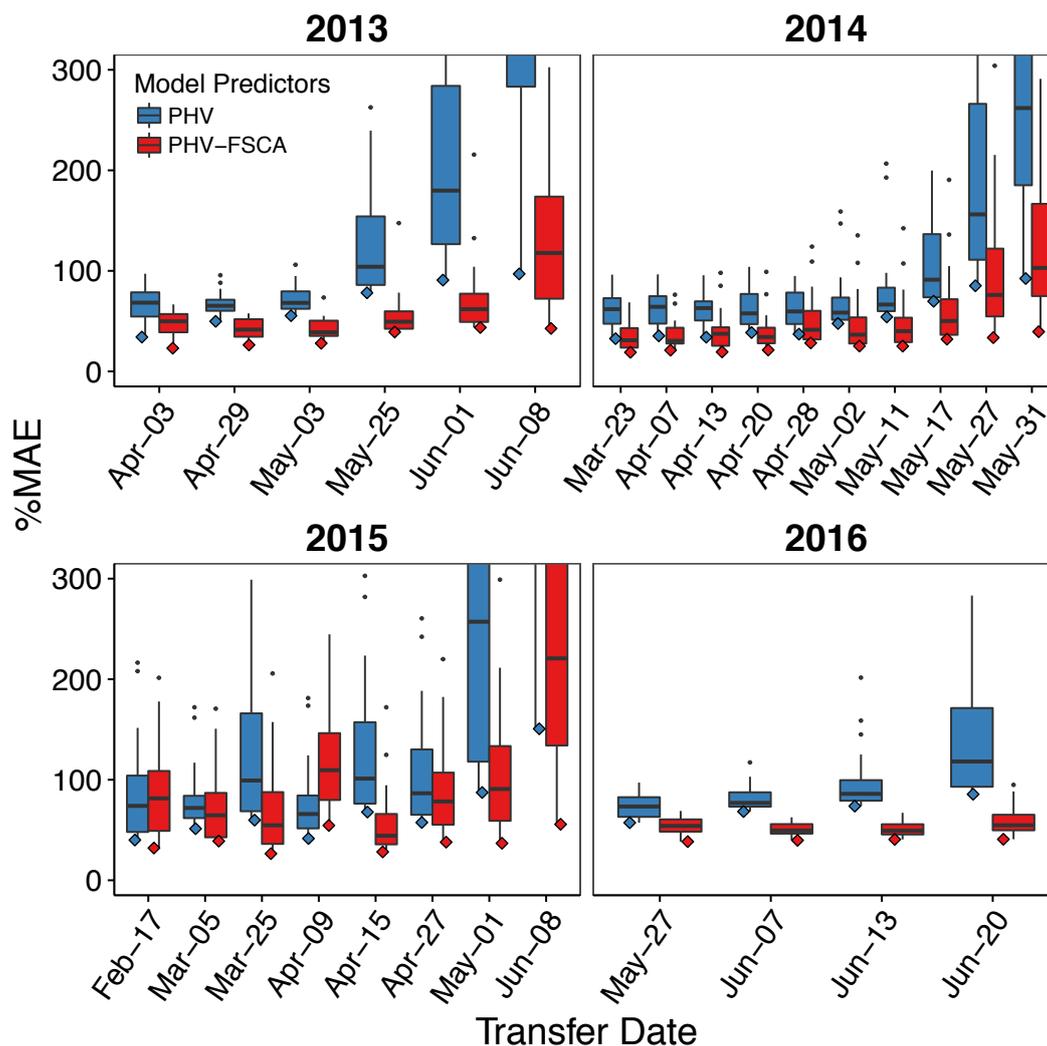
303

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

307

308 Figure 4 also clearly shows PHV-FSCA to exhibit the best, i.e. lowest, %MAE with transferred models
 309 compared to PHV. The mean best %MAE (mean of diamonds) for PHV-FSCA is 33% while the mean best %MAE
 310 for PHV is 63%. Particularly obvious in these panels is the upward distribution shift and larger range for PHV later in
 311 the season compared to only a minimal increase in %MAE for PHV-FSCA. The standard deviations of the best %MAE
 312 are 10% for PHV-FSCA and 26% for PHV.

313



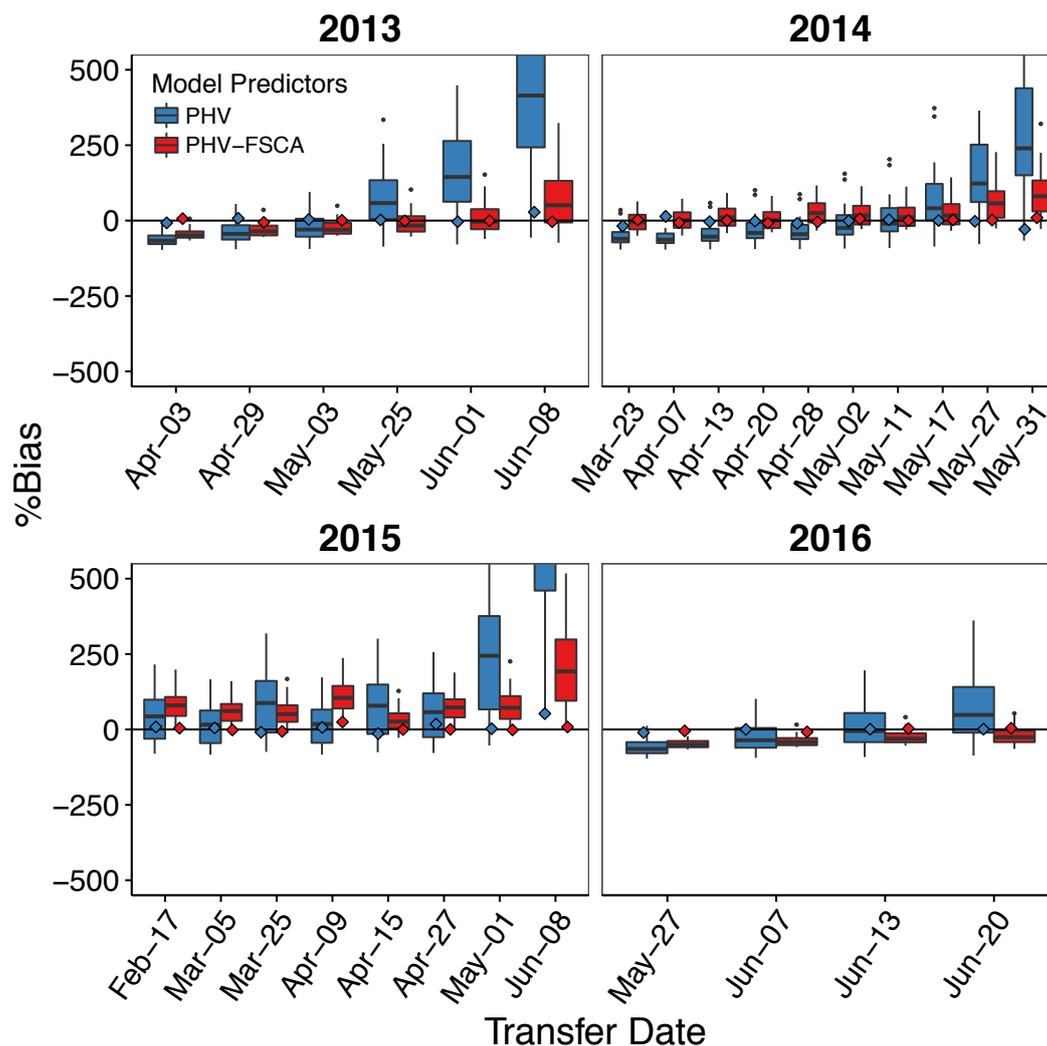
314

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

318

319 Figure 5 shows that the best transferred models (diamonds) for both all models exhibit close to zero bias. The
 320 mean best %Bias for PHV is 2% and for PHV-FSCA it is 1%. However, we note that the variability in %Bias increases
 321 more dramatically at the end of the season, especially for PHV. The standard deviations of the best (i.e. lowest) %Bias
 322 for PHV and PHV-FSCA are 15% and 7%, respectively.

323



324

325

326

327

Figure 5. The range of %Bias prediction errors for each transfer date from the transferred models. The diamond represents 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.

328

329

330

331

332

333

334

335

Following the demonstration of 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 6% of the split sample model with the mean 1% higher. The yearly mean magnitude of %Bias is actually lower,



336 i.e. better, in two years with the best transferred model compared to the split sample model, and the overall mean is
 337 the same.

338

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

	Split Sample Model			Best Transferred Model		
	r^2	%MAE	%Bias	r^2	%MAE	%Bias
2013	0.86	33	1	0.86	34	-1
2014	0.86	27	1	0.86	26	0
2015	0.83	33	1	0.82	39	3
2016	0.82	41	2	0.82	40	-1
341 Mean	0.85	32	1	0.85	33	1

342 **5.3 Evaluating the Selected Model Performance**

343 Table 4 summarizes and compares the yearly mean statistics from PHV-FSCA for the best models and the selected
 344 models transferred from another year. The model selection procedure results in similar yearly r^2 to the best models
 345 but increases in both %MAE and %Bias are apparent with the selected models. Compared to the best models, yearly
 346 mean %MAE for selected models increases between 3% and 18% and yearly absolute %Bias increases between 3%
 347 and 37%. The years 2013, 2014, and 2015 yielded increases in %MAE of 18%, 17%, and 16%, respectively for
 348 selected versus best models. In 2016 selected model %MAE exhibited an increase of only 3%. The absolute %Bias
 349 increases 3% in 2013, 6% in 2014, 37% in 2015, and 24% in 2016 for selected versus best models.

350

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

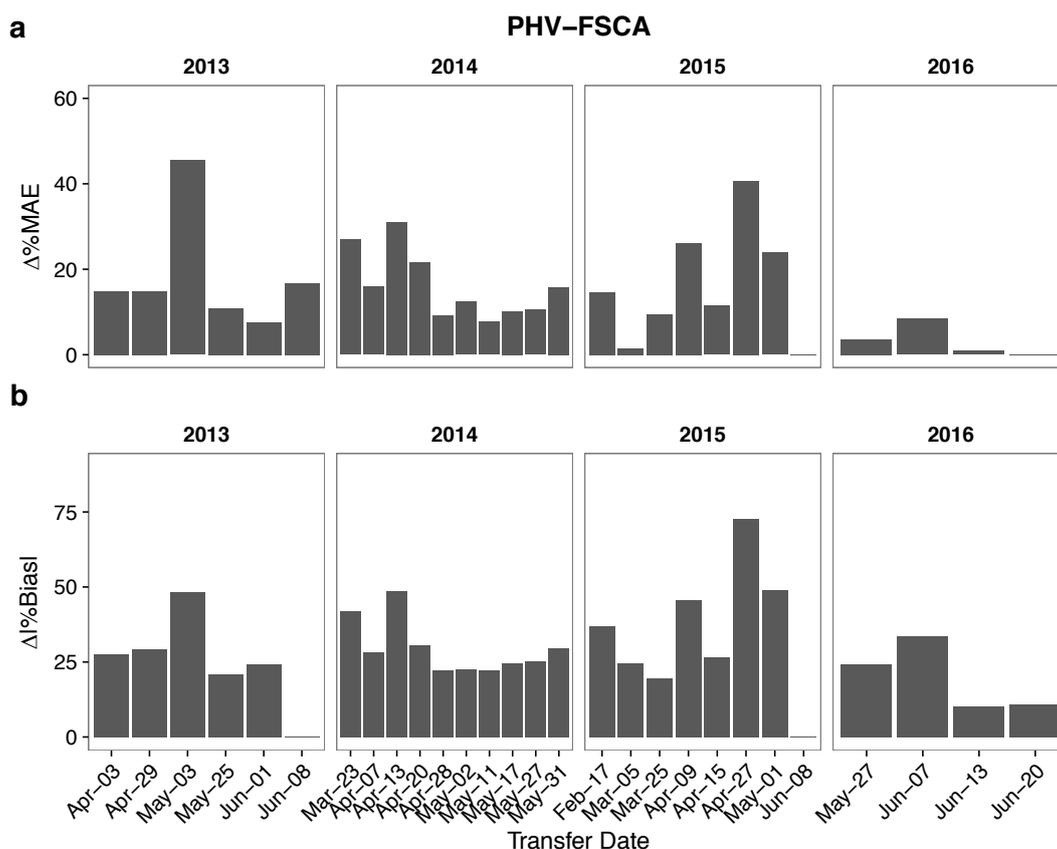
	Best Transferred Model			Selected Model		
	r^2	%MAE	%Bias	r^2	%MAE	%Bias
2013	0.86	34	-1	0.84	52	4
2014	0.86	26	0	0.84	43	6
2015	0.82	39	3	0.81	55	40
2016	0.82	40	-1	0.81	43	-25
356 Mean	0.85	33	1	0.83	48	11

357

358 Figure 6 shows the difference for each transfer date between the errors of the best models (diamonds in Figs.s
 359 4, 5) and the errors of the selected models (these are shown as open circles in Fig. 6). We focus on %MAE and %Bias
 360 from PHV-FSCA only because r^2 showed generally consistent performance for a given transfer date (Table 4;
 361 comparatively small vertical range of the purple boxplots in Fig. 3 compared with Figs. 4 and 5).



362 The mean difference in %MAE between the selected and best models is 15% and the selected model exhibited
 363 the same %MAE as the best model on only one date (Fig. 6a). The range in performance difference is between 0%
 364 and 45% with a mean of 15% and median of 12%. The differences in %MAE were generally lowest in 2016 with a
 365 mean of 3% and standard deviation of 4%. In contrast, 2013, 2014, and 2015 exhibited both higher mean differences
 366 (18%, 16%, and 18%, respectively) and higher standard deviations in %MAE (14%, 8%, and 14%, respectively).
 367 The best model had a lower absolute magnitude %Bias of between 0% and 73% with a mean of 29% and
 368 median of 26%. The selected model was the same as the best model on only one date (Fig. 6b). The yearly mean
 369 difference in error was consistently higher for %Bias than %MAE, with means of 25% in 2013, 30% in 2014, 34% in
 370 2015, and 20% in 2016. The standard deviation in the error difference was lowest in 2014 (9%) compared to 16% in
 371 2013, 22% in 2015, and 11% in 2016.



372
 373 **Figure 6. The increase in prediction error for %MAE (a) and %Bias (b) between the best model and the selected model.**

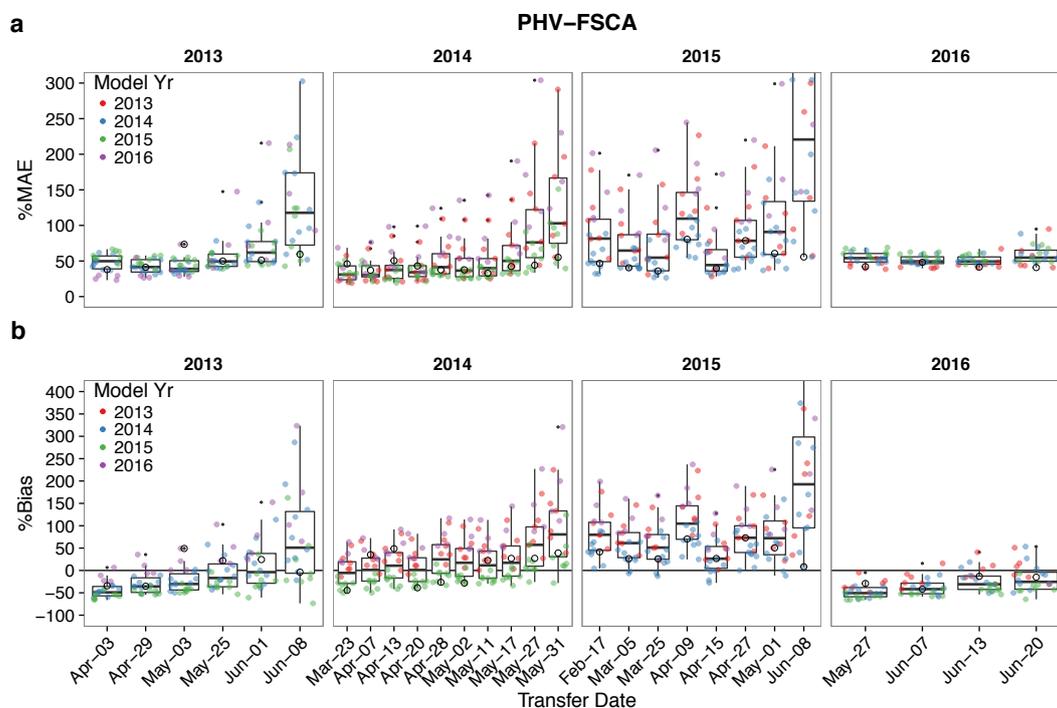
374
 375 We evaluate the prediction errors from different model years to see if there are any systematic differences in
 376 the predictions generated; i.e. do some years yield better predictions of other years? This allows us to determine how
 377 sensitive predictions for transfer dates will be given the existing ensemble of observations. In this context, we note



378 distinct differences in the predictive ability of models from different years, which has important implications for the
 379 ability of a model to transfer in time.

380 Figure 7a shows that, on average, models *from* 2015 produced the lowest %MAE and models from 2016
 381 produce predictions with the largest %MAE. The mean %MAE in 2016 is 118% compared to 79% with 2013 models,
 382 69% with 2014 models and 50% with 2015 models. We observe consistency in a year's ability to predict another year
 383 relative to the overall distribution, i.e. the colored dots are typically clustered within the range of the boxplots. We
 384 also note in 2016 that the best models always came from 2013, but in 2013 the best models only came from 2016 for
 385 the first three flights. In 2014, the bulk of the models around the median performance were from 2015 and vice-versa
 386 in 2015. In these two years, the poorest predictions were from 2013 and 2016.

387 Figure 7b show that models *from* 2016 also produce the largest %Bias with a mean of 97% compared with
 388 56% from 2013, 21% from 2014, and -26% from 2015. We again see consistency in the location of the colored dots
 389 relative to the boxplots thus suggesting years will consistently model the SWE distribution of certain other years
 390 better. In this regard, similar to %MAE, 2013 produces the lowest %Bias in 2016, but in 2013 the inverse is only true
 391 for the first three flights as the distribution shifts up.
 392



393
 394 **Figure 7. The prediction errors for PHV-FSCA for each date; %MAE (a) and %Bias (b). The colored dots are the prediction**
 395 **errors coded by model year and jittered to prevent over plotting. Open circles represent the selected model. The boxplots**
 396 **represent the entire distribution of prediction errors for each date. The boxes represent the interquartile range with vertical**
 397 **bars for the 5th and 95th percentiles. Small black dots are statistical outliers.**



398 **6 Discussion**

399 **6.1 SWE Distribution Modeling**

400 The relationship between snow covered area and SWE is well established as evident by the common use of depletion
401 curves in hydrologic modeling (Anderson, 1973; Clark et al., 2011; Liston, 1999; Luce et al., 1999). We invert the
402 idea of the depletion curve in this study by using the spatial distribution of fSCA to predict the spatial distribution of
403 SWE. The basis for this approach is relatively well established given that fSCA is sensitive to topographic complexity
404 and the spatial distribution of SWE (Donald et al., 1995; Fassnacht et al., 2016; Niu and Yang, 2007). In this regard,
405 model performance consistently improved when fSCA was used as a predictor, with an average %MAE decrease from
406 61% with PHV to 32% with PHV-FSCA. The utility of fSCA as an explainer of SWE distribution is not an
407 unprecedented result. König and Sturm (1998) showed that in situ measurements of SWE were correlated with fSCA
408 from aerial photographs. In addition, Marchand et al. (2005) suggested that the sub-grid standard deviation of
409 physiography could improve regression models of SWE distribution because it accounts for sub-grid snow depth
410 variability. Remotely sensed fSCA provides a means of capturing this sub-grid variability in SWE without requiring
411 higher resolution data to compute the variance of physiography for each pixel.

412 The statistics for the selected PHV-FSCA model reported in this study compare favorably to SWE
413 distribution statistics reported previously. Headwater catchment scale studies, based on intensive field data, have been
414 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;
415 Erxleben et al., 2002; López-Moreno and Nogués-Bravo, 2006; Molotch and Bales, 2005) compared to the mean r^2 of
416 0.83 for the selected models in this study. These papers, which cover only a few square kilometers, represent a far
417 more simplistic problem with regard to characterizing relationships between snow accumulation and physiographic
418 variables.

419 The results presented herein compare favorably with larger scale studies (i.e. $> 1000 \text{ km}^2$) of snow
420 distribution. Fassnacht et al. (2003) reported average yearly RMSE between 0.12 and 0.16 m and 0 m bias when cross-
421 validated with snow pillows. Harshburger et al. (2010) reported an average r^2 of 0.82 and RMSE of 0.05 m when
422 cross-validated with snow pillows. Schneider and Molotch (2016) reported a mean RMSE of 0.23 m and %Bias of
423 0.8% from snow surveys in the Upper Colorado River Basin. This is compared to a mean RMSE of 0.07 m and mean
424 %Bias of 11% for the selected model in this study. The favorable error statistics reported here are even more
425 encouraging when considering the differences in evaluation methods of these previous studies. In this regard, the error
426 values reported here are quite robust in that we compare against spatially explicit observations over relatively large
427 spatial scales (i.e. $> 1000 \text{ km}^2$). In contrast, the aforementioned works were evaluated against relatively sparse
428 observations (Fassnacht et al., 2003; Harshburger et al., 2010; Schneider and Molotch, 2016).

429 Bair et al. (2016) compared a retrospective SWE reconstruction to the same ASO observations (2013-2015)
430 and reported yearly mean %MAE between 20% and 31% and yearly mean %Bias between -11% and 10%. These
431 yearly statistics are better than those reported in this study (Tables 3, 4), but are the result of a much more complicated
432 energy balance model that can only be run after the snow has disappeared. The selected model in this study is a simple
433 linear regression that can be applied in real-time, thus we consider our results valuable for applications where real-
434 time estimates of SWE distribution are needed. Furthermore, we compare our selected model results to SWE estimates



435 from the U.S. National Weather Service's operational Snow Data Assimilation System (SNODAS). SNODAS
436 produces spatially distributed SWE estimates for the coterminous United States at 1 km by assimilating a physically
437 based model with SNOTEL observations and remotely sensed snow covered area. The SWE product from SNODAS
438 is the only high-resolution, gridded SWE product available at a daily time step for the continental United States and
439 is available from <http://nsidc.org> (Barrett, 2003). Previous work has shown the physically based model to perform
440 well at the point scale (Rutter et al., 2008) but suffer in alpine zones because it does not consider wind redistribution
441 (Clow et al., 2012). The yearly mean r^2 between ASO and SNODAS ranges from 0.04 in 2016 to 0.36 in 2015 with a
442 mean of 0.17. The yearly mean %MAE ranges from 120% in 2014 to 274% in 2013 with a mean of 199%. The yearly
443 mean %Bias ranges from -10% in 2016 to 236% in 2015. We refer the reader to "Selected Model" of Table 4 for the
444 PHV-FSCA error summary. In this regard, SNODAS exhibits a mean %MAE 4 times greater than that of PHV-FSCA
445 and a mean %Bias 8 times higher than PHV-FSCA. Both models poorly predicted the anomalous conditions in 2015.
446 While it is clear that the errors with PHV-FSCA are considerably lower than with SNODAS overall, SNODAS is a
447 complex system that attempts to capture the snow dynamics across the entire United States compared with PHV-
448 FSCA which was trained using a very specialized data set in the study region.

449 We also show that SWE distributions can be related to fSCA and physiography in one year and applied to
450 another year. The performance of PHV-FSCA was quite similar when applied to the date at which the model was
451 trained (i.e. discrete time models with a split sample) versus applying the model to other years (i.e. transferred models).
452 In this regard, we see a minimal decrease in prediction skill and minimal increase in prediction error when we compare
453 split sample models with the best transferred models. Recall that the split sample model was trained and tested on the
454 same day and the transferred model was trained in a different year from which the model was applied. Table 3 shows
455 the mean r^2 in the split sample model to be equivalent to that of the best transferred model. Moreover, the average
456 %MAE of the best transferred model exhibits only a 1% difference from the split sample model and the mean %Bias
457 exhibits no difference. Thus, if we are able to identify the best model for a given date of interest we would see minimal
458 degradation in predictive ability relative to a model derived from data acquired on the date of interest. However, we
459 see significant differences in predictive ability from models of different years (Fig. 7) and conclude that relationships
460 between SWE and physiography are only similar between specific years, not as uniformly as put forth by previous
461 studies (Erickson et al., 2005; Grünewald et al., 2013). Also, it is unclear as to the impact of climate non-stationarity
462 with respect to the ability to transfer models to future years. Even so, for each year in this dataset there exists a
463 corresponding year from which accurate predictions can be made.

464 The benefits of using fSCA as a predictor variable in the transferred models are particularly large at the end
465 of the season when the errors are highest (Figs. 3, 4, 5). For the last 2 dates of each year, the average difference in
466 %MAE between the best PHV model and best PHV-FSCA model was 54% compared to an average difference of 20%
467 for the other dates. The improvements seen by including fSCA as a predictor are noteworthy because the ensemble of
468 models trained using ASO observations could then be applied using remotely sensed fSCA from satellites. The degree
469 to which satellite-based fSCA will improve model performance toward the end of the snowmelt season will be partially
470 dictated by the accuracy of the fSCA data, which is subject to increasing uncertainty at low fSCA values (Painter et
471 al., 2009; Rittger et al., 2013). Optical fSCA products such as MODSCAG also suffer in forested areas since snow



472 cover is occluded by the canopy (Raleigh et al., 2013; Rittger et al., 2013). A viewable gap fraction correction is
473 typically used to extrapolate fSCA to the occluded portions of a pixel by assuming that fSCA is the same under the
474 canopy as it is in canopy gaps (Molotch and Margulis, 2008), and future studies should evaluate the sensitivity of
475 PHV-fSCA to this assumption. Cloud cover can also obscure a satellite's view of the snow and therefore making it
476 difficult to estimate the snow extent. However, Slater et al. (2013) showed that gaps of 5 or more consecutive days
477 are rare with MODIS thus suggesting that weekly estimates of SWE would be feasible. Lastly, omission and
478 commission errors of cloud identification can provide erroneous estimates of fSCA although Parajka and Blöschl
479 (2008) report an overall filtering accuracy of 96% in the Alps. Nonetheless, this is still an active area of research
480 (Dozier et al., 2008; Parajka and Bloeschl, 2008; Rittger et al., 2013; Xia et al., 2012).

481 **6.2 Considerations for Extending the ASO Record**

482 The value of ASO during the year flown is significant for water management because it provides high resolution SWE
483 information with low uncertainty compared to traditional estimation methods and therefore facilitates more accurate
484 water supply forecasts. The downside to ASO is the relatively high cost of operation compared to traditional
485 measurement campaigns. The work presented here provides a first step in realizing the value of ASO subsequent to
486 active operations. We find that the number of flights within a year affects the mean and variance of the predictions in
487 other years by <1%; this was quantified by iteratively selecting between 1 and the number of observations in a given
488 year 100 times and assessing the change in error. In other words, it is better to perform ASO once per year for 10 years
489 than 10 times in one year.

490 It is clear from our results that flights in one year do not necessarily transfer well to another year, e.g. 2016 to
491 2014 and 2016 to 2015 (Fig. 7). The model selection procedure relies on operational snow pillows to identify similar
492 patterns of SWE between historical dates (i.e. model dates) and the date of interest (i.e. transfer dates). The spatial
493 representativeness of these stations may have contributed to the general inability of these stations to select the best
494 historical date for a given date of interest. In this context, it is well established that these snow pillows may not
495 adequately represent the SWE of the surrounding terrain (Meromy et al., 2012; Molotch and Bales, 2006; Rice et al.,
496 2011). However, the minimal degradation in prediction performance when considering the best transferred model
497 should motivate improvements for identifying the dates from the past with the most similar SWE distribution. As
498 shown here, if these dates can be identified, the SWE can be modeled at accuracy levels that are equivalent to models
499 generated from ASO data collected on the date of interest. We also tested the similarity in remotely sensed fSCA as a
500 method for identifying the historical date with the most similar SWE distribution to the date of interest. We found that
501 anomalous SWE and fSCA distributions make this method less robust than using snow pillow data. However, a
502 completely remote sensing based approach would be useful in data sparse regions where ground stations do not exist.
503 A potentially robust remote sensing based method could be to track fSCA through time and select a similar SWE
504 distribution based on the trajectory of fSCA rather than a single snapshot of fSCA. It is also important to note that
505 flying ASO during a year with an anomalously low snowpack such as in 2015 in California does not necessarily reduce
506 the predictive capacity of models for future years. In our case, 2015 provided useful estimates of SWE distribution in



507 all the other years, including 2016, even though the models from the relatively wet year of 2016 did not transfer well
508 to the very dry year of 2015 (Fig. 7).

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

520 7 Conclusion

521 We estimated the relationships between SWE, physiography, and fSCA and show that the temporal consistency in
522 these relationships can be used to estimate SWE in years beyond the ASO observation record, with a mean r^2 of .85,
523 mean %MAE of 33%, and mean %Bias of 1%. The relationships transfer robustly in time with no degradation in r^2 or
524 %Bias and only 1% in %MAE when comparing predictions between models fit on the same day and models from a
525 different year. Models with fSCA as a predictor transfer better than those without, and we suggest that the inclusion
526 of fSCA provides information with regard to the variability of the SWE resulting from different accumulation
527 dynamics due to differences in terrain roughness. In this regard, the availability of satellite images of fSCA facilitate
528 the transfer of modeled relationships based on ASO observations to dates when no airborne snow depth measurements
529 exist. The crux of this proposition is in selecting the model to transfer. We offer a method with which to select a model
530 from another year based on the similarity in SWE distribution at existing snow pillows in the area. Comparison of the
531 best predictions and the selected predictions results in a mild decrease in r^2 (0.02) and moderate increases in %MAE
532 (15%) and %Bias (10%). The results presented above motivate further refinement in the technique used to select the
533 best model because if these dates can be identified then SWE can be modeled at accuracy levels equivalent to models
534 generated from ASO data collected on the day of interest. Lastly, although SWE distributions simulated in years with
535 anomalous SWE distributions (2014, 2015) had the highest errors, models from these years still yielded good
536 performance in 2013 and 2016. Thus, the benefit of ASO in anomalously dry years is two-fold: water managers receive
537 accurate information during a year that is difficult to model, but also these observed SWE distributions can be used to
538 simulate SWE distributions in future, less anomalous years. Overall, ASO provides an unprecedented observation of
539 the relationships between SWE, fSCA, and physiography. The ASO dataset facilitates improved understanding of
540 these relationships in both time and space and should lead to better information for water managers.

541



542 **8 Data Availability**

543 All data is available from the corresponding author at noah.molotch@colorado.edu.

544 **9 Acknowledgements**

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

548 **10 References**

- 549 Anderson, E.: Snow Accumulation and Ablation Model, US Department of Commerce, Silver Spring, MD. 1973.
- 550 Bair, E. H., Rittger, K. E., Davis, R. E., Painter, T. H. and Dozier, J.: Validating reconstruction of snow water
551 equivalent in California's Sierra Nevada using measurements from the NASA Airborne Snow Observatory, Water
552 Resour. Res., n/a–n/a, doi:10.1002/2016WR018704, 2016.
- 553 Balk, B. and Elder, E.: Combining binary decision tree and geostatistical methods to estimate snow distribution in a
554 mountain watershed, Water Resources Research, 36(1), 13–26, doi:10.1029/1999WR900251, 2000.
- 555 Barrett, A. P.: National Operational Hydrologic Remote Sensing Center SNOW Data Assimilation System (SNODAS)
556 Products at NSIDC, National Snow and Ice Data Center, Boulder, CO, USA. 2003.
- 557 Blöschl, G., Kirnbauer, R. and Gutknecht, D.: Distributed Snowmelt Simulations in an Alpine Catchment 1. Model
558 Evaluation on the Basis of Snow Cover Patterns, Water Resources Research, 27(12), 3171–3179,
559 doi:10.1029/91WR02250, 1991.
- 560 Bühler, Y., Marty, M., Egli, L., Veitinger, J., Jonas, T., Thee, P. and Ginzler, C.: Snow depth mapping in high-alpine
561 catchments using digital photogrammetry, The Cryosphere, 9(1), 229–243, doi:10.5194/tc-9-229-2015, 2015.
- 562 Carroll, S. S. and Cressie, N.: A comparison of geostatistical methodologies used to estimate snow water equivalent -
563 Reply, Journal of the American Water Resources Association, 33(1), 221–222, 1997.
- 564 Clark, M. P., Hendrikx, J., Slater, A. G., Kavetski, D., Anderson, B., Cullen, N. J., Kerr, T., Örn Hreinsson, E. and
565 Woods, R. A.: Representing spatial variability of snow water equivalent in hydrologic and land-surface models: A
566 review, Water Resources Research, 47(7), doi:10.1029/2011WR010745, 2011.
- 567 Clow, D. W., Nanus, L., Verdin, K. L. and Schmidt, J.: Evaluation of SNODAS snow depth and snow water equivalent
568 estimates for the Colorado Rocky Mountains, USA, edited by M. Pelto and K. RRichard, Hydrological Processes,
569 26(17), 2583–2591, doi:10.1002/hyp.9385, 2012.



- 570 Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V. and Boehner, J.:
571 System for Automated Geoscientific Analyses (SAGA) v. 2.1.4, Geoscientific Model Development, 8(7), 1991–2007,
572 doi:10.5194/gmd-8-1991-2015, 2015.
- 573 Cressie, N. A. C.: Statistics for spatial data, J. Wiley. 1993.
- 574 Deems, J. S., Fassnacht, S. R. and Elder, K. J.: Interannual Consistency in Fractal Snow Depth Patterns at Two
575 Colorado Mountain Sites, J. Hydrometeor, 9(5), 977–988, doi:10.1175/2008JHM901.1, 2008.
- 576 Deems, J. S., Painter, T. H. and Finnegan, D. C.: Lidar measurement of snow depth: a review, Journal of Glaciology,
577 59(215), 467–479, doi:10.3189/2013JoG12J154, 2013.
- 578 Donald, J. R., Soulis, E. D., Kouwen, N. and Pietroniro, A.: A Land Cover-Based Snow Cover Representation for
579 Distributed Hydrologic-Models, Water Resour. Res., 31(4), 995–1009, doi:10.1029/94WR02973, 1995.
- 580 Dozier, J.: Mountain hydrology, snow color, and the fourth paradigm, Eos Trans. AGU, 92(43), 373–374,
581 doi:10.1029/2011EO430001, 2011.
- 582 Dozier, J., Painter, T. H., Rittger, K. E. and Frew, J. E.: Time–space continuity of daily maps of fractional snow cover
583 and albedo from MODIS, Advances in Water Resources, 31(11), 1515–1526, doi:10.1016/j.advwatres.2008.08.011,
584 2008.
- 585 Dozier, J., Schneider, S. R. and McGinnis, D. F.: Effect of grain size and snowpack water equivalence on visible and
586 near-infrared satellite observations of snow, Water Resources Research, 17(4), 1213–1221,
587 doi:10.1029/WR017i004p01213, 1981.
- 588 Egli, L. and Jonas, T.: Hysteretic dynamics of seasonal snow depth distribution in the Swiss Alps, Geophys. Res. Lett.,
589 36(2), n/a–n/a, doi:10.1029/2008GL035545, 2009.
- 590 Elder, K., Dozier, J. and Michaelsen, J.: Snow accumulation and distribution in an Alpine Watershed, Water Resources
591 Research, 27(7), 1541–1552, doi:10.1029/91WR00506, 1991.
- 592 Elder, K., Rosenthal, W. and Davis, R. E.: Estimating the spatial distribution of snow water equivalence in a montane
593 watershed, Hydrological Processes, 12(10-11), 1793–1808, doi:10.1002/(SICI)1099-
594 1085(199808/09)12:10/11<1793::AID-HYP695>3.3.CO;2-B, 1998.
- 595 Erickson, T. A., Williams, M. W. and Winstral, A.: Persistence of topographic controls on the spatial distribution of
596 snow in rugged mountain terrain, Colorado, United States, Water Resour. Res., 41(4), W04014,
597 doi:10.1029/2003WR002973, 2005.
- 598 Erxleben, J., Elder, E. and Davis, R.: Comparison of spatial interpolation methods for estimating snow distribution in



- 599 the Colorado Rocky Mountains, *Hydrological Processes*, 16(18), 3627–3649, doi:10.1002/hyp.1239, 2002.
- 600 Fassnacht, S. R., Dressler, K. A. and Bales, R. C.: Snow water equivalent interpolation for the Colorado River Basin
601 from snow telemetry (SNOTEL) data, *Water Resour. Res.*, 39(8), 1208, doi:10.1029/2002WR001512, 2003.
- 602 Fassnacht, S. R., Sexstone, G. A., Kashipazha, A. H., Jasinski, M. F., Kampf, S. K., Thaden, Von, B. C. and López
603 Moreno, J. I.: Deriving snow-cover depletion curves for different spatial scales from remote sensing and snow
604 telemetry data, *Hydrological Processes*, 30(11), 1708–1717, doi:10.1002/hyp.10730, 2016.
- 605 Friedman, J., Hastie, T. and Tibshirani, R.: Regularization Paths for Generalized Linear Models via Coordinate
606 Descent, *Journal of Statistical Software*, 33(1), 1–22, 2010.
- 607 GDAL Development Team: GDAL - Geospatial Data Abstraction Library, Version 1.11.3. 2015.
- 608 Good, W. and Martinec, J.: Pattern recognition of air photographs for estimation of snow reserves, *Annals of*
609 *Glaciology*, 1987.
- 610 Grohmann, C. H., Smith, M. J. and Riccomini, C.: Multiscale Analysis of Topographic Surface Roughness in the
611 Midland Valley, Scotland, *Geoscience and Remote Sensing, IEEE Transactions on*, 49(1), 1200–1213,
612 doi:10.1109/TGRS.2010.2053546, 2011.
- 613 Grünewald, T., Stötter, J., Pomeroy, J. W., Dadic, R., Moreno Baños, I., Marturià, J., Spross, M., Hopkinson, C.,
614 Burlando, P. and Lehning, M.: Statistical modelling of the snow depth distribution in open alpine terrain, *Hydrol.*
615 *Earth Syst. Sci.*, 17(8), 3005–3021, doi:10.5194/hess-17-3005-2013, 2013.
- 616 Gutmann, E. D., Larson, K. M., Williams, M. W., Nievinski, F. G. and Zavorotny, V.: Snow measurement by GPS
617 interferometric reflectometry: an evaluation at Niwot Ridge, Colorado, *Hydrological Processes*, 26(19), 2951–2961,
618 2012.
- 619 Hall, D. K., Foster, J. L., Salomonson, V. V., Klein, A. G. and Chien, J. Y. L.: Development of a technique to assess
620 snow-cover mapping errors from space, *Geoscience and Remote Sensing, IEEE Transactions on*, 39(2), 432–438,
621 doi:10.1109/36.905251, 2001.
- 622 Harpold, A., Dettinger, M. and Rajagopal, S.: Defining snow drought and why it matters, *EOS - Earth & Space Science*
623 *News*, 98, doi:10.1029/2017EO068775, 2017.
- 624 Harshburger, B. J., Humes, K. S., Walden, V. P., Blandford, T. R., Moore, B. C. and Dezzani, R. J.: Spatial
625 interpolation of snow water equivalency using surface observations and remotely sensed images of snow-covered
626 area, *Hydrological Processes*, 24(10), 1285–1295, doi:10.1002/hyp.7590, 2010.
- 627 Hastie, T., Tibshirani, R. and Friedman, J.: *The Elements of Statistical Learning*, Springer New York, New York, NY.



- 628 2009.
- 629 Kimbauer, R. and Blöschl, G.: Wie ähnlich sind Ausaperungsmuster von Jahr zu Jahr (How similar are snow cover
630 patterns from year to year)? *Deutsche Wässerkundliche Mitteilungen*, 37(5/6), 113–121, 1994.
- 631 Koch, F., Prasch, M., Schmid, L., Schweizer, J. and Mauser, W.: Measuring Snow Liquid Water Content with Low-
632 Cost GPS Receivers, *Sensors* 2014, Vol. 14, Pages 20975–20999, 14(11), 20975–20999, doi:10.3390/s141120975,
633 2014.
- 634 König, M. and Sturm, M.: Mapping snow distribution in the Alaskan Arctic using aerial photography and topographic
635 relationships, *Water Resources Research*, 34(12), 3471–3483, doi:10.1029/98WR02514, 1998.
- 636 Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., Zeng, X., Yang,
637 Z.-L., Levis, S., Sakaguchi, K., Bonan, G. B. and Slater, A. G.: Parameterization Improvements and Functional and
638 Structural Advances in Version 4 of the Community Land Model, *J. Adv. Model. Earth Syst.*, 3(3),
639 doi:10.1029/2011MS000045, 2011.
- 640 Liston, G. E.: Interrelationships among Snow Distribution, Snowmelt, and Snow Cover Depletion: Implications for
641 Atmospheric, Hydrologic, and Ecologic Modeling, *Journal of Applied Meteorology*, 38(10), 1474–1487,
642 doi:10.1175/1520-0450(1999)038<1474:IASDSA>2.0.CO;2, 1999.
- 643 Litaor, M. I., Williams, M. and Seastedt, T. R.: Topographic controls on snow distribution, soil moisture, and species
644 diversity of herbaceous alpine vegetation, Niwot Ridge, Colorado, *J. Geophys. Res.*, 113(G2), n/a–n/a,
645 doi:10.1029/2007JG000419, 2008.
- 646 Livneh, B., Xia, Y., Mitchell, K. E., Ek, M. B. and Lettenmaier, D. P.: Noah LSM Snow Model Diagnostics and
647 Enhancements, *J. Hydrometeor.*, 11(3), 721–738, doi:10.1175/2009JHM1174.1, 2010.
- 648 López-Moreno, J. I. and Nogués-Bravo, D.: Interpolating local snow depth data: an evaluation of methods,
649 *Hydrological Processes*, 20(10), 2217–2232, doi:10.1002/hyp.6199, 2006.
- 650 López-Moreno, J. I., Revuelto, J., Fassnacht, S. R., Azorín-Molina, C., Vicente-Serrano, S. M., Morán-Tejeda, E. and
651 Sexstone, G. A.: Snowpack variability across various spatio-temporal resolutions, *Hydrological Processes*, 29(6),
652 1213–1224, doi:10.1002/hyp.10245, 2014.
- 653 Luce, C. H. and Tarboton, D. G.: The application of depletion curves for parameterization of subgrid variability of
654 snow, *Hydrological Processes*, 18(8), 1409–1422, doi:10.1002/hyp.1420, 2004.
- 655 Luce, C. H., Tarboton, D. G. and Cooley, K. R.: Sub-grid parameterization of snow distribution for an energy and
656 mass balance snow cover model, *Hydrological Processes*, 13(12-13), 1921–1933, doi:10.1002/(SICI)1099-
657 1085(199909)13:12/13<1921::AID-HYP867>3.0.CO;2-S, 1999.



- 658 Marchand, W. D. and Killingtveit, A.: Statistical probability distribution of snow depth at the model sub-grid cell
659 spatial scale, *Hydrological Processes*, 19(2), 355–369, doi:10.1002/hyp.5543, 2005.
- 660 Marshall, H.-P. and Koh, G.: FMCW radars for snow research, *Cold Regions Science and Technology*, 52(2), 118–
661 131, 2008.
- 662 Martinec, J. and Rango, A.: Areal distribution of snow water equivalent evaluated by snow cover monitoring, *Water*
663 *Resour. Res.*, 17(5), 1480–1488, 1981.
- 664 Meromy, L., Molotch, N. P., Link, T. E., Fassnacht, S. R. and Rice, R.: Subgrid variability of snow water equivalent
665 at operational snow stations in the western United States, *Hydrological Processes*, 27(17), 2383–2400,
666 doi:10.1002/hyp.9355, 2012.
- 667 Mizukami, N. and Perica, S.: Spatiotemporal Characteristics of Snowpack Density in the Mountainous Regions of the
668 Western United States, *Journal of Hydrometeorology*, 9(6), 1416–1426, doi:10.1175/2008JHM981.1, 2008.
- 669 Molotch, N. P. and Bales, R. C.: Scaling snow observations from the point to the grid element: Implications for
670 observation network design, *Water Resources Research*, 41(11), doi:10.1029/2005WR004229, 2005.
- 671 Molotch, N. P. and Bales, R. C.: SNOTEL representativeness in the Rio Grande headwaters on the basis of
672 physiographics and remotely sensed snow cover persistence, *Hydrological Processes*, 20(4), 723–739,
673 doi:10.1002/hyp.6128, 2006.
- 674 Molotch, N. P. and Margulis, S. A.: Estimating the distribution of snow water equivalent using remotely sensed snow
675 cover data and a spatially distributed snowmelt model: A multi-resolution, multi-sensor comparison, *Advances in*
676 *Water Resources*, 31(11), 1503–1514, doi:10.1016/j.advwatres.2008.07.017, 2008.
- 677 Molotch, N. P., Colee, M. T., Bales, R. C. and Dozier, J.: Estimating the spatial distribution of snow water equivalent
678 in an alpine basin using binary regression tree models: the impact of digital elevation data and independent variable
679 selection, *Hydrological Processes*, 19(7), 1459–1479, doi:10.1002/hyp.5586, 2005.
- 680 Niu, G.-Y. and Yang, Z.-L.: An observation-based formulation of snow cover fraction and its evaluation over large
681 North American river basins, *J. Geophys. Res.*, 112(D21), doi:10.1029/2007JD008674, 2007.
- 682 Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D.,
683 Rosero, E., Tewari, M. and Xia, Y.: The community Noah land surface model with multiparameterization options
684 (Noah-MP): 1. Model description and evaluation with local-scale measurements, *J. Geophys. Res.*, 116(D12), n/a–
685 n/a, doi:10.1029/2010JD015139, 2011.
- 686 Nolan, M., Larsen, C. F. and Sturm, M.: Mapping snow-depth from manned-aircraft on landscape scales at centimeter
687 resolution using Structure-from-Motion photogrammetry, *The Cryosphere Discuss.*, 9(1), 333–381, doi:10.5194/tcd-



- 688 9-333-2015, 2015.
- 689 NPS: Plants - Yosemite National Park, [online] Available from: <https://www.nps.gov/yose/learn/nature/plants.htm>
690 (Accessed 5 October 2016), 2016.
- 691 Painter, T. H., Berisford, D. F., Boardman, J. W., Bormann, K. J., Deems, J. S., Gehrke, F., Hedrick, A., Joyce, M.,
692 Laidlaw, R., Marks, D., Mattmann, C., McGurk, B., Ramirez, P., Richardson, M., Skiles, S. M., Seidel, F. C. and
693 Winstral, A.: The Airborne Snow Observatory: Fusion of scanning lidar, imaging spectrometer, and physically-based
694 modeling for mapping snow water equivalent and snow albedo, *Remote Sensing of Environment*, 184, 139–152,
695 doi:10.1016/j.rse.2016.06.018, 2016.
- 696 Painter, T. H., Rittger, K. E., McKenzie, C., Slaughter, P., Davis, R. E. and Dozier, J.: Retrieval of subpixel snow
697 covered area, grain size, and albedo from MODIS, *Remote Sensing of Environment*, 113(4), 868–879,
698 doi:10.1016/j.rse.2009.01.001, 2009.
- 699 Parajka, J. and Bloeschl, G.: Spatio-temporal combination of MODIS images - potential for snow cover mapping,
700 *Water Resources Research*, 44(3), doi:10.1029/2007WR006204, 2008.
- 701 Parsons, W. J. and Castle, G. H.: Aerial reconnaissance of mountain snow fields for maintaining up-to-date forecasts
702 of snow melt runoff during the melt season, *Western Snow Conference*, Reno, Nevada. 1959.
- 703 Potts, H. L.: Snow surveys and runoff forecasting from photographs, *Proceedings of the 5th Annual Western Interstate*
704 *Snow Survey Conference*, 1937.
- 705 Potts, H. L.: A photographic snow survey method of forecasting runoff, *Proceedings of the 12th Annual Western*
706 *Interstate Snow Survey Conference*, 1944.
- 707 R Core Team: *R: A Language and Environment for Statistical Computing*, Vienna, Austria. 2015.
- 708 Raleigh, M. S., Rittger, K. E., Moore, C. E., Henn, B., Lutz, J. A. and Lundquist, J. D.: Ground-based testing of
709 MODIS fractional snow cover in subalpine meadows and forests of the Sierra Nevada, *Mon. Wea. Rev.*, 128, 44–57,
710 doi:10.1016/j.rse.2012.09.016, 2013.
- 711 Revuelto, J., López-Moreno, J. I., Azorín-Molina, C. and Vicente-Serrano, S. M.: Topographic control of snowpack
712 distribution in a small catchment in the central Spanish Pyrenees: intra- and inter-annual persistence, *The Cryosphere*,
713 8(5), 1989–2006, doi:10.5194/tc-8-1989-2014, 2014.
- 714 Rice, R., Bales, R. C., Painter, T. H. and Dozier, J.: Snow water equivalent along elevation gradients in the Merced
715 and Tuolumne River basins of the Sierra Nevada, *Water Resour. Res.*, 47, doi:10.1029/2010WR009278, 2011.
- 716 Rittger, K. E., Painter, T. H. and Dozier, J.: Assessment of methods for mapping snow cover from MODIS, *Advances*



- 717 in Water Resources, 51, 367–380, doi:10.1016/j.advwatres.2012.03.002, 2013.
- 718 Rosenthal, W. and Dozier, J.: Automated Mapping of Montane Snow Cover at Subpixel Resolution from the Landsat
719 Thematic Mapper, Water Resources Research, 32(1), 115–130, doi:10.1029/95WR02718, 1996.
- 720 Rutter, N., Cline, D. and Li, L.: Evaluation of the NOHRSC snow model (NSM) in a one-dimensional mode, Journal
721 of Hydrometeorology, 9(4), 695–711, doi:10.1175/2008JHM861.1, 2008.
- 722 Salomonson, V. V. and Appel, I.: Estimating fractional snow cover from MODIS using the normalized difference
723 snow index, Remote Sensing of Environment, 89(3), 351–360, doi:10.1016/j.rse.2003.10.016, 2004.
- 724 Sappington, J. M., Longshore, K. M. and Thompson, D. B.: Quantifying Landscape Ruggedness for Animal Habitat
725 Analysis: A Case Study Using Bighorn Sheep in the Mojave Desert, Journal of Wildlife Management, 71(5), 1419–
726 1426, doi:10.2193/2005-723, 2007.
- 727 Schirmer, M., Wirz, V., Clifton, A. and Lehning, M.: Persistence in intra-annual snow depth distribution:
728 1. Measurements and topographic control, Water Resources Research, 47(9), W09516–n/a,
729 doi:10.1029/2010WR009426, 2011.
- 730 Schneider, D. and Molotch, N. P.: Real-Time Estimation of Snow Water Equivalent in the Upper Colorado River
731 Basin using MODIS-based SWE Reconstructions and SNOTEL data, Water Resources Research,
732 doi:10.1002/2016WR019067, 2016.
- 733 Slater, A. G., Barrett, A. P., Clark, M. P., Lundquist, J. D. and Raleigh, M. S.: Uncertainty in seasonal snow
734 reconstruction: Relative impacts of model forcing and image availability, Advances in Water Resources, 55(0), 165–
735 177, doi:10.1016/j.advwatres.2012.07.006, 2013.
- 736 Stewart, I. T., Cayan, D. R. and Dettinger, M. D.: Changes toward earlier streamflow timing across western North
737 America, Journal of Climate, 18(8), 1136–1155, 2004.
- 738 Sturm, M., Taras, B., Liston, G. E., Derksen, C., Jonas, T. and Lea, J.: Estimating Snow Water Equivalent Using Snow
739 Depth Data and Climate Classes, Journal of Hydrometeorology, 11(6), 1380–1394, doi:10.1175/2010JHM1202.1,
740 2010.
- 741 Trujillo, E., Ramirez, J. A. and Elder, K. J.: Scaling properties and spatial organization of snow depth fields in sub-
742 alpine forest and alpine tundra, Hydrological Processes, 23(11), 1575–1590, doi:10.1002/hyp.7270, 2009.
- 743 Veitinger, J., Sovilla, B. and Purves, R. S.: Influence of snow depth distribution on surface roughness in alpine terrain:
744 a multi-scale approach, The Cryosphere, 8(2), 547–569, doi:10.5194/tc-8-547-2014, 2014.
- 745 Xia, Q., Gao, X., Chu, W. and Sorooshian, S.: Estimation of daily cloud-free, snow-covered areas from MODIS based



746 on variational interpolation, *Water Resour. Res.*, 48(9), 1480–1488, doi:10.1029/WR017i005p01480, 2012.

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

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

751

752