Converting Snow Depth to Snow Water Equivalent Using Climatological Variables

David F. Hill[1], Elizabeth A. Burakowski[2], Ryan L. Crumley[3], Julia Keon[4], J. Michelle Hu[5], Anthony A. Arendt[6], Katreen Wikstrom Jones[7], Gabriel J. Wolken[8]

[1]Civil and Construction Engineering, Oregon State University, OR, USA
[2]Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, NH, USA
[3]Water Resources Graduate Program, Oregon State University, OR, USA
[4]Civil and Construction Engineering, Oregon State University, OR, USA
[5]Civil and Environmental Engineering, University of Washington
[6]Applied Physics Laboratory, University of Washington
[7]Alaska Division of Geological & Geophysical Surveys, Fairbanks, AK, USA
[8]Alaska Division of Geological & Geophysical Surveys, Fairbanks, AK, USA; International Arctic Research Center, University of Alaska Fairbanks, Fairbanks, AK, USA

Correspondence to: David F. Hill (david.hill@oregonstate.edu)

Abstract. We present a simple method that allows snow depth measurements to be converted to snow water equivalent (SWE) estimates. These estimates are useful to individuals interested in water resources, ecological function, and avalanche forecasting. They can also be assimilated into models to help improve predictions of total water volumes over large regions. The conversion of depth to SWE is particularly valuable since snow depth measurements are far more numerous than costlier and more complex SWE measurements. Our model regresses SWE against snow depth (h), day of water year (DOY) and climatological (30-year normal) values for winter (December, January, February) precipitation ($PPTWT$) and the difference (TD) between mean temperature of the warmest month and mean temperature of the coldest month, producing a power-law relationship. Relying on climatological normals rather than weather data for a given year allows our model to be applied at measurement sites lacking a weather station. Separate equations are obtained for the accumulation and the ablation phases of the snowpack. The model is validated against a large database of snow pillow measurements and yields a bias in SWE of less than 2 mm and a root-mean-squared-error (RMSE) in SWE of less than 60 mm. The model is additionally validated against two completely independent sets of data; one from western North America, and one from the northeast United States. Finally, the results are compared with three other models for bulk density that have varying degrees of complexity and that were built in multiple geographic regions. The results show that the model described in this paper has the best performance for the validation data sets.

1 Introduction

In many parts of the world, snow plays a leading-order role in the hydrological cycle (USACE, 1956; Mote et al., 2018). Accurate information about the spatial and temporal distribution of snow water equivalent (SWE) is useful to many stakeholders (water resource planners, avalanche forecasters, aquatic ecologists, etc.), but can be time consuming and expensive to obtain.

Snow pillows (Beaumont, 1965) are a well-established tool for measuring SWE at fixed locations. Figure 1 provides a conceptual sketch of the variation of SWE with time over a typical water year. A comparatively long accumulation phase is followed by a short ablation phase. While simple in operation, snow pillows are relatively large in size and they need to be installed prior to the onset of the season's snowfall. This limits their ability to be rapidly or opportunistically deployed. Additionally, snow pillow installations tend to require vehicular access, limiting their locations to relatively simple topography. Finally, snow pillow sites are not representative of the lowest or highest elevation bands within mountainous regions (Molotch and Bales, 2005). In the western United States (USA), the Natural Resources Conservation Service (NRCS) operates a large network of Snow Telemetry (SNOTEL) sites, featuring snow pillows. The NRCS also operates the smaller Soil Climate Analysis Network (SCAN) which provides the only, and very limited, snow pillow SWE measurements in the eastern USA.

SWE can also be measured manually, using a snow coring device that measures the weight of a known volume of snow to determine snow density (Church, 1933). These measurements are often one-off measurements, or in the case of 'snow courses' they are repeated weekly or monthly as a transect of measurements at a given location. The simplicity and portability of coring devices expand the range over which measurements can be collected, but it can be challenging to apply these methods to deep snowpacks due to the limited length of standard coring devices. Note that there are numerous different styles of coring devices, including the Adirondack sampler and the Mt. Rose / Federal sampler (Church and Marr, 1937). The NRCS operates a large network of snow course sites (USDA, 2011) in the western United States.

There are a number of issues that affect the accuracy of both snow pillow and snow coring measurements. With coring measurements, if the coring device is not carefully extracted, a portion of the core may fall out of the device. Or, snow may become compressed in the coring device during insertion. These effects have led to varying conclusions, with some studies (e.g., Sturm et al., 2010) showing a low SWE bias and other studies (e.g., Goodison, 1978) showing a high SWE bias. As noted by Johnson et al. (2015) a good rule of thumb is that coring devices are accurate to around ± 10%. Also, studies comparing different styles of snow samplers report statistically different results, suggesting that SWE measurements are sensitive to the design of the specific coring device, such as the presence of holes or slots, the device material, etc. (Beaumont and Work, 1963; Dixon and Boon, 2012). With snow pillows, some studies (e.g., Goodison et al., 1981) note that ice bridging can lead to low biases in measured SWE, with the snow surrounding the pillow partly supporting the snow over the pillow. Other studies (Johnson and Marks, 2004; Dressler et al., 2006; Johnson et al., 2015) note a more complex situation with SWE under-reported at times,

72   but over-reported at other times. Note that when snow pillow data are evaluated, they are most commonly compared

73   to coring measurements at the same location.

74

75   All methods of measuring SWE are challenged by the fact that SWE is a depth-integrated property of a snowpack.

76   This is why the snowpack must be weighed, in the case of a snow pillow, or a core must be extracted from the

77   surface to the ground. This measurement complexity makes it difficult to obtain SWE information with the spatial

78   and temporal resolution desired for watershed-scale studies. Other snowpack properties, such as the depth h, are

79   much easier to measure. For example, using a graduated device such as a meterstick or an avalanche probe to

80   measure the depth takes only seconds. Automating depth measurements at a fixed location can easily be done using

81   low-cost ultrasonic devices (Goodison et al., 1984; Ryan et al., 2008). High-spatial-resolution measurements of

82   snowpack depth are commonly made with Light Detection and Ranging (LIDAR). One example of this is the

83   Airborne Snow Observatory program (ASO; Painter et al., 2016). The comparatively high expense of airborne

84   LIDAR surveys typical limits measurements geographically (to a few basins) and temporally (weekly to monthly

85   interval).

86

87   Given the relative ease in obtaining depth measurements, it is common to use h as a proxy for SWE. Figure 1 shows

88   a conceptual sketch of the variation of SWE with h over a typical water year. Noting the arrows on the curve, we see

89   that SWE is multi-valued for each h. This is due to the fact that the snowpack increases in density throughout the

90   water year, producing a hysteresis loop in the curve. A large body of literature exists on the topic of how to convert

91   h to SWE. It is beyond the scope of this paper to provide a full review of these 'bulk density equations,' where the

92   density is given by $\rho_b = \text{SWE}/h$. Instead, we refer readers to the useful comparative review by Avanzi et al. (2015).

93   Here, we prefer to discuss a limited number of previous studies that illustrate the spectrum of methodologies and

94   complexities that can be used to determine ρ_b or SWE.

95

96   Many studies express ρ_b as an increasing function (often linear) of h. In some cases (e.g., Lundberg et al., 2006) a

97   second equation is added where ρ_b attains a constant value when a threshold h is exceeded. A single linear equation

98   captures the process of densification of the snowpack during the accumulation phase, but performs poorly during the

99   ablation phase, where depths are decreasing but densities continue to increase or approach a constant value.

100   Other approaches choose to parameterize ρ_b in terms of time, rather than h. Pistocchi (2016) provides a single

101   equation while Mizukami and Perica (2008) provide two sets of equations, one set each for early and late season.

102   Each set contains four equations, each of which is applicable to a particular 'cluster' of stations. This clustering was

103   driven by observed densification characteristics and the resulting clusters are relatively spatially discontinuous.

104   Jonas et al. (2009) take the idea of region- (or cluster-) specific equations and extend it further to provide

105   coefficients that depend on time and elevation as well. They use a simple linear equation for ρ_b in terms of h and the

106   slope and intercept of the equation are given as monthly values, with three elevation bins for each month (36 pairs of

107   coefficients). There is an additional contribution to the intercept (or 'offset') which is region-specific (one of 7

108   regions).

109

110 These classifications, whether based on region, elevation, or season, are valuable since they acknowledge that all

111 snow is not equal. McKay and Findlay (1971) discuss the controls that climate and vegetation exert on snow density,

112 and Sturm et al. (2010) address this directly by developing a snow density equation where the coefficients depend

113 upon the 'snow class' (5 classes). Sturm et al. (1995) explain the decision tree, based on temperature, precipitation,

114 and wind speed, that leads to the classification. The temperature metric is the 'cooling degree month' calculated

115 during winter months only. Similarly, only precipitation falling during winter months was used in the classification.

116 Finally, given the challenges in obtaining high quality, high-spatial-resolution wind information, vegetation

117 classification was used as a proxy. Using climatological values (rather than values for a given year), Sturm et al.

118 (1995) were able to develop a global map of snow classification.

119

120 There are many other formulations for snow density that increase in complexity and data requirements. Meloysund

121 et al. (2007) express ρ_b in terms of sub-daily measurements of relative humidity, wind characteristics, air pressure,

122 and rainfall, as well as h and estimates of solar exposure ('sun hours'). McCreight and Small (2014) use daily snow

123 depth measurements to develop their regression equation. They demonstrate improved performance over both Sturm

124 et al. (2010) and Jonas et al. (2009). However, a key difference between the McCreight and Small (2014) model and

125 the others listed above is that the former cannot be applied to a single snow depth measurement. Instead, it requires a

126 continuous time series of depth measurements at a fixed location. Further increases in complexity are found in

127 energy-balance snowpack models (SnowModel, Liston and Elder, 2006; VIC, Liang et al., 1994, DHSVM,

128 Wigmosta et al., 1994, others), many of which use multi-layer models to capture the vertical structure of the

129 snowpack. While the particular details vary, these models generally require high temporal-resolution time series of

130 many meteorological variables as input.

131

132 Despite the development of multi-layer energy-balance snow models, there is still a demonstrated need for bulk

133 density formulations and for vertically integrated data products like SWE. Pagano et al. (2009) review the

134 advantages and disadvantages of energy-balance models and statistical models and describe how the NRCS uses

135 SWE (from SNOTEL stations) and accumulated precipitation in their statistical models to make daily water supply

136 forecasts. If SWE information is desired at a location that does not have a SNOTEL station, and is not part of a

137 modeling effort, then bulk density equations and depth measurements are an excellent choice.

138

139 The present paper seeks to generalize the ideas of Mizukami and Perica (2008), Jonas et al. (2009), and Sturm et al.,

140 (2010). Specifically, our goal is to regress physical and environmental variables directly into the equations. In this

141 way, environmental variability is handled in a continuous fashion rather than in a discrete way (model coefficients

142 based on classes). The main motivation for this comes from evidence (e.g., Fig. 3 of Alford, 1967) that density can

143 vary significantly over short distances on a given day. Bulk density equations that rely solely on time completely

144 miss this variability and equations that have coarse (model coefficients varying over either vertical bins or horizontal

145 grids) spatial resolution may not fully capture it either.

146

Our approach is most similar to Mizukami and Perica (2008), Jonas et al. (2009), and Sturm et al., (2010) in that a minimum of information is needed for the calculations; we intentionally avoid approaches like Meloysund et al. (2007) and McCreight and Small (2014). This is because our interests are in converting h measurements to SWE estimates in areas lacking weather instrumentation. The following sections introduce the numerous data sets that were used in this study, outline the regression model adopted, and assess the performance of the model.

2 Methods

2.1 Data

2.1.1 Snow Depth and Snow Water Equivalent

In this section, we list sources of 1970-present snow data utilized for this study (Table 1). With regards to snow coring devices, we refer to them using the terminology preferred in the references describing the datasets.

2.1.1.1 USA NRCS Snow Telemetry and Soil Climate Analysis Networks

SNOTEL (Serreze et al., 1999; Dressler et al., 2006) and SCAN (Schaefer et al. 2007) stations in the contiguous United States (CONUS) and Alaska typically record sub-daily observations of h, SWE, and a variety of weather variables (Figure 2a). The periods of record are variable, but the vast majority of stations have a period of record in excess of 30 years. For this study, data from all SNOTEL sites in CONUS and Alaska and northeast USA SCAN sites (Figure 2b) were obtained with the exception of sites whose period of record data were unavailable online. Only stations with both SWE and h data were retained.

2.1.1.2 Canada (British Columbia) Snow Survey Data

Goodison et al. (1987) note that Canada has no national digital archive of snow observations from the many independent agencies that collect snow data and that snow data are instead managed provincially. The quantity and availability of the data vary considerably among the provinces. The Water Management Branch of the British Columbia (BC) Ministry of the Environment manages a comparatively dense network of Automated Snow Weather Stations (ASWS) that measure SWE, h, accumulated precipitation, and other weather variables (Figure 2a). For this study, data from all British Columbia ASWS sites were initially obtained. As with the NRCS stations, only ASWS stations with both SWE and h data were retained.

2.1.1.3 USA NRCS Snow Course / Aerial Marker Data

The snow survey program (USDA, 2008) dates to the 1930s and includes a large number of snow course and aerial marker sites (Figure 2c) in western North America. While the measurement frequency is variable, it is most commonly monthly. To generate a dataset for this study, data were extracted using the National Water and Climate

181 Center Report Generator 2.0. This allows filtering by time period, elevation band, and other elements. All sites with
182 data between 1980-2018 were included (Figure 2c).
183

184 **2.1.1.4 Northeast USA Data**
185 In addition to the data from the SCAN sites, snow data for this project from the northeast US come from two
186 networks and three research sites (Figure 2b). The Maine Cooperative Snow Survey (MCSS, 2018) network
187 includes h and SWE data collected by the Maine Geological Survey, the United States Geological Survey, and
188 numerous private contributors and contractors. MCSS snow data are collected using the Standard Federal or
189 Adirondack snow sampling tubes typically on a weekly to bi-weekly schedule throughout the winter and spring,
190 1951-present. The New York Snow Survey network data were obtained from the National Oceanic and Atmospheric
191 Administration's Northeast Regional Climate Center at Cornell University (NYSS, 2018). Similar to the MCSS,
192 NYSS data are collected using Standard Federal or Adirondack snow sampling tubes on weekly to bi-weekly
193 schedules, 1938-present.
194

195 The Sleepers River, Vermont Research Watershed in Danville, Vermont (Shanley and Chalmers, 1999) is a USGS
196 site that includes 15 stations with long-term weekly records of h and SWE collected using Adirondack snow tubes.
197 Most of the periods of record are 1981-present, with a few stations going back to the 1960s. The sites include
198 topographically flat openings in conifer stands, old fields with shrub and grass, a hayfield, a pasture, and openings in
199 mixed softwood-hardwood forests. The Hubbard Brook Experiment Forest (Campbell et al., 2010) has collected
200 weekly snow observations at the Station 2 rain gauge site, 1959-present. Measurement protocol collects ten samples
201 2 m apart along a 20 m transect in a hardwood forest opening about ¼ hectare in size. At each sample location along
202 the transect, h and SWE are measured using a Mt. Rose snow tube and the ten samples are averaged for each
203 transect. Finally, the Thompson Farm Research site includes a mixed hardwood forest site and an open pasture site
204 (Burakowski et al. 2013; Burakowski et al. 2015). Daily (from 2011-2018), at each site, a snow core is extracted
205 with an aluminum tube and weighed (tube + snow) using a digital hanging scale. The net weight of the snow is
206 combined with the depth and the tube diameter to determine ρ_b, similar to a Federal or Adirondack sampler.
207

208 **2.1.1.5 Chugach Mountains (Alaska) Data**
209 In the spring of 2018, we conducted three weeks of fieldwork in the Chugach mountains in coastal Alaska, near the
210 city of Valdez (Figure 2d-e). We measured h using an avalanche probe at 71 sites along elevational transects during
211 March, April, and May. The elevational transects ranged between 250 and 1100 m (net change along transect) and
212 were accessible by ski and snowshoe travel. At each site, we measured h in 8 locations within the surrounding 10
213 m^2, resulting in a total of 550+ snow depth measurements. These 71 sites were scattered across 8 regions in order to
214 capture spatial gradients that exist in the Chugach mountains as the wetter, more-dense maritime snow near the coast
215 gradually changes to drier, less dense snow on the interior side.
216

2.1.1.5 Data Pre-Processing

Figure 3 demonstrates that it is not uncommon for automated snow pillow measurements to become noisy or non-physical, at times reporting large depths when there is no SWE reported. This is different from instances when physically plausible, but very low densities might be reported; say in response to early season dry, light snowfalls. It was therefore desirable to apply some objective, uniform procedure to each station's dataset in order to remove clear outlier points, while minimizing the removal of valid data points. We recognize that there is no accepted standardized method for cleaning bivariate SWE-*h* data sets. While Serreze et al. (1999) offer a procedure for SNOTEL data in their appendix, it is relevant only for precipitation and SWE values, not *h*. Given the strong correlation between *h* and SWE, we instead choose to use common outlier detection techniques for bivariate data.

The Mahalanobis distance (MD; Maesschalck et al., 2000) quantifies how far a point lies from the mean of a bivariate distribution. The distances are in terms of the number of standard deviations along the respective principal component axes of the distribution. For highly correlated bivariate data, the MD can be qualitatively thought of as a measure of how far a given point deviates from an ellipse enclosing the bulk of the data. One problem is that the MD is based on the statistical properties of the bivariate data (mean, covariance) and these properties can be adversely affected by outlier values. Therefore, it has been suggested (e.g., Leys et al., 2018) that a 'robust' MD (RMD) be calculated. The RMD is essentially the MD calculated based on statistical properties of the distribution unaffected by the outliers. This can be done using the Minimum Covariance Determinant (MCD) method as first introduced by Rousseeuw (1984).

Once RMDs have been calculated for a bivariate data set, there is the question of how large an RMD must be in order for the data point to be considered an outlier. For bivariate normal data, the distribution of the square of the RMD is χ^2 (Gnanadesikan and Kettenring, 1972), with p (the dimension of the dataset) degrees of freedom. So, a rule for identifying outliers could be implemented by selecting as a threshold some arbitrary quantile (say 0.99) of χ_p^2. For the current study, a threshold quantile of 0.999 was determined to be an appropriate compromise in terms of removing obviously outlier points, yet retaining physically plausible results.

A scatter plot of SWE vs. *h* for the SNOTEL dataset from CONUS and AK reveals many non-physical points, mostly when a very large *h* is reported for a very low SWE (Figure 4a). Approximately 0.7% of the original data points were removed in the pre-processing described above, creating a more physically plausible scatter plot (Figure 4b). Note that the outlier detection process was applied to each station individually. The distribution of 'day of year' (*DOY*) values of removed data points was broad, with a mean of 160 and a standard deviation of 65. Note that the *DOY* origin is 1 October. The same procedure was applied to the BC snow pillow, NRCS snow course, and northeast USA data sets as well (not shown). Table 1 summarizes useful information about the numerous data sets described above and indicates the final number of data points retained for each. We acknowledge that our process inevitably removes some valid data points, but, as a small percentage of an already small 0.7% removal rate, we judged this to be acceptable.

254

255 Table 1: Summary of information about the datasets used in this study. Datasets in bold font were used to construct
256 the regression model. The numbers of stations and data points reflect the post-processed data.

Dataset Name	Dataset Type	Number of retained stations	Number and percentage of retained data points	Precision (*h* / SWE)
NRCS SNOTEL	**Snow pillow (SWE), ultrasonic (*h*)**	**791**	**1,900,000 (99.3%)**	**(0.5 in / 0.1 in)**
NRCS SCAN	Snow pillow (SWE), ultrasonic (*h*)	5	7094 (97.8%)	(0.5 in / 0.1 in)
British Columbia Snow Survey	**Snow pillow (SWE), ultrasonic (*h*)**	**31**	**61,000 (97.5%)**	**(1 cm / 1 mm)**
NRCS Snow Survey	Federal sampler / Aerial marker	1085	116,000 (99.6%)	(0.5 in / 0.1 in) for manual sampler (2 in / n/a) for aerial marker
Maine Geological Survey	Adirondack or Federal sampler (SWE and *h*)	431	28,000 (99.3%)	(0.5 in / 0.5 in)
Hubbard Brook (Station 2), NH	Mount Rose sampler (SWE and *h*)	1	704 (99.4%)	(0.1 in / 0.1 in)
Thompson Farm, NH	Snow core (SWE and *h*)	2	988 (99.4%)	0.5 in / 0.5 in)
Sleepers River, VT	Adirondack sampler	14	7214 (99.4%)	(0.5 in / 0.5 in)
New York Snow Survey	Adirondack or Federal sampler (SWE and *h*)	523	44,614 (98.2%)	(0.5 in / 0.5 in)
Chugach Mountains, AK	Avalanche probe (*h*)	71	71 (100%)	(1 cm)

257

2.1.2 Climatological Variables

259 30-year climate normals at 1 km resolution for North America were obtained from the ClimateNA project (Wang et
260 al., 2016). This project provides grids for minimum, maximum, and mean temperature, and total precipitation for a
261 given month. These grids are based on the PRISM normals (Daly et al., 1994) and are available for the periods
262 1961-1990 and 1981-2010. For this study, the more recent climatology was used. The ClimateNA project also
263 provides a wide array of derived bioclimatic variables, such as precipitation as snow (*PAS*), frost-free-period (*FFP*),
264 mean annual relative humidity (*RH*) and others. Wang et al. (2012) summarize these additional variables and how
265 they are derived. Figure 5 shows gridded maps of winter (sum of December, January, February) precipitation
266 (*PPTWT*) and the temperature difference (*TD*) between the mean temperature of the warmest month and the mean
267 temperature of the coldest month. The latter variable (*TD*) is a measure of continentality.

268

2.2 Regression Model

270 In order to demonstrate the varying degrees of influence of explanatory variables, several regression models were
271 constructed. In each case, the model was built by randomly selecting 50% of the paired SWE-*h* measurements from
272 the aggregated CONUS, AK, and BC snow pillow datasets. The model was then validated by applying it to the

273 remaining 50% of the dataset and comparing the modeled SWE to the observed SWE for those points. We

274 constructed a second version of the regression models by randomly selecting 50% of the snow pillow stations and

275 using all of the data from those stations. The model was then validated by applying it to the data from the remaining

276 50% of the stations. These two methods provided identical results, likely due to the very large sample size (N) of our

277 dataset. In all cases, the p values from the linear regression were 0, again due to the large sample size. Additional

278 validation was done with the northeast USA datasets (SCAN snow pillow and various snow coring datasets) and the

279 NRCS snow course dataset, which were completely left out of the model building process.

280

281 **2.2.1 One-Equation Model**

282 The simplest equation, and one that is supported by the strong correlation seen in the portions of Figure 3 when

283 SWE is present, is one that expresses SWE as a function of h. A linear model is attractive in terms of simplicity, but

284 this limits the snowpack to a constant density. An alternative is to express SWE as a power law, i.e.,

285

286 (1) $SWE = Ah^{a_1}$.

287

288 This equation can be log-transformed into

289

290 (2) $log_{10}(SWE) = log_{10}(A) + a_1 log_{10}(h)$

291

292 which immediately allows for simple linear regression methods to be applied. With both h and SWE expressed in

293 units of mm, the obtained coefficients are $(A, a_1) = (0.146, 1.102)$. Information on the performance of the model

294 will be deferred until the results section.

295

296 **2.2.2 Two-Equation Model**

297 Recall from Figures 1 and 4 that there is a hysteresis loop in the SWE-h relationship. During the accumulation

298 phase, snow densities are relatively low. During the ablation phase, the densities are relatively high. So, the same

299 snowpack depth is associated with two different SWEs, depending upon the time of year. The regression equation

300 given above does not resolve this difference. This can be addressed by developing two separate regression

301 equations, one for the accumulation (*acc*) and one for the ablation (*abl*) phase. This approach takes the form

302

303 (3) $SWE_{acc} = Ah^{a_1}; \quad DOY < DOY^*$

304

305 (4) $SWE_{abl} = Bh^{b_1}; \quad DOY \geq DOY^*$

306

307 where DOY is the number of days from the start of the water-year, and DOY^* is the critical or dividing day-of-water-

308 year separating the two phases. Put another way, DOY^* is the day of peak SWE. Interannual variability results in a

309 range of DOY^* for a given site. Additionally, some sites, particularly the SCAN sites in the northeast USA,

310   demonstrate multi-peak SWE profiles in some years. To reduce model complexity, however, we investigated the use

311   of a simple climatological (long term average) value of DOY^* at each site. For each snow pillow station, the average

312   DOY^* was computed over the period of record of that station. Analysis of all of the stations revealed that this

313   average DOY^* was relatively well correlated with the climatological mean April maximum temperature (the average

314   of the daily maximums recorded in April; $R^2 = 0.7$). However, subsequent regression analysis demonstrated that the

315   SWE estimates were relatively insensitive to DOY^* and the best results were actually obtained when DOY^* was

316   uniformly set to 180 for all stations. Again, with both SWE and h in units of mm, the regression coefficients turn out

317   to be $(A, a_1) = (0.150, 1.082)$ and $(B, b_1) = (0.239, 1.069)$.

318

319   As these two equations are discontinuous at DOY^*, they are blended smoothly together to produce the final two-

320   equation model

321

322   (5) $SWE = SWE_{acc} \frac{1}{2}(1 - tanh[0.01\{DOY - DOY^*\}]) +$

323    $SWE_{abl} \frac{1}{2}(1 + tanh[0.01\{DOY - DOY^*\}])$

324

325   The coefficient 0.01 in the tanh function controls the width of the blending window and was selected to minimize

326   the root mean square error of the model estimates.

327

328   **2.2.3 Two-Equation Model with Climate Parameters**

329   A final model was constructed by incorporating climatological variables. Again, the emphasis in this study is on

330   methods that can be implemented at locations lacking the time series of weather variables that might be available at

331   a weather or SNOTEL station. Climatological normals are unable to account for interannual variability, but they do

332   preserve the high spatial gradients in climate that can lead to spatial gradients in snowpack characteristics. Stepwise

333   linear regression was used to determine which variables to include in the regression. The initial list of potential

334   variables included was

335

336   (6) $SWE = f(h, z, PPTWT, PAS, TWT, TD, DOY, RH)$

337

338   where z is the elevation (m), $PPTWT$ is the winter (sum of December, January, February) precipitation (mm), PAS is

339   mean annual precipitation as snow (mm), TWT is the winter (December, January, February) mean temperature ($°C$),

340   TD is the difference between the mean temperature of the warmest month and the mean temperature of the coldest

341   month ($°C$), DOY is the day of water year, and RH is the relative humidity (%). In the stepwise regression,

342   explanatory variables were accepted only if they improved the adjusted R^2 value by 0.001. The result of the

343   regression yielded

344

345   (7) $SWE_{acc} = Ah^{a_1} PPTWT^{a_2} TD^{a_3} DOY^{a_4}; \quad DOY < DOY^*$

346

347   (8) $SWE_{abl} = Bh^{b_1}PPTWT^{b_2}TD^{b_3}DOY^{b_4}; \quad DOY \geq DOY^*$

348

349 or, in log-transformed format,

350

351   (9) $log_{10}(SWE_{acc}) = log_{10}(A) + a_1 log_{10}(h) + a_2 log_{10}(PPTWT) +$

352                               $a_3 log_{10}(TD) + a_4 log_{10}(DOY); \quad DOY < DOY^*$

353

354   (10) $log_{10}(SWE_{abl}) = log_{10}(B) + b_1 log_{10}(h) + b_2 log_{10}(PPTWT) +$

355                               $b_3 log_{10}(TD) + b_4 log_{10}(DOY); \quad DOY \geq DOY^*$

356

357 indicating that only snow depth, winter precipitation, temperature difference, and day of water year were relevant.

358 Manual tests of model construction with other variables included confirmed that Eqns. (7-8) yielded the best results.

359 These two SWE estimates for the individual (*acc* and *abl*) phases of the snowpack were then blended with Eqn. (5)

360 to produce a single equation for SWE spanning the entire water year. The obtained regression coefficients were

361 $(A, a_1, a_2, a_3, a_4) = (0.0533, \ 0.9480, \ 0.1701, -0.1314, 0.2922)$ and $(B, b_1, b_2, b_3, b_4) = (0.0481, 1.0395,$

362 $0.1699, -0.0461, 0.1804)$. The physical interpretation of these coefficients is straightforward. For example, both a_2

363 and b_2 are greater than zero. So, for two locations with equal *h*, *DOY*, and *TD*, the location with greater *PPTWT* will

364 have a greater SWE and therefore density. These locations are typically maritime climates with wetter, denser snow.

365 In contrast, both a_3 and b_3 are less than zero. Therefore, for two locations with equal *h*, *DOY*, and *PPTWT*, the

366 location with greater TD (a more continental climate) will have a lower density, which is again an expected result.

367 These trends are similar in concept to Sturm et al. (2010), whose discrete snow classes (based on climate classes)

368 indicate which snow will densify more rapidly.

369 **3 Results**

370 A comparison of the three regression models (one-equation model, Eq. (2); two-equation model, Eqs. (3-5); multi-

371 variable two-equation model, Eqs. (5, 7-8)) is provided in Figure 6. The left column shows scatter plots of modeled

372 SWE to observed SWE for the validation data set with the 1:1 line shown in black. The right column shows

373 distributions of the model residuals. The vertical lines in the right column show the mean error, or model bias.

374 Visually, it is clear that the one-equation model performs relatively poorly with a large negative bias. This large

375 negative bias is partially overcome by the two-equation model (middle row, Figure 6). The cloud of points is closer

376 to the 1:1 line and the vertical black line indicating the mean error is closer to zero. In the final row of Figure 6, we

377 see that the multi-variable two-equation model yields the best result by far. The residuals are now evenly distributed

378 with a small bias. Several metrics of performance for the three models, including R^2 (Pearson coefficient), bias, and

379 root-mean-square-error (RMSE), are provided in Table 2. Figure 7 shows the distribution of model residuals for the

380 multi-variable two-equation model as a function of DOY.

381

Table 2: Summary of performance metrics for the three regression models presented in Section 2.2.

Model	R^2	Bias (mm)	RMSE (mm)
One-equation	0.946	-19.5	102
Two-equation	0.962	-5.1	81
Multi-variable two-equation	0.978	-1.2	59

It is useful to also consider the model errors in a non-dimensional way. Therefore, an RMSE was computed at each station location and normalized by the winter precipitation (*PPTWT*) at that location. Figure 8 shows the probability density function of these normalized errors. The average RMSE is approximately 15% of *PPTWT* with most values falling into the range of 5-30%. The spatial distribution of these normalized errors is shown in Figure 9. For the SNOTEL stations, it appears there is a slight regional trend, in terms of stations in continental climates (Rockies) having larger relative errors than stations in maritime climates (Cascades). The British Columbia stations also show higher relative errors.

3.1 Results for Snow Classes

A key objective of this study is to regress climatological information in a continuous rather than a discrete way. The work by Sturm et al. (2010) therefore provides a valuable point of comparison. In that study, the authors developed the following equation for density ρ_b

(11) $\quad \rho_b = (\rho_{max} - \rho_0)\left[1 - e^{(-k_1 h - k_2 DOY)}\right] + \rho_0$

where ρ_0 is the initial density, ρ_{max} is the maximum or 'final' density (end of water year), k_1 and k_2 are coefficients, and *DOY* in this case begins on January 1. This means that their *DOY* for October 1 is -92. The coefficients vary with snow class and the values determined by Sturm et al. (2010) are shown in Table 3.

Table 3: Model parameters by snow class for Sturm et al. (2010).

Snow Class	ρ_{max}	ρ_0	k_1	k_2
Alpine	0.5975	0.2237	0.0012	0.0038
Maritime	0.5979	0.2578	0.0010	0.0038
Prairie	0.5941	0.2332	0.0016	0.0031
Tundra	0.3630	0.2425	0.0029	0.0049
Taiga	0.2170	0.2170	0.0000	0.0000

To make a comparison, the snow class for each SNOTEL and British Columbia snow survey (Rows 1 and 3 of Table 1) site was determined using a 1-km snow class grid (Sturm et al., 2010). The aggregated dataset from these stations was made up of 27% Alpine, 14% Maritime, 10% Prairie, 11% Tundra, and 38% Taiga data points. Equation (11) was then used to estimate snow density (and then SWE) for every point in the validation dataset described in Section 2.2. Figure 10 compares the SWE estimates from the Sturm model and from the current multi-variable, two-equation model (Equations 5, 7-8). The upper left panel of Figure 10 shows all of the data, and the remaining panels show the results for each snow class. In all cases, the current model provides better estimates (narrow cloud of points; closer to the 1:1 line). Plots of the residuals by snow class are provided in Figure 11, giving an indication of the bias of

413 each model for each snow class. Summaries of the model performance, broken out by snow class, are given in Table
414 4. The current model has smaller biases and RMSEs for each snow class.
415
416 Table 4: Comparison of model performance by Sturm et al. (2010) and the current study.

Model	Sturm et al. (2010)			Multi-variable two-equation model		
Snow Class	R^2	Bias (mm)	RMSE (mm)	R^2	Bias (mm)	RMSE (mm)
All Data	0.928	-29.2	111	0.978	-1.2	59
Alpine	0.973	10.1	55	0.978	-2.7	48
Maritime	0.968	-16.8	109	0.975	-7.8	95
Prairie	0.967	18.7	56	0.971	-0.7	45
Tundra	0.956	-10.5	82	0.974	-2.9	59
Taiga	0.943	-80.0	151	0.978	2.6	54

417
418 **3.2 Comparison to Pistocchi (2016)**
419 In order to provide an additional comparison, the simple model of Pistocchi (2016) was also applied to the validation
420 dataset. His model calculates the bulk density as
421
422 (12) $\rho_b = \rho_0 + K(DOY + 61),$
423
424 where ρ_0 has a value of 200 kg m^{-3} and K has a value of 1 kg m^{-3}. The DOY for this model has its origin at
425 November 1. Application of this model to the validation dataset yields a bias of 55 mm and an RMSE of 94 mm.
426 These results are comparable to the Sturm et al. (2010) model, with a larger bias but smaller RMSE.
427
428 **3.3 Comparison to Jonas et al. (2009)**
429 A final point of comparison can be provided by the model of Jonas et al. (2009). The full version of that model
430 contains region-specific offset parameters that are not relevant to North America, so the following partial version of
431 the model is used (their Eq. 4):
432
433 (13) $\rho_b = ah + b,$
434
435 where the parameters (a, b) vary with elevation and month, as given by Table 5. Note that coefficients are not given
436 for every month. Application of the Jonas et al. (2009) model to the snow pillow dataset yields a bias of -5 mm and
437 an RMSE of 69 mm. These results are not directly comparable to those of the current model (Table 2, row 3) since
438 the Jonas et al. (2009) model is unable to compute results for several months of the year. To make a direct
439 comparison to the current model, it is necessary to first remove those data points (about 5%). When this is done, the
440 current model yields a bias of -0.3 mm and an RMSE of 59 mm.
441
442 Table 5: Model coefficients (a, b) for the Jonas et al. (2009) model.

Month	$z > 2000\ m$	$2000\ m > z > 1400\ m$	$1400\ m > z$
January	$(206, 52)$	$(208, 47)$	$(235, 31)$

February	(217, 46)	(218, 52)	(279, 9)
March	(272, 26)	(281, 31)	(333, 3)
April	(331, 9)	(354, 15)	(347, 25)
May	(378, 21)	(409, 29)	(413, 19)
June	(452, 8)	n/a	n/a
July	(470, 15)	n/a	n/a
August	n/a	n/a	n/a
September	n/a	n/a	n/a
October	n/a	n/a	n/a
November	(206, 47)	(183, 35)	(149, 37)
December	(203, 52)	(190, 47)	(201, 26)

443

3.4 Results for Northeast USA

The regression equations in this study were developed using a large collection of snow pillow sites in CONUS, AK, and BC. The snow pillow sites are limited to locations west of approximately W 105° (Figure 2a). By design, the data sets from the northeastern USA (Section 2.1.1.3) were left as an entirely independent validation set. These northeastern sites are geographically distant from the training data sets, subject to a very different climate, largely use different methods (snow coring, with the exception of the SCAN network) and are generally at much lower elevations than the western sites, providing an interesting opportunity to test how robust the current model is.

Figure 12 graphically summarizes the datasets and the performance of the multi-variable two-equation model of the current study. The RMSE values are comparable to those found for the western stations, but, given the comparatively thinner snowpacks in the northeast, represent a larger relative error (Table 5). The bias of the model is consistently positive, in contrast to the western stations where the bias was negligible. Note that Table 5 also includes results from the application of the other three models discussed. Sturm et al. (2010) cannot be applied to several of the datasets since their available 1 km snowclass dataset cuts off at -71.6° longitude. The current model and the Jonas et al. (2009) model perform better than the other two models, with the current model generally outperforming the Jonas et al. (2009) model. The two datasets where the Jonas et al. (2009) model has a slightly better performance are the two smallest datasets (less than 1000 measurements; see Table 1).

Table 5: Performance metrics for various models applied to the northeastern USA datasets. Bold font is used to highlight the model with the best performance for each dataset.

	Multi-variable, two-equation model		Sturm et al. (2010)		Jonas et al. (2009)		Pistocchi (2015)	
Dataset Name	Bias (mm)	RMSE (mm)	Bias (mm)	RMSE (mm)	Bias (mm)	RMSE (mm)	Bias (mm)	RMSE (mm)
Maine Geological Survey, ME	**13.1**	**34.0**	n/a	n/a	25.1	46.0	59.2	77.1
Hubbard Brook (Station 2), NH	21.8	66.6	34.2	76.9	**19.4**	**65.4**	52.0	90.8
Thompson Farm, NH	7.1	20.2	n/a	n/a	**5.6**	**19.9**	20.4	32.3
NRCS SCAN	**-1.2**	**39.2**	8.4	45.0	-2.8	40.6	23.4	56.9
Sleepers River, VT	**14.4**	**28.2**	36.5	48.9	20.4	33.5	55.8	67.1
New York Snow Survey	**14.8**	**31.2**	21.0	49.3	16.3	33.0	41.3	56.1

464

3.5 Results for NRCS Snow Course / Aerial Marker Data

The NRCS snow course and aerial marker data were also left out of the model building process so they provide an additional and completely independent comparison of the various models considered. Recall that these data come from snow course (coring measurements) and aerial surveys, which are different measurement methods than the snow pillows which provided the data for construction of the current regression model. Table 6 shows the results and demonstrates that the current model has the best performance.

Table 6: Performance metrics for various models applied to the NRCS snow course and aerial marker dataset. Bold font is used to highlight the model with the best performance.

	Multi-variable, two-equation model		Sturm et al. (2010)		Jonas et al. (2009)		Pistocchi (2015)	
Dataset Name	Bias (mm)	RMSE (mm)	Bias (mm)	RMSE (mm)	Bias (mm)	RMSE (mm)	Bias (mm)	RMSE (mm)
NRCS Snow Course / Aerial Marker	**0**	**59**	-24	123	24	72	71	99

4 Discussion

The results presented in this study show that the regression equation described by equations (5, 7-8) is an improvement (lower bias and RMSE) over other widely used bulk density equations. The key advantage is that the current method regresses in relevant parameters directly, rather than using discrete bins (for snow class, elevation, month of year, etc.), each with its own set of model coefficients. The comparison (Figs. 10-11; Table 4) to the model of Sturm et al. (2010) reveals a peculiar behavior of that model for the Taiga snow class, with a large negative bias in the Sturm estimates. Inspection of the coefficients provided for that class (Table 3) shows that the model simply predicts that $\rho_b = \rho_{max} = 0.217$ for all conditions.

When our multi-variable two-equation model, developed solely from western North American data, is applied to northeast USA locations, it produces SWE estimates with smaller RSME values and larger biases than the western stations. When comparing the SWE-h scatter plots of the SNOTEL data (Figure 4b) to those of the east coast data sets (left column; Figure 12), it is clear that the northeast data generally have more scatter. This is confirmed by computing the correlation coefficients between SWE and h for each dataset. It is unclear if this disparity in correlation is related to measurement methodology or is instead a 'signal to noise' issue. Comparing Figures 4 and 12 shows the considerable difference in snowpack depth between the western and northeastern data sets. When the western dataset is filtered to include only measurement pairs where $h < 1.5$ m, the correlation coefficient is reduced to a value consistent with the northeast datasets. This suggests that the performance of the current (or other) regression model is not as good at shallow snowpack depths. This is also suggested upon examination of the time series of observed $\rho_b = SWE/h$ for a given season at a snow pillow site. Very early in the season, when the depths are small, the density curve has a lot of variability. Later in the season, when depths are greater, the density curve becomes much smoother. Very late in the season, when depths are low again, the density curve becomes highly variable again.

498   Measurement precision and accuracy affect the construction and use of a regression model. Upon inspection of the

499   snow pillow data, it was observed that the precision of the depth measurements was approximately 25 mm and that

500   of the SWE measurements was approximately 2.5 mm. To test the sensitivity of the model coefficients to the

501   measurement precision, the depth values in the training dataset were randomly perturbed by +/- 25 mm and the SWE

502   values were randomly perturbed by +/- 2.5 mm and the regression coefficients were recomputed. This process was

503   repeated numerous times and the mean values of the perturbed coefficients were obtained. These adjusted

504   coefficients were then used to recompute the SWE values for the validation data set and the bias and RMSE were

505   found to be -10.5 mm and 72.7 mm. This represents a roughly 10% increase in RMSE, but a considerable increase in

506   bias magnitude (see Table 4 for the original values). This sensitivity of the regression analysis to measurement

507   precision underscores the need to have high-precision measurements for the training data set. Regarding accuracy,

508   random and systematic errors in the paired SWE - h data used to construct the regression model will lead to

509   uncertainties in SWE values predicted by the model. As noted in the introduction, snow pillow errors in SWE

510   estimates do not follow a simple pattern. Additionally, they are complicated by the fact that the errors are often

511   computed by comparing snow pillow data to coring data, which itself is subject to error. Lacking quantitative

512   information on the distribution of snow pillow errors, we are unable to quantify the uncertainty in the SWE

513   estimates.

514

515   Another important consideration has to do with the uncertainty of depth measurements that the model is applied to.

516   For context, one application of this study is to crowd-sourced, opportunistic snow depth measurements from

517   programs like the Community Snow Observations (CSO; Hill et al., 2018) project. In the CSO program,

518   backcountry recreational users submit depth measurements, typically taken with an avalanche probe, using a

519   smartphone in the field. The measurements are then converted to SWE estimates which are assimilated into

520   snowpack models. These depth measurements are 'any time, any place' in contrast to repeated measurements from

521   the same location, like snow pillows or snow courses. Most avalanche probes have cm-scale graduated markings, so

522   measurement precision is not a major issue. A larger problem is the considerable variability in snowpack depth that

523   can exist over short (meter scale) distances. The variability of the Chugach avalanche probe measurements was

524   assessed by taking the standard deviation of 8 h measurements per site. The average of this standard deviation over

525   the sites was 22 cm and the average coefficient of variation (standard deviation normalized by the mean) over the

526   sites was 15%. This variability is a function of the surface roughness of the underlying terrain, and also a function of

527   wind redistribution of snow. Propagating this uncertainty through the regression equations yields a slightly higher

528   (16%) uncertainty in the SWE estimates. CSO participants can do three things to ensure that their recorded depth

529   measurements are as representative as possible. First, avoid measurements in areas of significant wind scour or

530   deposition. Second, avoid measurements in terrain likely to have significant surface roughness (rocks, fallen logs,

531   etc.). Third, take several measurements and average them.

532

533   Expansion of CSO measurements in areas lacking SWE measurements can increase our understanding of the

534   extreme spatial variability in snow distribution and the inherent uncertainties associated with modeling SWE in

535   these regions. It could also prove useful for estimating watershed-scale SWE in regions like the northeastern USA,

536   which is currently limited to five automated SCAN sites with historical SWE measurements for only the past two

537   decades. Additionally, historical snow depth measurements are more widely available in the Global Historical

538   Climatology Network (GHCN-Daily; Menne et al. 2012), with several records extending back to the late 1800s.

539   While many of the GHCN stations are confined to lower elevations with shallower snow depths, the broader

540   network of quality-controlled snow depth data paired with daily GHCN temperature and precipitation measurements

541   could potentially be used to reconstruct SWE in the eastern US given additional model development and refinement.

542   **5 Conclusions**

543   We have developed a new, easy to use method for converting snow depth measurements to snow water equivalent

544   estimates. The key difference between our approach and previous approaches is that we directly regress in

545   climatological variables in a continuous fashion, rather than a discrete one. Given the abundance of freely available

546   climatological norms, a depth measurement tagged with coordinates (latitude and longitude) and a time stamp is

547   easily and immediately converted into SWE.

548

549   We developed this model with data from paired SWE-h measurements from the western United States and British

550   Columbia. The model was tested against entirely independent data (primarily snow course; some snow pillow) from

551   the northeastern United States and was found to perform well, albeit with larger biases and root-mean-squared-

552   errors. The model was tested against other well-known regression equations and was found to perform better. The

553   model was also tested against a large dataset of independent snow course and aerial marker measurements from

554   western North America. For this second independent test, the current model outperformed the other models

555   considered.

556

557   This model is not a replacement for more sophisticated snow models that evolve the snowpack based on high

558   frequency (e.g., daily or sub-daily) weather data inputs. The intended purpose of this model is to constrain SWE

559   estimates in circumstances where snow depth is known, but weather variables are not, a common issue in sparsely

560   instrumented areas in North America.

561   **6 Acknowledgements**

562   Support for this project was provided by NASA (NNX17AG67A). R. Crumley acknowledges support from the

563   CUAHSI Pathfinder Fellowship. E. Burakowski acknowledges support from NSF (MSB-ECA #1802726). We thank

564   M. Sturm, A. Winstral and a third anonymous referee for their careful and thoughtful reviews of this manuscript.

565   **7 Data Access**

566   Numerous online datasets were used for this project and were obtained from the following locations:

567   1. NRCS Snow Telemetry: https://www.wcc.nrcs.usda.gov/snow/SNOTEL-wedata.html

568   2. NRCS Soil Climate Analysis Network: https://www.wcc.nrcs.usda.gov/scan/

569   3.   British Columbia Automated Snow Weather Stations:
570   https://www2.gov.bc.ca/gov/content/environment/air-land-water/water/water-science-data/water-data-
571   tools/snow-survey-data/automated-snow-weather-station-data
572   4.   Maine Cooperative Snow Survey: https://mgs-maine.opendata.arcgis.com/datasets/maine-snow-survey-data
573   5.   New York Snow Survey: http://www.nrcc.cornell.edu/regional/snowsurvey/snowsurvey.html
574   6.   Sleepers River Research Watershed. Snow data not available online; request data from contact at:
575   https://nh.water.usgs.gov/project/sleepers/index.htm
576   7.   Hubbard Brook Experimental Forest: https://hubbardbrook.org/d/hubbard-brook-data-catalog
577   8.   Climatological Data: https://adaptwest.databasin.org/pages/adaptwest-climatena
578   9.   NRCS Snow Course / Aerial Marker Data: https://wcc.sc.egov.usda.gov/reportGenerator/
579
580
581   A Matlab function for calculating SWE based on the results is this paper has been made publicly available at Github

582   (https://github.com/communitysnowobs/snowdensity).

583

584 **References**

585

586 Alford, D.: Density variations in alpine snow, J. Glaciol., 6(46), 495-503,

587 https://doi.org/10.3189/S0022143000019717, 1967.

588

589 Avanzi, F., De Michele, C., and Ghezzi, A.: On the performances of empirical regressions for the estimation of bulk

590 snow density, Geogr. Fis. Dinam. Quat., 38, 105-112, doi:10.4461/GFDQ.2015.38.10, 2015.

591

592 Beaumont, R.: Mt. Hood pressure pillow snow gage, J. Appl. Meteorol., 4, 626-631, https://doi.org/10.1175/1520-

593 0450(1965)004<0626:MHPPSG>2.0.CO;2 1965.

594

595 Beaumont, R., and Work, R.: Snow sampling results from three samplers, Hydrol. Sci. J., 8(4), 74-78,

596 https://doi.org/10.1080/02626666309493359, 1963.

597

598 Burakowski, E.A., Ollinger, S., Lepine, L., Schaaf, C.B., Wang, Z., Dibb, J.E., Hollinger, D.Y., Kim, J.-H., Erb, A.,

599 and Martin, M.E.: Spatial scaling of reflectance and surface albedo over a mixed-use, temperate forest landscape

600 during snow-covered periods, Remote Sens. Environ., 158, 465-477, https://doi.org/10.1016/j.rse.2014.11.023,

601 2015.

602

603 Burakowski, E.A., Wake, C.P., Stampone, M., and Dibb, J.: Putting the Capital 'A' in CoCoRAHS: An

604 Experimental Program to Measure Albedo using the Community Collaborative Rain Hail and Snow (CoCoRaHS)

605 Network, Hydrol. Process., 27(21), 3024-3034, https://doi.org/10.1002/hyp.9825, 2013.

606

607 Campbell, J., Ollinger, S., Flerchinger, G., Wicklein, H., Hayhoe, K., and Bailey, A.: Past and projected future

608 changes in snowpack and soil frost at the Hubbard Brook Experimental Forest, New Hampshire, USA, Hydrol.

609 Process., 24, 2465-2480, https://doi.org/10.1002/hyp.7666, 2010.

610

611 Church, J.E.: Snow surveying: its principles and possibilities, Geogr. Rev., 23(4), 529-563, DOI: 10.2307/209242,

612 1933.

613

614 Church, J.E., and Marr, J.C.: Further improvement of snow-survey apparatus, Transactions of the American

615 Geophysical Union, 18(2), 607-617, 10.1029/TR018i002p00607, 1937.

616

617 Daly, C., Neilson, R., and Phillips, D.: A statistical-topographic model for mapping climatological precipitation over

618 mountainous terrain, J. Appl. Meteorol., 33, 140-158, https://doi.org/10.1175/1520-

619 0450(1994)033<0140:ASTMFM>2.0.CO;2, 1994.

620

621   Dixon, D., and Boon, S.: Comparison of the SnowHydro sampler with existing snow tube designs, Hydrol. Process.,
622   26(17), 2555-2562, https://doi.org/10.1002/hyp.9317, 2012.

623

624   Dressler, K., Fassnacht, S., and Bales, R.: A comparison of snow telemetry and snow course measurements in the
625   Colorado River basin, J. Hydrometeorol., 7, 705-712, https://doi.org/10.1175/JHM506.1, 2006.

626

627   Goodison, B.: Accuracy of snow samplers for measuring shallow snowpacks: An update, Proceedings of the 35th
628   Annual Eastern Snow Conference, Hanover, NH, 36-49, 1978.

629

630   Goodison, B., Ferguson, H., and McKay, G.: Measurement and data analysis. The Handbook of Snow: Principles,
631   Processes, Management, and Use, D. Gray and D. Male, Eds., Pergamon Press, 191-274, 1981.

632

633   Goodison, B., Wilson, B., Wu., K, and Metcalfe, J.: An inexpensive remote snow-depth gauge: An assessment,
634   Proceedings of the 52nd Annual Western Snow Conference, Sun Valley, ID, 188-191, 1984.

635

636   Goodison, B., Glynn, J., Harvey, K., and Slater, J.: Snow Surveying in Canada: A Perspective, Can. Water Resour.
637   J., 12(2), 27-42, https://doi.org/10.4296/cwrj1202027, 1987.

638

639   Gnanadesikan, R., and Kettenring, J.: Robust estimates, residuals, and outlier detection with multiresponse data,
640   Biometrics, 28, 81-124, DOI: 10.2307/2528963, 1972.

641

642   Hill, D.F., Wolken, G. J., Wikstrom Jones, K., Crumley, R., and Arendt, A.: Crowdsourcing snow depth data with
643   citizen scientists, Eos, 99, https://doi.org/10.1029/2018EO108991, 2018.

644

645   Johnson, J., and Marks, D.: The detection and correction of snow water equivalent pressure sensor errors, Hydrol.
646   Proc., https://doi.org/10.1002/hyp.5795, 2004.

647

648   Johnson, J., Gelvin, A., Duvoy, P., Schaefer G., Poole, G., and Horton, G.: Performance characteristics of a new
649   electronic snow water equivalent sensor in different climates, Hydrol. Proc., DOI: 10.1002/hyp.10211, 2015.

650

651   Jonas, T., Marty, C., and Magnusson, M.: Estimating the snow water equivalent from snow depth measurements, J.
652   Hydrol., 378, 161-167, https://doi.org/10.1016/j.jhydrol.2009.09.021, 2009.

653

654   Leys, C., Klein, O., Dominicy, Y., and Ley, C.: Detecting multivariate outliers: use a robust variant of the
655   Mahalanobis distance, J. Exp. Soc. Psychol., 74, 150-156, https://doi.org/10.1016/j.jesp.2017.09.011, 2018.

656

657 Liang, X., Lettermaier, D., Wood, E., and Burges, S.: A simple hydrologically based model of land surface water
658 and energy fluxes for general circulation models, J. Geophys. Res. Atmos., 99(D7), 14,415-14,428,
659 https://doi.org/10.1029/94JD00483, 1994.

661 Liston, G., and Elder, K.: A distributed snow evolution modeling system (SnowModel), J. Hydrometerol., 7, 1259-
662 1276, https://doi.org/10.1175/JHM548.1, 2006.

664 Lundberg, A., Richardson-Naslund, C., and Andersson, C.: Snow density variations: consequences for ground
665 penetrating radar, Hydrol. Process., 20, 1483-1495, https://doi.org/10.1002/hyp.5944, 2006.

667 Maine Geological Survey: Maine Cooperative Snow Survey Dataset,
668 https://www.maine.gov/dacf/mgs/hazards/snow_survey/, 2018.

670 McKay, G., and Findlay, B., 1971: Variation of snow resources with climate and vegetation in Canada, Proceedings
671 of the 39th Western Snow Conference, Billings, MT, 17-26, 1971.

673 De Maesschalck, R., Jouan-Rimbaud, D., and Massart, D.: The Mahalanobis distance, Chemometr. Intell. Lab. Syst.,
674 50(1), 1-18, https://doi.org/10.1016/S0169-7439(99)00047-7, 2000.

676 McCreight, J., and Small, E.: Modeling bulk density and snow water equivalent using daily snow depth
677 observations, The Cryosphere, 8, 521-536, https://doi.org/10.5194/tc-8-521-2014, 2014.

679 Meloysund, V., Leira, B., Hoiseth, K., and Liso, K.: Predicting snow density using meterological data, Meteorol.
680 Appl., 14, 413-423, https://doi.org/10.1002/met.40, 2007.

682 Menne, M.J., I. Durre, R.S. Vose, B.E. Gleason, and Houston, T.G.: An overview
683 of the Global Historical Climatology Network-Daily Database, J. Atmos. Ocean. Technol., 29, 97-910,
684 doi:10.1175/JTECH-D-11-00103.1, 2012.

686 Molotch, N.P., and Bales, R.C.: SNOTEL representativeness in the Rio Grande headwaters on the basis of
687 physiographics and remotely sensed snow cover persistence, Hydrol. Process., 20(4), 723-739,
688 https://doi.org/10.1002/hyp.6128, 2006.

690 Mote, P., Li, S., Letternaier, D., Xiao, M., and Engel, R.: Dramatic declines in snowpack in the western US," npj
691 Clim. Atmos. Sci., 1(2), 1-6, doi:10.1038/s41612-018-0012-1, 2018.

692

693   Mizukami, N., and Perica, S.: Spatiotemporal characteristics of snowpack density in the mountainous regions of the
694   western United States, J. Hydrometeorol., 9, 1416-1426, https://doi.org/10.1175/2008JHM981.1, 2008.
695
696   New York Snow Survey, NOAA, Northeast Regional Climate Center at Cornell University, 2018.
697
698   Pagano, T., Garen, D., Perkins, T., and Pasteris, P.: Daily updating of operational statistical seasonal water supply
699   forecasts for the western U.S., J. Am. Water Resour. Assoc., 45(3), 767-778, https://doi.org/10.1111/j.1752-
700   1688.2009.00321.x, 2009.
701
702   Painter, T., Berisford, D., Boardman, J., Bormann, K., Deems, J., Gehrke, F., Hedrick, A., Joyce, M., Laidlaw, R.,
703   Marks, D., Mattmann, C., Mcgurk, B., Ramirez, P., Richardson, M., Skiles, S., Seidel, F., and Winstral, A.: The
704   Airborne Snow Observatory: fusion of scanning lidar, imaging spectrometer, and physically-based modeling for
705   mapping snow water equivalent and snow albedo, Remote Sens. Environ., 184, 139-152,
706   doi:10.1016/j.rse.2016.06.018, 2016.
707
708   Pistocchi, A.: Simple estimation of snow density in an Alpine region, J. Hydrol. Reg. Stud., 6, 82-89,
709   http://dx.doi.org/10.1016/j.ejrh.2016.03.004, 2016.
710
711   Rousseeuw, P.: Least Median of Squares Regression, J. Am. Stat. Assoc., 79, 871-880, DOI:
712   10.1080/01621459.1984.10477105, 1984.
713
714   Ryan, W., Doesken, N., and Fassnacht, S.: Evaluation of Ultrasonic Snow Depth Sensors for U.S. Snow
715   Measurements, J. Atmos. Ocean. Technol., 25, 667-684, https://doi.org/10.1175/2007JTECHA947.1, 2008.
716
717   Schaefer, G., Cosh, M., and Jackson, T.: The USDA Natural Resources Conservation Service Soil Climate Analysis
718   Network (SCAN), J. Atmos. Ocean. Technol., 24, 2073-2077, https://doi.org/10.1175/2007JTECHA930.1, 2007.
719
720   Serreze, M., Clark, M., Armstrong, R., McGinnis, D., and Pulwarty, R.: Characteristics of the western United States
721   snowpack from snowpack telemetry (SNOTEL) data, Water Resour. Res., 35(7), 2145-2160,
722   https://doi.org/10.1029/1999WR900090, 1999.
723
724   Shanley, J., and Chalmers, A.: The effect of frozen soil on snowmelt runoff at Sleepers River, Vermont, Hydrol.
725   Process., 13(12-13), 1843-1857, https://doi.org/10.1002/(SICI)1099-1085(199909)13:12/13<1843::AID-
726   HYP879>3.0.CO;2-G, 1999.
727
728   Sturm, M., Holmgren, J., and Liston, G.: A seasonal snow cover classification system for local to global
729   applications, J. Clim., 8, 1261-1283, https://doi.org/10.1175/1520-0442(1995)008<1261:ASSCCS>2.0.CO;2, 1995.

730

731 Sturm, M., Taras, B., Liston, G.E., Derksen, C., Jonas, T., and Lea, J.: Estimating snow water equivalent using snow

732 depth data and climate classes, J. Hydrometeorol., 11, 1380-1394, https://doi.org/10.1175/2010JHM1202.1, 2010.

733

734 U.S. Army Corps of Engineers: Snow hydrology: Summary report of the snow investigations of the North Pacific

735 Division, 437pp., 1956.

736

737 U.S. Department of Agriculture: The History of Snow Survey and Water Supply Forecasting. Interviews With U.S.

738 Department of Agriculture Pioneers, D. Helms, S. Phillips and P. Reich (eds.), Natural Resources Conservation

739 Service, 2008.

740

741 U.S. Department of Agriculture: Snow Survey and Water Supply Forecasting. National Engineering Handbook Part

742 622, Water and Climate Center, Natural Resources Conservation Service, 2011.

743

744 Wang, T., Hamann, A., Spittlehouse, D.L., and Murdock, T.: ClimateWNA - High-Resolution Spatial Climate Data

745 for Western North America, J. Appl. Meteorol. Climatol., 51, 16-29, https://doi.org/10.1175/JAMC-D-11-043.1,

746 2012.

747

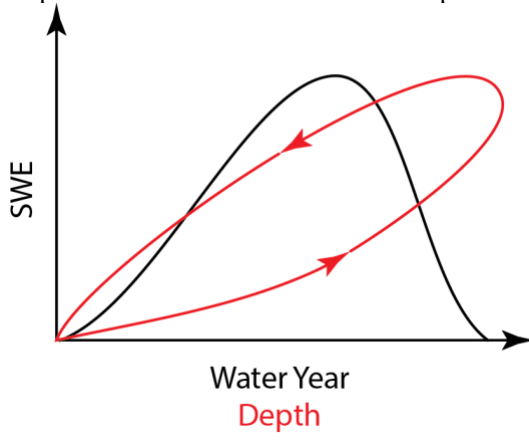
748 Wang, T., Hamann, A., Spittlehouse, D.L., and Carroll, C: Locally downscaled and spatially customizable climate

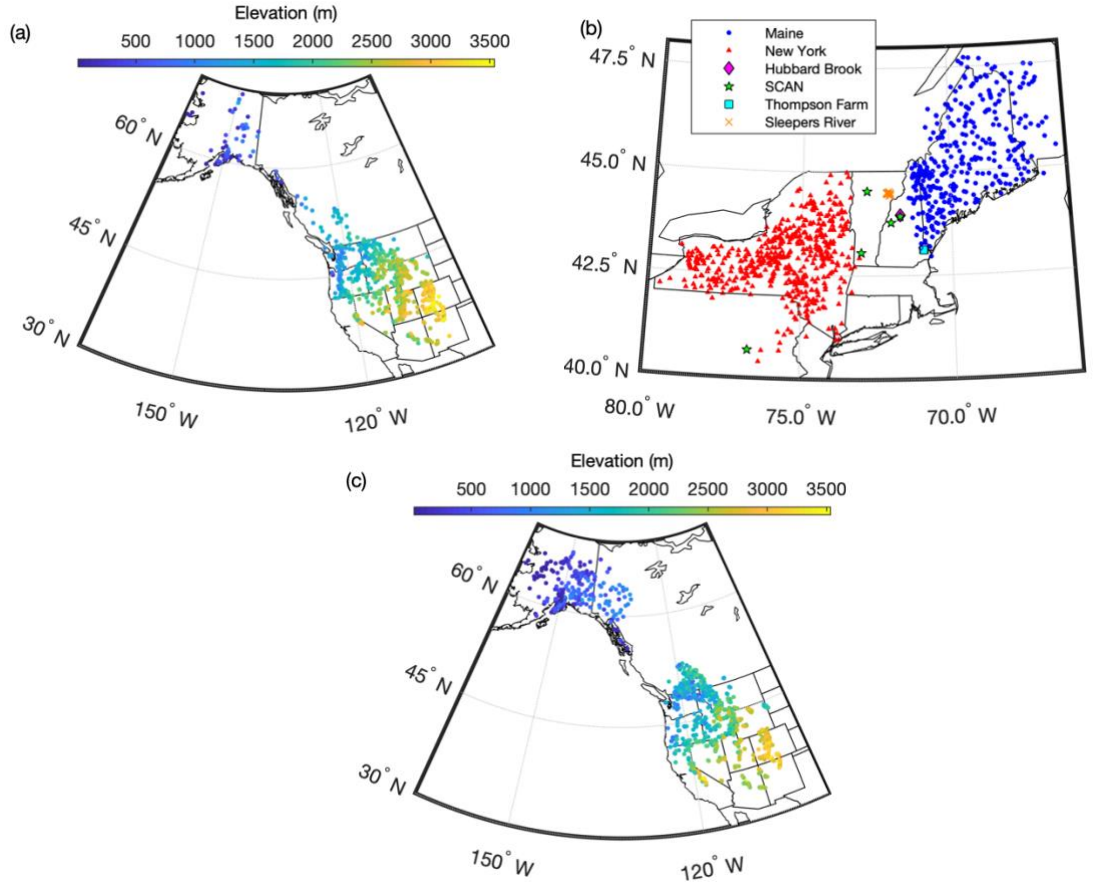
749 data for historical and future periods for North America, PLoS One, 11, DOI:10.1371/journal.pone.0156720.

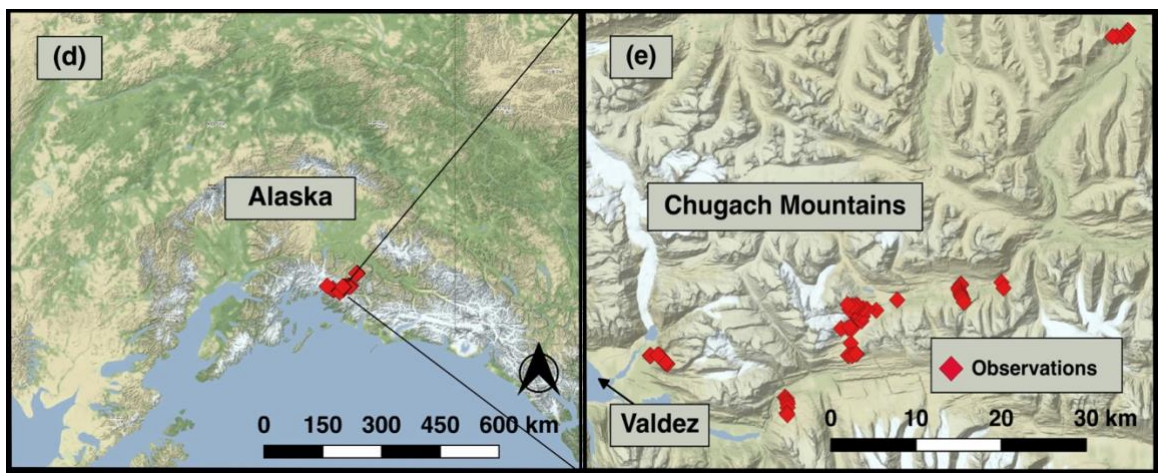
750

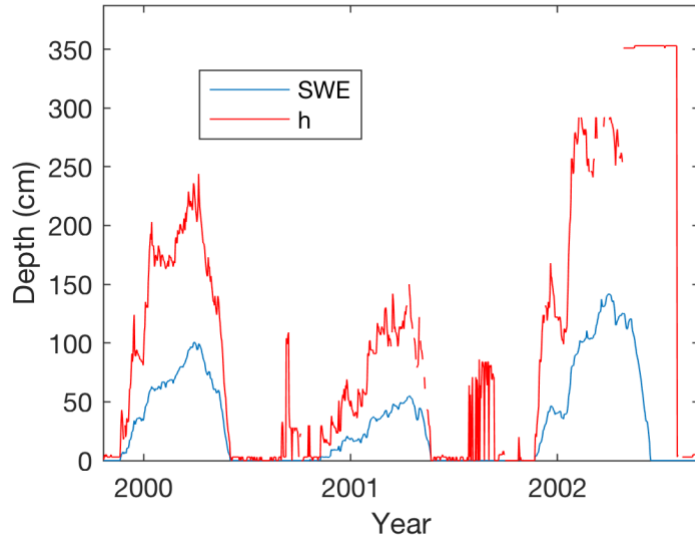
751 Wigmosta, M.S., Vail, L., and Lettenmaier, D.: A distributed hydrology-vegetation model for complex

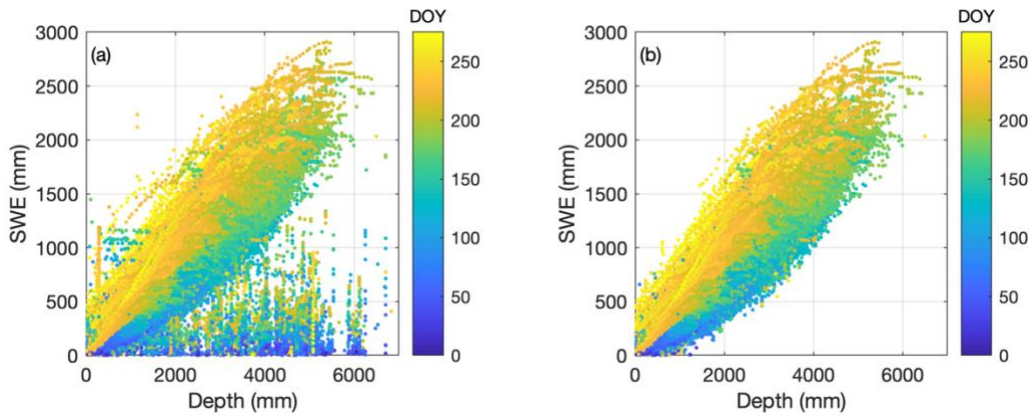
752 terrain, Water Resour. Res., 30, 1665-1679, https://doi.org/10.1029/94WR00436, 1994

753   Figure 1: Conceptual sketch of the evolution of snow water equivalent (SWE) over the course of a water year (black
754   line). Also shown is the evolution of SWE with snowpack depth over a water year (red line). Note the hysteresis
755   loop due to the densification of the snowpack.

756

757 Figure 2: Distribution of measurement locations used in this study. (a) Western USA and Canada snow pillow
758 locations, with colors indicating station elevation in meters. (b) Northeast USA snow pillow and snow course
759 locations, with stations colored according to data source. (c) Western North America snow course and aerial marker
760 locations, with colors indicating station elevation in meters. (d, e) Measurement sites in the Chugach Mountains,
761 southcentral Alaska.
762

763

764

765

766 Figure 3: Sample time series of SWE and *h* from the Rex River (WA) SNOTEL station. Observations of *h* at times
767 when SWE is zero are likely spurious.

768

769 Figure 4: Scatter plot of SWE vs. *h* for the complete SNOTEL dataset before (a) and after (b) removing data points,
770 following the method described in Section 2.1.1.5. Symbols are colored by 'day of water year' (*DOY*; October 1 is
771 the origin).
772

773

774 Figure 5: Gridded maps of winter (December, January, February) precipitation (PPTWT) and temperature difference
775 (TD) between mean of warmest month and mean of coldest month) for North America. Maps are for the 1981-2010
776 climatological period.

777

778
779
780
781

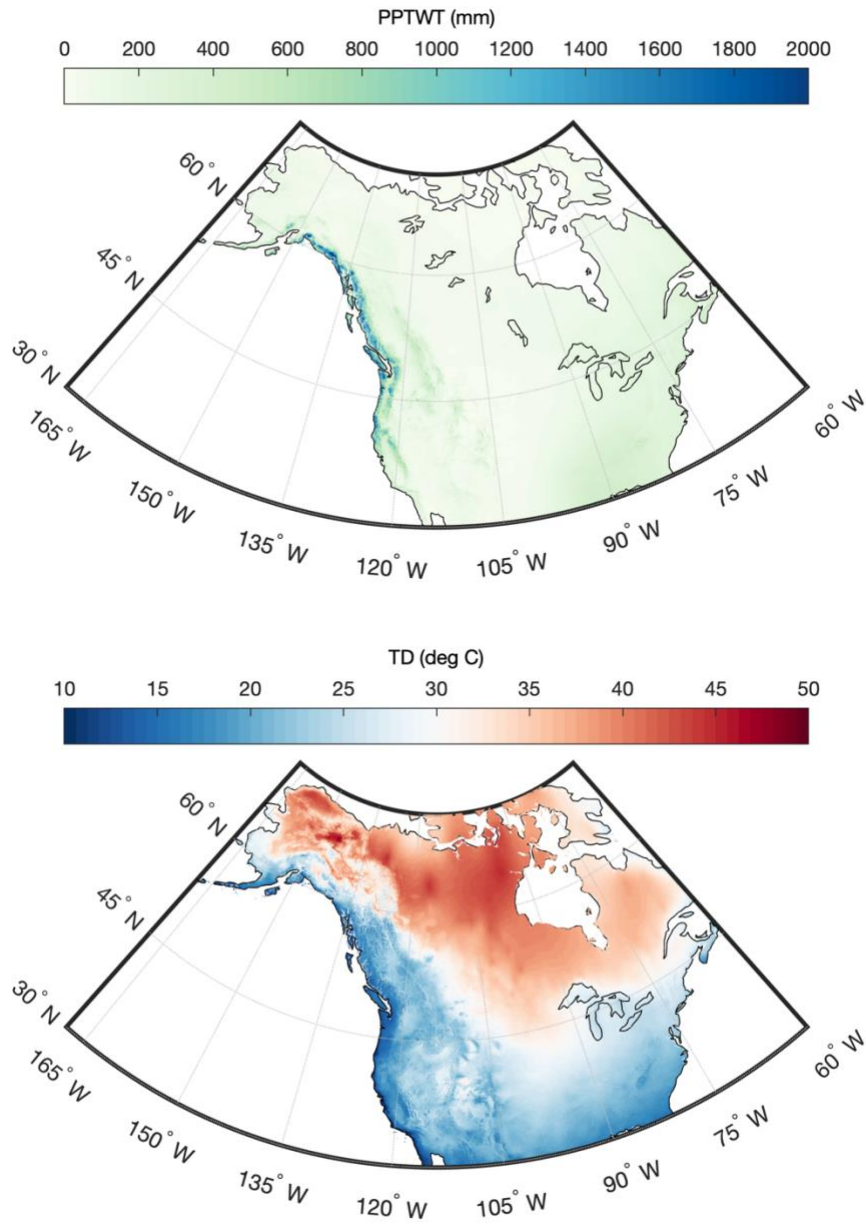
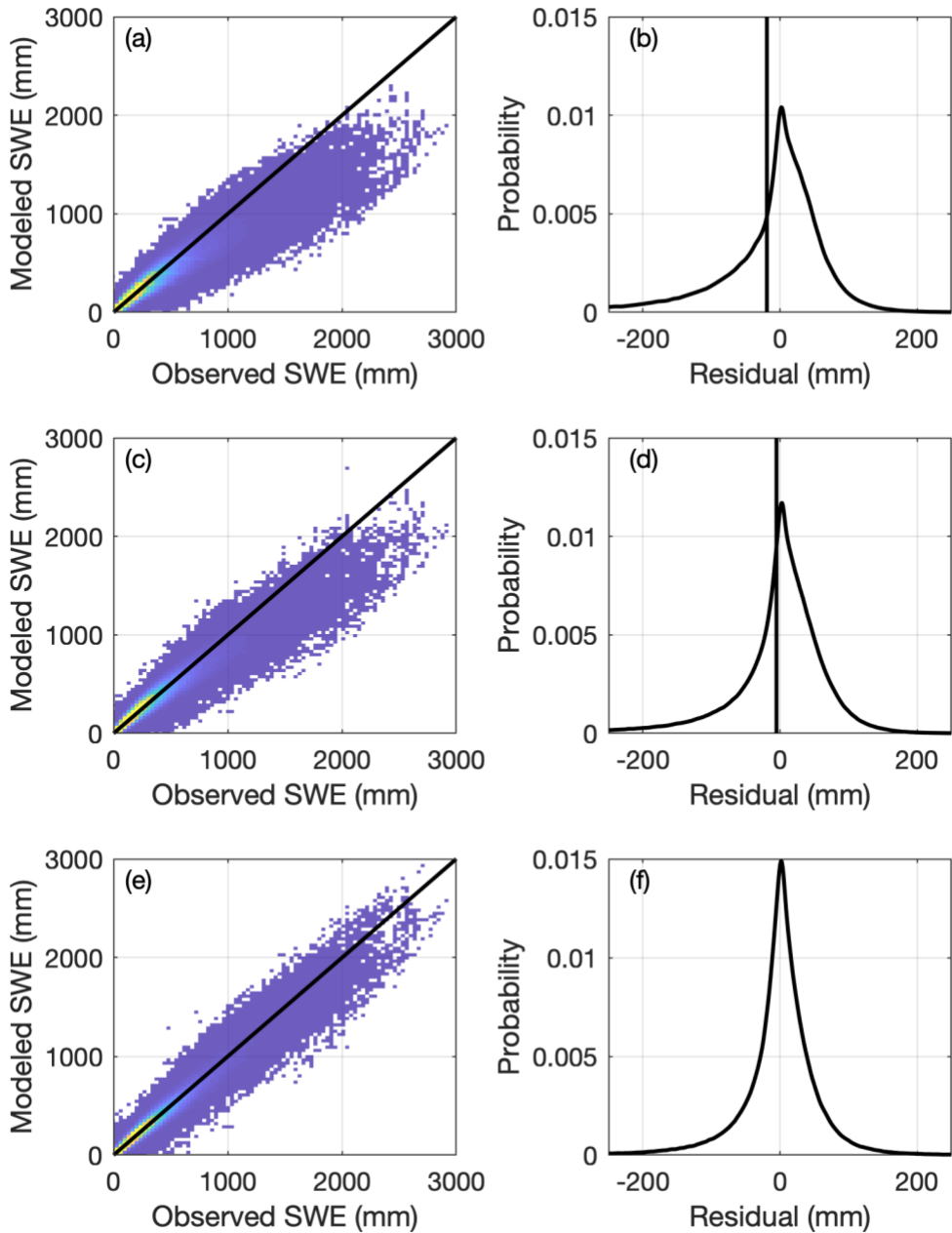
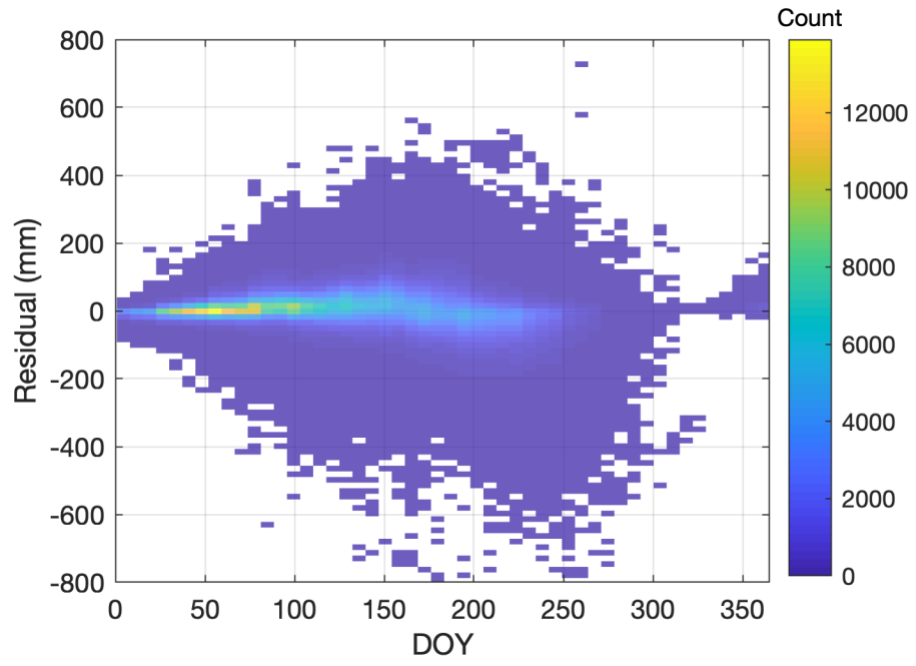
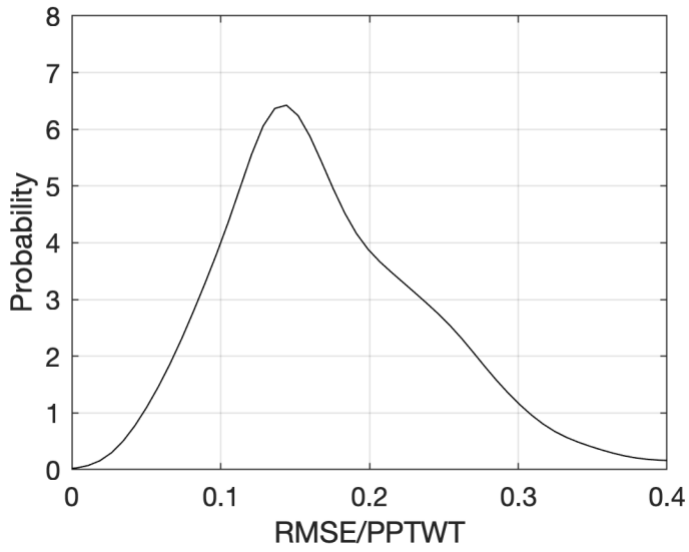
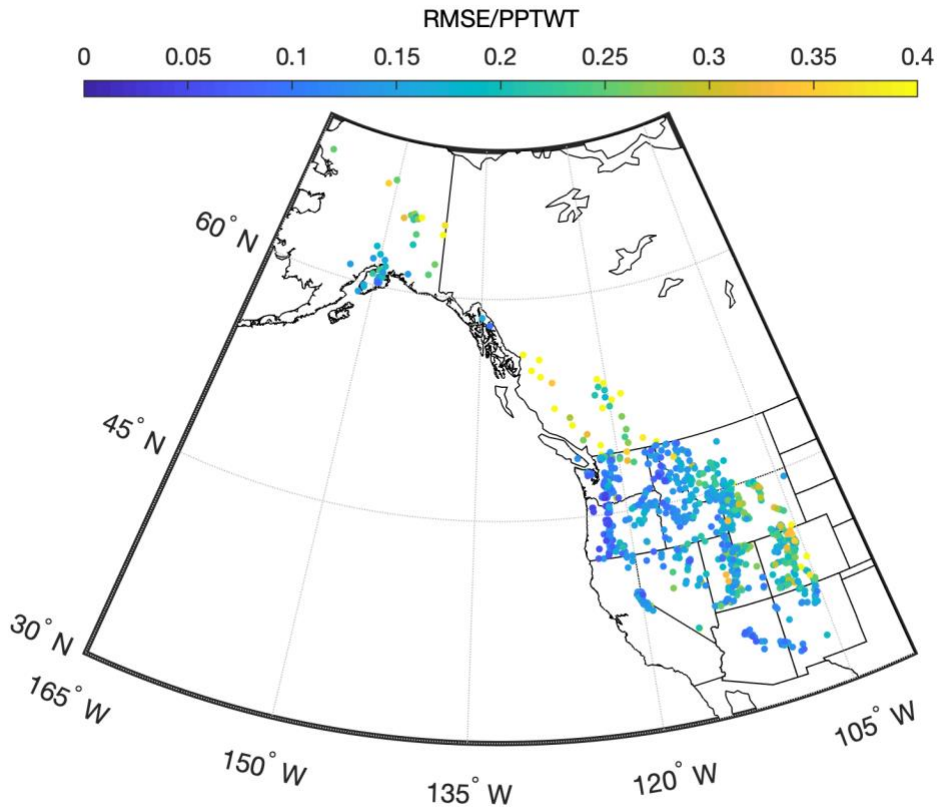
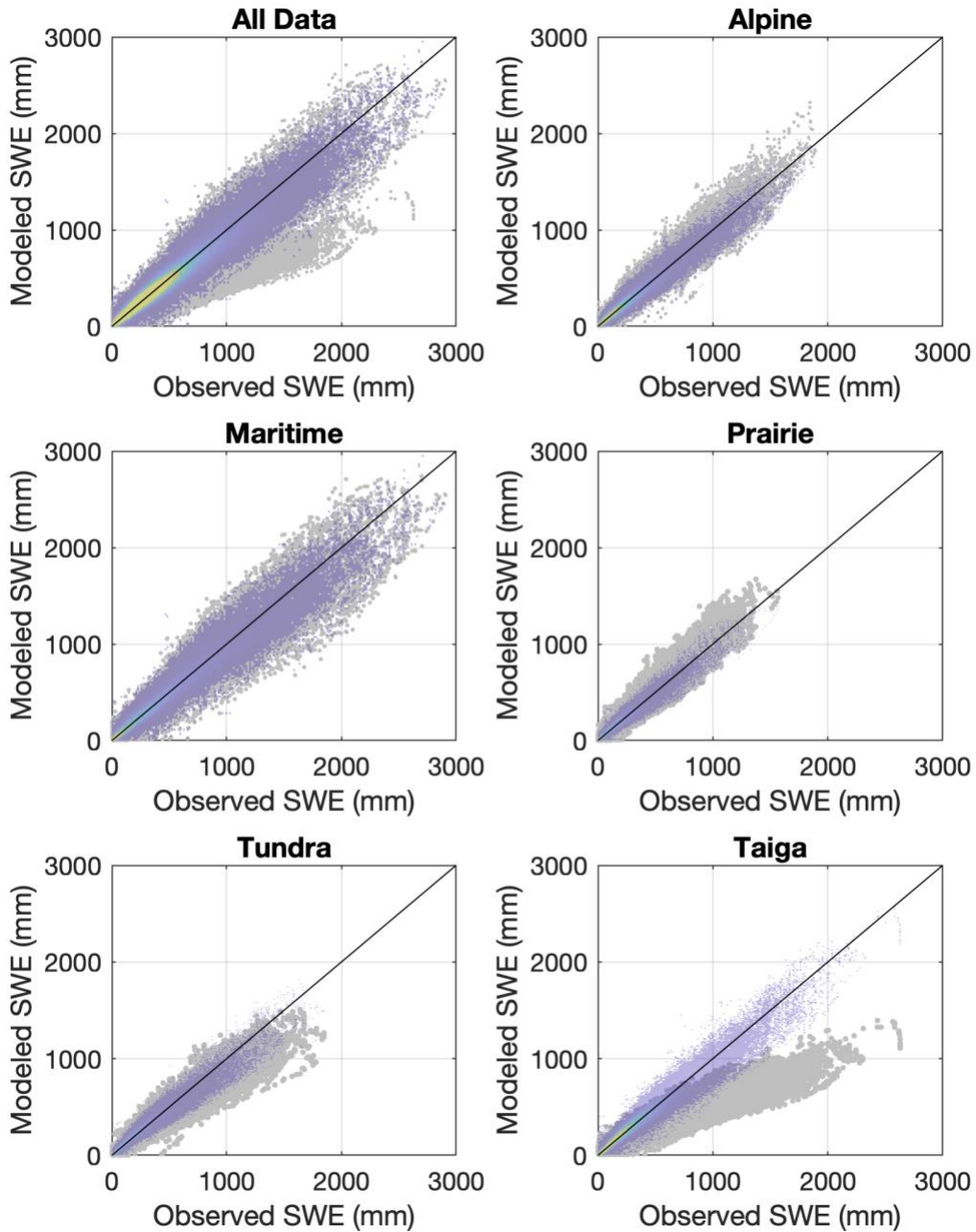
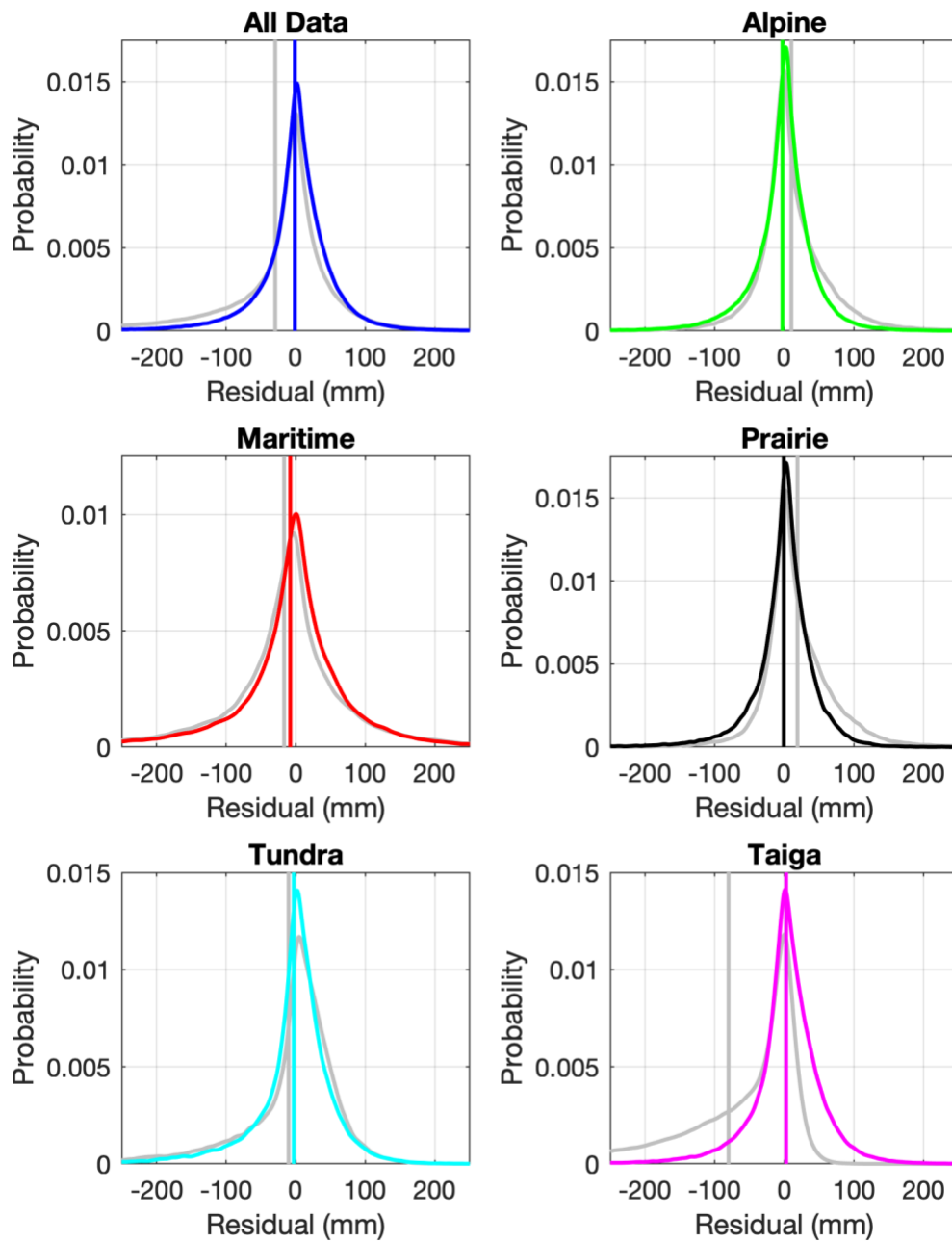



Figure 6: Two-dimensional histograms (heat maps; left column) of modeled vs. observed SWE and probability density functions (right column) of the residuals for three simple models applied to the CONUS, AK, and BC snow pillow data. Warmer colors in the heat maps indicate greater density. The vertical lines in the right column indicate the location of the mean residual, or bias. Top row (a-b): One-equation model (Section 2.2.1). Middle row (c-d): Two-equation model (Section 2.2.2). Bottom row (e-f): Multi-variable two-equation model (Section 2.2.3).

789 Figure 7: Heat map of SWE residuals as a function of *DOY* for the application of the multi-variable two-equation
790 model to the western North America snow pillow validation dataset.
791

792
793
794

795   Figure 8: Probability density function of snow pillow station root-mean-square error (RMSE) normalized by station
796   winter precipitation (*PPTWT*) for the application of the multi-variable two-equation model to the western North
797   America snow pillow validation dataset.

798

799 Figure 9: Spatial distribution of snow pillow station root-mean-square error (RMSE) normalized by station winter
800 precipitation (PPTWT) for the application of the multi-variable two-equation model to the western North America
801 snow pillow validation dataset.

802
803

804 Figure 10: Comparison of the multi-variable, two-equation model of the current study with the model of Sturm et al.
805 (2010), applied to the western North America snow pillow validation dataset. The subpanels show modeled SWE vs.
806 observed SWE for all of the data binned together, as well as for the data broken out by the snow classes identified by
807 Sturm et al. (1995). The gray symbols show the Sturm result and the transparent heat maps (warmer colors indicate
808 greater density) show the current result.
809

810

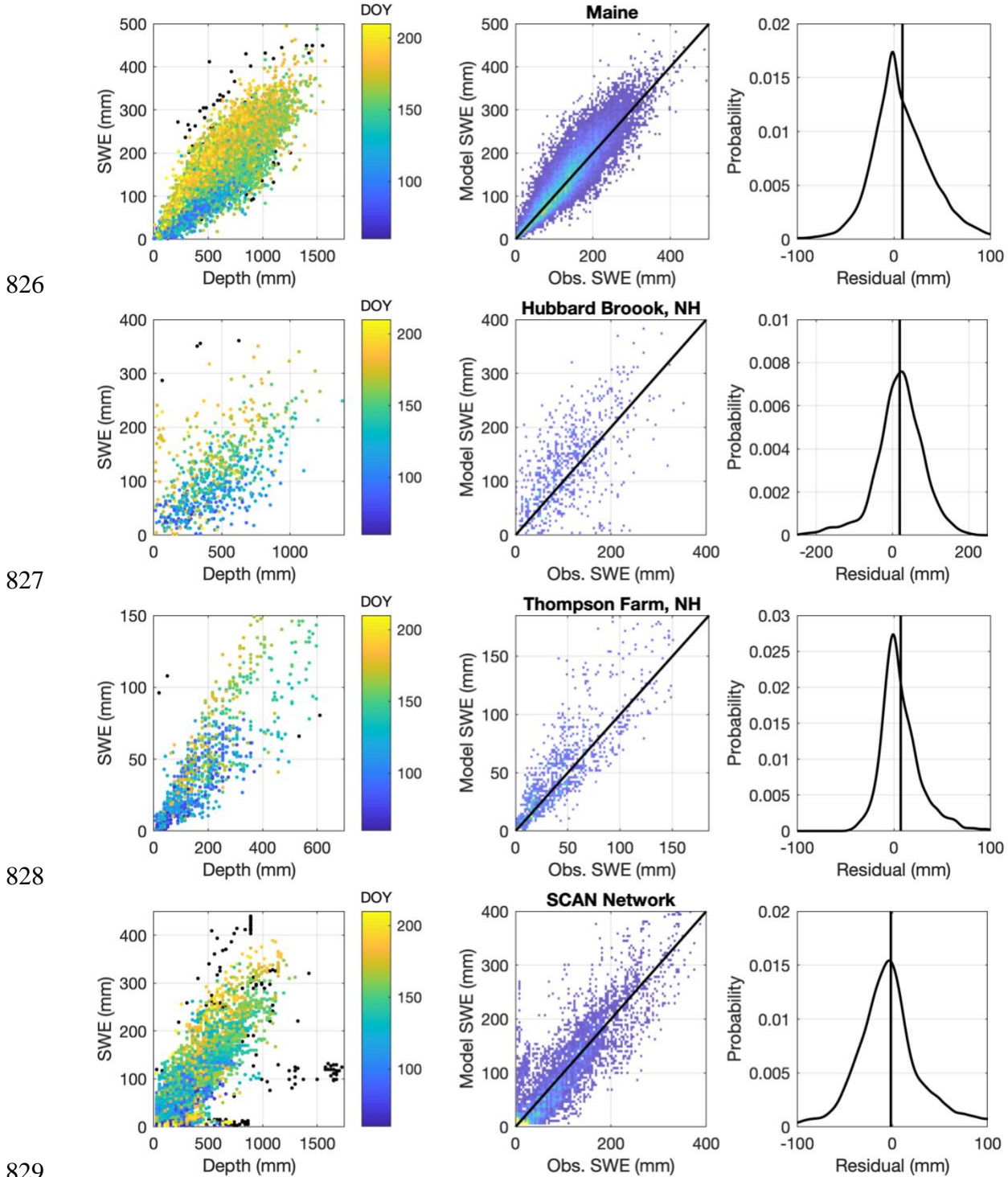
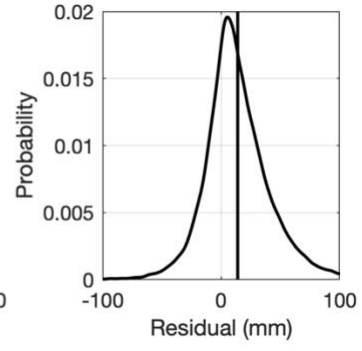
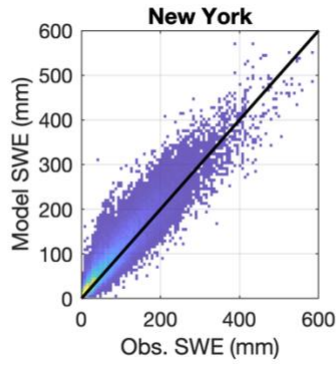
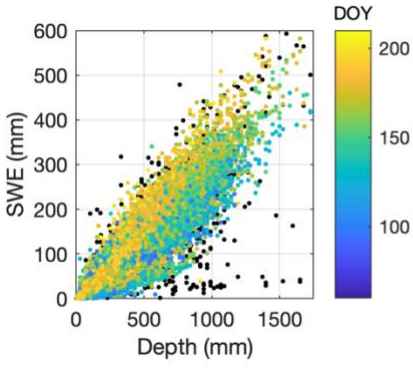
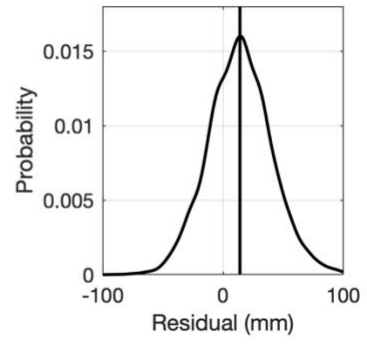
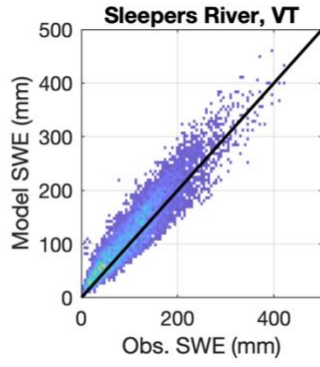
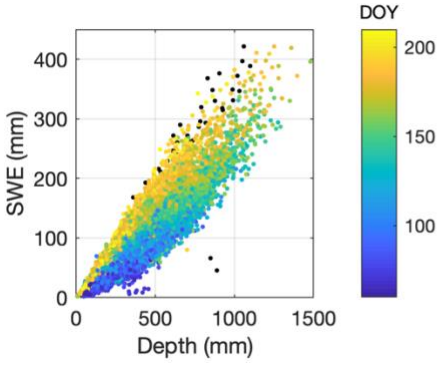
811 Figure 11: Comparison of the multi-variable, two-equation model of the current study with the model of Sturm et al.
812 (2010), applied to the western North America snow pillow validation dataset. The subpanels show probability
813 density functions of the residuals of the model fits for all of the data binned together, as well as for the data broken
814 out by the snow classes identified by Sturm et al. (1995). The gray lines show the Sturm result and the colored lines
815 show the current result. The vertical lines show the mean error, or the model bias, for both the Sturm and the current
816 result.
817

818

819   Figure 12: Results from application of the multi-variable, two-equation model to numerous northeast USA datasets.
820   The left column shows the SWE-*h* data for each dataset. Note that the black symbols are points removed by the
821   outlier detection procedure discussed in section 2.1.1.4. The remaining symbols are colored by *DOY*. The middle
822   panel plots heat maps of the model estimates of SWE against the observations of SWE with the 1:1 line included.
823   Warmer colors indicate higher densities. The right panel shows probability density functions of the model residuals,
824   with the vertical line indicating the mean error, or bias. Individual rows correspond to individual data sets and are
825   labeled.

826

827

828

829

830

831
832
833

834 Figure 13: Results from application of the multi-variable, two-equation model to the NRCS snow course / aerial
835 marker dataset. The left column shows the SWE-*h* data for each dataset. Note that the black symbols are points
836 removed by the outlier detection procedure discussed in section 2.1.1.5. The remaining symbols are colored by
837 *DOY*. The middle panel plots heat maps of the model estimates of SWE against the observations of SWE with the
838 1:1 line included. Warmer colors indicate higher densities. The right panel shows probability the density function of
839 the model residuals, with the vertical line indicating the mean error, or bias.
840

841
842