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 and climatological (30-year normal) values for mean annual precipitation (*MAP*) and mean February temperature ($\bar{T}_{F_{mean}}$), 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, which introduces 'day of water year' (*DOY*) as an additional variable. The model is validated against a large database of snow pillow measurements and yields a bias in SWE of less than 0.5 mm and a root-mean-squared-error (RMSE) in SWE of approximately 65 mm. When the errors are investigated on a station-by-station basis, the average RMSE is about 5% of the *MAP* at each station. The model is additionally validated against a completely independent set of data from the northeast United States. Finally, the results are compared with 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 set.

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 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 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).

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 $\pm$ 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,

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

72   to coring measurements at the same location.

73

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

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

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

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

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

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

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

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

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

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

84   interval).

85

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

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

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

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

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

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

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

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

94

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

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

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

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

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

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

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

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

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

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

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

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

107   regions).

108

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

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

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

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

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

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

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

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

117 (1995) were able to develop a global map of snow classification.

118

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

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

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

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

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

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

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

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

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

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

129 many meteorological variables as input.

130

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

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

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

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

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

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

137

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

139 (2010). Specifically, our goal is to regress physical and environmental variables directly into the equations. In this

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

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

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

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

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

145

146   Our approach is most similar to Mizukami and Perica (2008), Jonas et al. (2009), and Sturm et al., (2010) in that a

147   minimum of information is needed for the calculations; we intentionally avoid approaches like Meloysund et al.

148   (2007) and McCreight and Small (2014). This is because our interests are in converting h measurements to SWE

149   estimates in areas lacking weather instrumentation. The following sections introduce the numerous data sets that

150   were used in this study, outline the regression model adopted, and assess the performance of the model.

151   **2 Methods**

152

153   **2.1 Data**

154

155   **2.1.1 Snow Depth and Snow Water Equivalent**

156   In this section, we list sources of 1970-present snow data utilized for this study (Table 1). With regards to snow

157   coring devices, we refer to them using the terminology preferred in the references describing the datasets.

158

159   **2.1.1.1 USA NRCS Snow Telemetry and Soil Climate Analysis Networks**

160   SNOTEL (Serreze et al., 1999; Dressler et al., 2006) and SCAN (Schaefer et al. 2007) stations in the contiguous

161   United States (CONUS) and Alaska typically record sub-daily observations of h, SWE, and a variety of weather

162   variables (Figure 2a-b). The periods of record are variable, but the vast majority of stations have a period of record

163   in excess of 30 years. For this study, data from all SNOTEL sites in CONUS and Alaska and northeast USA SCAN

164   sites were obtained with the exception of sites whose period of record data were unavailable online. Only stations

165   with both SWE and h data were retained.

166

167   **2.1.1.2 Canada (British Columbia) Snow Survey Data**

168   Goodison et al. (1987) note that Canada has no national digital archive of snow observations from the many

169   independent agencies that collect snow data and that snow data are instead managed provincially. The quantity and

170   availability of the data vary considerably among the provinces. The Water Management Branch of the British

171   Columbia (BC) Ministry of the Environment manages a comparatively dense network of Automated Snow Weather

172   Stations (ASWS) that measure SWE, h, accumulated precipitation, and other weather variables (Figure 2a). For this

173   study, data from all British Columbia ASWS sites were initially obtained. As with the NRCS stations, only ASWS

174   stations with both SWE and h data were retained.

175

176   **2.1.1.3 Northeast USA Data**

177   In addition to the data from the SCAN sites, snow data for this project from the northeast US come from two

178   networks and three research sites (Figure 2b). The Maine Cooperative Snow Survey (MCSS, 2018) network

179   includes h and SWE data collected by the Maine Geological Survey, the United States Geological Survey, and

180   numerous private contributors and contractors. MCSS snow data are collected using the Standard Federal or

181   Adirondack snow sampling tubes typically on a weekly to bi-weekly schedule throughout the winter and spring,

182   1951-present. The New York Snow Survey network data were obtained from the National Oceanic and Atmospheric

183   Administration's Northeast Regional Climate Center at Cornell University (NYSS, 2018). Similar to the MCSS,

184   NYSS data are collected using Standard Federal or Adirondack snow sampling tubes on weekly to bi-weekly

185   schedules, 1938-present.

186

187   The Sleepers River, Vermont Research Watershed in Danville, Vermont (Shanley and Chalmers, 1999) is a USGS

188   site that includes 15 stations with long-term weekly records of h and SWE collected using Adirondack snow tubes.

189   Most of the periods of record are 1981-present, with a few stations going back to the 1960s. The sites include

190   topographically flat openings in conifer stands, old fields with shrub and grass, a hayfield, a pasture, and openings in

191   mixed softwood-hardwood forests. The Hubbard Brook Experiment Forest (Campbell et al., 2010) has collected

192   weekly snow observations at the Station 2 rain gauge site, 1959-present. Measurement protocol collects ten samples

193   2 m apart along a 20 m transect in a hardwood forest opening about ¼ hectare in size. At each sample location along

194   the transect, h and SWE are measured using a Mt. Rose snow tube and the ten samples are averaged for each

195   transect. Finally, the Thompson Farm Research site includes a mixed hardwood forest site and an open pasture site

196   (Burakowski et al. 2013; Burakowski et al. 2015). Daily (from 2011-2018), at each site, a snow core is extracted

197   with an aluminum tube and weighed (tube + snow) using a digital hanging scale. The net weight of the snow is

198   combined with the depth and the tube diameter to determine ρ_b, similar to a Federal or Adirondack sampler.

199

200   **2.1.1.4 Chugach Mountains (Alaska) Data**

201   In the spring of 2018, we conducted three weeks of fieldwork in the Chugach mountains in coastal Alaska, near the

202   city of Valdez (Figure 2c-d). We measured h using an avalanche probe at 71 sites along elevational transects during

203   March, April, and May. The elevational transects ranged between 250 and 1100 m (net change along transect) and

204   were accessible by ski and snowshoe travel. At each site, we measured h in 8 locations within the surrounding 10

205   m^2, resulting in a total of 550+ snow depth measurements. These 71 sites were scattered across 8 regions in order to

206   capture spatial gradients that exist in the Chugach mountains as the wetter, more-dense maritime snow near the coast

207   gradually changes to drier, less dense snow on the interior side.

208

209   **2.1.1.5 Data Pre-Processing**

210   Figure 3 demonstrates that it is not uncommon for automated snow depth measurements to become noisy or non-

211   physical, at times reporting large depths when there is no SWE reported. This is different from instances when

212   physically plausible, but very low densities might be reported; say in response to early season dry, light snowfalls. It

213   was therefore desirable to apply some objective, uniform procedure to each station's dataset in order to remove clear

214   outlier points, while minimizing the removal of valid data points. We recognize that there is no accepted

215   standardized method for cleaning bivariate SWE-h data sets. While Serreze et al. (1999) offer a procedure for

216   SNOTEL data in their appendix, it is relevant only for precipitation and SWE values, not h. Given the strong

217   correlation between h and SWE, we instead choose to use common outlier detection techniques for bivariate data.

218

219   The Mahalanobis distance (MD; Maesschalck et al., 2000) quantifies how far a point lies from the mean of a

220   bivariate distribution. The distances are in terms of the number of standard deviations along the respective principal

221   component axes of the distribution. For highly correlated bivariate data, the MD can be qualitatively thought of as a

222   measure of how far a given point deviates from an ellipse enclosing the bulk of the data. One problem is that the MD

223   is based on the statistical properties of the bivariate data (mean, covariance) and these properties can be adversely

224   affected by outlier values. Therefore, it has been suggested (e.g., Leys et al., 2018) that a 'robust' MD (RMD) be

225   calculated. The RMD is essentially the MD calculated based on statistical properties of the distribution unaffected

226   by the outliers. This can be done using the Minimum Covariance Determinant (MCD) method as first introduced by

227   Rousseeuw (1984).

228

229   Once RMDs have been calculated for a bivariate data set, there is the question of how large an RMD must be in

230   order for the data point to be considered an outlier. For bivariate normal data, the distribution of the square of the

231   RMD is χ^2 (Gnanadesikan and Kettenring, 1972), with p (the dimension of the dataset) degrees of freedom. So, a

232   rule for identifying outliers could be implemented by selecting as a threshold some arbitrary quantile (say 0.99) of

233   χ_p^2. For the current study, a threshold quantile of 0.999 was determined to be an appropriate compromise in terms of

234   removing obviously outlier points, yet retaining physically plausible results.

235

236   A scatter plot of SWE vs. *h* for the source SNOTEL dataset from CONUS and AK reveals many non-physical

237   points, mostly when a very large *h* is reported for a very low SWE (Figure 4a). Approximately 0.7% of the original

238   data points were removed in the pre-processing described above, creating a more physically plausible scatter plot

239   (Figure 4b). Note that the outlier detection process was applied to each station individually. The distribution of 'day

240   of year' (*DOY*) values of removed data points was broad, with a mean of 160 and a standard deviation of 65. Note

241   that the *DOY* origin is 1 October. The same procedure was applied to the BC and northeast USA data sets as well

242   (not shown). Table 1 summarizes useful information about the numerous data sets described above and indicates the

243   final number of data points retained for each. We acknowledge that our process inevitably removes some valid data

244   points, but, as a small percentage of an already 0.7% removal rate, we judged this to be acceptable.

245

246   Table 1: Summary of information about the datasets used in this study. Datasets in bold font were used to construct
247   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)
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)

248

2.1.2 Climatological Variables

250 30-year climate normals at 800 m (nominal) resolution for CONUS and for the period 1981-2010 were obtained

251 from the PRISM website (Daly et al., 1994). PRISM normals for British Columbia (BC), Canada, were obtained

252 from the ClimateBC project (Wang et al., 2012), also for the 1981-2010 period. Finally, PRISM normals for Alaska

253 (AK) were obtained from the Integrated Resource Management Applications (IRMA) Portal run by the National

254 Park Service. The AK normals are for the 1971-2000 period and have a slightly coarser resolution (approximately

255 1.5 km). Figure 5 shows gridded maps of mean annual precipitation (MAP) and mean February Temperature ($\overline{T}_F$)

256 for these three climate products, plotted together. Other temperature products (max and min temperatures; other

257 months) were obtained as well, but are not shown.

258

2.2 Regression Model

260 In order to demonstrate the varying degrees of influence of explanatory variables, several regression models were

261 constructed. In each case, the model was built by randomly selecting 50% of the paired SWE-h measurements from

262 the aggregated CONUS, AK, and BC snow pillow datasets. The model was then validated by applying it to the

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

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

265 were completely left out of the model building process.

266

2.2.1 One-Equation Model

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

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

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

271

272 (1) $\quad SWE = Ah^{a_1}$.

273

274 This equation can be log-transformed into

275

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

277

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

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

280   will be deferred until the results section.

281

282   **2.2.2 Two-Equation Model**

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

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

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

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

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

288

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

290

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

292

293   where DOY is the number of days from the start of the water-year (October 1 is the origin), and DOY^* is the critical

294   or dividing day-of-water-year separating the two phases. Put another way, DOY^* is the day of peak SWE.

295   Interannual variability results in a range of DOY^* for a given site. Additionally, some sites, particularly the SCAN

296   sites in the northeast USA, demonstrate multi-peak SWE profiles in some years. To reduce model complexity,

297   however, we investigated the use of a simple climatological (long term average) value of DOY^*. For each snow

298   pillow station, the average DOY^* was computed over the period of record of that station. Analysis of all of the

299   stations revealed that this average DOY^* was relatively well correlated with the climatological mean April maximum

300   temperature (the average of the daily maximums recorded in April; $R^2 = 0.7$). However, subsequent regression

301   analysis demonstrated that the SWE estimates were relatively insensitive to DOY^* and the best results were actually

302   obtained when DOY^* was uniformly set to 180 for all stations. Again, with both SWE and h in units of mm, the

303   regression coefficients turn out to be $(A, a_1) = (0.150, 1.082)$ and $(B, b_1) = (0.239, 1.069)$.

304

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

306   equation model

307

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

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

310

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

312   the root mean square error of the model estimates.

313

314   **2.2.3 Two-Equation Model with Climate Parameters**

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

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

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

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

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

320   variables included was

321

322   (6) $SWE = f\left(h, z, MAP, \bar{T}_{J_{min}}, \bar{T}_{J_{mean}}, \bar{T}_{J_{max}}, \bar{T}_{F_{min}}, \bar{T}_{F_{mean}}, \bar{T}_{F_{max}}, \bar{T}_{M_{min}}, \bar{T}_{M_{mean}}, \bar{T}_{M_{max}}, \bar{T}_{A_{min}}, \bar{T}_{A_{mean}}, \bar{T}_{A_{max}}\right)$

323

324   where z is the elevation (m), *MAP* is the mean annual precipitation (mm) and the temperatures (°C) represent the

325   mean of minimum, mean, and maximum daily values for the months January through April (J, F, M, A). For

326   example, $\bar{T}_{J_{min}}$ is the climatological normal of the average of the daily minimum temperatures observed in January.

327   In the stepwise regression, explanatory variables were accepted if they improved the adjusted R^2 value by 0.001.

328   The result of the regression yielded

329

330   (7) $SWE_{acc} = Ah^{a_1}MAP^{a_2}\left(\bar{T}_{F_{mean}} + 30\right)^{a_3}; \quad DOY < DOY^*$

331

332   (8) $SWE_{abl} = Bh^{b_1}MAP^{b_2}\left(\bar{T}_{F_{mean}} + 30\right)^{b_3}; \quad DOY \geq DOY^*$

333

334   or, in log-transformed format,

335

336   (9) $log_{10}(SWE_{acc}) = log_{10}(A) + a_1 log_{10}(h) +$

337    $a_2 log_{10}(MAP) + a_3 log_{10}\left(\bar{T}_{F_{mean}} + 30\right); \quad DOY < DOY^*$

338

339   (10) $log_{10}(SWE_{abl}) = log_{10}(B) + b_1 log_{10}(h) +$

340    $b_2 log_{10}(MAP) + b_3 log_{10}\left(\bar{T}_{F_{mean}} + 30\right); \quad DOY \geq DOY^*$

341

342   indicating that only snow depth, mean annual precipitation and mean February temperature were relevant. Manual

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

344   above equations, note that an offset is added to the temperature in order to avoid taking the log of a negative

345   number. These two SWE estimates for the individual (*acc* and *abl*) phases of the snowpack are then blended with

346   Eqn. (5) to produce a single equation for SWE spanning the entire water year. The obtained regression coefficients

347   were $(A, a_1, a_2, a_3) = (0.0128, 1.070, 0.132, 0.506)$ and $(B, b_1, b_2, b_3) = (0.0271, 1.038, 0.201, 0.310)$. The

348   physical interpretation of these coefficients is straightforward. If a_1 and b_1 were equal to unity, then the density,

349   given by (SWE/h), would be a constant at a given location. Since they are greater than unity, they capture the effect

350   that snow density increases as depth increases. Turning to the coefficients on the climate variables, both a_2 and b_2

351   are greater than zero. So, for two locations with equal depth, equal temperature characteristics, but different

352   precipitation characteristics, the regression model predicts that the wetter location (larger *MAP*) will have a greater

353   density. Finally, regarding temperature, both a_3 and b_3 are greater than zero. Therefore, for two locations with equal

354   depth, equal precipitation characteristics, but different temperature characteristics, the regression model predicts that

355   the warmer location (larger $\bar{T}_{F_{mean}}$) will have a greater density. These trends are similar in concept to Sturm et al.

356   (2010), whose snow classes (based on climate classes) indicate which snow will densify more rapidly.

357   **3 Results**

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

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

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

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

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

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

364   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

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

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

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

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

369

370   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	0.972	-0.5	67

371

372   It is useful to also consider the model errors in a non-dimensional way. Therefore, an RMSE was computed at each

373   station location and normalized by the mean annual maximum SWE (SWE_{max}) at that location. Figure 8 shows the

374   probability density function of these normalized errors. The average RMSE is approximately 11% of SWE_{max}, with

375   most falling into the range of 5-25%. The spatial distribution of these normalized errors is shown in Figure 9. For

376   the SNOTEL stations, it appears there is a slight regional trend, in terms of stations in continental climates (northern

377   Rockies) having smaller relative errors than stations in maritime climates (Cascades). The British Columbia stations

378   also show higher relative errors.

379

380 **3.1 Results for Snow Classes**

381 A key objective of this study is to regress climatological information in a continuous rather than a discrete way. The

382 work by Sturm et al. (2010) therefore provides a valuable point of comparison. In that study, the authors developed

383 the following equation for density ρ_b

384

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

386

387 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,

388 and DOY in this case begins on January 1. This means that their DOY for October 1 is -92. The coefficients vary

389 with snow class and the values determined by Sturm et al. (2010) are shown in Table 3.

390

391 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

392

393 To make a comparison, the snow class for each SNOTEL and British Columbia snow survey (Rows 1 and 3 of Table

394 1) site was determined using a 1-km snow class grid (Sturm et al., 2010). The aggregated dataset from these stations

395 was made up of 27% Alpine, 14% Maritime, 10% Prairie, 11% Tundra, and 38% Taiga data points. Equation (11)

396 was then used to estimate snow density (and then SWE) for every point in the validation dataset described in Section

397 2.2. Figure 10 compares the SWE estimates from the Sturm model and from the present multi-variable, two-equation

398 model (Equations 5, 7-8). The upper left panel of Figure 10 shows all of the data, and the remaining panels show the

399 results for each snow class. In all cases, the current model provides better estimates. Plots of the residuals by snow

400 class are provided in Figure 11, giving an indication of the bias of each model for each snow class. Summaries of the

401 model performance, broken out by snow class, are given in Table 4.

402

403 Table 4: Comparison of model performance by Sturm et al. (2010) and the present 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	0.972	-0.5	67
Alpine	0.973	10.1	55	0.971	-0.3	55
Maritime	0.968	-16.8	109	0.970	-4.5	105
Prairie	0.967	18.7	56	0.965	-0.2	51
Tundra	0.956	-10.5	82	0.969	-6.1	67
Taiga	0.943	-80.0	151	0.971	2.4	62

404

405 **3.2 Comparison to Pistocchi (2016)**

406 In order to provide one additional comparison, the simple model of Pistocchi (2016) was also applied to the
407 validation dataset. His model calculates the bulk density as

408

409 (12) $\rho_b = \rho_0 + K(DOY + 61),$

410

411 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
412 November 1. Application of this model to the validation dataset yields a bias of 55 mm and an RMSE of 94 mm.
413 These results are comparable to the Sturm et al. (2010) model, with a larger bias but smaller RMSE.

414
415 **3.3 Results for Northeast USA**
416 The regression equations in this study were developed using a large collection of SNOTEL sites in CONUS, AK,
417 and BC. The snow pillow sites are limited to locations west of approximately W 105° (Figure 2a). By design, the
418 data sets from the northeastern USA (Section 2.1.1.3) were left as an entirely independent validation set. These
419 northeastern sites are geographically distant from the training data sets, are subject to a very different climate, and
420 are generally at much lower elevations than the western sites, providing an interesting opportunity to test how robust
421 the present model is.

422

423 Figure 12 graphically summarizes the datasets and the performance of the multi-variable two-equation model of the
424 current study. The RMSE values are comparable to those found for the western stations, but, given the
425 comparatively thinner snowpacks in the northeast, represent a larger relative error (Table 5). The bias of the model
426 is consistently positive, in contrast to the western stations where the bias was negligible.

427

428 Table 5: Performance metrics for the multi-variable two-equation model applied to various northeastern USA
429 datasets.

Dataset Name	R^2	Bias (mm)	RMSE (mm)
Maine Geological Survey, ME	0.91	8.9	33.3
Hubbard Brook (Station 2), NH	0.63	18.9	64.2
Thompson Farm, NH	0.85	7.1	21.6
NRCS SCAN	0.87	-1.8	38.7
Sleepers River, VT	0.93	14.0	29.7
New York Snow Survey	0.93	13.8	31.2

430

431 **4 Discussion**
432 The results presented in this study show that the regression equation described by equations (5, 7-8) is an
433 improvement (lower bias and RMSE) over other widely used bulk density equations. The key advantage is that the
434 present method regresses in relevant physical parameters directly, rather than using discrete bins (for snow class,
435 elevation, month of year, etc.), each with its own set of model coefficients. The comparison (Figs. 10-11; Table 4) to
436 the model of Sturm et al. (2010) reveals a peculiar behavior of that model for the Taiga snow class, with a large

437 negative bias in the Sturm estimates. Inspection of the coefficients provided for that class (Table 3) shows that the

438 model simply predicts that $\rho_b = \rho_{max} = 0.217$ for all conditions.

439

440 When our multi-variable two-equation model, developed solely from western North American data, is applied to

441 northeast USA locations, it produces SWE estimates with smaller RSME values and larger biases than the western

442 stations. When comparing the SWE-h curves of the SNOTEL data (Figure 4b) to those of the east coast data sets

443 (left column; Figure 12), it is clear that the northeast data generally have more scatter. This is confirmed by

444 computing the correlation coefficients between SWE and h for each dataset. It is unclear if this disparity in

445 correlation is related to measurement methodology or is instead a 'signal to noise' issue. Comparing Figures 4 and

446 12 shows the considerable difference in snowpack depth between the western and northeastern data sets. When the

447 western dataset is filtered to include only measurement pairs where $h < 1.5$ m, the correlation coefficient is reduced

448 to a value consistent with the northeast datasets. This suggests that the performance of the current (or other)

449 regression model is not as good at shallow snowpack depths. This is also suggested upon examination of the time

450 series of observed $\rho_b = SWE/h$ for a given season at a snow pillow site. Very early in the season, when the depths

451 are small, the density curve has a lot of variability. Later in the season, when depths are greater, the density curve

452 becomes much smoother. Very late in the season, when depths are low again, the density curve becomes highly

453 variable again.

454

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

470 estimates.

471

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

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

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

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

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

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

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

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

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

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

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

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

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

485 (16%) uncertainty in the SWE estimates. CSO participants can do three things to ensure that their recorded depth

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

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

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

489

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

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

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

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

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

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

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

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

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

499 **5 Conclusions**

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

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

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

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

504 easily and immediately converted into SWE.

505

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

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

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

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

510

This model is not a replacement for more sophisticated snow models that evolve the snowpack based on high frequency (e.g., daily or sub-daily) weather data inputs. The intended purpose of this model is to constrain SWE estimates in circumstances where snow depth is known, but weather variables are not, a common issue in sparsely instrumented areas in North America.

7 Data Access

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

1. NRCS Snow Telemetry: https://www.wcc.nrcs.usda.gov/snow/SNOTEL-wedata.html
2. NRCS Soil Climate Analysis Network: https://www.wcc.nrcs.usda.gov/scan/
3. British Columbia Automated Snow Weather Stations: https://www2.gov.bc.ca/gov/content/environment/air-land-water/water/water-science-data/water-data-tools/snow-survey-data/automated-snow-weather-station-data
4. Maine Cooperative Snow Survey: https://mgs-maine.opendata.arcgis.com/datasets/maine-snow-survey-data
5. New York Snow Survey: http://www.nrcc.cornell.edu/regional/snowsurvey/snowsurvey.html
6. Sleepers River Research Watershed. Snow data not available online; request data from contact at: https://nh.water.usgs.gov/project/sleepers/index.htm
7. Hubbard Brook Experimental Forest: https://hubbardbrook.org/d/hubbard-brook-data-catalog
8. CONUS PRISM Data: http://www.prism.oregonstate.edu/
9. British Columbia PRISM Data: http://climatebcdata.climatewna.com/
10. Alaska PRISM Data: https://irma.nps.gov/Portal/

A Matlab function for calculating SWE based on the results is this paper has been made publicly available at Github (https://github.com/communitysnowobs/snowdensity).

References

Alford, D.: Density variations in alpine snow, J. Glaciol., 6(46), 495-503, https://doi.org/10.3189/S0022143000019717, 1967.

Avanzi, F., De Michele, C., and Ghezzi, A.: On the performances of empirical regressions for the estimation of bulk snow density, Geogr. Fis. Dinam. Quat., 38, 105-112, doi:10.4461/GFDQ.2015.38.10, 2015.

Beaumont, R.: Mt. Hood pressure pillow snow gage, J. Appl. Meteorol., 4, 626-631, https://doi.org/10.1175/1520-0450(1965)004<0626:MHPPSG>2.0.CO;2 1965.

Beaumont, R., and Work, R.: Snow sampling results from three samplers, Hydrol. Sci. J., 8(4), 74-78, https://doi.org/10.1080/02626666309493359, 1963.

Burakowski, E.A., Ollinger, S., Lepine, L., Schaaf, C.B., Wang, Z., Dibb, J.E., Hollinger, D.Y., Kim, J.-H., Erb, A., and Martin, M.E.: Spatial scaling of reflectance and surface albedo over a mixed-use, temperate forest landscape during snow-covered periods, Remote Sens. Environ., 158, 465-477, https://doi.org/10.1016/j.rse.2014.11.023, 2015.

Burakowski, E.A., Wake, C.P., Stampone, M., and Dibb, J.: Putting the Capital 'A' in CoCoRAHS: An Experimental Program to Measure Albedo using the Community Collaborative Rain Hail and Snow (CoCoRaHS) Network, Hydrol. Process., 27(21), 3024-3034, https://doi.org/10.1002/hyp.9825, 2013.

Campbell, J., Ollinger, S., Flerchinger, G., Wicklein, H., Hayhoe, K., and Bailey, A.: Past and projected future changes in snowpack and soil frost at the Hubbard Brook Experimental Forest, New Hampshire, USA, Hydrol. Process., 24, 2465-2480, https://doi.org/10.1002/hyp.7666, 2010.

Church, J.E.: Snow surveying: its principles and possibilities, Geogr. Rev., 23(4), 529-563, DOI: 10.2307/209242, 1933.

Church, J.E., and Marr, J.C.: Further improvement of snow-survey apparatus, Transactions of the American Geophysical Union, 18(2), 607-617, 10.1029/TR018i002p00607, 1937.

Daly, C., Neilson, R., and Phillips, D.: A statistical-topographic model for mapping climatological precipitation over mountainous terrain, J. Appl. Meteorol., 33, 140-158, https://doi.org/10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2, 1994.

574 Deeb, E., Marshall, H.-P., Forster, R., Jones, C., Hiemstra, C., and Siqueira, P.: Supporting NASA SnowEx remote

575 sensing strategies and requirements for L-band interferometric snow depth and snow water equivalent estimation,

576 Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium, Fort Worth, TX, 1395-

577 1396, DOI: 10.1109/IGARSS.2017.8127224, 2017.

578

579 Dixon, D., and Boon, S.: Comparison of the SnowHydro sampler with existing snow tube designs, Hydrol. Process.,

580 26(17), 2555-2562, https://doi.org/10.1002/hyp.9317, 2012.

581

582 Dressler, K., Fassnacht, S., and Bales, R.: A comparison of snow telemetry and snow course measurements in the

583 Colorado River basin, J. Hydrometeorol., 7, 705-712, https://doi.org/10.1175/JHM506.1, 2006.

584

585 Fick, S., and Hijmans, R.: WorldClim2: new 1-km spatial resolution climate surfaces for global land areas, Int. J.

586 Climatol., 37, 4302-4315, https://doi.org/10.1002/joc.5086, 2017.

587

588 Goodison, B.: Accuracy of snow samplers for measuring shallow snowpacks: An update, Proceedings of the 35th

589 Annual Eastern Snow Conference, Hanover, NH, 36-49, 1978.

590

591 Goodison, B., Ferguson, H., and McKay, G.: Measurement and data analysis. The Handbook of Snow: Principles,

592 Processes, Management, and Use, D. Gray and D. Male, Eds., Pergamon Press, 191-274, 1981.

593

594 Goodison, B., Wilson, B., Wu., K, and Metcalfe, J.: An inexpensive remote snow-depth gauge: An assessment,

595 Proceedings of the 52nd Annual Western Snow Conference, Sun Valley, ID, 188-191, 1984.

596

597 Goodison, B., Glynn, J., Harvey, K., and Slater, J.: Snow Surveying in Canada: A Perspective, Can. Water Resour.

598 J., 12(2), 27-42, https://doi.org/10.4296/cwrj1202027, 1987.

599

600 Gnanadesikan, R., and Kettenring, J.: Robust estimates, residuals, and outlier detection with multiresponse data,

601 Biometrics, 28, 81-124, DOI: 10.2307/2528963, 1972.

602

603 Hill, D.F., Wolken, G. J., Wikstrom Jones, K., Crumley, R., and Arendt, A.: Crowdsourcing snow depth data with

604 citizen scientists, Eos, 99, https://doi.org/10.1029/2018EO108991, 2018.

605

606 Johnson, J., and Marks, D.: The detection and correction of snow water equivalent pressure sensor errors, Hydrol.

607 Proc., https://doi.org/10.1002/hyp.5795, 2004.

608

609 Johnson, J., Gelvin, A., Duvoy, P., Schaefer G., Poole, G., and Horton, G.: Performance characteristics of a new

610 electronic snow water equivalent sensor in different climates, Hydrol. Proc., DOI: 10.1002/hyp.10211, 2015.

611

612   Jonas, T., Marty, C., and Magnusson, M.: Estimating the snow water equivalent from snow depth measurements, J.

613   Hydrol., 378, 161-167, https://doi.org/10.1016/j.jhydrol.2009.09.021, 2009.

614

615   Kang, D.H., and Barros, A.P.: Full-system testing in laboratory conditions of an L-band snow sensor system for in

616   situ monitoring of snow-water content, IEEE Trans. Geosci. Remote Sens., 49(3), 908-919,

617   10.1109/TGRS.2010.2072786, 2011.

618

619   Leinss, S., Wiesmann, A., Lemmetyinen, J., and Hajnsek, I.: Snow water equivalent measured by differential

620   interferometry, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 8(8), 3773-3790, 10.1109/JSTARS.2015.2432031,

621   2015.

622

623   Leys, C., Klein, O., Dominicy, Y., and Ley, C.: Detecting multivariate outliers: use a robust variant of the

624   Mahalanobis distance, J. Exp. Soc. Psychol., 74, 150-156, https://doi.org/10.1016/j.jesp.2017.09.011, 2018.

625

626   Liang, X., Lettermaier, D., Wood, E., and Burges, S.: A simple hydrologically based model of land surface water

627   and energy fluxes for general circulation models, J. Geophys. Res. Atmos., 99(D7), 14,415-14,428,

628   https://doi.org/10.1029/94JD00483, 1994.

629

630   Liston, G., and Elder, K.: A distributed snow evolution modeling system (SnowModel), J. Hydrometerol., 7, 1259-

631   1276, https://doi.org/10.1175/JHM548.1, 2006.

632

633   Lundberg, A., Richardson-Naslund, C., and Andersson, C.: Snow density variations: consequences for ground

634   penetrating radar, Hydrol. Process., 20, 1483-1495, https://doi.org/10.1002/hyp.5944, 2006.

635

636   Maine Geological Survey: Maine Cooperative Snow Survey Dataset,

637   https://www.maine.gov/dacf/mgs/hazards/snow_survey/, 2018.

638

639   McKay, G., and Findlay, B., 1971: Variation of snow resources with climate and vegetation in Canada, Proceedings

640   of the 39th Western Snow Conference, Billings, MT, 17-26, 1971.

641

642   De Maesschalck, R., Jouan-Rimbaud, D., and Massart, D.: The Mahalanobis distance, Chemometr. Intell. Lab. Syst.,

643   50(1), 1-18, https://doi.org/10.1016/S0169-7439(99)00047-7, 2000.

644

645   McCreight, J., and Small, E.: Modeling bulk density and snow water equivalent using daily snow depth

646   observations, The Cryosphere, 8, 521-536, https://doi.org/10.5194/tc-8-521-2014, 2014.

647

648 Meloysund, V., Leira, B., Hoiseth, K., and Liso, K.: Predicting snow density using meterological data, Meteorol.

649 Appl., 14, 413-423, https://doi.org/10.1002/met.40, 2007.

650

651 Menne, M.J., I. Durre, R.S. Vose, B.E. Gleason, and Houston, T.G.: An overview

652 of the Global Historical Climatology Network-Daily Database, J. Atmos. Ocean. Technol., 29, 97-910,

653 doi:10.1175/JTECH-D-11-00103.1, 2012.

654

655 Molotch, N.P., and Bales, R.C.: SNOTEL representativeness in the Rio Grande headwaters on the basis of

656 physiographics and remotely sensed snow cover persistence, Hydrol. Process., 20(4), 723-739,

657 https://doi.org/10.1002/hyp.6128, 2006.

658

659 Mote, P., Li, S., Letternaier, D., Xiao, M., and Engel, R.: Dramatic declines in snowpack in the western US," npj

660 Clim. Atmos. Sci., 1(2), 1-6, doi:10.1038/s41612-018-0012-1, 2018.

661

662 Mizukami, N., and Perica, S.: Spatiotemporal characteristics of snowpack density in the mountainous regions of the

663 western United States, J. Hydrometeorol., 9, 1416-1426, https://doi.org/10.1175/2008JHM981.1, 2008.

664

665 New York Snow Survey, NOAA, Northeast Regional Climate Center at Cornell University, 2018.

666

667 Pagano, T., Garen, D., Perkins, T., and Pasteris, P.: Daily updating of operational statistical seasonal water supply

668 forecasts for the western U.S., J. Am. Water Resour. Assoc., 45(3), 767-778, https://doi.org/10.1111/j.1752-

669 1688.2009.00321.x, 2009.

670

671 Painter, T., Berisford, D., Boardman, J., Bormann, K., Deems, J., Gehrke, F., Hedrick, A., Joyce, M., Laidlaw, R.,

672 Marks, D., Mattmann, C., Mcgurk, B., Ramirez, P., Richardson, M., Skiles, S., Seidel, F., and Winstral, A.: The

673 Airborne Snow Observatory: fusion of scanning lidar, imaging spectrometer, and physically-based modeling for

674 mapping snow water equivalent and snow albedo, Remote Sens. Environ., 184, 139-152,

675 doi:10.1016/j.rse.2016.06.018, 2016.

676

677 Pistocchi, A.: Simple estimation of snow density in an Alpine region, J. Hydrol. Reg. Stud., 6, 82-89,

678 http://dx.doi.org/10.1016/j.ejrh.2016.03.004, 2016.

679

680 Rousseeuw, P.: Least Median of Squares Regression, J. Am. Stat. Assoc., 79, 871-880, DOI:

681 10.1080/01621459.1984.10477105, 1984.

682

683 Ryan, W., Doesken, N., and Fassnacht, S.: Evaluation of Ultrasonic Snow Depth Sensors for U.S. Snow

684 Measurements, J. Atmos. Ocean. Technol., 25, 667-684, https://doi.org/10.1175/2007JTECHA947.1, 2008.

685

686   Schaefer, G., Cosh, M., and Jackson, T.: The USDA Natural Resources Conservation Service Soil Climate Analysis
687   Network (SCAN), J. Atmos. Ocean. Technol., 24, 2073-2077, https://doi.org/10.1175/2007JTECHA930.1, 2007.
688

689   Serreze, M., Clark, M., Armstrong, R., McGinnis, D., and Pulwarty, R.: Characteristics of the western United States
690   snowpack from snowpack telemetry (SNOTEL) data, Water Resour. Res., 35(7), 2145-2160,
691   https://doi.org/10.1029/1999WR900090, 1999.
692

693   Shanley, J., and Chalmers, A.: The effect of frozen soil on snowmelt runoff at Sleepers River, Vermont, Hydrol.
694   Process., 13(12-13), 1843-1857, https://doi.org/10.1002/(SICI)1099-1085(199909)13:12/13<1843::AID-
695   HYP879>3.0.CO;2-G, 1999.
696

697   Sokol, J., Pultz, T., and Walker, A.: Passive and active airborne microwave remote sensing of snow cover," Int.
698   J. Remote. Sens., 24, 5327-5344, https://doi.org/10.1080/0143116031000115076, 2003.
699

700   Sturm, M., Holmgren, J., and Liston, G.: A seasonal snow cover classification system for local to global
701   applications, J. Clim., 8, 1261-1283, https://doi.org/10.1175/1520-0442(1995)008<1261:ASSCCS>2.0.CO;2, 1995.
702

703   Sturm, M., Taras, B., Liston, G.E., Derksen, C., Jonas, T., and Lea, J.: Estimating snow water equivalent using snow
704   depth data and climate classes, J. Hydrometeorol., 11, 1380-1394, https://doi.org/10.1175/2010JHM1202.1, 2010.
705

706   U.S. Army Corps of Engineers: Snow hydrology: Summary report of the snow investigations of the North Pacific
707   Division, 437pp., 1956.
708

709   Vuyovich, C., Jacobs, J., and Daly, S.: Comparison of passive microwave and modeled estimates of total watershed
710   SWE in the continental United States, Water Resour. Res., 50(11), 9088-9012,
711   https://doi.org/10.1002/2013WR014734, 2014.
712

713   Wang, T., Hamann, A., Spittlehouse, D.L., and Murdock, T.: ClimateWNA - High-Resolution Spatial Climate Data
714   for Western North America, J. Appl. Meteorol. Climatol., 51, 16-29, https://doi.org/10.1175/JAMC-D-11-043.1,
715   2012.
716

717   Wigmosta, M.S., Vail, L., and Lettenmaier, D.: A distributed hydrology-vegetation model for complex
718   terrain, Water Resour. Res., 30, 1665-1679, https://doi.org/10.1029/94WR00436, 1994

719   Figure 1: Conceptual sketch of the evolution of snow water equivalent (SWE) over the course of a water year (black

720   line). Also shown is the evolution of SWE with snowpack depth over a water year (red line). Note the hysteresis

721   loop due to the densification of the snowpack.
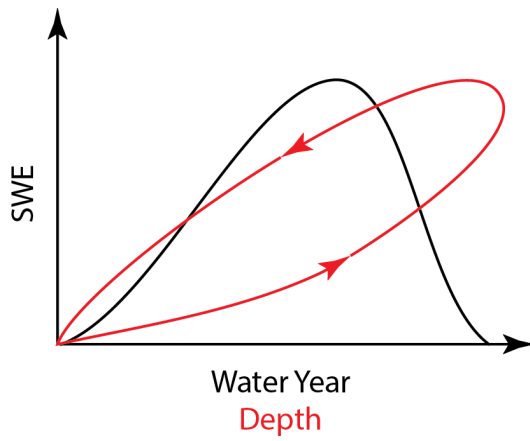
722

723 Figure 2: Distribution of measurement locations used in this study. (a) Western USA and Canada station locations,
724 with colors indicating station elevation in meters. (b) Northeast USA locations, with stations colored according to
725 data source. (c, d) Measurement sites in the Chugach Mountains, southcentral Alaska.
726
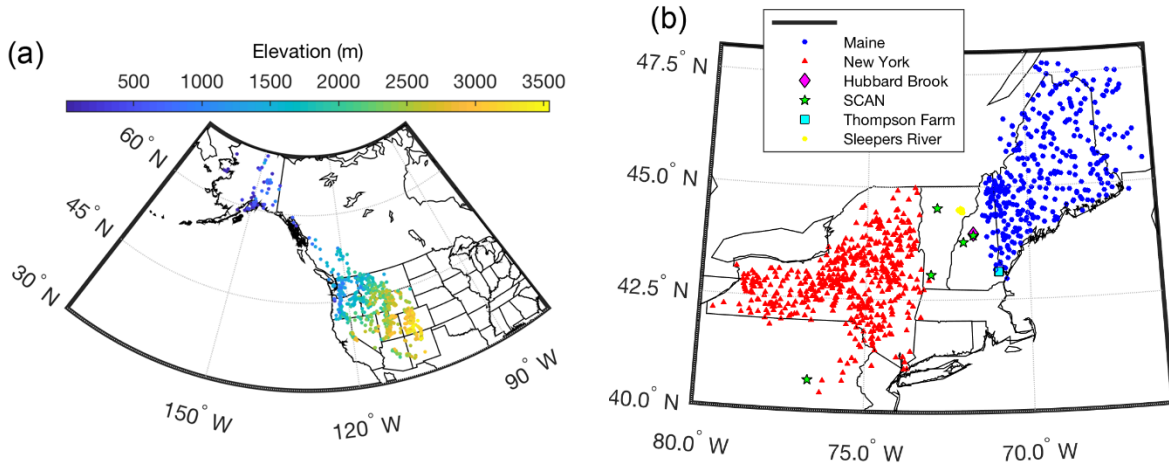
727
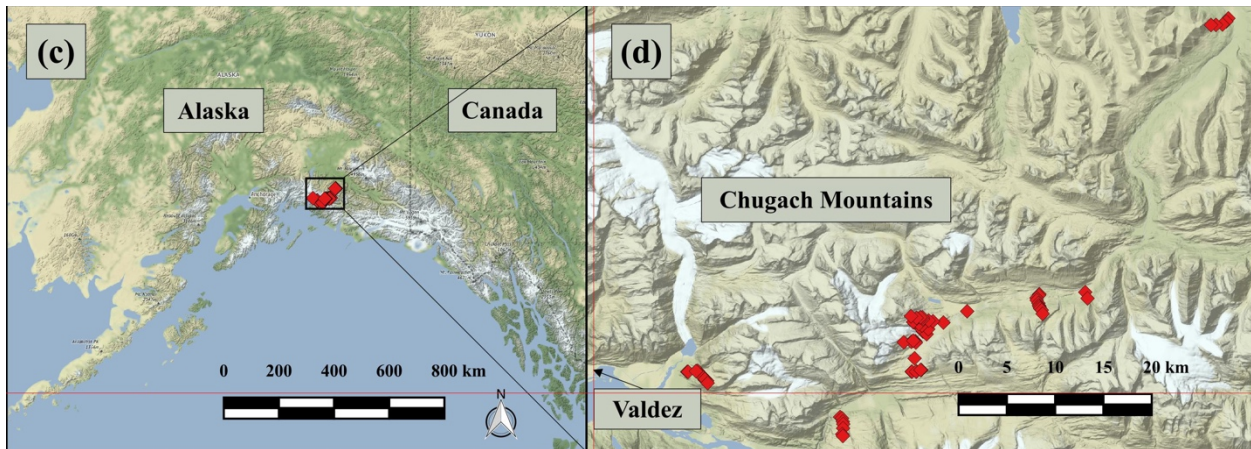
728

729

730   Figure 3: Sample time series of SWE and h from the Rex River (WA) SNOTEL station. Observations of h at times

731   when SWE is zero are likely spurious.
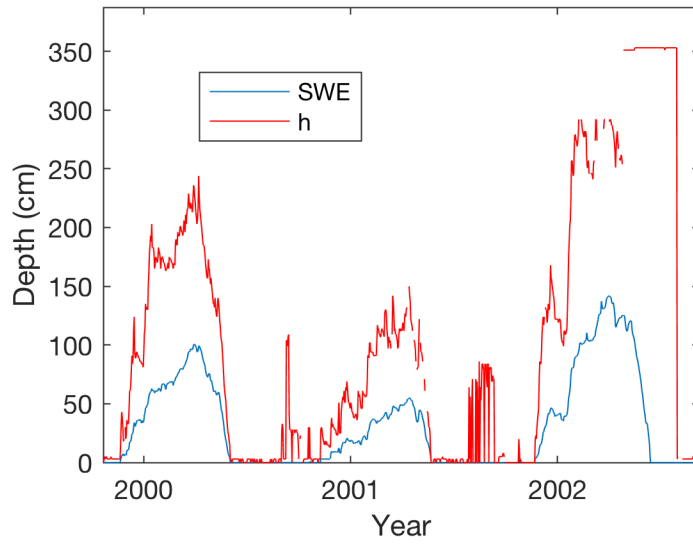
732

733   Figure 4: Scatter plot of SWE vs. *h* for the complete SNOTEL dataset before (a) and after (b) removing data points.

734   Symbols are colored by 'day of water year' (*DOY*; October 1 is the origin).

735
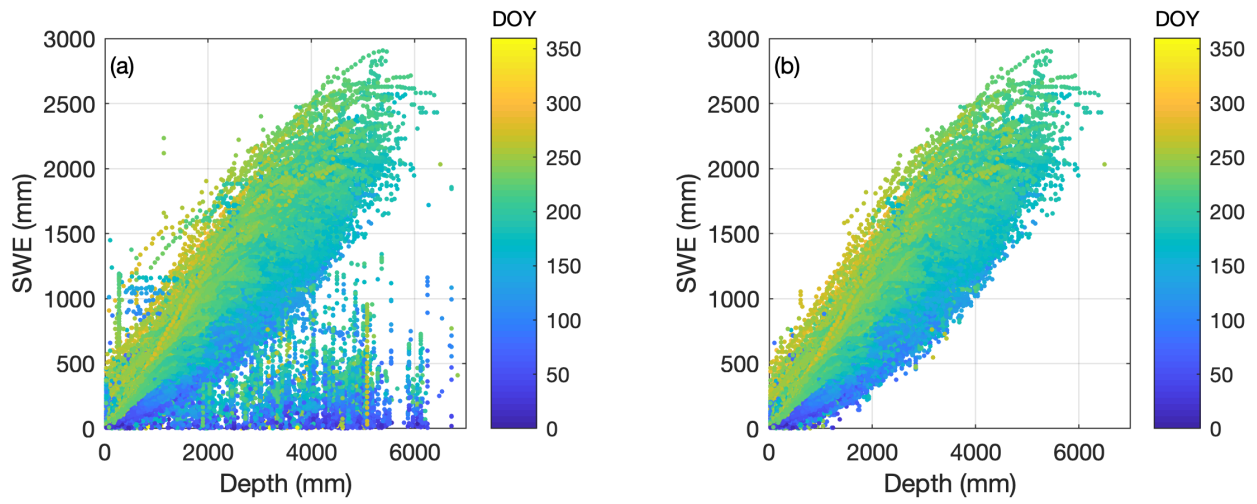
736

737   Figure 5: Gridded maps of mean annual precipitation (MAP) and mean February temperature $\left(\overline{T}_{F_{mean}}\right)$ for the study

738   regions. Climate normals are from the PRISM data set (1981-2010 for CONUS and British Columbia; 1971-2000

739   for Alaska).
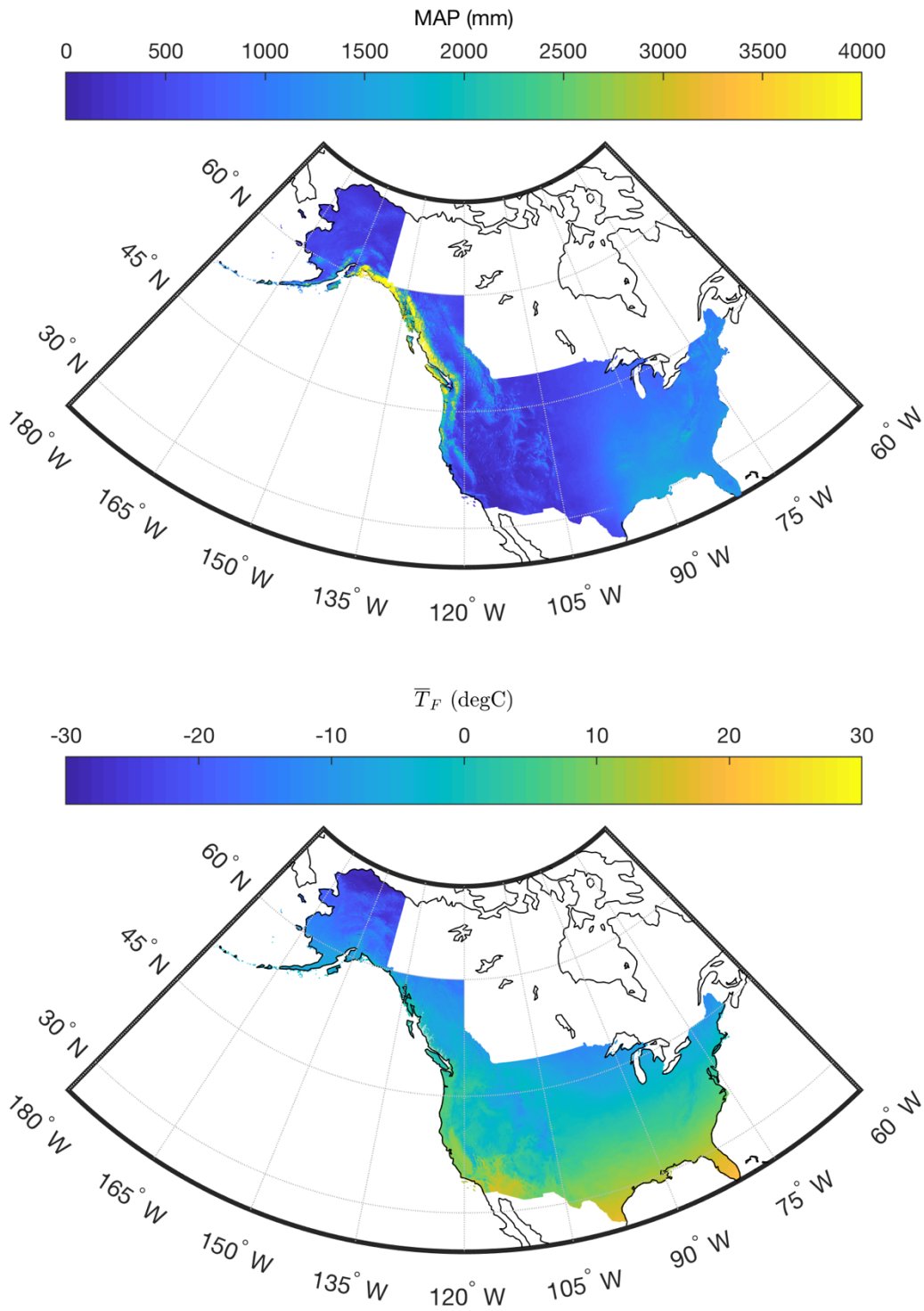
740

741
742

743 Figure 6: Two-dimensional histograms (heat maps; left column) of modeled vs. observed SWE and probability
744 density functions (right column) of the residuals for three simple models applied to the CONUS, AK, and BC snow
745 pillow data. Warmer colors in the heat maps indicate greater density. The vertical lines in the right column indicate
746 the location of the mean residual, or bias. Top row (a-b): One-equation model (Section 2.2.1). Middle row (c-d):
747 Two-equation model (Section 2.2.2). Bottom row (e-f): Multi-variable two-equation model (Section 2.2.3).
748
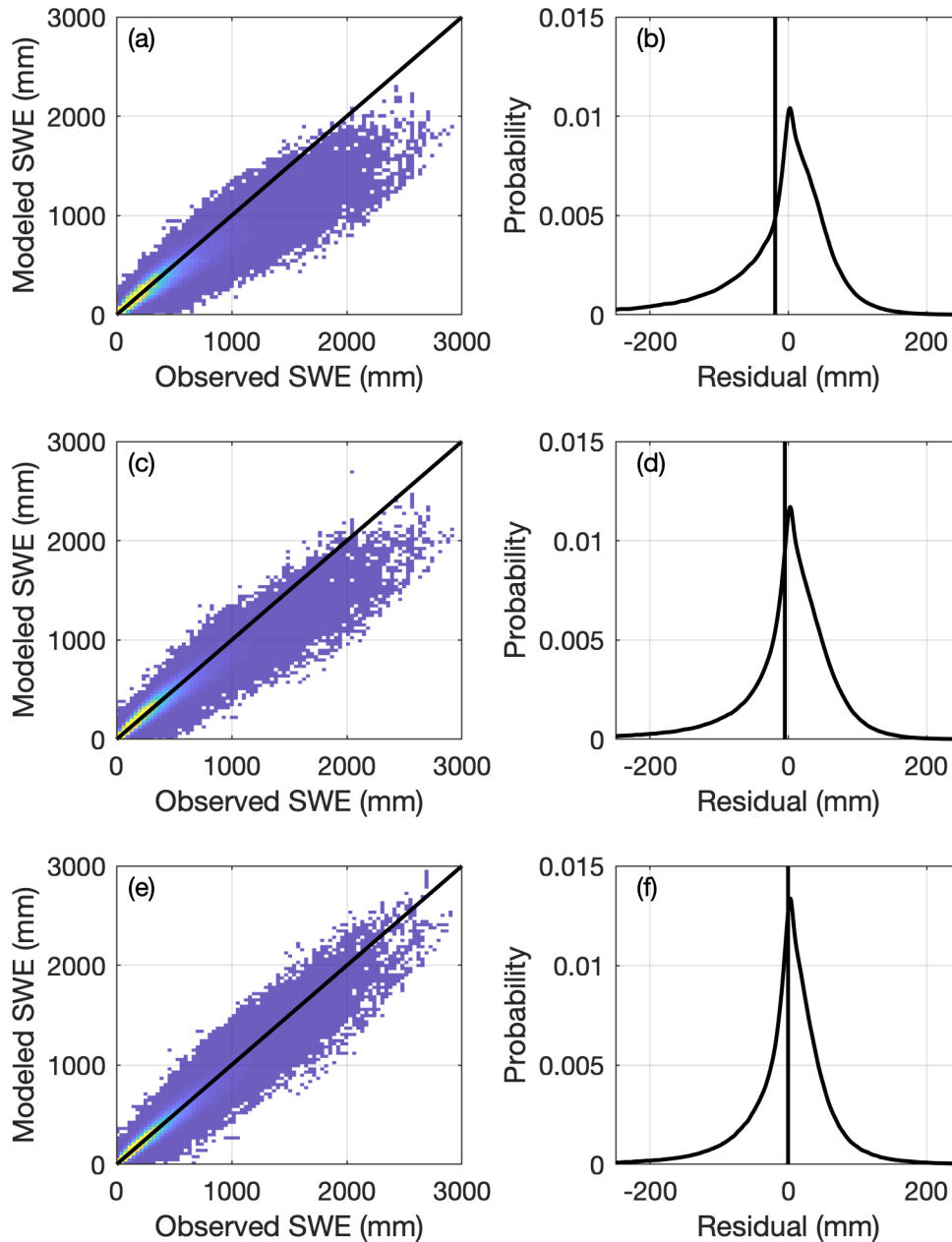
749

750   Figure 7: Heat map of SWE residuals as a function of *DOY*.
751
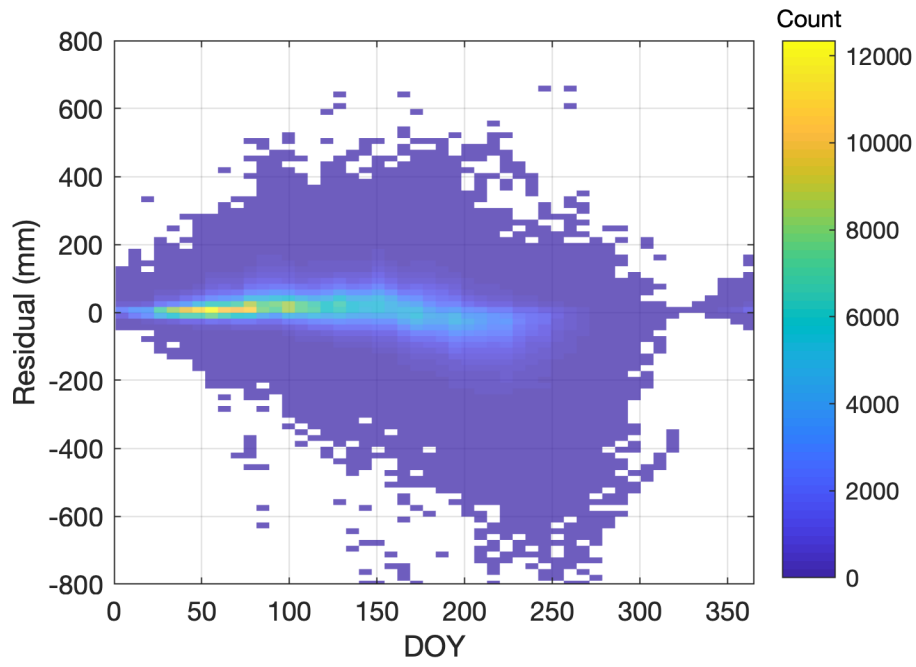
752

753   Figure 8: Probability density function of snow pillow station root-mean-square error (RMSE) normalized by station

754   mean annual maximum SWE.
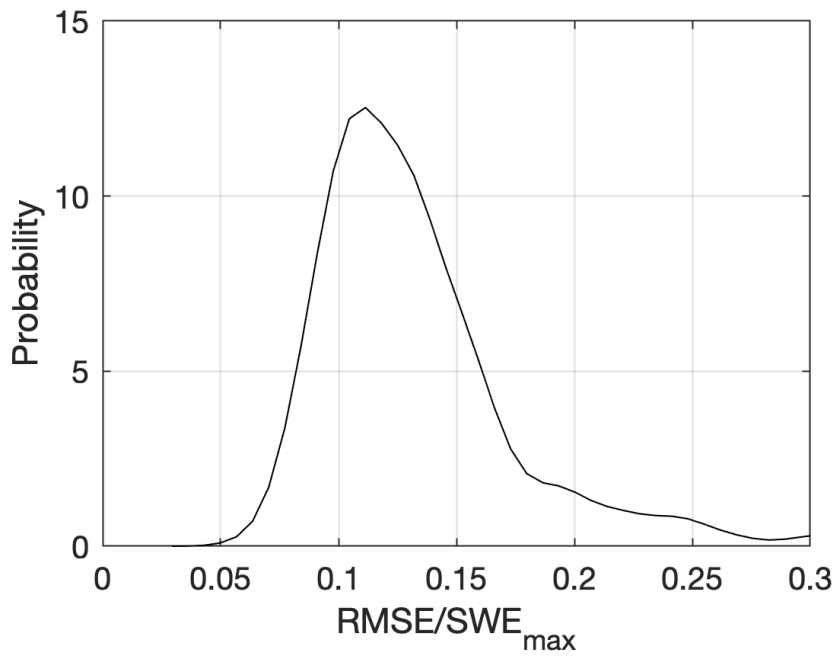
755

756 Figure 9: Spatial distribution of snow pillow station root-mean-square error (RMSE) normalized by station mean
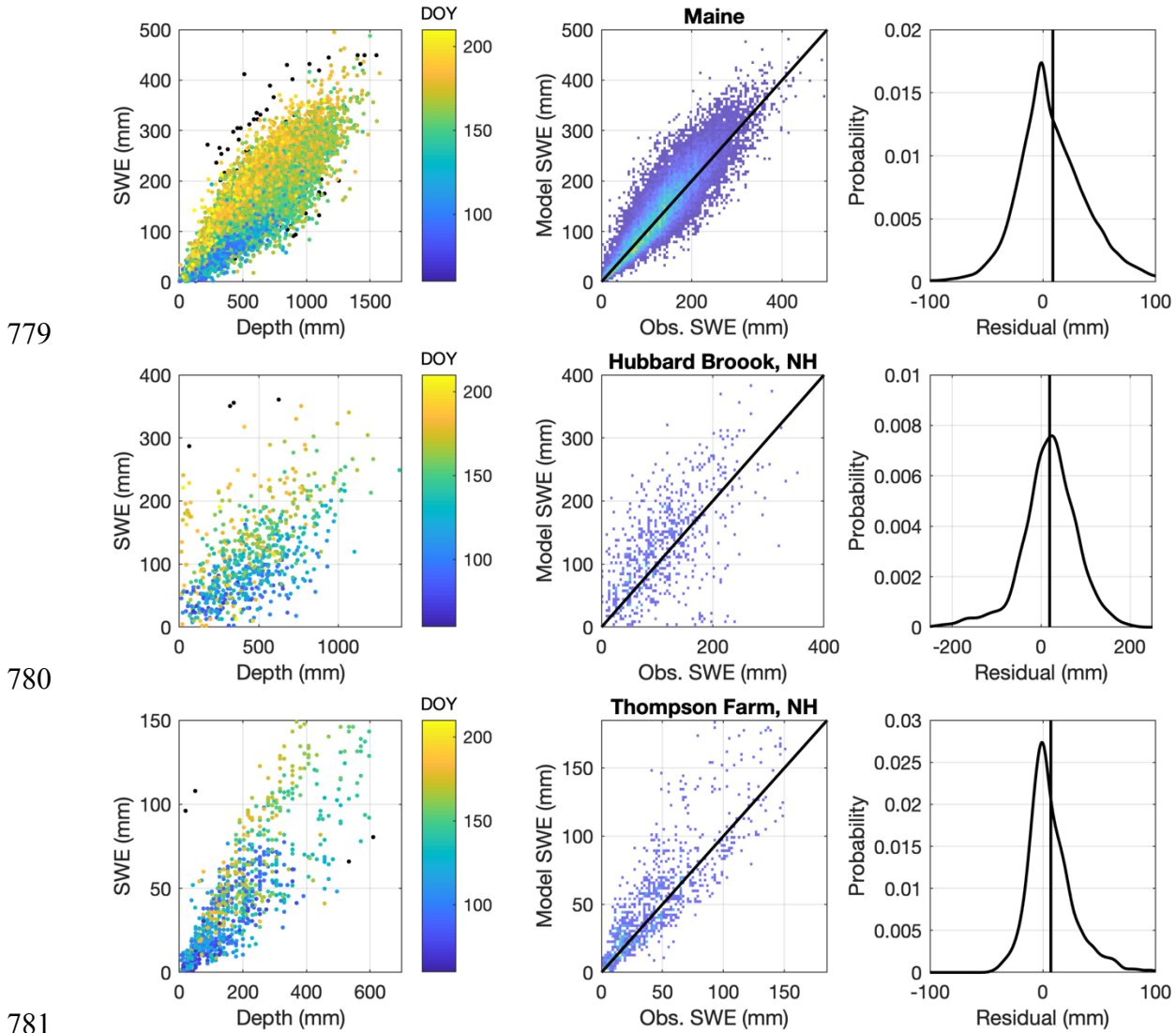
757 annual maximum SWE.

758
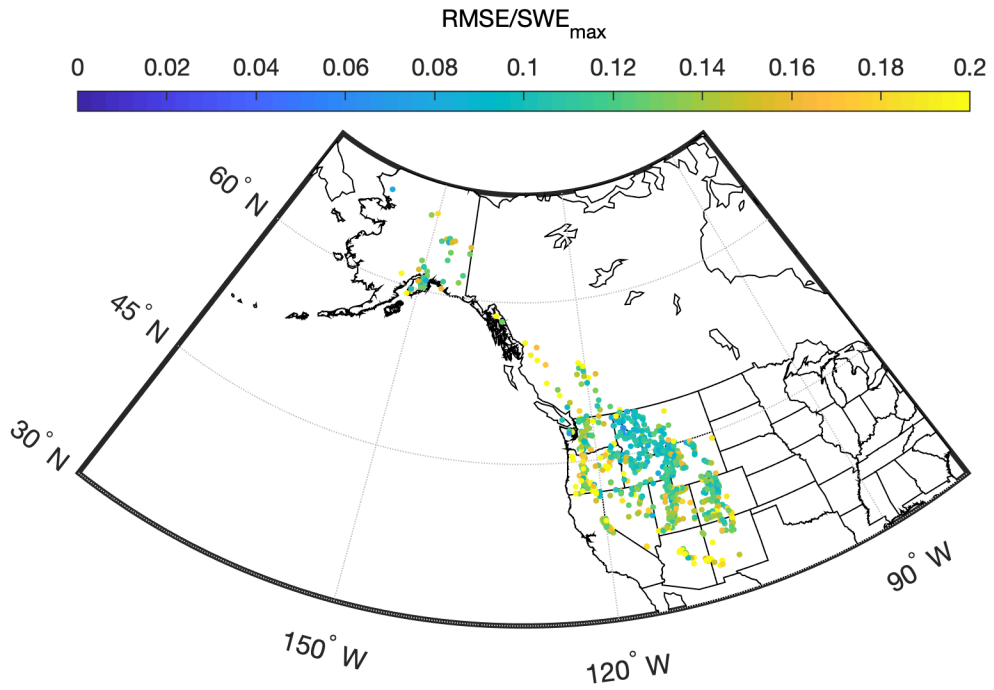
759

760 Figure 10: Comparison of the multi-variable, two-equation model of the present study with the model of Sturm et al.

761 (2010). The subpanels show modeled SWE vs. observed SWE for all of the data binned together, as well as for the

762 data broken out by the snow classes identified by Sturm et al. (1995). The gray symbols show the Sturm result and

763 the transparent heat maps (warmer colors indicate greater density) show the current result. The models are being

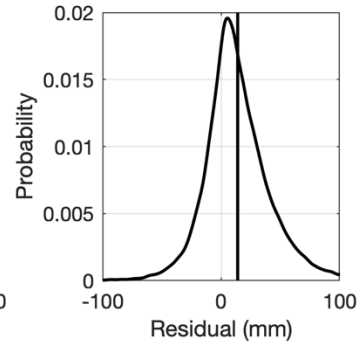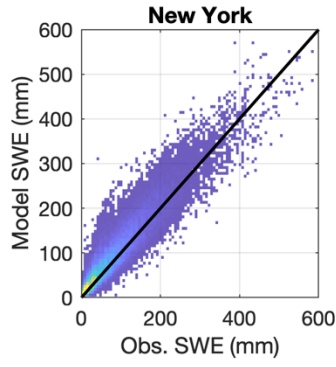764 applied to the validation data set (50% of the aggregated snow pillow data for CONUS, AK, and BC).
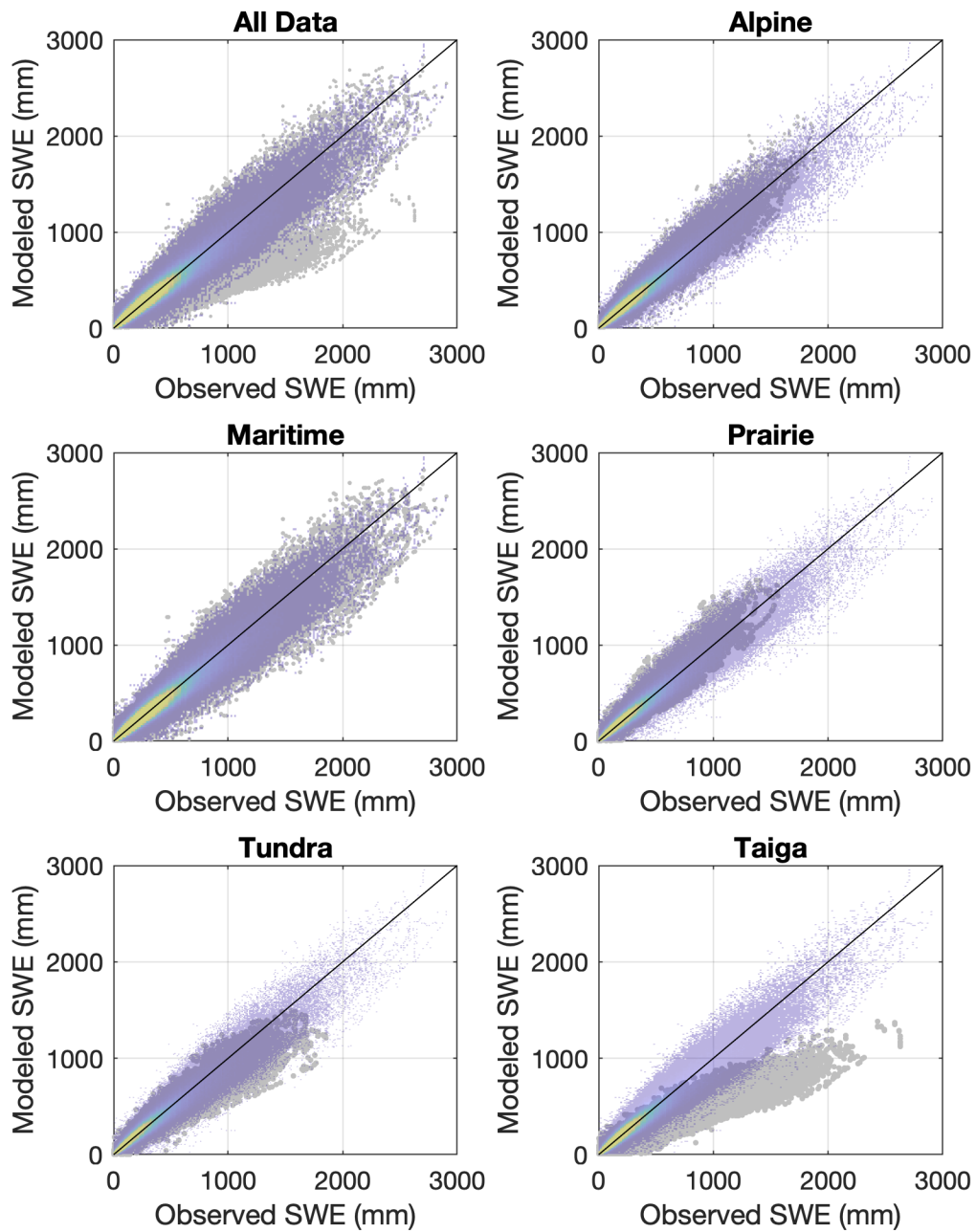
765

766   Figure 11: Comparison of the multi-variable, two-equation model of the present study with the model of Sturm et al.

767   (2010). The subpanels show probability density functions of the residuals of the model fits for all of the data binned

768   together, as well as for the data broken out by the snow classes identified by Sturm et al. (1995). The gray lines

769   show the Sturm result and the colored lines show the current result. The vertical lines show the mean error, or the

770   model bias, for both the Sturm and the current result. The models are being applied to the validation data set (50% of

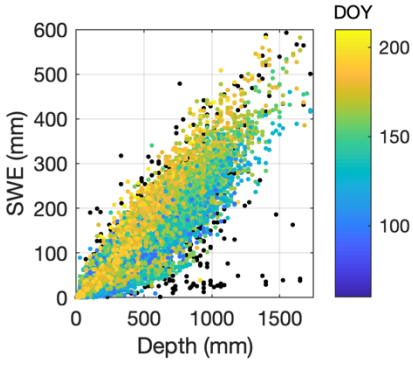771   the aggregated snow pillow data for CONUS, AK, and BC).
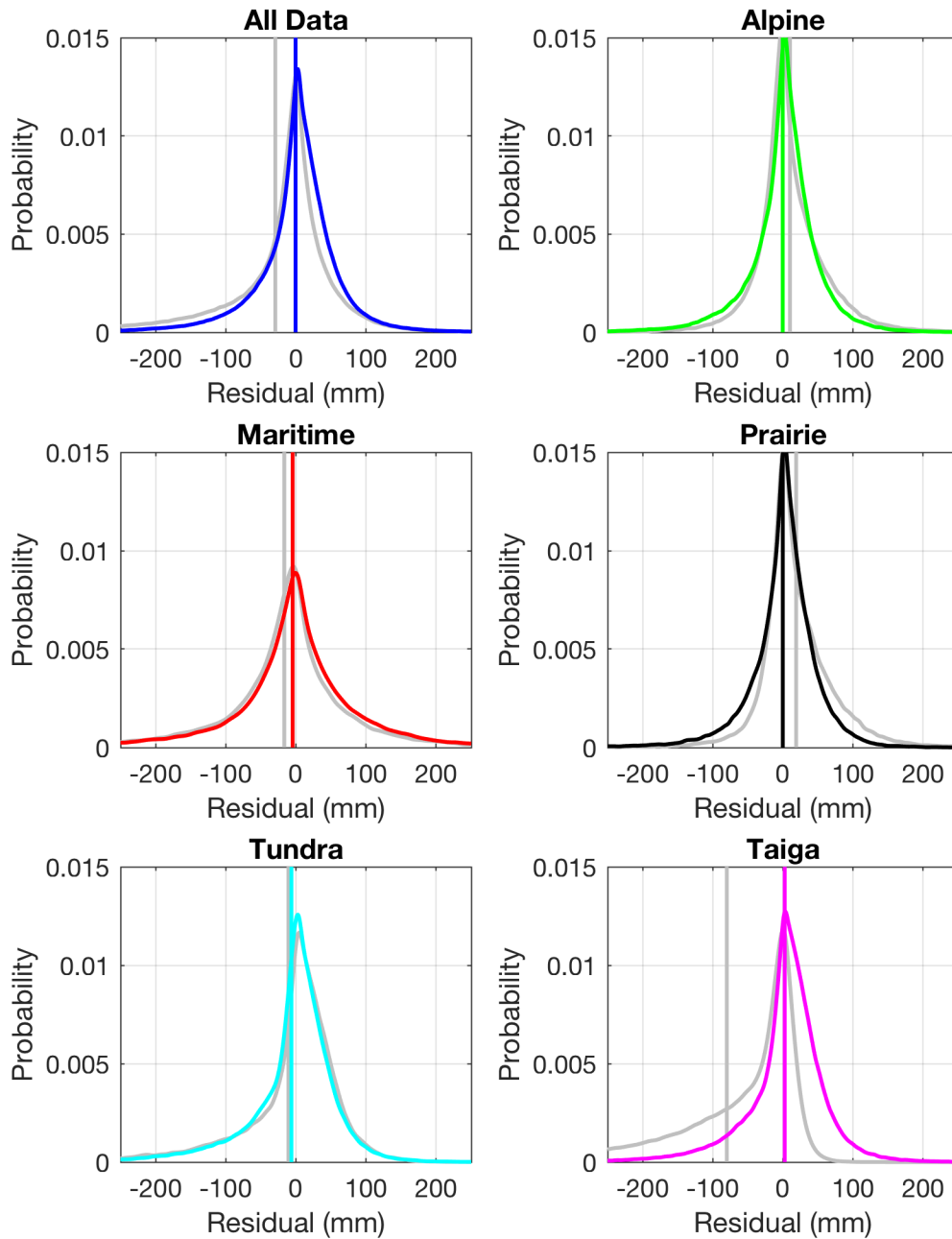
772

773 Figure 12: Results from application of the multi-variable, two-equation model to numerous east coast datasets. The
774 left column shows the SWE-*h* data for each dataset. Note that the black symbols are points removed by the outlier
775 detection procedure discussed in section 2.1.1.4. The remaining symbols are colored by DOY. The middle panel
776 plots heat maps of the model estimates of SWE against the observations of SWE with the 1:1 line included. Warmer
777 colors indicate higher densities. The right panel shows probability density functions of the model residuals, with the
778 vertical line indicating the mean error, or bias. Individual rows correspond to individual data sets and are labeled.
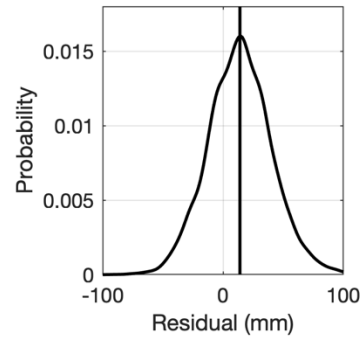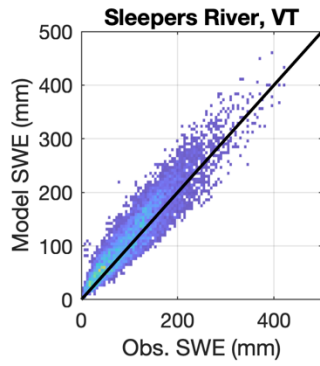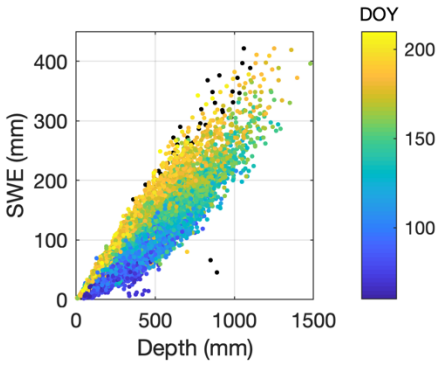
779

780

781

782

783

784
785
786

34