Improving Snow Deposition Magnitude and Heterogeneity Using Historic Snow Patterns in the California, USA, Sierra Nevada

Justin Matthew Pflug

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2021

Reading Committee:

Jessica D. Lundquist, Chair

Bart Nijssen

David Shean

Program Authorized to Offer Degree:

Civil and Environmental Engineering

University of Washington

Abstract

Improving Snow Deposition Magnitude and Heterogeneity Using Historic Snow Patterns in the California, USA, Sierra Nevada

Justin Matthew Pflug

Chair of the Supervisory Committee:

Dr. Jessica D. Lundquist

Civil and Environmental Engineering

Mountainous snow-covered landscapes in the Western United States behave like natural reservoirs, storing water during cold winter periods and sustaining snowmelt-driven streamflow vital for agriculture, municipalities, hydropower generation, and local ecosystems. In these regions, the timing and duration of spring snowmelt, and the resulting streamflow, are driven by both total snow volume and its spatial distribution across the landscape. Yet, our ability to model mountainous snow magnitude at hillslope spatial scales (< 100 m resolution) is hindered by uncertainties in snowfall and misrepresentations of snow processes like wind redistribution, preferential deposition, and avalanching. Fortunately, snow deposition in mountainous landscapes is driven by the interaction between prevailing snowstorm characteristics and static features like terrain and vegetation, often resulting in interannually repeatable snow distribution patterns.

This dissertation investigated the value of repeatable snow patterns in the California Sierra Nevada using an unprecedented collection of ground-based, airborne, and satellite-based snow

observations. We investigated how historic information about snow distribution and real-time snowpack observations could be combined to infer snow magnitude at hillslope spatial scales using both statistical and numerical modeling approaches. In Chapter 2, we began by calculating snow depth pattern repeatability at 25 m spatial resolution using a set of 47 airborne lidar snow depth observations in the Upper Tuolumne river watershed spanning water-years 2013 through 2019. Our results showed that normalized snow depth patterns between observation dates with similar relative amounts of snow accumulation and depletion, similar seasonal timing, and similar snow extents, were well-correlated in space (median $r > 0.84$). This pattern repeatability could be used to infer watershed-scale snow depth distribution using the relationship between a snow depth pattern from a different year and a small subset of real-time observations covering < 4% of the watershed.

In Chapter 3, airborne lidar snow depth patterns were used to downscale 25 m snow deposition, as a substitution for more complex snow accumulation processes, from coarser-scale (6 km) snowfall estimates. Snow models were used to evaluate the accuracy of pattern-based snow deposition downscaling in the Upper Tuolumne river watershed. Snow input downscaled using snow depth patterns resulted in simulations with improved snow depth spatial heterogeneity ($0.52 < r < 0.76$), as compared to more-common statistical methods that downscaled snow input using terrain elevation ($r = 0.27$). However, snow input downscaled using historic snow patterns were still subject to snowfall biases from atmospheric models, which can be as large as $\pm 60\%$ in mountainous terrain. Winter losses (snowmelt and sublimation), snow density spatial variability, and interannual changes to snow depth patterns also influenced the spatial variability of the snow depth patterns and the resulting accuracy of pattern-based snow downscaling.

In the last phase of this dissertation (Chapter 4), we focused on correcting the pattern-based snow downscaling strategy using snow accumulation patterns from a 31-year snow water equivalent (SWE) reanalysis. Using the relationship between historic snow accumulation patterns, and ground observations of SWE accumulation, watershed-scale mean snowfall was inferred to within $\pm$13% accuracy (interquartile range) between 1985 – 2016 in the Upper Tuolumne, Upper Kings, and Sagehen Creek Sierra Nevada watersheds. Additionally, as compared to snow depth simulations that downscale snowfall from a popular meteorological product (NLDAS2) using terrain elevation ($r = 0.54$, on average), simulations inferring snow input using the relationship between historic snow accumulation patterns and real-time ground observations had significantly improved snow depth spatial distribution ($r = 0.76$). These results show that, where and when interannually repeatable snow patterns emerge, historic snow information can be used to overcome model and forcing data constraints and improve representations of snow magnitude and hillslope spatial heterogeneity in a way that is practical for both operational and scientific purposes.

Acknowledgements

I would like to express my thanks and gratitude to the graduate supervisory committee for their support, guidance, and feedback. In particular, I want to thank Jessica Lundquist, who has encouraged my research interests and advocated for me since my arrival at the University of Washington. I have Jessica to thank for my research, field experiences, and professional connections that I have fostered both within, and outside of, the University of Washington. I would also like to thank committee members Bart Nijssen and David Shean, who both provided feedback on this research, and research preceding this dissertation. I would also like to thank researchers and professionals who have at one point hosted me. These individuals include 1) Glen Liston, who trained me on the model that I have used extensively in this research, 2) Bill Baccus, who advised me on work analyzing glacier decline in the Olympic Mountains, 3) Tobias Jonas, who provided collaboration opportunities at the Swiss Institute for Snow and Avalanche Research, and 4) Steve Margulis, who advised me on the snowpack reanalysis used in Chapter 4 of this dissertation. Finally, I would like to thank Lynn McMurdie, who is serving as the Graduate School Representative.

Financial support for this research came from the NASA Earth and Space Sciences Fellowship. Airborne lidar data was provided by the Airborne Snow Observatory. I would particularly like to acknowledge Kathryn Bormann, who assisted with processing the lidar surveys and making the data available. I would also like to thank Adrian Harpold, who provided access to some of the lidar surveys used in this research. WRF data was provided by Mimi Hughes, and Sierra Nevada snowpack reanalyses were provided by Steve Margulis. I would also like to recognize Adam Winstral, and other anonymous journal reviewers, for their thoughtful feedback on Chapters 2 and 3 of this dissertation.

I would finally like to thank the members of the Mountain Hydrology Research Group at the University of Washington, who have been my research family, providing mentorship and friendship through the ups and downs of graduate school. Specifically, I would like to thank Ryan Currier, Steven Pestana, Cassie Lumbrazo, Dylan Reynolds, Nicoleta Cristea, Max Mozer, Giulia Mazzotti, Teddy Thorson, Colin Butler, Karl Lapo, Victoria Ly, Eric Keenan, Brian Henn, Hanna Besso, Joe Ammatelli, Andri Gunnarsson, Michelle Hu, Mark Raleigh, and Calista Moore. Most of all, I would like to express my gratitude to my family, and especially my partner, Maggie, who have stood by me and supported my graduate school journey.

Published Material

At the time of writing, Chapter 2 of this research has been published, and Chapter 3 is in revisions. Both chapters were submitted to *Water Resources Research*. I would like to acknowledge John Wiley and Sons, publisher of *Water Resources Research*, for granting permissions to reproduce the published and submitted articles. The citation for the published article is:

Pflug, J.M., Lundquist, J.D., 2020. Inferring distributed snow depth by leveraging snow pattern repeatability: Investigation using 47 lidar observations in the Tuolumne watershed, Sierra Nevada, California. Water Resources Research 56, e2020WR027243. https://doi.org/10.1029/2020WR027243

Table of Contents

List of Tables

List of Figures

Chapter 1 : Introduction

Mountainous snowpack is a vital water resource for over a billion people worldwide (Dozier et al., 2016), and is a major driver of wildlife habitat (e.g., Barsugli et al., 2020; Liston et al., 2016) and heat and mass exchanges between the land surface and atmosphere (e.g., Aas et al., 2017; Mott et al., 2017). Additionally, spring and summer streamflow is sensitive to both snow volume and snow spatial distribution in mountainous basins. For instance, total snow volume near peak snowpack timing is important for forecasting the amount of snow-fed streamflow expected for a given year, and the water resource allocation to downstream municipalities and water-rights holders (Li et al., 2019; Sturm et al., 2017). Snowpack spatial heterogeneity also influences the timing and duration of snowmelt as deeper snow deposits require more energy to bring to ripe, melting conditions (Brauchli et al., 2017; Egli et al., 2012; Freudiger et al., 2017; Liston, 2004; Luce et al., 1998; Lundquist and Dettinger, 2005). Therefore, better representations of winter snow distribution, and the resulting timing and duration of spring snowmelt, may improve reservoir and hydropower operations.

Snow magnitude and spatial distribution is driven by the superposition of both snow accumulation and depletion. At fine spatial scales (< 10 m), processes like snow creep, wind-redistribution, forest canopy interception, and preferential deposition control the heterogeneity of snow accumulation (Clark et al., 2011; Currier and Lundquist, 2018; Dadic et al., 2010a; Filhol and Sturm, 2019; Moeser et al., 2015; Winstral et al., 2002). At larger spatial scales (100 m to several kilometers), snow accumulation is driven more by coarser-scale features like elevation, slope, and terrain aspect, which influence orographic precipitation gradients, rain shadow effects, and snow avalanching (Bernhardt and Schulz, 2010; Clark et al., 2011; Elder et al., 1991; McGrath et al., 2018; Minder et al., 2008; Roe and Baker, 2006). Snow depletion is driven largely by radiation (Cristea et al., 2014; Marks and Dozier, 1992; Mazurkiewicz et al., 2008),

which like snow accumulation, is influenced by time-static features like terrain and vegetation. Although the spatial heterogeneity of both snow accumulation and depletion are important for snow modeling, the spatial heterogeneity of snow accumulation in mountainous landscapes can influence spring snowmelt and snow distribution as much as, or more than, the spatial heterogeneity of snowmelt energy (Egli et al., 2012; Luce et al., 1998).

Despite the importance of snow accumulation, most models struggle to represent snow deposition in mountainous terrain, often resulting in biased snowpacks with simulated snow depth and snow water equivalent (SWE) that is too spatially homogeneous (Brauchli et al., 2017; Hedrick et al., 2018; Vögeli et al., 2016). These struggles stem from issues with the models themselves, or the meteorological data used to force simulations. For instance, snow processes like wind-redistribution, preferential deposition, avalanching, and canopy interception are not represented by all models. Even for models that do include some of these processes, process representations are often heavily parameterized, resulting in model setups that perform well in one location, but are not easily transferred between domains with different snow conditions (e.g., Clark et al., 2015; Essery et al., 2013; Mendoza et al., 2015). However, even a model with perfect representations of all snow accumulation processes may be too computationally expensive for practical purposes. For instance, to explicitly model wind-redistribution, simulations would have to be performed at spatial scales smaller than the wind-drifts themselves (< 10 m), which would dramatically increase the computational cost of watershed- or regional-scale simulations. Despite this, the accuracy of simulations are typically constrained on a first-order by the accuracy of meteorological forcing data like precipitation, which is the most-cited meteorological driver of snow model errors (Günther et al., 2019; Lundquist et al., 2019; Raleigh et al., 2015; Wayand et al., 2013).

Fortunately, snow accumulation and depletion patterns are driven by the interaction between local meteorology and static features (e.g., terrain and vegetation), which at seasonal timescales, have been seen by many to result in interannually repeatable snow patterns (Deems et al., 2008; López-Moreno et al., 2017; Schirmer et al., 2011; Schirmer and Lehning, 2011; Sturm and Wagner, 2010a; Winstral et al., 2002; Woodruff and Qualls, 2019). This dissertation focuses on snow distribution patterns, and how information about snow distribution in historic periods can be useful in later periods. We specifically focused on the California, USA, Sierra Nevada mountain range where we have a unique library of snow depth observations from airborne lidar surveys (ALS) (Painter et al., 2016) and a multidecadal SWE reanalyses (Margulis et al., 2016a).

In Chapter 2 of this dissertation, we used the relationship between normalized snow depth maps from different years, and snow depth observations over just a portion (< 4%) of a mountainous watershed, to infer domain-scale snow depth mean and standard deviation, and the resulting distribution of snow depth at 25 m posting. Additionally, we evaluated the causes of interannual snow depth pattern differences and demonstrated the impact of total snow mass and snow spatial heterogeneity on simulated spring snowmelt. Expanding on these results in Chapter 3, we investigated how snow depth patterns from near peak-snowpack timing in different years could be used to downscale 25 m snow deposition, as a proxy for unmodeled snow accumulation processes, from 6 km mean snowfall simulated by an atmospheric model. Using snow depth simulations and an ensemble-based approach, we diagnosed and ranked model errors driven by 1) atmospheric model snowfall biases, 2) the impact of winter snowmelt, winter snow sublimation, and spatially variable snow density on snow depth patterns, and 3) interannual changes in snow depth patterns. Finally, in Chapter 4, we investigated how representative SWE accumulation patterns from a multidecadal SWE reanalysis could 1) be identified using real-time

information about snow cover, 2) be used to infer watershed-scale mean snowfall from ground observations, and 3) simulate 90 m snow deposition magnitude. In Chapter 5, we discussed how our methods used in the previous chapters, which focused only on the California Sierra Nevada, could be tested globally, including mountainous regions with varying levels of snow pattern repeatability and a wider variety of snow observations. Our methods here were focused on improving estimates of total snow volume and snow spatial heterogeneity for both operational and scientific purposes, in a fashion that was practical using models and data at our disposal today.

Chapter 2 : Inferring distributed snow depth by leveraging snow pattern repeatability: Investigation using 47 lidar observations in the Tuolumne watershed, Sierra Nevada, California

Justin M. Pflug[1] and Jessica D. Lundquist[1]

[1] Department of Civil and Environmental Engineering, University of Washington, Seattle, WA, USA.

Note: This chapter has been published in its current form as an article in Water Resources Research (Pflug and Lundquist, 2020); the only differences are in section, figure, and table numbering. It is used here by permission of John Wiley and Sons.

2.1. Abstract

Snow distribution is controlled by the interaction between local meteorology and static features like topography and vegetation. The resulting spatial pattern of snow in mountainous terrain is often repeatable and can be used to infer snowpack distribution in periods when observations are limited. This study uses a library of airborne lidar surveys (ALS) in California's Tuolumne watershed to analyze snow patterns at extents (1650 km^2), resolutions (25 m), and temporal scales (47 ALS observations over 7 years) unmatched by previous research. Distributed snow depth was inferred from snow depth observations covering a portion of the domain (< 4%) at a period near peak-snowpack timing, and snow distribution patterns from different years. Snow patterns from different years differed as a function of snow extents, variability, and interannual noise (r = 0.30 to 0.90). However, matching criteria like seasonal timing and snow extents were able to identify pairs of snow patterns with increased spatial agreement (median r > 0.84). Distributed snow depth inferred using a strip of observations (< 4% coverage), and a well-correlated snow pattern ($r \geq 0.90$) had a mean absolute error (MAE) of 0.22 m and snow volume error of -6%. Distributed snow depth inferred using patterns of reduced accuracy (r < 0.80) were often too homogeneous, thereby increasing MAE and decreasing the duration of the simulated snowmelt season. This work has applications in water management, where distributed snow depth observations in watersheds with interannual snow pattern repeatability could decrease the extent of observations necessary in future years.

2.2. Introduction

The spatial distribution of snow in mountainous terrain affects the timing of snowmelt, thereby influencing water supply for an estimated billion people worldwide who rely on mountainous snowpack water resources (Dozier et al., 2016). In California alone, over 62% of water resources for municipalities and agriculture come from Sierra Nevada seasonal snow melt (Li et al., 2017).

This runoff provides water to agribusiness annually valued at $47 billion (*CDFA*, 2017) and major cities, including 4 municipalities with gross domestic products (GDP) within the top 15 in the United States (*BEA*, 2017). Streamflow forecasts are also important to watershed managers who are tasked with optimizing hydropower generation while managing downstream flood risks. Accurate determination of distributed snowpack volume and the expected timing of water resources are therefore of immense economic value at both local and global scales.

To meet the demand for better streamflow forecasts, California state has invested in airborne lidar surveys (ALS) (Painter et al., 2016). ALS observes snow depth at spatial extents and resolutions unmatched by in-situ and satellite observations (Blöschl, 1999; Cao et al., 2017; Currier et al., 2019; Kirchner et al., 2014), thereby providing valuable insights into watershed hydrology (Hedrick, 2018; Henn et al., 2016a; Margulis et al., 2019a), drivers of snowpack deposition (Currier and Lundquist, 2018), and snowpack observation strategies (Oroza et al., 2016). Although valuable, distributed snow depth observations, such as those provided by ALS, are rare in montane watersheds globally. Recent studies have therefore focused on how interannually repeatable snow distribution patterns (e.g., Sturm and Wagner, 2010; Vögeli et al., 2016) and snow distribution hysteresis (e.g., Egli et al., 2012; Liston, 2004) can be used to infer snowpack distribution in locations and during periods where snow depth observations are limited.

In this study, we used a library of ALS data to investigate snow patterns at unprecedented temporal scales (7 snow seasons), spatial extents (1650 km^2), and spatial resolutions (25 m). Specifically, we tested the extent to which distributed snow depth in a California Sierra Nevada watershed at a period near peak-snowpack could be inferred using snow depth observations covering only a portion (< 4%) of the domain at that time, and a snow distribution pattern from a

different water year. We evaluated the accuracy of inferred snow depth using common metrics (mean absolute error and snow volume percent error), but also tested how differences in snow patterns impacted our estimates of distributed snow depth and the resulting snowmelt using a simple radiation-based model. Finally, snow pattern repeatability between all ALS snow patterns from different water years were evaluated using matching criteria, such as proximity to peak-snowpack timing (from continuous point observations of SWE) and snow-covered extents (from ALS), to identify snow patterns that were more likely to match. We discussed how historic snow patterns observed from distributed datasets could provide valuable insights into snowpack spatial distribution at times and locations where snow depth observations may be limited in frequency and/or spatial extent. Although this work focused on one watershed in California, we hypothesize that the methodology and results are likely transferable to many montane regions with socially and economically valuable snow water resources.

2.3. Background

Snow distribution patterns have been described by spatial variations in topography, vegetation, and wind fields (Currier and Lundquist, 2018; Elder et al., 1991; McGrath et al., 2018; Schmidt et al., 2009; Trujillo et al., 2007) which influence snow distribution at both hillslope (1 – 100 m) and watershed (100 – 10,000 m) spatial scales (Clark et al., 2011). On annual timescales, these patterns have often been found to be repeatable in space and insensitive to snowfall magnitude (Deems et al., 2008; Sturm and Wagner, 2010a; Vögeli et al., 2016). For instance, in an Alaskan arctic domain, Sturm and Wagner (2010) noted that the spatial distribution of snow was repeatable across 14 gridded snow courses at times near peak-snowpack between 1990 and 2009. Fine-resolution (1 m) observations of snow depth at times near peak-snowpack also revealed repeatable snow patterns in the Colorado Rocky Mountains (Deems et al., 2008) and Swiss Alps (Schirmer et al., 2011; Schirmer and Lehning, 2011), although each of

these studies compared only two snow seasons. Snow depth spatial patterns observed by Schirmer et al. (2011) at the conclusion of the accumulation season in the Swiss Alps were more repeatable than snow patterns from individual snowfall events, as the characteristics of these snowfall events varied. Snow-extent repeatability was also found across 18 snow melt seasons in the Upper Snake River Basin in Idaho, using observations from the MODerate Resolution Imaging Spectroradiometer (MODIS) (Woodruff and Qualls, 2019).

The presence of repeatable snow patterns suggested that shortcomings in snow observations and simulations could be improved by using historic snow distribution patterns. Sturm and Wagner (2010) used normalized snow patterns from different years to identify simulations with more accurate snowpack distribution from an ensemble of simulations with various blowing-snow parameters. Vögeli et al. (2016) also used information about historic snow patterns to adjust incoming precipitation as a proxy for physical processes that were unrepresented by simulations. Subgrid snow distributions, wherein snow distribution at spatial scales finer than those of simulations or observations (e.g., He and Ohara, 2019; Liston, 2004; Skaugen, 1999), have also been used to implicitly mimic the fine-scale variability of snow without attempting to physically resolve complex processes like wind-redistribution. Recently, subgrid distributions have been assimilated with simulations and satellite-observations of snow-covered area (SCA) to retroactively reanalyze snowpack in the California Sierra Nevada and high mountain Asia (Margulis et al., 2016a, 2019b). A review of snow distribution studies, both explicit and subgrid, is presented in Clark et al. (2011).

Here, we used ALS observations over the Tuolumne watershed in California state (Section 2.4) to resolve snow depth spatial patterns. We normalize snow patterns using a common statistical distribution (like implicit methods from Liston (2004) and Skaugen (1999)), but

explicitly evaluate snow depth heterogeneity at 25 m resolutions. This domain is typically subjected to relatively few (less than 10) winter storms that have little variance in spatial trajectory and orientation (Lundquist et al., 2015, 2010). Therefore, incoming storm tracks interact with the terrain in a repeatable fashion, generating snow patterns as a function of elevation, preferential deposition (preferred placement of snowfall from terrain/vegetation-influenced wind-fields), and wind-redistribution. ALS observations were collected from peak-snowpack timing throughout the snowmelt season and were therefore subject to both, spatially unique, snow accumulation and melt patterns. Currier and Lundquist (2018) noted statistically different snow patterns on north-, and south-facing sides of forested stands, which were described by differences in incoming radiation. Since spring snowmelt in the California Sierra Nevada is dominated by radiation (Marks and Dozier, 1992), we expect melt patterns influenced by shading from the unchanging terrain to be spatially repeatable at interannual timescales.

2.4. Study site

The Tuolumne River basin is in the California Sierra Nevada mountain range within Yosemite National Park. The basin ranges from 700 – 3900 m in elevation, covering approximately 1650 square kilometers, with 36% forest coverage (Figure 2.1). River discharge in this basin is dominated by snowmelt, with 30 – 40% of the annual snowmelt from elevations of 3000 m and above (Rice et al., 2011). To our knowledge, minimal forest and watershed management have occurred in this region, leaving the terrain and vegetation, and resulting drivers of snow deposition and melt patterns, similar between water-year 2013 (WY2013) and water-year 2019 (WY2019). While several forest-fires have occurred in this region, most were at elevations below and outside of the study domain.

In addition to the full Tuolumne study domain, an upper-elevation subdomain (6.25 x 6.25 km) was also selected (Figure 2.1, yellow square). To date, most snow-pattern studies have focused on areas that are smaller in extent and primarily above the seasonal snow-line (e.g., Deems et al., 2008; Egli and Jonas, 2009; Schirmer and Lehning, 2011), therefore predominantly covered by snow and impacted by snow-absence to a lesser degree. The subdomain used in this study (2430 – 3230 m elevation, 19% forested), was chosen to mimic these conditions.

2.5. Data

ALS snow depth observations were collected and processed by the Airborne Snow Observatory (ASO) (Painter et al., 2016). Snow depth rasters were provided at 3 m resolution and are accurate to within $\pm$0.08 m, or better, in both forested and open areas (Currier et al., 2019; Painter et al., 2016). We used 47 ALS observations of snow depth between WY2013 and WY2019. To our knowledge, this extent of ALS coverage and frequency is unusual, making this the most ALS-observed region in the world.

ALS 3 m snow depth gridcells that were adjacent or nearby each other were likely impacted by the same physical processes or features (e.g. snow drift, forest-edge, ridgeline cornice, etc.). Currier and Lundquist (2018) noted a high amount of spatial autocorrelation (Moran's index between 0.75 and 0.87) for the 3 m resolution lidar data in the Tuolumne watershed. Variogram analysis at hillslope scales (< 100 m) showed that the reach at which the variogram leveled out was approximately 25 m in length and changed little for different locations across the Tuolumne watershed. However, the variability of snow depth across the full Tuolumne study domain consistently increased with longer reaches as variability was driven primarily by elevation. This result was consistent with findings from Clark et al. (2011) who found that snow depth variance across a large mountainous domain was largest when differences in elevation were maximized.

ALS observations in this study were regridded to 25 m spatial resolution to reduce spatial autocorrelation and isolate snowpack deposition and melt processes at hillslope scales (< 100 m), all the while preserving elevation-driven snow depth gradients.

The California Department of Water Resources Dana Meadows snow pillow (37.9°N, 119.3°W, 2966 m) (Figure 2.1) was used to identify the proximity to the date of peak snow water equivalent (SWE) for each ALS observation in each snow season. Proximity to peak-snowpack timing was measured in days, where periods prior to peak-snowpack timing were negative, and periods after peak-snowpack timing were positive.

Water-year 2014 6-km Weather Research and Forecasting (WRF; Skamarock, 2008) modeled surface fields of air temperature, relative humidity, wind speed, wind direction, and precipitation were used to force snowmelt simulations. Boundary conditions were from the North American regional reanalysis climate model (NARR; Mesinger et al., 2006). WRF was chosen because surface forcing data from distributed atmospheric models have demonstrated benefits over statistically distributed point-measurements in complex, mountainous terrain (Currier et al., 2017; Havens et al., 2019; Lundquist et al., 2019; Wayand et al., 2013). A statistical interpolation routine (Barnes, 1964; Liston and Elder, 2006a) was used to downscale 6-km WRF forcing fields to the model grid (25 m) (see Section 2.6.3).

2.6. Methods

Our methods focused on the capability to infer distributed snow depth using snow depth observations subsampled across only a portion of the study domain and a snow pattern from a different water year. We started by testing how ALS snow depth observations conformed to common statistical distributions, and how statistical distributions could be used to normalize snow depth observations in years with different snow depth magnitudes (Section 2.6.1). We then

investigated how a strip of snow depth observations (covering < 4% of the domain) from a period near peak-snowpack timing in one year could be related to snow patterns from a different year, and in the process, infer the mean, standard deviation, and distribution of snow depth in regions without observations (Section 2.6.2). In addition to widely used metrics (mean absolute error and snow volume percent error) we tested how modeled basin-wide snowmelt changed when initialized with peak-snowpack snow depth maps inferred from snow patterns of various accuracies (Section 2.6.3). Although we inferred distributed snow depth on only one date (near peak-snowpack in 2014) in Sections 2.6.2 and 2.6.3, we also tested snow pattern repeatability between all pairs of ALS-observed snow patterns from two different years for both the full Tuolumne study domain and the upper-elevation subdomain (Section 2.6.4).

2.6.1. Snow patterns

While snow depth distributions have been analyzed using normal (e.g., Elder et al., 1991; Marchand and Killingtveit, 2005), lognormal (e.g., Jonas et al., 2009; Liston, 2004), and gamma (e.g., Egli et al., 2012; Egli and Jonas, 2009; Skaugen, 1999) distributions, the shape of the ALS snow depth distributions agreed most-closely with normal distributions (Appendix A: Section 2.10). Snow patterns in periods with different snow depth magnitudes were normalized by mapping ALS snow depth observations to normal distributions using the standardized depth value (SDV) from Sturm and Wagner (2010),

$$\mathrm{SDV_i} = \frac{\mathrm{d_i} - \mu_\mathrm{d}}{\sigma_\mathrm{d}}, \tag{2.1}$$

where $\mathrm{d_i}$ is the observed snow depth at gridcell i, $\mu_\mathrm{d}$is the domain mean snow depth and σ_d is the domain snow depth standard deviation. Snow depth at any given gridcell ($\mathrm{d_i}$) was therefore put in context with the domain-wide snow conditions (this is analogous to the Z-variate in normal

distribution statistics). While we used SDV patterns here, we acknowledge that snow in different types of terrain, landscapes, and geographical regions may be best represented by different distributions and/or transformations.

SDV patterns were skewed from normal distributions by snow absent gridcells. This was particularly true for the ALS observations used here, which were collected primarily throughout the snowmelt season. For any single ALS observation, snow extents identified regions where snow disappeared earlier as a function of snowmelt and/or reduced snowfall. Since the extent and spatial orientation of snow-free regions in mountainous watersheds are meaningful for forecasting streamflow magnitude and timing (Clark et al., 2006; Hall et al., 2012), snow distribution patterns were calculated, including snow-absent gridcells, at all gridcells where snow occurred for at least one of the 47 ALS observations.

2.6.2. Snow depth spatial extrapolation

Provided snow depth observations across only a portion (< 4%) of the domain and an SDV pattern (Equation 2.1) from a different water year, the relationship between snow depth and SDV patterns could be used to infer snow depth at gridcells where observations were not available. The workflow used in this study is as follows:

1. Two ALS snow depth observations from different water years were selected from the library of 47 ALS observations. One observation was assumed to be representative of *real time* snow depth while the other was assumed to be from a *different year*.
2. Snow depth observations were subsampled from the *real time* ALS observation (Figure 2.2a).
3. An SDV pattern (Equation 2.1) was calculated using the full ALS snow depth observation from the *different year* (Figure 2.2b).
4. The subsampled snow depth observations (step 2) and corresponding SDV values (step 3) were regressed. Domain-wide *real time* snow depth standard deviation was calculated from

the slope of the regression (change in snow depth per change in SDV), and the mean was calculated from the snow depth at which the regression had an SDV of zero (Figure 2.2c).

5. An *inferred depth map* was calculated (Figure 2.2d) by rearranging Equation 2.1, substituting the snow pattern from step 3 and the mean and standard deviation calculated by the regression in step 4. Snow depth observations from step 2 were used explicitly in the *inferred depth map.*

The spatial heterogeneity of snowpack prior to melt often controls streamflow as much as, or more than, the spatial heterogeneity of incoming energy and resulting snowmelt (Egli et al., 2012; Luce et al., 1998). Therefore, we focused our efforts on inferring distributed snow depth in the full Tuolumne study domain at a period near peak-snowpack timing on 7 April 2014 using SDV patterns from all other ALS observations in different water years. We did so by using a strip of snow depth observations (~1000 m wide) covering < 4% of the full Tuolumne study domain (Figure 2.2a). Although we focused on inferred depth maps generated using one type (strip) of observation on 7 April 2014, we also investigated inferred depth maps using different snow depth observations (Section 2.8.2). We realize that in practice, *real time* snow depth could only be inferred using historic snow patterns from previous years. However, we use patterns from all years, both before WY2014 and after WY2014, to utilize the full library of ALS observations. We focus on WY2014 since we have reliable WRF-simulated surface meteorology at our disposal for this domain and water year (Section 2.5).

The accuracy of each inferred snow depth map was evaluated using the total snow volume percent-error,

$$V_{err} = \frac{\sum_i^N (d_i^* - d_i)}{\sum_i^N d_i} \times 100, \qquad (2.2)$$

and mean absolute error,

$$\mathrm{MAE} = \frac{\sum_i^N |d_i^* - d_i|}{N}, \tag{2.3}$$

where d_i^* is the 7 April 2014 inferred snow depth (Figure 2.2d), d_i is the 7 April 2014 ALS observed snow depth, and N is the number of 25 m gridcells across the Tuolumne domain (N = 1,811,893). We choose these metrics since they represent the extent to which we were able to capture total snow volume and heterogeneity, both of which are important to water resource managers.

2.6.3. Snowmelt modeling

Although the metrics above (Equation 2.2 and Equation 2.3) were representative of inferred depth map accuracy, the spatial distribution of snow depth errors is hydrologically important, as gridcells with different slopes, elevation, and vegetation receive different amounts of radiation and melt at different rates. To test the sensitivity of basin-wide snowmelt to inferred snow depth accuracy, we set up a simple radiation-based snowmelt model using atmospheric forcing from 6 km WRF surface fields, distributed to the 25 m grid using the MicroMet meteorological preprocessing routine (Liston and Elder, 2006a). Using a relationship fit between cumulative radiation and snowmelt at the Dana Meadows snow pillow (which implicitly represents the surface energy balance), cumulative potential snowmelt was calculated at every gridcell in the full Tuolumne study domain (N = 1,811,893) between 7 April 2014 and 30 September 2014. Spatiotemporal snowmelt was then modeled, initialized with snow on 7 April 2014 from 1) the ALS observation, and 2) the inferred depth maps calculated in Section 2.6.2. All 7 April 2014 snow depth was converted to SWE assuming constant snow density from Dana Meadows (0.40 m SWE, per 0.98 m depth observed by ALS at the snow pillow). For timesteps where potential melt exceeded the remaining SWE, melt at that

timestep was constrained to the remaining SWE and was snow-free for all remaining timesteps. We acknowledge that this snowmelt model is simpler than modeling approaches that explicitly solve the snow-surface energy balance. However, this simple snowmelt model illustrated 1) the interplay between snow heterogeneity and radiation variability, and 2) how errors in snow patterns and the resulting inferred depth maps influenced snowmelt magnitude and timing. The snowmelt model is discussed more in Section 2.11 (Appendix B).

2.6.4. Snow pattern repeatability

Inferred snow depth maps (Section 2.6.2) were perfect (mean absolute error of 0.0 m and snow-volume percent error of 0%) if and only if the SDV patterns (Equation 2.1) over the full Tuolumne domain on the dates of the real time observation (Figure 2.2a), and observation from a different water year (Figure 2.2b), were identical. Only ALS observations with the same number of snow-free gridcells and an identical ratio between snow depth mean and standard deviation (μ_d/σ_d) have SDV probability distribution functions (PDFs) that are the same. Differences in snow extents and/or μ_d/σ_d resulted in SDV distributions that were skewed from each other by different amounts (Section 2.10). Even if μ_d/σ_d and snow extents between two ALS observations were identical, interannual pattern noise (i.e., spatial shifting of features like snow drifts) also decreased the agreement between SDV patterns and the resulting accuracy of inferred depth maps. The degree to which two SDV patterns agreed was best measured using Pearson's coefficient of correlation,

$$r(SDV_a, SDV_b) = r(SDV_b, SDV_a) = \frac{\left(\sum SDV_{ai}SDV_{bi} - N\mu_{SDV_a}\mu_{SDV_b}\right)}{(N-1)\sigma_{SDV_a}\sigma_{SDV_b}} = \frac{\left(\sum SDV_{ai}SDV_{bi}\right)}{N-1}, \quad (2.4)$$

where N is the number of gridcells (N = 1,811,893 for the full Tuolumne study domain and N = 62,500 for the upper-Tuolumne subdomain), and SDV_a and SDV_b are SDV maps for two

different ALS observations. Any disagreement in snow extents and/or μ_d/σ_d resulted in departures from the 1:1 relationship and decreases in r. This is analogous to quantile-quantile (Q-Q) plotting (e.g., Adams et al., 2018). Interannual noise also increased the spread in the relationship between two SDV patterns and decreased the resulting r. Of all statistical metrics considered, only r exhibited the ability to represent differences in μ_d/σ_d, differences in the number of snow-free gridcells, and snow pattern noise, in a fashion that was standardized (e.g., between -1 and +1) for all comparisons of two SDV patterns.

We used time (in days) since peak-SWE occurred at the Dana Meadows snow pillow as a criterion to search the ALS library for the pattern, from a different water year, with the closest corresponding seasonal timing (Figure 2.3b). This search was performed using each SDV pattern as the reference. Instances when the SDV patterns of closest-corresponding timing differed by more than 14 days were excluded from the analysis. Snow-covered extent from the ALS observations was also used as a criterion to identify ALS observations with similar conditions. ALS observations with similar snow-covered extents were found by evaluating the percentage of the domain's gridcells that agreed using snow presence/absence binary maps (assuming snow presence for gridcells with snow depth $\geq$ 0.01 m). The ALS observation from a different water year most-closely matching the snow extent for each observation was identified (Figure 2.3c). Instances where the SDV patterns of best-corresponding snow extent agreed by less than 75% were excluded from the analysis. The dates of best-matching snow extents for the full Tuolumne domain and upper-elevation subdomain were sometimes different since snow extents for these domains were calculated separately.

The SDV patterns with the closest-corresponding timing or extents did not always identify the pair of SDV patterns from two different water years with the best pattern repeatability (r). For

example, for an SDV pattern on 3 April 2013 in the upper-elevation subdomain, the ALS observation with the best corresponding timing fell on 5 March 2015, while the best corresponding snow extent was on 23 March 2014 (Figure 2.3b and Figure 2.3c). The date of best r occurred on 1 April 2016 (Figure 2.3d) and thus was not identified by either the observation of best timing or extent. Therefore, for each ALS observation, the date analogue, or date with the best r (analogous snowpack) in a different water year, was identified to represent how repeatable a snow pattern could potentially be for every SDV pattern. Correlation coefficients (r) for ALS maps of similar seasonal timing, similar snow extents, and analogous snowpack (analogues) were then compared to the range in r for all possible comparisons of two SDV patterns from different water years.

2.7. Results

2.7.1. 7 April 2014 inferred snow depth

Inferred depth maps were generated using a strip of observations on 7 April 2014 (Figure 2.2a) and SDV patterns from different water years (Figure 2.2b). The mean absolute error (MAE) of inferred depth maps on 7 April 2014 improved when using SDV patterns from dates with better pattern repeatability (r; relative to the SDV pattern on 7 April 2014) (Figure 2.4a). In fact, inferred depth map MAE improved with r when generating inferred depth maps on any date (Figure 2.5). To illustrate the causes of inferred depth map errors, we focused on inferred depth maps generated using SDV patterns from three different dates (Figure 2.4, stars). These included SDV patterns from periods near peak-snowpack timing in a year with similar snowpack (16 April 2016), abnormally shallow snow (5 March 2015; ~34% normal from snow course measurements between 1970 and 2019), and abnormally deep snow (1 April 2017; ~209% normal).

Inferred snow depth accuracy was driven by the agreement between the strip of snow depth observations on 7 April 2014 and corresponding SDV values from different years (Figure 2.5 and Table 2.1). The regressions calculated between 7 April 2014 snow depth and corresponding SDV values from 16 April 2016 and 5 March 2015 approached the minimum SDV values (corresponding to zero snow depth in the patterns) at nonzero snow depth (Figure 2.6, scatterplots). This resulted in inferred depth maps that were nonzero at all gridcells where the SDV pattern had been snow-free. This was driven largely by snow extents on 7 April 2014 that were larger than snow extents on 16 April 2016 (r = 0.91) or 5 March 2015 (r = 0.80). Conversely, the 1 April 2017 SDV pattern (r = 0.83) had snow extents that were 5% larger than snow extents on 7 April 2014, and when regressed, reached zero snow depth at an SDV value of -1.63, which was greater than the minimum SDV value in the 1 April 2017 SDV pattern (-1.75) (Figure 2.6, bottom row). As a result, all gridcells in the inferred depth map with SDV pattern values less than -1.63 were assigned a value of zero snow, thereby reducing snow-covered extent by 4% as compared to 1 April 2017. Surprisingly, regions of mean snow depth on 7 April 2014 and 5 March 2015 agreed closely, and errors in the inferred snow depth map (generated using the SDV pattern from 5 March 2015) were balanced across the Tuolumne domain ($V_{err} = 2\%$). However, mean absolute error improved from 0.32 m to 0.22 m when inferring distributed snow depth using SDV maps from 5 March 2015 (r = 0.80) and 16 April 2016 (r = 0.91), respectively. Errors in the inferred depth map generated using the 16 April 2016 SDV pattern were caused by differences in snow accumulation on the West and East portions of the domain between WY2014 and WY2016 (Figure 2.6, middle row). Inferred depth maps generated using SDV patterns from 5 March 2015 and 1 April 2017 were too homogeneous, thereby overestimating snow depth in regions with shallow snow and underestimating snow depth in regions with deep snow (Figure

2.6). Sources of snow pattern disagreements between these dates are discussed more in Section 2.8.1.

We inferred distributed snow depth across the large Tuolumne study domain, which included regions both above and below the seasonal snowline and contained large variations in snow extent, mean snow depth, and snow depth variability. We found that the ratio between snow depth mean and standard deviation (μ_d/σ_d) changed nonlinearly with melt or snow extents as snow depth mean was often greater-reduced than snow depth standard deviation. This is typically true in complex terrain since melt patterns are spatially different than snowfall patterns, resulting in patchy snow-cover after melt with relatively large standard deviations with respect to the mean. However, the relationship between snow-covered area and μ_d/σ_d followed a predictable trajectory (Figure 2.7). Therefore, any two ALS observations from seasons with typical snowfall and similar snow extents often had similar μ_d/σ_d, and as a result, more-similar SDV patterns. The trajectory of the relationship shown in Figure 2.7 suggested that snow patterns at a period with less-similar snow conditions could be adjusted to infer a snow pattern on a different date provided information about how snow-cover and/or μ_d/σ_d varied. This is discussed in Section 2.8.3.

2.7.2. Snowmelt sensitivity to inferred depth map accuracy

The snowmelt simulation initiated with ALS-observed 7 April 2014 snow depth (baseline simulation) agreed with melt-season ALS observations of snow depth (Figure 2.8). Simulated snowmelt initiated with 7 April 2014 inferred depth maps generated from SDV patterns with high pattern repeatability (r) approached the baseline simulation (Figure 2.9a). Inferred depth maps generated using SDV patterns with reduced pattern repeatability ($r < 0.80$) trended towards more homogenous snow conditions, thereby increasing early-season snowmelt, decreasing late-

season snowmelt, and reducing the length of the snowmelt season. Many of the inferred depth maps generated in Section 2.7.1 were from regressions that were non-zero at the minimum SDV value (SDV = $-\mu_d/\sigma_d$). These inferred depth maps prescribed a constant snow depth to all snow-free gridcells in the pattern, resulting in simulations that exceeded daily snowmelt of 8000 acre-feet per day as snow was melting at low elevations that should have been snow-free on 7 April 2014 (Figure 2.9a). Snowmelt simulated from the inferred depth map using the SDV pattern from 16 April 2016 (r = 0.91) best-matched the late-season snowmelt of the baseline simulation, while snowmelt simulated from the inferred depth map generated from the 1 April 2017 SDV pattern (r = 0.83) agreed best between early-April and mid-May. Snowmelt simulated from the inferred depth map using the SDV pattern from 5 March 2015 (r = 0.80) diverged from the baseline simulation for most periods of the snowmelt season. Although simulated snowmelt from inferred depth maps of reduced pattern repeatability (r) diverged from the baseline simulation, total seasonal snowmelt (Figure 2.9b, maximum) was instead a function of snow volume percent error which was sometimes best for simulations of only moderate (r ≈ 0.75 to 0.85) pattern-repeatability as snow depth errors were balanced to different degrees (Figure 2.4b).

2.7.3. Interannual snow pattern repeatability

Mean absolute error (Figure 2.5) and deviations from the baseline snowmelt simulation (Figure 2.9) were clearly related to the pattern repeatability (r) between the SDV pattern on 7 April 2014 and the SDV pattern used to generate each inferred depth map. Using the library of ALS observations, we found that r varied substantially for all possible comparisons of two SDV patterns from different water years (Figure 2.10, all pairs). For the full Tuolumne study domain, the interquartile range (IQR) of *r* varied from 0.63 to 0.82, with a median of 0.74. The IQR of r for the upper-elevation subdomain varied between 0.57 and 0.81, with a median of 0.70. In both the full Tuolumne domain and upper-elevation subdomain, r between SDV patterns from 16

August 2017 and periods closer to peak-snowpack timing in different water years were smaller, and statistically different, than r of all other comparisons of two SDV patterns (Figure 2.10, all pairs, outliers). This was not surprising given the abnormally deep snow depth in WY2017 (~209% of normal) and the extreme late-season timing (117 days after peak-snowpack timing at the Dana Meadow snow pillow) of the 16 August 2017 observation.

The distribution of pattern repeatability (r) for SDV maps from different water years with similar proximity to peak-snowpack timing (best timing; within 14 days) was better, and statistically different ($p < 0.001$), than the distribution of r for all comparisons of two SDV maps (all pairs) (Figure 2.10). This was true for both the full Tuolumne domain and the upper-elevation subdomains. As compared to all pairs of two SDV maps, median r for the comparisons with similar timing improved by 0.12 (from 0.74 to 0.86) and 0.15 (from 0.70 to 0.85) for the full Tuolumne domain and upper-elevation subdomain, respectively. Most of the outlying comparisons (Figure 2.10, best timing, outliers) in both domains were comparisons between SDV patterns in abnormally deep WY2017 and abnormally shallow WY2015. The remaining outliers compared SDV patterns from periods near peak-snowpack timing in abnormally shallow WY2015 with SDV patterns from WY2014 and WY2016. Therefore, proximity to peak-snowpack timing at a point (such as the Dana Meadows snow pillow) may not always be indicative of domain-wide snow conditions, particularly if located at points where seasonal variations in temperature, precipitation, wind, vegetation, and other drivers of snow evolution are not representative of conditions (snow extent, magnitude, etc.) across larger areas. Further, snow patterns on reference dates, such as April 1st, which is commonly used to represent peak-SWE timing in the Western United States (e.g., Cayan, 1996; Kapnick and Hall, 2012; McCabe and Dettinger, 2002; Mote, 2006; Mote et al., 2005), may be less-repeatable in water years where

deviations from typical temperature and precipitation influence the amount and phase (snow, rain, or mixed-precipitation) of precipitation. This is particularly true given expected changes in peak-SWE timing with a warming climate (Margulis et al., 2016b; Rasmussen et al., 2011).

Comparisons of SDV patterns in different years with similar snow extents (best extent; within 75% agreement) also had pattern repeatability (r) that was better, and statistically different ($p < 0.001$), than r for all comparisons in different water years (Figure 2.10). However, the spread in r for SDV comparisons of similar extents (within 75% agreement) was not statistically different than the spread in r for comparisons of similar timing (within 14 days) for either the full Tuolumne domain or upper-elevation subdomain. Relative to all comparisons of two SDV patterns, median r for the comparisons with similar snow extents improved by 0.13 (from 0.74 to 0.87) and 0.16 (from 0.70 to 0.86) for the full Tuolumne domain and upper-elevation subdomain, respectively. Improvements in r were more apparent for the full Tuolumne subdomain than for the upper-elevation subdomain. This was expected since the upper-elevation subdomain was predominantly snow-covered for longer portions of the snow season, thereby decreasing the utility of snow-extent as a matching criterion. Comparisons with statistically worse r values (outliers) compared SDV patterns with similar extents between WY2015 and WY2017.

The median pattern repeatability (r) of the comparisons of greatest r (analogues from different snow years) was 0.89 and 0.88 for the full Tuolumne domain and upper-elevation subdomain, respectively, and ranged little (IQR < 0.04). Analogues had pattern repeatability that was greater and statistically different than comparisons of similar timing (within 14 days) for both the full Tuolumne domain and upper-elevation subdomain ($p < 0.001$). However, comparisons of analogues were not statistically different from comparisons of similar snow extents (within 75% agreement). Outlying r values (Figure 2.10, analogue, outliers) compared SDV patterns from 5

March 2015, 9 April 2015, 3 March 2017, 17 July 2017, 27 July 2017, and 16 August 2017 with SDV patterns form other dates.

The analysis for SDV patterns of similar timing (within 14 days) included pairs of SDV patterns from different water years with average seasonal timing that ranged from before peak-snowpack (~20 days before) to the late melt season (~100 days after peak-snowpack). However, no statistically significant relationship existed between pattern repeatability (r) and the timing (days since peak-snowpack) at which these SDV patterns were observed. In other words, r was not sensitive to whether comparisons of two SDV patterns of similar timing were from periods before, near, or after peak-snowpack timing. The same was true for SDV patterns of matching extents, which ranged from ~10% to ~100% spatial coverage.

The full Tuolumne subdomain, which accumulated and melted snow in large part as a function of elevation, tended to have larger pattern repeatability (r) than the upper-elevation subdomain (Figure 2.10). This does not suggest that inferred depth maps generated in the upper-elevation subdomain were less-accurate than inferred depth maps in the full Tuolumne domain. Instead, the upper-elevation subdomain was weighted more-heavily on finer-scale processes (snow drifting, preferential deposition, avalanching, etc.), and was therefore different from the full Tuolumne domain which was driven largely by snow extents and elevation. In fact, very little difference in MAE was seen for inferred depth maps generated across the full Tuolumne domain and upper-elevation subdomain (Figure 2.11) using observations of similar percent-coverage.

2.8. Discussion

In Section 2.7, we demonstrated how spatially distributed snow depth could be inferred at a period near peak-snowpack timing using normalized snow patterns that increased in agreement (r) for periods of similar timing and/or snow extents. In the following sections, we expand on

these results by discussing sources of snow pattern disagreements in snow seasons with less-typical precipitation (Section 2.8.1) and the sensitivity of inferred depth maps to the location and number of snow depth observations (Section 2.8.2). We finish by hypothesizing the effects that this work may have on water management practices (Section 2.8.3).

2.8.1. Snow pattern repeatability for abnormal snow seasons

Abnormally shallow WY2015 (~34% normal) and abnormally deep WY2017 (~209% normal) displayed SDV patterns that were consistently different from each other, and sometimes different from similarly timed (within 14 days) patterns in other years (Section 2.7.3). Snow extents for WY2017 were nearly 15% larger than snow extents in WY2015 at peak-snowpack timing. Additionally, WY2015 SDV patterns from nearer to peak-snowpack timing had extents that were more similar to WY2017 SDV patterns from the mid-melt season, thus differing in the drivers (accumulation and melt) of snow distribution. This suggested that combining matching criteria (such as seasonal timing and snow extents), although more restrictive, could better identify similar snow patterns.

The ratio between peak-snowpack snow depth mean and standard deviation (μ_d/σ_d) across the full Tuolumne domain on 7 April 2014 disagreed by 9% and -20% as compared to 1 April 2017 and 5 March 2015, respectively, and was worse when calculating μ_d/σ_d in regions of only overlapping snow cover (Table 2.1). This showed that differences in μ_d/σ_d were not solely due to differences in snow extents, but also interannual changes in snow patterns. Although variograms suggested that snow depth variability plateaued at reaches of ~25 m in this region (Currier and Lundquist, 2018), we realized that the spatial reach of snow features could change with differences in precipitation, wind, and/or snow volume as the terrain was smoothed by the snowpack to different degrees (e.g., Filhol and Sturm, 2019; Schirmer and Lehning, 2011). To

test this, we calculated snow depth variograms at the periods nearest peak-snowpack timing in each water year (Figure 2.12) using the native-resolution (3 m) ALS observations within a 1 km^2 region of similar elevation (~2700 – 2800 m) in the upper-elevation subdomain. Variograms illustrated the distance at which the variability of snow depth increased within snow processes (wind drifts, cornices, etc.) and then plateaued as multiple snow features were included and domain-average variability was approached. Results found that gridcells of 25 m resolution (Figure 2.12, dashed line) were adequate to represent the process-scale (~85-90% of the distance at which the variogram plateaued) of snow features in most snow seasons. However, the snow depth variogram on 5 March 2015 plateaued almost immediately at spatial reaches of less than 10 meters while the variogram on 1 April 2017 did not plateau at all between 3 and 60 m distances. Physically, this meant that 7 April 2017 smoothed the local snow landscape to a degree that was unable to be matched by the reduced snowpack volume on 5 March 2015. The correlation length scale of snow processes was therefore different for these two snow seasons and changed how snow features were being resolved within, or across, (a) 25 m gridcell(s). This influenced both relative snow depth variability (μ_d/σ_d) and snow pattern noise.

2.8.2. Inferred depth map sensitivity to observation location and extent

Instead of a strip of observations (Figure 2.2a), we could imagine snow survey-like samples that capture 25 m mean snow depth at random locations around the watershed. Therefore, inferred depth maps were also generated by regressing snow depth on 7 April 2014 and an SDV pattern from a different water year using 5, 10, 25, and 100 randomly located 25 m gridcells (Figure 2.2a). The mean absolute error of the 7 April 2014 inferred depth map typically decreased as the number of observations increased (Figure 2.4a). Even for SDV patterns with high pattern repeatability (r), some level of interannual pattern noise existed in the relationship between snow depth and the corresponding SDV values. Therefore, snow depth mean and

standard deviation inferred by a regression of few observations (e.g., 5, 10, or 25) risked being skewed by this noise. Regressions from more observations (strip or 100 randomly located observations) better captured the central tendency of the relationship between snow depth observations and SDV patterns.

The locations of snow depth observations (Figure 2.2a) were also important for inferred depth maps. For instance, 25 observations from a wide spread of SDV values (e.g., $\pm$2 SDV) were more informative for the regression (Figure 2.2c) than the same number of observations with a tightly grouped spread of SDV values (e.g., $\pm$0.2 SDV). Therefore, snow observation stations, which often oversample snow depth from similar elevations and landcover/terrain types (Groisman and Legates, 1994; Henn et al., 2016a; Hughes et al., 2017; Lundquist et al., 2019), may be poorly suited to generate inferred depth maps. The mean absolute error (MAE) of 7 April 2014 inferred depth maps generated using 100 randomly located gridcells (covering ~0.005% of the domain) was only 0.03 m larger, on average, than the MAE from inferred depth maps generated using the strip of observations (56,866 gridcells, ~3% of the domain). When testing different locations for the 7 April 2014 strip of observations, the mean absolute error of inferred depth maps generated using only 100 randomly located gridcells was sometimes better than inferred depth maps generated using the strip (not shown). This was due to the fact that strips of observations, which were grouped in space, ran the risk of 1) disproportionately observing one type of snow depth (deep/shallow) and a small spread of SDV values, and 2) oversampling snow depth in a region where snow patterns differed by more than average amounts. This was particularly true for the full Tuolumne domain, where snow depth variability is governed largely by elevation.

2.8.3. Applications and future research

Our results are encouraging from an operational standpoint, as patterns from previous seasons could be used to infer distributed snowpack in following seasons. Although the expansive library of ALS observations allowed us to look at snow pattern repeatability across a large domain for various periods of snow seasons, ALS observations of this spatial and temporal extent (47 observations across 7 snow seasons) are unusual. We also acknowledge that most snow-producing storms follow storm-tracks from the southwest in the Tuolumne watershed (Lundquist et al., 2015, 2010), thereby increasing the likelihood of persistent snow patterns as storms interact with the terrain in a repeatable fashion. While common in the coastal mountain ranges of the western U.S., repeatable storm tracks are not guaranteed in all snow-covered mountainous watersheds. Therefore, snow and meteorological pattern repeatability should be investigated in other snow landscapes to test the applicability of this method at global scales. While we focus on pattern repeatability between ALS observations, future pattern repeatability studies could leverage the use of multiple observations from satellite platforms, airborne and UAV instruments, and ground-based methods.

The 7 water years covered by the library of ALS observations (WY2013 through WY2019) amazingly included 1) abnormally shallow WY2015 that was characterized by a snow volume return period (estimated average time between two seasons with similar snow volume) of over 600 years (Margulis et al., 2016b) and was among the shallowest historic snowpacks estimated from tree-ring SWE reconstruction (Belmecheri et al., 2016), and 2) abnormally deep WY2017 with the second-largest April 1 SWE ever recorded in this region (Hedrick et al., 2020). The library of observations allowed us to investigate how extreme years differed from snow distribution patterns from seasons with more typical snowpack. Results found that snow patterns and the process-scale of snow features were less interannually repeatable 1) between seasons

with extreme low and high snowfall (WY2015 and WY2017), and 2) between seasons with high/low snowfall and more typical snowfall. Although not as abnormally snowy as WY2017, the SDV pattern from 24 March 2019 (~179% normal snow depth) had a pattern repeatability (r) of 0.89 versus the 1 April 2017 SDV pattern, and agreed closer than any other SDV pattern from the ALS library. The 24 March 2019 and 1 April 2017 variograms also tended to plateau more gradually than variograms from peak-snowpack timing in other snow seasons (Figure 2.12), suggesting that the snow processes were also similar in these two seasons. Therefore, multi-decadal snow reanalyses that reconstruct snowpack using satellite observations of snow extent (Margulis et al., 2016a), may provide the opportunity to expand the library of snow patterns and include more seasons of less-typical snowfall. We hypothesize that this will enhance our ability to identify snow patterns with good pattern repeatability in extreme snow seasons. This would be particularly valuable since traditional hydrologic forecasting tools may be most uncertain in seasons with less-typical snowfall.

The trajectory of the relationship between snow extents and the ratio of snow depth mean and standard deviation (μ_d/σ_d) (Figure 2.7) suggested that snow patterns in the full Tuolumne study domain could also be inferred by adjusting a snow pattern from a period with less-representative conditions using information about how the pattern would change with melt, snowfall, snow cover, or other forcings. For instance, Sturm and Wagner (2010) found that snow depth mean increased linearly with snow depth standard deviation throughout accumulation. Egli and Jonas, (2009) instead found a power-law relationship between snow depth mean and standard deviation throughout accumulation, with melt-season transformations explained by spatially homogeneous snowmelt. In the full Tuolumne domain, we found that changes in snow depth distributions throughout snowmelt were unable to be explained by homogeneous snowmelt. Instead, the

relationship between μ_d/σ_d and snow extents were best-approximated using a third-order polynomial (Figure 2.7). Future work should investigate how gridcell-by-gridcell SDV values evolve with changes in μ_d/σ_d, snow extents, and the resulting domain-wide snow depth distribution.

Seasonal snow timing and snow-covered extents (which we used as matching criteria) may not always be available, or reliable, for those tasked with managing water resources. For instance, determining peak-snowpack timing may be difficult for periods and watersheds without clear separations between snowpack accumulation and melt. Determining snow extents using space-based platforms would also be challenging in forested regions where canopy obscures observations. Although widely dispersed snow depth observations were preferable as compared to observations that were grouped in space, a strip of observations from airborne or space-based platforms may be more feasible to organize from an operations standpoint. However, SDV patterns from previous years could nevertheless be used to inform where observations occur (at points with a wide spread in SDV values) during periods of similar timing and/or snow extents.

2.9. Conclusions

Snow depth patterns over mountainous regions are generated by the superposition of spatially unique snow accumulation and snowmelt patterns. These patterns contribute to the timing of downstream water resources. Fortunately, snow distribution patterns are often repeatable as the snowfall deposition and snowmelt patterns are driven by consistent interactions between the local meteorology and terrain. Here, we tested 1) if repeatable snow patterns emerged at fine spatial resolutions (25 m) from an unprecedented library of snow depth observations (47 ALS observations over 7 snow seasons) over the California state Tuolumne watershed, and 2) how

these patterns could be used to infer snow depth in regions, and during periods, without distributed snow depth observations.

Normalized snow depth patterns from different years varied as a function of snow accumulation and melt processes with a spatial coefficient of correlation (r) ranging from approximately 0.30 to 0.90. Peak-snowpack timing from a snow pillow within the domain and snow-covered extents from the ALS data could be used as matching criteria to identify snow patterns from two different seasons that were more likely to have similar snowpack distribution (median $r > 0.84$). This was true for the full Tuolumne study domain (1650 km^2, 700 – 3900 m elevation) and for a smaller upper-elevation study domain (39 km^2, 2430 – 3230 m elevation). Distributed snow depth was inferred across the full Tuolumne domain by relating snow depth observations and snow patterns (from different years) across only a portion (< 4%) of the domain. As the snow pattern repeatability (r) of the snow pattern improved, the mean absolute error of inferred depth maps and simulated melt-season (April – August) snowmelt also improved. Distributed snow depth inferred on 7 April 2014 using a strip of snow depth observations (on that date) and a snow pattern from a time near peak-snowpack in a different season had a mean absolute error of only 0.22 m and had snow volume that was biased by less than $\pm$10%.

Operationally, historic snow patterns could provide a basis in which to determine ideal locations for snow observations in future snow campaigns. Although we focus on ALS observations at 25 m resolution here, future research should consider the use of various snow depth observations from satellite platforms, airborne and UAV instruments, and ground-based methods, at various spatial resolutions. Future research should also consider the use of long-term

(multi-decadal) snow datasets which may increase the chances of identifying matching snow patterns in water years with atypical snowfall and melt.

2.10. Appendix A: Snowpack distribution analysis

Normal, lognormal, and gamma distributions have all been employed to represent snowpack spatial distributions in mountainous terrain (e.g., Egli and Jonas, 2009; Elder et al., 1991; Jonas et al., 2009; Liston, 2004; Marchand and Killingtveit, 2005; Skaugen, 1999). Each of these theoretical distributions were fit to empirical snow depth observed by ALS (47 collections over 7 snow seasons) across the full Tuolumne domain and subdomain. The fitting algorithm for each distribution optimized (minimized) the maximum difference between the theoretical (normal, lognormal, or gamma) and observed cumulative distribution function (CDF). Lognormal distributions, which were undefined at zero, were transposed by +1m of snow depth at all gridcells prior to fitting. Results found that empirical snow depth distributions observed by ALS converged more-closely with normal distributions than lognormal or gamma distributions, across the full Tuolumne domain and subdomain (Figure 2.13).

Although SDV patterns (Equation 2.1) assumed normally distributed snow depth, non-normal snow patterns could still be used to infer distributed snow depth if the departures from normality were similar. For example, SDV patterns across the full Tuolumne domain were inherently skewed by the prevalence of snow-free gridcells, resulting in large departures from normal distributions in regions corresponding to zero snow depth ($SDV = -\mu_d/\sigma_d$). Assuming that domain-wide snow depth between two observations have different means, but retain identical zero-snow prevalence and an identical ratio between mean snow depth and snow depth standard deviation (μ_d/σ_d), the resulting SDV probability distributions will be identical. Now, assuming a change in the ratio between mean snow depth and snow depth standard deviation (μ_d/σ_d) and/or zero-snow prevalence, the SDV patterns between the two ALS snow depth observations

will no longer agree, resulting in mismatching SDV probability distributions. Even if SDV probability distributions generated from two different ALS snow depth observations are identical, SDV patterns could still differ if the gridcell-by-gridcell placement of deep/shallow snow depth disagrees. Although a variety of metrics were tested, only Pearson's coefficient of correlation (Equation 2.4) was representative of all the pattern errors mentioned above. In fact, the coefficient of correlation between SDV patterns on different dates was strongly related to the mean absolute error of the inferred depth map, where the trajectory of the relationship was controlled by the mean snow depth at the date of the inferred snow depth map (Figure 2.5a).

2.11. Appendix B: Snowmelt model

To illustrate the interplay between the heterogeneity of snow depth and melt-season radiation, a simplified snowmelt model was setup wherein 1) melt-season average radiation patterns were generated, 2) the relationship between cumulative radiation and snowmelt was calculated at the Dana Meadows snow pillow, and 3) daily distributed snowmelt was calculated using the average radiation pattern (step 1) and relationship from step 2. We acknowledge that more-robust modeling efforts, which explicitly solve the snowpack energy-balance and resulting snow melt, could be used. However, the simple snowmelt model used here was sufficiently able to illustrate the impact on snowmelt timing from inferred depth maps of various accuracies.

The MicroMet meteorological preprocessing routine (Liston and Elder, 2006a) was used to generate spatiotemporal estimates of incoming longwave and shortwave using a 25 m digital elevation model and 6 km WRF surface fields (air temperature, relative humidity, wind speed, wind direction, and precipitation). Since WRF has noted issues with shortwave radiation in this region (Lapo et al., 2017), shortwave radiation was calculated within MicroMet using basic solar geometry and cloud fractions calculated from air temperature, relative humidity, and dew point temperature (Walcek, 1994). Estimates of longwave, were calculated using the cloud fraction, air

temperature, and the longwave scheme from Iziomon et al. (2003). Daily-average shortwave and longwave were output from the model simulations and then aggregated across the melt season (April – August 2014).

Daily cumulative radiation (shortwave plus longwave) was calculated using MicroMet at daily timesteps (between 7 April and 30 September 2014) for the 25 m gridcell including the Dana Meadows snow pillow. A melt factor was inferred using the relationship between cumulative daily snowmelt (observed) and cumulative daily energy (modeled) at this gridcell. The empirical relationship between cumulative daily energy (modeled) and cumulative daily snowmelt (observed) at the Dana Meadows snow pillow was approximately 10% of the latent heat of fusion for ice. This suggests that approximately 90% of the daily total incoming energy was lost to reflection off the snow surface (in the case of shortwave) or went towards warming the snowpack to isothermal, melting conditions before snowmelt occurred and water percolated through the snowpack.

Daily total radiation (shortwave plus longwave) was distributed in space and time from Dana Meadows to each gridcell across the full Tuolumne domain using,

$$Q_i^t = \left(\frac{Q_i}{Q_{DM_i}}\right) Q_{DM}^t, \tag{2.5}$$

where Q_i is daily-average radiation at gridcell *i* and Q_{DM_i} is daily-average total radiation at the gridcell containing the Dana Meadows snow pillow. Q_{DM}^t is daily cumulative modeled radiation for the gridcell containing the Dana Meadows snow pillow at day *t.* As an example, gridcells where daily-average radiation was 10% larger than daily-average radiation at the Dana Meadows snow pillow will always be prescribed daily radiation that is 10% larger than the radiation modeled at Dana Meadows. Using the relationship between cumulative energy and cumulative snowmelt at Dana Meadows, snowmelt was calculated at each model gridcell at each daily

timestep throughout the melt season. For timesteps and gridcells where cumulative melt exceeded the remaining SWE, melt at that timestep was constrained to the remaining SWE, and was assumed to be snow-free for all remaining timesteps. Daily domain-wide snowmelt was then summed for each timestep.

We acknowledge that while this method was computationally efficient, we assumed that domain-wide snow density was constant, when in fact, this is expected to vary in space. Additionally, we assume that all snow depths will melt the same amount per given unit of energy when deeper snowpacks will take more energy to warm than shallower snowpacks. Finally, we acknowledge that the spatial pattern of shortwave and longwave energy will change throughout the melt season with changes in lapse rates and solar angles. We expect that more robust representations of snowmelt to increase the sensitivity of snowmelt to changes in distributed snow depth. However, the snowmelt model initiated with ALS-observed 7 April 2014 snow depth tracked ALS observations of snow depth throughout snowmelt closely (Figure 2.8). Although the simplified model neglected snow mass lost to sublimation or evapotranspiration, snow depth evolution showed that radiation effects were the dominant driver of snow depletion, as has been shown in other Sierra Nevada studies (e.g., Marks and Dozier, 1992). While more robust modeling efforts could be employed, the baseline simulation initiated with ALS-observed 7 April 2014 snow depth was sufficient to illustrate how errors in the7 April 2014 inferred depth maps would influence modeled snowmelt.

Table 2.1. Snow pattern agreement and the accuracy of 7 April 2014 inferred snow depth generated using SDV patterns from three different dates (rows). r is the coefficient of correlation between SDV patterns on 7 April 2014 and the SDV pattern from each date. MAE and V_{err} are the inferred depth map mean absolute error (Equation 2.3) and volume percent error (Equation 2.2), respectively. ΔSCA is the difference in snow covered area as compared to 7 April 2014. $\Delta\left(\frac{\mu_d}{\sigma_d}\right)$ is the difference (in percent) in the ratio between snow depth mean and standard deviation on 7 April 2014, and each date (row), where values in parenthesis are the same calculation in only regions of overlapping snow extents. Min(d^*) is the minimum inferred snow depth on 7 April 2014.

Date	**r**	**MAE [m]**	V_{err}	ΔSCA	$\Delta\left(\frac{\mu_d}{\sigma_d}\right)$	**Min(d^*) [m]**
16 Apr 2016	0.91	0.22	-6%	-4%	-8% (-2%)	0.06
1 Apr 2017	0.83	0.30	-9%	+5%	+9% (+17%)	0.00
5 Mar 2015	0.80	0.32	+2%	-9%	-20% (-21%)	0.39

Table 2.2. Statistics corresponding to Figure 2.10. The minimum presented in this table represents the minimum whisker of the boxplot, excluding outliers.

	Domain	*No. of combos.*	*Median r*	*IQR*			*Min.*	*Max.*	*No. of outliers*
				Lower	*Upper*	*Range*			
All pairs	Full	901	0.74	0.63	0.82	0.19	0.33	0.92	6
	Subdomain	901	0.70	0.57	0.81	0.24	0.22	0.91	13
Best timing	Full	39	0.86	0.82	0.87	0.05	0.74	0.92	4
	Subdomain	39	0.85	0.76	0.88	0.12	0.64	0.91	2
Best extent	Full	38	0.87	0.84	0.90	0.06	0.80	0.92	3
	Subdomain	43	0.86	0.79	0.88	0.09	0.73	0.90	2
Analogue	Full	47	0.89	0.87	0.90	0.03	0.82	0.92	2
	Subdomain	47	0.88	0.86	0.90	0.04	0.80	0.91	6

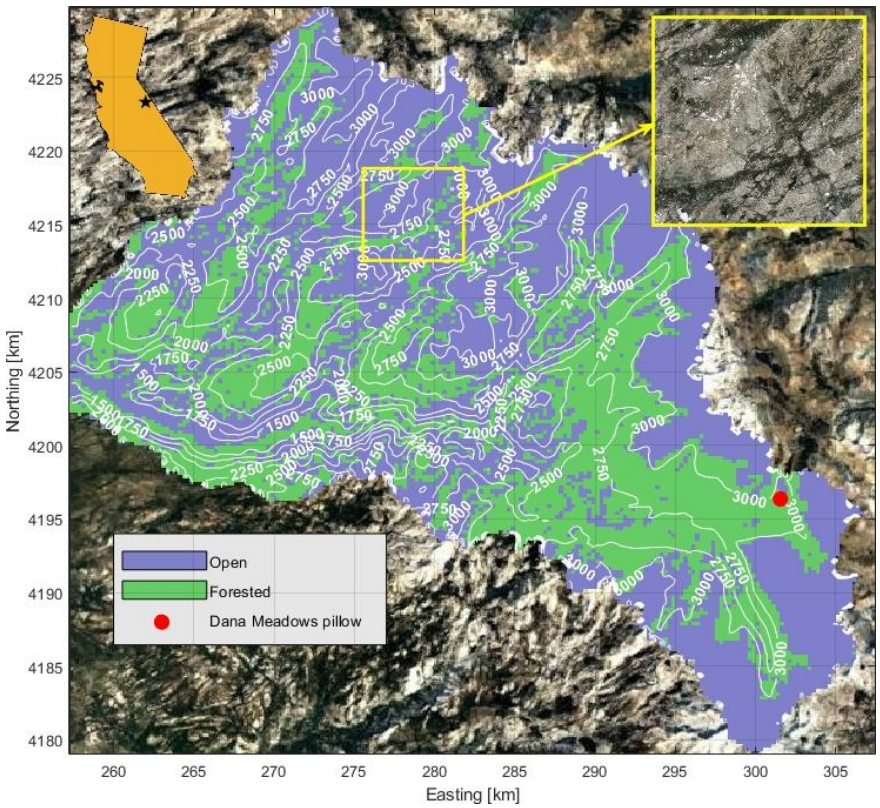

Figure 2.1. Tuolumne, CA study domain (black star on orange California state delineation). Green gridcells represent forest cover (ALS-observed vegetation greater than 5 m in height), and white contours represent elevation (m; defined from the ALS snow-free observation). The red dot indicates the location of the Dana Meadows snow pillow. Imagery from the National Agriculture Imagery Program (NAIP) is shown for the surrounding regions and for the upper-elevation subdomain (yellow boxed region).

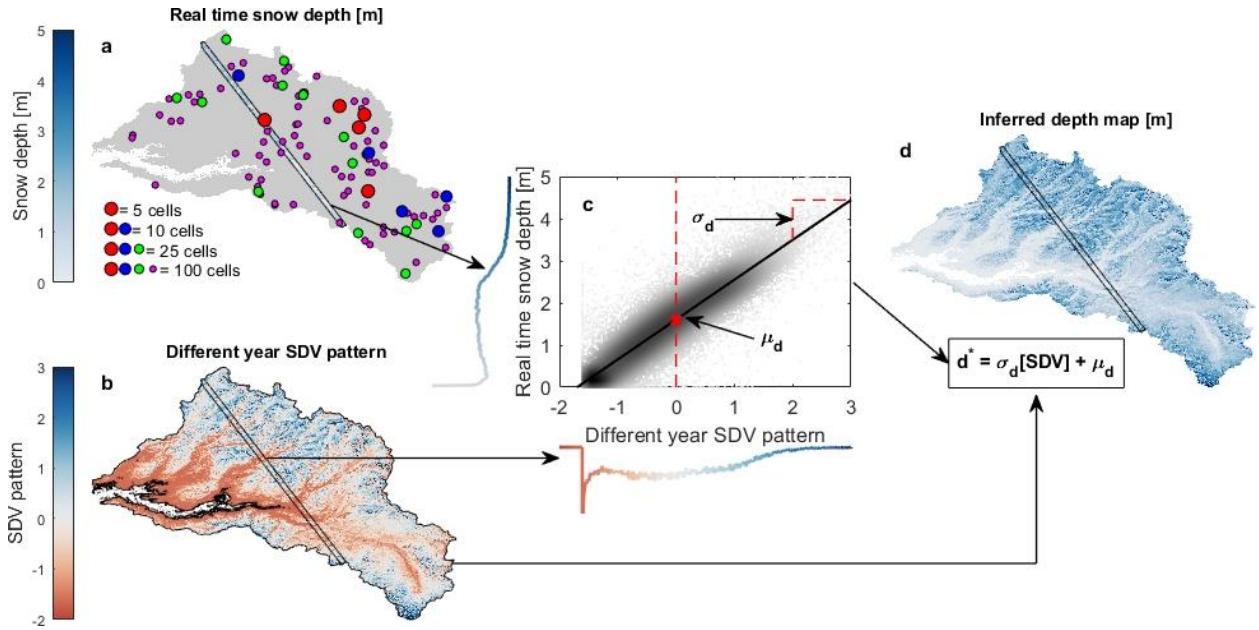

Figure 2.2. Snow depth spatial extrapolation workflow. Snow depth subsampled across a portion of the Tuolumne study domain (a) and the corresponding SDV pattern from a different snow season (b) were regressed to estimate domain-wide snow depth mean (μ_d) and standard deviation (σ_d) (c). Using the snow pattern, μ_d, and σ_d, an inferred depth map was calculated (d). Inferred depth maps were generated on 7 April 2014 using a strip of snow depth observations (example shown here) and observations (a) from 5 (red dots), 10 (red and blue dots), 25 (red, blue, and green dots), and 100 (all dots) randomly located gridcells.

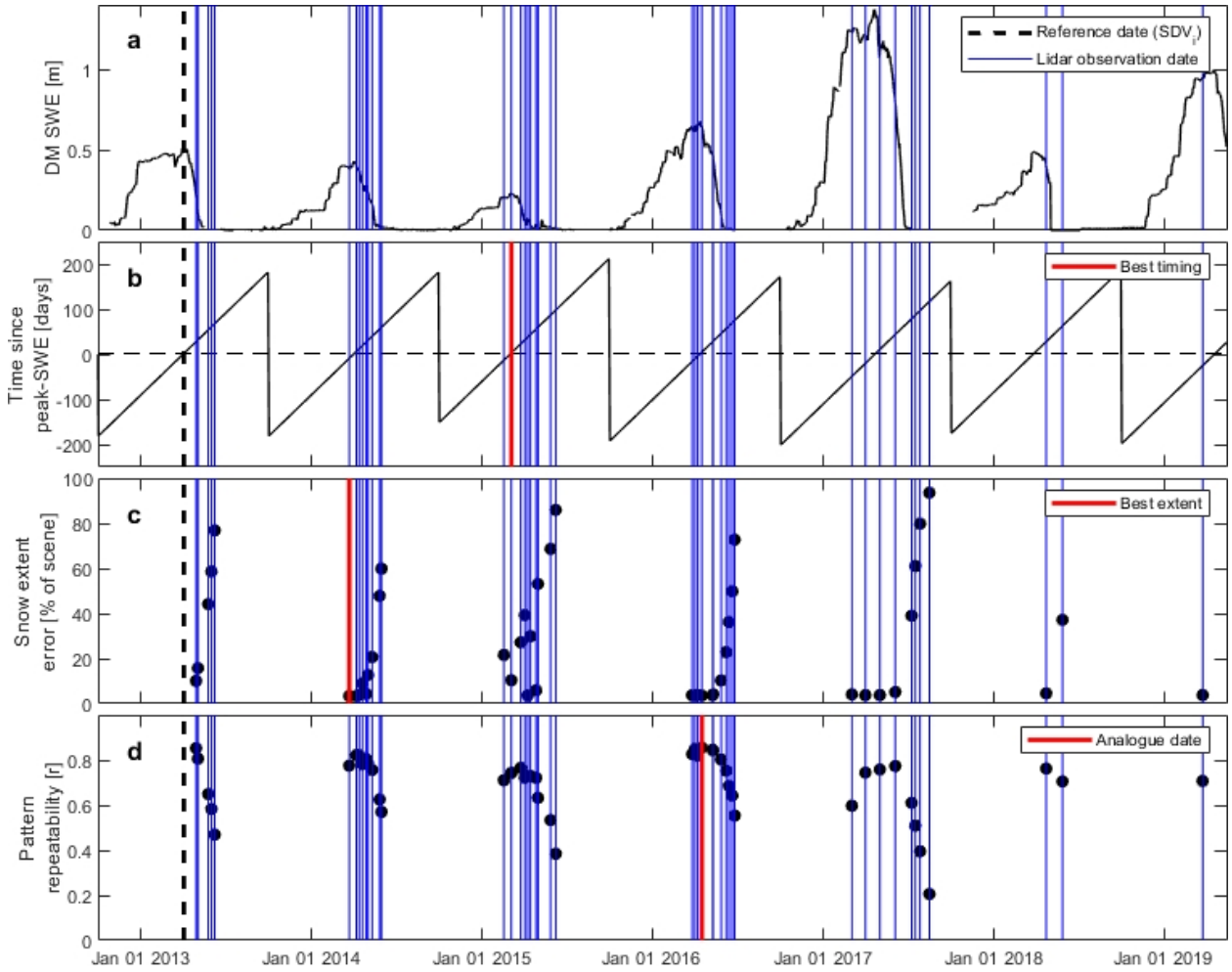

Figure 2.3. ALS observations (blue, vertical lines) are referenced in time (x-axis) versus SWE evolution (a) and days since peak-SWE timing at the Dana Meadows (DM) pillow (b). Using the SDV map from 3 April 2013 in the upper-elevation subdomain (black, vertical dashed line), snow extent error (c) and pattern repeatability (d) were evaluated at every other ALS observation date (black points). The ALS observation in a different water year with the timing, extent, and repeatability (r) that best corresponded to the reference SDV map are shown by the vertical red lines (subplots b, c, and d). This process was repeated using each ALS observation as the reference.

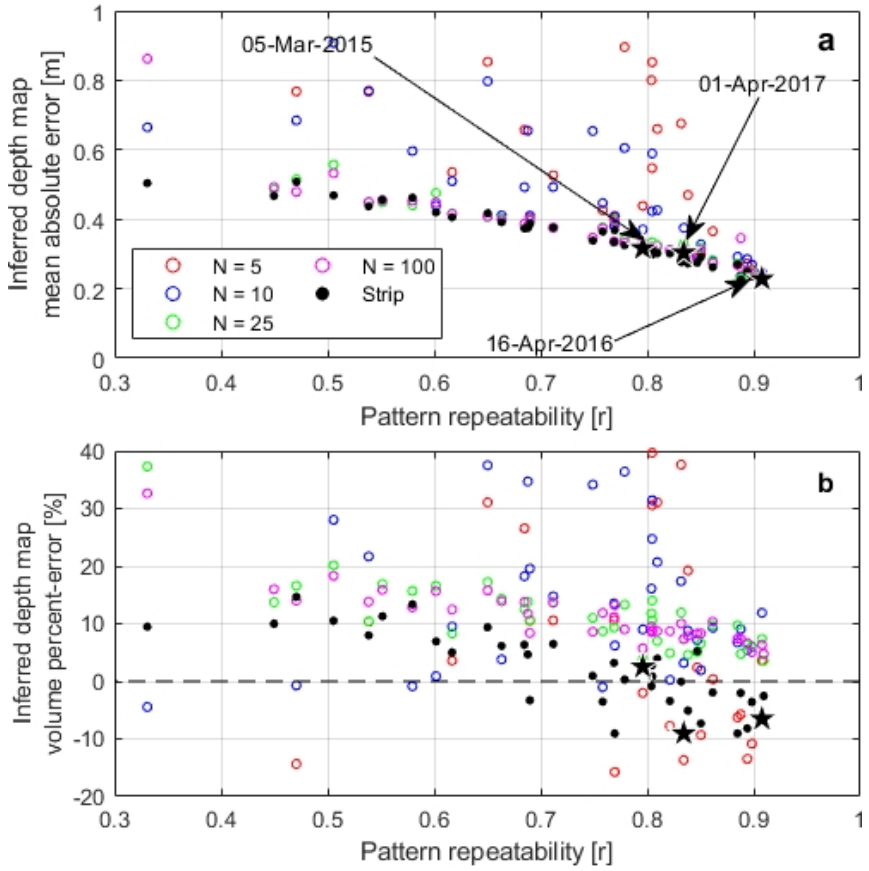

Figure 2.4. Mean absolute error (a) and volume percent-error (b) of the inferred depth maps generated using a strip of observations on 7 April 2014 and SDV patterns from other water years (black dots). Metrics (y-axes) are plotted versus the pattern repeatability (r) between the SDV pattern on 7 April 2014 and the SDV pattern from the different water year. Colored dots represent inferred depth maps generated using 5, 10, 25, and 100 randomly located observations. Starred points with labels are the primary focus of Section 2.7.1.

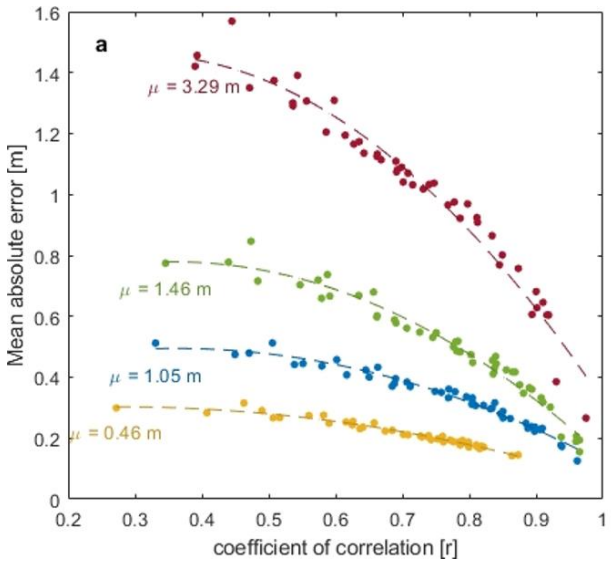

Figure 2.5. Mean absolute error of inferred depth maps plotted versus the spatial coefficient of correlation between the SDV maps used to subsample snow depth observations (Figure 2.2a) and define the SDV pattern (Figure 2.2b). Inferred depth maps are generated on four different dates (colors) using SDV patterns from all other dates. Mean snow depth of the four different dates are presented in the plot near the corresponding data.

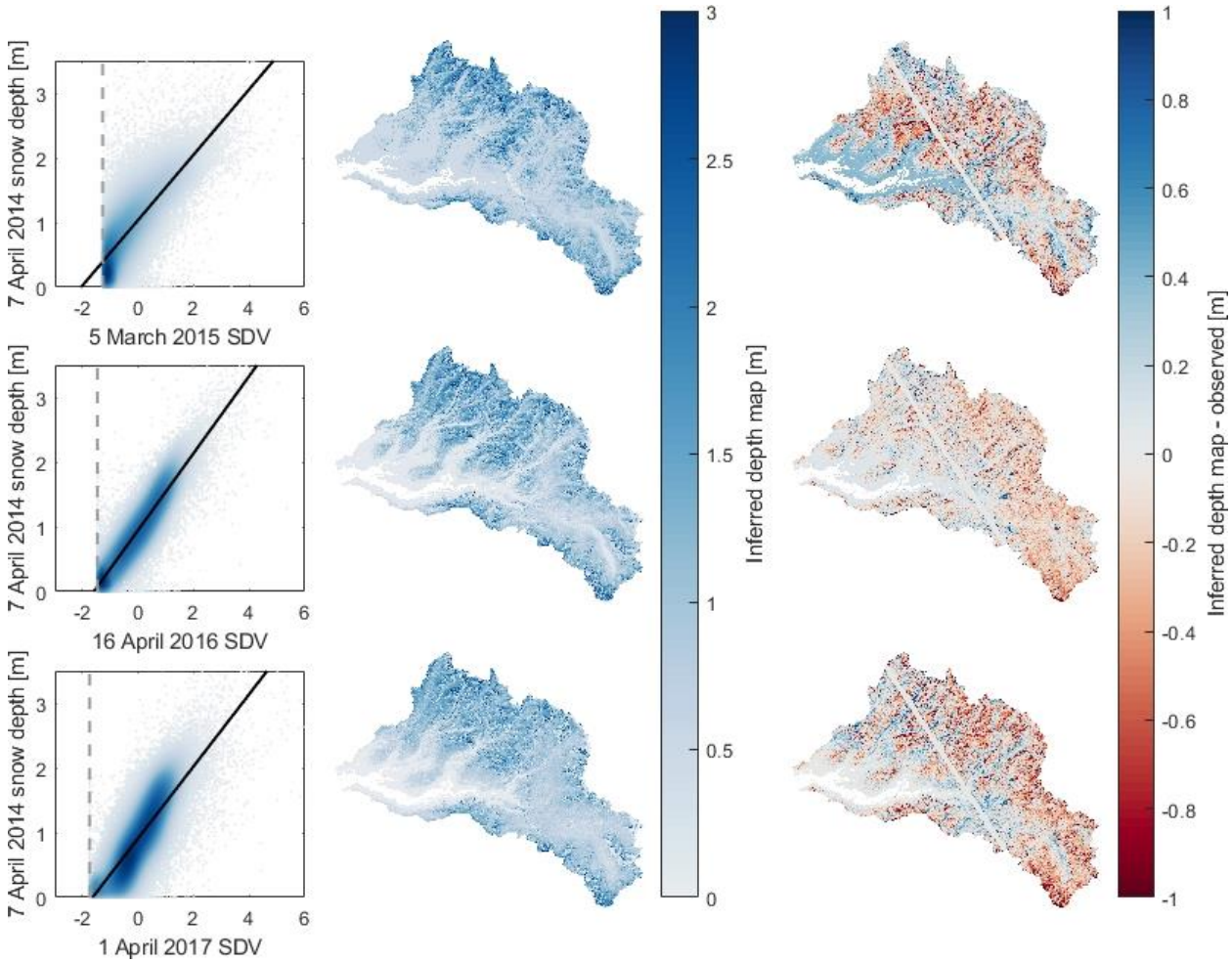

Figure 2.6. Regressions (left column) between the strip of snow depth observations on 7 April 2014 and the corresponding SDV values from three different dates (rows, x-axis labels). The inferred depth maps are shown in the middle column with the snow depth errors for each inferred depth map in the rightmost column. Vertical dashed lines in the scatterplots (left column) represent the minimum SDV value for each SDV pattern. Darker points in the scatterplots represent higher point density.

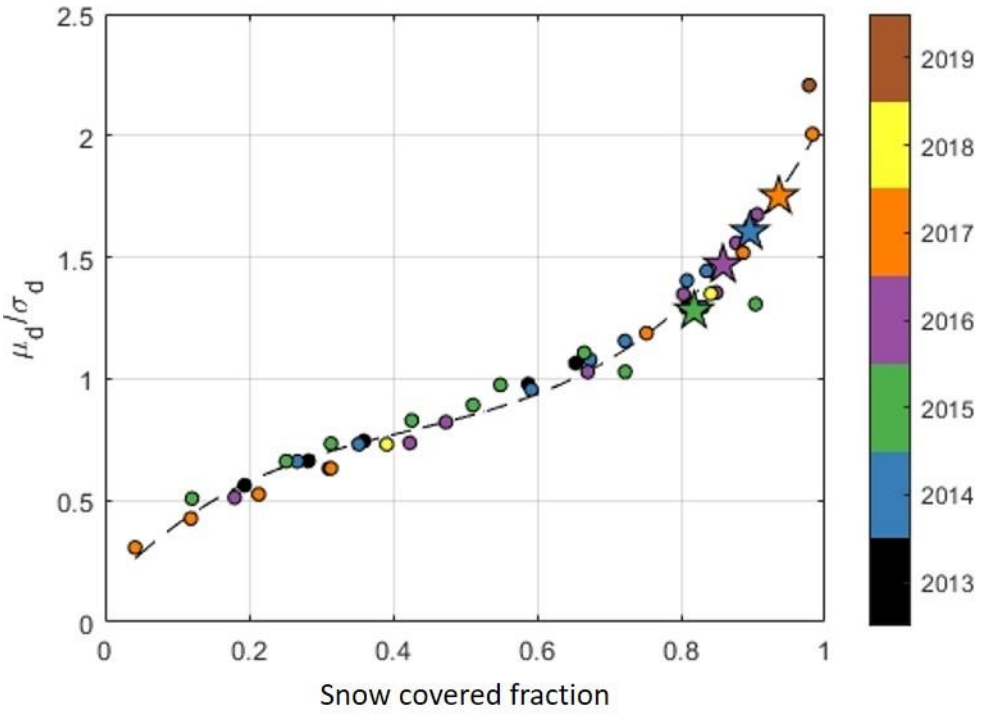

Figure 2.7. Relationship between snow covered fraction and the ratio between snow depth mean and standard deviation (μ_d/σ_d) across the full Tuolumne study domain for each ALS observation (dots), colored by water year. The approximate fit is shown by the dashed line. Starred markers correspond to 7 April 2014, and the dates of SDV patterns focused on in Section 2.7.1 (5 March 2015, 16 April 2016, and 1 April 2017).

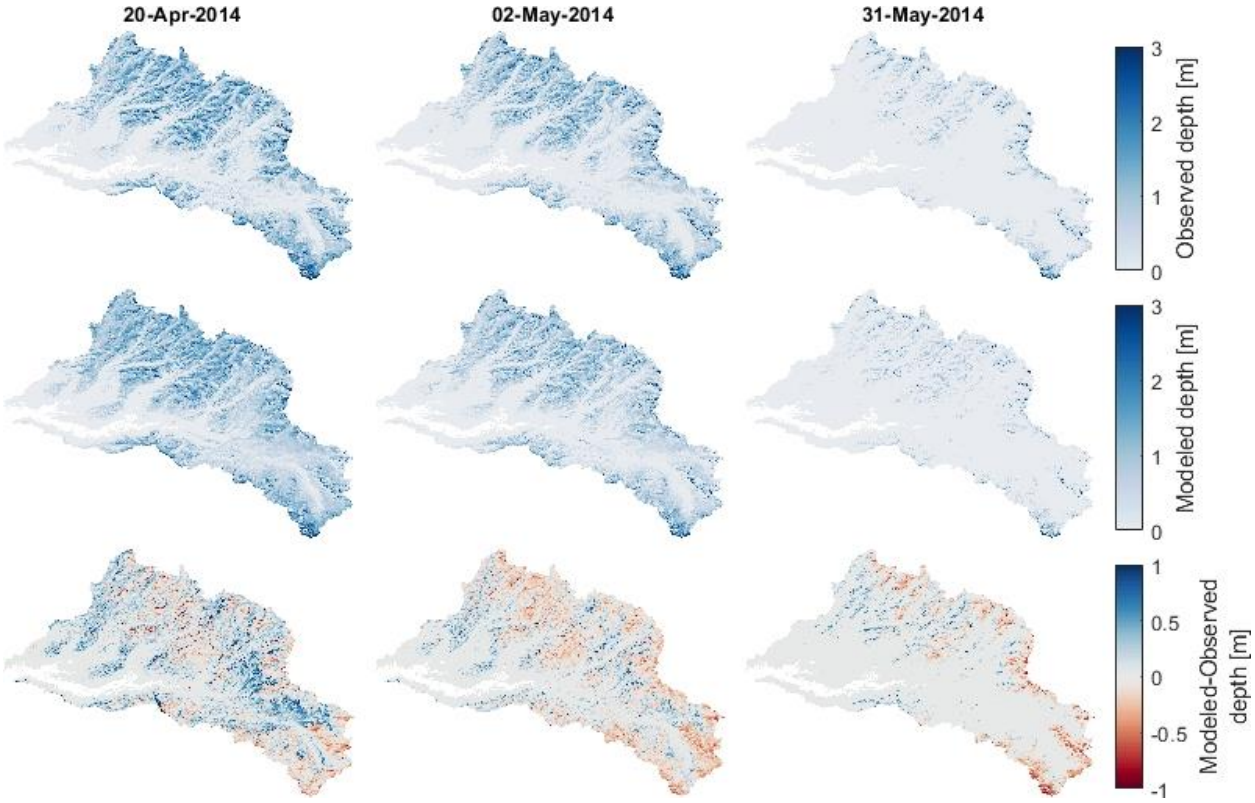

Figure 2.8. LS-observed (top row) and modeled (middle row) snow depth for three different dates (columns) in the 2014 melt-season. The modeled simulation was initiated with 7 April 2014 snow depth from the ALS observation. The bottom row shows the errors in modeled snow depth (modeled-observed).

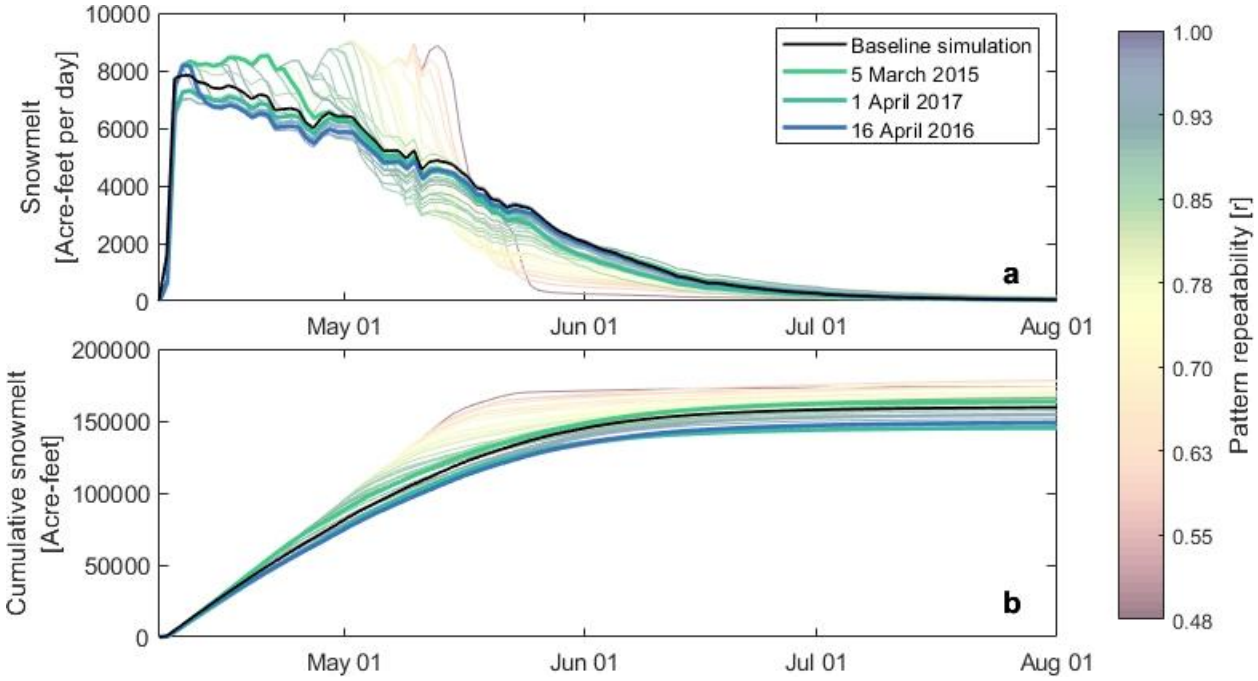

Figure 2.9. Daily snowmelt summed across the full Tuolumne study domain between 7 April 2014 and 30 September 2014 (a), initiated with 7 April 2014 inferred depth maps generated using SDV patterns with different pattern repeatabilities (r; with respect to 7 April) (colors). Results are shown versus a baseline simulation (black line) initiated with the 7 April 2014 ALS snow depth. The cumulative snowmelt from subplot a is shown in subplot b. Thicker lines with saturated colors correspond to the three inferred depth maps focused on in Section 2.7.1 (legend).

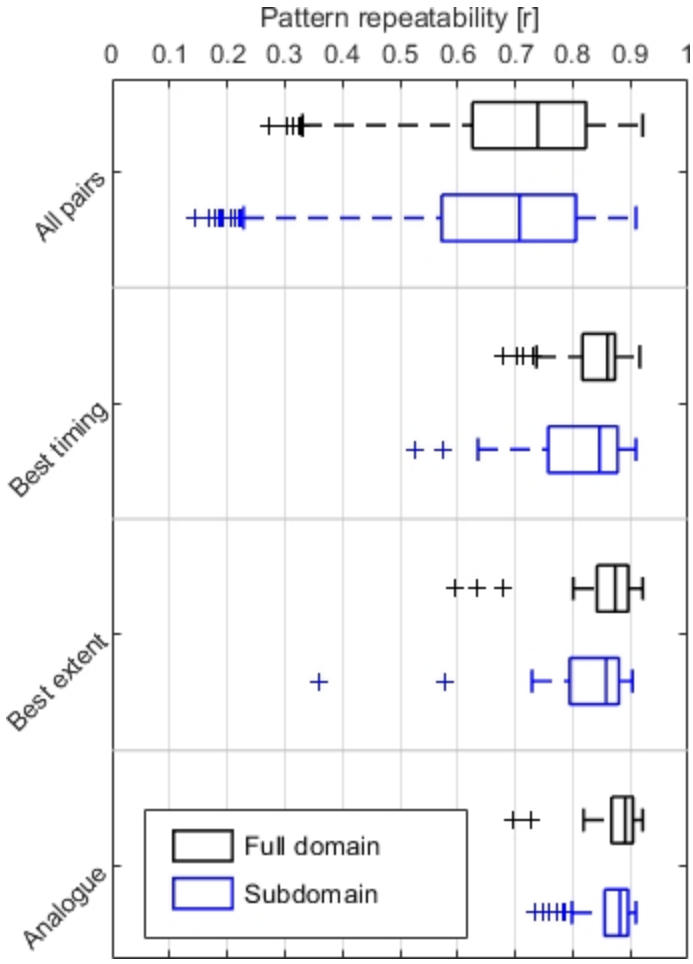

Figure 2.10. Pattern repeatability (r) between SDV maps from different water years including: all possible comparisons of two SDV patterns (all pairs), comparisons of similar seasonal timing (best timing; within 14 days), comparisons of similar snow extents (best extent; within 75% agreement), and the comparisons of highest pattern repeatability (analogue). Results are shown for the full Tuolumne domain (black) and upper-elevation subdomain (blue). Statistics for this plot are presented in Table 2.2.

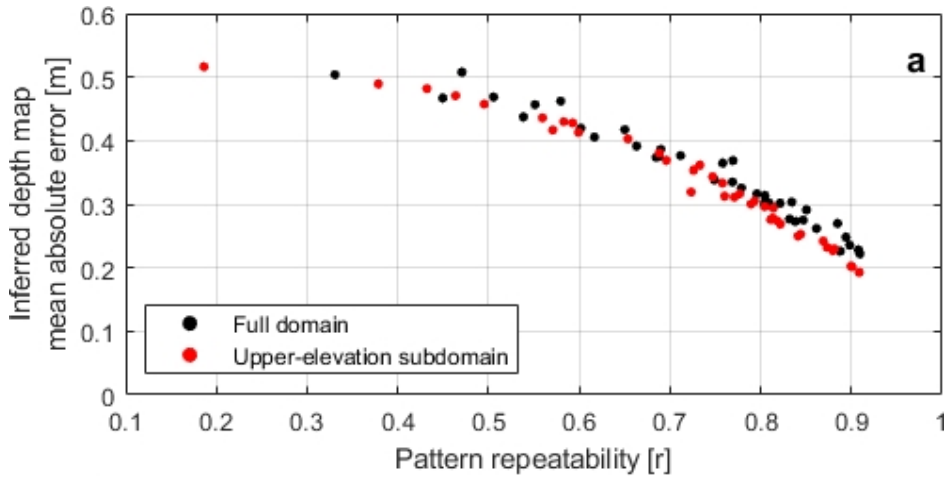

Figure 2.11. Mean absolute error (a) of the inferred depth maps generated using a strip of observations (~3 % coverage) on 7 April 2014 and SDV patterns from other water years in both the full Tuolumne domain (black dots) and upper-elevation subdomain (red dots). Mean absolute error (y-axis) is plotted versus the pattern repeatability (r) between the SDV pattern on 7 April 2014 and the SDV pattern from a different water year used to generate each inferred depth map.

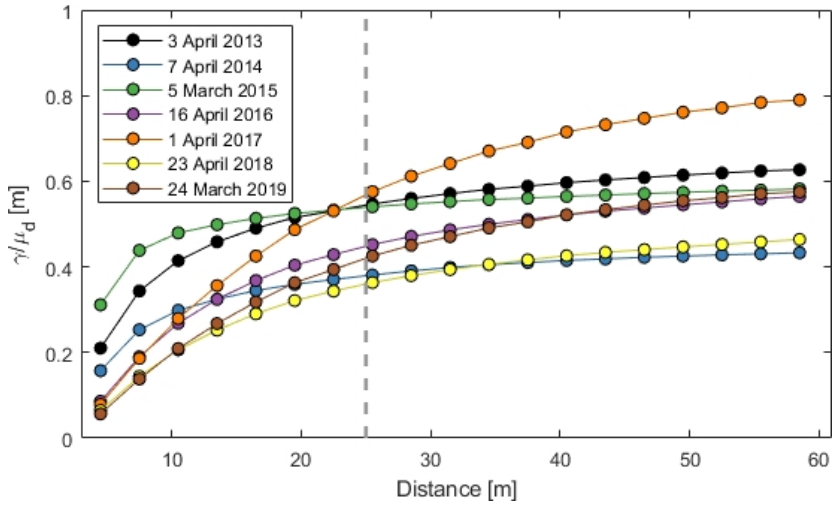

Figure 2.12. Snow depth semivariance divided by mean snow depth (γ/μ_d) for the 3 m ALS observations in the upper-elevation subdomain at periods near peak-snowpack timing in 7 different snow seasons (legend). The vertical dashed line represents the resolution that ALS observations were regridded to in this experiment (25 m).

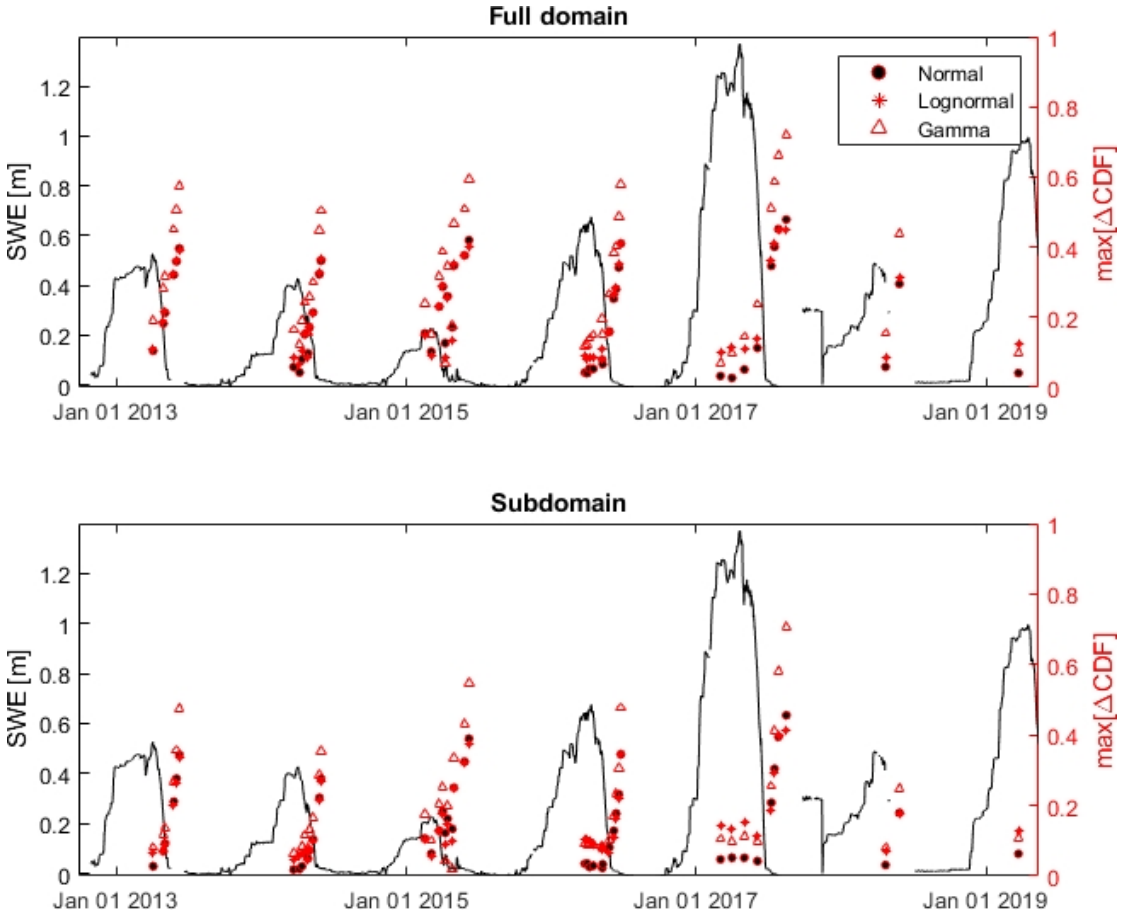

Figure 2.13. The maximum departure between the empirical CDF and fitted normal, lognormal, and gamma distributions on each ALS observation date (scatter points, right y-axis) for the full Tuolumne domain (top) and upper-elevation subdomain (bottom). For reference, scatter points are plotted versus observed SWE evolution at the Dana meadows snow pillow (SWE, black line, left y-axis).

Chapter 3 : Downscaling snow deposition using historic snow depth patterns: diagnosing limitations from snowfall biases, winter snow losses, and interannual snow pattern repeatability

J.M. Pflug[1], M. Hughes[2], J.D. Lundquist[1]

[1] Department of Civil and Environmental Engineering, University of Washington, Seattle, WA, USA. [2] National Oceanic and Atmospheric Administration, Earth System Research Laboratories, Physical Sciences Laboratory, Boulder, CO, USA.

Note: This chapter is being revised for resubmission to Water Resources Research.

3.1. Abstract

Repeatable snow depth patterns have been identified between years with similar meteorological characteristics in many regions. This suggests that snow patterns from previous years may be used to downscale snow deposition as a substitution for mismodeled or unrepresented snow accumulation processes. Here, we tested a popular pattern-based snow deposition downscaling routine which assumes 1) a spatially consistent relationship between snow deposition and snow depth, 2) interannual snow pattern repeatability, and 3) unbiased coarser-resolution snowfall. We investigated these assumptions, and future avenues for improvement, in water-year 2014 over the California Tuolumne River Watershed. 6 km snowfall from an atmospheric model was downscaled to 25 m resolution using 1) a common terrain-based downscaling method, and 2) pattern-based downscaling using snow depth patterns from 7 different years. In this region, snow depth patterns were influenced by winter snow ablation, resulting in pattern-based snow deposition and simulated snow depth that was too spatially heterogeneous. However, snow depth simulated using terrain-based downscaling was too spatially homogeneous, and less spatially correlated ($r = 0.27$), than simulations with pattern-based downscaling using snow depth patterns from the simulation season ($r = 0.76$), or from a different year ($r = 0.52$). For all downscaling experiments, the spread and magnitude of peak-snowpack snow depth errors were driven more by atmospheric model snowfall biases than different downscaling methods. In order of most- to least-importance, future research should focus on bias-correcting coarse-scale snowfall estimates, correcting snow deposition patterns for winter snow ablation, and identifying the historic periods of most-similar snow accumulation.

3.2. Introduction

The timing and magnitude of mountain snowmelt is of vital importance for flood management, hydropower generation, local ecology, and recreation (Lundquist and Dettinger, 2005; Mote et al., 2003). During the late-summer period (July – September) , snowmelt is a dominant source of streamflow in many mountainous basins (e.g., Brauchli et al., 2017; Masiokas et al., 2006; Siderius et al., 2013; Vano et al., 2015). This streamflow is driven disproportionately by snow that remains late into the melt season as a result of vegetation, terrain, and wind-enhanced snowfall and snow redistribution (Brauchli et al., 2017; Egli et al., 2012; Luce et al., 1998; Lundquist et al., 2005; Seyfried and Wilcox, 1995). In fact, the spatial heterogeneity of snow accumulation typically controls spring snow depth distribution as much as, or more than, the spatial heterogeneity of incoming energy and the resulting snowmelt (Egli et al., 2012; Luce et al., 1998).

Models struggle to represent the magnitude and spatial heterogeneity of snow water equivalent (SWE) accumulation in mountainous terrain (e.g., Brauchli et al., 2017; Clark et al., 2011; Freudiger et al., 2017). Precipitation is commonly cited as a dominant source of this error (Günther et al., 2019; Raleigh et al., 2015; Wayand et al., 2013). Atmospheric models, which are sometimes used to provide estimates of snowfall, are at coarse resolutions (3-50 km) and are thereby unable to represent finer-scale snow accumulation in mountain terrain (Helbig and van Herwijnen, 2017; Strachan et al., 2016; Winstral et al., 2019). To better represent the heterogeneity of fine-scale snowfall, snow precipitation can be downscaled using methods that account for terrain effects like elevation-driven airflow ascent and the associated increases in condensation (e.g., Bavay and Egger, 2014; Liston and Elder, 2006a), as well as wind-exposure and terrain-driven preferential deposition (e.g., Dadic et al., 2010; Winstral et al., 2013). However, increases in snowfall may not be consistent across all slopes at all times (Minder et al.,

2008; Roe and Baker, 2006) and fine-scale wind speed and wind direction are uncertain (Musselman et al., 2015; Reynolds et al., 2021). A host of spatial-interpolation methods (e.g., Bavay and Egger, 2014; Daly et al., 2008; Liston and Elder, 2006a) have also been used to infer how precipitation is deposited in spaces between observations. Yet, spatial interpolation methods have issues in mountainous regions where precipitation is spatially variable, observations are sparse (Groisman and Legates, 1994; Henn et al., 2016b; Hughes et al., 2017; Lundquist et al., 2019), and considerable amounts of terrain are at elevations higher than precipitation gauges (Lundquist et al., 2019; Wayand et al., 2013). Finally, downscaling methods assume unbiased coarser-scale precipitation, when in mountainous areas, precipitation magnitude and phase from atmospheric models and gridded datasets are often biased (e.g., Liu and Margulis, 2019; Viste and Sorteberg, 2015; Wayand et al., 2013).

Fortunately, snow deposition is driven by the interaction between local weather and static features like terrain and vegetation, resulting in snowpack spatial distribution patterns that have interannually repeatable characteristics (Deems et al., 2008; Mendoza et al., 2020; Pflug and Lundquist, 2020; Sturm and Wagner, 2010a). This suggests that observations of snowpack distribution from previous seasons could be used to inform how snow should be deposited in real-time. In a study by Vögeli et al. (2016), snow deposition was increased and decreased on a gridcell-by-gridcell basis to mimic the spatial increases and decreases from a distributed snow depth pattern. In doing so, snow input was downscaled as a proxy for unmodeled snow accumulation processes (e.g., preferential deposition, wind-redistribution, etc.). In a following study, Brauchli et al. (2017) found that pattern-based downscaling extended the snowmelt season by a month and greatly improved modeled streamflow. Other studies have used the same strategy

to improve modeled snow distribution and compare interannual snow depth patterns (Gerber et al., 2019; Haberkorn et al., 2017).

The approach described above (Vögeli et al., 2016) could downscale snowfall from a coarser-resolution atmospheric product using a normalized snow depth pattern from a prior year. However, the success of such an approach depends on a relationship between snow deposition (SWE input from snowfall and snow redistribution) and snow depth that is spatially consistent, with a snow depth pattern that is influenced by only snow accumulation. In other words, pattern-based downscaling assumes that the snow depth pattern used to downscale snow deposition has 1) spatially homogeneous snow density, and 2) is not influenced by winter snowmelt or snow sublimation. Pattern-based snow deposition downscaling also requires an accurate estimate of coarse-resolution snowfall (from which to downscale), and interannually repeatable snow depth patterns. We tested these dependencies in California's Tuolumne River watershed in a year with below-normal snowfall (water-year 2014) using airborne lidar snow depth patterns from near peak-snowpack timing in 7 different years and mean snowfall from the 6 km Weather Research and Forecasting (WRF) atmospheric model. Snow depth simulations were performed using a distributed model called SnowModel (Liston and Elder, 2006b), with snow deposition downscaled using both pattern-based downscaling (Section 3.3.1) and a widely used terrain-based downscaling method (Section 3.3.2). In this paper, we address three overarching issues:

1. **The relationship between snow deposition and snow depth near peak-snowpack timing:** How accurate are snow depth simulations when forced with snow deposition downscaled using 1) pattern-based downscaling (Vögeli et al., 2016), and 2) terrain-based downscaling (Liston and Elder, 2006a)? How much do winter snow losses (snowmelt and

sublimation) and snow density influence the relationship between snow deposition and snow depth near peak-snowpack timing?

2. **The interannual consistency of snow patterns:** How accurate are snow depth simulations when downscaling snow deposition using pattern-based downscaling and a library of snow depth patterns from different years? What additional error could be expected from picking a less-representative pattern?
3. **The reliability of coarse-resolution snowfall:** What sort of biases in modeled snow depth do we expect from typical, literature-reported biases in snowfall from atmospheric models?

For each downscaling method (pattern-based and terrain-based), we generated ensembles of simulations with different snow deposition multipliers (constant in space and time). These ensembles were used to 1) diagnose the magnitude and origins of model errors for each downscaling, and 2) represent different atmospheric model snowfall biases. We combined our results to assess and compare the dominant sources of modeled snow depth uncertainty: the snow deposition downscaling method, interannual snow pattern repeatability, or atmospheric model snowfall biases. This careful analysis and synthesis provides guidance on how to best prioritize future research and implement the idea of snow-pattern repeatability in practical operations using tools at our disposal today.

3.3. Background

3.3.1. Pattern-based snow deposition downscaling

Atmospheric models can provide snow model forcing data exceeding the availability of observations in complex terrain (Lundquist et al. 2019). Unfortunately, downscaling snow deposition from the scales of atmospheric models (>1 km) to scales relevant to hydrology (< 100 m; Clark et al., 2011) remains challenging. Pattern-based downscaling adjusts snow deposition

as a proxy for the combined effects of snow accumulation processes (e.g., preferential deposition, avalanching, wind-redistribution, etc.) by assuming that snow deposition mimics the spatial heterogeneity of a snow depth observation (Vögeli et al., 2016). This method downscales snow deposition (*p*, in meters SWE) at each time (*t*) and gridcell (*x,y*),

$$p_{x,y}^{t} = V_{x,y} \times \mu_{p}^{t}, \tag{3.1}$$

where μ_{p}^{i} is domain-mean snowfall (spatially) at each time step (*t*), and $V_{x,y}$ is a spatially distributed snow deposition downscaling factor (constant in time). $V_{x,y}$ is defined from a distributed snow depth observation,

$$V_{x,y} = \frac{d_{x,y}}{\mu_{d}}, \tag{3.2}$$

where $d_{x,y}$ is snow depth at gridcell (*x,y*), and μ_{d} is domain-mean snow depth (spatially). These equations operate under three major assumptions: 1) every snowfall event deposits snow with spatial heterogeneity identical to that of a distributed snow depth observation, 2) an *X*-percent change in snow deposition (relative to mean snowfall; μ_{p}^{t}) will result in an *X*-percent change in snow depth (relative to mean snow depth; μ_{d}), and 3) mean snowfall (μ_{p}^{t}) is unbiased. The pattern-based downscaling also implies that the snow depth observation used to calculate $V_{x,y}$ was influenced only by spatial differences in total snow accumulation. Therefore, gridcells with no snow in the pattern ($V_{x,y}$ = 0.0) would be provided no snow deposition. Gridcells with mean snow depth in the pattern ($V_{x,y}$ = 1.0) would be provided snow deposition at every time step (*t*) identical to μ_{p}^{t}. This forms a linear relationship ($d_{x,y} = m \sum_{t} p_{x,y}^{t}$), of slope $m = \mu_{d}/\mu_{p}$, between snow depth ($d_{x,y}$) and cumulative snow deposition ($\sum_{t} p_{x,y}^{t}$) at every gridcell.

Pattern-based snow deposition downscaling would be most-beneficial operationally if snow deposition downscaling factors ($V_{x,y}$) were interannually repeatable, and therefore could be used

to downscale real-time snow deposition using $V_{x,y}$ from previous years. Interannual snow pattern repeatability has been observed in the Rocky Mountains (Deems et al., 2008; Winstral et al., 2013; Woodruff and Qualls, 2019), Swiss Alps (Schirmer et al., 2011; Schirmer and Lehning, 2011; Vögeli et al., 2016), Alaskan interior (Sturm and Wagner, 2010a), Spanish Pyrenees (López-Moreno et al., 2017), and California Sierra Nevada (Pflug and Lundquist, 2020). In fact, for the Tuolumne watershed investigated here, snow depth observations (at 25 m spatial resolution) from different years, but with similar snow extents and/or seasonal timing, were spatially correlated (median $r > 0.84$). However, for snow deposition downscaling factors ($V_{x,y}$) from two different years to match, snow depth must not only be spatially correlated, but also have identical normalization. Equation 3.2 normalizes snow depth patterns using only mean snow depth, and therefore assumes a constant ratio between mean snow depth and snow depth standard deviation. This assumption is in-line with studies by Sturm and Wagner (2010) and Liston (2004), who assumed a linear relationship between snow magnitude and variability (or constant coefficient of variation). However, Pflug and Lundquist (2020) noted that the snow depth coefficient of variation changed nonlinearly with changes in snow extents, but agreed closely between lidar observations from years with more-similar meteorological conditions and similar snow extents. Here, we tested snow deposition downscaled using $V_{x,y}$ calculated from peak-snowpack snow depth distribution in 7 years, including years with similar ($r = 0.89$) and less-similar ($r = 0.71$) snow depth patterns.

3.3.2. Terrain-based snow deposition downscaling

To date, most modeling approaches downscale snow deposition from coarser-resolution atmospheric models using static controls like the terrain (e.g., Bavay and Egger, 2014; Liston and Elder, 2006a). Here, we focus on MicroMet (MM; Liston and Elder, 2006a), which has been used to distribute precipitation in complex terrain (e.g., Hiemstra et al., 2002; Liston et al., 2007;

Sproles et al., 2013) and is conceptually similar to other terrain-based spatial interpolation methods (e.g., Bavay and Egger, 2014; Daly et al., 2008). Provided precipitation magnitude from a coarser-resolution dataset (such as an atmospheric model), MicroMet interpolates both elevation and precipitation from the dataset to the model grid using inverse-distance weighted (IDW) spatial interpolation (e.g., Barnes, 1964; Koch et al., 1983). Then, to represent changes in precipitation as a function of orographic effects, MicroMet downscales precipitation (p) at each gridcell (x,y) and model time-step (t) using a nonlinear precipitation lapse rate,

$$p^{t}_{x,y} = p^{t}_{x_0,y_0}\left[\frac{1+\beta^{t}\left(z_{x,y}-z_{x_0,y_0}\right)}{1-\beta^{t}\left(z_{x,y}-z_{x_0,y_0}\right)}\right], \tag{3.3}$$

where z_{x_0,y_0} and $p^{t}_{x_0,y_0}$ represent the elevation (m) and precipitation ($mm\ hr^{-1}$), respectively, from the IDW-interpolated precipitation data. $z_{x,y}$ represents the true elevation of each model gridcell and β^{t} is a precipitation lapse rate (km^{-1}), which is specific to each calendar month.

3.3.3. Coarser-resolution precipitation and snowfall biases

The pattern-based and terrain-based snow deposition downscaling methods both rely on unbiased coarse-resolution snowfall. However, snowfall in mountainous watersheds is often uncertain. These uncertainties can stem from uncertainties in the observations used to measure precipitation at storm-event timescales and generate distributed precipitation patterns (e.g., Lundquist et al., 2015, 2010; Trapero et al., 2009; Ye et al., 2012). In fact, gridded precipitation products routinely decrease in accuracy at points further away from observations, where precipitation magnitude depends heavily on spatial interpolation methods (Currier et al., 2017; Gutmann et al., 2012; Hiemstra et al., 2006). Snowfall is also commonly partitioned from rainfall using surface temperature and humidity, the spatial distribution of which can be uncertain in mountainous terrain (Feld et al., 2013; Minder et al., 2010; Wayand et al., 2016). Additionally, common precipitation-phase partitioning methods do not fully capture the atmospheric

conditions influencing precipitation phase (Wayand et al., 2016), and therefore have calibrations that vary across different climates (Jennings et al., 2018).

As opposed to observation-based precipitation datasets, atmospheric models account for processes like orographic precipitation, thermal convection, and cloud microphysics, which are driven in part by synoptic-scale processes and interactions with the underlying terrain (orographic lift, flow separation, etc.). Precipitation from these atmospheric models are beginning to surpass the accuracy of observations in mountainous terrain (Lundquist et al., 2019). However, atmospheric models are still subject to uncertainties. These uncertainties are often demonstrated by atmospheric model sensitivities to different boundary conditions, convections schemes, microphysical partitioning methods, and other parameterizations (Eyring et al., 2016; Hughes et al., 2017; Jankov et al., 2009; Kim et al., 2021; Morales et al., 2018; Taylor et al., 2012). In this study, to test the impact of snowfall biases, we manually imposed snowfall biases consistent with errors from atmospheric models and gridded precipitation products in the Tuolumne region (Henn et al., 2018b; Hughes et al., 2017; Lundquist et al., 2015).

3.4. Data and Modeling

We compared the accuracy of terrain-based and pattern-based snow deposition downscaling in the California Sierra Nevada Tuolumne River watershed. Airborne lidar scanning (ALS) snow depth observations ($\pm$ 0.08 m accuracy; Currier et al. (2019) and Painter et al. (2016)) were used to both define snow deposition downscaling factors (Equation 3.2) and evaluate snow depth simulations (discussed later in this section). We subset 3 m ALS observations to across three 6.25-by-6.25 km subdomains with different snow accumulation processes (Figure 3.1). These domains were approximately the size of atmospheric model gridcells (detailed in the next paragraph), and included terrain with 1) steep slopes with a variety of different aspects (near

Stubblefield canyon), 2) terrain with diverse forest-canopy densities (Matterhorn canyon), and 3) wind-drifted snow bisected by a North/South oriented canyon (Lyell canyon). Pflug and Lundquist (2020) and Currier and Lundquist (2018) both noted the length-scale of snow accumulation variability in this region was 25 m. To reduce spatial autocorrelation, ALS datasets were regridded from 3 m to 25 m spatial resolution. Although this spatial resolution smoothed the effects of snow processes at spatial reaches less than 25 m, regridding: 1) isolated the dominant source(s) of snow deposition (e.g., Clark et al., 2011; Deems et al., 2008; Mendoza et al., 2020), and 2) increased the likelihood of interannually repeatable patterns (Deems et al., 2008; Pflug and Lundquist, 2020). However, we expect that the pattern-based downscaling investigated here (Equation 3.1) could also be applied at finer wind-drift spatial scales (< 10 m).

6 km Weather Research and Forecasting (WRF Version 3.6; Skamarock, 2008) surface fields of air temperature, relative humidity, wind speed, wind direction, and precipitation were used to force snow depth simulations. Boundary conditions for the WRF simulations were obtained from the North American Regional Reanalysis climate model (NARR; Mesinger et al., 2006). The double-moment Thompson microphysics scheme (Thompson et al., 2008, 2004), which best resolves precipitation phase in mountainous terrain (Currier et al., 2017; Hughes et al., 2017; Wayand et al., 2016), was used to partition the precipitation phase at each hourly time-step. This atmospheric model configuration was identical to that of the "*N.Thom"* setup from Hughes et al. (2017). Air temperature, relative humidity, wind speed, and wind direction were downscaled to the ALS grid (25 m resolution) using the MicroMet meteorological distribution routine (Liston and Elder, 2006a). Since WRF has noted issues with shortwave radiation in this region (Lapo et al., 2017), this was calculated within MicroMet using basic solar geometry and cloud fractions calculated from air temperature, relative humidity, and dew point temperature (Walcek, 1994).

Estimates of longwave radiation were calculated using the cloud fraction, air temperature, and the longwave scheme from Iziomon et al. (2003).

SnowModel (Liston and Elder, 2006b) was used to simulate distributed snowpack at hourly time steps and 25 m spatial resolution. Simulations included liquid water percolation adaptations from Pflug et al. (2019), with the melting and non-melting snow albedo decay functions from Sproles et al. (2013). Snowpack simulations were constrained to a maximum of 6 layers, which was found by Pflug et al. (2019) to acceptably represent snow-layer temperature evolution, and the internal movement of liquid snowmelt. SnowModel includes representations of snow redistribution by wind (Liston et al., 2007). However, simulated snowpack redistribution was disabled. Therefore, SWE was influenced only by precipitation (both snow and rain), snowmelt, and snow sublimation. These processes are included in the overwhelming majority of models with snowpack representations. By downscaling snow deposition using pattern-based downscaling, we increased and decreased snow deposition into each gridcell as a proxy for unmodeled snow accumulation and redistribution processes (including the disabled wind-redistribution).

Simulations were restricted to water-year 2014, when WRF outputs and ALS observations overlapped. Additionally, snowpack in water-year 2014 had spatially variable snow density and accumulation-season snowmelt, and thereby provided to opportunity to evaluate pattern-based snow deposition downscaling in a domain and season with less-idealized conditions, as compared to previous studies (e.g., Brauchli et al., 2017; Haberkorn et al., 2017; Vögeli et al., 2016). To check model performance, SWE was simulated in water-year 2014 at the Dana Meadows snow pillow (Figure 3.1). Snowfall at Dana Meadows was defined using the observed increases in SWE, with all other forcing from WRF. Modeled SWE evolution matched

observations closely (temporal coefficient of correlation of 0.99) (Figure 3.2). Snow depth modeled at Dana Meadows also agreed closely with nearby snow depth observations, suggesting that modeled SWE and snow density, and the resulting snow depth, were realistic.

3.5. Methods

Snow depth simulations were performed in water-year 2014 with snowfall from WRF downscaled using terrain-based (Equation 3.3) and pattern-based (Equation 3.1) snow deposition downscaling. For each downscaling, an ensemble of snow depth simulations was generated by perturbing snow deposition by a set of constant (in time and space) snow deposition multipliers (ε) (Figure 3.3, box #1) (described in Section 3.5.1). We first compared snow depth simulations using snow deposition downscaled with 1) MicroMet terrain-based lapse rates (MM), and 2) a snow deposition downscaling factor ($V_{x,y}$) from the simulation season (7 April 2014; V14). In Section 3.5.2, we describe how we compared the two downscaling, and identified the sources and magnitudes of model errors (relating to issue #1 in the introduction) using 1) the default ensemble member ($\varepsilon = 1.0$, constant in space and time), and 2) the ensemble members (ε_{adj}, constant in time) that best matched the 7 April 2014 ALS observation on a gridcell-by-gridcell basis (Figure 3.3, box #3). In Section 3.5.3, we describe how this analysis was expanded to include $V_{x,y}$ calculated from a year with spatially correlated snow depth (16 April 2016; r = 0.89), and $V_{x,y}$ from near peak-snowpack timing in other years ($0.71 \leq r \leq 0.84$) (relating to issue #2 in the introduction). In Section 3.5.4, we focus on snow depth simulated by individual ensemble members with different ε, representative of common atmospheric model snowfall biases (relating to issue #3 in the introduction). Finally, in Section 3.5.5, we describe how we compared the sources of model errors (downscaling methods, interannual pattern repeatability, and coarse-scale snowfall biases), both individually, and in combination.

3.5.1. Model ensemble generation

Equation 3.3 was first used to independently downscale both rainfall and snowfall (partitioned by WRF) to the 25 m model grid. To ensure that mean snowfall was identical for all downscaling (pattern-based and terrain-based), domain mean snowfall (μ_p^t) was calculated from the snowfall downscaled using Equation 3.3 over each domain and time step. Pattern-based snow deposition downscaling (Equation 3.1) was then calculated from the product of μ_p^t and snow deposition downscaling factors ($V_{x,y}$) calculated from ALS snow depth patterns from near peak snowpack timing in 7 different years (2013 through 2019). The pattern-based downscaling was designed to adjust snow deposition (SWE input) as a proxy for multiple snow accumulation processes (wind-redistribution, preferential deposition, avalanching, etc.). However, these processes likely did not affect the deposition of rainfall in the same way. Following methods from Vögeli et al. (2016), we used the rainfall distributed by the terrain-based downscaling in all simulations.

Constant snow deposition multipliers (ε; Figure 3.3) (constant in space and time) were used to generate an ensemble of simulations for each downscaling. Henn et al. (2018a) used a distribution of precipitation multipliers with a standard deviation of ~0.26 to represent basin-average snowfall uncertainty across the full Tuolumne watershed. Across the smaller domains here, which were approximately the size of the WRF pixels, we expected snowfall biases to be larger (e.g., Lundquist et al., 2019). Additionally, snow into any 25 m gridcell could be biased by not only WRF snowfall biases, but also errors with snow deposition heterogeneity. To account for these errors, the standard deviation of ε was set at 0.52 (twice the standard deviation used by Henn et al. (2018a)), discretized at 50 values, centered at $\varepsilon = 1.00$, and spread by equal probability between the 5% and 95% confidence intervals. ε was extended at the lower boundary (by increments of 0.02, to 0.05) and upper boundary (by increments of 0.1, to 2.9), to encompass

snow depth observations that diverged by large amounts from the simulations. Snow deposition downscaled using terrain-based lapse-rates (Equation 3.3) was adjusted,

$$\mathbf{P}^{\mathbf{t}}_{\mathbf{x,y}} = \mathrm{p}^{\mathrm{t}}_{\mathrm{x,y}} \times \varepsilon, \tag{3.4}$$

and used to generate an ensemble of simulations (Figure 3.3, MM). The simulations with pattern-based snow deposition downscaling (Equation 3.1) were also adjusted by ε,

$$\mathbf{P}^{\mathbf{t}}_{\mathbf{x,y}} = \mathrm{V}_{\mathrm{x,y}} \times \mu^{\mathrm{t}}_{\mathrm{p}} \times \varepsilon. \tag{3.5}$$

3.5.2. Comparing pattern-based and terrain-based snow deposition downscaling

We started by comparing snow depth simulated by the default ensemble members ($\varepsilon = 1.0$, constant in space and time) (Figure 3.3, box #3, red line) using 1) terrain-based downscaling (MM) and 2) pattern-based downscaling with a snow deposition downscaling factor ($\mathrm{V}_{\mathrm{x,y}}$) defined from the ALS observation nearest peak-snowpack from the same year as the simulation (V14; 7 April 2014). The snow depth spatial coefficient of correlation (r), median percent bias (of all individual gridcells), mean absolute error (μ_{err}), and coefficient of variation (CoV) percent-error were used to evaluate the accuracy of simulated snow depth versus the 7 April 2014 ALS observation.

To diagnose the sources and magnitudes of downscaling errors, we determined the MM and V14 ensemble members that agreed closest with the 7 April 2014 target ALS observations on a gridcell-by-gridcell basis ($\varepsilon_{\mathrm{adj}}$, constant in time) (Figure 3.3, box #3, black line). Since both MM and V14 downscaling were provided identical domain mean snowfall ($\mu^{\mathrm{t}}_{\mathrm{p}}$), identical model forcing, and the same snow model, differences in the distributions of $\varepsilon_{\mathrm{adj}}$ between the different downscaling (e.g., Figure 3.4a) were driven only by differences in snow deposition. Since the V14 simulation uses the same ALS observation (from 7 April 2014) to define the snow deposition downscaling factor ($\mathrm{V}_{\mathrm{x,y}}$) and pick the best-performing ensemble members, model

errors are not driven by interannual differences in $V_{x,y}$ (issue #2 from the introduction). Instead, the distribution of ε_{adj} necessary to correct the V14 pattern-based downscaling could be attributed to the combination of 1) the impacts of winter snow losses (snowmelt and snow sublimation) and snow density on $V_{x,y}$ (issue #1 from the introduction), and 2) biases in μ_p^t (issue #3 from the introduction). V14 ε_{adj} that were used more-frequently (e.g., Figure 3.4a, ε_{adj_1} and ε_{adj_2}) identified portions of the modeling domain that exhibited similar relationships between cumulative snow deposition and snow depth (Figure 3.4b, blue dashed lines). Here, we evaluated the V14 most frequently used ε_{adj} and their connections with winter snow losses and snow density spatial variability, and the degree to which the relationship between cumulative snow deposition and snow depth differed from the linear relationship (discussed in Section 3.3.1) assumed by the pattern-based downscaling method (Figure 3.4b, solid line).

For the small portion (< 4%) of each domain with snow-absence on 7 April 2014, ensemble members using different snow deposition multipliers could reproduce snow-free conditions by simulating snow disappearance at different dates. In the results, we excluded gridcells where snow depth in the 7 April 2014 observation was less than the uncertainty of the ALS observations (0.08 m). Excluding these gridcells excluded comparisons between matching snow-free simulations and observations, and thereby made our metrics of model performance on 7 April 2014 (coefficient of correlation, median bias, mean absolute error, and coefficient of variation), more critical.

3.5.3. Identifying errors driven by interannual changes in snow patterns

The pattern-based downscaling would be most-useful if snow deposition downscaling factors ($V_{x,y}$) from previous years could be used to downscale snow deposition in real-time. Therefore, the methods from Section 3.5.2 were repeated using the snow deposition downscaling factor

($V_{x,y}$) defined from near peak-snowpack timing in other years. We first focused on pattern-based snow deposition downscaling using $V_{x,y}$ from 16 April 2016 (V16), which had different snow depth magnitude (~42% larger), but spatially correlated snow depth (r = 0.89). $V_{x,y}$ calculated from the ALS observation on this date represented the best-case-scenario for pattern-based snow deposition downscaling using $V_{x,y}$ from a different year (given this set of ALS observations). However, we also evaluated the impact on snow depth simulations when downscaling snow deposition using $V_{x,y}$ from years with less-similar meteorological conditions and less-spatially correlated ($0.71 \leq r \leq 0.84$) snow depth patterns.

3.5.4. Identifying errors driven by coarse-scale snowfall biases

Individual ensemble members with different snow deposition multipliers (ε; constant in time and space) were also used to evaluate the impact on modeled snow depth for each downscaling (terrain-based and pattern-based) provided a bias to domain mean snowfall. In other words, since all downscaling methods were provided identical μ_p^t (Section 3.5.1), ε influenced the average snow mass into each domain by the same amount, yet the spatial heterogeneity of snow deposition was still dictated by the downscaling strategy. For instance, the ensemble member with $\varepsilon = 1.25$ (Equation 3.5) was used to represent snow depth that would be simulated given a +25% bias to the WRF snowfall. ε values were selected to encompass the range of snowfall biases observed for atmospheric models and precipitation products in this region (e.g., Henn et al., 2018b; Hughes et al., 2017; Lundquist et al., 2015), ranging between $\pm 60\%$.

3.5.5. Combining and isolating errors driven by downscaling methods, interannual snow pattern repeatability, and coarse-scale snowfall biases

We finished by comparing the sources of snow deposition downscaling errors detailed in Sections 3.5.2 – 3.5.4 above, both individually, and in combination. We used two-way analysis of variance (ANOVA) to compare peak-snowpack modeled snow depth errors across simulations

with different snowfall biases (ε) against simulations with different downscaling methods (pattern-based versus terrain-based) and snow deposition downscaling factors ($V_{x,y}$) from different years. These results were used to disentangle the combined errors from different snow deposition downscaling (pattern-based versus terrain-based), snowfall biases, and interannual changes to snow depth patterns.

3.6. Results

3.6.1. Comparing pattern-based and terrain-based snow deposition downscaling

7 April 2014 snow depth simulated using pattern-based and terrain-based downscaling (default ensemble members, $\varepsilon = 1.00$) differed by large amounts. Snow depth simulated using terrain-based downscaling (MM) was only biased by -5% (-0.06 m), 14% (0.14 m), and -8% (-0.07 m) across the Stubblefield, Matterhorn, and Lyell domains, respectively (Table 3.1). However, the spatial distribution of snow depth was too homogeneous (CoV biased between -13 and -47%) with poor snow depth spatial distribution ($r < 0.38$) (Figure 3.5). As compared to the simulation with terrain-based downscaling, the simulation with pattern-based downscaling using the snow deposition downscaling factor ($V_{x,y}$) from the same year as the simulation (V14) was biased by similar magnitudes and had a 7 April 2014 snow depth coefficient of correlation (r) that improved from 0.28 to 0.76, and a mean absolute error that improved from 0.38 to 0.24 (on average, across the three domains). However, the V14 default ensemble member tended to overaccumulate and underaccumulate in deep and shallow-snow regions, respectively, resulting in a modeled snow depth coefficient of variation that was high-biased by greater than 20% in each domain (Table 3.1).

The best-performing ensemble members (ε_{adj}) for the simulations with terrain-based (MM) and pattern-based (V14) downscaling reproduced the 7 April 2014 target snow depth observations (Table 3.1) using very different adjustments to snow deposition (Figure 3.6a and

3.6b). The MM ensemble members that best-reproduced the 7 April 2014 snow depth (ε_{adj}) were widely distributed. In fact, no single ε_{adj} occurred across more than 5% of any domain (Figure 3.6a). This differed from the V14 best-performing ensembles, which matched 7 April 2014 snow depth using unique ensemble members with much higher frequency (Figure 3.6b). For each domain, these increases in V14 ε_{adj} frequency were not centered about a single ε_{adj} value, but instead resembled multi-modal distributions about different ε_{adj}, each with narrow peaks (relative to the distribution of ε_{adj} for the MM ensemble). This suggested that large portions of each domain had V14 snow deposition that was able to be corrected with a smaller number of snow deposition multipliers, each corresponding to unique issues with the pattern-based snow deposition downscaling. Recall that since the V14 simulation used the same ALS observation (from 7 April 2014) to define the snow deposition downscaling factor ($V_{x,y}$) and pick the best-performing ensemble members, V14 model errors (Figure 3.5) were attributable to errors with $V_{x,y}$ and domain mean snowfall (μ_p^t) biases (Section 3.5.2). In the remainder of this section, we focus on the sources of $V_{x,y}$ errors. μ_p^t biases are discussed in Section 3.6.3.

Pattern-based snow deposition downscaling assumes that the snow deposition downscaling factor ($V_{x,y}$) is influenced only by snow accumulation processes and gridcells with identical $V_{x,y}$ receive the same amount of snow deposition. However, spatially variable winter snow losses (snowmelt and sublimation) and snow density can manipulate $V_{x,y}$. For instance, assuming two gridcells with equal snow depth (equal $V_{x,y}$), and equal SWE and snow density, a gridcell influenced by snowmelt and/or snow sublimation would require more snow deposition to achieve that snow depth than a gridcell with no winter snow losses. Similarly, assuming two gridcells with equal snow depth (equal $V_{x,y}$) and no winter snow losses, a gridcell with denser snow would require more snow deposition to achieve that snow depth than a gridcell with less-dense snow. In

the Tuolumne domains in water-year 2014, the best-performing V14 ensemble members simulated spatially variable snow density (Figure 3.7, d-f) and large amounts of winter snowmelt and sublimation (Figure 3.7, j-l) prior to 7 April 2014. This resulted in a gridcell-by-gridcell relationship between cumulative snow deposition and 7 April 2014 snow depth that disagreed with the linear relationship assumed by pattern-based snow deposition downscaling (Figure 3.7, solid lines), but conformed more with several different linear relationships. These linear relationships corresponded well with the peak ε_{adj} frequencies from Figure 3.6b (e.g., conceptual diagram in Figure 3.4). For example, winter snow losses occurred prior to 7 April 2014 in all three domains (Figure 3.7, m-o, red shading). In these gridcells, snow losses reduced snow depth, reduced $V_{x,y}$ values calculated from the 7 April 2014 ALS observation, reduced snow deposition (relative to μ_p^t), and simulated shallower snow depths that melted more-readily. Put another way, since 7 April 2014 $V_{x,y}$ was influenced by winter snowmelt and snow sublimation, the V14 simulation "double-counted" winter snow losses in 1) $V_{x,y}$, and 2) the snow model. At gridcells with winter snow losses, the relationship between cumulative snow deposition and 7 April 2014 snow depth was approximated best (Figure 3.7, g-i) by the rightmost peak in ε_{adj} frequency (Figure 3.6b) corresponding to larger snow deposition multipliers.

Winter snow losses influenced $V_{x,y}$ not only in gridcells with snowmelt and snow sublimation, but also in gridcells without winter snow losses. For instance, in the Lyell canyon domain, lower-elevation snowmelt reduced domain mean snow depth (μ_d) to a value smaller than what it would have been if snow depths were influenced by snow accumulation only. As a result, snow depths in gridcells with minimal winter snow losses were deeper, relative to μ_d, than they should have been. This resulted in high-biased $V_{x,y}$ (Equation 3.2) and snow deposition. The relationship between cumulative snow deposition and 7 April 2014 snow depth in these gridcells was

approximated best using the snow deposition multipliers of 0.85, 0.74, and 0.82 in the Stubblefield, Matterhorn, and Lyell domains, respectively (Figure 3.7, g-i). These snow deposition multipliers reduced snow deposition (relative to the default ensemble member, ε = 1.0) and corresponded well with peaks in ε_{adj} frequency from Figure 3.6b. Approximately 4% of the Stubblefield domain, 13% of the Matterhorn domain, and 8% of the Lyell domain also contained gridcells on mid- to high-elevation north-facing slopes (Figure 3.7, m-o, blue shading) that had less winter snow settling, less-dense snow, and thereby required less snow deposition to achieve snow depths identical to gridcells in more-exposed locations (Figure 3.7, a-c, red boxes). This effect was also noted by Vögeli et al. (2016) and Brauchli et al. (2017) between north and south-facing slopes in the Swiss Alps. These gridcells approached 7 April 2014 snow depth observations well using the smallest ε_{adj} from Figure 3.6b.

Above, we evaluated the combined effect of winter snow losses (snowmelt and snow sublimation) and spatially variable snow density on $V_{x,y}$, which resulted in pattern-based snow deposition downscaling that was too heterogeneous (Figure 3.5). However, using the best-performing V14 ensemble members ε_{adj}, we also can separate and compare the impact of winter snow losses and snow density spatial-variability on the heterogeneity of the snow deposition downscaling factor ($V_{x,y}$). For instance, $V_{x,y}$ calculated using the 7 April 2014 SWE map from the best-performing V14 ensemble members had corrected snow deposition that was not influenced by the spatial variability of snow density. As compared to $V_{x,y}$ calculated using 7 April 2014 snow depth, $V_{x,y}$ calculated from 7 April 2014 SWE had a coefficient of variation that was ~4% smaller (Figure 3.8). However, $V_{x,y}$ calculated using the cumulative snow deposition from the best-performing V14 ensemble members, which was not influenced by either snow density spatial variability or winter snow losses, had a coefficient of variation 31 – 40% smaller

than $V_{x,y}$ calculated from 7 April 2014 snow depth (Figure 3.8). In other words, in this year and location, the overestimation of snow depth variability for the pattern-based downscaling method (Figure 3.5 and Table 3.1) was driven much more (~8 times more) by winter snow losses than 7 April 2014 snow density spatial variability.

3.6.2. Identifying errors driven by interannual changes in snow patterns

The snow deposition downscaling factors ($V_{x,y}$, calculated using ALS snow depth) from 7 April 2014 and 16 April 2016 were remarkably similar (r = 0.89). In fact, it was difficult to visually distinguish differences in $V_{x,y}$ between the two dates (Figure 3.9). Although mean snow depth on 16 April 2016 was approximately 42% greater (on average, across the domains), the ALS observations were both within a week of peak-snowpack timing (defined from the Dana Meadows snow pillow) in their respective seasons, snow extents agreed to within 8%, and snow depth coefficients of variation agreed to within 5%. Overall, differences between 7 April 2014 and 16 April 2016 $V_{x,y}$ were not explained well by any terrain characteristic ($|r| < 0.20$). However, some of the larger $V_{x,y}$ differences in the Matterhorn domain were in regions with forest cover (green regions in Figure 3.1), suggesting that snow-forest processes may have been different. In the Lyell domain, $V_{x,y}$ differences were larger on many Northeastern-facing slopes and in the Northeast corner of the domain where avalanches and snow-sloughing are more common. The default snow depth ensemble member (ε = 1.00) on 7 April 2014 using pattern-based downscaling and $V_{x,y}$ from 16 April 2016 (V16) had a mean absolute error of 0.41 m (71% larger than V14), and a coefficient of correlation of 0.44 (28% smaller than V14), on average across the three domains.

Some of the best-performing V16 ensemble members (Figure 3.6c) used ε_{adj} from the edges of the ensemble-space (ε_{adj} = 0.05 or ε_{adj} = 2.90), showing that the lidar observations were

sometimes outside of the model ensemble-space. This occurred at gridcells with shallow snow depth and small snow deposition downscaling factors ($V_{x,y}$), where large and small ε_{adj} changed snow deposition and the resulting snow depth by small amounts (typically < 0.05 m). Despite these issues, the best-performing ensemble members still reproduced 7 April 2014 snow depth well (Table 3.1). In addition to the errors with pattern-based downscaling driven by snow density spatial variability and winter snow losses, the best-performing V16 ensemble members also had to correct for the spatial differences in $V_{x,y}$ between 16 April 2016 and 7 April 2014. As a result, the V16 ensemble required a wider range ε_{adj} to approach the 7 April 2014 ALS observation, smoothing the unique spikes in ε_{adj} frequency seen for the V14 ensemble (Figure 3.6).

Snow patterns ($V_{x,y}$) varied from the 7 April 2014 snow pattern most in seasons with snowfall magnitude and snow extents unlike 2014 (Figure 3.9) (Pflug and Lundquist, 2020). For example, water-year 2015 had the shallowest snow volume ever recorded or estimated in the Tuolumne watershed (Belmecheri et al., 2016; Margulis et al., 2016b). As a result, larger portions of the model domains were snow-free during the 5 March 2015 ALS observation, reducing mean snow depth (μ_d) and elevating $V_{x,y}$ values in gridcells where snow existed. Snow deposition downscaled using this snow pattern simulated snow depth that was too heterogeneous, increasing the interquartile range (IQR) of simulated snow depth errors at peak snowpack timing (Figure 3.10, $\varepsilon = 1.0$). The same was true for the simulations with pattern-based snow deposition downscaling using the more-heterogeneous $V_{x,y}$ from 3 April 2013. Conversely, $V_{x,y}$ calculated from ALS observations of deep snowpacks on 1 April 2017 and 24 March 2019 had less snow depth variability relative to domain mean snow depth (Figure 3.10), and thereby simulated snow depth that was more spatially homogeneous than pattern-based downscaling using other $V_{x,y}$.

3.6.3. Identifying errors driven by coarse-scale snowfall biases

Cumulative WRF snowfall prior to 7 April 2014 was biased by only small amounts (Section 3.6.1). However, we know that 6 km snowfall from atmospheric models can be biased in mountainous regions. In Figure 3.10, we show the distribution of snow depth errors for various constant snow deposition multipliers (ε, constant in space and time). ε ranged between 0.40 and 1.60 to represent a $\pm$60% error in mean snowfall, consistent with previous studies in this region (Henn et al., 2018b; Hughes et al., 2017; Lundquist et al., 2015). 7 April 2014 median snow depth biases typically increased linearly with ε for all terrain-based and pattern-based downscaling methods. For $\varepsilon \pm 10\%$ (Figure 3.10, zoomed portion), the IQR of errors driven by errors with the different snow deposition downscaling were more than twice as large as the median snow depth biases driven by domain mean snowfall. Even provided a +10% bias in domain mean snowfall (ε = 1.10), all simulations had an IQR that included gridcells with low-biased snow depth. This was particularly true for the simulations with pattern-based downscaling using the spatially heterogeneous snow deposition downscaling factor ($V_{x,y}$) calculated from the low-snow 2015 ALS observations. For $\pm 25\%$ biases (ε values of 0.75 and 1.25), the upper and lower boundaries (respectively) of the IQR of snow depth errors were closer to 0 m, suggesting that the magnitude of snow depth errors driven by both pattern-based and terrain-based downscaling were comparable in size to snow depth errors driven by snowfall biases. This is discussed more in Section 3.6.4.

We tested the impact of snowfall biases by prescribing identical snow deposition multipliers (ε) to all snowfall events. However, errors in snowfall for single snow events may still have a sizable impact on modeled snow depth. For instance, in this study, WRF overestimated a spring snowfall event (25/26 April 2014), causing snow depth volume biases upwards of 30% through much of the snowmelt season (Figure 3.11). Even though the spatial heterogeneity of snow

influenced the rate and timing of spring snowmelt (e.g., Freudiger et al., 2017; Luce et al., 1998; Pflug and Lundquist, 2020), this spring snowfall bias caused departures from observed snow volume that were larger than the late-spring difference between simulations with different snow deposition downscaling. The spring snowfall event also biased the cold content of late-season snowpack and delayed the rate of spring snowmelt (relative to observations), causing differences between modeled and observed snow depth volume that grew over time for all simulations. Currently, springtime precipitation magnitude and phase errors are a known issue for atmospheric models in mountainous terrain (Aas et al., 2017; Wrzesien et al., 2015). As shown by Figure 3.11, these errors could affect snowmelt modeling and forecasting, even in cases where peak snow volume and distribution are modeled well.

3.6.4. Combining and isolating errors driven by downscaling methods, interannual snow pattern repeatability, and coarse-scale snowfall biases

The combination of snowfall biases, and the impact of winter snowmelt, winter sublimation, and spatially variable snow density on the snow deposition downscaling factor ($V_{x,y}$), could have compounding effects on snow depth modeled using the pattern-based snow deposition downscaling. For example, assuming a domain mean snowfall bias of +50% ($\varepsilon = 1.50$), snow deposited using the pattern-based downscaling was not uniformly biased (in space) but was instead biased as a multiplicative factor of $V_{x,y}$ at each gridcell (Equation 3.5). Given this snowfall bias ($\varepsilon = 1.50$), $V_{x,y}$ from a year with a more heterogeneous snow pattern (e.g., 2015) would multiplicatively increase the heterogeneity of simulated snow depth (Figure 3.10). As opposed to pattern-based downscaling, the snow depth heterogeneity in simulations with terrain-based downscaling was sensitive only to the land surface elevation and parameterized lapse-rates (Equation 3.3). Therefore, the IQR of MM snow depth errors was insensitive to ε (Figure 3.10). Despite this, two-way analysis of variance (ANOVA) determined that the IQR of snow depth

errors were significantly different ($p < 0.01$) across ensemble members with different mean snowfall biases (ε, ranging between 0.40 and 1.60), but not for simulations with different downscaling methodologies, or pattern-based downscaling using $V_{x,y}$ from different years ($p > 0.30$). Therefore, in these domains, the bias and range of modeled snow depth errors driven by common snowfall biases dominated, even when assuming the snow depth used to downscale snow deposition was only available from years with less-typical snowpack (e.g., 2015 and 2017) and the range of snowfall biases (ε) were more modest (between 0.75 and 1.25).

Simulations with terrain-based downscaling were less-sensitive to snowfall biases than simulations with pattern-based downscaling. For instance, $V_{x,y}$ defined from ALS snow depth observations (Figure 3.9) commonly varied from 0 (-100% adjustment from mean snowfall) to greater than 2 (+200% adjustment from mean snowfall). However, the spatial pattern of terrain-based downscaling dictated by elevation and β^t (Equation 3.3) adjusted snowfall from WRF by no more than $\pm$30%, resulting in simulated snow depth that was far too spatially homogeneous. In fact, the heterogeneity of snow was so poorly-captured that the IQR of snow depth errors for simulations with terrain-based downscaling was nearly equal to the IQR of snow depth observed by ALS on 7 April 2014 (0.69 m). Additionally, there was little difference in the IQR of snow depth errors across the $\pm$60% range of mean snowfall biases (ε) tested in Figure 3.10. Put another way, we would expect a similar IQR of model errors on 7 April 2014 provided simulated snow depth that was equal to mean snow depth at all model gridcells.

Since snow depth simulated using terrain-based snow deposition downscaling did not attempt to capture the more-complex drivers of snow accumulation (e.g., wind-redistribution, preferential deposition, avalanching, etc.), and as a result simulated snow depth that was too spatially homogeneous, spring snow melted more rapidly than observations and simulations with pattern-

based downscaling (Figure 3.11). Interestingly, the simulation with pattern-based downscaling using $V_{x,y}$ from 1 April 2017, which was more spatially homogeneous (coefficient of variation 14% smaller than $V_{x,y}$ calculated from the 7 April 2014 ALS observation), matched ALS observed snow volume preceding the spring snowfall event better than any other simulation. This was driven by the fact that 1 April 2017 snow depth observation had much deeper snow depth that was less-influenced by winter snowmelt and snow density spatial variability, and as a result, had $V_{x,y}$ spatial variability (Figure 3.9) that agreed closely with melt- and density-corrected $V_{x,y}$ on 7 April 2014 (Figure 3.8, solid black line). This suggests that interannual changes in winter snowmelt, snow sublimation, and snow density were major sources of interannual differences in $V_{x,y}$.

3.7. Discussion and recommendations for future work

Since interannual changes to snowmelt, snow sublimation, and snow density were partially responsible for interannual changes to $V_{x,y}$, future efforts should focus on correcting or mitigating for these effects. For instance, Vögeli et al. (2016) calculated $V_{x,y}$ from earlier in the winter snow season, thus reducing the amount of winter snow losses. Vögeli et al. (2016) also calculated $V_{x,y}$ across a region with less snow density spatial variability. However, early-winter snowmelt occurred in each of the Tuolumne domains between mid-December and February water-year 2014. $V_{x,y}$ calculated from too early in the winter season may also decrease the likelihood of interannual snow pattern repeatability since winter-integrated snow accumulation patterns tend to be more repeatable than snow accumulation patterns driven by a small number of snowstorms (Schirmer et al., 2011). Pattern-based downscaling could have also been performed including only gridcells with similar snow densities and negligible winter snow losses. However,

including only these regions (unshaded regions in Figure 3.7, m-o) would neglect snowpack that is hydrologically important for streamflow forecasting in this region.

Some snow deposition downscaling methods more-explicitly attempt to account for effects like wind-driven snow preferential deposition and redistribution (e.g., Dadic et al., 2010; Sturm and Wagner, 2010; Winstral et al., 2013) and larger-scale meteorological processes like air temperature gradients, rain shadows, and coastal effects (e.g., Daly et al., 2008; Livneh et al., 2015; Thornton et al., 1997). Contrary to these approaches, the pattern-based snow deposition downscaling method tested here downscaled snow deposition to implicitly mimic the spatial increases and decreases in snow depth patterns resulting from multiple complex processes including wind-redistribution, preferential deposition, and avalanching. This approach is advantageous since mountain meteorological forcings at the spatial scales of these processes are uncertain (e.g., Musselman et al., 2015; Reynolds et al., 2021). Additionally, few snow models represent these complex snow processes, and for those that do, performing simulations with spatial resolutions and snow-layering schemes relevant to processes like wind-redistribution and avalanching is computationally expensive. In this study, we downscaled snow deposition using a unique library of ALS snow depth observations (7 years) at spatial resolutions much finer (25 m) than the commonly used gridded meteorological products discussed above ($\geq$ 800 m). These lidar snow depth observations also have great accuracy ($\pm$ 0.08 m) in high-elevation terrain where gridded products routinely decrease in reliability (Lundquist et al., 2019; Wayand et al., 2013).

Future snow-focused satellite missions (e.g., Kim et al., 2017) are expected to increase the frequency and extent of snow pattern observations, and interannual snow pattern repeatability has been observed in many locations (Deems et al., 2008; Pflug and Lundquist, 2020; Sturm and

Wagner, 2010; Woodruff and Qualls, 2019). Therefore, we expect that pattern-based snow deposition downscaling could still be a valuable approach for the snow community, becoming even more valuable after correcting for the issues detailed in this manuscript. For example, since errors in the linear relationship between snow depth and the snow deposition were dominated by winter snow losses, time-continuous SWE products from previous years (Barrett, 2003; Margulis et al., 2019, 2016a; Rittger et al., 2016) may be better suited to calculate snow deposition downscaling factors. These SWE products could be used to calculate the gridcell-by-gridcell daily cumulative increases in SWE, including the SWE lost to winter snowmelt and snow sublimation. Snow deposition downscaling factors calculated using cumulative increases in SWE would not be influenced by snow density spatial variability nor winter snow losses. Additionally, these continuous SWE products are often calculated using satellite observations and distributed point observations, which are more common than airborne lidar and have multidecadal data-records, which may increase the likelihood of identifying a historic snow pattern with conditions representative of snowpack in any given year. This may be particularly true in years with less-typical meteorological conditions and snow patterns. Snowfall biases from atmospheric models could also be corrected by improving atmospheric model performance, or by bias correcting snowfall using real-time observations. For instance, subgrid snow depletion simulated by the WRF land surface model is a known issue in mountainous regions (Aas et al., 2017; Wrzesien et al., 2015), including the Tuolumne domain investigated here. Aas et al. (2017) found that correcting WRF subgrid snow depletion improved air temperature and latent heat fluxes between the land surface and atmosphere, changing spring precipitation magnitude by up to 50% in a Nordic region. This approach should also be investigated in the California Sierra Nevada. Provided a reliable snow deposition pattern, strategically located field observations (e.g., Oroza

et al., 2016; Pflug and Lundquist, 2020) may be also able to bias-correct coarser-scale estimates of snowfall from atmospheric models like WRF.

3.8. Conclusions

Snow accumulation in mountainous terrain is highly variable, and the distribution at fine spatial scales (<25 m) impacts the timing and duration of spring snowmelt, and related ecology and water operations. While most current models do not accurately account for the multitude of processes that influence snow spatial heterogeneity, snow accumulation patterns can be remarkably similar interannually (Plug and Lundquist 2021). Here, we downscaled 6 km WRF estimates of mean snowfall to 25 m using lidar-observed snow depth patterns from the California Tuolumne River watershed near peak-snowpack timing in 7 different years. We compared the pattern-based snow deposition downscaling with a popular terrain-based snow deposition downscaling routine, using lidar observations as evaluation, and assessed the magnitude of snow depth errors driven by domain mean snowfall biases, different snow deposition downscaling methods, and interannual changes to snow depth patterns.

Fine-scale snow depth magnitude and spatial distribution simulations were most-influenced by mean snowfall estimates in complex terrain, the accuracy of which is among the greatest limitations for mountain snow modeling (Lundquist et al., 2019; Raleigh et al., 2015). At peak-snowpack timing, the influence of common Sierra Nevada snowfall biases ($\pm$60%) on the magnitude and spread of snow depth errors was larger, and significantly different from ($p < 0.01$) errors contributed by two different snow deposition downscaling approaches (terrain-based and pattern-based) and by interannual changes to snow depth patterns (for pattern-based downscaling). Even provided unbiased winter snowfall, snowfall biases from a single spring snow event caused simulated snowmelt rates and spring snow volume to diverge from observations. Again, this divergence was larger than variations driven by 1) simulations with

different downscaling, and 2) simulations downscaling snow deposition using snow depth patterns more- and less-similar to snow depth patterns from the simulation year (Figure 3.11).

Snow depth simulated at peak-snowpack timing using pattern-based snow deposition downscaling was better correlated with snow depth observations ($0.52 < r < 0.76$) than simulations with more-commonly-used terrain-based downscaling ($r = 0.27$). However, snow deposition downscaled using snow depth patterns often simulated snow depths that were too spatially heterogeneous. These errors were driven mostly by winter snowmelt and snow sublimation, which influenced peak-snowpack snow depth and the gridcell-by-gridcell relationship to cumulative snow deposition prior to that date. Although spatially variable snow density also influenced the relationship between snow deposition and snow depth, its impact on the spatial heterogeneity of snow deposition in this domain was about 8 times smaller than winter snowmelt and snow sublimation (Figure 3.8). Interannual changes in snowmelt, snow sublimation, and snow density were also partially responsible for interannual changes in snow depth patterns, and the resulting snow input when using pattern-based snow deposition downscaling.

We determined that the largest sources of pattern-based snow deposition downscaling errors, listed in order from most- to least-impactful, included 1) coarse-scale snowfall biases from atmospheric models, 2) snow depth patterns that were influenced by processes other than snowfall (like snowmelt, snow sublimation, and snow density), and 3) interannual changes to snow accumulation patterns. Although pattern-based snow deposition downscaling adjusts snow input as a proxy for physical processes that are a current focus of much model development, pattern-based downscaling may provide the ability to improve modeled snowpack accuracy for the majority of models without complex snow deposition and redistribution routines, in a

computationally efficient manner, using tools and observations that we have at our disposal today. Future efforts should therefore focus on strategies that could bias-correct snowfall provided from atmospheric models and correct snow patterns for processes that change the relationship between observed peak snow depth and winter snow deposition, like snowmelt, snow sublimation, and snow density.

Table 3.1. Statistics for snow depth simulated on 7 April 2014 (at noon) using terrain-based (MM) and pattern-based snow deposition downscaling factors ($V_{x,y}$) from 7 April 2014 (V14) and 16 April 2016 (V16). Statistics are shown for the default ensemble member (ε = 1.0) and for the best-performing ensemble members (ε_{adj}) in the Stubblefield, Matterhorn, and Lyell domains (S, M, and L, respectively).

Snowfall downscaling	*Ensemble members*	*Coefficient of correlation (r)*			*Median bias (%)*			*Mean absolute error (m)*			*Coefficient of variation bias (%)*		
	domain:	S	M	L	S	M	L	S	M	L	S	M	L
Terrain-based [MM]	$\varepsilon = 1.0$	0.24	0.21	0.37	-5	14	-8	0.42	0.42	0.30	-47	-32	-13
	ε_{adj}	0.99	0.99	0.99	0	0	0	0.01	0.02	0.01	0	0	0
Pattern-based [V14]	$\varepsilon = 1.0$	0.83	0.81	0.63	-6	15	11	0.20	0.27	0.26	27	21	20
	ε_{adj}	0.99	0.99	0.99	0	0	0	0.01	0.01	0.01	0	0	2
Pattern-based [V16]	$\varepsilon = 1.0$	0.64	0.53	0.46	-5	19	10	0.32	0.48	0.43	33	32	38
	ε_{adj}	0.98	0.95	0.94	0	0	0	0.03	0.04	0.06	6	9	20

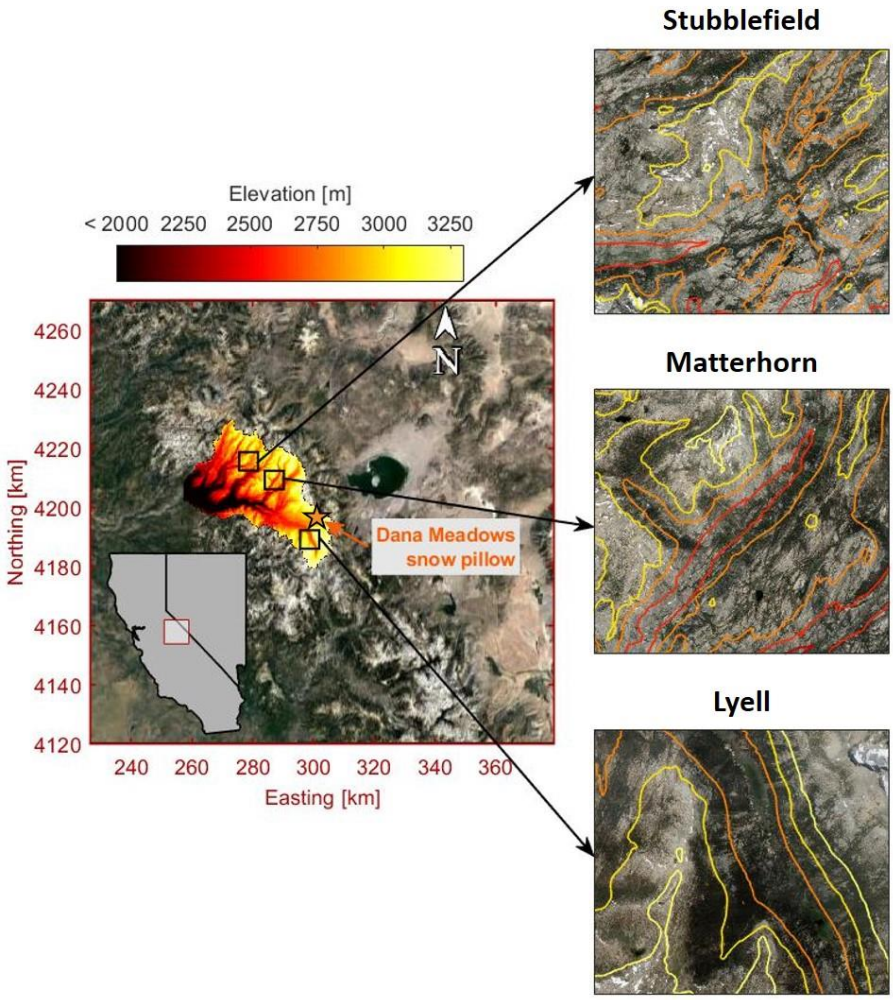

Figure 3.1. Elevation (colorbar) for the Tuolumne watershed (left) and modeling domains (right, contours). The location of the Dana Meadows snow pillow is shown in the leftmost plot.

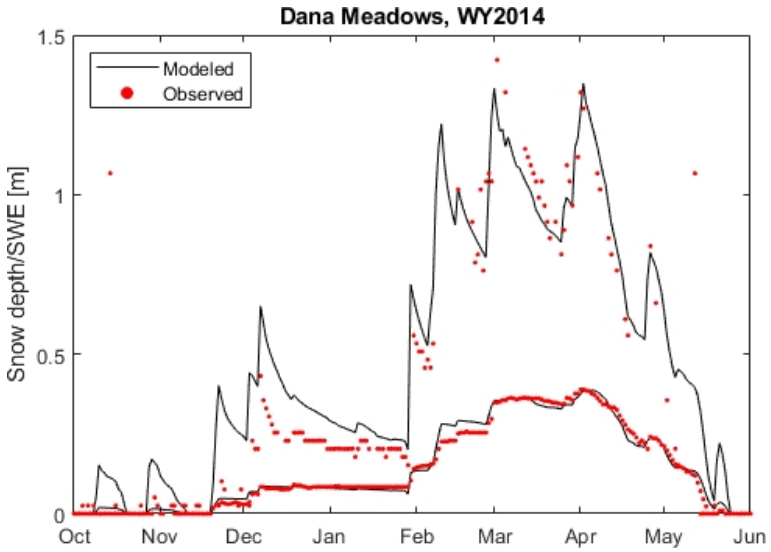

Figure 3.2. Modeled (line) and observed (dotted) snow depth and SWE at the Dana Meadows snow pillow in water-year 2014.

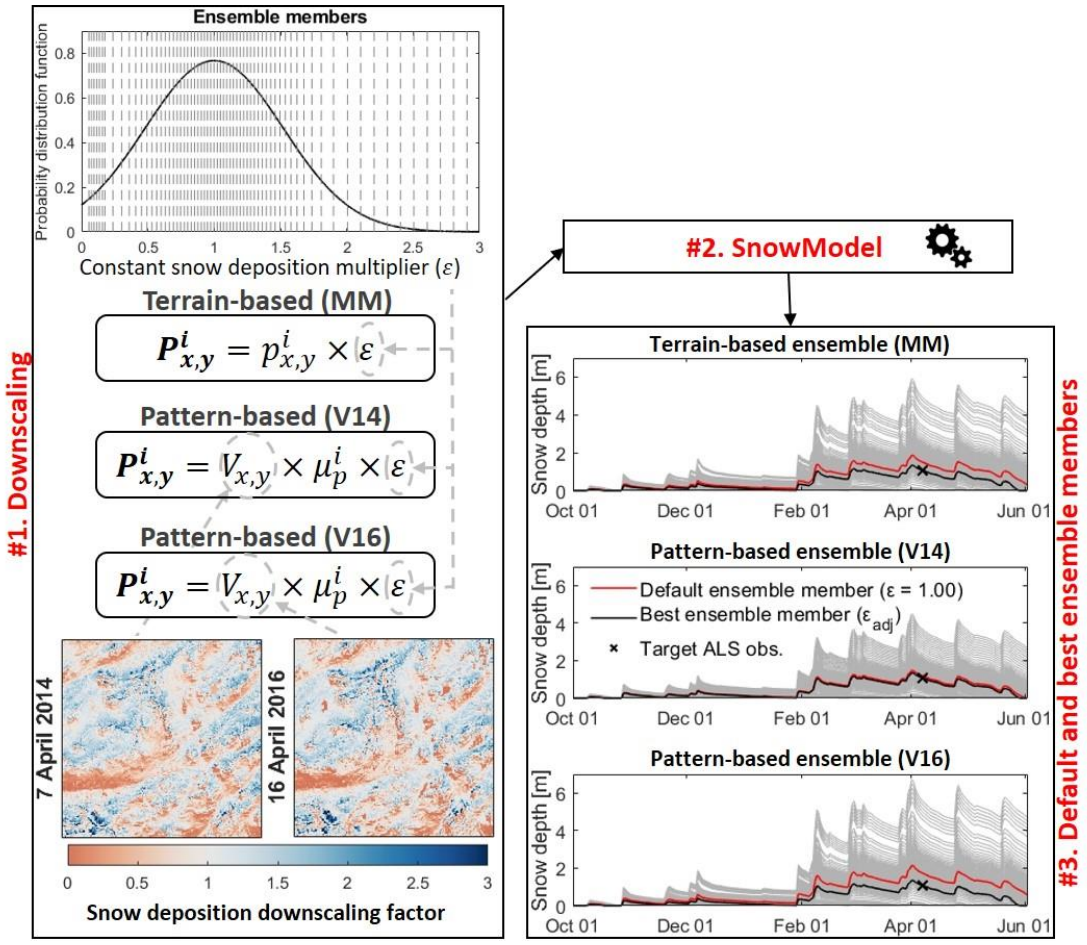

Figure 3.3. Snowfall from WRF in water-year 2014 was downscaled (box #1) using 1) terrain-based MicroMet lapse rates (MM), 2) a snow deposition downscaling factor ($V_{x,y}$) from an ALS observation on 7 April 2014 (V14), and 3) $V_{x,y}$ from an ALS observation on 16 April 2016 (V16). An ensemble of simulations was performed for each downscaling using a normally distributed set of constant (in both space and time) snow deposition multipliers (ε, dashed lines). In addition to the default ensemble member (box #3, red line), the ensemble member best-reproducing the 7 April 2014 (box #3, black line) target ALS observation (box #3, scatter point) was selected for each model gridcell. Ensembles were also performed using $V_{x,y}$ from 2013, 2015, 2017, 2018, and 2019 (not shown).

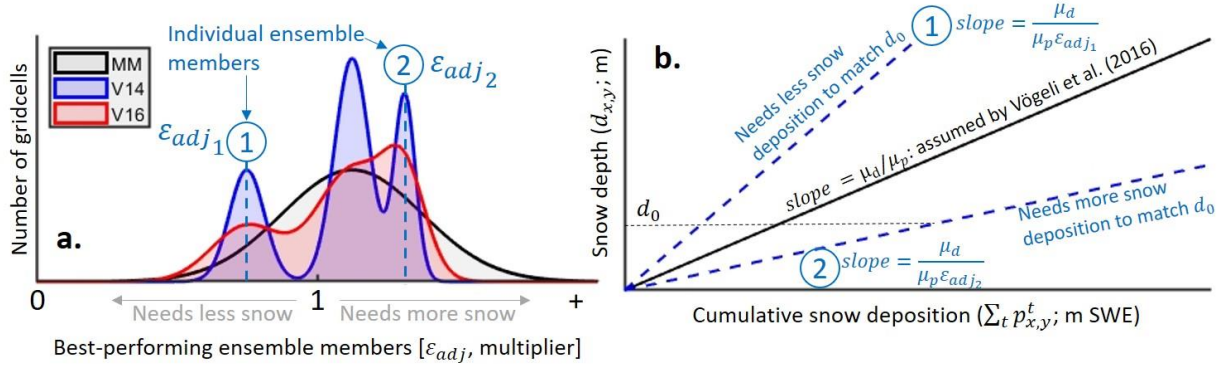

Figure 3.4. Conceptual diagram showing the distribution of the best-performing ensemble members (ε_{adj}) that corrected gridcell-by-gridcell snow depth errors for simulations with terrain-based downscaling (MM), and pattern-based downscaling (V14 and V16) (a). The linear relationship between cumulative snow deposition and snow depth assumed by the pattern-based downscaling (b, solid line) is compared versus the linear relationships for two different ε_{adj} (b, dashed lines).

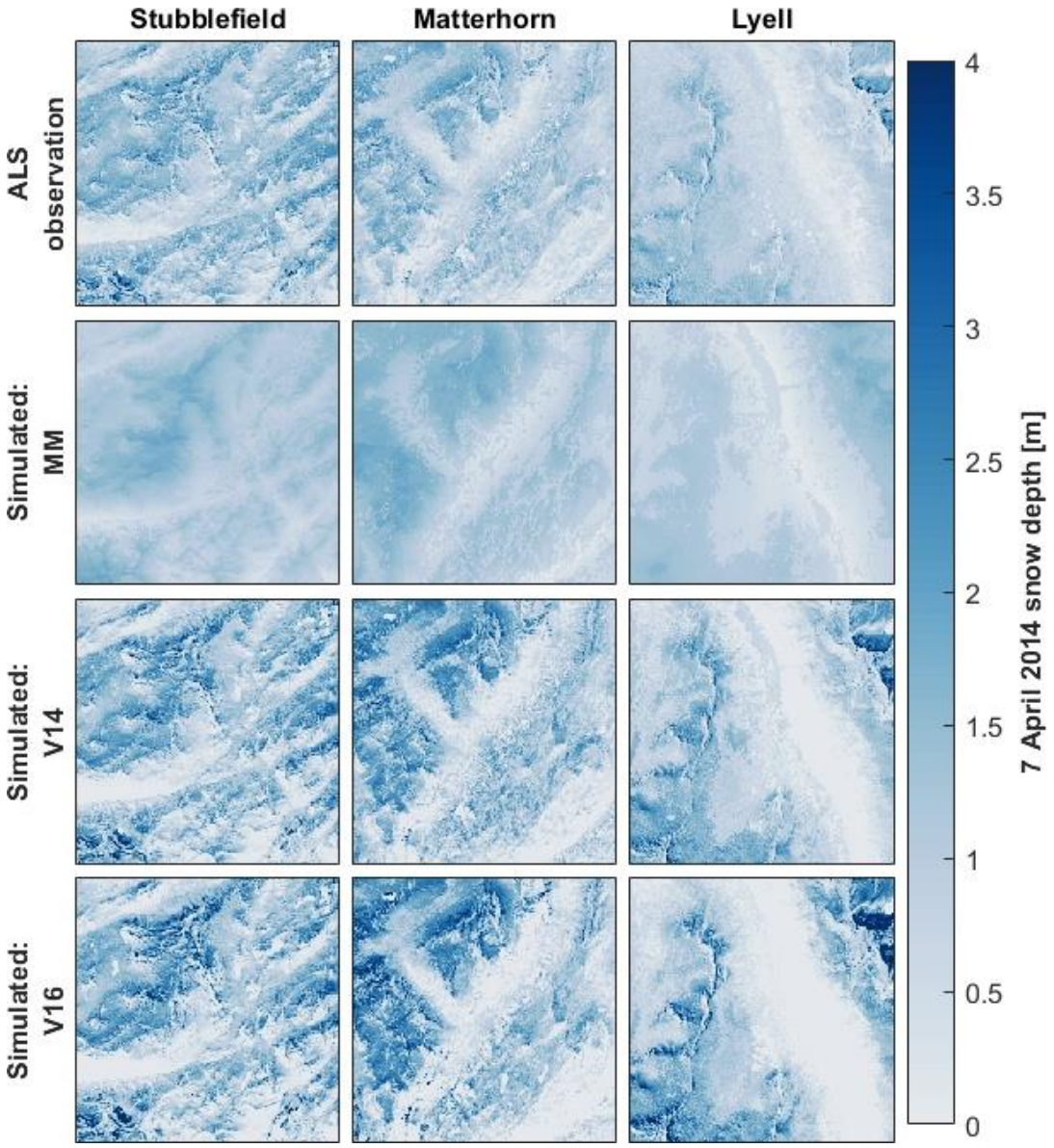

Figure 3.5. 7 April 2014 snow depth (colorbar) for each domain (columns) observed by ALS (top) and simulated using snow deposition ($\varepsilon = 1.0$) downscaled using terrain-based downscaling (MM, second row) and pattern-based snow deposition downscaling factors ($V_{x,y}$) from 7 April 2014 (V14, third row) and 16 April 2016 (V16, fourth row).

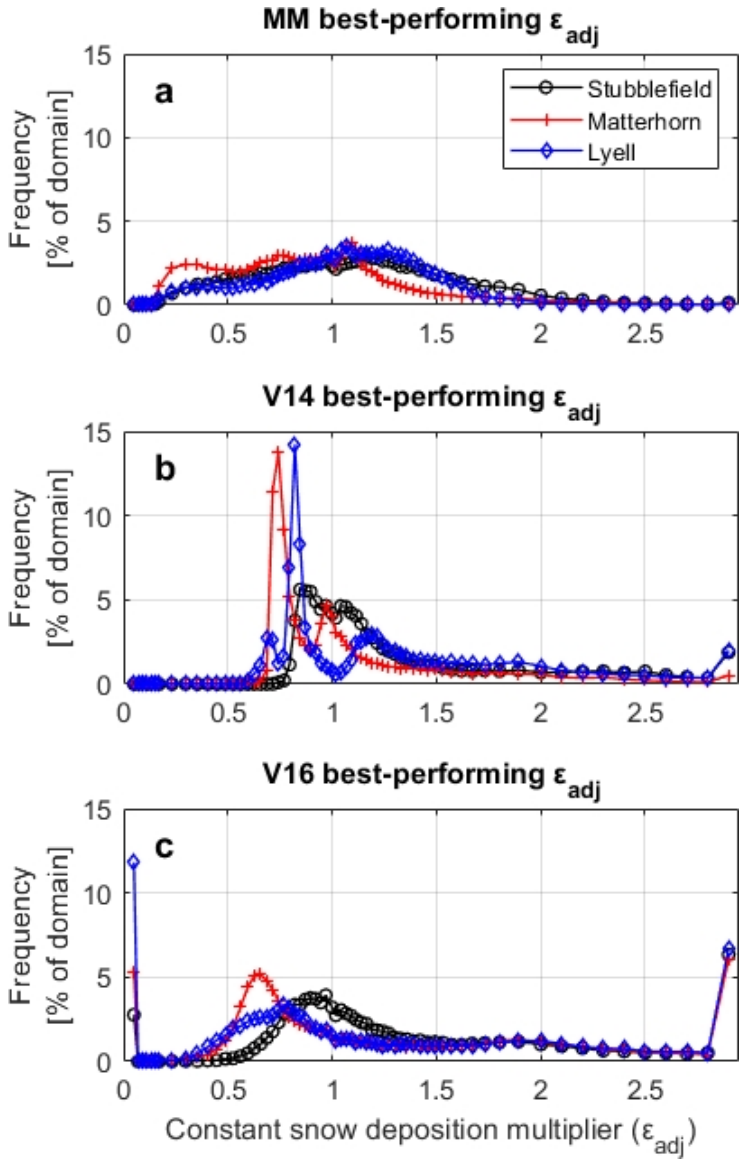

Figure 3.6. Frequency (percent of each domain's gridcells) of the ensemble-members (ε_{adj}, constant in time) best-reproducing the 7 April 2014 target ALS observation in each domain (line color and marker). Results are for the terrain-based snow deposition downscaling (MM, a), and pattern-based downscaling using the snow deposition downscaling factor ($V_{x,y}$) from 7 April 2014 (V14, b), and $V_{x,y}$ from 16 April 2016 (V16, c).

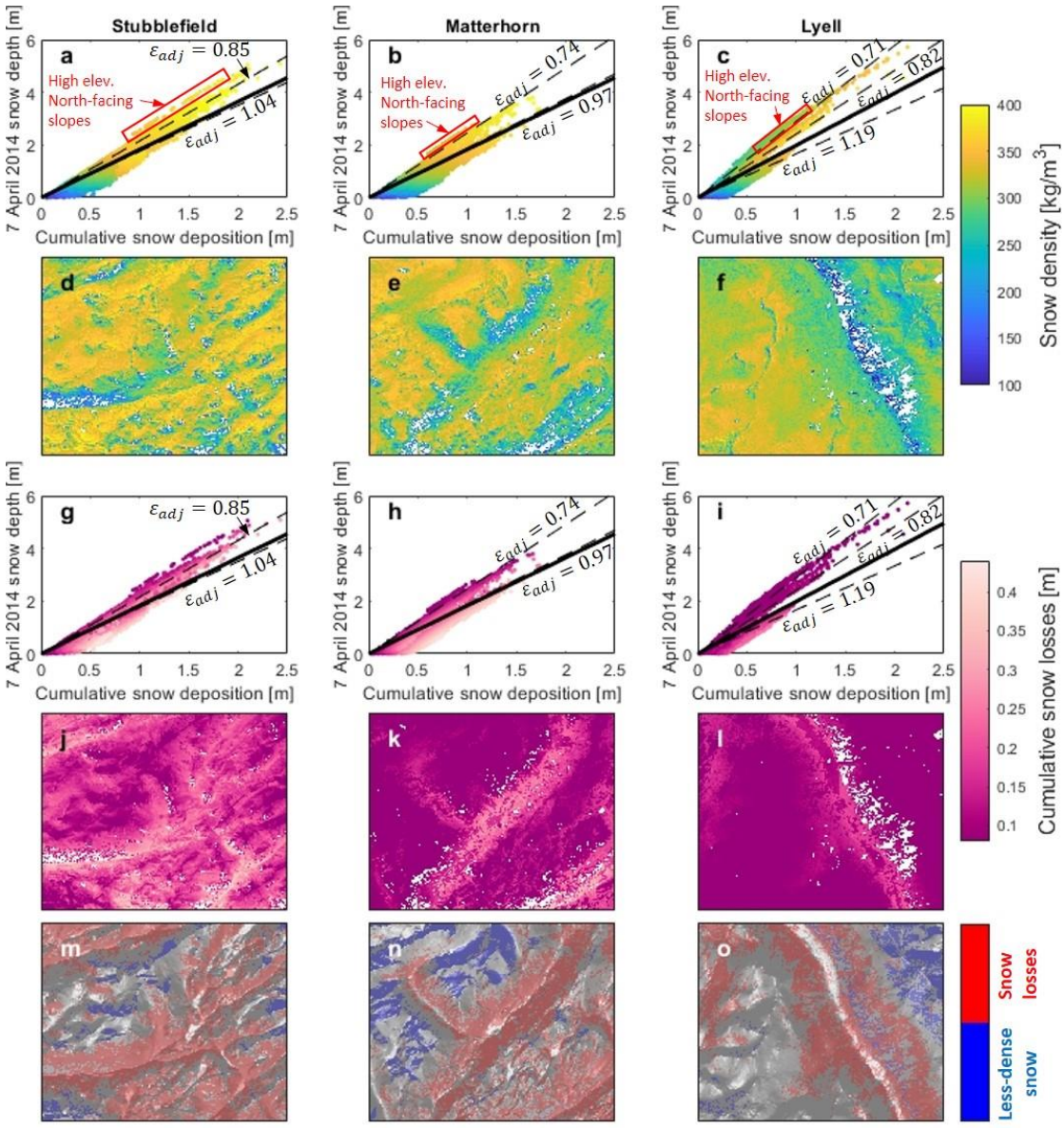

Figure 3.7. The impact of 7 April 2014 snow density (a-f) and winter snow losses to snowmelt and snow sublimation (g-l) is shown for the best-performing V14 ensemble members in each domain (columns). The scatter plots in subplots a-c and g-i are identical, but colored in accordance to the spatial data in the subplots below them. For visualization purposes, the 25 m pixels that diverged the most from the pattern-based snow deposition downscaling assumptions were plotted over a Hillshaded map of each domain (m-o).

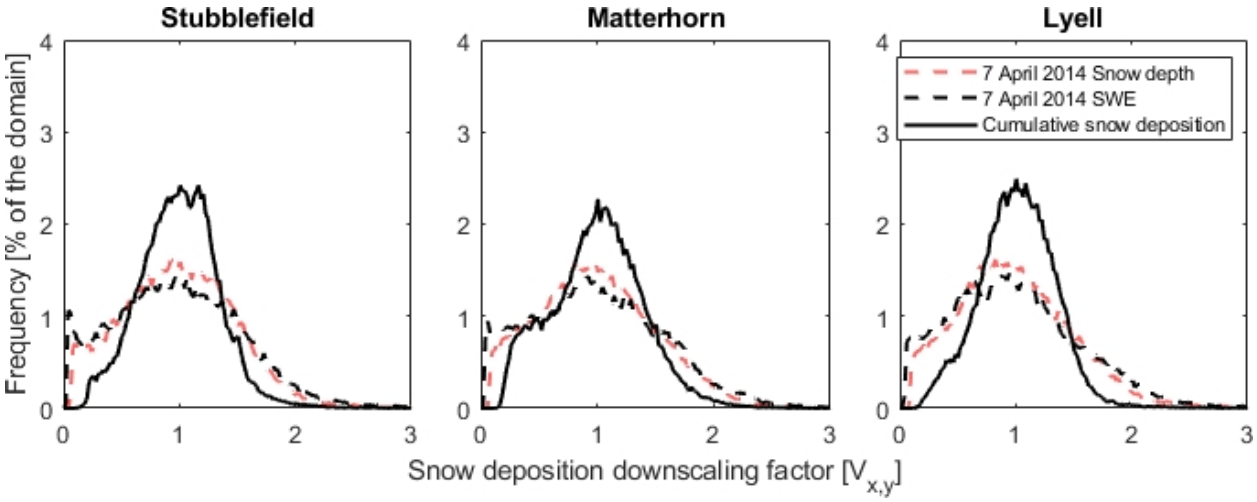

Figure 3.8. The distribution of snow deposition downscaling factors ($V_{x,y}$) calculated using 7 April 2014 snow depth (red, dashed), 7 April 2014 SWE (black, dashed), and cumulative snow deposition prior to 7 April 2014 (black, solid). Distributions shown by black lines were calculated using the V14 best-performing ensemble members (ε_{adj}).

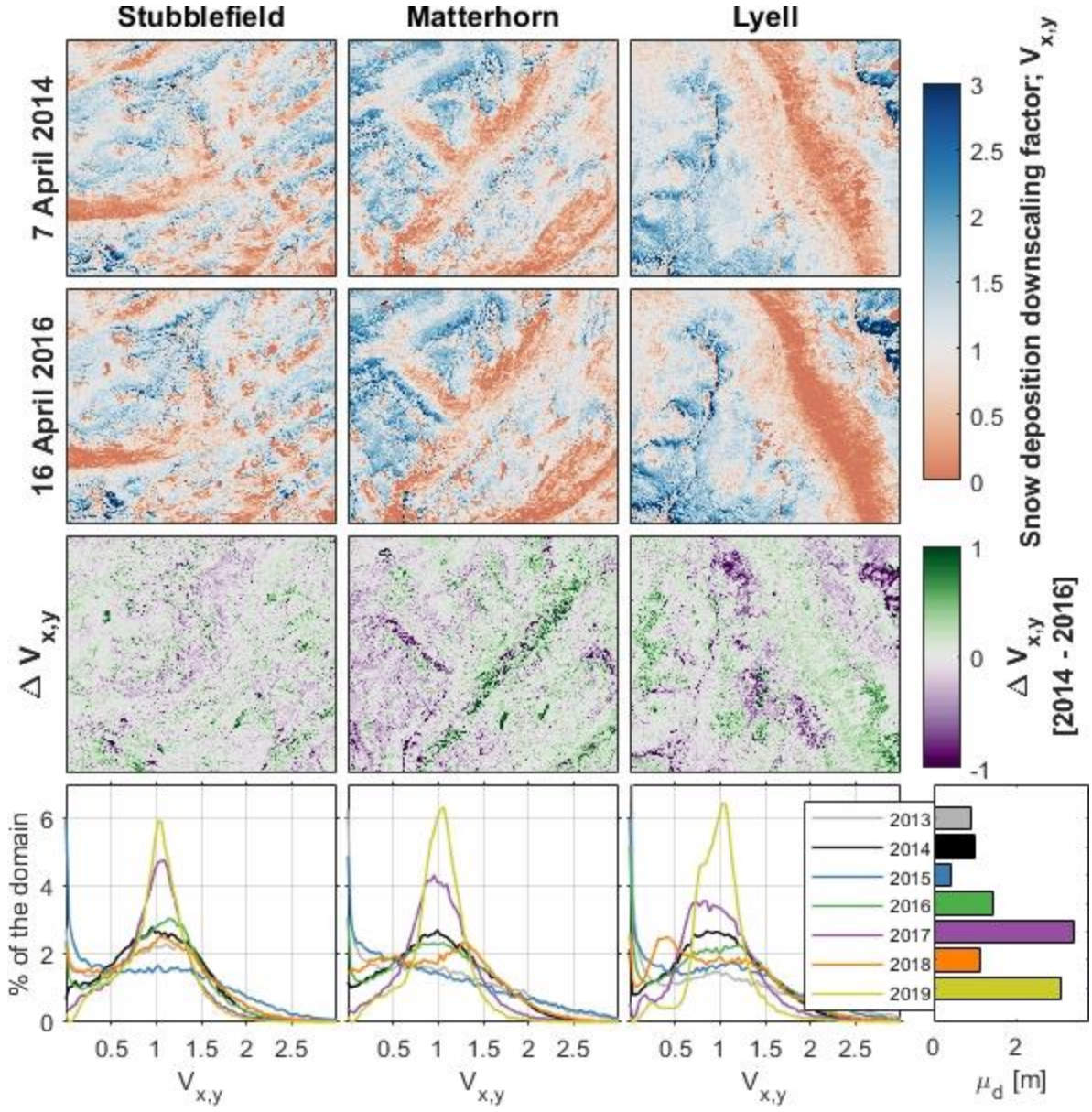

Figure 3.9. Snow deposition downscaling factors ($V_{x,y}$) calculated from ALS observations on 7 April 2014 (top row) and 16 April 2016 (second row) in the three domains (columns). Spatial differences in $V_{x,y}$ (7 April 2014 – 16 April 2016) are shown in the third row. The histogram of $V_{x,y}$, and the mean snow (μ_d, including all three domains) are also compared for ALS observations nearest peak-snowpack timing for 7 years (bottom row).

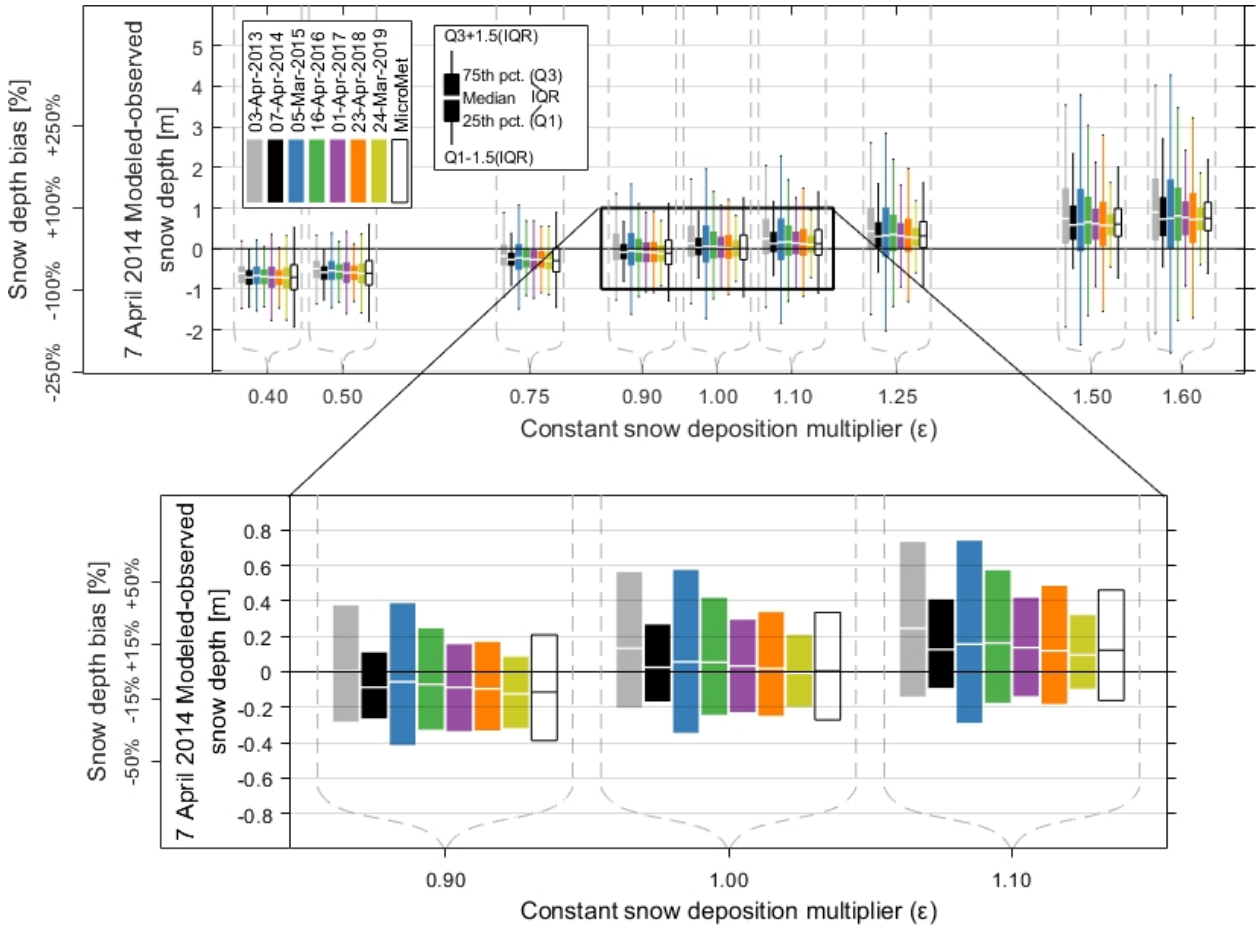

Figure 3.10. Simulated snow depth errors for all domains on 7 April 2014. Snow deposition was downscaled using terrain-based MicroMet lapse rates (white with black outline) and pattern-based downscaling using $V_{x,y}$ from ALS observations in different years (colors) across a range of constant (in space and time) snowfall biases (x-axis).

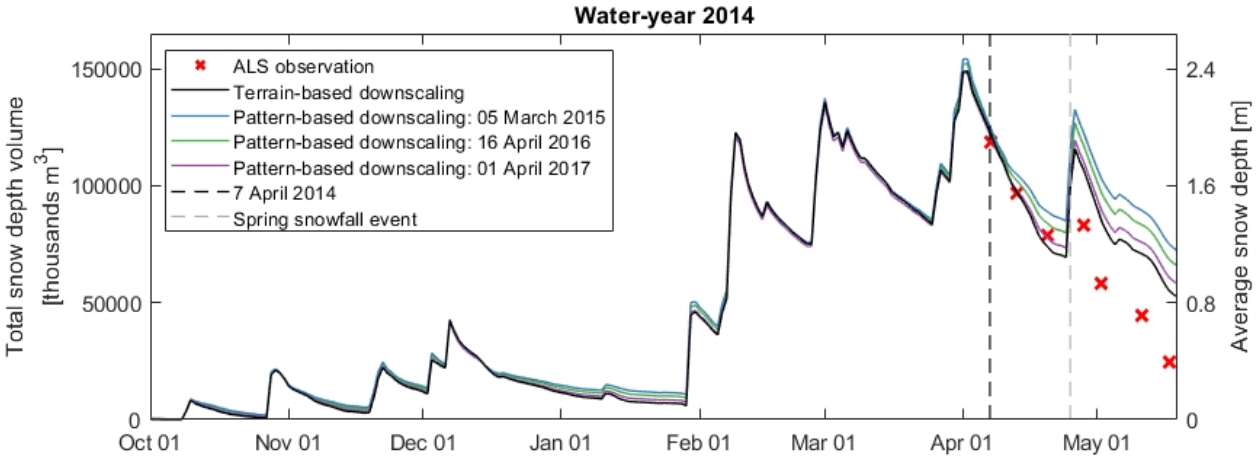

Figure 3.11. Water-year 2014 total (summed) snow depth volume across the Stubblefield, Matterhorn, and Lyell domains. Snow depth was simulated using terrain-based snow deposition downscaling (black line) and pattern-based downscaling using snow depth patterns from 2015, 2016, and 2017 (colored lines) ($\varepsilon = 1.0$, constant in space and time). Simulated snow depth was compared versus airborne lidar observations (markers). 7 April 2014 and a spring snowfall event are also marked.

Chapter 4 : Inferring watershed-scale mean snow magnitude and distribution using multidecadal snow reanalysis patterns and snow pillow observations

J.M. Pflug[1], S.A. Margulis[2], J.D. Lundquist[1]

[1] Department of Civil and Environmental Engineering, University of Washington, Seattle, WA, USA. [2] Department of Civil and Environmental Engineering, University of California Los Angeles, Los Angeles, CA, USA.

4.1. Abstract

The magnitude and spatial heterogeneity of snow deposition are difficult to model in mountainous terrain. Here, we investigated how snow patterns from a 32-year (1985 – 2016) snow reanalysis in the Tuolumne, Kings, and Sagehen Creek California Sierra Nevada watersheds could be used to improve simulations of winter snow deposition (water input from snowfall and redistribution). Remotely-sensed fractional snow-covered area from dates following peak-snowpack timing were used to identify dates from different years with similar snow accumulation and depletion patterns. Historic snow accumulation patterns were then used to 1) relate snow accumulation observed by snow pillows to watershed-scale estimates of mean snowfall, and 2) estimate 90 m snow deposition. Finally, snow deposition fields were used to force snow simulations, the accuracy of which were evaluated versus airborne lidar snow depth observations. Except for water-year 2015, which had the shallowest snow estimated in the Sierra Nevada, snow accumulation and depletion patterns identified from historic dates with spatially correlated fractional snow-covered area agreed on average, with absolute differences of less than 10%. Watershed-scale mean snowfall inferred from the relationship between historic snow accumulation patterns and snow pillow observations had a $\pm$13% interquartile range of biases between 1985 and 2016. Finally, simulations using historic snow accumulation patterns and snow accumulation from snow pillows had snow depth coefficients of correlations and mean absolute errors that improved by 70% and 27%, respectively, as compared to simulations using more common forcing and snow deposition downscaling techniques. This work demonstrates the real-time benefits of satellite-era snow reanalyses in mountainous regions with uncertain snowfall magnitude and spatial heterogeneity.

4.2. Introduction

The spatial heterogeneity of snow regulates the rate of snow cover depletion and resulting mass and energy fluxes between the land surface and atmosphere (Aas et al., 2017; Minder et al., 2016; Mott et al., 2017). In mountainous terrain, the spatial pattern of snow accumulation often controls the timing and rate of spring snowmelt as much, or more than, the heterogeneity of melt energy (Egli et al., 2012; Luce et al., 1998; Lundquist and Dettinger, 2005). However, the physical processes that control how snow accumulates at a variety of spatial scales (e.g., orographic precipitation gradients, preferential deposition, wind-redistribution, avalanching, etc.) are difficult to model, particularly in mountainous regions where the meteorology and snowpack characteristics that drive snow accumulation and redistribution processes are spatially variable (Musselman et al., 2015; Reynolds et al., 2021).

Fortunately, snow distribution is the result of superimposed snow accumulation and depletion processes, many of which display interannually repeatable characteristics (Deems et al., 2008; Marks and Dozier, 1992; Pflug and Lundquist, 2020; Schirmer et al., 2011; Schirmer and Lehning, 2011; Sturm and Wagner, 2010; Vögeli et al., 2016; Woodruff and Qualls, 2019). Vögeli et al. (2016) demonstrated how coarser-resolution snowfall could be downscaled in mountainous terrain using a snow depth distribution pattern from a different year to scale snow input as a substitution for more complex processes like wind redistribution and preferential deposition. Pflug et al. (2021) examined this method in the California, USA, Tuolumne watershed during a shallow-snow year. They downscaled snowfall from a 6 km atmospheric model to 25 m spatial resolution using snow depth patterns from airborne lidar observations in 7 different years. This work identified multiple issues that could occur for pattern-based snow deposition downscaling. Ordered from most- to least-impacting, these issues included: 1) snowfall biases from atmospheric models, 2) snow depth patterns that were influenced not only

by snow accumulation, but also by snow depletion (snowmelt and snow sublimation) and spatially variable snow density, and 3) interannual changes to snow depth patterns.

Here, we test whether historic snow patterns from a daily snow water equivalent (SWE) product spanning 32 years (Margulis et al., 2016a) could rectify the issues listed above. Specifically, we use information about satellite-observed snow cover to identify historic dates with similar snow patterns (corresponding to issue 3). On these historic dates, we calculate normalized snow accumulation patterns from daily cumulative increases in SWE, thereby ignoring the effect of snowmelt, snow sublimation, and snow density (corresponding to issue 2). We then use the relationship between historic snow accumulation patterns, and SWE accumulation from snow pillow observations, to infer mean snowfall at watershed-scale spatial extents (corresponding to issue 1) across 3 different Sierra Nevada watersheds. We ask three overarching questions:

1. How can real-time information about snow cover identify the best-matching historic snow accumulation and depletion patterns from a multidecadal (1985 – 2016) snow reanalysis?
2. Can context provided from historic snow accumulation patterns be used to infer watershed-scale estimates of mean snowfall from current-season snow pillow observations?
3. How accurate is simulated snow depth when forced with snow pillow observed snow accumulation (question 2) and historic snow accumulation patterns (question 1)?

This work applies historic snow accumulation patterns in real-time modeling scenarios. These methods are therefore of practical use for both scientific and operational purposes.

4.3. Background

4.3.1. Interannual snow pattern repeatability

The spatial distribution of snow accumulation is driven by processes that tend to have time-repeatable characteristics. At fine spatial scales (< 10 m), snow creep and snow-particle

bouncing fill depressions and smooth the snow surface (Filhol and Sturm, 2019). At similar spatial scales, wind redistribution and preferential deposition deposit snow in regions alee of obstructions (e.g., vegetation and ridges) (Dadic et al., 2010; Liston et al., 2007; Mott et al., 2010) and snow is intercepted in the forest canopy (Hedstrom and Pomeroy, 1998; Lundquist et al., 2013; Moeser et al., 2015). At larger spatial scales (~100-10,000 m), elevation is commonly cited as the dominant driver of spatial variations in snow deposition (e.g., Balk and Elder, 2000; Clark et al., 2011; Elder et al., 1991; McGrath et al., 2018). Along with elevation, the slope and aspect of terrain can control the rate and occurrence of orographic uplifting, condensation, and resulting precipitation (Minder et al., 2008; Mott et al., 2014; Roe and Baker, 2006), and the amount of snow a slope can hold before sloughing to lower-elevation less-steep terrain (e.g., Bernhardt and Schulz, 2010). Many of the processes listed above are driven by the relationship between terrain and local meteorology, which at the conclusion of winter snow accumulation, often result in interannually repeatable snow distribution (Deems et al., 2008; Hiemstra et al., 2002; McGrath et al., 2018; Pflug and Lundquist, 2020; Schirmer et al., 2011; Schirmer and Lehning, 2011; Sturm and Wagner, 2010; Vögeli et al., 2016).

Snow cover and snow depth spatial distribution is influenced not only by snow accumulation, but also snowmelt and sublimation. In the California Sierra Nevada, Marks and Dozier (1992) found that incoming energy was dominated (66 – 90% of total energy) by shortwave and longwave radiation. In sunny regions like the California Sierra Nevada, solar shading and air temperature gradients are largely influenced by the terrain and vegetation, resulting in radiation patterns that vary little from year to year. Although both snow accumulation and depletion are influenced by static terrain features, they are influenced in different ways, resulting in snow accumulation and depletion patterns that can be uncorrelated in space.

4.3.1. Snowfall spatial patterns

Snow pillow observations, which calculate SWE from the weight of overlying snowpack, have been used to evaluate the accuracy of gridded precipitation products (e.g., Hughes et al., 2017; Lundquist et al., 2015), and characterize interannual differences in meteorological conditions (e.g., Montoya et al., 2014; Nayak et al., 2010; Pflug and Lundquist, 2020). However, even a large sample of snow pillow observations may not necessarily be representative of snowfall at watershed-spatial extents (Dressler et al., 2006; Harshburger et al., 2010; Meromy et al., 2013; Molotch and Bales, 2005). To address the disconnect between point (snow pillow) and watershed spatial scales, gridded precipitation products (e.g., Albers et al., 1996; Daly et al., 1997; Livneh et al., 2015; Thornton et al., 2014) are widely used. These approaches use the relationship between point observations and the coarser-scale gridded products that observations are within, to spatially interpolate gridded estimates of precipitation (e.g., Hamlet and Lettenmaier, 2005; Maurer et al., 2002). However, these approaches assume time-consistent precipitation patterns when years with anomalous meteorological conditions may diverge from these patterns by larger amounts. For instance, Lundquist et al. (2015) found that normalized precipitation products were unable to capture precipitation gradients associated specifically with post-cold-frontal storms in the California Sierra Nevada. Gutmann et al. (2012) also noted that normalized precipitation products averaged from past periods may be less-likely to represent changes to precipitation patterns in future climates. Additionally, precipitation products are sometimes extrapolated from ground observations, which are often scarce and preferentially located in lower-elevation easily accessible locations (Groisman and Legates, 1994; Henn et al., 2016; Hughes et al., 2017; Lundquist et al., 2019). As a result, precipitation products often decrease in accuracy at locations further from observations (Gutmann et al., 2012; Hiemstra et al., 2006), particularly at high elevations (Wayand et al., 2013).

4.4. Study domains

We focused on domains where airborne lidar snow depth data overlapped with the 32-year SWE reanalysis used to derive snow patterns (both datasets are detailed in Section 4.5.1). These domains included the Upper Kings, Upper Tuolumne, and Sagehen Creek watersheds, located at approximately 36.8°N, 38.1°N, and 39.4°N, respectively, in the California Sierra Nevada (Figure 4.1). Domain extents were restricted to only gridcells with airborne lidar snow depth observations. As a result, the Sagehen domain (73 km^2) was less than 7% the size of the Tuolumne (1166 km^2) and Kings (1107 km^2) domains. As opposed to the Tuolumne and Kings domains, which ranged from approximately 1430 – 4300 m in elevation, Sagehen Creek elevation only varied by 867 m (between 1833 and 2700 m). Forested regions covered approximately 28%, 43%, and 93% of the Upper Kings, Upper Tuolumne, and Sagehen Creek domains, respectively.

4.5. Data and snow model

4.5.1. Data

Daily SWE data were extracted from 58 snow pillows, maintained by the California Data Exchange Center (CDEC) and Natural Resources Conservation Service (Snow Telemetry stations; SNOTEL). Figure 4.1 displays the distribution of these snow pillows and the year in which data became available for each. Cumulative snow deposition was calculated in each water-year (from 1 October) using daily cumulative increases in SWE from each snow pillow. Cumulative increases in SWE across gaps of 2 days or less were filled by assuming equal snowfall on each day. Periods following SWE data gaps exceeding 2 days in length were excluded.

Airborne lidar survey (ALS) snow depth observations collected by the Airborne Snow Observatory (Painter et al., 2016) were used to evaluate the accuracy of snow simulations performed in each domain. To our knowledge, no other region in the world has been observed by

airborne lidar as frequently as the Upper Tuolumne watershed, making this watershed a foundational area for snow research (e.g., Cannistra et al., 2021; Currier and Lundquist, 2018; Hedrick et al., 2018; Henn et al., 2016; Margulis et al., 2019; Pflug and Lundquist, 2020). We also included ALS observations collected over the Upper Kings and Sagehen Creek watersheds, observed in water-years 2015 and 2016, respectively. For comparison purposes, airborne lidar observations were regridded from their native spatial resolution (3 m) to the SWE reanalysis 90 m grid. At these spatial resolutions, ALS snow depth was likely accurate to within $\pm$0.05 m (Margulis et al., 2019), and larger-scale terrain characteristics like elevation, slope, and aspect were the dominant drivers of snow spatial variability (e.g., Clark et al., 2011; McGrath et al., 2018; Trujillo et al., 2007).

The Sierra Nevada Snow Reanalysis (Margulis et al., 2016a) was used to calculate snow accumulation and snow depletion patterns. This product provides daily estimates of SWE and fractional snow-covered area (fSCA) at 90 m spatial resolution over the California Sierra Nevada from 1985 – 2016. A fully Bayesian approach is used to constrain prior estimates of SWE and fSCA from a coupled land surface model (Xue et al., 1991) and snow depletion curve representation of sub-grid snow distribution (Liston, 2004), to generate posterior estimates of fSCA (and the resulting SWE) that most-closely correspond to Landsat fSCA observations. This product exhibited high accuracy versus snow pillow observations (Margulis et al., 2016a), and has since been used to appraise the value of SWE information on streamflow forecasting (Li et al., 2019), and characterize the hydrologic impact of an extreme snow drought (Margulis et al., 2016b). Readers are referenced to Margulis et al. (2016a) and Margulis et al. (2015) for more information. Posterior SWE estimates were extracted for the 90 m gridcells 1) overlapping snow pillows, and 2) across each of the study domains (Figure 4.1). Because the reanalysis

retroactively estimates SWE, this product is not available in real-time. However, the reanalysis provides spatiotemporally continuous estimates of SWE, which could be used to diagnose the daily increases (from snow deposition) and decreases (from snowmelt and sublimation) in SWE from previous years. For the domains investigated here, the reanalysis matched ALS snow depth patterns and fSCA observations well (Appendix C), suggesting a high degree of accuracy.

4.5.2. Snow model

Real-time modeling scenarios were performed in this study, downscaling snow deposition using historic snow accumulation patterns defined from the SWE reanalysis (more in Section 4.6). Snow was simulated at 90 m spatial resolution using SnowModel (Liston and Elder, 2006a), with snowpack liquid water percolation additions from Pflug et al. (2019). This model has been applied in mountainous terrain (e.g., Hiemstra et al., 2002; Reynolds et al., 2021; Sturm and Wagner, 2010) where SnowModel has the unique capability of representing snow redistribution via wind. However, the reanalysis used in this study retroactively inferred 90 m SWE using Landsat observations of snow cover evolution and the associated snowmelt (Margulis et al., 2015). Put another way, the reanalysis likely included the impact of multiple snow processes, including wind redistribution, on SWE spatial patterns. Therefore, wind redistribution was disabled in SnowModel to avoid double-counting this process. However, the impact of wind-redistribution (which typically influences snow distribution at spatial scales of 10 m or less) on 90 m mean snow depth was likely minimal (e.g., Clark et al., 2011; Mott et al., 2010; Trujillo et al., 2007).

Meteorological forcings for the snow simulations were provided from 1/8° (~12 km) hourly air temperature, relative humidity, and wind speed from the North American Land Data Assimilation System Phase 2 (NLDAS2) (Xia et al., 2012). This forcing was identical to the meteorological forcing used by the SWE reanalysis and is much coarser than the 90 m

simulations performed in this study. Therefore, air temperature, relative humidity, wind speed, and wind direction were downscaled to the model grid using the MicroMet meteorological distribution routine (Liston and Elder, 2006b). 90 m Land-surface elevation and vegetation was regridded from the 30 m Shuttle Radar Topography Mission (SRTM), and 30 m United States Geological Survey landcover classes (Davidson and McKerrow, 2016), respectively. Hourly shortwave radiation and longwave radiation were calculated within SnowModel using cloud fractions calculated from air temperature and dew-point temperature, with terrain-shading (for shortwave radiation) and the cloudy-sky longwave calculation from Iziomon et al. (2003). Except for snow deposition (discussed in Section 4.6), all simulations were provided identical forcing.

Snow depth simulations were performed for three water-years coincident with years with ALS observations. These ALS were used to evaluate simulation accuracy. Simulations included years with lower-than average snowfall (water-year 2014, ALS existed for the Upper Tuolumne only), snow that was among the shallowest ever recorded or estimated in the Sierra Nevada (water-year 2015, Upper Tuolumne and Upper Kings domains), and a year with approximately average snowfall (water-year 2016, Upper Tuolumne and Sagehen Creek domains). To ensure reasonable model accuracy, model parameters (namely, albedo decay parameters and air temperature lapse rates) were calibrated using forcing and observations from snow pillow observations (Figure 4.2) in each simulation year.

4.6. Methods

Methods were split into three phases corresponding to the questions numbered in the introduction. First, cumulative daily increases and decreases in SWE from the reanalysis were used to calculate snow accumulation and depletion patterns (Figure 4.3, historic date) in all years (1985 – 2016). We then investigated how remotely-sensed fSCA observations could be used to

search the reanalysis (Figure 4.3, top) for a period, in a different year, with similar snow accumulation and depletion patterns (more on this in Section 4.6.1). In Section 4.6.2, we calculated domain mean cumulative snowfall two ways: 1) using the relationship between observed increases in SWE at snow pillows and a snow accumulation pattern from a historic date, and 2) mean snowfall provided from NLDAS2 (Figure 4.3, simulation year). Domain mean cumulative snowfall, and historic snow accumulation patterns, were finally used to derive snow deposition at 90 m spatial resolution. The snow deposition fields were used as forcing for snow simulations, which were compared versus ALS snow depth observations in water-years 2014, 2105, and 2016 (more on this in Section 4.6.3).

4.6.1. Snow pattern identification and comparison

Using the reanalysis, daily cumulative increases in SWE (SWE^+) and decreases in SWE (SWE^-) were calculated on a gridcell-by-gridcell (x) basis,

$$SWE^+(x, doy, yr) = \sum_{t=1}^{t=doy} \frac{\Delta SWE_{x,t,yr}}{\Delta t}, where/when\ \frac{\Delta SWE_{x,t,yr}}{\Delta t} > 0, \quad (4.1a)$$

$$SWE^-(x, doy, yr) = \sum_{t=1}^{t=doy} \frac{\Delta SWE_{x,t,yr}}{\Delta t}, where/when\ \frac{\Delta SWE_{x,t,yr}}{\Delta t} < 0, \quad (4.1b)$$

where *doy* is the day of the water-year (*doy* = 1 on October 1st), *yr* is the water-year, and Δt is equal to one day. To compare snow accumulation and depletion between years with different snowfall magnitudes, we normalized both snow accumulation and depletion patterns,

$$Pattern^+(x, doy, yr) = \frac{SWE^+(x, doy, yr)}{\mu_{SWE^+}(doy, yr)}, \quad (4.2a)$$

$$Pattern^-(x, doy, yr) = \frac{SWE^-(x, doy, yr)}{\mu_{SWE^+}(doy, yr)}, \quad (4.2b)$$

where μ_{SWE^+} is spatially averaged SWE^+, or domain mean cumulative snowfall, to date, from October 1st in a given year (Figure 4.3, shaded plane). We normalized by μ_{SWE^+} since both SWE^+ and SWE^- are sensitive to μ_{SWE^+}. For instance, the timing and amount of snow accumulation is a major influence on winter snowmelt as snowpacks with less SWE typically have less cold content and therefore take less energy to bring to ripe, melting conditions.

The similarity between the snow accumulation and depletion patterns from two dates (doy_1 and doy_2) in different years (yr_1 and yr_2) was calculated using the mean absolute error (MAE) of all gridcells (x), including both the accumulation and depletion patterns,

$$MAE_{1,2} = mean \begin{bmatrix} abs|Pattern^+(x, doy_1, yr_1) - Pattern^+(x, doy_2, yr_2)|; \\ abs|Pattern^-(x, doy_1, yr_1) - Pattern^-(x, doy_2, yr_2)| \end{bmatrix}. \quad (4.3)$$

Since Equation 4.3 compares normalized snow accumulation and depletion, MAE represents the average departure between snow accumulation and depletion on two dates, relative to mean snow accumulation on each date. For example, for MAE = 0.10, snow accumulation and depletion on two dates differed by an amount, on average, that was 10% the domain-mean cumulative snowfall (μ_{SWE^+}). This can also be conceptualized as the mean absolute difference of both snow accumulation and depletion, in meters, between two years with μ_{SWE^+} = 1.00 m (approximately the average annual snowfall in the Upper Tuolumne watershed). For snow patterns to match ($MAE \to 0$), SWE^+ and SWE^- from two dates must be spatially correlated and have identical coefficients of variation.

Despite the time-static drivers of both SWE^+ and SWE^- (Section 4.3.1), the spatial dissimilarity of how snow accumulates and depletes suggests that periods from different years with similar relative amounts of snow accumulation and depletion may be more likely to display similar snow distribution. In fact, Pflug and Lundquist (2020) found that the snow depth coefficient of variation in the Tuolumne watershed was well-explained (r = 0.93) by the fraction

of the domain covered with snow (snow-covered fraction, SCF). This suggested that real-time observations of snow cover could be used to identify dates with similar snow spatial variability from previous years. Since the reanalysis assimilates Landsat fSCA observations (Section 4.5.1), fSCA modeled by the reanalysis typically matched fSCA observations closely throughout early spring snowmelt (Appendix C). Therefore, fSCA from the reanalysis was extracted from dates with cloud-free Landsat observations between peak-SWE and mid-snowmelt (Figure 4.4, gray lines), and was used to approximate 90 m observations of fSCA. MAE (Equation 4.3) was then calculated on each Landsat observation date versus other dates over the historical record selected using three criteria:

1. Periods following peak snow volume from all other years (n = 31) with matching SCF.
2. The date from a different year (n = 1) with the best spatially correlated fSCA.
3. The date from a different year (n = 1) with the smallest MAE.

In practice, criteria 1 and 2 could use real-time satellite observations to select historic dates. Criterion 3 was instead used as a baseline to identify how repeatable snow accumulation and depletion patterns could be for each domain and Landsat observation date. In this study, we searched the reanalysis using these three criteria, including years both before and after each Landsat observation date. Results were representative of the degree to which dates with similar snow patterns could be identified today given the interannual variability of snow in the 32-year reanalysis (Figure 4.4). We used Landsat observation dates since these observations could easily be coarsened to the resolution of the reanalysis (90 m). However, we acknowledge that this represents a conservative case as more-frequent (~daily) fSCA observations from the MODerate-resolution Imaging Spectrometer (MODIS) could also be used to query the reanalysis for similar historic periods.

4.6.2. Domain mean cumulative snowfall calculations

The SWE reanalysis used in this study was highly accurate (Appendix C). In fact, Margulis et al. (2016a) found that the mean SWE error versus Sierra Nevada snow pillow observations was only 0.03 m. However, snow pillow observed daily cumulative increases in (SWE^{+}_{pillow}) disagreed by larger amounts, relative to daily cumulative increases in SWE from the reanalysis (Figure 4.5, left). This was a result of both 1) differences in the spatiotemporal evolution of snow accumulation at point-scales (snow pillows) and 90 m gridcells (from the reanalysis), and 2) noise in the snow pillow daily SWE observations, which sometimes appeared as spurious snowfall events. Here, we used a multilinear regression (Appendix D) to adjust SWE accumulation observed by the snow pillows for annually-persistent differences between point-scale (snow pillow) and 90 m (SWE reanalysis) cumulative increases in SWE. Using only SWE^{+}_{pillow} and the number of snowfall events as predictive variables (both of which are observable in real-time), SWE accumulation at snow pillows estimated by the multilinear regression was biased versus the reanalysis by less than 2% (r = 0.94), with randomly distributed differences (Figure 4.5, right) and accuracies similar to those reported by Margulis et al. (2016a).

Historic snow accumulation patterns ($Pattern^{+}$) were used to relate SWE^{+}_{pillow} point-observations to watershed-scale estimates of mean snowfall (Figure 4.3, option #1). For each year (1985 – 2016), we:

1. Extracted the cumulative daily increases in SWE from all snow pillows (SWE^{+}_{pillow}) within 1° (latitude/longitude) of each modeling domain and identified the first cloud-free Landsat observation date following peak-SWE timing.
2. Transformed SWE^{+}_{pillow} using the multilinear regression (Figure 4.5), trained across all other years (n = 31, excluding the current year).

3. Selected historic snow accumulation patterns ($Pattern^+$) from other years using the selection criteria from Section 4.6.1.
4. Calculated estimates of μ_{SWE^+} using the ratio between SWE^+_{pillow} (step 2 above) and $Pattern^+$ (step 3).

Additionally, μ_{SWE^+} was calculated directly from the SWE reanalysis in each domain. Since the SWE reanalysis was unbiased (Margulis et al., 2016a), μ_{SWE^+} from the reanalysis was assumed to be an unbiased estimate of total snow accumulation. Therefore, μ_{SWE^+} from the reanalysis was used to assess the accuracy of median μ_{SWE^+} calculated from snow pillows (listed steps above).

Finally, μ_{SWE^+} was also calculated from NLDAS2 (Section 4.5.1). However, the coarse-resolution of NLDAS2 (1/8°) caused gridded precipitation estimates to include terrain beyond the irregular borders of each domain. Therefore, MicroMet (Liston and Elder, 2006b) was used to downscale precipitation from NLDAS2 to the 90 m reanalysis grid within the boundaries of each domain. This approach uses a combination of inverse-distance weighted spatial interpolation and terrain-based precipitation lapse rates, and is conceptually similar to other popular terrain-based precipitation downscaling methods (Bavay and Egger, 2014; Daly et al., 2008; Thornton et al., 1997). Precipitation was then partitioned into rainfall and snowfall using air temperature and a 2°C threshold (corresponding to the precipitation partition used by the snow reanalysis). NLDAS2 μ_{SWE^+} was then calculated from the average (spatially) cumulative snowfall in each domain (Figure 4.3, option #2).

4.6.3. Snow depth simulations

90 m snow deposition fields (SWE^+) were used to force snow simulations. SWE^+ was calculated using different combinations of snow accumulation patterns ($Pattern^+$, Section 4.6.1), pillow-estimated domain mean cumulative snowfall (μ_{SWE^+}, Section 4.6.2), NLDAS-

estimated μ_{SWE^+} (Section 4.6.2), and NLDAS2 snowfall distributed using MicroMet (previous paragraph). Simulations included:

1. **NL2+R:** daily increases in μ_{SWE^+} from NLDAS2, distributed in space using (multiplied by) $Pattern^+$ from the historic date with the best-correlated fSCA (criterion 2 from Section 4.6.1).
2. **NL2+Best:** daily increases in μ_{SWE^+} from NLDAS2 distributed in space using $Pattern^+$ from the historic date with the smallest MAE (criterion 3 from Section 4.6.1).
3. **SNO+R:** daily increases in median μ_{SWE^+} estimated from snow pillows using $Pattern^+$ from the historic date with the best-correlated fSCA, and distributed in space using the same pattern.
4. **SNO+Best:** daily increases in median μ_{SWE^+} estimated from snow pillows using $Pattern^+$ from the historic date with the smallest MAE, and distributed in space using the same pattern.
5. **NL2+MM:** Precipitation from NLDAS2, downscaled using terrain-based precipitation lapse rates (MicroMet; Liston and Elder, 2006b).

Following methods from Vögeli et al. (2016), cumulative snow deposition (SWE^+) for simulations 1 – 4 was assumed to be the product of daily increases in domain mean cumulative snowfall (μ_{SWE^+}, constant in space but variable in time) and a snow accumulation pattern ($Pattern^+$, variable in space but constant in time) (Figure 4.3). Simulation 5 was instead used as a baseline to represent popular modeling approaches, which commonly downscale precipitation using statistical approaches related to terrain characteristics like elevation. Except for snow input, all simulations were provided identical forcing and were performed at hourly timesteps by partitioning daily snowfall at each gridcell equally across all hours. The accuracy of each

simulation was evaluated versus ALS snow depth observations in the Tuolumne (water-years 2013 – 2016), Kings (water-year 2015), and Sagehen Creek (water-year 2016) domains. Although the reanalysis exhibited high levels of accuracy, we evaluated snow depth simulations versus lidar observations since these observations were independent of $Pattern^{+}$ and $\mu_{SWE^{+}}$ calculated in Sections 4.6.1 and 4.6.2.

4.7. Results

4.7.1. Snow pattern identification and comparison

The spatial variability of SWE (coefficient of variation, CoV), exhibited an interannually repeatable relationship with snow covered fraction (SCF) in each domain (Figure 4.6). This was particularly apparent for the larger Tuolumne and Kings domains, suggesting that information about snow cover, which could be available in real-time from satellite observations, could be used to identify historic periods with similar SWE spatial variability. However, dates from years with different meteorological conditions could approach similar SCF with different relative amounts of snow accumulation and depletion. For instance, as demonstrated by Figure 4.6, the SWE coefficient of variation (CoV) tended to be larger, regardless of the SCF ($0.10 \leq SCF \leq 1.00$), in years with smaller peak-SWE volume. Years with smaller peak-SWE snow volume that were characterized by higher-elevation snow lines and patchier snow cover (e.g., Pflug and Lundquist, 2020), resulting in more SWE spatial variability, relative to mean SWE. Therefore, although SCF was clearly related to SWE CoV in each domain and year, remotely-sensed observations of domain-scale SCF (such as those provided from coarser-resolution satellite observations) may be an insufficient searching criterion (criterion 1 from Section 4.6.1) to identify periods, from previous seasons, that approach similar SWE spatial distributions with similar snow accumulation and depletion patterns. In fact, although the median MAE (Equation 3) between Landsat observation dates and the dates of matching SCF in all other years (n >

1900) was 0.12, 0.13, and 0.12 for the Tuolumne, Kings, and Sagehen domains, respectively, outliers from 0.20 to 0.40 were common (Figure 4.7). To put this another way, the mean absolute differences between snow accumulation and depletion between dates with matching SCF could be 20 – 40% of the magnitude of cumulative snowfall. Comparatively, MAE calculated versus the historic dates of best spatially correlated 90 m fSCA (Figure 4.7, red), which accounted for the spatial increases and decreases in snow cover, tended to identify dates with similar relative amounts of snow accumulation and depletion and had medians of 0.09, 0.11, and 0.11 for the Tuolumne, Kings, and Sagehen domains, respectively. Most-notably, the distribution of MAE were significantly improved ($p < 0.01$) as compared to historic dates selected using SCF (Figure 4.7).

The median MAE of the dates with the smallest MAE (Figure 4.7, blue) (criterion 3 from Section 4.6.1) was 0.09, 0.09, and 0.07 for Tuolumne, Kings, and Sagehen, respectively. However, a number of outliers (MAE > 0.12) still existed. These outliers compared $Pattern^{+}$ and $Pattern^{-}$ from water-year 2015 Landsat observation dates versus other low-snow years like 2013 and 2014. Although SWE volume in water-year 2015 resembled water-years 2013 and 2014 (Figure 4.4), the magnitude of SWE, timing of snowfall, and spatial distribution of SWE in water-year 2015 was unlike any other year in the 32-year record (Figure 4.8). In fact, water-year 2015 in the Sierra Nevada had return period (estimated duration between two years with similar snow volume) of over 600 years (Margulis et al., 2016b), and was among the most snow-scarce years calculated from tree-ring SWE reconstructions (Belmecheri et al., 2016). In Sagehen Creek, water-year 1993 also exhibited unique SWE volumes that were much greater than any other year (Figure 4.4). However, even given the dramatic departure from the typical relationship between SCF and SWE CoV in water-year 1993 in Sagehen Creek (Figure 4.6), the MAE

between the 4 May 1993 Landsat observation date and the date with the smallest MAE (13 April 2010) was only 0.07. Surprisingly, although these two dates were from different parts of the snowmelt season in years with different SWE magnitudes, $Pattern^{+}$ and $Pattern^{-}$ between the two dates were spatially correlated (r = 0.98) and the relative amounts of snow accumulation and depletion (ratio of total snow accumulation to total snow depletion) agreed to within 5%.

The difference in MAE between the dates of spatially correlated fSCA (Figure 4.7, red), and the dates of best MAE (Figure 4.7, blue), were larger in Sagehen Creek than the other two domains. In other words, it was less likely to identify a date with historically representative snow accumulation and depletion patterns using fSCA as a searching criterion. This was attributed to the smaller size and smaller range of elevations in Sagehen Creek which had fSCA that depleted more homogeneously in space and rapidly in time, and as a result, was less-indicative of annual differences in snow accumulation and depletion. For instance, relative to a Landsat observation date from 2 March 2008, the date of best-correlated fSCA (3 April 2006) had a MAE of 0.15, but the date with the best MAE (4 March 2004) had a MAE that was 0.09 better (MAE = 0.06). Despite this, the spatial correlation of fSCA between the Landsat observation date, and 3 April 2006 (spatially correlated fSCA) and 4 March 2004 (best MAE), was almost identical at 0.99 and 0.98, respectively. In Section 4.8.1, we discussed how differences in the size of domains influenced the ability to identify representative historic patterns from the reanalysis.

4.7.2. Domain mean cumulative snowfall calculations

Domain mean cumulative snowfall ($\mu_{SWE^{+}}$) estimated from the relationship between snow pillow observations and historic snow patterns (Figure 4.3, option #1) closely matched $\mu_{SWE^{+}}$ estimated from the reanalysis. In fact, on 1 April, a reference date widely used to approximate Western U.S. peak-SWE timing (e.g., Cayan, 1996; Kapnick and Hall, 2012; McCabe and Dettinger, 2002; Mote, 2006; Mote et al., 2005), median $\mu_{SWE^{+}}$ estimated from snow pillow

observations was biased by less than 8% versus μ_{SWE^+} from the reanalysis, on average, with both positive and negative annual biases (Figure 4.9). The interquartile range of 1 April μ_{SWE^+} biases varied by less than $\pm$13% across the 32-year record, well within the errors expected from coarse-scale atmospheric models and gridded precipitation products (e.g., Henn et al., 2016; Hughes et al., 2017; Lannoy et al., 2010). The accuracy of μ_{SWE^+} was influenced less by differences in $Pattern^+$ between the historic dates with spatially correlated fSCA (SNO+R) and the dates with smallest MAE (SNO+Best), and more by uncertainties associated with inferring μ_{SWE^+} from the relationship between SWE^+_{pillow} and $Pattern^+$. These uncertainties likely included improper comparisons between SWE accumulation at points (snow pillows) and 90 m (reanalysis grid) scales, and historic snow accumulation patterns ($Pattern^+$) that were accurate on average (had small MAE), but were erroneous at the 90 m gridcells containing snow pillow observations. Because of these uncertainties, median μ_{SWE^+} estimated from multiple snow pillows tended to improve as the number of snow pillows increased. This was demonstrated in Figure 4.9 in the Tuolumne and Kings domains from periods prior to 1995, when in those periods, fewer snow pillow observations existed, and median μ_{SWE^+} estimated from snow pillow observations and historic snow accumulation patterns varied in accuracy by amounts greater than the accuracies between 1996 – 2016.

NLDAS2 μ_{SWE^+} was typically smaller than μ_{SWE^+} estimated from the reanalysis (e.g., Figure 4.10). This finding was consistent with findings from multiple studies, who have noted low-biases in mountainous NLDAS2 precipitation (Henn et al., 2018; Lannoy et al., 2010; Luo et al., 2003; Margulis et al., 2015; Pan et al., 2003). For the years shown in Figure 4.10, end-of-year μ_{SWE^+} from NLDAS2 agreed closer with the reanalysis than μ_{SWE^+} estimated from snow pillows only in low-snow water-year 2015. However, this was less-likely due to improvements in

NLDAS2, but instead due to errors in snow-pillow estimated μ_{SWE^+} driven by snow accumulation patterns in 2015 that were unlike any other year in the reanalysis (Section 4.7.1).

4.7.3. Snow depth simulations

The five snow deposition fields listed in Section 4.6.3 were used to simulate snow depth. For the Kings domain in water-year 2015, historic $Pattern^+$ selected from the dates of spatially correlated fSCA (criterion 2 from Section 4.6.1) and best MAE (criterion 3 from Section 4.6.1) were from the same year (2013), and only 4 days apart, resulting in identical 1) $Pattern^+$, and 2) snow pillow estimated domain mean cumulative snowfall (μ_{SWE^+}) (Figure 4.10). The same was true for the water-year 2016 simulation in Sagehen Creek, which identified $Pattern^+$ using the two criteria from only 3 days apart in 2004. In the Tuolumne domain, the historic $Pattern^+$ identified in each year (2014 – 2016) using the two criteria came from different years, resulting in μ_{SWE^+} estimated from snow pillows that differed from each other (Figure 4.10, bottom), albeit by less than 15%.

For simulations with μ_{SWE^+} from NLDAS2, snow deposition downscaled in space using snow accumulation patterns (NL2+R and NL2+Best) improved the simulated snow depth coefficient of correlation at peak snowpack timing by 0.13 (41%), on average, versus the simulations using the more common terrain-based downscaling (NL2+MM) (Table 4.1). This was improved even further by the SNO+R and SNO+Best simulations, which used snow pillow estimated μ_{SWE^+}, and were 0.22 better-correlated with ALS observed snow depth (70% better) than the NL2+MM simulations (Table 4.1, Figure 4.11). However, in water-year 2015, which had snow accumulation patterns that were unlike any other year, simulations using snow pillow estimated μ_{SWE^+} overestimated snow depth in the Upper Kings and Upper Tuolumne watersheds by +68% (0.11 m) and +73% (0.29 m) respectively.

A major source of error for the NL2+MM simulations was the assumption that snow input at spatial scales smaller than that of NLDAS2 (1/8° resolution) only increased and decreased with increases and decreases in elevation (Figure 4.12a). However, in the Tuolumne watershed in water-year 2016, elevations greater than 3100 m (~21% of the domain) had SWE that tended to decrease, on average, with increases in elevation (Figure 4.12c). This has also been noted by others in this region (e.g., Feld et al., 2013; Kirchner et al., 2014), and has been attributed to rainshadow effects from the nearby Clark's mountain range and the depletion of precipitable water at elevations below the Sierra Crest. The 2016 SNO+R simulation in the Tuolumne watershed which used $Pattern^{+}$ from 29 March 1989, implicitly represented decreases in Tuolumne snow deposition with increases in elevation above 3100 m (Figure 4.12b), mimicking the distribution of SWE estimated from the 1 April 2016 ALS observation significantly better. Importantly, the improvement in snow deposition for the SNO+R simulation was most-prominent at high-elevation regions above precipitation observations, where precipitation products often tend to decrease in accuracy (Wayand et al., 2013). The NL2+MM simulation also simulated snow that was far too spatially homogeneous (Figure 4.12a). Conversely, the SNO+R simulation (Figure 4.12b) simulated SWE with spatial heterogeneity that closer-matched SWE estimated from the ALS observation (Figure 4.12c). This SWE heterogeneity has been shown by many to influence the rate, timing, and duration of spring snowmelt (e.g., Egli et al., 2012; Freudiger et al., 2017; Luce et al., 1998; Lundquist and Dettinger, 2005), and therefore may influence water management decisions.

4.8. Discussion

Our results showed that satellite-era snow reanalyses could be combined with point snow measurements to estimate current-year snow magnitude and distribution with accuracies surpassing more common snow modeling approaches. In this section, we expand on the results,

testing the ability to identify historic dates with similar snow accumulation and depletion patterns based on changes to 1) the temporal length of the reanalysis data record, and 2) the spatial extents of domains (section 4.8.1). In Section 4.8.2, we discuss the differences between snow deposition processes in the three domains, and the changes to SWE accumulation patterns between winter and spring periods. Finally, in Section 4.8.3, we discuss how future research could address some of the opportunities and shortcomings demonstrated by this work.

4.8.1. Sensitivities to record length and domain size

The SWE reanalysis spanned years with significant differences in SWE volume. Here, we tested how data records over shorter periods influenced the ability to identify historically similar snow patterns in each domain. For each Landsat observation date (Figure 4.4, vertical gray lines), SWE from the reanalysis was extracted, at random, from 3, 7, and 15 other years. We then identified the dates of best-correlated fSCA from each SWE record (3, 7, and 15 years in duration), and calculated the MAE (Equation 4.3). These steps were repeated 10 times for each Landsat observation date to smooth spurious results. As expected, the median and spread of MAE decreased as the number of years in the data record increased (Figure 4.13, left column). In other words, as the length of the SWE data record increased, it became more likely to identify a period, from a different year, with similar snow accumulation and depletion patterns. This was particularly true in the Tuolumne and Kings domains for comparisons with moderate-to-poor MAE (CDF > 0.50), which most-commonly used fSCA from years with less-typical SWE volume to search the reanalysis for a period with similar snow patterns. However, even for 3-year SWE records, 80% of the comparisons had a MAE of less than 0.15. Therefore, interannual snow accumulation and depletion pattern repeatability existed, even between years with less-similar meteorological conditions. Therefore, snow accumulation patterns identified from shorter SWE reanalyses could likely downscale snow deposition in space better than simulations

downscaling precipitation employing popular terrain-based precipitation lapse rates (e.g., NL2+MM).

As opposed to the Tuolumne and Kings domains (both of which had over 130,000 90 m gridcells and spanned more than 2000 m of elevation), fSCA was a less-reliable standard to search the reanalysis for dates with similar snow patterns in Sagehen Creek (9,039 gridcells, ranging ~800 m in elevation). To better understand the influence of basin size, we selected a point within each domain, at random, and then extracted gridcells about that point to subsample domains with areas of 0.81 km^2 (10-by-10 gridcells), 8.1 km^2 (100-by-100 gridcells), and 81 km^2 (1000-by-1000 gridcells). Using these subsampled domains, the date of best-correlated fSCA (criterion 2 from Section 4.6.1) and best MAE (criterion 3) from a different year was identified for each Landsat observation date (Figure 4.4, gray lines). This test was repeated, using domains centered around 10 different randomly located points. Using the dates of best-correlated fSCA, the spread of MAE grew as domain size decreased (Figure 4.13, right). However, the best MAE improved as domain size decreased. In summary, as domain size decreased, $Pattern^+$ and $Pattern^-$ became more interannually repeatable, but it became less-likely to identify these dates using fSCA spatial correlation as a searching criterion.

Given the increased likelihood of identifying similar snow patterns from longer data-records (Figure 4.13, left), we would anticipate that snow patterns identified from the reanalysis would only improve with time as time progresses and the reanalysis record grows. However, even given the wide range of winter conditions over the 32-year data record, water-year 2015 exhibited snow accumulation patterns that were unlike any other year. Although some other years exhibited similar peak-SWE volume (Figure 4.4), water-year 2015 snowfall was exclusively at higher elevations and occurred much earlier in the season than was typical. Although water-year

2015 was unique, it is possible that SWE reanalysis estimates in this season could increase the likelihood of scaling snow deposition in future seasons with similar departures from normal conditions.

Rather than the subjective lidar extents used in this study, snow accumulation and depletion patterns could also be calculated across various spatial extents with the goal of identifying the regions that best-describe interannual differences in snow accumulation patterns as a function of meteorological characteristics like the sequence, frequency, timing, and orientation of snowstorms (e.g., Brandt et al., 2019; Pflug and Lundquist, 2020; Schirmer et al., 2011). Historic snow accumulation patterns could also be identified using fSCA observations from spatial extents larger than the domain across which snow patterns are calculated. The sensitivities to the spatial extents of snow patterns, and the spatial extents at which fSCA is used as a searching criteria, should be topics of future research.

4.8.2. Drivers of snow deposition

Both the Upper Tuolumne and Upper Kings watersheds (Figure 4.14, right column, blue lines) exhibited decreases in mean snow accumulation in upper-elevation gridcells. This was contrary to the Sagehen Creek domain, which was heavily forested, flatter by comparison, and had average winter snow deposition that increased with elevation. Therefore, the spatial distribution of snow deposition downscaled using terrain-based lapse rates (NL2+MM) performed better in Sagehen Creek with a peak-snowpack snow depth coefficient of correlation of 0.73, only 0.08 worse than simulations downscaling snow deposition using snow accumulation patterns from the reanalysis. At 90 m spatial resolution, snow deposition heterogeneity was less-likely influenced by processes like wind-redistribution and preferential deposition (e.g., Clark et al., 2011; Mott et al., 2010; Trujillo et al., 2007), and avalanches were too rare to greatly influence domain-scale snow distribution. Therefore, snow accumulation

patterns may have been driven predominantly by meteorologic effects, like orographic uplifting (Minder et al., 2008; Roe and Baker, 2006), barrier jets (Lundquist et al., 2010), rain shadow and flow-separation effects (Feld et al., 2013), and cloud microphysical processes (e.g., Hughes et al., 2017; Jankov et al., 2009; Morrison et al., 2009; Thompson et al., 2008). While some of these processes are represented by modern atmospheric models, the spatial resolution of atmospheric model setups are coarser (often > 1 km) than the 90 m simulations performed here. By downscaling snow deposition using historic snow patterns, we adjusted snow input as a proxy for these processes, increasing snow deposition heterogeneity in the process.

Downscaling snow deposition spatial heterogeneity using snow accumulation patterns assumed that all snowfall events in a given year deposited snow with an identical spatial pattern. However, snow deposition changes across individual snowfall events. This may be particularly true between winter and spring periods when the convective and synoptic drivers of precipitation vary more. In all three domains, 32-year average spring (April 1 through July 1) snow accumulation patterns were more spatially heterogeneous than winter (October 1 through April 1) snow accumulation patterns (Figure 4.14). Therefore, in later spring periods, historic snow accumulation patterns calculated over winter periods may 1) bias estimates of domain mean snowfall inferred from snow pillow observations, and 2) underestimate the spatial heterogeneity of snow deposition. Although snow accumulation patterns changed between winter and spring periods, the snow accumulation patterns were spatially correlated (coefficient of correlation of 0.91, 0.87, and 0.90 for Tuolumne, Kings, and Sagehen, respectively), suggesting that adjustments could be made to winter snow accumulation patterns to account for seasonal changes to the elevation of mean snow deposition, and the spatial heterogeneity of the snow accumulation.

4.8.3. Considerations and future paths of research

Very few regions have as many snow observations (both from ALS and snow pillows) than the Sierra Nevada domains investigated here. In Figure 4.9, the uncertainty of domain mean cumulative snowfall (μ_{SWE^+}) worsened in years prior to 1995 with fewer snow pillow observations. Therefore, other global domains with comparable or fewer snow pillow observations may be less successful at inferring μ_{SWE^+}. Additionally, the California Sierra Nevada receive most moisture from Pacific storms with trajectories from the southwest, whereas many mountainous regions have more varied storm patterns. As a result, snow accumulation patterns in the California Sierra Nevada may be more repeatable than those in other mountainous regions. Future research should investigate the methods used here in less-densely gauged watersheds with more interannual storm variability.

Here, we were constrained to define snow patterns at the 90 m spatial resolution of the snow reanalysis. The spatial resolution of Landsat fSCA observations used by the reanalysis make it unlikely that the snow reanalysis could be performed at spatial scales ($\leq$ 10 m) corresponding to the snow processes that are among the most-difficult to model today (e.g., wind-redistribution, preferential deposition, etc.). However, recent constellations of small-size satellites (CubeSats) have demonstrated the ability to observe snow covered area at ~3 m spatial resolution at nearly-daily temporal resolutions in mountainous terrain (Cannistra et al., 2021). While these observations have only become available within the past decade, and therefore have a limited data record, snow cover information observed by CubeSats may provide the opportunity to estimate SWE evolution at snow drift spatial resolutions.

4.9. Conclusions

Mean snowfall and the spatial distribution of snow accumulation are commonly misrepresented in complex terrain. We investigated how historical snow accumulation patterns

from a Sierra Nevada multidecadal (1985 – 2016) reanalysis could be identified in real-time scenarios, and how those snow accumulation patterns could be combined with real-time ground measurements to infer mean snowfall magnitude and snow deposition spatial heterogeneity. The dissimilarity between superimposed snow accumulation and depletion patterns, and the resulting distribution of fractional snow-covered area (fSCA) between peak-SWE timing and mid spring snowmelt, could be used to search the reanalysis for a date, from a different year, with similar relative amounts of snow accumulation and depletion. This was particularly true for larger domains with spatially variable elevation, where the spatial pattern of fSCA was more indicative of annual differences in snow accumulation and depletion. For the Sierra Nevada domains tested here, normalized snow accumulation and depletion patterns between Landsat observation dates and historic dates with spatially correlated fSCA, agreed by approximately 9 – 11% (on average). In all cases, historic patterns queried from longer portions of the 32-year reanalysis were more accurate than patterns queried from random subsets of shorter duration. This was especially true for seasons with less-typical snowfall magnitude. However, water-year 2015 was the driest snow year ever recorded or estimated in the California Sierra Nevada (Belmecheri et al., 2016; Margulis et al., 2016b), resulting in snow accumulation patterns that were unlike any other year.

April 1 domain mean cumulative snowfall between 1985 – 2016, inferred from the relationship between snow pillow-observed cumulative increases in SWE and historic snow accumulation patterns, was biased by $\pm$13% (32-year interquartile range). This spread of biases was better than biases identified for atmospheric models and gridded precipitation products in this region (e.g., Henn et al., 2016; Hughes et al., 2017). Additionally, there was little difference in domain mean snowfall estimated using snow pillow observations and historic snow

accumulation patterns from the best-matching historic dates, and the historic dates with spatially correlated fSCA, the latter of which could be identified using real-time satellite observations.

The snow depth coefficient of correlation at peak-SWE timing was 0.22 better (70% better), on average, for simulations using historic snow accumulation patterns to both 1) infer domain-mean snowfall from snow pillows, and 2) downscale 90 m snow deposition, than simulations using precipitation from NLDAS2 and a common terrain-based snow deposition downscaling routine. Snow deposition downscaled using historic snow accumulation patterns captured a decrease in mean snowfall at the upper-elevations of the Tuolumne and Kings watersheds which was not represented by simulations downscaling snow deposition using terrain-based precipitation lapse rates. The results here suggest that water operations in California may benefit from merging multidecadal reanalyses with existing or expanded snow pillow networks in the Sierra Nevada.

4.9. Appendix C: Snow reanalysis accuracy

The spatial variability of snow depth is driven more by the spatial variability of SWE than snow density. Therefore, assuming increases and decreases in snow depth should correspond to similar relative increases and decreases in SWE, we can compare the accuracy of the SWE reanalysis versus ALS snow depth observations between 2013 and 2016. ALS snow depth observations were compared to the reanalysis SWE on coincident dates using the spearman-rank coefficient of correlation (Figure 4.15) across the full Upper Tuolumne extents. Results found strong spatial correlations (r_s > 0.80) between the datasets for periods with snow-covered fractions (SCF) between 0.40 and 0.90. In fact, reanalysis accuracy seemed to peak at periods of approximately 0.60 SCF. This was expected since the reanalysis calculates SWE using satellite-observations of fractional snow-covered area (fSCA) and is therefore provided the most information on dates when larger portions of the domain have intermediate fSCA (between 0 and

1). In this study, we used fSCA from the reanalysis, sampled at Landsat observation periods between peak-SWE timing and mid-spring snowmelt (corresponding well with SCF between 0.90 and 0.40), to identify historic dates with similar snow patterns (Section 4.6.1). It is important to note that although the spearman-rank coefficient of correlation (r_s) between ALS snow depth and the reanalysis SWE decreases below SCF of 0.40 (Figure 4.15), this is likely driven in part by spatial variability in snow density, which is elevated in late-spring periods between portions of the domain with larger differences in cumulative melt energy. Across the 4 years of overlapping ALS data, regions where the reanalysis SWE differed from the ALS snow depth observations sometimes differed in a consistent fashion from season to season (not pictured). While this could be due to underlying snow density patterns, interannually-persistent errors in snow reconstructions and snow reanalyses may be able to be corrected with one, or few, ALS observations

Although the SWE reanalysis used in this study has demonstrated good performance , accuracy has primarily been investigated versus snow pillow observations preferentially located in regions with no, or reduced, forest cover. Given how forest canopy can obstruct Landsat observations of snow on the land surface, we expected less-accurate SWE reanalyses in forested regions. To test this, we compared fSCA from the reanalysis versus fSCA derived using high-resolution lidar point clouds in the Sagehen Creek domain, 93% of which was covered by forests. Results found that fSCA calculated from the reanalysis matched fSCA calculated from lidar well (Figure 4.16), with no perceivable difference in performance between forested and non-forested pixels.

4.10. Appendix D: Snow pillow multilinear regression

The fine-scale spatial heterogeneity of snow deposition made SWE accumulation from snow pillows (SWE^{+}_{pillow}) difficult to compare with 90 m SWE accumulation from the reanalysis

(SWE^+). However, given that finer-scale snow accumulation can exhibit interannually-repeatable characteristics, snow accumulation at a point was likely to be related the snow accumulation at larger spatial scales encompassing that point. We tested this by fitting a multilinear relationship, using SWE^+_{pillow} and other snow pillow explanatory variables, to reproduce SWE^+ (response variable) between 1985 – 2016. Tested explanatory variables included seasonal SWE information that could be provided in real-time from snow pillows (number of snowfall events, number of snowmelt events, average/maximum size snowfall/snowmelt events, days since snowpack initiation, and the ratio between cumulative snowfall and snowmelt), and terrain-differences (elevation, exposure, northness, and slope) between the 30 m SRTM pixels and 90 m reanalysis pixels containing each snow pillow.

Only the number of snowfall events exhibited significant explanatory power. In other words, the disagreement between SWE^+ and SWE^+_{pillow} (Figure 4.5, left) was well-explained by a multiplier, specific to each snow pillow, that changed with the number of snowfall events (Figure 4.5, right, r = 0.94, mean absolute error of 104 mm, and median bias of 1.93%). Although differences between the terrain at snow pillows and averaged across the 90 m reanalysis pixel likely influenced the agreement between of SWE^+ and SWE^+_{pillow}, the multilinear regression including terrain explanatory variables (elevation, exposure, northness, and slope) improved the multilinear regression by amounts much smaller than the accuracy expected from the reanalysis. The sensitivity of the multilinear regression to the number of snowfall events suggested 1) that the number of snowfall events physically influenced snow deposition at point-scales and 90 m scales differently, and/or 2) differences between SWE^+ and SWE^+_{pillow} were different between years with deep and shallow snow, and/or 3) noise in SWE observations from snow pillow, or increases in snow from processes like wind-redistribution, could be registered as snowfall events,

when in reality, no snowfall actually occurred. Here, we used the snow accumulation from the reanalysis as the response variable, and thus attributed the disagreements only to the snow pillows. However, we also acknowledge that the disagreement between SWE^{+} and SWE^{+}_{pillow} were likely driven in part by errors in the reanalysis. However, snow accumulation transformed using the multilinear regression were similar in accuracy to errors in the reanalysis SWE noted by Margulis et al. (2016a).

In Section 4.6.2, we derived domain mean snowfall in each year, using the multilinear regression trained in all other years. For example, given snow pillow observations from water-year 2000, the multilinear regression was trained using the relationship between SWE^{+} (response variable), and SWE^{+}_{pillow} and the number of snowfall events (explanatory variables), at each snow pillow from 1985 – 1999 and 2001 – 2016. We also tested the multilinear relationship trained using only the historic years with the most-similar snow patterns (Section 4.6.1, criterion 3). However, SWE^{+} predicted by the multilinear regression was much better when trained using all 30 other years, as opposed to trained in only the year(s) with most-similar snowpack.

Table 4.1. Snow depth accuracy for each simulation versus the ALS observation closest peak-snowpack timing in each domain and year. The best performing simulation for each statistic is highlighted.

		Statistic			
Domain, year	**Simulation**	**Coeff. of correlation [r]**	**Mean abs. error [m]**	**Bias [%]**	**Snow extent error [%]**
Tuolumne, WY2014	**NL2 + MM**	0.57	0.53	-49%	+2%
	NL2 + R	0.66	0.60	-63%	-4%
	NL2 + Best	0.67	0.57	-60%	-6%
	SNO + R	0.81	0.25	-12%	+4%
	SNO + Best	0.81	0.23	-3%	+3%
Tuolumne, WY2015	**NL2 + MM**	0.37	0.30	-43%	-39%
	NL2 + R	0.46	0.26	-34%	-40%
	NL2 + Best	0.47	0.25	-23%	-39%
	SNO + R	0.65	0.38	+73%	+12%
	SNO + Best	0.64	0.39	+73%	+6%
Tuolumne, WY2016	**NL2 + MM**	0.50	0.78	-36%	-6%
	NL2 + R	0.74	0.69	-34%	-5%
	NL2 + Best	0.75	0.68	-33%	-5%
	SNO + R	0.77	0.53	+20%	+0%
	SNO + Best	0.77	0.45	+9%	-1%
Kings, WY2015	**NL2 + MM**	0.08	0.30	+5%	-31%
	NL2 + R	0.22	0.27	+16%	-46%
	NL2 + Best	0.22	0.27	+16%	-46%
	SNO + R	0.30	0.32	+68%	+43%
	SNO + Best	0.30	0.32	+68%	+43%
Sagehen, WY2016	**NL2 + MM**	0.73	0.48	-37%	-8%
	NL2 + R	0.81	0.46	-39%	-7%
	NL2 + Best	0.81	0.46	-39%	-7%
	SNO + R	0.81	0.31	+10%	-2%
	SNO + Best	0.81	0.31	+10%	-2%

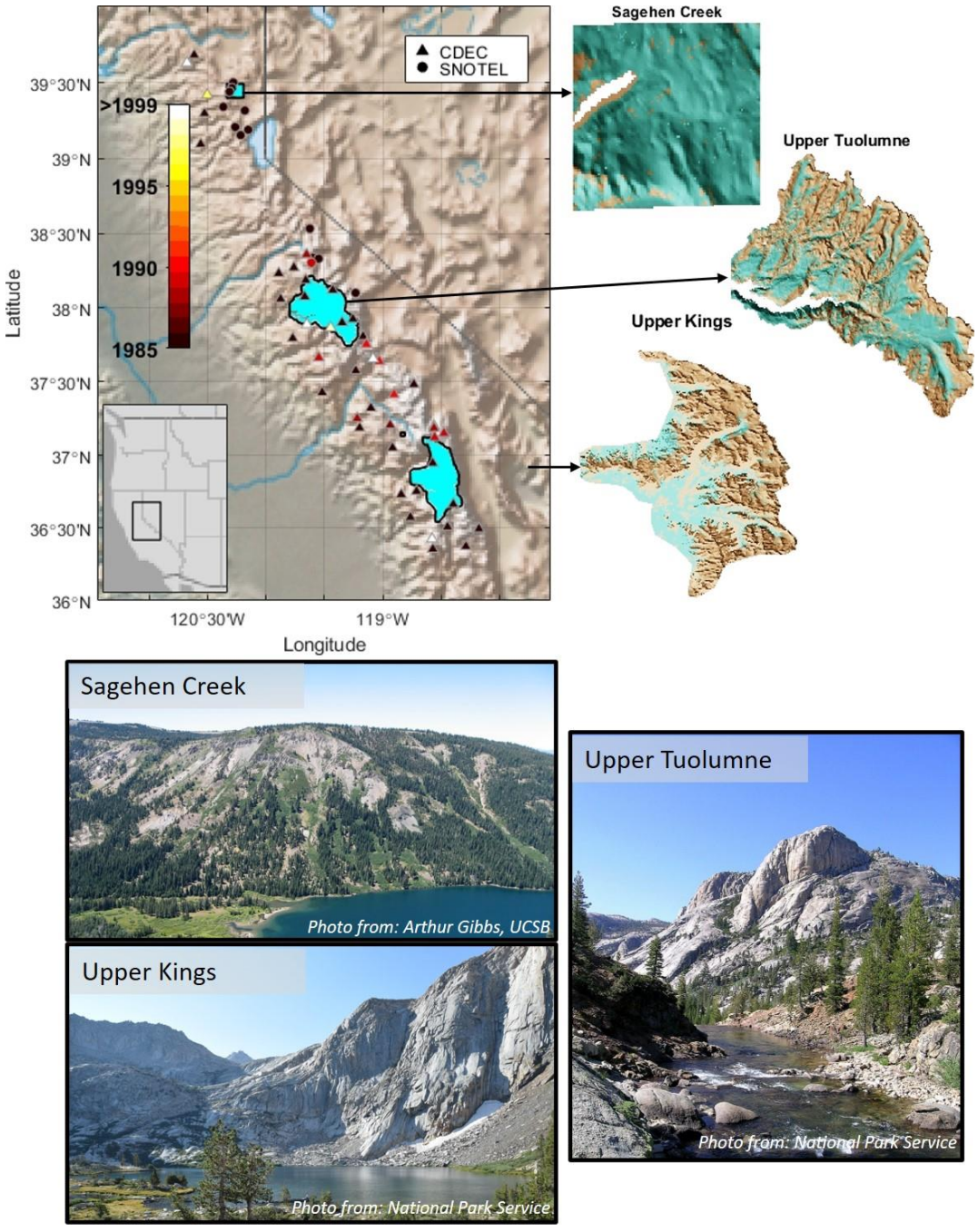

Figure 4.1. Sierra Nevada study domains. Green cells (top, right) represent 90 m pixels with ≥ 50% forest cover. The spatial location of these domains and 58 snow pillows (markers) are shown in the leftmost plot. Snow pillows are colored by the year in which SWE data first became available (colorbar).

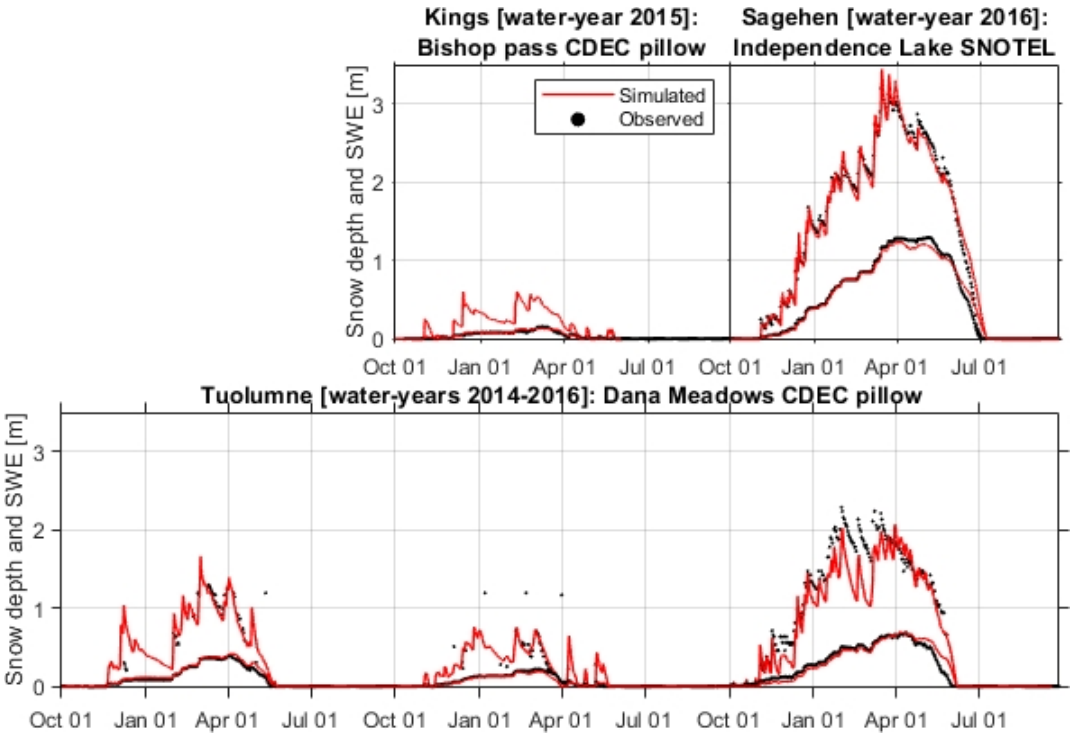

Figure 4.2. Hourly simulated snow depth and SWE (red line) calibrated versus daily observations (black dots) at snow pillows located within each of the three domains (subplots). For calibration purposes, simulations were forced with snowfall prescribed from the daily increases in SWE observed at each snow pillow.

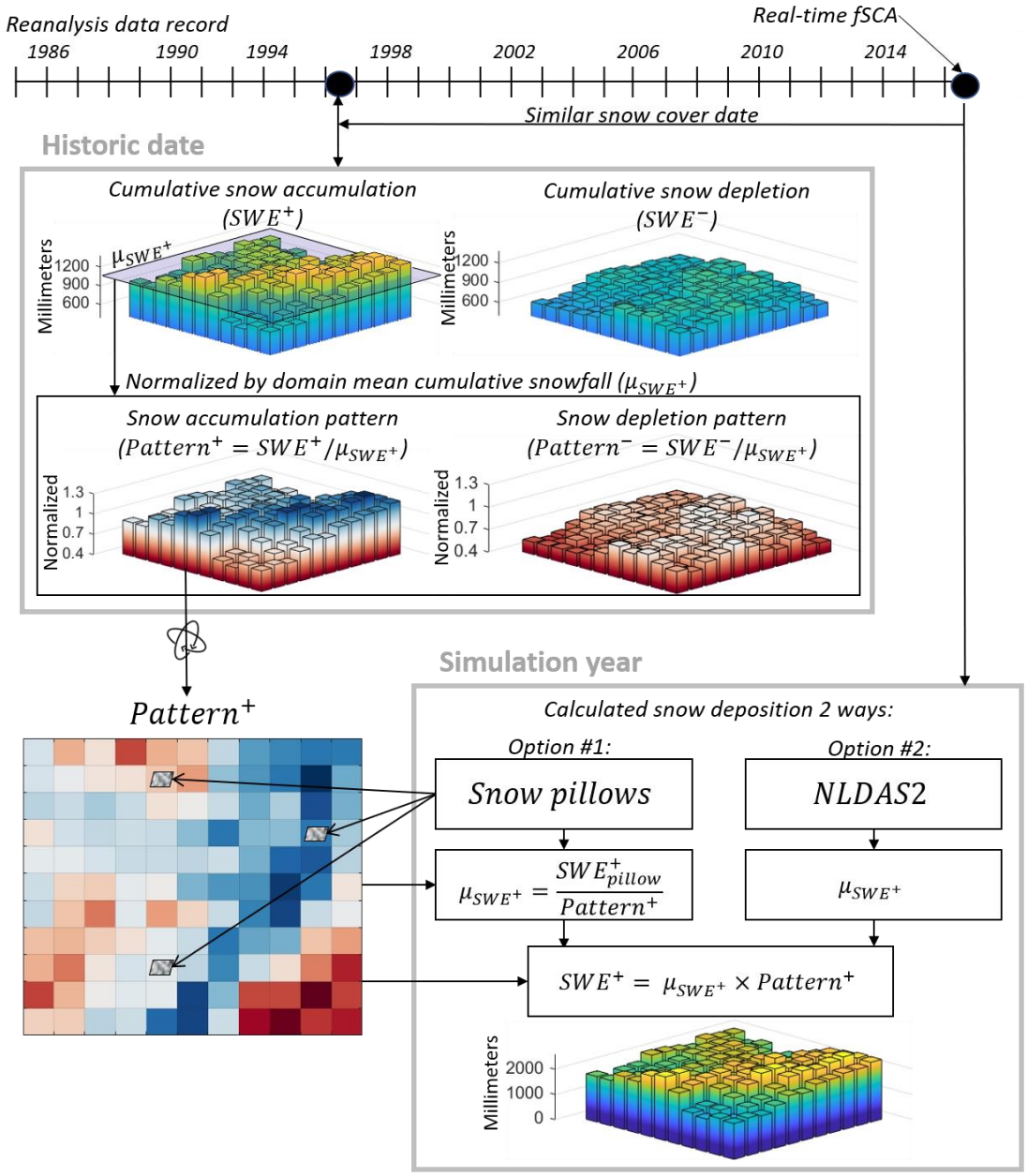

Figure 4.3. A historic date with similar snow cover was selected from the reanalysis data record (top). On this date, cumulative increases (SWE^+) and decreases (SWE^-) in SWE were calculated. Both were normalized by the domain mean cumulative snowfall (μ_{SWE^+}, spatial average of SWE^+, shaded plane), to generate snow accumulation ($Pattern^+$) and depletion ($Pattern^-$) patterns. In the simulation year, μ_{SWE^+} was calculated from 1) the ratio between cumulative increases in SWE observed by snow pillows (SWE^+_{pillow}) and the $Pattern^+$ at gridcells containing those snow pillows (option #1), and 2) NLDAS2 (option #2). Finally, SWE^+ was calculated in the simulation year from the product of μ_{SWE^+} and $Pattern^+$.

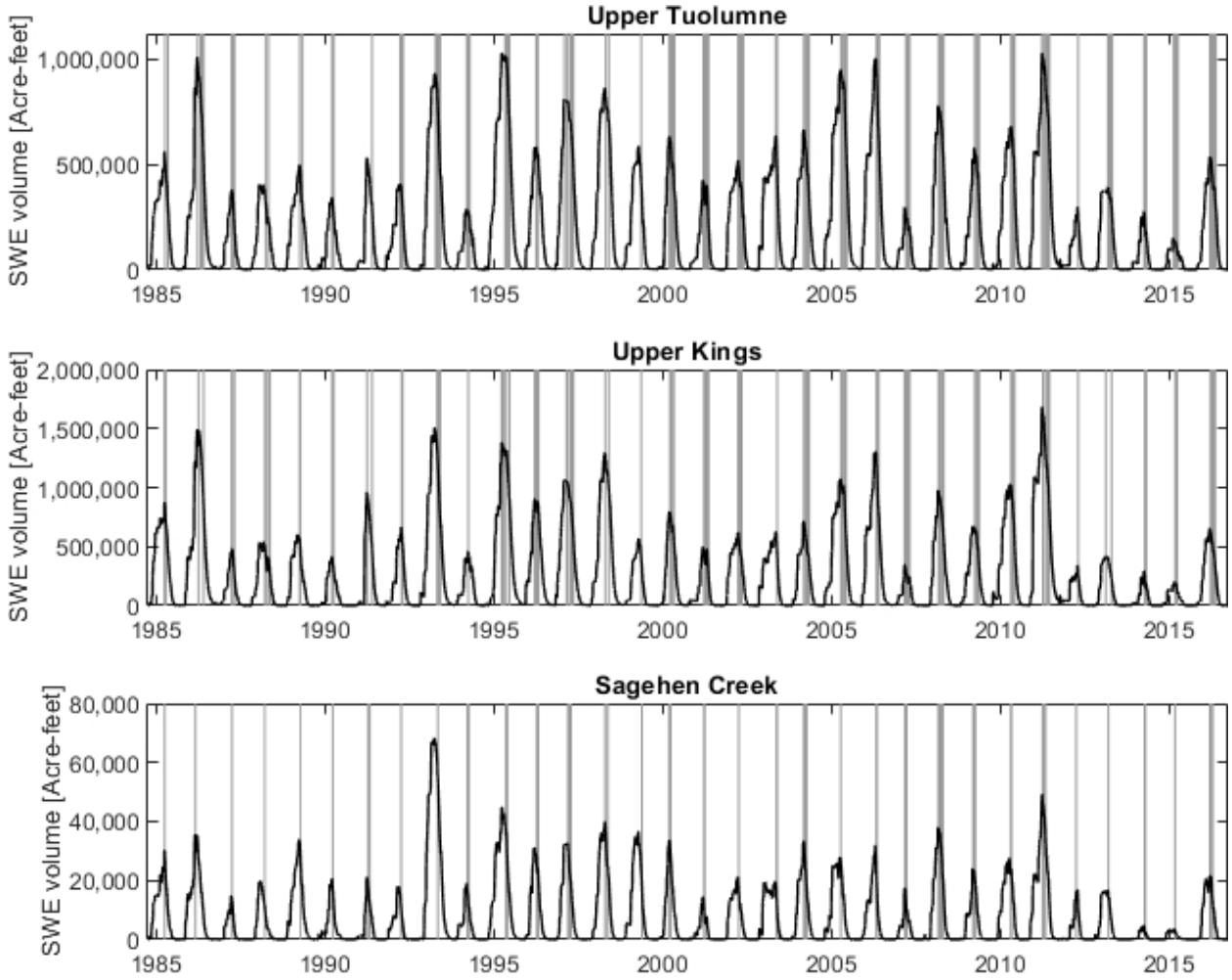

Figure 4.4. SWE volume calculated from the reanalysis. Vertical gray lines represent cloud-free Landsat observation dates, limited to dates between peak-SWE timing and midway through spring snowmelt. Note that y-axis limits change for each domain (subplot).

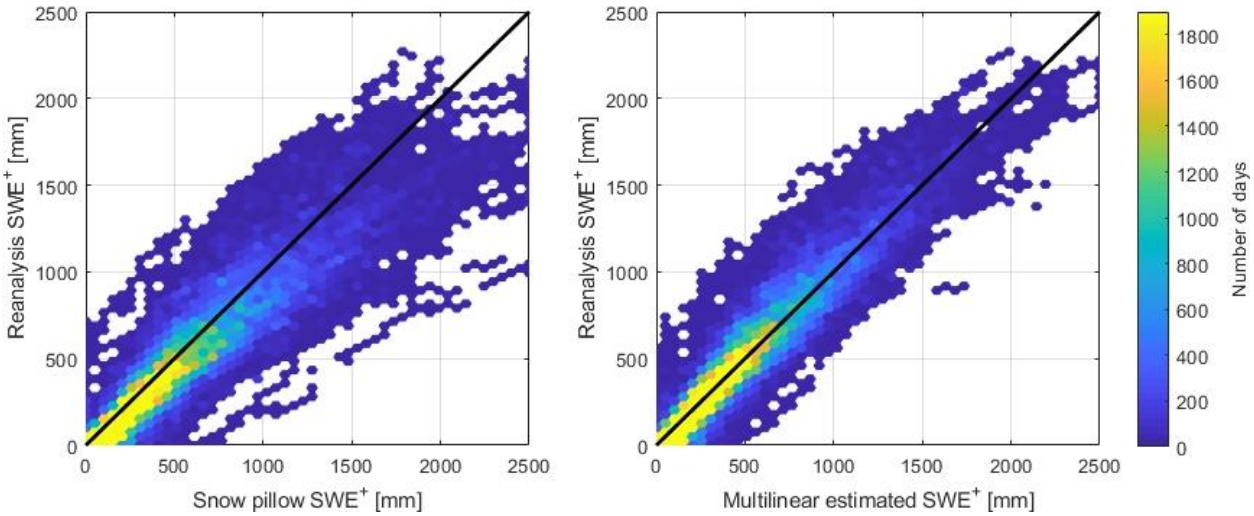

Figure 4.5. Density scatter plots showing cumulative increases in SWE (SWE^{+}) between 1 October and 1 May in water-years 1985 – 2016 from the reanalysis (y-axes), plotted versus observations at snow pillows (left), and snow pillow observations rectified using a multilinear regression (right).

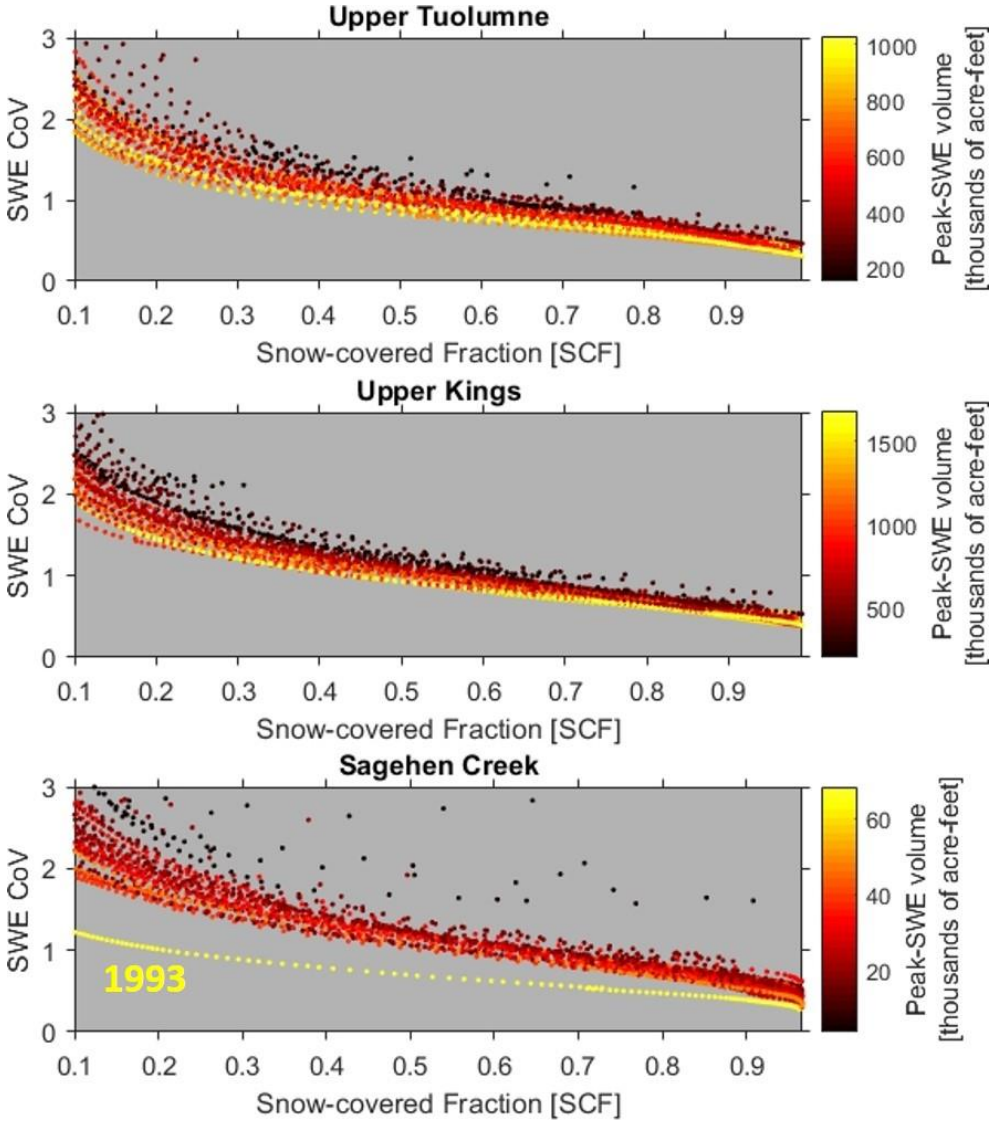

Figure 4.6. The relationship between snow covered fraction (x-axis) and the SWE coefficient of variation (CoV, y-axis) throughout spring snowmelt in 32 different years with different peak-snowpack SWE volumes (colorbar) in each domain (subplots).

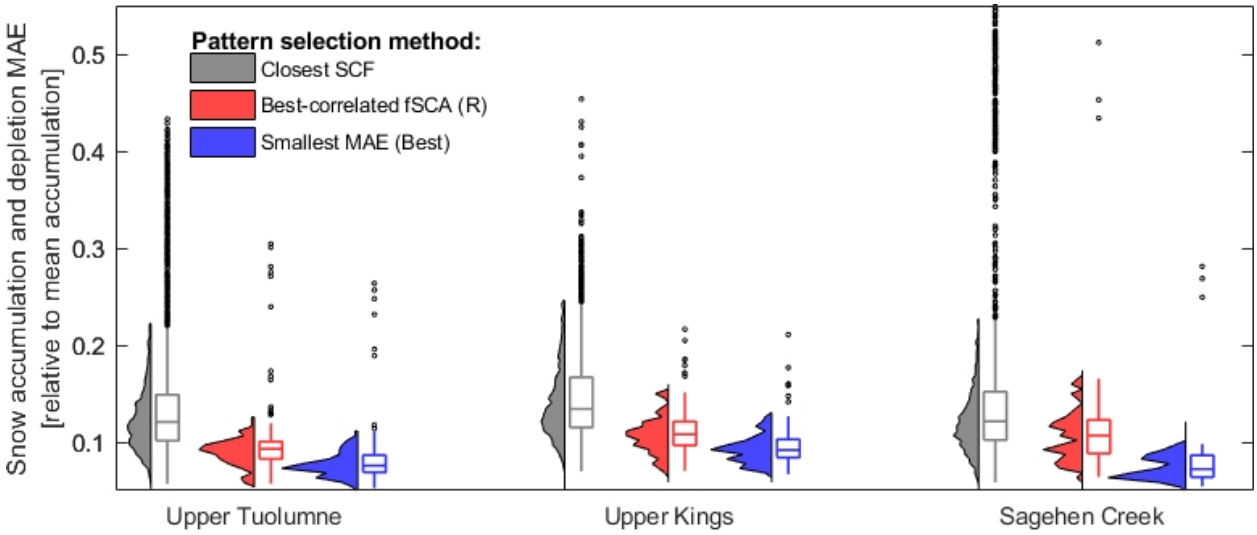

Figure 4.7. MAE (Equation 3) calculated between Landsat observation dates and dates, in different years, with matching SCF (gray), spatially-correlated fSCA (red), and the dates of smallest MAE (blue).

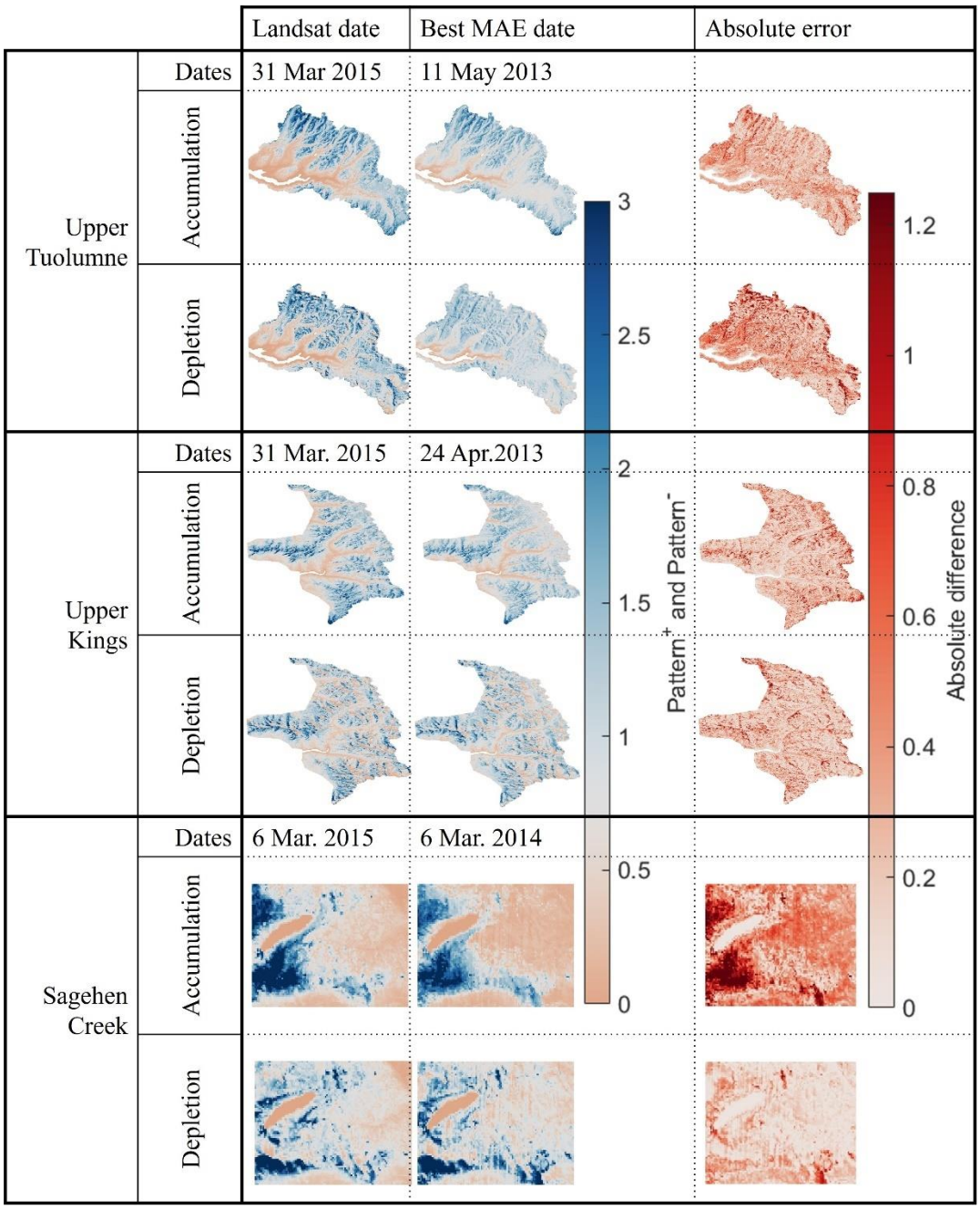

Figure 4.8. Snow accumulation and depletion patterns compared between water-year 2015 Landsat observation dates and the dates with the best-matching patterns in each domain.

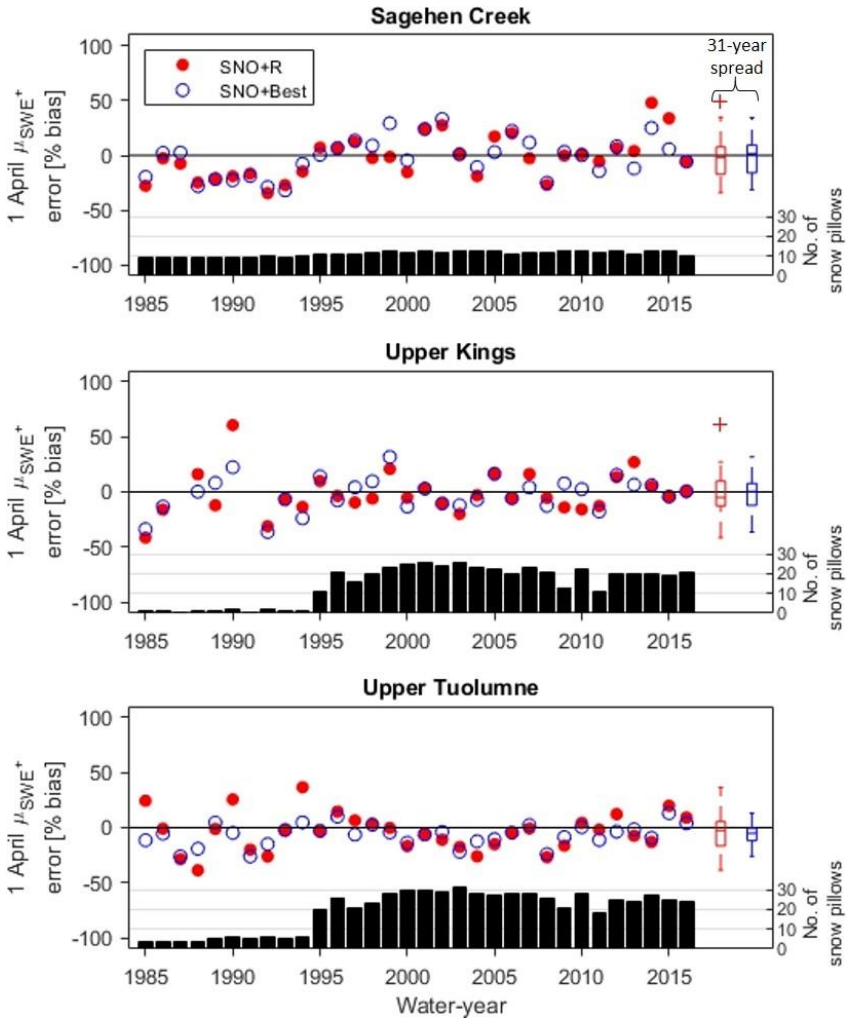

Figure 4.9. Biases in 1 April μ_{SWE^+} inferred from the relationship between snow pillow observations in each year and $Pattern^+$ from 1) dates with spatially-correlated fSCA (red), and 2) the dates of smallest MAE (blue). The number of snow pillow observations in each year are shown by the bar plots (corresponding to the right y-axis). The 32-year spread of 1 April μ_{SWE^+} biases are shown by the rightmost boxplots.

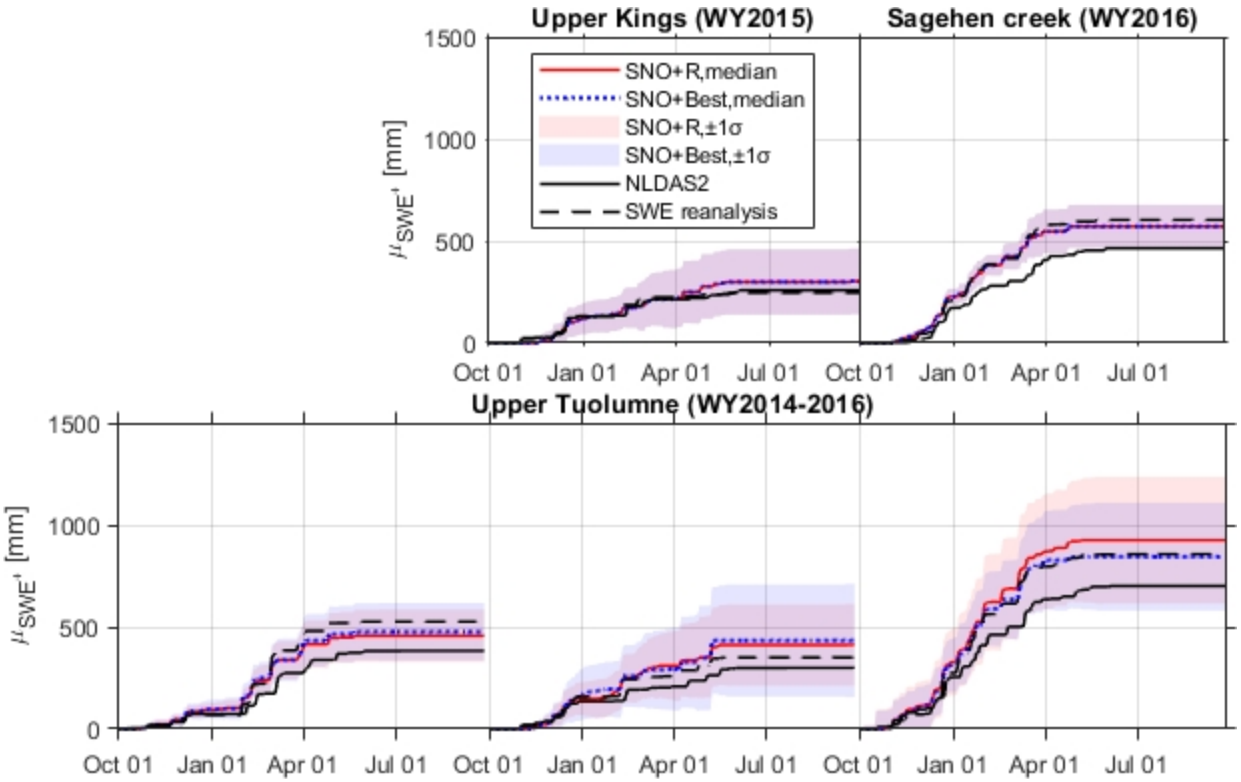

Figure 4.10. Domain mean cumulative snowfall (μ_{SWE+}, y-axis) in the years and domains with ALS observations. In addition to the median, the ± 1 standard deviation in μ_{SWE+} calculated from snow pillows is shown by the shaded regions.

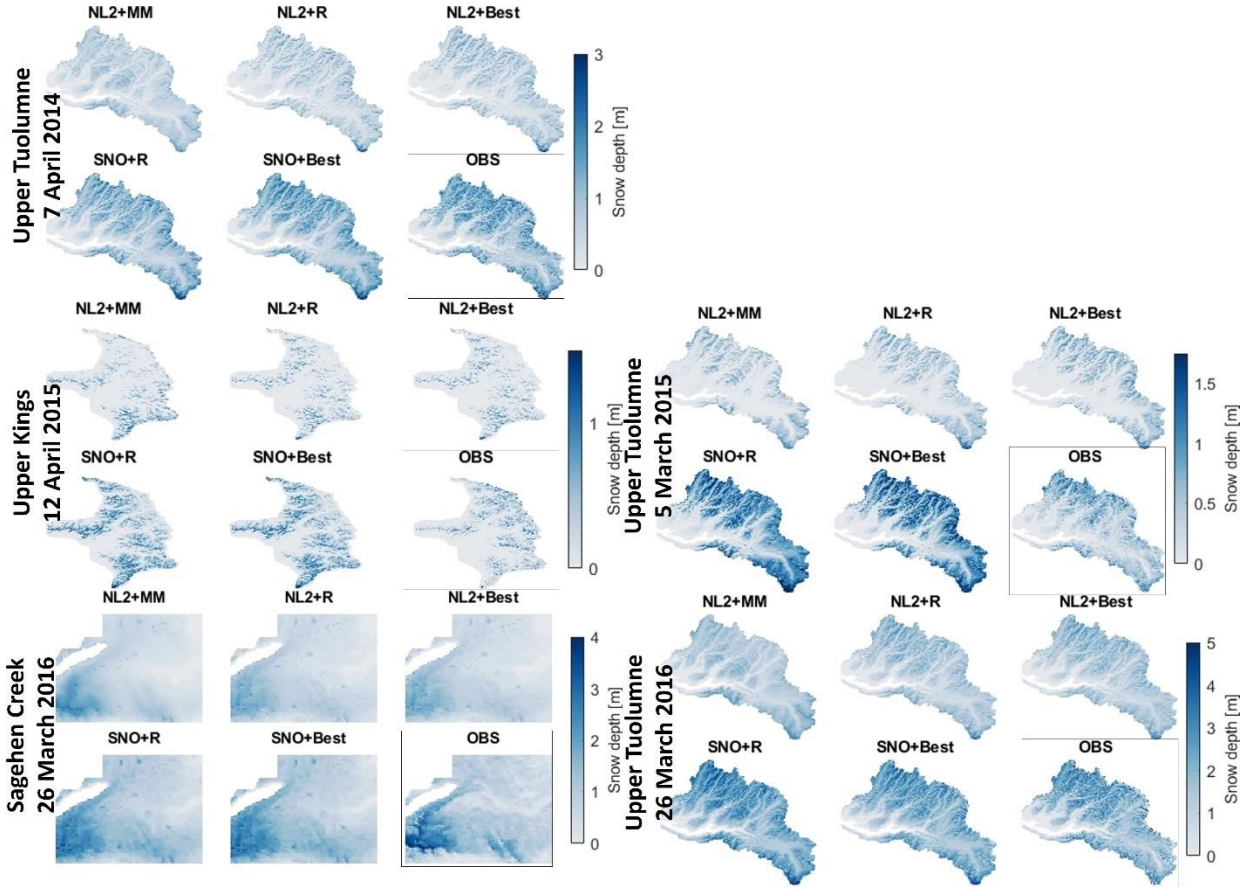

Figure 4.11. Snow depth modeled and observed (OBS; boxed) near peak-snowpack timing in each domain and year. To easily visualize snow depth, the colorbar limits were adjusted for each set of simulations.

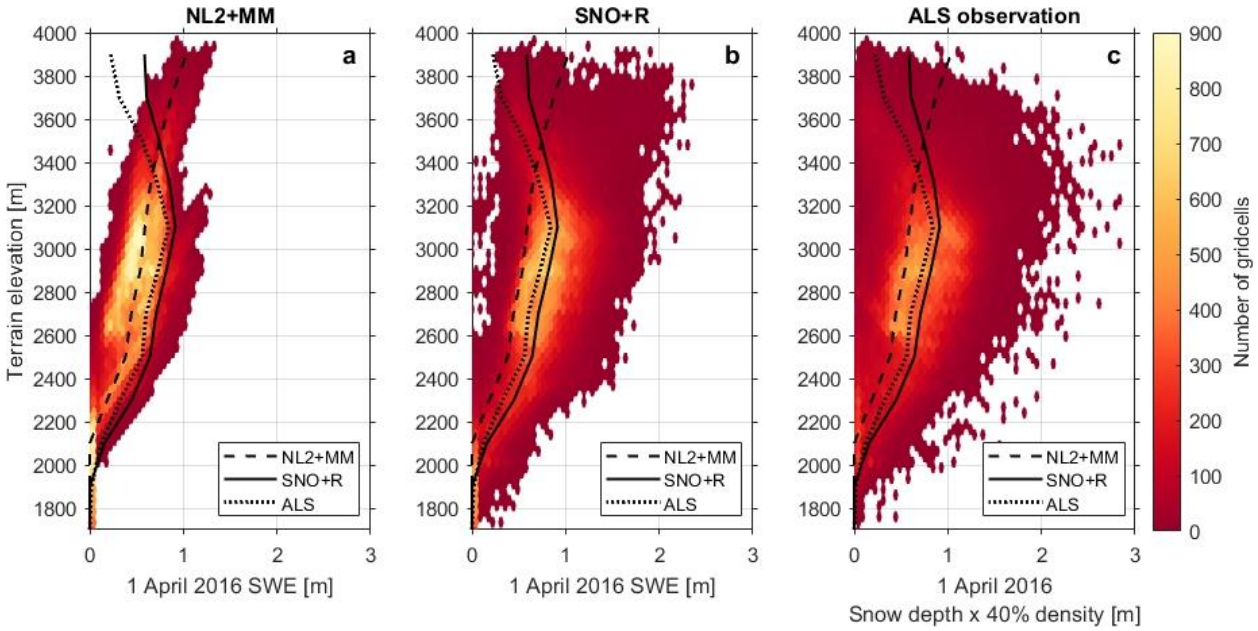

Figure 4.12. Density scatter plot showing the relationship between Tuolumne elevation and SWE on 1 April 2016 for the NL2+MM (a) and SNO+R (b) simulations. SWE on 1 April 2016 was also estimated using ALS snow depth multiplied by a spatially constant snow density averaged from surrounding snow pillows (c). The median SWE values across increasing elevation bands were plotted for comparison in each subplot.

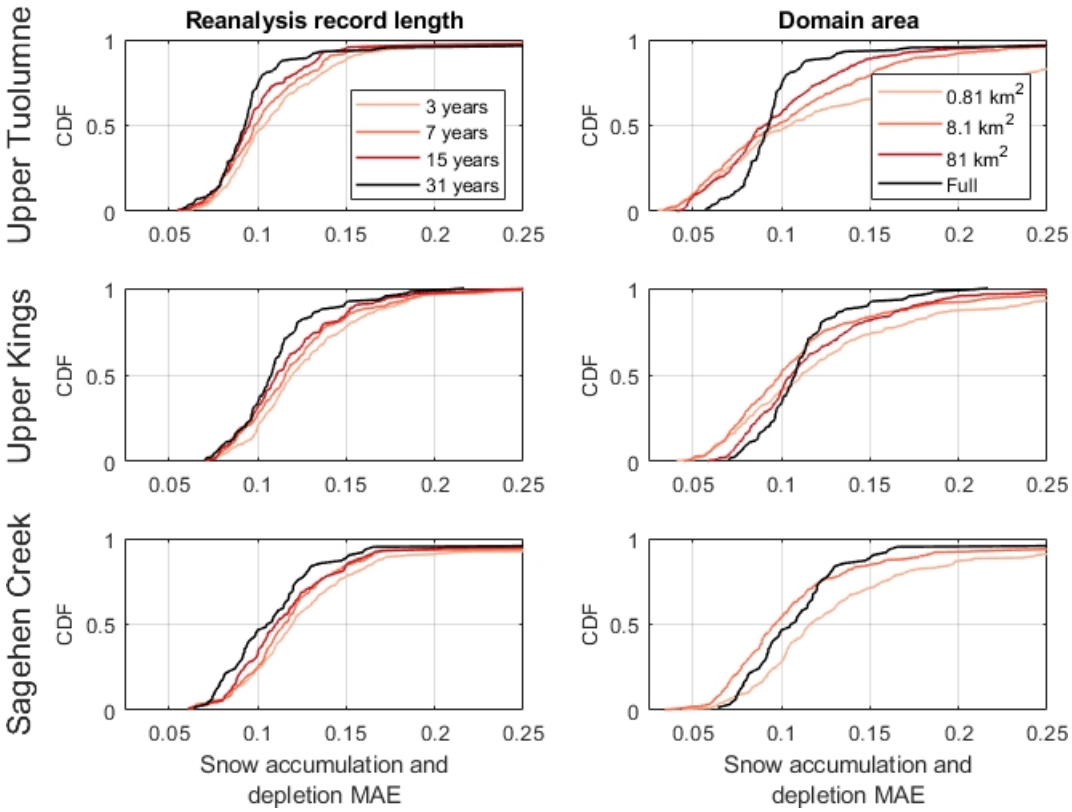

Figure 4.13. The cumulative distribution function (CDF) of MAE between Landsat observation dates and the dates of best-correlated fSCA. The distributions are for data records of different lengths (left) and domains with different areas (right) in all three domains (rows). The 81 km^2 domain was excluded for Sagehen Creek since it was larger than the full domain (bottom right).

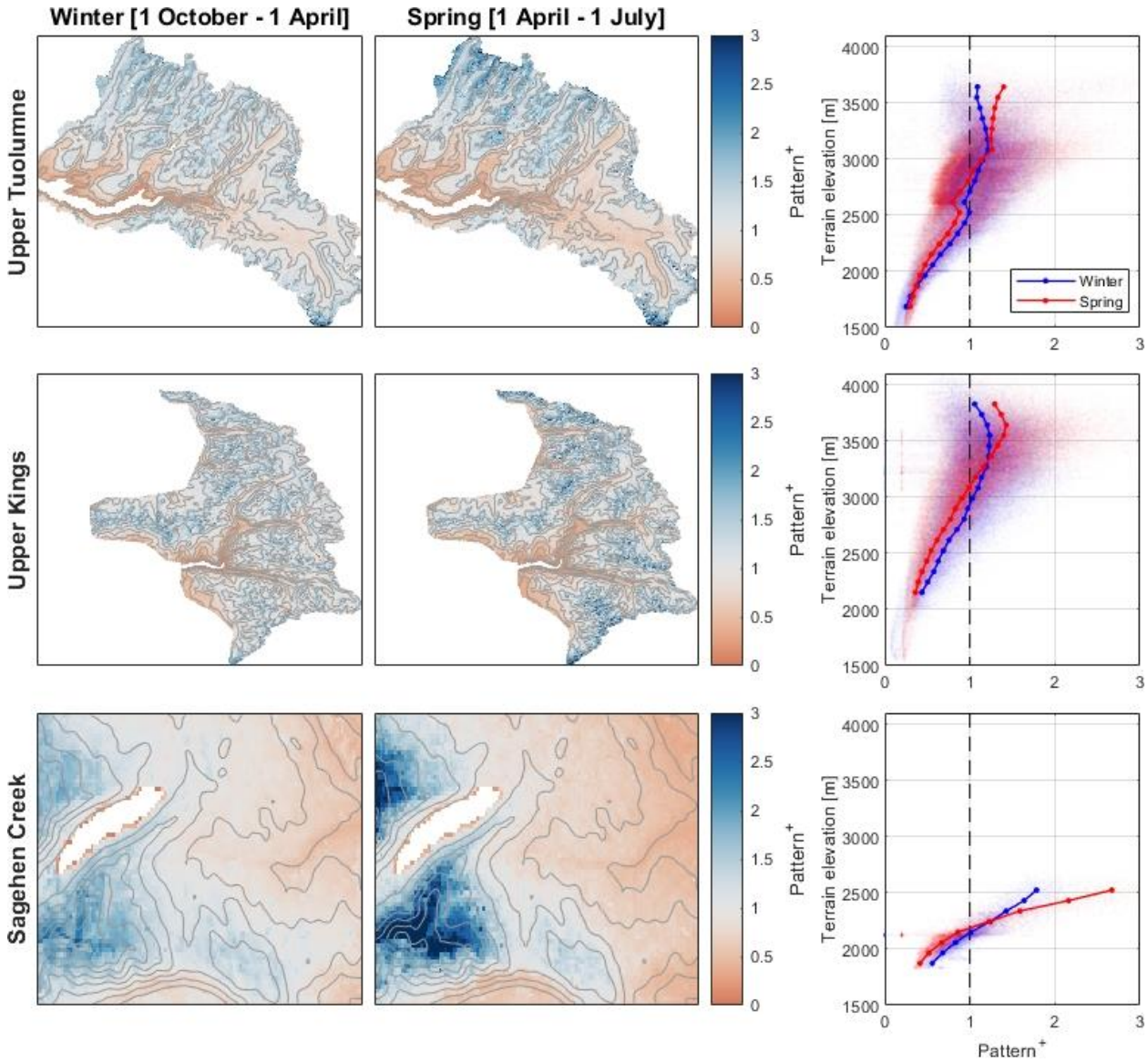

Figure 4.14. 32-year average winter (left) and spring (center) snow accumulation patterns over each domain (rows). The median relationship between elevation and winter (blue) and spring (red) snow accumulation patterns are compared in the rightmost column, with the vertical dashed line representing mean snowfall ($Pattern^{+}$ = 1.0).

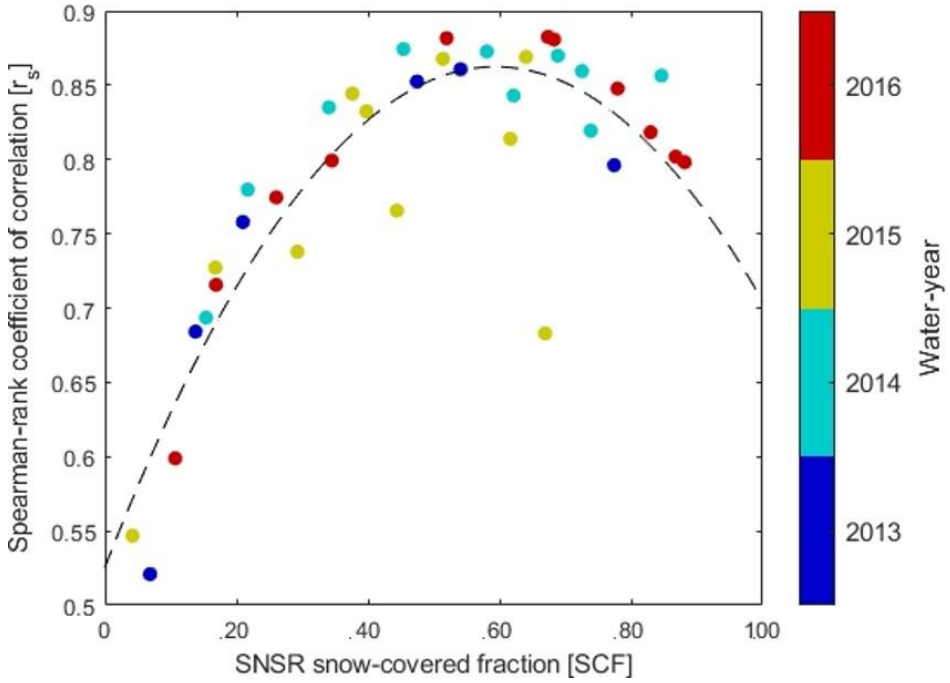

Figure 4.15. Spearman rank coefficient of correlation (y-axis) between airborne lidar snow depth and SWE from the reanalysis in the four years with overlapping data (colorbar) in the Tuolumne watershed. The relationship between the coefficient of correlation and snow-covered fraction (x-axis) is approximated using a polynomial function (dashed).

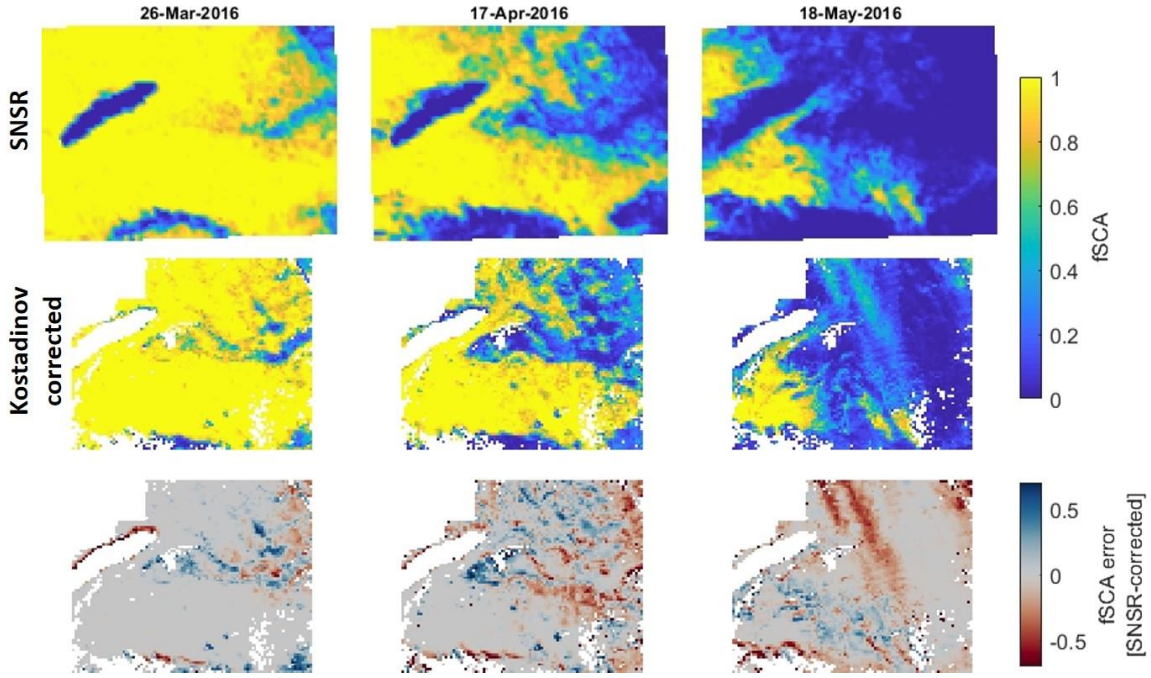

Figure 4.16. Sierra Nevada Snow Reanalysis (SNSR) fSCA (top) and fSCA derived using lidar observations by Kostadinov et al. (2019) (middle) in the Sagehen Creek watershed. The difference between the two estimates is shown between the two products (bottom row) on each date (columns).

Chapter 5 : Conclusions

Snow patterns from historic periods can be used to improve estimates of total snow volume and snow distribution in mountainous domains. Chapter 2 showed that distributed snow depth could be inferred (mean absolute error of 0.22 m, with snow volume biased by 6%) by leveraging the relationship between a historically representative snow depth pattern and an observation across only a portion (< 4%) of a watershed. Since current land-surface and snow models are more challenged by representations of snow accumulation than snow depletion, pattern-based snow depth spatial extrapolation could be used to determine peak-snowpack snow depth magnitude and distribution. When paired with a snowmelt model, pattern-based snow depth spatial extrapolation could improve simulated snowmelt (Figure 2.9) and the resulting streamflow. Provided a sufficient set of historic snow depth patterns from which to choose from, the pattern-based snow depth spatial extrapolation used in Chapter 2 may provide the ability to economize airborne lidar scanning (ALS) surveys. For instance, in areas with repeatable snow depth patterns, distributed snow depth could be 1) inferred in a single watershed using less-costly reduced ALS coverage, or 2) inferred across larger regions by collecting ALS observations with partial coverage over multiple watersheds.

Chapter 2 also demonstrated that snow depth maps inferred from the relationship between real-time snow depth observations and historic snow depth patterns improved as the number and variability of real-time snow depth observations increased (Figure 2.4). However, snow depth observations to the extents required for reliable snow depth spatial extrapolation may be challenged by inaccessibility and avalanche hazards in snowy mountainous terrain. Additionally, even reduced ALS observations may be economically expensive, owing to the logistical costs of these surveys (jet fuel, airplane hangar fees, instrument calibration, data processing, etc.). Therefore, in Chapter 3, we downscaled 25 m snow deposition from 6 km mean snowfall

provided by an atmospheric model snowfall using historic snow depth patterns. Pattern-based snow deposition downscaling was used as forcing for simulated snow depth ensembles, and were compared versus simulations using a more-common terrain-based snow deposition downscaling approach. In the Tuolumne River watershed, using a set of peak-snowpack ALS observations between 2013 and 2019, the impact of interannual snow depth pattern-repeatability on snow deposition downscaling was relatively small as compared to 1) atmospheric model snowfall biases, and 2) winter snow depletion (snowmelt and snow sublimation) and snow density influences on snow depth patterns (Figure 3.10).

In Chapter 4, we addressed the issues with the snow depth pattern-based snow deposition downscaling routine discussed above. Specifically, we downscaled snow deposition using SWE accumulation patterns, which neglected winter snow depletion and snow density variability, from a 31-year daily SWE product reanalyzed using a land surface model and satellite observations of snow cover depletion. We showed that the relationship between scattered snow-pillow observations of SWE accumulation and historic snow patterns identified from dates of spatially correlated fractional snow covered area (fSCA), were able to derive watershed-scale estimates of mean snowfall with accuracies ($\pm$13% biases in 1 April cumulative snowfall) surpassing more common atmospheric models and gridded precipitation products (like WRF an NLDAS2) (Figure 4.9). Additionally, snow accumulation patterns identified from the SWE reanalysis better-represented snow deposition heterogeneity than terrain-based snow deposition downscaling, even capturing a decrease in mean SWE in the upper-elevations of the Upper Tuolumne and Upper Kings watersheds (Figure 4.12).

The methods presented in this dissertation focused on improving estimates of snow magnitude and spatial heterogeneity by implicitly representing more complex snow accumulation processes.

Currently, models (atmospheric, land-surface, and coupled land-atmosphere models) and dynamic downscaling routines are both being developed to more-physically represent storm-specific variations in hillslope (< 100 m) snow accumulation. However, the ability to resolve spatiotemporal snow mass and distribution with accuracies important for hydrologic forecasting, ecologic refugia, avalanche mitigation, and wildlife habitat, is still lacking. Here, we improved spatiotemporal snow depth and SWE estimates using practical, easy-to-implement strategies and tools that we have at our disposal today. We think that the methods presented in Chapter 4, driven by the lessons learned in Chapters 2 and 3, provide the most practical and promising path forward for mountainous snow modeling. However, this work was focused in the California Sierra Nevada, where storms come predominantly from the Pacific. Therefore, snow deposition in this region may have more interannually prevailing snow accumulation patterns than regions with more-variable snowstorm characteristics. Additionally, the accuracy of watershed-scale mean snowfall inferred from the relationship between snow pillow observations and historic snow accumulation patterns tended to decrease in years with fewer snow pillow observations. Therefore, snow pattern repeatability and the methods presented in Chapter 4 should be investigated in other, less-densely gauged mountainous regions with more-variable storm patterns. Chapter 4 also used a multidecadal snow reanalysis (Margulis et al., 2016a) that was at spatial resolutions (90 m) coarser than some of the snow accumulation processes (e.g., wind-redistribution, preferential deposition, etc.) that influence snow heterogeneity and the resulting spring snowmelt. Therefore, future work should consider SWE datasets reanalyzed in time and space from finer-resolution snow cover observations (e.g., Cannistra et al., 2021).

Finally, this dissertation demonstrated the value of snow modeling efforts that fuse information from different models, observations, and snow deposition downscaling strategies.

Future efforts should strive to represent spatiotemporal snow evolution not solely with physically based models, but instead with model frameworks that combine all the information and tools at our disposal. For instance, although snow models are increasingly being developed to include representations of wind-redistribution, historic snow depth patterns have demonstrated the ability to improve parameterizations of simulated wind-drifting in the Alaskan arctic (Sturm and Wagner, 2010a). Snow patterns identified from snow reanalyses (like those evaluated in Chapter 4) could also provide a basis for snow spatial extrapolation (methods from Chapter 2) using other snow depth observations, such as altimetry from the IceSat-2 satellite platform (Abdalati et al., 2010). Strategies like these that optimize the trade-space between model representations of snow processes, ground- and remotely sensed snow observations, and implicit substitutions for mismodeled, unmodeled, or unobserved processes are best-suited to 1) improve real-time operational and scientific modeling, and 2) identify the gaps of greatest importance for future generations of snow models and snow observation campaigns.

Chapter 6 : References

Aas, K.S., Gisnås, K., Westermann, S., Berntsen, T.K., 2017. A Tiling Approach to Represent Subgrid Snow Variability in Coupled Land Surface–Atmosphere Models. J. Hydrometeor. 18, 49–63. https://doi.org/10.1175/JHM-D-16-0026.1

Abdalati, W., Zwally, H.J., Bindschadler, R., Csatho, B., Farrell, S.L., Fricker, H.A., Harding, D., Kwok, R., Lefsky, M., Markus, T., Marshak, A., Neumann, T., Palm, S., Schutz, B., Smith, B., Spinhirne, J., Webb, C., 2010. The ICESat-2 Laser Altimetry Mission. Proceedings of the IEEE 98, 735–751. https://doi.org/10.1109/JPROC.2009.2034765

Adams, M.S., Bühler, Y., Fromm, R., 2018. Multitemporal Accuracy and Precision Assessment of Unmanned Aerial System Photogrammetry for Slope-Scale Snow Depth Maps in Alpine Terrain. Pure Appl. Geophys. 175, 3303–3324. https://doi.org/10.1007/s00024-017-1748-y

Albers, S.C., McGinley, J.A., Birkenheuer, D.L., Smart, J.R., 1996. The Local Analysis and Prediction System (LAPS): Analyses of Clouds, Precipitation, and Temperature. Weather and Forecasting 11, 273–287. https://doi.org/10.1175/1520-0434(1996)011<0273:TLAAPS>2.0.CO;2

Balk, B., Elder, K., 2000. Combining binary decision tree and geostatistical methods to estimate snow distribution in a mountain watershed. Water Resources Research 36, 13–26. https://doi.org/10.1029/1999WR900251

Barnes, S.L., 1964. A Technique for Maximizing Details in Numerical Weather Map Analysis. Journal of Applied Meteorology 3, 396–409.

Barrett, A.P., 2003. National operational hydrologic remote sensing center snow data assimilation system (SNODAS) products at NSIDC. National Snow and Ice Data Center, Cooperative Institute for Research in Environmental Sciences, Boulder, CO.

Barsugli, J.J., Ray, A.J., Livneh, B., Dewes, C.F., Heldmyer, A., Rangwala, I., Guinotte, J.M., Torbit, S., 2020. Projections of Mountain Snowpack Loss for Wolverine Denning Elevations in the Rocky Mountains. Earth's Future 8, e2020EF001537. https://doi.org/10.1029/2020EF001537

Bavay, M., Egger, T., 2014. MeteoIO 2.4. 2: a preprocessing library for meteorological data. Geoscientific Model Development 7, 3135–3151.

BEA, 2017. . Bureau of Economic Analysis, United States.

Belmecheri, S., Babst, F., Wahl, E.R., Stahle, D.W., Trouet, V., 2016. Multi-century evaluation of Sierra Nevada snowpack. Nature Climate Change 6, 2–3. https://doi.org/10.1038/nclimate2809

Bernhardt, M., Schulz, K., 2010. SnowSlide: A simple routine for calculating gravitational snow transport. Geophys. Res. Lett. 37, L11502. https://doi.org/10.1029/2010GL043086

Blöschl, G., 1999. Scaling issues in snow hydrology. Hydrological Processes 13, 2149–2175. https://doi.org/10.1002/(SICI)1099-1085(199910)13:14/15<2149::AID-HYP847>3.0.CO;2-8

Brandt, T., Cannon, F., Dozier, J., 2019. Assessing WRF's Seasonal Quantitative Precipitation Estimates (QPE) in California's Sierra Nevada Using the Airborne Snow Observatory. AGUFM 2019, H33P-2239.

Brauchli, T., Trujillo, E., Huwald, H., Lehning, M., 2017. Influence of Slope-Scale Snowmelt on Catchment Response Simulated With the *Alpine3D* Model. Water Resources Research 53, 10723–10739. https://doi.org/10.1002/2017WR021278

Cannistra, A.F., Shean, D.E., Cristea, N.C., 2021. High-resolution CubeSat imagery and machine learning for detailed snow-covered area. Remote Sensing of Environment 258, 112399. https://doi.org/10.1016/j.rse.2021.112399

Cao, Q., Painter, T.H., Currier, W.R., Lundquist, J.D., Lettenmaier, D.P., 2017. Estimation of Precipitation over the OLYMPEX Domain during Winter 2015/16. J. Hydrometeor. 19, 143–160. https://doi.org/10.1175/JHM-D-17-0076.1

Cayan, D.R., 1996. Interannual climate variability and snowpack in the western United States. Journal of Climate 9, 928–948.

CDFA (California Department of Food and Agriculture), 2017. , California Agricultural Production Statistics. Sacramento, California.

Clark, M.P., Hendrikx, J., Slater, A.G., Kavetski, D., Anderson, B., Cullen, N.J., Kerr, T., Hreinsson, E.Ö., Woods, R.A., 2011. Representing spatial variability of snow water equivalent in hydrologic and land-surface models: A review. Water Resources Research 47. https://doi.org/10.1029/2011WR010745

Clark, M.P., Nijssen, B., Lundquist, J., Kavetski, D., Rupp, D.E., Woods, R.A., Freer, J.E., Gutmann, E.D., Wood, A.W., Brekke, L.D., Arnold, J.R., Gochis, D.J., Rasmussen, R.M., 2015. A unified approach for process-based hydrologic modeling: 1. Modeling concept. Water Resources Research 51, 2498–2514. https://doi.org/10.1002/2015WR017198

Clark, M.P., Slater, A.G., Barrett, A.P., Hay, L.E., McCabe, G.J., Rajagopalan, B., Leavesley, G.H., 2006. Assimilation of snow covered area information into hydrologic and land-surface models. Advances in Water Resources 29, 1209–1221. https://doi.org/10.1016/j.advwatres.2005.10.001

Cristea, N.C., Lundquist, J.D., Loheide, S.P., Lowry, C.S., Moore, C.E., 2014. Modelling how vegetation cover affects climate change impacts on streamflow timing and magnitude in the snowmelt-dominated upper Tuolumne Basin, Sierra Nevada. Hydrological Processes 28, 3896–3918. https://doi.org/10.1002/hyp.9909

Currier, W.R., Lundquist, J.D., 2018. Snow Depth Variability at the Forest Edge in Multiple Climates in the Western United States. Water Resources Research 54, 8756–8773. https://doi.org/10.1029/2018WR022553

Currier, W.R., Pflug, J., Mazzotti, G., Jonas, T., Deems, J.S., Bormann, K.J., Painter, T.H., Hiemstra, C.A., Gelvin, A., Uhlmann, Z., Spaete, L., Glenn, N.F., Lundquist, J.D., 2019. Comparing Aerial Lidar Observations With Terrestrial Lidar and Snow-Probe Transects From NASA's 2017 SnowEx Campaign. Water Resources Research 55, 6285–6294. https://doi.org/10.1029/2018WR024533

Currier, W.R., Thorson, T., Lundquist, J.D., 2017. Independent Evaluation of Frozen Precipitation from WRF and PRISM in the Olympic Mountains. Journal of Hydrometeorology 18, 2681–2703.

Dadic, R., Mott, R., Lehning, M., Burlando, P., 2010a. Parameterization for wind–induced preferential deposition of snow. Hydrological Processes 24, 1994–2006. https://doi.org/10.1002/hyp.7776

Dadic, R., Mott, R., Lehning, M., Burlando, P., 2010b. Wind influence on snow depth distribution and accumulation over glaciers. J. Geophys. Res. 115, F01012. https://doi.org/10.1029/2009JF001261

Daly, C., Halbleib, M., Smith, J.I., Gibson, W.P., Doggett, M.K., Taylor, G.H., Curtis, J., Pasteris, P.P., 2008. Physiographically sensitive mapping of climatological temperature

and precipitation across the conterminous United States. International Journal of Climatology 28, 2031–2064. https://doi.org/10.1002/joc.1688
Daly, C., Taylor, G., Gibson, W., 1997. The PRISM approach to mapping precipitation and temperature, in: 10th Conference on Applied Climatology, American Meteorological Society, 20-23 October, Reno NV. pp. 10–12.
Davidson, A., McKerrow, A., 2016. GAP/LANDFIRE National Terrestrial Ecosystems 2011. https://doi.org/10.5066/F7ZS2TM0
Deems, J.S., Fassnacht, S.R., Elder, K.J., 2008. Interannual Consistency in Fractal Snow Depth Patterns at Two Colorado Mountain Sites. J. Hydrometeor. 9, 977–988. https://doi.org/10.1175/2008JHM901.1
Dozier, J., Bair, E.H., Davis, R.E., 2016. Estimating the spatial distribution of snow water equivalent in the world's mountains. Wiley Interdisciplinary Reviews: Water 3, 461–474. https://doi.org/10.1002/wat2.1140
Dressler, K.A., Fassnacht, S.R., Bales, R.C., 2006. A Comparison of Snow Telemetry and Snow Course Measurements in the Colorado River Basin. Journal of Hydrometeorology 7, 705–712. https://doi.org/10.1175/JHM506.1
Egli, L., Jonas, T., 2009. Hysteretic dynamics of seasonal snow depth distribution in the Swiss Alps. Geophysical Research Letters 36. https://doi.org/10.1029/2008GL035545
Egli, L., Jonas, T., Grünewald, T., Schirmer, M., Burlando, P., 2012. Dynamics of snow ablation in a small Alpine catchment observed by repeated terrestrial laser scans. Hydrol. Process. 26, 1574–1585. https://doi.org/10.1002/hyp.8244
Elder, K., Dozier, J., Michaelsen, J., 1991. Snow accumulation and distribution in an Alpine Watershed. Water Resour. Res. 27, 1541–1552. https://doi.org/10.1029/91WR00506
Essery, R., Morin, S., Lejeune, Y., B Ménard, C., 2013. A comparison of 1701 snow models using observations from an alpine site. Advances in Water Resources, Snow–Atmosphere Interactions and Hydrological Consequences 55, 131–148. https://doi.org/10.1016/j.advwatres.2012.07.013
Eyring, V., Bony, S., Meehl, G.A., Senior, C.A., Stevens, B., Stouffer, R.J., Taylor, K.E., 2016. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geoscientific Model Development 9, 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016
Feld, S.I., Cristea, N.C., Lundquist, J.D., 2013. Representing atmospheric moisture content along mountain slopes: Examination using distributed sensors in the Sierra Nevada, California. Water Resources Research 49, 4424–4441. https://doi.org/10.1002/wrcr.20318
Filhol, S., Sturm, M., 2019. The smoothing of landscapes during snowfall with no wind. Journal of Glaciology 65, 173–187. https://doi.org/10.1017/jog.2018.104
Freudiger, D., Kohn, I., Seibert, J., Stahl, K., Weiler, M., 2017. Snow redistribution for the hydrological modeling of alpine catchments. WIREs Water 4, e1232. https://doi.org/10.1002/wat2.1232
Gerber, F., Mott, R., Lehning, M., 2019. The Importance of Near-Surface Winter Precipitation Processes in Complex Alpine Terrain. Journal of Hydrometeorology 20, 177–196. https://doi.org/10.1175/JHM-D-18-0055.1
Groisman, P.Y., Legates, D.R., 1994. The Accuracy of United States Precipitation Data. Bull. Amer. Meteor. Soc. 75, 215–228. https://doi.org/10.1175/1520-0477(1994)075<0215:TAOUSP>2.0.CO;2

Günther, D., Marke, T., Essery, R., Strasser, U., 2019. Uncertainties in Snowpack Simulations—Assessing the Impact of Model Structure, Parameter Choice, and Forcing Data Error on Point-Scale Energy Balance Snow Model Performance. Water Resources Research 55, 2779–2800. https://doi.org/10.1029/2018WR023403

Gutmann, E.D., Rasmussen, R.M., Liu, C., Ikeda, K., Gochis, D.J., Clark, M.P., Dudhia, J., Thompson, G., 2012. A Comparison of Statistical and Dynamical Downscaling of Winter Precipitation over Complex Terrain. J. Climate 25, 262–281. https://doi.org/10.1175/2011JCLI4109.1

Haberkorn, A., Wever, N., Hoelzle, M., Phillips, M., Kenner, R., Bavay, M., Lehning, M., 2017. Distributed snow and rock temperature modelling in steep rock walls using Alpine3D. The Cryosphere 11, 585–607. https://doi.org/10.5194/tc-11-585-2017

Hall, D.K., Foster, J.L., DiGirolamo, N.E., Riggs, G.A., 2012. Snow cover, snowmelt timing and stream power in the Wind River Range, Wyoming. Geomorphology, Geospatial Technologies and Geomorphological Mapping Proceedings of the 41st Annual Binghamton Geomorphology Symposium 137, 87–93. https://doi.org/10.1016/j.geomorph.2010.11.011

Hamlet, A.F., Lettenmaier, D.P., 2005. Production of Temporally Consistent Gridded Precipitation and Temperature Fields for the Continental United States. J. Hydrometeor. 6, 330–336. https://doi.org/10.1175/JHM420.1

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

Havens, S., Marks, D., FitzGerald, K., Masarik, M., Flores, A.N., Kormos, P., Hedrick, A., 2019. Approximating Input Data to a Snowmelt Model Using Weather Research and Forecasting Model Outputs in Lieu of Meteorological Measurements. J. Hydrometeor. 20, 847–862. https://doi.org/10.1175/JHM-D-18-0146.1

He, S., Ohara, N., 2019. Modeling Subgrid Variability of Snow Depth Using the Fokker-Planck Equation Approach. Water Resources Research 55, 3137–3155. https://doi.org/10.1029/2017WR022017

Hedrick, A., 2018. Investigating the Annual Water Balance of a High-Altitude Watershed Using Near-Real Time Lidar Data Integration into a Physically Based Snowmelt Model. Boise State University Theses and Dissertations. https://doi.org/10.18122/td/1476/boisestate

Hedrick, A.R., Marks, D., Havens, S., Robertson, M., Johnson, M., Sandusky, M., Marshall, H.-P., Kormos, P.R., Bormann, K.J., Painter, T.H., 2018. Direct Insertion of NASA Airborne Snow Observatory-Derived Snow Depth Time Series Into the iSnobal Energy Balance Snow Model. Water Resources Research 54, 8045–8063. https://doi.org/10.1029/2018WR023190

Hedrick, A.R., Marks, D., Marshall, H.-P., McNamara, J., Havens, S., Trujillo, E., Sandusky, M., Robertson, M., Johnson, M., Bormann, K.J., Painter, T.H., 2020. From drought to flood: A water balance analysis of the Tuolumne River basin during extreme conditions (2015–2017). Hydrological Processes 34, 2560–2574. https://doi.org/10.1002/hyp.13749

Hedstrom, N.R., Pomeroy, J.W., 1998. Measurements and modelling of snow interception in the boreal forest. Hydrological Processes 12, 1611–1625. https://doi.org/10.1002/(SICI)1099-1085(199808/09)12:10/11<1611::AID-HYP684>3.0.CO;2-4

Helbig, N., van Herwijnen, A., 2017. Subgrid parameterization for snow depth over mountainous terrain from flat field snow depth. Water Resour. Res. 53, 1444–1456. https://doi.org/10.1002/2016WR019872

Henn, B., Clark, M.P., Kavetski, D., McGurk, B., Painter, T.H., Lundquist, J.D., 2016a. Combining snow, streamflow, and precipitation gauge observations to infer basin-mean precipitation: TUOLUMNE SNOW, STREAMFLOW, AND PRECIPITATION. Water Resour. Res. 52, 8700–8723. https://doi.org/10.1002/2015WR018564

Henn, B., Clark, M.P., Kavetski, D., McGurk, B., Painter, T.H., Lundquist, J.D., 2016b. Combining snow, streamflow, and precipitation gauge observations to infer basin-mean precipitation. Water Resour. Res. 52, 8700–8723. https://doi.org/10.1002/2015WR018564

Henn, B., Clark, M.P., Kavetski, D., Newman, A.J., Hughes, M., McGurk, B., Lundquist, J.D., 2018a. Spatiotemporal patterns of precipitation inferred from streamflow observations across the Sierra Nevada mountain range. Journal of Hydrology 556, 993–1012. https://doi.org/10.1016/j.jhydrol.2016.08.009

Henn, B., Newman, A.J., Livneh, B., Daly, C., Lundquist, J.D., 2018b. An assessment of differences in gridded precipitation datasets in complex terrain. Journal of Hydrology 556, 1205–1219. https://doi.org/10.1016/j.jhydrol.2017.03.008

Hiemstra, C.A., Liston, G.E., Pielke, R.A., Birkenheuer, D.L., Albers, S.C., 2006. Comparing Local Analysis and Prediction System (LAPS) Assimilations with Independent Observations. Weather and Forecasting 21, 1024–1040. https://doi.org/10.1175/WAF961.1

Hiemstra, C.A., Liston, G.E., Reiners, W.A., 2002. Snow Redistribution by Wind and Interactions with Vegetation at Upper Treeline in the Medicine Bow Mountains, Wyoming, U.S.A. Arctic, Antarctic, and Alpine Research 34, 262–273. https://doi.org/10.1080/15230430.2002.12003493

Hughes, M., Lundquist, J.D., Henn, B., 2017. Dynamical downscaling improves upon gridded precipitation products in the Sierra Nevada, California. Clim Dyn. https://doi.org/10.1007/s00382-017-3631-z

Iziomon, M.G., Mayer, H., Matzarakis, A., 2003. Downward atmospheric longwave irradiance under clear and cloudy skies: Measurement and parameterization. Journal of Atmospheric and Solar-Terrestrial Physics 65, 1107–1116. https://doi.org/10.1016/j.jastp.2003.07.007

Jankov, I., Bao, J.-W., Neiman, P.J., Schultz, P.J., Yuan, H., White, A.B., 2009. Evaluation and Comparison of Microphysical Algorithms in ARW-WRF Model Simulations of Atmospheric River Events Affecting the California Coast. Journal of Hydrometeorology 10, 847–870. https://doi.org/10.1175/2009JHM1059.1

Jennings, K.S., Winchell, T.S., Livneh, B., Molotch, N.P., 2018. Spatial variation of the rain–snow temperature threshold across the Northern Hemisphere. Nature Communications 9. https://doi.org/10.1038/s41467-018-03629-7

Jonas, T., Marty, C., Magnusson, J., 2009. Estimating the snow water equivalent from snow depth measurements in the Swiss Alps. Journal of Hydrology 378, 161–167. https://doi.org/10.1016/j.jhydrol.2009.09.021

Kapnick, S., Hall, A., 2012. Causes of recent changes in western North American snowpack. Clim Dyn 38, 1885–1899. https://doi.org/10.1007/s00382-011-1089-y

Kim, R.S., Kumar, S., Vuyovich, C., Houser, P., Lundquist, J., Mudryk, L., Durand, M., Barros, A., Kim, E.J., Forman, B.A., Gutmann, E.D., Wrzesien, M.L., Garnaud, C., Sandells, M.,

Marshall, H.-P., Cristea, N., Pflug, J.M., Johnston, J., Cao, Y., Mocko, D., Wang, S., 2021. Snow Ensemble Uncertainty Project (SEUP): quantification of snow water equivalent uncertainty across North America via ensemble land surface modeling. The Cryosphere 15, 771–791. https://doi.org/10.5194/tc-15-771-2021

Kirchner, P.B., Bales, R.C., Molotch, N.P., Flanagan, J., Guo, Q., 2014. LiDAR measurement of seasonal snow accumulation along an elevation gradient in the southern Sierra Nevada, California. Hydrology and Earth System Sciences 18, 4261–4275. https://doi.org/10.5194/hess-18-4261-2014

Koch, S.E., desJardins, M., Kocin, P.J., 1983. An Interactive Barnes Objective Map Analysis Scheme for Use with Satellite and Conventional Data. J. Climate Appl. Meteor. 22, 1487–1503. https://doi.org/10.1175/1520-0450(1983)022<1487:AIBOMA>2.0.CO;2

Kostadinov, T.S., Schumer, R., Hausner, M., Bormann, K.J., Gaffney, R., McGwire, K., Painter, T.H., Tyler, S., Harpold, A.A., 2019. Watershed-scale mapping of fractional snow cover under conifer forest canopy using lidar. Remote Sensing of Environment 222, 34–49. https://doi.org/10.1016/j.rse.2018.11.037

Lannoy, G.J.M.D., Reichle, R.H., Houser, P.R., Arsenault, K.R., Verhoest, N.E.C., Pauwels, V.R.N., 2010. Satellite-Scale Snow Water Equivalent Assimilation into a High-Resolution Land Surface Model. Journal of Hydrometeorology 11, 352–369. https://doi.org/10.1175/2009JHM1192.1

Lapo, K.E., Hinkelman, L.M., Sumargo, E., Hughes, M., Lundquist, J.D., 2017. A critical evaluation of modeled solar irradiance over California for hydrologic and land surface modeling. Journal of Geophysical Research: Atmospheres 122, 299–317. https://doi.org/10.1002/2016JD025527

Li, D., Lettenmaier, D.P., Margulis, S.A., Andreadis, K., 2019. The Value of Accurate High-Resolution and Spatially Continuous Snow Information to Streamflow Forecasts. Journal of Hydrometeorology 20, 731–749. https://doi.org/10.1175/JHM-D-18-0210.1

Li, D., Wrzesien, M.L., Durand, M., Adam, J., Lettenmaier, D.P., 2017. How much runoff originates as snow in the western United States, and how will that change in the future? Geophysical Research Letters 44, 6163–6172. https://doi.org/10.1002/2017GL073551

Liston, G.E., 2004. Representing Subgrid Snow Cover Heterogeneities in Regional and Global Models. J. Climate 17, 1381–1397. https://doi.org/10.1175/1520-0442(2004)017<1381:RSSCHI>2.0.CO;2

Liston, G.E., Elder, K., 2006a. A meteorological distribution system for high-resolution terrestrial modeling (MicroMet). Journal of Hydrometeorology 7, 217–234.

Liston, G.E., Elder, K., 2006b. A Distributed Snow-Evolution Modeling System (SnowModel). J. Hydrometeor. 7, 1259–1276. https://doi.org/10.1175/JHM548.1

Liston, G.E., Haehnel, R., Sturm, M., A Hiemstra, C., Stuefer, S., D. Tabler, R., 2007. Instruments and Methods Simulating complex snow distributions in windy environments using SnowTran-3D. Journal of Glaciology 53. https://doi.org/10.3189/172756507782202865

Liston, G.E., Perham, C.J., Shideler, R.T., Cheuvront, A.N., 2016. Modeling snowdrift habitat for polar bear dens. Ecological Modelling 320, 114–134. https://doi.org/10.1016/j.ecolmodel.2015.09.010

Liu, Y., Margulis, S.A., 2019. Deriving Bias and Uncertainty in MERRA-2 Snowfall Precipitation Over High Mountain Asia. Front. Earth Sci. 7. https://doi.org/10.3389/feart.2019.00280

Livneh, B., Bohn, T.J., Pierce, D.W., Munoz-Arriola, F., Nijssen, B., Vose, R., Cayan, D.R., Brekke, L., 2015. A spatially comprehensive, hydrometeorological data set for Mexico, the U.S., and Southern Canada 1950–2013. Scientific Data 2, 1–12. https://doi.org/10.1038/sdata.2015.42

López-Moreno, J., Revuelto, J., Alonso-González, E., Sanmiguel-Vallelado, A., Fassnacht, S., Deems, J., Morán-Tejeda, E., 2017. Using very long-range terrestrial laser scanner to analyze the temporal consistency of the snowpack distribution in a high mountain environment. Journal of Mountain Science 14, 823–842. https://doi.org/10.1007/s11629-016-4086-0

Luce, C.H., Tarboton, D.G., Cooley, K.R., 1998. The influence of the spatial distribution of snow on basin-averaged snowmelt. Hydrological Processes 12, 1671–1683. https://doi.org/10.1002/(SICI)1099-1085(199808/09)12:10/11<1671::AID-HYP688>3.0.CO;2-N

Lundquist, J., Hughes, M., Gutmann, E., Kapnick, S., 2019. Our skill in modeling mountain rain and snow is bypassing the skill of our observational networks. Bull. Amer. Meteor. Soc. https://doi.org/10.1175/BAMS-D-19-0001.1

Lundquist, J.D., Dettinger, M.D., 2005. How snowpack heterogeneity affects diurnal streamflow timing. Water Resources Research 41. https://doi.org/10.1029/2004WR003649

Lundquist, J.D., Dettinger, M.D., Cayan, D.R., 2005. Snow-fed streamflow timing at different basin scales: Case study of the Tuolumne River above Hetch Hetchy, Yosemite, California. Water Resources Research 41.

Lundquist, J.D., Dickerson-Lange, S.E., Lutz, J.A., Cristea, N.C., 2013. Lower forest density enhances snow retention in regions with warmer winters: A global framework developed from plot-scale observations and modeling. Water Resources Research 49, 6356–6370. https://doi.org/10.1002/wrcr.20504

Lundquist, J.D., Hughes, M., Henn, B., Gutmann, E.D., Livneh, B., Dozier, J., Neiman, P., 2015. High-Elevation Precipitation Patterns: Using Snow Measurements to Assess Daily Gridded Datasets across the Sierra Nevada, California. J. Hydrometeor. 16, 1773–1792. https://doi.org/10.1175/JHM-D-15-0019.1

Lundquist, J.D., Minder, J.R., Neiman, P.J., Sukovich, E., 2010. Relationships between Barrier Jet Heights, Orographic Precipitation Gradients, and Streamflow in the Northern Sierra Nevada. Journal of Hydrometeorology 11, 1141–1156. https://doi.org/10.1175/2010JHM1264.1

Luo, L., Robock, A., Mitchell, K.E., Houser, P.R., Wood, E.F., Schaake, J.C., Lohmann, D., Cosgrove, B., Wen, F., Sheffield, J., Duan, Q., Higgins, R.W., Pinker, R.T., Tarpley, J.D., 2003. Validation of the North American Land Data Assimilation System (NLDAS) retrospective forcing over the southern Great Plains. Journal of Geophysical Research: Atmospheres 108. https://doi.org/10.1029/2002JD003246

Marchand, W.-D., Killingtveit, Å., 2005. Statistical probability distribution of snow depth at the model sub-grid cell spatial scale. Hydrological Processes 19, 355–369. https://doi.org/10.1002/hyp.5543

Margulis, S.A., Cortés, G., Girotto, M., Durand, M., 2016a. A Landsat-Era Sierra Nevada Snow Reanalysis (1985–2015). J. Hydrometeor. 17, 1203–1221. https://doi.org/10.1175/JHM-D-15-0177.1

Margulis, S.A., Cortés, G., Girotto, M., Huning, L.S., Li, D., Durand, M., 2016b. Characterizing the extreme 2015 snowpack deficit in the Sierra Nevada (USA) and the implications for

drought recovery. Geophysical Research Letters 43, 6341–6349. https://doi.org/10.1002/2016GL068520
Margulis, S.A., Fang, Y., Li, D., Lettenmaier, D.P., Andreadis, K., 2019a. The Utility of Infrequent Snow Depth Images for Deriving Continuous Space-Time Estimates of Seasonal Snow Water Equivalent. Geophysical Research Letters 46, 5331–5340. https://doi.org/10.1029/2019GL082507
Margulis, S.A., Girotto, M., Cortés, G., Durand, M., 2015. A Particle Batch Smoother Approach to Snow Water Equivalent Estimation. J. Hydrometeor. 16, 1752–1772. https://doi.org/10.1175/JHM-D-14-0177.1
Margulis, S.A., Liu, Y., Baldo, E., 2019b. A Joint Landsat- and MODIS-Based Reanalysis Approach for Midlatitude Montane Seasonal Snow Characterization. Front. Earth Sci. 7. https://doi.org/10.3389/feart.2019.00272
Marks, D., Dozier, J., 1992. Climate and energy exchange at the snow surface in the Alpine Region of the Sierra Nevada: 2. Snow cover energy balance. Water Resources Research 28, 3043–3054. https://doi.org/10.1029/92WR01483
Masiokas, M.H., Villalba, R., Luckman, B.H., Quesne, C.L., Aravena, J.C., 2006. Snowpack Variations in the Central Andes of Argentina and Chile, 1951–2005: Large-Scale Atmospheric Influences and Implications for Water Resources in the Region. Journal of Climate 19, 6334–6352. https://doi.org/10.1175/JCLI3969.1
Maurer, E.P., Wood, A.W., Adam, J.C., Lettenmaier, D.P., Nijssen, B., 2002. A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States. J. Climate 15, 3237–3251. https://doi.org/10.1175/1520-0442(2002)015<3237:ALTHBD>2.0.CO;2
Mazurkiewicz, A.B., Callery, D.G., McDonnell, J.J., 2008. Assessing the controls of the snow energy balance and water available for runoff in a rain-on-snow environment. Journal of Hydrology 354, 1–14. https://doi.org/10.1016/j.jhydrol.2007.12.027
McCabe, G.J., Dettinger, M.D., 2002. Primary Modes and Predictability of Year-to-Year Snowpack Variations in the Western United States from Teleconnections with Pacific Ocean Climate. J. Hydrometeor. 3, 13–25. https://doi.org/10.1175/1525-7541(2002)003<0013:PMAPOY>2.0.CO;2
McGrath, D., Sass, L., O'Neel, S., McNeil, C., Candela, S.G., Baker, E.H., Marshall, H.P., 2018. Interannual snow accumulation variability on glaciers derived from repeat, spatially extensive ground-penetrating radar surveys. The Cryosphere 12, 3617–3633. https://doi.org/10.5194/tc-12-3617-2018
Mendoza, P.A., Clark, M.P., Barlage, M., Rajagopalan, B., Samaniego, L., Abramowitz, G., Gupta, H., 2015. Are we unnecessarily constraining the agility of complex process-based models? Water Resources Research 51, 716–728. https://doi.org/10.1002/2014WR015820
Mendoza, P.A., Musselman, K.N., Revuelto, J., Deems, J.S., López-Moreno, J.I., McPhee, J., 2020. Interannual and Seasonal Variability of Snow Depth Scaling Behavior in a Subalpine Catchment. Water Resources Research 56, e2020WR027343. https://doi.org/10.1029/2020WR027343
Meromy, L., Molotch, N.P., Link, T.E., Fassnacht, S.R., Rice, R., 2013. Subgrid variability of snow water equivalent at operational snow stations in the western USA. Hydrological Processes 27, 2383–2400. https://doi.org/10.1002/hyp.9355

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P.C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E.H., Ek, M.B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D., Shi, W., 2006. North American Regional Reanalysis. Bull. Amer. Meteor. Soc. 87, 343–360. https://doi.org/10.1175/BAMS-87-3-343

Minder, J.R., Durran, D.R., Roe, G.H., Anders, A.M., 2008. The climatology of small-scale orographic precipitation over the Olympic Mountains: Patterns and processes. Quarterly Journal of the Royal Meteorological Society 134, 817–839. https://doi.org/10.1002/qj.258

Minder, J.R., Letcher, T.W., Skiles, S.M., 2016. An evaluation of high-resolution regional climate model simulations of snow cover and albedo over the Rocky Mountains, with implications for the simulated snow-albedo feedback. Journal of Geophysical Research: Atmospheres 121, 9069–9088. https://doi.org/10.1002/2016JD024995

Minder, J.R., Mote, P.W., Lundquist, J.D., 2010. Surface temperature lapse rates over complex terrain: Lessons from the Cascade Mountains. Journal of Geophysical Research: Atmospheres 115. https://doi.org/10.1029/2009JD013493

Moeser, D., Stähli, M., Jonas, T., 2015. Improved snow interception modeling using canopy parameters derived from airborne LiDAR data. Water Resources Research 51, 5041–5059.

Molotch, N.P., Bales, R.C., 2005. Scaling snow observations from the point to the grid element: Implications for observation network design. Water Resources Research 41. https://doi.org/10.1029/2005WR004229

Montoya, E.L., Dozier, J., Meiring, W., 2014. Biases of April 1 snow water equivalent records in the Sierra Nevada and their associations with large-scale climate indices. Geophysical Research Letters 41, 5912–5918. https://doi.org/10.1002/2014GL060588

Morales, A., Morrison, H., Posselt, D.J., 2018. Orographic Precipitation Response to Microphysical Parameter Perturbations for Idealized Moist Nearly Neutral Flow. Journal of the Atmospheric Sciences 75, 1933–1953. https://doi.org/10.1175/JAS-D-17-0389.1

Morrison, H., Thompson, G., Tatarskii, V., 2009. Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one-and two-moment schemes. Monthly weather review 137, 991–1007.

Mote, P.W., 2006. Climate-Driven Variability and Trends in Mountain Snowpack in Western North America. J. Climate 19, 6209–6220. https://doi.org/10.1175/JCLI3971.1

Mote, P.W., Hamlet, A.F., Clark, M.P., Lettenmaier, D.P., 2005. Declining mountain snowpack in western North America. Bulletin of the American Meteorological Society 86, 39–49. https://doi.org/10.1175/BAMS-86-1-39

Mote, P.W., Parson, E.A., Hamlet, A.F., Keeton, W.S., Lettenmaier, D.P., Mantua, N., Miles, E.L., Peterson, D.W., Peterson, D.L., Slaughter, R., Snover, A.K., 2003. Preparing for climate change: The water, salmon, and forests of the Pacific Northwest. Climatic Change 61, 45–88.

Mott, R., Schirmer, M., Bavay, M., Grünewald, T., Lehning, M., 2010. Understanding snow-transport processes shaping the mountain snow-cover. The Cryosphere 4, 545–559. https://doi.org/10.5194/tc-4-545-2010

Mott, R., Schlögl, S., Dirks, L., Lehning, M., 2017. Impact of Extreme Land Surface Heterogeneity on Micrometeorology over Spring Snow Cover. J. Hydrometeor. 18, 2705–2722. https://doi.org/10.1175/JHM-D-17-0074.1

Mott, R., Scipión, D., Schneebeli, M., Dawes, N., Berne, A., Lehning, M., 2014. Orographic effects on snow deposition patterns in mountainous terrain. Journal of Geophysical Research: Atmospheres 119, 1419–1439. https://doi.org/10.1002/2013JD019880

Musselman, K.N., Pomeroy, J.W., Essery, R.L.H., Leroux, N., 2015. Impact of windflow calculations on simulations of alpine snow accumulation, redistribution and ablation. Hydrological Processes 29, 3983–3999. https://doi.org/10.1002/hyp.10595

Nayak, A., Marks, D., Chandler, D.G., Seyfried, M., 2010. Long-term snow, climate, and streamflow trends at the Reynolds Creek Experimental Watershed, Owyhee Mountains, Idaho, United States. Water Resources Research 46. https://doi.org/10.1029/2008WR007525

Oroza, C.A., Zheng, Z., Glaser, S.D., Tuia, D., Bales, R.C., 2016. Optimizing embedded sensor network design for catchment-scale snow-depth estimation using LiDAR and machine learning: OPTIMIZING SNOW SENSOR PLACEMENTS. Water Resour. Res. 52, 8174–8189. https://doi.org/10.1002/2016WR018896

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

Pan, M., Sheffield, J., Wood, E.F., Mitchell, K.E., Houser, P.R., Schaake, J.C., Robock, A., Lohmann, D., Cosgrove, B., Duan, Q., Luo, L., Higgins, R.W., Pinker, R.T., Tarpley, J.D., 2003. Snow process modeling in the North American Land Data Assimilation System (NLDAS): 2. Evaluation of model simulated snow water equivalent. Journal of Geophysical Research: Atmospheres 108. https://doi.org/10.1029/2003JD003994

Pflug, J.M., Liston, G.E., Nijssen, B., Lundquist, J.D., 2019. Testing Model Representations of Snowpack Liquid Water Percolation Across Multiple Climates. Water Resources Research 55, 4820–4838. https://doi.org/10.1029/2018WR024632

Pflug, J.M., Lundquist, J.D., 2020. Inferring distributed snow depth by leveraging snow pattern repeatability: Investigation using 47 lidar observations in the Tuolumne watershed, Sierra Nevada, California. Water Resources Research 56, e2020WR027243. https://doi.org/10.1029/2020WR027243

Raleigh, M.S., Lundquist, J.D., Clark, M.P., 2015. Exploring the impact of forcing error characteristics on physically based snow simulations within a global sensitivity analysis framework. Hydrol. Earth Syst. Sci. 19, 3153–3179. https://doi.org/10.5194/hess-19-3153-2015

Rasmussen, R., Liu, C., Ikeda, K., Gochis, D., Yates, D., Chen, F., Tewari, M., Barlage, M., Dudhia, J., Yu, W., Miller, K., Arsenault, K., Grubišić, V., Thompson, G., Gutmann, E., 2011. High-Resolution Coupled Climate Runoff Simulations of Seasonal Snowfall over Colorado: A Process Study of Current and Warmer Climate. J. Climate 24, 3015–3048. https://doi.org/10.1175/2010JCLI3985.1

Reynolds, D.S., Pflug, J.M., Lundquist, J.D., 2021. Evaluating Wind Fields for Use in Basin-Scale Distributed Snow Models. Water Resources Research 57, e2020WR028536. https://doi.org/10.1029/2020WR028536

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

Rittger, K., Bair, E.H., Kahl, A., Dozier, J., 2016. Spatial estimates of snow water equivalent from reconstruction. Advances in Water Resources 94, 345–363. https://doi.org/10.1016/j.advwatres.2016.05.015

Roe, G.H., Baker, M.B., 2006. Microphysical and Geometrical Controls on the Pattern of Orographic Precipitation. J. Atmos. Sci. 63, 861–880. https://doi.org/10.1175/JAS3619.1

Schirmer, M., Lehning, M., 2011. Persistence in intra-annual snow depth distribution: 2. Fractal analysis of snow depth development. Water Resources Research 47. https://doi.org/10.1029/2010WR009429

Schirmer, M., Wirz, V., Clifton, A., Lehning, M., 2011. Persistence in intra-annual snow depth distribution: 1. Measurements and topographic control. Water Resources Research 47.

Schmidt, S., Weber, B., Winiger, M., 2009. Analyses of seasonal snow disappearance in an alpine valley from micro- to meso-scale (Loetschental, Switzerland). Hydrological Processes 23, 1041–1051. https://doi.org/10.1002/hyp.7205

Seyfried, M.S., Wilcox, B.P., 1995. Scale and the Nature of Spatial Variability: Field Examples Having Implications for Hydrologic Modeling. Water Resources Research 31, 173–184. https://doi.org/10.1029/94WR02025

Siderius, C., Biemans, H., Wiltshire, A., Rao, S., Franssen, W.H.P., Kumar, P., Gosain, A.K., van Vliet, M.T.H., Collins, D.N., 2013. Snowmelt contributions to discharge of the Ganges. Science of The Total Environment, Changing water resources availability in Northern India with respect to Himalayan glacier retreat and changing monsoon patterns: consequences and adaptation 468–469, S93–S101. https://doi.org/10.1016/j.scitotenv.2013.05.084

Skamarock, W.C., 2008. A Description of the Advanced Research WRF Version 3 (NCAR Technical Note No. NCAR/TN-475+STR).

Skaugen, T., 1999. Estimating the mean areal snow water equivalent by integration in time and space. Hydrological Processes 13, 2051–2066. https://doi.org/10.1002/(SICI)1099-1085(199909)13:12/13<2051::AID-HYP881>3.0.CO;2-U

Sproles, E.A., Rittger, K., Nolin, A.W., Painter, T.H., 2013. Climate change impacts on maritime mountain snowpack in the Oregon Cascades. Hydrol. Earth Syst. Sci. 17, 2581–2597.

Strachan, S., Kelsey, E.P., Brown, R.F., Dascalu, S., Harris, F., Kent, G., Lyles, B., McCurdy, G., Slater, D., Smith, K., 2016. Filling the Data Gaps in Mountain Climate Observatories Through Advanced Technology, Refined Instrument Siting, and a Focus on Gradients. mred 36, 518–527. https://doi.org/10.1659/MRD-JOURNAL-D-16-00028.1

Sturm, M., Goldstein, M.A., Parr, C., 2017. Water and life from snow: A trillion dollar science question. Water Resources Research 53, 3534–3544. https://doi.org/10.1002/2017WR020840

Sturm, M., Wagner, A.M., 2010a. Using repeated patterns in snow distribution modeling: An Arctic example. Water Resour. Res. 46, W12549. https://doi.org/10.1029/2010WR009434

Sturm, M., Wagner, A.M., 2010b. Using repeated patterns in snow distribution modeling: An Arctic example. Water Resources Research 46.

Taylor, K.E., Stouffer, R.J., Meehl, G.A., 2012. An overview of CMIP5 and the experiment design. Bulletin of the American Meteorological Society 93, 485–498.

Thompson, G., Field, P.R., Rasmussen, R.M., Hall, W.D., 2008. Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization. Mon. Wea. Rev. 136, 5095–5115. https://doi.org/10.1175/2008MWR2387.1

Thompson, G., Rasmussen, R.M., Manning, K., 2004. Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part I: Description and Sensitivity Analysis. Mon. Wea. Rev. 132, 519–542. https://doi.org/10.1175/1520-0493(2004)132<0519:EFOWPU>2.0.CO;2

Thornton, P.E., Running, S.W., White, M.A., 1997. Generating surfaces of daily meteorological variables over large regions of complex terrain. Journal of Hydrology, Aggregate Description of Land-Atmosphere Interactions 190, 214–251. https://doi.org/10.1016/S0022-1694(96)03128-9

Thornton, P.E., Thornton, M.M., Mayer, B.W., Wilhelmi, N., Wei, Y., Devarakonda, R., Cook, R.B., 2014. Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 2. Oak Ridge National Lab.(ORNL), Oak Ridge, TN (United States).

Trapero, L., Bech, J., Rigo, T., Pineda, N., Forcadell, D., 2009. Uncertainty of precipitation estimates in convective events by the Meteorological Service of Catalonia radar network. Atmospheric Research, 4th European Conference on Severe Storms 93, 408–418. https://doi.org/10.1016/j.atmosres.2009.01.021

Trujillo, E., Ramírez, J.A., Elder, K.J., 2007. Topographic, meteorologic, and canopy controls on the scaling characteristics of the spatial distribution of snow depth fields. Water Resources Research 43. https://doi.org/10.1029/2006WR005317

Vano, J.A., Nijssen, B., Lettenmaier, D.P., 2015. Seasonal hydrologic responses to climate change in the Pacific Northwest. Water Resources Research 51, 1959–1976. https://doi.org/10.1002/2014WR015909

Viste, E., Sorteberg, A., 2015. Snowfall in the Himalayas: An uncertain future from a little-known past. https://doi.org/10.5194/tc-9-1147-2015

Vögeli, C., Lehning, M., Wever, N., Bavay, M., 2016. Scaling Precipitation Input to Spatially Distributed Hydrological Models by Measured Snow Distribution. Front. Earth Sci. 4. https://doi.org/10.3389/feart.2016.00108

Walcek, C.J., 1994. Cloud Cover and Its Relationship to Relative Humidity during a Springtime Midlatitude Cyclone. Mon. Wea. Rev. 122, 1021–1035. https://doi.org/10.1175/1520-0493(1994)122<1021:CCAIRT>2.0.CO;2

Wayand, N.E., Hamlet, A.F., Hughes, M., Feld, S.I., Lundquist, J.D., 2013. Intercomparison of Meteorological Forcing Data from Empirical and Mesoscale Model Sources in the North Fork American River Basin in Northern Sierra Nevada, California. Journal of Hydrometeorology 14, 677–699. https://doi.org/10.1175/JHM-D-12-0102.1

Wayand, N.E., Stimberis, J., Zagrodnik, J., Mass, C.F., Lundquist, J.D., 2016. Improving simulations of precipitation phase and snowpack at a site subject to cold air intrusions: Snoqualmie Pass, WA. Journal of Geophysical Research: Atmospheres 121, 9929–9942. https://doi.org/10.1002/2016JD025387

Winstral, A., Elder, K., Davis, R.E., 2002. Spatial Snow Modeling of Wind-Redistributed Snow Using Terrain-Based Parameters. J. Hydrometeor. 3, 524–538. https://doi.org/10.1175/1525-7541(2002)003<0524:SSMOWR>2.0.CO;2

Winstral, A., Magnusson, J., Schirmer, M., Jonas, T., 2019. The Bias-Detecting Ensemble: A New and Efficient Technique for Dynamically Incorporating Observations Into Physics-

Based, Multilayer Snow Models. Water Resources Research 55, 613–631. https://doi.org/10.1029/2018WR024521
Winstral, A., Marks, D., Gurney, R., 2013. Simulating wind-affected snow accumulations at catchment to basin scales. Advances in Water Resources, Snow–Atmosphere Interactions and Hydrological Consequences 55, 64–79. https://doi.org/10.1016/j.advwatres.2012.08.011
Woodruff, C.D., Qualls, R.J., 2019. Recurrent Snowmelt Pattern Synthesis using Principal Component Analysis of Multi-Year Remotely Sensed Snow Cover. Water Resources Research 55, 6869–6885. https://doi.org/10.1029/2018WR024546
Wrzesien, M.L., Pavelsky, T.M., Kapnick, S.B., Durand, M.T., Painter, T.H., 2015. Evaluation of snow cover fraction for regional climate simulations in the Sierra Nevada. International Journal of Climatology 35, 2472–2484. https://doi.org/10.1002/joc.4136
Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, L., Alonge, C., Wei, H., Meng, J., Livneh, B., Lettenmaier, D., Koren, V., Duan, Q., Mo, K., Fan, Y., Mocko, D., 2012. Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. Journal of Geophysical Research: Atmospheres 117. https://doi.org/10.1029/2011JD016048
Xue, Y., Sellers, P.J., Kinter, J.L., Shukla, J., 1991. A Simplified Biosphere Model for Global Climate Studies. J. Climate 4, 345–364. https://doi.org/10.1175/1520-0442(1991)004<0345:ASBMFG>2.0.CO;2
Ye, B., Yang, D., Ma, L., 2012. Effect of precipitation bias correction on water budget calculation in Upper Yellow River, China. Environ. Res. Lett. 7, 025201. https://doi.org/10.1088/1748-9326/7/2/025201