Water Resources Research

RESEARCH ARTICLE
10.1029/2021WR029999

Special Section:
Advances in remote sensing, measurement, and simulation of seasonal snow

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], and 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

Key Points:

- Interannually repeatable snow patterns could downscale snow deposition in space as a proxy for unmodeled snowfall and snow redistribution
- Winter snow losses and snow density spatial variability high-biased the spatial heterogeneity of pattern-based snow deposition downscaling
- Modeled snow depth accuracy was more limited by common mean snowfall biases than different snow deposition downscaling approaches

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

J. M. Pflug,
jpflug@uw.edu

Citation:

Pflug, J. M., Hughes, M., & Lundquist, J. D. (2021). Downscaling snow deposition using historic snow depth patterns: Diagnosing limitations from snowfall biases, winter snow losses, and interannual snow pattern repeatability. *Water Resources Research*, *57*, e2021WR029999. https://doi.org/10.1029/2021WR029999

Received 15 MAR 2021
Accepted 14 JUL 2021

Abstract Repeatable snow depth patterns have been identified in many regions between years with similar meteorological characteristics. This suggests that snow patterns from previous years could adjust snow deposition in space as a substitution for unmodeled snow processes. Here, we tested a pattern-based snow deposition downscaling routine which assumes (a) a spatially consistent relationship between snow deposition and snow depth, (b) interannually repeatable snow patterns, and (c) unbiased mean 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 snow depth patterns from seven different years, and was compared to a more common terrain-based downscaling method. Snow depth patterns were influenced not only by snow accumulation, but also snowmelt, snow sublimation, and snow density, resulting in pattern-based snow deposition downscaling that was too spatially heterogeneous. However, snow depth simulated using terrain-based downscaling was too spatially homogeneous, and less spatially correlated with observations ($r = 0.27$), than simulations with pattern-based downscaling using snow depth patterns from the simulation season ($r = 0.76$), or from a different year ($r = 0.52$). Overall, modeled snow depth errors at peak-snowpack timing 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 losses and snow density spatial variability, and identifying the historic periods of most-similar snow accumulation.

1. Introduction

The timing and magnitude of mountain snowmelt is of vital importance for flood management, hydropower generation, local ecology, and recreation (Lundquist & 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 & 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 & 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 & Egger, 2014; Liston & 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 & 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 & Egger, 2014; Daly et al., 2008; Liston & 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 & Legates, 1994; Henn et al., 2016; 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 data sets are often biased (e.g., Liu & Margulis, 2019; Viste & Sorteberg, 2015; Wayand et al., 2013).

Fortunately, snow deposition is driven by the interaction between prevailing snowstorm characteristics and static features like terrain and vegetation, resulting in snowpack spatial distribution patterns that often have interannually repeatable characteristics (Deems et al., 2008; Mendoza et al., 2020; Pflug & Lundquist, 2020; Sturm & Wagner, 2010). 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 the relationship between snow deposition (total SWE input from snowfall and snow redistribution) and snow depth that is identical at all gridcells. In other words, pattern-based downscaling assumes that the snow depth pattern used to downscale snow deposition has (a) spatially homogeneous snow density, and (b) is not influenced by winter snowmelt or snow sublimation. Pattern-based snow deposition downscaling also requires an unbiased 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 seven 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 & Elder, 2006b), with snow deposition downscaled using both pattern-based downscaling (Section 2.1) and a widely used terrain-based downscaling method (Section 2.2). In this paper, we address three overarching issues:

1. The relationship between snow deposition and snow depth near peak-snowpack timing: How much do winter snow losses (snowmelt and sublimation) and snow density influence the relationship between snow deposition and snow depth, and how does this affect the accuracy of snow depth simulations using pattern-based snow deposition downscaling (Vögeli et al., 2016)? How much better or worse is pattern-based downscaling as compared to a popular terrain-based downscaling method (Liston & Elder, 2006a)?
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 seven different years? What errors 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 in the California Sierra Nevada?

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 diagnose the origins and magnitudes of model errors for each downscaling (pattern-based and terrain-based). We then compared 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 provided guidance on how to best prioritize future research and implement the idea of snow-pattern repeatability in practical operations.

2. Background

2.1. Pattern-Based Snow Deposition Downscaling

Atmospheric models can provide meteorological forcing data exceeding the availability of observations in a 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 (SWE input) 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),

$$p^t_{x,y} = V_{x,y} \times \mu^t_p, \tag{1}$$

where μ^t_p 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,

$$V_{x,y} = \frac{d_{x,y}}{\mu_d}, \tag{2}$$

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: (a) every snowfall event deposits snow with spatial heterogeneity identical to that of a distributed snow depth observation, (b) an X-percent change in snow deposition (relative to mean snowfall; μ^t_p) will result in an X-percent change in snow depth (relative to mean snow depth; μ_d), and (c) mean snowfall (μ^t_p) 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 snow accumulation. In other words, 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^t_p. This assumes a linear relationship ($d_{x,y} = m \sum_t p^t_{x,y}$), of slope $m = \mu_d \,/\, \sum_t \mu^t_p$, between the snow depth ($d_{x,y}$) and cumulative snow deposition ($\sum_t p^t_{x,y}$) 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 & Qualls, 2019), Swiss Alps (Schirmer et al., 2011; Schirmer & Lehning, 2011; Vögeli et al., 2016), Alaskan interior (Sturm & Wagner, 2010), Spanish Pyrenees (López-Moreno et al., 2017), and California Sierra Nevada (Pflug & 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 $r > 0.84$) (Pflug & Lundquist, 2020). 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 2 normalizes snow depth patterns using only mean snow depth, and therefore assumes that matching snow depth patterns have identical coefficients of variation (ratios between snow depth standard deviation and mean snow depth). 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 (i.e., constant coefficient of variation). Pflug and Lundquist (2020) noted that the ratio between Tuolumne mean snow depth and snow depth standard deviation 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 seven years, including years with similar ($r = 0.89$) and less-correlated ($r = 0.71$) snow depth patterns.

2.2. 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 & Egger, 2014; Liston & Elder, 2006a). Here, we focus on MicroMet (MM; Liston & Elder, 2006a), which has been used to distribute precipitation in a 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 & Egger, 2014; Daly et al., 2008). Provided precipitation magnitude from a coarser-resolution data set (such as an atmospheric model), MM smooths both elevation and precipitation from the data set 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, MM downscales precipitation (p) at each gridcell (x,y) and model time-step (t) using a nonlinear precipitation lapse rate,

$$p^t_{x,y} = \left[\frac{1 + \beta^t\left(z_{x,y} - z_{x_0,y_0}\right)}{1 - \beta^t\left(z_{x,y} - z_{x_0,y_0}\right)}\right] \times p^t_{x_0,y_0}, \tag{3}$$

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 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.

2.3. 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. Uncertainties can stem from errors in the observations (e.g., gauge-undercatch) 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). Gridded precipitation products also tend to 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). Even if precipitation magnitude is correct, snowfall is 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 data sets, 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 observation-based data sets in mountainous terrain (Lundquist et al., 2019). However, atmospheric models are still subject to uncertainties. These uncertainties are often demonstrated by sensitivities to different boundary conditions, convections schemes, microphysical partitioning methods, and other parameterizations in atmospheric models (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 demonstrated for atmospheric models in the Tuolumne region (Henn, Newman, et al., 2018; Hughes et al., 2017; Lundquist et al., 2015).

3. 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 (Painter et al., 2016) were used to both define snow deposition downscaling factors (Equation 2) and evaluate snow depth simulations (discussed later in this section). We subset ALS observations across three 6.25-by-6.25 km subdomains (Figure 1) including (a) steep slopes with a variety of different aspects (near

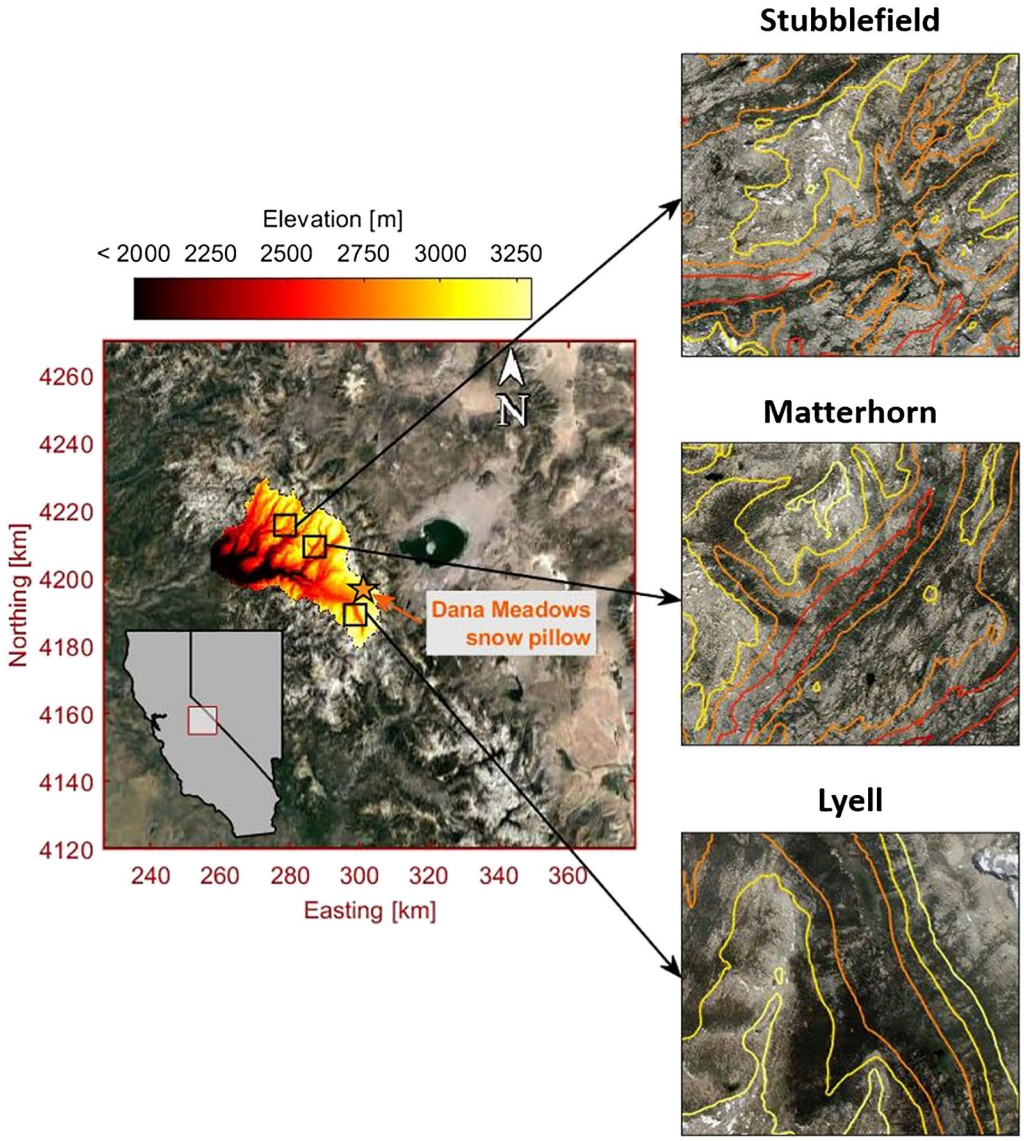

Figure 1. Elevation (colorbar) for the Tuolumne watershed (left) and modeling domains (right, contours). The location of the Dana Meadows snow pillow is shown in the leftmost plot.

Stubblefield canyon), (b) terrain with spatially variable vegetation (Matterhorn canyon), and (c) a North/South oriented canyon with wind-drifted snow and avalanching (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 data sets were regridded from 3 m (native-resolution of the ALS observations) to 25 m spatial resolution. Although this spatial resolution smoothed the effects of snow processes at spatial reaches less than 25 m, regridding: (a) isolated the dominant source(s) of snow deposition (e.g., Clark et al., 2011; Deems et al., 2008; Mendoza et al., 2020), and (b) increased the likelihood of interannually repeatable patterns (Deems et al., 2008; Pflug & Lundquist, 2020). However, we expect that the pattern-based downscaling investigated here (Equation 1) could also be applied at finer wind-drift spatial scales (<10 m).

6 km 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., 2004, 2008),

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. Air temperature, relative humidity, wind speed, and wind direction were downscaled to the ALS grid (25 m resolution) using the MM meteorological preprocessing routine (Liston & Elder, 2006a). Since WRF has noted issues with shortwave radiation in this region (Lapo et al., 2017), this was calculated within MM using basic solar geometry and cloud fractions calculated from air temperature, relative humidity, and dew point temperature (Walcek, 1994). Estimates of longwave radiation, calculated using the cloud fraction, air temperature, and the longwave scheme from Iziomon et al. (2003) (Figure S1c) were used since these estimates were more accurate than the WRF longwave radiation at a point near the Tuolumne watershed (Figures S1a and S1b).

SnowModel (Liston & 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 nonmelting snow albedo decay functions from Sproles et al. (2013). Snowpack simulations were constrained to a maximum of six 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). Although simulations with wind-redistribution slightly outperformed those without, their results were very similar (Figure S2). In this work, all modeled snowpack redistribution was disabled. Therefore, SWE was influenced only by precipitation (both snow and rain), snowmelt, and 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. To check model performance, SWE was simulated in water-year 2014 at the Dana Meadows snow pillow (Figure 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 S3). 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.

4. Methods

Snow depth simulations were performed in water-year 2014 with snowfall from WRF downscaled using terrain-based (Equation 3) and pattern-based (Equation 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 2, box #1) (described in Section 4.1). We first compared snow depth simulations using snow deposition downscaled with (a) MicroMet terrain-based downscaling (MM, Section 2.2), and (b) a pattern-based snow deposition downscaling factor ($V_{x,y}$, Section 2.1) from the simulation season (April 7, 2014; V14). In Section 4.2, we described how we compared the two downscaling, and identified the sources and magnitudes of model errors (relating to issue #1 in the introduction) using (a) the default ensemble member ($\varepsilon = 1.0$, constant in space and time), and (b) the ensemble members (ε_{adj}, constant in time) that best matched a April 7, 2014 ALS snow depth observation on a gridcell-by-gridcell basis (Figure 2, box #3). In Section 4.3, we described how this analysis was expanded to include $V_{x,y}$ calculated from a different year with spatially correlated snow depth (April 16, 2016; $r = 0.89$), and $V_{x,y}$ calculated from snow depth observations in other years ($0.71 \leq r \leq 0.84$) (relating to issue #2 in the introduction). In Section 4.4, we focused 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 4.5, we described how we compared the sources of model errors (downscaling methods, interannual pattern repeatability, and coarse-scale snowfall biases), both individually, and in combination.

4.1. Model Ensemble Generation

Equation 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

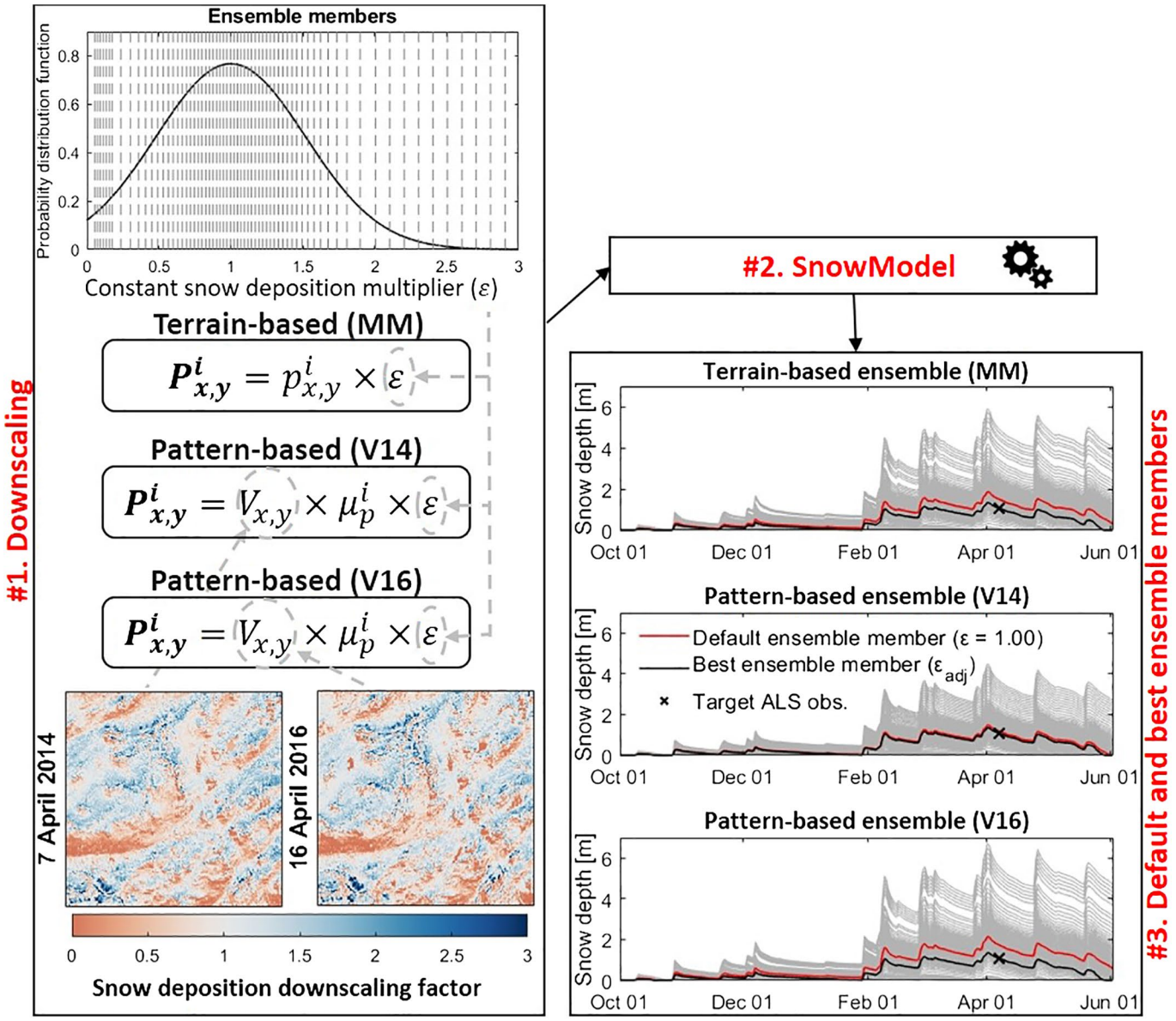

Figure 2. Snow deposition was downscaled (box #1) using 1) terrain-based MicroMet lapse rates (MM), 2) a snow deposition downscaling factor ($V_{x,y}$) from April 7, 2014 (V14), and 3) $V_{x,y}$ from airborne lidar scanning (ALS) snow depth on April 16, 2016 (V16). An ensemble of simulations was performed 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 April 7, 2014 (box #3, black line) target ALS observation (box #3, scatter point) was selected for each model gridcell.

terrain-based), domain mean snowfall (μ^t_p) was calculated from the snowfall downscaled using Equation 3 over each domain and time step. Pattern-based snow deposition downscaling (Equation 1) was then calculated from the product of μ^t_p and snow deposition downscaling factors ($V_{x,y}$) calculated from ALS snow depth patterns from near peak snowpack timing in seven 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 rainfall in the same way. Following methods from Vögeli et al. (2016), rainfall was distributed using only terrain-based downscaling (Equation 3) in all simulations.

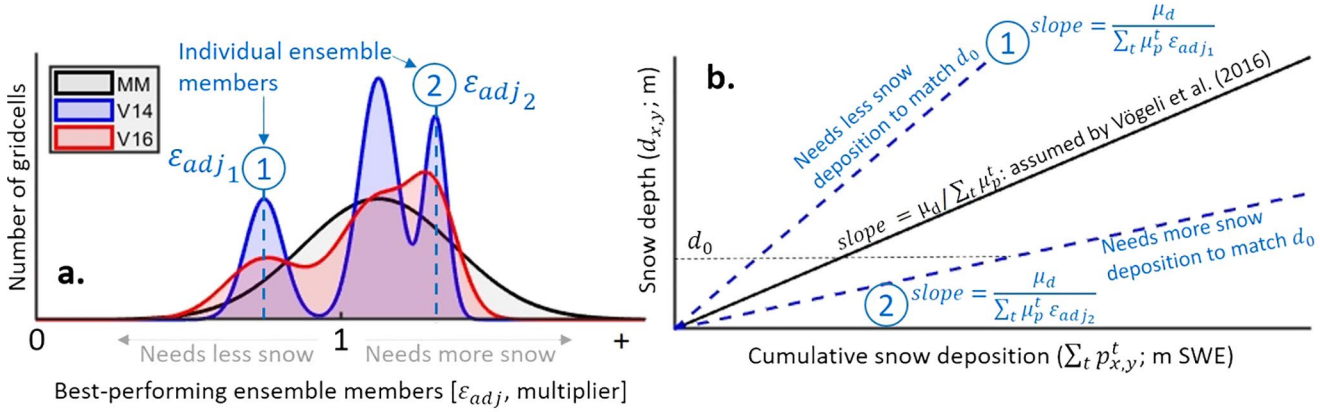

Figure 3. Conceptual diagram (a) showing the distribution of the best-performing ensemble members (ε_{adj}) for simulations with terrain-based downscaling (MM), and pattern-based downscaling (V14 and V16). The linear relationship between cumulative snow deposition and snow depth assumed by the pattern-based downscaling (b, solid line) is compared versus the linear relationships for two different ε_{adj} (b, dashed lines).

Constant snow deposition multipliers (ε; Figure 2) (constant in space and time) were used to generate an ensemble of simulations for each downscaling. Henn, Clark, et al. (2018) 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 modeling domains here (Figure 1), 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, Clark, et al., 2018), 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 snow deposition downscaling (Equation 3) was adjusted,

$$P_{x,y}^t = p_{x,y}^t \times \varepsilon, \tag{4}$$

and used to generate an ensemble of simulations (Figure 2, MM). The simulations with pattern-based snow deposition downscaling (Equation 1) were also adjusted by ε,

$$P_{x,y}^t = V_{x,y} \times \mu_p^t \times \varepsilon. \tag{5}$$

4.2. 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 2, box #3, red line) using (a) terrain-based downscaling (MM) and (b) pattern-based downscaling with a snow deposition downscaling factor ($V_{x,y}$) defined from the ALS observation nearest peak-snowpack from the same year as the simulation (V14; April 7, 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 April 7, 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 April 7, 2014 target ALS observations on a gridcell-by-gridcell basis (ε_{adj}, constant in time, variable in space) (Figure 2, box #3, black line). Since both MM and V14 simulations were provided identical domain mean snowfall (μ_p^t), identical model forcing, and the same snow model, differences in the distributions of ε_{adj} between the different downscaling (e.g., Figure 3a) were driven only by differences in snow deposition. Additionally, since the V14 simulation used the same ALS observation (from April 7, 2014) to define the snow deposition downscaling factor ($V_{x,y}$) and pick the best-performing

ensemble members (ε_{adj}), model errors were 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 (a) spatial differences in $V_{x,y}$ driven by processes other than snow accumulation (like snowmelt, snow sublimation, and snow density spatial variability), and (b) biases in μ_p^t. V14 ε_{adj} that were used more-frequently (e.g., Figure 3a, ε_{adj1} and ε_{adj2}) identified portions of the modeling domain that exhibited similar relationships between cumulative snow deposition and snow depth (Figure 3b, blue dashed lines). Here, we evaluated the V14 most frequently used ε_{adj} and their connections with winter snow losses (snowmelt and snow sublimation), connections with snow density spatial variability, and the degree to which the relationships between cumulative snow deposition and snow depth differed from the linear relationship (discussed in Section 2.1) assumed by the pattern-based downscaling method (Figure 3b, solid line).

For the small portion (<4%) of each domain with snow-absence on April 7, 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 April 7, 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 April 7, 2014 (coefficient of correlation, median bias, mean absolute error, and CoV error), more critical.

4.3. 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 4.2 were repeated using the snow deposition downscaling factors ($V_{x,y}$) defined from ALS observations near peak-snowpack timing in other years (2013, and 2015–2019). We first focused on pattern-based snow deposition downscaling using $V_{x,y}$ from April 16, 2016 (V16, Figure 3a), 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.

4.4. 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 of domain mean snowfall biases on simulations with terrain-based downscaling, and simulations with pattern-based downscaling using $V_{x,y}$ from different years. In other words, since all downscaling methods were provided identical μ_p^t (Section 4.1), 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 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, Newman, et al., 2018; Hughes et al., 2017; Lundquist et al., 2015), ranging between ±60%.

4.5. 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 4.2–4.4, 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 vs. terrain-based) and snow deposition downscaling factors ($V_{x,y}$) from different years. These results were used to disentangle the combined errors

Table 1
Statistics for Snow Depth Simulated on April 7, 2014 (at noon) Using Terrain-Based (MM) and Pattern-Based Snow Deposition Downscaling Factors ($V_{x,y}$) From April 7, 2014 (V14) and April 16, 2016 (V16)

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

Note. Statistics are shown for the default ensemble member ($\varepsilon = 1.0$) and for the best-performing ensemble members (ε_{adj}) in the Stubblefield, Matterhorn, and Lyell domains (S, M, and L, respectively).

from different snow deposition downscaling (pattern-based vs. terrain-based), snowfall biases, and interannual changes to snow depth patterns.

5. Results

5.1. Comparing Pattern-Based and Terrain-Based Snow Deposition Downscaling

April 7, 2014 snow depth simulated using pattern-based and terrain-based downscaling (default ensemble members, $\varepsilon = 1.00$ at all gridcells) 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 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 4). 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 but had a April 7, 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 CoV that was high-biased by greater than 20% in each domain (Table 1). The causes for this are detailed in the remainder of this section.

The best-performing ensemble members (ε_{adj}) for the simulations with terrain-based (MM) and pattern-based (V14) downscaling reproduced the April 7, 2014 target snow depth observations (Table 1) using very different adjustments to snow deposition (Figures 5a and 5b). The MM ensemble members that best-reproduced the April 7, 2014 snow depth (ε_{adj}) were widely distributed. In fact, no single ε_{adj} (given the set of ε used in this study) occurred across more than 5% of any domain (Figure 5a). This showed that the spatial heterogeneity of snow accumulation was misrepresented by the simulation with terrain-based downscaling, requiring a large range of snow deposition adjustments to match ALS observed snow depth on April 7, 2014. This differed from the V14 best-performing ensembles, which matched April 7, 2014 snow depth using unique ensemble members much more frequently (Figure 5b). 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 downscaling). In other words, large portions of each domain had snow deposition that was able to be corrected with a smaller number of snow deposition multipliers, each corresponding to unique issues with the V14 pattern-based snow deposition downscaling. Since the V14 simulation used the same ALS observation (from April 7, 2014) to define the snow deposition downscaling factor ($V_{x,y}$) and pick the best-performing ensemble members,

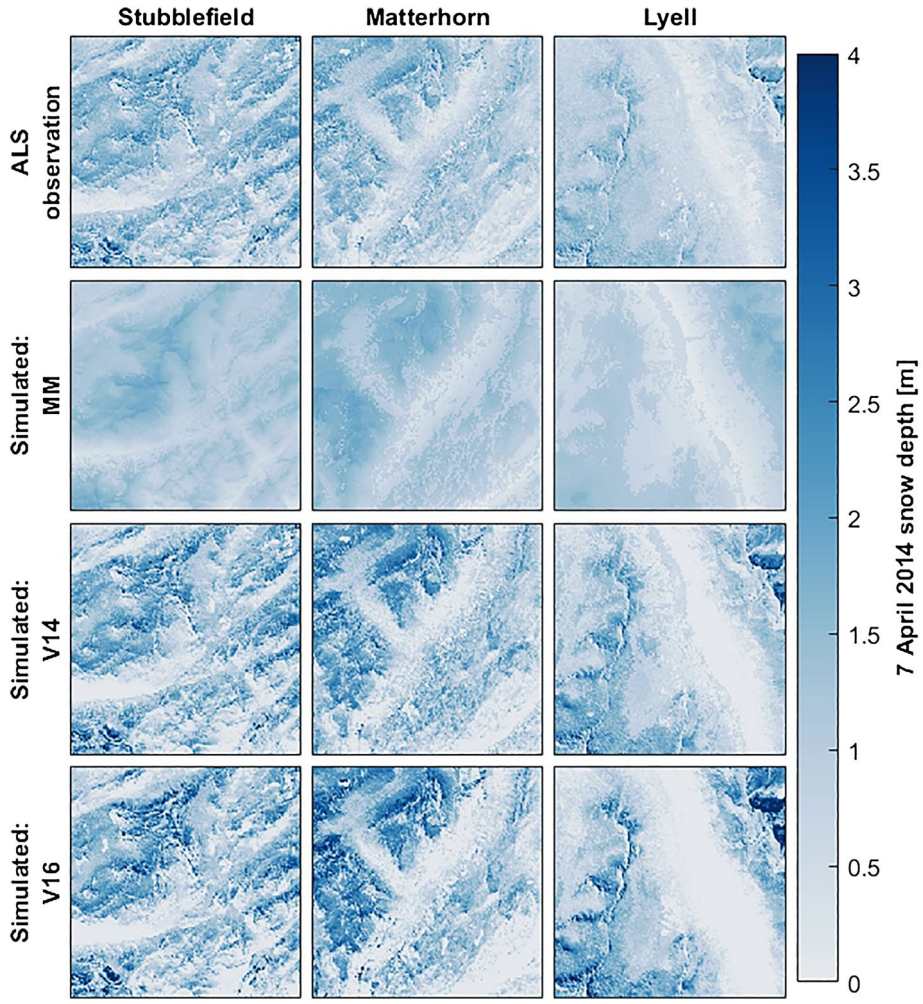

Figure 4. April 7, 2014 snow depth (colorbar) for each domain (columns) observed by airborne lidar scanning (ALS) (top) and simulated using snow deposition ($\varepsilon = 1.0$) downscaled using terrain-based downscaling (MM, second row) and pattern-based snow deposition downscaling factors ($V_{x,y}$) from April 7, 2014 (V14, third row) and April 16, 2016 (V16, fourth row).

V14 model errors were attributable to (a) errors with $V_{x,y}$, and (b) biases in domain mean snowfall (μ_p^t) (Section 4.2). In the remainder of this section, we focus on the sources of $V_{x,y}$ errors. μ_p^t biases are discussed in Section 5.3.

Pattern-based snow deposition downscaling assumes that snow depth and the resulting snow deposition downscaling factor ($V_{x,y}$) is influenced only by spatial differences in 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 and snow density (and thus equal $V_{x,y}$ and SWE), a gridcell that had snowmelt and/or snow sublimation would require more snow deposition to achieve that snow depth than a gridcell with no 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 had spatially variable snow density (Figures 6d–6f) and winter snowmelt and sublimation (snow losses) prior to April 7, 2014 (Figures 6j–6l). This resulted in a gridcell-by-gridcell relationship between cumulative snow deposition and April 7, 2014 snow depth that differed from the linear relationship assumed by pattern-based snow deposition downscaling (Figure 6, solid lines), but conformed more with several different linear relationships. These linear relationships corresponded well with the peak ε_{adj} frequencies from Figure 5b (e.g., conceptual diagram in Figure 3).

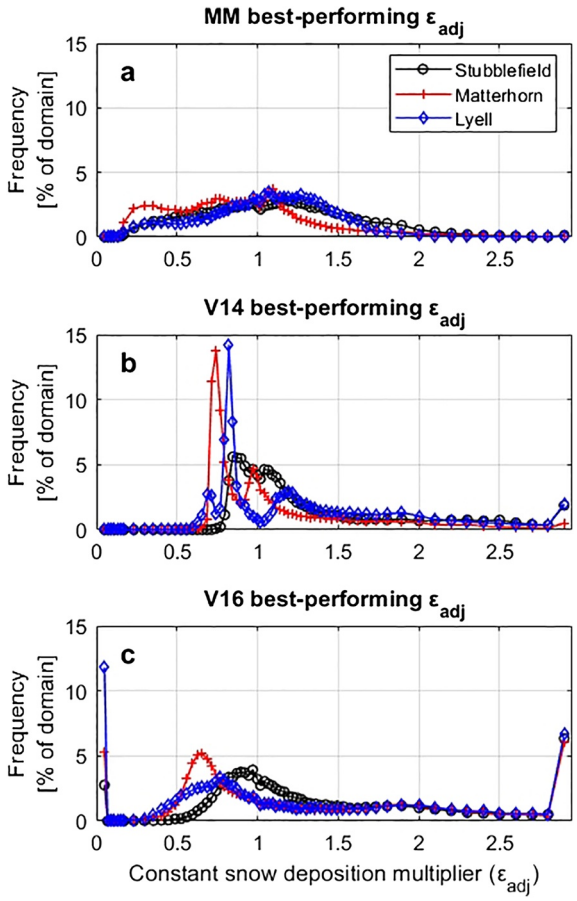

Figure 5. Frequency (percent of each domain's gridcells) of the ensemble-members (ε_{adj}) best-reproducing the April 7, 2014 target airborne lidar scanning (ALS) observation in each domain (line color and marker). Results are for the terrain-based snow deposition downscaling (MM, (a), and pattern-based downscaling using the snow deposition downscaling factor ($V_{x,y}$) from April 7, 2014 (V14, b), and $V_{x,y}$ from April 16, 2016 (V16, c). The spatial map of ε_{adj} are shown in Figure S4.

For example, winter snow losses occurred prior to April 7, 2014 in all the three domains (Figures 6j–6l). In these gridcells, snow losses reduced snow depth observed by ALS, reduced $V_{x,y}$, reduced snow deposition, and resulted in shallower simulated snow depths that melted more readily. Put another way, since April 7, 2014 $V_{x,y}$ was influenced by winter snowmelt and snow sublimation, the V14 simulation double-counted winter snow losses in both (a) $V_{x,y}$, and (b) the snow simulation. At gridcells with larger winter snow losses (Figures 6m–6o, red-shaded), the relationship between cumulative snow deposition and April 7, 2014 snow depth was approximated best (Figures 6g–6i) by the rightmost peak in ε_{adj} frequency (Figure 5b) 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, gridcells with no winter snow losses appeared deeper, relative to μ_d, than they should have been. This resulted in high-biased $V_{x,y}$ and snow deposition. The relationship between cumulative snow deposition and April 7, 2014 snow depth in these gridcells was approximated best using snow deposition multipliers of 0.85, 0.74, and 0.82 in the Stubblefield, Matterhorn, and Lyell domains, respectively (Figures 6g–6i). These snow deposition multipliers corresponded well with peaks in ε_{adj} frequency from Figure 5b. 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 (Figures 6m–6o, blue shading) that had no winter snow losses, less snow settling, less-dense snow, and thereby required less snow deposition to achieve snow depths identical to gridcells in more-exposed locations (Figures 6a–6c, 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 April 7, 2014 snow depth observations best using the smallest ε_{adj} from Figure 5b.

The combined impact of winter snow losses (snowmelt and snow sublimation) and spatially variable snow density on snow depth resulted in $V_{x,y}$ and simulated snow depth that was too spatially heterogeneous (Figure 4). Here, using the best-performing V14 ensemble members (Figure 6), we also calculated $V_{x,y}$ using simulated April 7, 2014 SWE ($V_{x,y}$ = $SWE_{x,y}$ / μ_{SWE}). In doing so, $V_{x,y}$ was no longer influenced by snow density spatial variability. As compared to $V_{x,y}$ calculated using April 7, 2014 snow depth, $V_{x,y}$ calculated from April 7, 2014 SWE had a CoV that was only ~4% smaller (Figure 7). However, we also calculated $V_{x,y}$ using the cumulative snow deposition of the best-performing V14 ensemble members ($V_{x,y} = \sum_t p^t_{x,y} / \mu_{\sum p}$), which was not influenced by either snow density spatial variability or winter snow losses. $V_{x,y}$ calculated using April 7, 2014 cumulative snow deposition had a CoV 31%–40% smaller than $V_{x,y}$ calculated from April 7, 2014 snow depth (Figure 7). Therefore, for pattern-based downscaling, the overaccumulation and underaccumulation of snow in deep and shallow snow regions (Figure 4 and Table 1) was driven much more (~8 times more) by winter snow losses than snow density spatial variability.

5.2. Identifying Errors Driven by Interannual Changes in Snow Patterns

The snow deposition downscaling factors ($V_{x,y}$, calculated using ALS snow depth observations) from April 7, 2014 and April 16, 2016 were remarkably similar ($r = 0.89$). In fact, it was difficult to visually distinguish

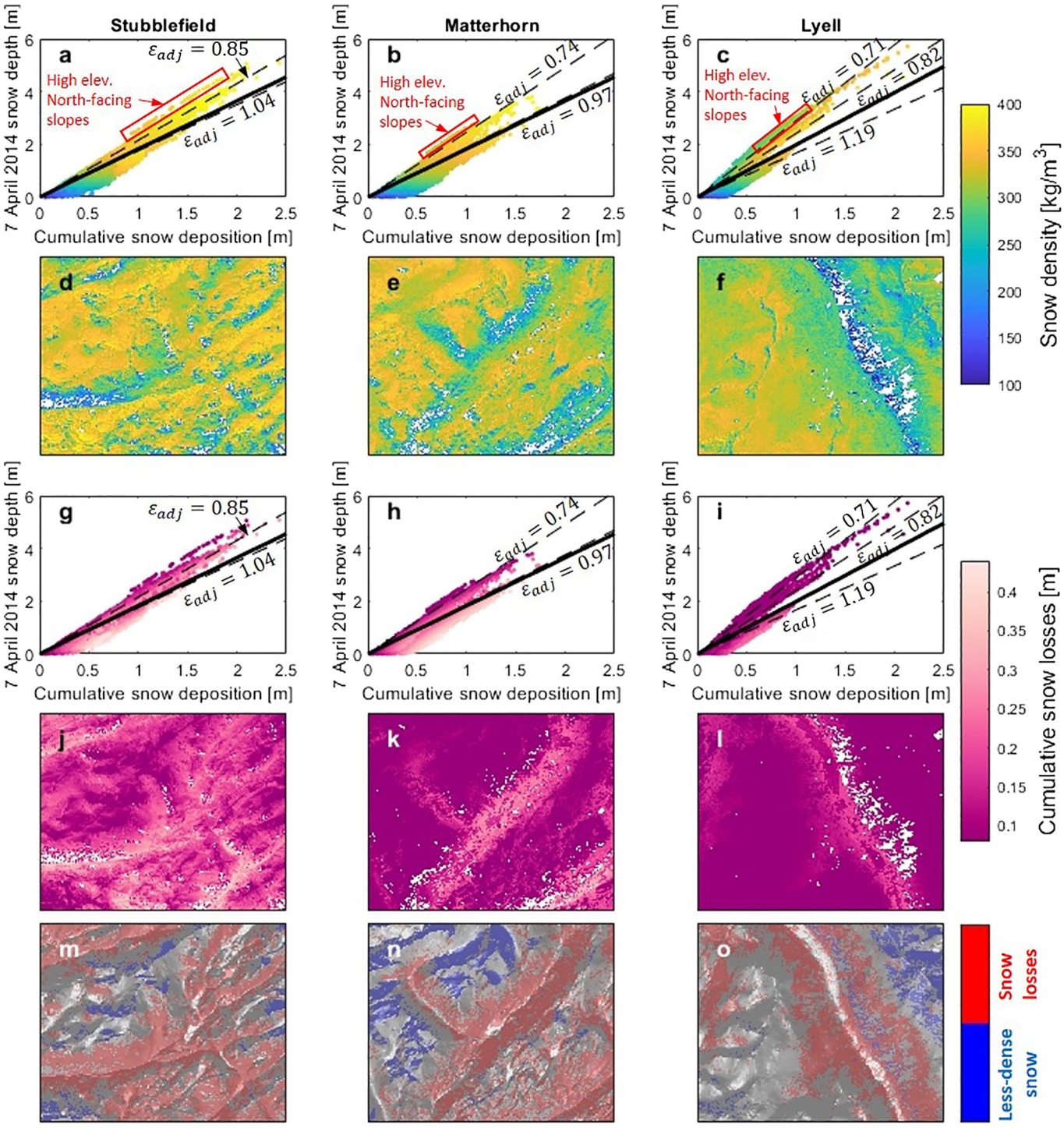

Figure 6. April 7, 2014 snow density (a–f) and winter snow losses (g–l) for the best-performing V14 ensemble members in each domain (columns). The scatter plots in subplots (a–c) and (g–i) are identical but colored in accordance with the spatial data in the subplots below them. The 25 m gridcells that diverged the most from the pattern-based snow deposition downscaling assumptions were colored over a hillshaded map of each domain (m–o).

differences in $V_{x,y}$ between the two dates (Figure 8). Although mean snow depth on April 16, 2016 was ~42% greater on average, the ALS observations were both within a week of peak-snowpack timing in their respective seasons, snow extents agreed to within 8%, and snow depth coefficients of variation agreed to

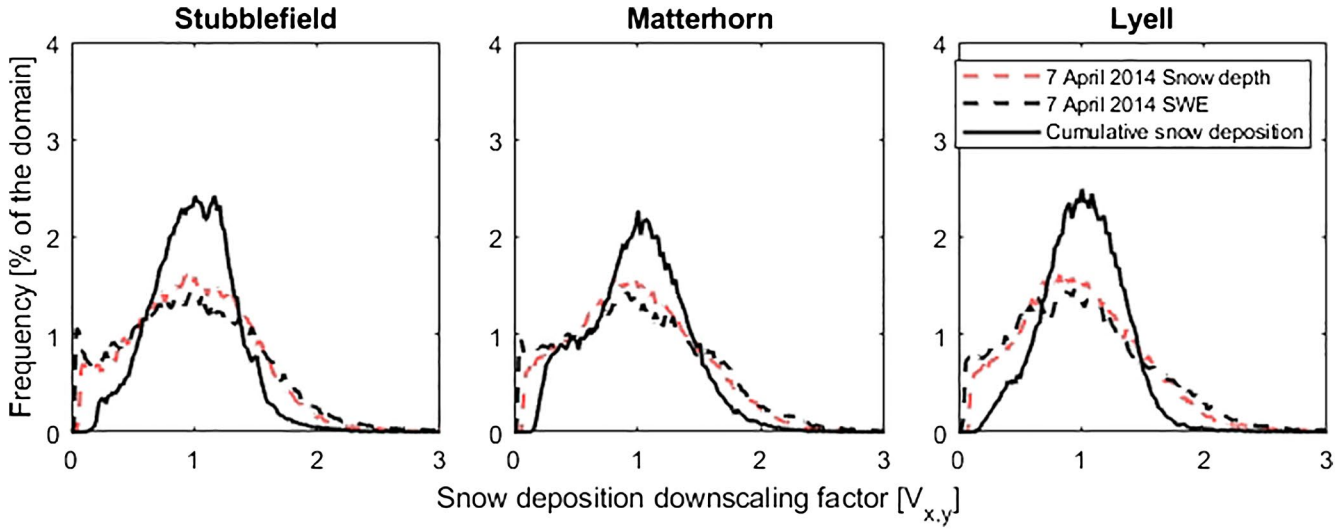

Figure 7. The distribution of snow deposition downscaling factors ($V_{x,y}$) calculated using April 7, 2014 snow depth (red, dashed), April 7, 2014 snow water equivalent (SWE) (black, dashed), and cumulative snow deposition (cumulative increases in SWE) prior to April 7, 2014 (black, solid). Distributions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

within 5%. Overall, the difference between April 7, 2014 and April 16, 2016 $V_{x,y}$ was not well-explained 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 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 ($\varepsilon = 1.00$) using pattern-based downscaling and $V_{x,y}$ from April 16, 2016 (V16) had a April 7, 2014 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 5c) used ε_{adj} from the edges of the ensemble-space ($\varepsilon_{adj} = 0.05$ or $\varepsilon_{adj} = 2.90$), showing that the lidar snow depth observations were sometimes outside of the snow depth 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 April 7, 2014 snow depth well (Table 1). In addition to the errors with pattern-based downscaling driven by snow density spatial variability and winter snow losses (Section 5.1), the best-performing V16 ensemble members also had to correct for the spatial differences in snow patterns between April 16, 2016 and April 7, 2014. As a result, the V16 ensemble required a wider range ε_{adj} to approach the April 7, 2014 ALS observation, smoothing the unique spikes in ε_{adj} frequency seen for the V14 ensemble (Figure 5).

Snow deposition downscaling factors ($V_{x,y}$) varied most from $V_{x,y}$ on April 7, 2014 in seasons with snowfall magnitude and snow extents unlike 2014 (Pflug & 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, Cortés, Girotto, Huning, et al., 2016). As a result, larger portions of the model domains were snow-free during the March 5, 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 9, $\varepsilon = 1.0$). The same was true for the simulations with pattern-based snow deposition downscaling using $V_{x,y}$ from April 3, 2013. Conversely, $V_{x,y}$ calculated from ALS observations of deep snowpacks on April 1, 2017 and March 24, 2019 had less snow depth variability relative to domain mean snow depth (Figure 8), and thereby simulated snow depth that was more spatially homogeneous than pattern-based downscaling using other $V_{x,y}$. The sources of interannual snow depth pattern differences in this region are discussed more in Pflug and Lundquist (2020).

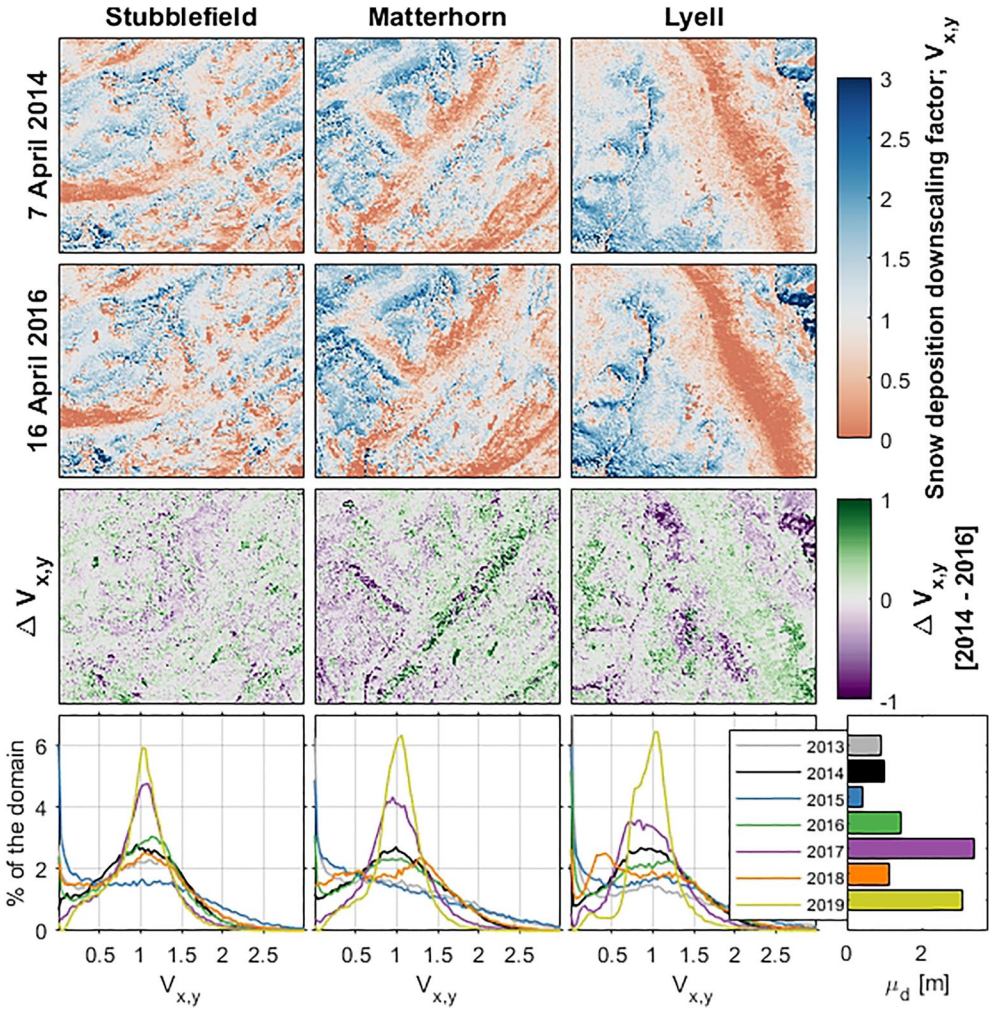

Figure 8. Snow deposition downscaling factors ($V_{x,y}$) calculated from airborne lidar scanning (ALS) observations on April 7, 2014 (top row) and April 16, 2016 (second row) in the three domains (columns). Spatial differences in $V_{x,y}$ (April 7, 2014–April 16, 2016) are shown in the third row. The distribution of $V_{x,y}$, and the mean snow depth (μ_d, including all three domains) are also compared for ALS observations nearest peak-snowpack timing for 7 years (bottom row).

5.3. Identifying Errors Driven by Coarse-Scale Snowfall Biases

Cumulative WRF snowfall prior to April 7, 2014 was biased by only small amounts, matching total snow volume on April 7, 2014 closely (Section 5.1). However, we know that 6 km snowfall from atmospheric models can be biased in mountainous regions. In Figure 9, 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 ±60% error in mean snowfall, consistent with previous studies in this region (Henn, Newman, et al., 2018; Hughes et al., 2017; Lundquist et al., 2015). April 7, 2014 median snow depth biases typically increased linearly with ε for all terrain-based and pattern-based downscaling methods. For $\varepsilon \pm 10\%$ Figure 9, 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. Therefore, even provided a ±10% bias in domain mean snowfall, all simulations had an IQR that included snow depth biases in the opposite direction (positive/negative) of the snowfall bias. However, for ±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 uncertainties with both the pattern-based and terrain-based

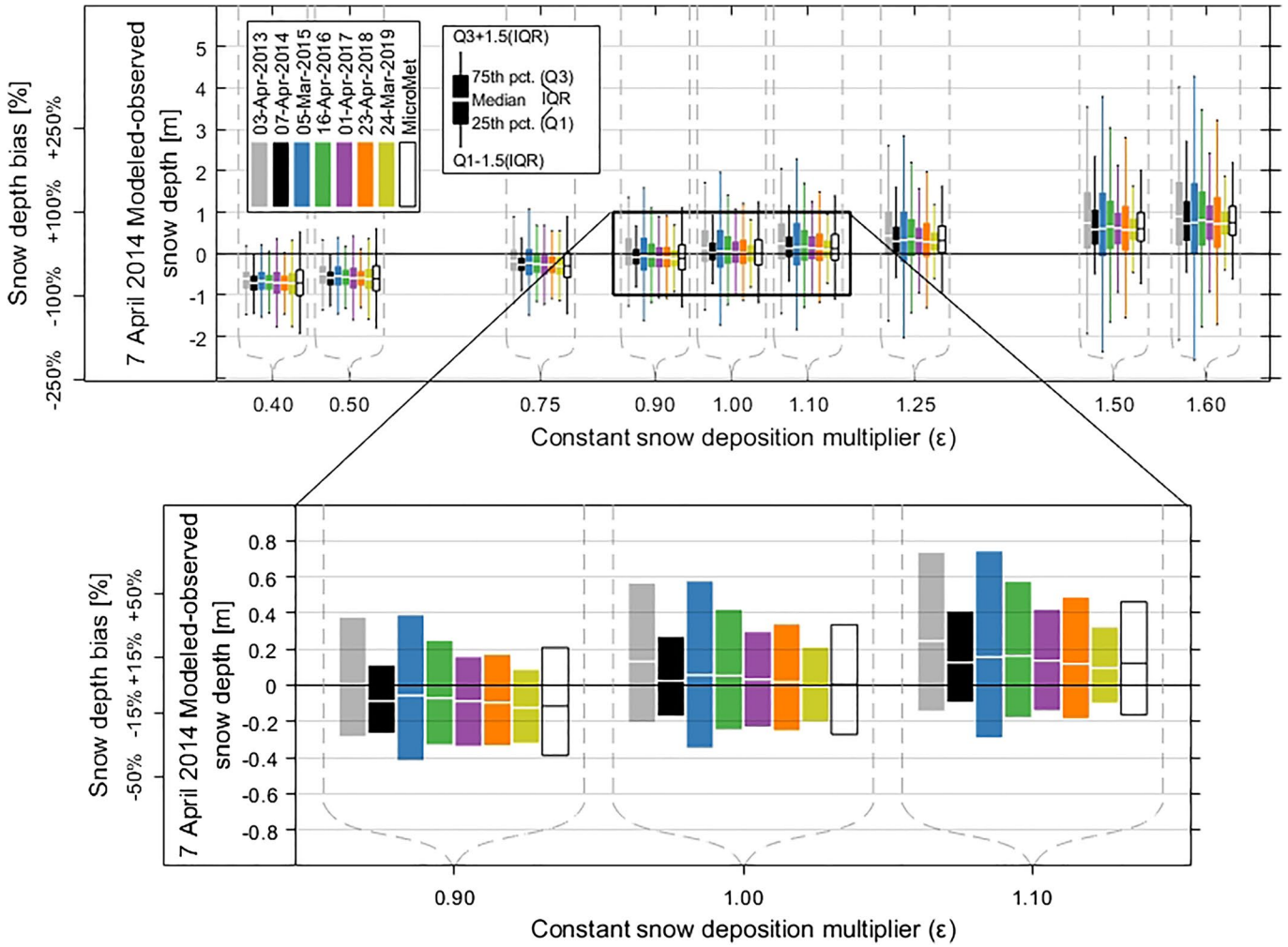

Figure 9. Snow depth simulation errors including all domains on April 7, 2014. Snow deposition was downscaled using terrain-based MicroMet lapse rates (white with black outline) and pattern-based downscaling using $V_{x,y}$ from airborne lidar scanning (ALS) observations in different years (colors) across a range of constant (in space and time) snowfall biases (*x*-axis).

downscaling were comparable in size to snow depth errors driven by snowfall biases. The combined effect of $V_{x,y}$ errors and snowfall biases are discussed in Section 5.4.

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 (April 25/26, 2014), causing snow depth volume biases upwards of 30% through much of the snowmelt season (Figure 10). Even though the spatial heterogeneity of snow influences the rate and timing of spring snowmelt (e.g., Freudiger et al., 2017; Luce et al., 1998; Pflug & Lundquist, 2020), this spring snowfall bias caused model departures from observed snow volume that were larger than the late-spring differences 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 10, these errors could affect snowmelt modeling and forecasting, even in cases where peak snowpack volume and spatial distribution is modeled well, and spring snowmelt energy is accurate.

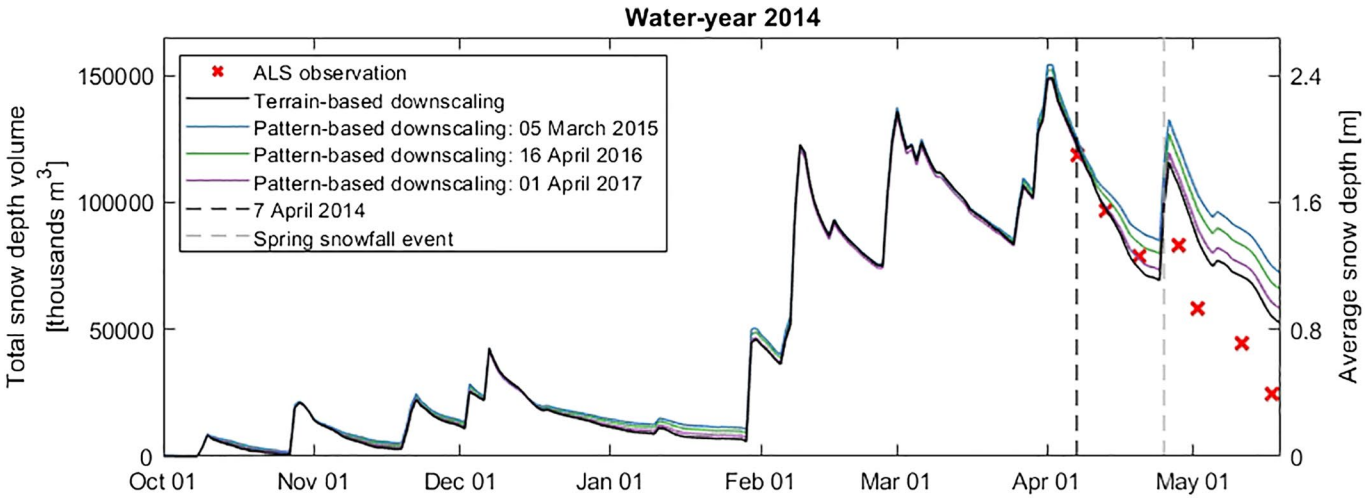

Figure 10. Total (summed) snow depth volume (left *y*-axis) and average snow depth (right *y*-axis) including the Stubblefield, Matterhorn, and Lyell domains. Snow depth was simulated using terrain-based snow deposition downscaling (black line) and pattern-based downscaling using snow deposition downscaling factors from peak-snowpack in 2015, 2016, and 2017 (colored lines) (default ensemble members, $\varepsilon = 1.0$). Simulated snow volume was compared versus airborne lidar observations (markers). April 7, 2014 and a spring snowfall event are also marked.

5.4. 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 the accuracy of 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 $V_{x,y}$ and ε at each gridcell (Equation 5). Given this snowfall bias ($\varepsilon = 1.50$), $V_{x,y}$ from a year with a too heterogeneous snow depth pattern (e.g., 2015) would multiplicatively increase the heterogeneity of simulated snow depth (Figure 9). However, two-way analysis of variance (ANOVA) determined that the IQR of snow depth errors were significantly different ($p < 0.01$) across simulations with different mean snowfall biases (ε, ranging between 0.40 and 1.60), but not for simulations using $V_{x,y}$ from different years ($p > 0.30$). Therefore, in these domains in water-year 2014, the magnitude and spread of modeled snow depth errors were dominated by common snowfall biases, even when assuming $V_{x,y}$ used to downscale snow deposition was only available from years with less-similar 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 8) 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) adjusted snowfall from WRF by no more than ±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 April 7, 2014 (0.69 m). Additionally, there was little difference in the IQR of snow depth errors across the ±60% range of mean snowfall biases (ε) tested in Figure 9. Put another way, we would expect a similar IQR of model errors on April 7, 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 10). Interestingly,

the simulation with pattern-based downscaling using $V_{x,y}$ from April 1, 2017, which was more spatially homogeneous (CoV 14% smaller than $V_{x,y}$ calculated from the April 7, 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 April 1, 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 8) that agreed closely with melt- and density-corrected $V_{x,y}$ on April 7, 2014 (Figure 7, solid black line). This suggests that interannual changes in winter snowmelt, snow sublimation, and snow density were major sources of interannual differences in $V_{x,y}$.

6. 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 snow depth observations 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 Figures 6m–6o) 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 & 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 (≥800 m). These lidar snow depth observations also have great accuracy (±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 & Lundquist, 2020; Sturm & Wagner, 2010; Woodruff & 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, Cortés, Girotto, & Durand et al., 2016; Margulis et al., 2019; 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 & Lundquist, 2020) may be also able to bias-correct coarser-scale estimates of snowfall from atmospheric models like WRF.

7. Conclusion

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 (Pflug & 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 seven 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 were most influenced by mean snowfall estimates in a 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 (±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 (a) simulations with different downscaling, and (b) simulations downscaling snow deposition using snow depth patterns more- and less-similar to snow depth patterns from the simulation year (Figure 10).

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 eight times smaller than winter snowmelt and snow sublimation (Figure 7). 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 (a) coarse-scale snowfall biases from atmospheric models, (b) snow depth patterns that were influenced by processes other than snowfall (like snowmelt, snow sublimation, and snow density), and (c) 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.

Data Availability Statement

The 25 m regridded lidar, 6 km WRF data, and model outputs from this study are archived in the Hydroshare public data repository at https://www.hydroshare.org/resource/0b41fc78e4f044b096b67f3197b8640c/

Acknowledgments
The authors gratefully acknowledge funding support from NASA grants NNX17AL59G, 80NSSC18K1326, and 80NSSC20K1293, and the Airborne Snow Observatory team for processing and sharing large lidar datasets.

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. *Journal of Hydrometeorology*, *18*, 49–63. https://doi.org/10.1175/JHM-D-16-0026.1

Barnes, S. L. (1964). A technique for maximizing details in numerical weather map analysis. *Journal of Applied Meteorology*, *3*, 396–409. https://doi.org/10.1175/1520-0450(1964)003<0396:ATFMDI>2.0.CO;2

Barrett, A. P. (2003). *National Operational Hydrologic Remote Sensing Center Snow Data Assimilation System (SNODAS) products at NSIDC*. Boulder, CO. National Snow and Ice Data Center, Cooperative Institute for Research in Environmental Sciences.

Bavay, M., & Egger, T. (2014). MeteoIO 2.4. 2: A pre-processing library for meteorological data. *Geoscientific Model Development*, *7*, 3135–3151. https://doi.org/10.5194/gmd-7-3135-2014

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

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

Clark, M. P., Hendrikx, J., Slater, A. G., Kavetski, D., Anderson, B., Cullen, N. J., et al. (2011). Representing spatial variability of snow water equivalent in hydrologic and land-surface models: A review. *Water Resources Research*, *47*, W07539. https://doi.org/10.1029/2011WR010745

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., 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. https://doi.org/10.1175/jhm-d-17-0026.1

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

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., et al. (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

Deems, J. S., Fassnacht, S. R., & Elder, K. J. (2008). Interannual Consistency in Fractal Snow Depth Patterns at Two Colorado Mountain Sites. *Journal of Hydrometeorology*, *9*, 977–988. https://doi.org/10.1175/2008JHM901.1

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. *Hydrological Processes*, *26*, 1574–1585. https://doi.org/10.1002/hyp.8244

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

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, E. D. P. Y., & Legates, D. R. (1994). The accuracy of United States precipitation data. *Bulletin of the American Meteorological Society*, *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., et al. (2012). A comparison of statistical and dynamical downscaling of winter precipitation over complex terrain. *Journal of 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

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

Henn, B., Clark, M. P., Kavetski, D., McGurk, B., Painter, T. H., & Lundquist, J. D. (2016). Combining snow, streamflow, and precipitation gauge observations to infer basin-mean precipitation. *Water Resources Research*, *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. (2018). 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. (2018). 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. *Climate Dynamics*, *55*, 111–129. 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

Kim, E., Gatebe, C., Hall, D., Newlin, J., Misakonis, A., Elder, K., et al. (2017). NASA's snowex campaign: Observing seasonal snow in a forested environment. In:*IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Presented at the 2017 IEEE International Geoscience and Remote Sensing Symposium* (pp. 1388–1390). IGARSS. https://doi.org/10.1109/IGARSS.2017.8127222

Kim, R. S., Kumar, S., Vuyovich, C., Houser, P., Lundquist, J., Mudryk, L., et al. (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

Koch, S. E., desJardins, M., & Kocin, P. J. (1983). An interactive barnes objective map analysis scheme for use with satellite and conventional data. *Journal of Climate and Applied Meteorology*, *22*, 1487–1503. https://doi.org/10.1175/1520-0450(1983)022<1487:AIBOMA>2.0.CO;2

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

Liston, G. E. (2004). Representing subgrid snow cover heterogeneities in regional and global models. *Journal of 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. https://doi.org/10.1175/jhm486.1

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

Liston, G. E., Haehnel, R., Sturm, M., A Hiemstra, C., Stuefer, S., & Tabler, D. 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

Liu, Y., & Margulis, S. A. (2019). Deriving Bias and Uncertainty in MERRA-2 Snowfall Precipitation Over High Mountain Asia. *Frontiers of Earth Science*, *7*. https://doi.org/10.3389/feart.2019.00280

Livneh, B., Bohn, T. J., Pierce, D. W., Munoz-Arriola, F., Nijssen, B., Vose, R., et al. (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. 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*. https://doi.org/10.1029/2004wr003933

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. *Journal of Hydrometeorology*, *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

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. *Bulletin of the American Meteorological Society*, *100*, 2473–2490. https://doi.org/10.1175/BAMS-D-19-0001.1

Margulis, S. A., Cortés, G., Girotto, M., & Durand, M. (2016). A Landsat-Era Sierra Nevada Snow Reanalysis (1985–2015). *Journal of Hydrometeorology*, *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. (2016). 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., Liu, Y., & Baldo, E. (2019). A Joint Landsat- and MODIS-Based Reanalysis Approach for Midlatitude Montane Seasonal Snow Characterization. *Frontiers of Earth Science*, *7*. https://doi.org/10.3389/feart.2019.00272

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

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

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., et al. (2006). North American Regional Reanalysis. *Bulletin of the American Meteorological Society*, *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., Mote, P. W., & Lundquist, J. D. (2010). Surface temperature lapse rates over complex terrain: Lessons from the Cascade Mountains. *Journal of Geophysical Research*, *115*. https://doi.org/10.1029/2009JD013493

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

Mote, P. W., Parson, E. A., Hamlet, A. F., Keeton, W. S., Lettenmaier, D. P., Mantua, N., et al. (2003). Preparing for climate change: The water, salmon, and forests of the Pacific Northwest. *Climatic Change*, *61*, 45–88. https://doi.org/10.1023/a:1026302914358

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

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 Resources Research*, *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., et al. (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

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. *Hydrology and Earth System Sciences*, *19*, 3153–3179. https://doi.org/10.5194/hess-19-3153-2015

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

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. *Journal of the Atmospheric Sciences*, *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*. https://doi.org/10.1029/2010wr009426

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., et al. (2013). Snowmelt contributions to discharge of the Ganges. In:*The 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* (pp. 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.

Sproles, E. A., Rittger, K., Nolin, A. W., & Painter, T. H. (2013). Climate change impacts on maritime mountain snowpack in the Oregon Cascades. *Hydrology and Earth System Sciences*, *17*, 2581–2597. https://doi.org/10.5194/hess-17-2581-2013

Strachan, S., Kelsey, E. P., Brown, R. F., Dascalu, S., Harris, F., Kent, G., et al. (2016). Filling the data gaps in mountain climate observatories through advanced technology, refined instrument siting, and a focus on gradients. *Mountain research and development*, *36*, 518–527. https://doi.org/10.1659/MRD-JOURNAL-D-16-00028.1

Sturm, M., & Wagner, A. M. (2010). Using repeated patterns in snow distribution modeling: An Arctic example. *Water Resources Research*, *46*, W12549. https://doi.org/10.1029/2010WR009434

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. https://doi.org/10.1175/bams-d-11-00094.1

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. *Monthly Weather Review*, *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. *Monthly Weather Review*, *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

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

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. *The Cryosphere*, *9* (pp. 1147–1167). 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. *Frontiers of Earth Science*, *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. *Monthly Weather Review*, *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., 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

Ye, B., Yang, D., & Ma, L. (2012). Effect of precipitation bias correction on water budget calculation in Upper Yellow River, China. *Environmental Research Letters*, *7*, 025201. https://doi.org/10.1088/1748-9326/7/2/025201