JGR Atmospheres

RESEARCH ARTICLE
10.1029/2019JD030823

Can Convection-Permitting Modeling Provide Decent Precipitation for Offline High-Resolution Snowpack Simulations Over Mountains?

Cenlin He[1,2], Fei Chen[2], Michael Barlage[2], Changhai Liu[2], Andrew Newman[2], Wenfu Tang[3], Kyoko Ikeda[2], and Roy Rasmussen[2]

[1]Advanced Study Program, National Center for Atmospheric Research, Boulder, CO, USA, [2]Research Applications Laboratory, National Center for Atmospheric Research, Boulder, CO, USA, [3]Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA

Key Points:
- Convection-permitting WRF results capture the observed precipitation, SWE, and snow depth in western and northern portions of WUS mountains
- All datasets produce large negative bias in precipitation, SWE, and snow depth in interior WUS mountains with fair performance for WRF
- Convection-permitting modeling has added values in providing precipitation for snow simulations over mountains without enough observations

Supporting Information:
- Supporting Information S1

Correspondence to:
C. He,
cenlinhe@ucar.edu

Citation:
He, C., Chen, F., Barlage, M., Liu, C., Newman, A., Tang, W., et al. (2019). Can convection-permitting modeling provide decent precipitation for offline high-resolution snowpack simulations over mountains?. *Journal of Geophysical Research: Atmospheres*, *124*, 12,631–12,654. https://doi.org/10.1029/2019JD030823

Received 14 APR 2019
Accepted 17 NOV 2019
Accepted article online 22 NOV 2019
Published online 5 DEC 2019

Abstract Accurate precipitation estimates are critical to simulating seasonal snowpack evolution. We conduct and evaluate high-resolution (4-km) snowpack simulations over the western United States (WUS) mountains in Water Year 2013 using the Noah with multi-parameterization (Noah-MP) land surface model driven by precipitation forcing from convection-permitting (4-km) Weather Research and Forecasting (WRF) modeling and four widely used high-resolution datasets that are derived from statistical interpolation based on in situ measurements. Substantial differences in the precipitation amount among these five datasets, particularly over the western and northern portions of WUS mountains, significantly affect simulated snow water equivalent (SWE) and snow depth (SD) but have relatively limited effects on snow cover fraction (SCF) and surface albedo. WRF generally captures observed precipitation patterns and results in an overall best-performed SWE and SD in the western and northern portions of WUS mountains, where the statistically interpolated datasets lead to underpredicted precipitation, SWE, and SD. Over the interior WUS mountains, all the datasets consistently underestimate precipitation, causing significant negative biases in SWE and SD, among which the results driven by the WRF precipitation show an average performance. Further analysis reveals systematic positive biases in SCF and surface albedo across the WUS mountains, with similar bias patterns and magnitudes for simulations driven by different precipitation datasets, suggesting an urgent need to improve the Noah-MP snowpack physics. This study highlights that convection-permitting modeling with proper configurations can have added values in providing decent precipitation for high-resolution snowpack simulations over the WUS mountains in a typical ENSO-neutral year, particularly over observation-scarce regions.

1. Introduction

Snowpack is a critical component of climate and weather systems, because of its high albedo, low roughness length, low thermal conductivity, and consequently its control and modulation in surface water and energy balance and feedback to the atmosphere (e.g., Barlage et al., 2010; Betts et al., 2014; Chen, Barlage, et al., 2014; Flanner et al., 2011; Minder et al., 2016, 2018). Mountain snowpack, in particular, plays an important role in hydrological applications by influencing snowmelt-induced streamflow and hence freshwater availability, irrigation for agriculture, potential energy for hydropower, river flows for fish spawning, and natural hazards such as droughts and floods (e.g., Bales et al., 2006; Barnett et al., 2005; Clark et al., 2011; Essery et al., 2009; Ullrich et al., 2018). For instance, in the western United States (WUS) mountainous areas, snowpack accounts for about 70% of the total runoff and supplies water demands for millions of people (Li et al., 2017; Mote et al., 2018). However, the WUS snowpack has experienced widespread declines in the past decades (Mote et al., 2005; Mote et al., 2018; Pederson et al., 2011) and is projected to continue decreasing under a future warming climate (e.g., Cayan et al., 2010; Rasmussen et al., 2011; Godsey et al., 2014; Li et al., 2017; Berg & Hall, 2017; Gergel et al., 2017; Musselman et al., 2017; Huning & AghaKouchak, 2018; Rhoades, Ullrich, & Zarzycki, 2018, Rhoades, Jones, & Ullrich, 2018; Marshall et al., 2019; Sun et al., 2019). Therefore, it is imperative to accurately simulate snowpack evolution in order to support weather and hydrological forecasts, climate modeling and projection, and water resource management, especially over mountainous regions.

Snowpack simulations are tightly related to both snow process modeling and atmospheric forcing (e.g., Barlage et al., 2010; Chen, Barlage, et al., 2014; He, Takano, & Liou, 2017; Henn et al., 2018; Minder et al., 2016). Precipitation, one primary driver for snowpack change and variability, has been shown to significantly contribute to the uncertainty in snow prediction and associated hydrological applications over mountains (Gutmann et al., 2012; Henn et al., 2018; Lundquist et al., 2015; Mote et al., 2005; Scalzitti et al., 2016). Moreover, it is particularly challenging to obtain accurate precipitation estimates with enough spatiotemporal coverage and resolution over data-scarce complex terrain due to complicated terrain–flow–precipitation interactions (Garvert et al., 2007; Houze, 2012; Mott et al., 2014). Previous studies further suggested that a fine horizontal resolution (e.g., less than ~10 km) is one of the key factors contributing to adequate simulations of precipitation and hence snowpack over mountains (Ikeda et al., 2010; Leung & Qian, 2003; Rasmussen et al., 2011).

Enormous efforts have been made to obtain accurate and high-resolution precipitation estimates across the continental United States. Direct ground-based measurements are usually considered to be the most accurate but limited in spatial distribution. Particularly in the WUS mountainous areas, in situ (gauge) measurements are far too sparse to represent the strong spatial variability of precipitation (Jing et al., 2017; Livneh et al., 2015) and may suffer from undercatch or other sampling errors (Bales et al., 2006; Clark & Slater, 2006; Yang et al., 1998), while ground-based radar measurements show limited effectiveness due to terrain blocking (Henn et al., 2018; Westrick et al., 1999). Thus, it is not adequate to use the point-scale precipitation measurements for regional snowpack studies over mountains. Additionally, satellite retrievals using passive or active remote sensing techniques are able to provide gridded precipitation distributions with large spatial coverage but lack sufficient spatiotemporal accuracy and resolution particularly over complex terrain (Ebert et al., 2007; Hou et al., 2014).

One common practice in hydrological modeling is to produce high-resolution gridded precipitation estimates by using statistical interpolation techniques based on ground-based (typically gauge) measurements with topographical and microclimate adjustments over mountains (e.g., Daly et al., 2008; Livneh et al., 2015; Newman et al., 2015; Xia et al., 2012). However, these gridded precipitation products are still associated with large uncertainty over complex terrain, because of biases from statistical interpolation methods, sparse measurement networks at higher elevations, and measurements that are often cited in canopy-free locations (Daly et al., 2008; Gutmann et al., 2012; Lundquist et al., 2015). As a result, the uncertainties and differences across datasets are expected to be larger in mountains with higher elevations (Henn et al., 2018; Jing et al., 2017). More importantly, the terrain-aware statistical interpolation methods depend on current/historical precipitation patterns and mainly account for the topographical and microclimate effects, which miss other key factors and physical processes that control precipitation distributions, such as mountain waves, atmospheric dynamics interacting with cloud microphysics, and local frontal systems (Garvert et al., 2007; Gutmann et al., 2012). This could introduce artifacts in interpolation processes under current/historical climate and further prevents an accurate precipitation projection in future climate, where mountain-precipitation dynamics and interactions may change (Gutmann et al., 2012; Rasmussen et al., 2011).

Recently, significant progress has been made in high-resolution (particularly convection-permitting) modeling of precipitation over complex terrain (Ikeda et al., 2010; Liu et al., 2017; Prein et al., 2015; Rasmussen et al., 2014), with the advantages of improved topographical features, physical representation of mountain-precipitation interactions, and avoided errors from convective parameterizations. For example, convection-permitting ($\leq$4 km) model simulations with proper configurations have been shown to be able to reasonably capture the distribution and amount of seasonal precipitation, snowfall, and hence snowpack evolution over the WUS mountains (Ikeda et al., 2010; Jing et al., 2017; Wang et al., 2018). However, the model performance at high resolution also relies on the choice of physical parameterizations, including microphysics schemes (e.g., cloud nucleation and growth), macrophysics schemes (e.g., cloud cover and longwave feedbacks), convection schemes (e.g., rain and snow formation), boundary layer schemes (e.g., atmosphere-land coupling), and land surface schemes (e.g., snow–rain partitioning, snow cover fraction, snow aging, and albedo evolution). Liu et al. (2011) found that cloud microphysics schemes have important effects on high-resolution precipitation modeling over the WUS mountains. A review paper by Lundquist et al. (2019) suggested that well-configured high-resolution atmospheric models can simulate total annual precipitation and snowfall better than gridded datasets derived from gauge measurements over

mountains, whereas model estimates of precipitation can still vary significantly due to different configurations, which require optimization and improvement by using a wide range of observations. Similarly, recent studies (O'Brien et al., 2016; Rhoades, Ullrich, Zarzycki, Johansen, et al., 2018) also indicated that increasing spatial resolution of global climate models may not necessarily improve mountain precipitation and hydroclimate simulations, which depends on the selection of model physical schemes.

The progress in convection-permitting modeling naturally leads to a question: can convection-permitting simulations provide viable precipitation forcing for simulating mountain snowpack? A systematic assessment of this question will further advance the understanding and quantification of the precipitation uncertainty and its impact on snowpack modeling in complex terrain. Therefore, this study seeks to (1) test the hypothesis that convection-permitting modeling with proper configurations is capable of producing adequately accurate precipitation for high-resolution snowpack simulations over mountains, which is as good as (if not better than) statistically interpolated observational datasets, (2) assess the differences across various precipitation products and associated impacts on snowpack simulations in various WUS subregions, and (3) highlight implications for future model improvements of snowpack physics in the context of precipitation uncertainty. The WUS mountains are selected for this study, because they are particularly important for water resource management and require accurate weather and hydrological forecasts to instill resilience in water conveyance, especially as the climate continues to change.

2. Method and Data

2.1. Model Descriptions and Simulations

We conduct high-resolution (4-km) snowpack simulations over the WUS (Figure 1) using a widely used land surface model (LSM), Noah with multi-parameterization (Noah-MP; Niu et al., 2011), through the high-resolution land data assimilation (HRLDAS) framework (Chen et al., 2007). The HRLDAS system was developed at the National Center for Atmospheric Research (NCAR) for the purpose of initializing land-state variables to support high-resolution LSM simulations and the coupling with WRF. Essentially, it is an uncoupled land surface modeling system, which integrates the observation/reanalysis-based atmospheric forcing data, fine-scale static surface fields (e.g., soil texture), and time-varying surface characteristics (e.g., vegetation) to drive Noah-MP LSM simulations. Thus, Noah-MP is the heart of the HRLDAS framework/infrastructure.

We drive the Noah-MP simulations using hourly atmospheric forcing data. We use five different precipitation forcing datasets (see section 2.2) to assess the precipitation uncertainty and its impact on snowpack simulations. The non-precipitation atmospheric forcing conditions (i.e., humidity, temperature, pressure, wind, downward solar, and longwave radiation) are obtained from convection-permitting (4-km) WRF model simulations (Liu et al., 2017). More details of the WRF simulations are provided in section 2.2. We perform this study for Water Year 2013 (October 2012–September 2013), an ENSO-neutral year with normal annual precipitation over the WUS (i.e., close to the regional climatological mean value; see https://www.ncdc.noaa.gov/cag/), to avoid ENSO-related uncertainty/biases from atmospheric forcing data. Our analyses particularly focus on October to June when mountain snow is present in the water year. We spin up the Noah-MP model for 5 years before official simulations (at the 4-km WRF forcing grids) and archive hourly results. The Noah-MP simulations are conducted on the NCAR's Cheyenne supercomputer using 288 CPU cores (on eight computing nodes), which totally take ~50,000 core hours and output ~4.2 TB data for all simulation cases in this study. To highlight the precipitation impact, we use the typical default Noah-MP model physics for all simulations, which has been well documented in Niu et al. (2011). The uncertainty associated with model physics is discussed in section 4 and will be investigated in detail in a follow-up study.

Here we briefly summarize the important snow-related Noah-MP model features. The partitioning of precipitation into rainfall and snowfall is a function of surface air temperature following the Jordan (1991) parameterization. Noah-MP accounts for snow interception and throughfall by vegetation canopy with distinction between water and snow/ice. The snow interception processes include loading and unloading of snowfall, frost/sublimation, melting of intercepted snow and refreezing of meltwater, and evaporation/dew (Niu & Yang, 2004). The loading rate is affected by the maximum loading capacity and snowfall rate, while the unloading rate is a function of wind speed and canopy temperature. The radiative transfer throughout canopy and intercepted snow is computed using the two-stream radiative transfer

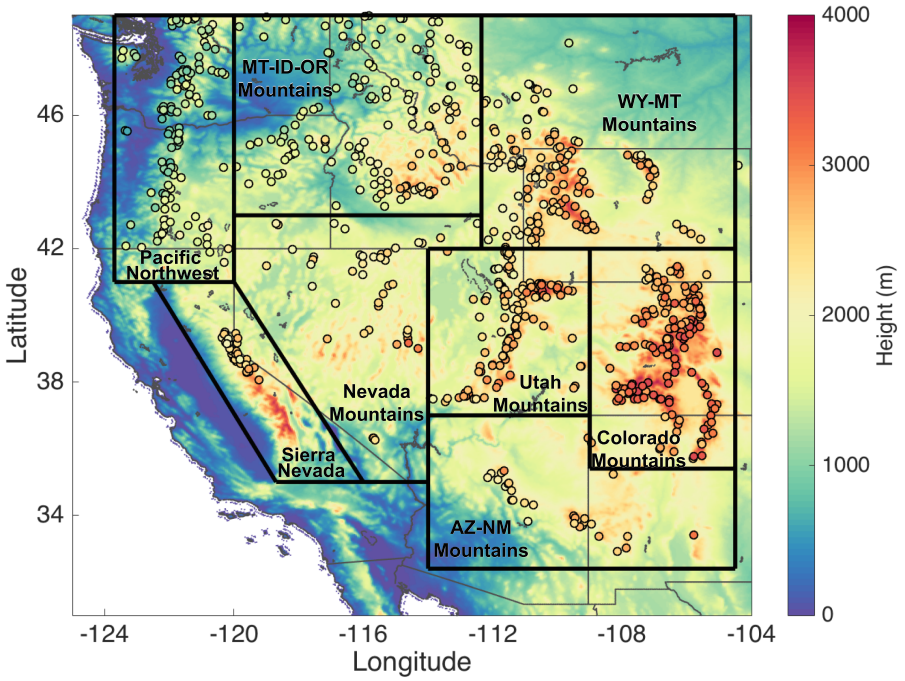

Figure 1. The topography over the western United States with eight subregions, including the Pacific Northwest mountains, Sierra Nevada mountains, Montana-Idaho-Oregon (MT-ID-OR) mountains, Wyoming-Montana (WY-MT) mountains, Nevada mountains, Utah mountains, Colorado mountains, and Arizona-New Mexico (AZ-NM) mountains. Also shown are the Snowpack Telemetry site locations (circles). Areas only include mountains (i.e., above subregional mean elevations) are shown in Figure S1.

approximation (Dickinson, 1983) based on prescribed snow and canopy optical properties. Noah-MP includes a multilayer (up to three) snowpack on the ground, depending on snow depth. The ground snow cover fraction is parameterized as a function of snow density and depth (Niu & Yang, 2007). The snow density, snow depth, snow temperature, snow ice, and water content are modeled dynamically by accounting for snow aging and melting as well as layer compaction. The snowpack albedo (for direct and diffuse radiation) is computed using the Canadian LAnd Surface Scheme (Verseghy, 1991), which accounts for fresh snow albedo and snow aging (i.e., albedo decaying with time). The ground albedo is then calculated as the area-weighted average of bare soil and snow albedos. There are eight soil types/textures with prescribed albedos (0.10–0.27 and 0.05–0.15 for dry and saturated soil visible albedos, respectively, with doubled values for near-infrared [NIR] albedos) in Noah-MP. We use the soil type with medium albedos (i.e., 0.18 and 0.09 for dry and saturated soil visible albedos, respectively) in this study. The overall surface albedo is determined by both ground albedo and canopy radiative transfer calculations. The soil and canopy cover distributions depend on the spatiotemporal distributions of land cover types obtained from the 1-km Moderate Resolution Imaging Spectroradiometer (MODIS) land cover dataset (available at http://www2.mmm.ucar.edu/wrf/users/download/get_sources_wps_geog.html).

2.2. Precipitation Forcing

We use four publicly available gridded precipitation datasets derived from statistical interpolation based on in situ measurements across the continental United States, including the Parameter-elevation Regressions on Independent Slopes Model (PRISM) dataset (Daly et al., 2008; hereinafter PRISM), the North American Land Data Assimilation System (NLDAS) version 2 dataset (Xia et al., 2012; hereinafter NLDAS2), the Livneh et al. (2015) dataset (hereinafter L15), and the Newman et al. (2015) dataset (hereinafter N15). For comparison, we also use the precipitation produced by convection-permitting WRF simulations (Liu et al., 2017; hereinafter WRF). These datasets reflect important independent efforts to estimate precipitation over the WUS mountains and have been widely used in previous studies (Henn et al., 2018; Jing et al., 2017; Lundquist et al.,

Table 1
Main Methodological Assumptions in Deriving the Precipitation Datasets Used in this Study to Evaluate the Western United States

Dataset[a]	Data type	Resolution	Observation source	Interpolation method	Topographic correction	Uncertainty or quality control
PRISM (Daly et al., 2008)	observation-based statistical interpolation	daily 4 km	multiple networks (e.g., NRCS SNOTEL, NWS COOP)	inverse distance, cluster weighting	empirical precipitation-elevation regression accounting for coastal proximity, rain shadows, and topographic aspects	10 to >30% errors for annual values in high-elevation WUS terrains
NLDAS2 (Xia et al., 2012)	observation-based statistical interpolation	hourly 1/8° (~12 km)	NOAA CPC	optimal interpolation	scaled to match PRISM 1961–1990 monthly climatology	uncertainty from statistical interpolation and measurement sparsity
L15 (Livneh et al., 2015)	observation-based statistical interpolation	daily 1/16° (~6 km)	NWS COOP	inverse distance	scaled to match PRISM 1981–2010 monthly climatology, a fixed lapse rate, MTCLIM[b]-derived elevation dependency	only sites with >20 years of valid data
N15 (Newman et al., 2015)	observation-based statistical interpolation	daily 1/8° (~12 km)	multiple networks (e.g., NRCS SNOTEL, NWS COOP)	probabilistic interpolation	multiple linear regression using topographic predictors (elevation, aspect, and location)	uncertainty range generated from 100-member ensemble
WRF (Liu et al., 2017)	convection-permitting model simulation	daily 4 km	—	—	physically modeled	uncertainty from model physics; no bias from convective parameterization

[a]The PRISM dataset is available at http://www.prism.oregonstate.edu/. The NLDAS2 dataset is available at https://ldas.gsfc.nasa.gov/nldas/NLDAS2forcing.php. The Livneh et al. (2015) dataset is available at ftp://livnehpublicstorage.colorado.edu/public/Livneh.2016.Dataset/. The Newman et al. (2015) dataset is available at https://doi.org/10.5065/D6TH8JR2. The 4-km WRF dataset (Liu et al., 2017) is available at https://rda.ucar.edu/datasets/ds612.0/. [b]MTCLIM: mountain microclimate simulator (Bohn et al., 2013).
Abbreviations: COOP=Cooperative Observer Program; NLDAS=North American Land Data Assimilation System; NOAA=National Oceanographic and Atmospheric Administration; NRCS=Natural Resources Conservation Service; NWS=National Weather Service; PRISM=Parameter elevation Regressions on Independent Slopes Model; SNOTEL=Snowpack Telemetry; WRF=Weather Research and Forecasting; WUS=western United States.

2015; Musselman et al., 2018). Here we briefly describe each dataset and summarize major data features in Table 1.

We use the 4-km daily PRISM dataset that applies an empirical statistical model to interpolate climate variables (e.g., precipitation and temperature) into fine-scale grids based on a variety of ground measurement networks. For precipitation measurements, the Snowpack Telemetry (SNOTEL) network operated by the Natural Resources Conservation Service (NRCS) and the National Weather Service (NWS) Cooperative Observer Program (COOP) gauge network are the two primary sources. The statistical interpolation adopts an empirical precipitation-elevation regression for topographic correction, which also accounts for coastal proximity, rain shadows, and topographic aspects (Daly et al., 2008). The errors of PRISM precipitation over high-elevation WUS terrain are estimated to be 10 to >30% in annual values and are attributable to data interpolation process, measurement data quality, and sparsity of measurement sites over mountains (Daly et al., 2008). Recent evaluations (Strachan & Daly, 2017) further revealed that the 4-km PRISM estimates of snowfall are significantly underestimated at several high-elevation sites over the semiarid eastern Sierra Nevada mountains and overestimated at some low-elevation sites, because of surface temperature biases.

The L15 and NLDAS2 precipitation datasets are available at daily 1/16° (~6 km) and hourly 1/8° (~12 km) resolutions, respectively. Both datasets are produced through statistical interpolation relying on ground measurements and are scaled to match the 30-year PRISM climatology (1981–2010 for L15 and 1961–1990 for NLDAS2) for each month as further topographic correction. The data sources for L15 are exclusively the NWS COOP gauge measurements that have >20 years of valid data (Livneh et al., 2015), while NLDAS2 uses the daily gauge-based precipitation analyses from the National Oceanographic and Atmospheric Administration (NOAA) Climate Prediction Center and is then disaggregated to hourly

frequency based on the NWS Doppler radar-based analyses (Xia et al., 2012). We note that L15 has some similarities to PRISM, due to the use of NWS COOP measurements and the scaling of L15 to match the PRISM monthly climatology. However, there are still many differences between L15 and PRISM, including the use of more data sources in PRISM, different data quality control procedures, and different interpolation/mapping algorithms. This could lead to different daily, monthly, and seasonal precipitation variations for individual years in the two datasets, which is reflected by our following analysis (section 3.1).

The N15 dataset with a daily 1/8° (~12 km) resolution is developed through probabilistic interpolation (Clark & Slater, 2006) based on several measurement networks, including the NRCS SNOTEL and NWS COOP gauge data. The topographic correction, independent of PRISM, is then performed via multiple linear regression using topographic predictors (i.e., elevation, aspect, and location). To explicitly quantify the uncertainty related to precipitation-topography regression, a 100-member ensemble of gridded precipitation is further generated by using spatially correlated random fields (Clark & Slater, 2006; Newman et al., 2015). In this study, we use the 5, 50, and 95% ensemble members (hereinafter N15-5%, N15-50%, and N15-95%) based on their total accumulated precipitation from October to June in each grid in order to demonstrate the 90% uncertainty range from the N15 ensemble dataset.

Compared with the statistically interpolated datasets, the convection-permitting WRF simulations are advantageous in physically and dynamically representing terrain–flow–precipitation interactions and hence orographic precipitation as well as projecting future precipitation changes in complex terrain. Particularly over the high-elevated mountains with rather scarce ground measurements, the WRF precipitation can avoid uncertainties/biases from statistical interpolation methods (e.g., missing key physical links and processes) and measurement sparsity/unrepresentativeness. Here we use the 4-km hourly WRF precipitation (Liu et al., 2017) to assess whether it could outperform the preceding statistically interpolated observational datasets in driving snowpack simulations over the WUS mountains. Specifically, the 4-km WRF simulations employed the WRF version 3.4.1 with 51 vertical levels up to 50 hPa and the following key physics schemes: the Yonsei University planetary boundary layer scheme (Hong et al., 2006), the Thompson cloud microphysics scheme (Thompson & Eidhammer, 2014), the revised Monin–Obukhov surface layer scheme (Jimenez et al., 2012), the Noah-MP land surface scheme (Niu et al., 2011), the Rapid Radiative Transfer Model for General Circulation Models shortwave and longwave radiation scheme (Iacono et al., 2008), and no cumulus parameterizations. Spectral nudging was additionally applied with several other upgrades in model parameterizations, including an improved lake water temperature treatment, a new microphysics-based snow–rain partition, modified snow cover and patchy snow treatments, and the inclusion of heat transport from precipitation to ground. The WRF model simulations were driven by the initial, boundary, and nudging meteorological fields from 6-hourly 0.7° ERA-Interim reanalysis data (Dee et al., 2011). More details are provided by Liu et al. (2017). This WRF precipitation dataset has also been used in previous studies (e.g., Jing et al., 2017; Musselman et al., 2017; Wang et al., 2018). We should note that the WRF precipitation may be sensitive to model configurations such as cloud microphysics (e.g., Liu et al., 2011). Thus, WRF simulations with the same spatial resolution but different model physical schemes may result in substantially different precipitation estimates (Lundquist et al., 2019). The model configuration of the 4-km WRF simulations used in this study has been properly optimized by Liu et al. (2017) to generate relatively accurate meteorological fields in the continental United States during 2000–2013.

For a consistent spatiotemporal resolution in all simulations, the statistically interpolated precipitation datasets are resampled into the 4-km WRF forcing grids using bilinear interpolation following previous studies (Henn et al., 2018; Jing et al., 2017). For datasets with daily frequency, we further disaggregate them into hourly frequency by applying the WRF hourly variation. Specifically, we scale the WRF hourly precipitation by the ratio of daily total precipitation in the target dataset to that in the WRF dataset. As such, the target dataset keeps its daily total precipitation unchanged but has hourly variation patterns similar to the WRF precipitation. For the days when there is nonzero precipitation in the target dataset but zero precipitation in the WRF dataset, we use the monthly averaged WRF hourly variation or annually averaged variation (if there is still zero WRF precipitation throughout the month).

2.3. Observations

We use a suite of observational datasets (see Table 2 for summary) to evaluate simulated snowpack properties over the WUS mountains. For SWE and precipitation, we use the NRCS SNOTEL in situ daily measurements

Table 2
Observational Datasets Used for Model Evaluation in this Study

Variable	Dataset name[a]	Dataset type	Resolution	Quality control or note	Reference[b]
SWE	SNOTEL	in situ	daily 762 sites	potential overestimates due to drifting snow	Serreze et al. (1999)
SWE	UA	reanalysis	daily 4 km	assimilation of in situ snow data and PRISM temperature & precipitation data	Broxton et al. (2016)
SD	SNODAS	reanalysis	daily 1 km	assimilation of snow data from ground, airborne, and satellite observations	Carroll et al. (2001)
SD	UA	reanalysis	daily 4 km	assimilation of in situ snow data and PRISM temperature & precipitation data	Broxton et al. (2016)
SCF	MODIS/Terra (MOD10C1)	satellite	daily 0.05°	quality flag ≤2 ("OK" or better), cloud cover < 20%	Hall and Riggs (2016)
SCF	MODIS/Aqua (MYD10C1)	satellite	daily 0.05°	quality flag ≤2 ("OK" or better), cloud cover < 20%	Hall and Riggs (2016)
SCF	IMS	satellite in situ	daily 4 km	mapping from satellite and in situ data	National Ice Center (2008)
ALBD	MODIS (MCD43C3)	satellite	daily 0.05°	quality flag ≤3 ("≤25% fill values" or better)	Schaaf and Wang (2015)
T2M	PRISM	reanalysis	daily 4 km	statistical interpolation based on in situ data	Daly et al. (2008)
PRCP	SNOTEL	in situ	daily 762 sites	10–15% precipitation undercatch	Serreze et al. (1999)

[a]See text for dataset full names. [b]The SNOTEL observations are available at https://www.wcc.nrcs.usda.gov/snow/. The SNODAS data is available at http://nsidc.org/data/G02158/. The University of Arizona (UA) snow products are available at https://climate.arizona.edu/data/UA_SWE/. The IMS snow cover data is available at https://nsidc.org/data/G02156#. The MODIS snow cover fraction datasets are available at https://nsidc.org/data/MOD10C1 and https://nsidc.org/data/MYD10C1, respectively. The MODIS surface albedo product is available at https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mcd43c3_v006.
Abbreviations: ALBD=Surface Albedo; IMS= Ice Mapping System; MODIS= Moderate Resolution Imaging Spectroradiometer; PRCP=Precipitation; PRISM=Parameter elevation Regressions on Independent Slopes Model; SCF=Snow Cover Fraction; SD=Snow Depth; SNODAS=Snow Data Assimilation System; SNOTEL=Snowpack Telemetry; SWE=Snow Water Equivalent; T2M=Surface 2-m Temperature; UA=University of Arizona.

(Serreze et al., 1999) at ~800 sites (with typical elevations of 600–3600 m) (Figure 1), which has often served as "ground truth" in the region. The SWE and precipitation are measured by snow pillows and gauges, respectively. The SNOTEL gauges are typically in small forest clearings to reduce wind-induced snow drift, with about 10–15% snowfall/precipitation undercatch (Yang et al., 1998). The snow pillow measurements may suffer from overestimation caused by drifting snow mainly for dry snow over high-elevation sites (Meyer et al., 2012).

For SD, we use the daily 1-km product from the NOAA National Operational Hydrologic Remote Sensing Center Snow Data Assimilation System (SNODAS), a modeling and data assimilation system for the continental United States. It assimilates snow data from ground, airborne, and satellite measurements into a physically based mass- and energy-balance National Operational Hydrologic Remote Sensing Center snow model. Detailed information has been documented in Carroll et al. (2001). Hedrick et al. (2015) showed that SNODAS tends to underestimate SD in areas with deep snowpack and dense forests in Colorado Rocky Mountains and has relatively larger biases in alpine regions than forested regions. Similar uncertainty and biases in the SNODAS SWE product over the same mountainous region were also found by Clow et al. (2012). Wrzesien et al. (2017) indicated that SNODAS is generally consistent with some other reference datasets in SWE estimates over Sierra Nevada mountains.

To further supplement model assessment of SWE and SD, we also use the recent University of Arizona (UA) daily 4-km reanalysis dataset. The product is generated by assimilating in situ SWE and SD measurements and the 4-km PRISM temperature and precipitation data across the continental United States. Detailed descriptions have been provided in Broxton et al. (2016).

For snow cover fraction (SCF), we use the MODIS Terra and Aqua Level 3 Version 6 products with a daily 0.05° (~5 km; climate modeling grid) resolution (MOD10C1 and MYD10C1). The dataset is derived based on the Normalized Difference Snow Index snow cover from the MODIS daily 500-m grid dataset. Detailed descriptions are provided in Hall and Riggs (2016). Since extensive cloud cover could degrade the SCF accuracy, we use the data with less than 20% cloud cover and quality flags of ≤2 ("OK" or better) in order to

obtain enough spatiotemporal data coverage with adequate quality. For additional independent SCF observations, we also use the NOAA Interactive Multisensor Snow and Ice Mapping System (IMS) daily 4-km snow cover product (National Ice Center, 2008), which is derived from the analysis of various observations including in situ and satellite data.

For surface albedo, we use the MODIS Level 3 Version 6 product with a daily 0.05° (~5 km; climate modeling grid) resolution (MCD43C3), including the direct (black sky, local solar noon) and diffuse (white sky) albedos at visible (0.3–0.7 μm), NIR (0.7–5.0 μm), and shortwave (0.3–5.0 μm) bands. Detailed data descriptions are presented in Schaaf and Wang (2015). To achieve both adequate data quality and spatiotemporal coverage, we use the data with quality flags of ≤3 ("≤25% fill values and ≤75% full inversions" or better).

For surface air temperature, we use the 4-km daily PRISM product derived from statistical interpolation based on various measurement networks (e.g., NWS COOP) with topographic correction (Daly et al., 2008). This dataset has been widely used to evaluate model simulations of surface temperature (Liu et al., 2017; Scalzitti et al., 2016). Strachan and Daly (2017) evaluated the 4-km PRISM surface temperature product by comparing with in situ observations over the semiarid eastern Sierra Nevada mountains. They found a consistent cold bias in daily temperature minimum associated with elevation throughout all seasons and sites, while the cold bias in daily temperature maximum varies with season and is associated with solar exposure and heat loading.

To compare observations and model simulations at a consistent spatial resolution, we resample model results into observational grids/sites using bilinear interpolation following previous studies (e.g., Liu et al., 2017). For detailed analysis, we further divide the entire WUS mountainous areas into eight subregions (see Figures 1 and S1), including the Pacific Northwest mountains, Sierra Nevada mountains, Montana-Idaho-Oregon (MT-ID-OR) mountains, Wyoming-Montana (WY-MT) mountains, Nevada mountains, Utah mountains, Colorado mountains, and Arizona-New Mexico (AZ-NM) mountains.

3. Results and Discussions

3.1. Precipitation

Figure 2 presents the spatial differences in accumulated precipitation throughout October to June (when mountain snow is present in the water year) between the WRF and statistically interpolated precipitation datasets (i.e., PRISM, NLDAS2, L15, N15-5%, and N15-95%) over the WUS. In general, the precipitation variation across these datasets is substantial, particularly over mountains, with strong spatial heterogeneity. Compared with the PRISM, NLDAS2, and L15 precipitation, the WRF precipitation is much higher (by up to ~400 mm accumulated during the period) in the Pacific Northwest, Sierra Nevada, and WY-MT mountains (in contrast to the much lower WRF precipitation near the coast), which also tends to be at the high end of the N15 90% uncertainty range in these mountains (Figure S2), while NLDAS2 shows the lowest precipitation in the three subregional mountains. On the contrary, WRF tends to have lower accumulated precipitation than the other datasets in the interior WUS mountainous areas (i.e., Nevada, Utah, Colorado, and AZ-NM), with subregional mean differences <50 mm (Figures 2 and S2). The pattern of precipitation differences over the MT-ID-OR mountains is more complicated, with many small-scale features (e.g., patchy distribution of positive and negative differences), likely reflecting uncertainties from complex topography–precipitation interactions. We find that L15 generally agrees with PRISM in the pattern and subregional mean values (differences less than ~20 mm) of accumulated precipitation across the WUS mountains, while the N15 90% uncertainty range covers the majority of variations across the other datasets (Figure S2). Moreover, previous studies (Henn et al., 2018; Jing et al., 2017) suggested that the disagreement among the statistically interpolated precipitation datasets is expected to be larger over the WUS mountains with higher elevations due to the lack of in situ measurements (used for the dataset development). Overall, the subregional mean precipitation from different datasets is relatively consistent in the interior WUS mountainous areas but largely differs in the western and northern portions of WUS mountainous areas (i.e., Pacific Northwest, Sierra Nevada, MT-ID-OR, and WY-MT), where the accumulated precipitation tends to be high.

We further evaluate the precipitation datasets against SNOTEL observations over the WUS mountains. WRF well captures the observed accumulated precipitation over the MT-ID-OR and WY-MT mountains with

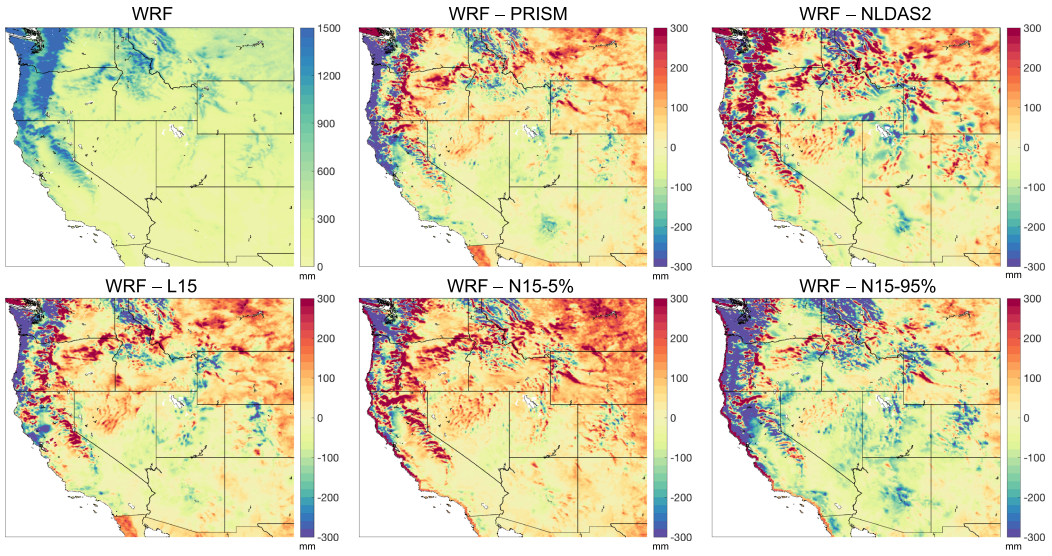

Figure 2. Spatial distributions of accumulated precipitation (unit: mm) throughout October 2012 to June 2013 from the Weather Research and Forecasting (WRF) dataset, and accumulated precipitation differences (unit: mm) between the WRF and other datasets (i.e., PRISM, NLDAS2, L15, N15-5%, and N15-95%; see Table 1 and text for details).

differences <5% but slightly (~10%) overestimates over the Pacific Northwest and Sierra Nevada mountains mainly due to the too strong precipitation during early winter (Figure 3). The precipitation overestimate has also been seen in the 4-km WRF 13-year climatological simulations (Liu et al., 2017) and simulations in other individual years (Chen, Liu, et al., 2014). We should note that the SNOTEL precipitation measurements typically suffer from ~10% undercatch errors (Yang et al., 1998). However, the statistically interpolated datasets (PRISM, NLDAS2, and L15) tend to underestimate the precipitation in the four subregions (Figure 3), with slight (<10%) biases for PRISM but substantial (up to >20%) biases for NLDAS2 (Table S1). Currier et al. (2017) suggested that the PRISM estimates of frozen precipitation tend to be biased low during cold seasons in the Olympic Mountains (one of the Pacific Northwest mountainous areas). The L15 precipitation is consistent with observations in the Pacific Northwest and WY-MT mountains but is largely (~30%) underestimated in the Sierra Nevada mountains. Lundquist et al. (2015) also found that the

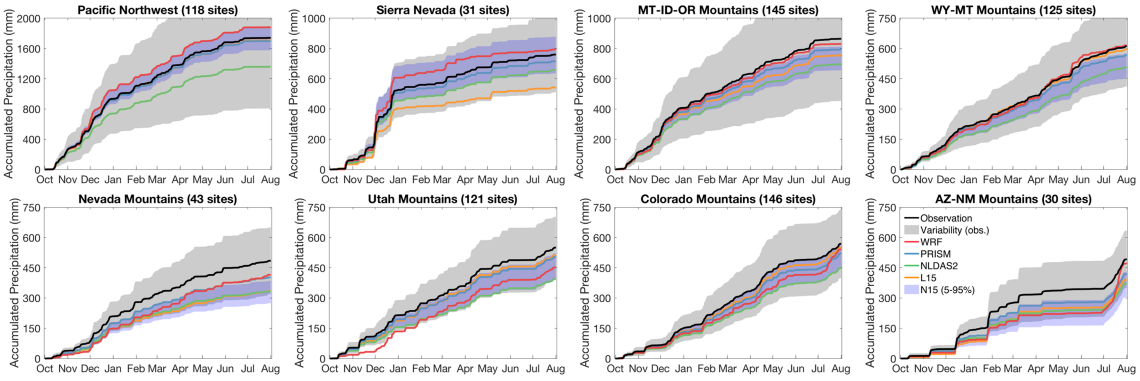

Figure 3. Accumulated precipitation during Water Year 3013 averaged across the SNOTEL sites in each subregion (see Figure 1) from the daily SNOTEL observations (black lines) and different precipitation datasets used in this study, including the WRF (red lines), PRISM (blue lines), NLDAS2 (green lines), L15 (orange lines), and 90% uncertainty range of N15 (purple shaded areas) datasets (see Table 1 and text for details). Also shown is the spatial variability (one standard deviation) of observations in each subregion (gray-shaded area). Note that the SNOTEL observations show strong spatial heterogeneity/variability. The figure legend is shown in the last panel.

statistically interpolated datasets tend to miss some individual storm events in the Sierra Nevada, leading to ~20% underprediction in annual precipitation. Strachan and Daly (2017) showed that the PRISM snowfall estimates are underestimated at some high-elevation sites in the semiarid Sierra Nevada mountains. The N15 90% uncertainty range covers the observations (very close to N15-50%) in the Pacific Northwest and Sierra Nevada mountains but slightly underestimates in the MT-ID-OR and WY-MT mountains (Figure 3).

In contrast, all the datasets underestimate the accumulated precipitation in the interior WUS mountainous areas (Nevada, Utah, Colorado, and AZ-NM) compared with SNOTEL measurements (Figure 3). Specifically, the WRF precipitation is significantly (15–30%) lower than observations in the four subregions. This is due to the inadequate precipitation primarily occurring during mid-autumn to winter, likely caused by insufficient moisture transport from the west coast to the interior WUS regions (Rutz et al., 2014) and/or biases from orographic and cloud microphysical effects on precipitation in WRF (Jing et al., 2017). The performance of statistically interpolated datasets varies notably. NLDAS2 has the largest (20–30%) deviation from observations across the four subregions (Table S1), whereas PRISM shows the smallest (10–20%) negative bias primarily because of its use of SNOTEL measurements in the dataset development. The PRISM bias may come from the statistical interpolation (Chen, Liu, et al., 2014). Daly et al. (2008) estimated that the error of PRISM annual precipitation could be 20–30% over mountains with larger uncertainties in winter than in summer. The L15 precipitation is relatively close (bias <10%) to observations in the Utah and Colorado mountains but is largely (~30%) underestimated in the Nevada and AZ-NM mountains. Interestingly, even the N15-95% precipitation is biased low in the four subregions. This highlights an urgent need to improve precipitation estimates in these mountainous areas (Risser et al., 2019; Timmermans et al., 2019).

Table S1 summarizes the performance statistics for each precipitation dataset compared with SNOTEL measurements, including normalized mean bias (NMB), root mean square error (RMSE), and correlation coefficient (r). Overall, the WRF precipitation shows the smallest NMB in the Sierra Nevada, MT-ID-OR, WY-MT, and Nevada mountains and gives an average performance in the remaining four subregions. Among the statistically interpolated datasets, PRISM (NLDAS2) generally performs the best (worst) with the lowest (highest) NMB and RMSE in the majority of WUS mountains. All the datasets reveal good temporal correlation ($r > 0.9$) with observations. We also note that averaged over the entire WUS mountains, the WRF precipitation indicates the best performance with the smallest NMB (−3.4%) and RMSE (4.9 mm). This, however, might be misleading because it is mainly due to an offset of positive and negative biases in different subregions. Further calculations of the normalized mean absolute value of bias show that PRISM has the lowest bias (7.5%) averaged over the entire WUS mountains, while WRF, L15, and N15-50% have comparable biases (~15%).

3.2. Snow Water Equivalent

To assess the precipitation impact on snowpack simulations, we first compared SWE simulations driven by different precipitation datasets with SNOTEL observations. We find that the WRF-driven model results capture the observed SWE over the western and northern portions of the WUS mountainous areas (Pacific Northwest, Sierra Nevada, MT-ID-OR, and WY-MT) with differences of <50 mm, which overall outperform the results based on the statistically interpolated datasets (Figure 4). This is primarily because of the relatively well-performed WRF precipitation in the four subregions. However, the model results reveal too slow snow ablation in the Sierra Nevada mountains starting around mid-March, which also occurs in the results driven by the other precipitation datasets, likely due to the model deficiency in snow ablation physics (Chen, Barlage, et al., 2014; Chen, Liu, et al., 2014) and/or cold biases in surface temperature in this area (see discussions below). The results based on PRISM, NLDAS2, and L15 precipitation systematically underestimate SWE in the four subregions throughout winter and spring, particularly over the Pacific Northwest and Sierra Nevada mountains (by up to 150–300 mm), which is probably caused by the underpredicted precipitation (Figure 3). Among the three datasets, the PRISM-driven and NLDAS2-driven results have the smallest (up to 17%) and largest (up to 47%) negative biases, respectively, across the four subregions (Figure 4 and Table S1). In the interior WUS mountainous areas (Nevada, Utah, Colorado, and AZ-NM), the SWE simulations driven by both WRF and statistically interpolated precipitation capture the observed temporal patterns but are consistently lower than observations by up to ~150 mm (Figure 4), due to the strong precipitation underestimates (Figure 3). Overall, the PRISM-driven results are the closest to observations in the four

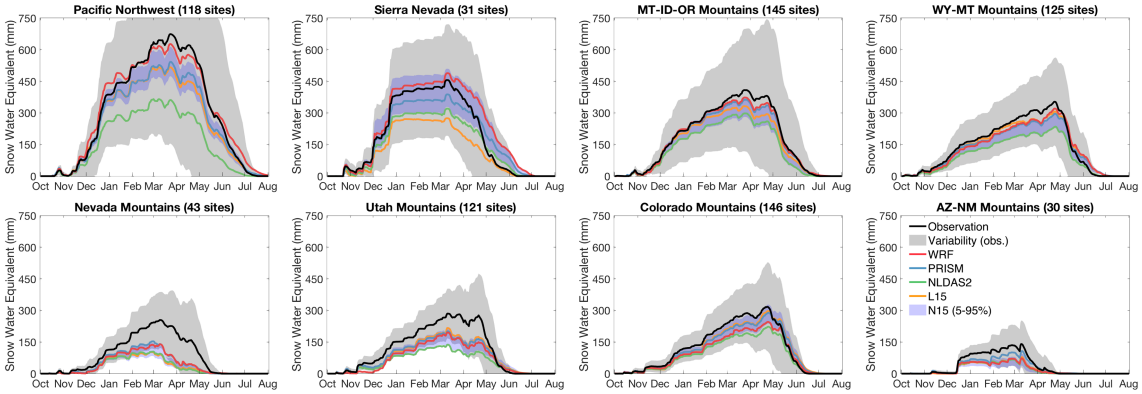

Figure 4. Same as Figure 3, but for daily snow water equivalent (SWE) from the SNOTEL observations and model simulations driven by different precipitation datasets. Note that the SNOTEL observations show strong spatial heterogeneity/variability in each subregion.

subregions, while the WRF-driven results are similar to those based on PRISM in the Nevada and Utah mountains but moderately worse in the Colorado and AZ-NM mountains. We note that the preceding model bias patterns across the eight subregions are valid and robust even when averaged over the entire subregional mountainous areas (instead of only the SNOTEL sites) by comparing with the UA gridded SWE reanalysis data (Figure S3), except that the model underestimates are weaker due to lower UA SWE values relative to SNOTEL observations (Figure S4).

Further analysis shows that model results driven by different precipitation datasets tend to underestimate the peak SWE values and dates across the WUS mountains (Figure 5). This could be due to the insufficient precipitation and/or deficient model treatment of interactions between canopy and snow (Chen, Liu, et al., 2014). We find that the WRF-driven results show the smallest biases for peak SWE (<40 mm) and dates (within ~10 days) in the western and northern portions of the WUS mountainous areas, followed by the PRISM-driven results. In the interior WUS mountainous areas, the PRISM-driven results generally outperform the others, with biases of 40–110 mm for peak SWE and less than ~20 days for peak dates, followed by

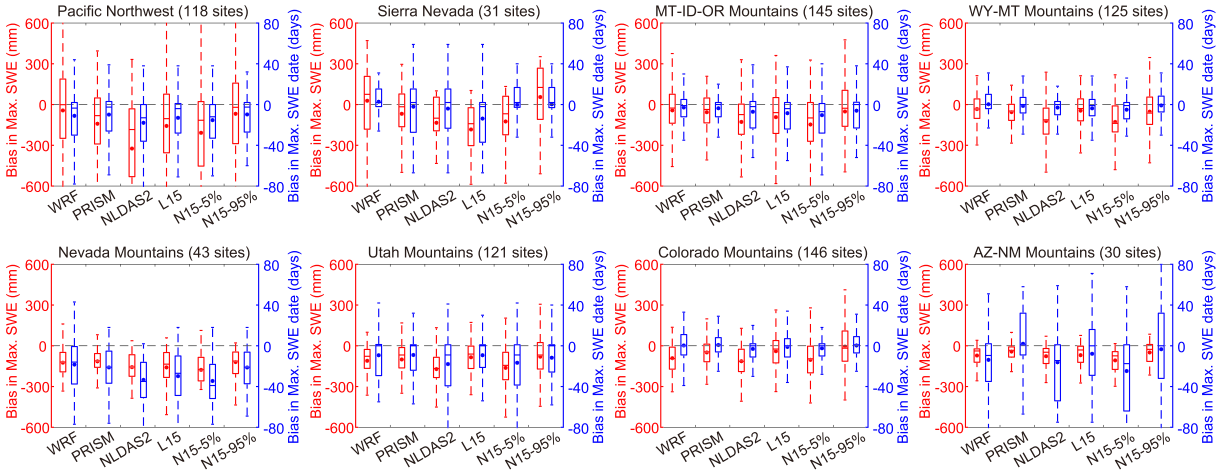

Figure 5. Box plots for model biases (simulations minus observations) in maximum snow water equivalent (SWE) values (red) and date (blue) averaged over the SNOTEL sites in each subregion (see Figure 1) during Water Year 2013. Model simulations are driven by different precipitation datasets, including the WRF, PRISM, NLDAS2, L15, N15-5%, and N15-95% datasets (see Table 1 and text for details). The box includes the means (circles), medians (middle bars), interquartile ranges (25th–75th percentiles), and maxima/minima (±1.5 × interquartile ranges; whiskers). Negative biases in SWE peak date means that model results reach peaks earlier than observations.

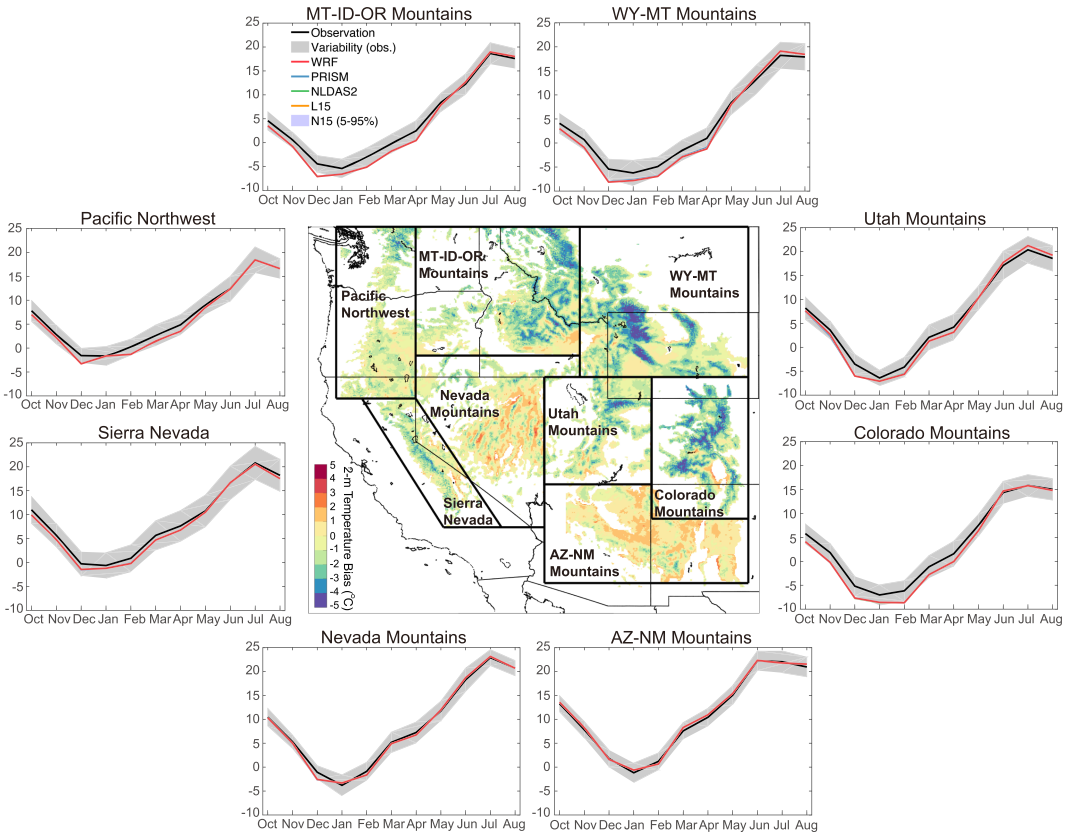

Figure 6. Central panel: Spatial distributions of surface (2-m) temperature biases (simulations driven by the WRF precipitation minus PRISM temperature observations) averaged throughout October 2012 to June 2013 over mountainous areas (i.e., above subregional mean elevations; see also Figure S1). Spatial distributions of model biases based on the other precipitation datasets are shown in Figure S5. Surrounding panels: Monthly mean surface (2-m) temperature (unit: °C) during October to August averaged over mountainous areas in each subregion from PRISM temperature observations (black lines) and model results driven by different precipitation datasets, including the WRF (red lines), PRISM (blue lines), NLDAS2 (green lines), L15 (orange lines), and 90% uncertainty range of N15 (purple shaded areas) datasets (see Table 1 and text for details). Also shown is the spatial variability (one standard deviation) of observations in each subregion (gray-shaded area). Note that model simulations are all overlapped and the purple shaded areas are not visible, due to rather small differences. The daily results are shown in Figure S6.

the results based on the WRF and L15 precipitation. Averaged over the entire WUS mountains, the WRF-driven results perform the best in peak SWE values (bias of about −64 mm) and as good as the PRISM-driven results in peak SWE dates (bias of about −5 days).

In addition to precipitation, surface temperature may also play an important role in SWE simulations through altering snow melting and precipitation partitioning into snowfall and rainfall (Scalzitti et al., 2016). Comparisons between model temperature and PRISM records indicate that model results driven by different precipitation datasets consistently have cold biases (by up to 3 °C) over the WUS mountains throughout October to May, except in the Nevada and AZ-NM mountains (Figures 6 and S5). This is a common problem that also occurs in previous WRF-driven Noah-MP LSM simulations either for other individual years (Chen, Liu, et al., 2014) or a long-term (13 years) climatology (Liu et al., 2017). However, since the cold bias mostly occurs below the freezing/melting point temperature (0 °C), except in the Sierra Nevada, it only has a negligible impact on snowfall partitioning and snow melting in the model. While in the Sierra Nevada mountains, the cold bias could contribute to excessive snowfall and too weak snow ablation (depicted in Figure 4). We should note that the surface (2-m) temperature differences among different simulations are negligible (Figures 6 and S6) mainly due to the use of the same non-precipitation

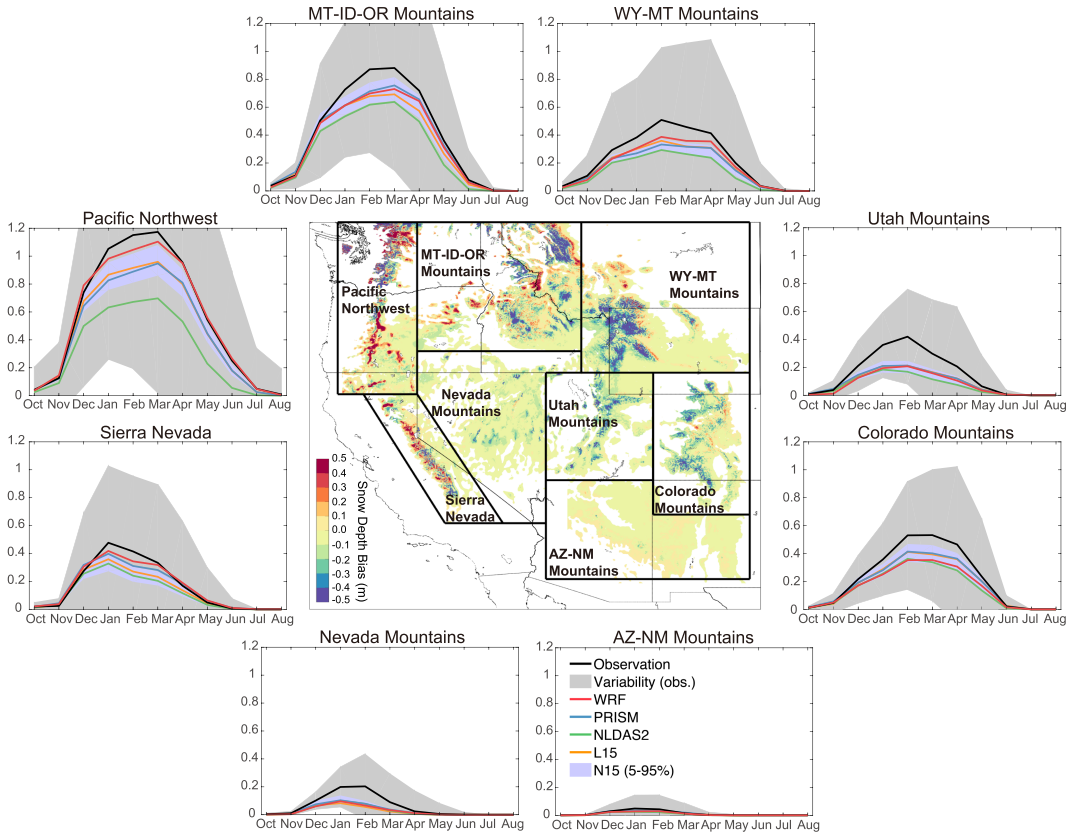

Figure 7. Same as Figure 6, but for snow depth (unit: m). Observations are from the SNODAS data. Spatial distributions of model biases based on the other precipitation datasets are shown in Figure S9. The daily results are shown in Figure S10.

atmospheric forcing data (particularly the temperature forcing exerted at a 10-m height), although the surface (2-m) temperature is dynamically simulated in Noah-MP based on land surface processes and conditions.

Besides, the snow–rain partitioning parameterization may also affect snowfall and hence SWE simulations in Noah-MP. We further examine the influence of the other two snow–rain partitioning parameterizations in Noah-MP, which are functions of surface air temperature only, including the biosphere-atmosphere transfer scheme (BATS) and a simple scheme assuming all precipitation as snowfall when surface air temperature is less than water freezing point (273.16 K). Compared to the standard snowfall simulation using the Jordan (1991) scheme, the spatiotemporal distribution of accumulated snowfall over the WUS mountains using the BATS scheme is very similar, with a few percent differences (Figures S7 and S8), but the simple scheme leads to much lower snowfall particularly in the western and northern portions of the WUS mountainous areas. Thus, using the Jordan (1991) and BATS schemes is expected to have consistent snowpack simulation results, whereas using the simple scheme could result in larger underestimates of SWE over the WUS mountains. Niu et al. (2011) pointed out that over regions where surface air temperature varies around the freezing point, snowpack simulations could be sensitive to the choice of snow–rain partitioning schemes.

Overall, different precipitation datasets show important effects on SWE simulations, including temporal evolution, peak SWE values, and dates, leading to large variations across model results. Model simulations all capture the temporal pattern of observations in the WUS mountains with similarly high correlation coefficients ($r > 0.85$). The WRF-driven results are among the top two best performed simulations over the Pacific Northwest, Sierra Nevada, MT-ID-OR, WY-MT, and Nevada mountains with small NMB and RMSE (Table S1), while the results based on the PRISM and/or L15 precipitation outperform the others in the

remaining subregions. We note that the WRF-driven SWE simulations have the smallest NMB (−9.3%) and RMSE (26.8 mm) averaged over all the subregions, but it should be interpreted with caution since the overestimates in the Sierra Nevada mountains partially offset the large underestimates in the interior WUS mountainous areas.

3.3. Snow Depth

Figure 7 shows the spatiotemporal distributions of SD from model results driven by different precipitation datasets compared with SNODAS observations. Due to the model underestimates of SWE (section 3.2), the simulated SD is consistently biased low across the WUS mountains during December to May, with stronger underestimates at higher elevations (Figures 7 and S9). The spatial pattern of modeled SD biases is similar for different precipitation datasets, with differences in their magnitudes (Figure S9). The biases in the AZ-NM mountains are small (<0.02 m) relative to the other subregions because of the small snow amount. Different precipitation datasets have larger impacts on SD simulations in the western and northern portions of the WUS mountainous areas (Pacific Northwest, Sierra Nevada, MT-ID-OR, and WY-MT), with variations of up to ~0.4 m. The WRF-driven results show the smallest biases (up to ~0.2 m) in the four subregions, followed by the PRISM-driven results, while the NLDAS2-driven results have the largest biases (up to ~0.5 m). On the contrary, model results driven by different precipitation datasets are similar (differences <0.1 m) in the interior WUS mountainous areas (Nevada, Utah, Colorado, and AZ-NM), with slightly better performance for the PRISM-driven results (Figures 7 and S10). We note that the results based on the N15-95% precipitation also do not capture the high observed SD across the eight subregions. Moreover, even with the slightly overestimated SWE in the Sierra Nevada mountains (Figure 4), the WRF-driven simulations still underestimate SD in mid-winter and early spring, suggesting model deficiency in snowpack compaction and metamorphism (Kim & Park, 2019). We find that the aforementioned SD bias patterns are valid and robust when comparing with the UA SD reanalysis data (Figures S11–S13), except that the model underestimates are smaller because of the lower UA SD values relative to SNODAS data.

Overall, model simulations all capture the temporal patterns of observed SD in different subregions ($r >$ 0.9), though with systematic underestimates. The WRF-driven results present the best performed SD simulations over the Pacific Northwest, Sierra Nevada, and WY-MT mountains with the minimum NMB and RMSE, while the PRISM-driven results have the smallest NMB and RMSE in the remaining subregions (Table S1), where the WRF-driven results show a fair performance. Averaged over the entire WUS mountains, the WRF-driven SD simulations show the minimum NMB (−16.7%) and RMSE (0.06 m), followed by results based on PRISM and L15, whereas NLDAS2 leads to the largest NMB (−38.9%) and RMSE (0.13 m).

3.4. Snow Cover Fraction

In contrast to the model underestimates in SWE and SD, the SCF simulations are systematically higher than the MODIS/Terra observations across the WUS mountains mainly during October to May, with similar bias patterns and magnitudes for model results driven by different precipitation datasets (Figures 8, S14, and S15). Thus, precipitation has relatively small (differences <0.1) effects on SCF simulations in the WUS during winter and spring compared with the large effects on SWE and SD simulations, which may depend on the LSM (see discussions below). We find that the SCF biases are generally smaller in the Nevada and Utah mountains than in the other subregions. The model biases tend to be small when the SCF reaches the peak in January, while they are particularly large during snow accumulation and ablation periods (Figure 8). Further analyses with the MODIS/Aqua and IMS SCF observations reveal that these model bias patterns are valid and robust (Figures S16–S21), with slightly stronger and weaker model overestimates for comparisons with the MODIS/Aqua and IMS data, respectively.

To understand the counterintuitive SCF and SD bias patterns (i.e., underestimated SD but overestimated SCF), we investigated the SCF–SD relationship from observations and model simulations, since SCF is parameterized as a function of SD in the model. We find that the SCF increases too fast with the increasing SD when SD is lower than 0.1 m in the model relative to observations (Figure 9), while the SCF–SD parameterization (Niu & Yang, 2007) used in Noah-MP also tends to produce higher SCF compared with other widely used parameterizations (Figure 9). As a result, this leads to the model overestimates of SCF and points towards urgent improvement of SCF formulation, which will be investigated in our future study.

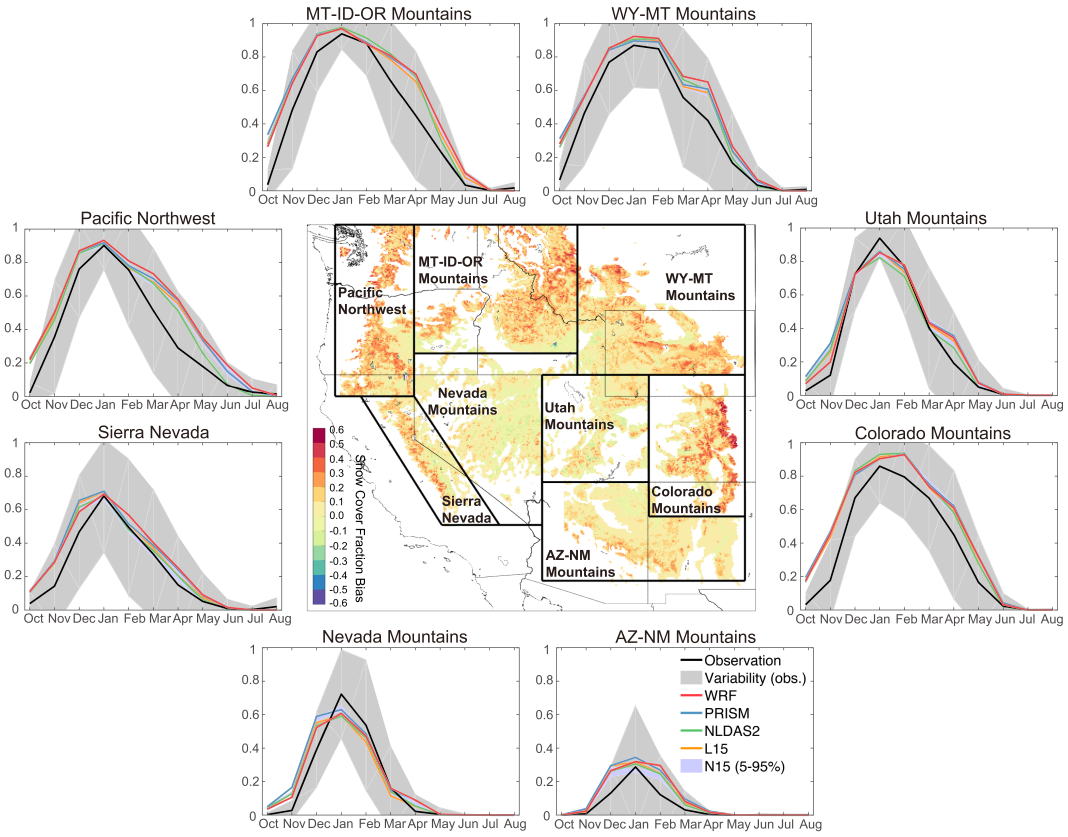

Figure 8. Same as Figure 6, but for snow cover fraction. Observations are from the MODIS/Terra satellite products. Spatial distributions of model biases based on the other precipitation datasets are shown in Figure S14. The daily results are shown in Figure S15.

Furthermore, we also note that when SD is higher than 0.1 m, the modeled SCF is insensitive to SD and very close to 1.0. This explains the relatively small precipitation impacts on the modeled SCF (as opposed to the large impacts on SD and SWE) during winter and spring when SD is larger than 0.1 m (Figure 7).

Overall, model simulations have consistent temporal patterns with the observed SCF over the WUS mountains ($r > 0.9$) but with systematic overestimates. Model results based on different precipitation datasets show similar performance in each subregion, with NMB of about 20–30% and RMSE of about 0.1–0.2 in the majority of subregions, except for smaller NMB (<16%) over the Nevada and Utah mountains and larger NMB (up to 83%) over the AZ-NM mountains (Table S1). In contrast to the poor performance of the NLDAS2-driven SWE and SD simulations, the NLDAS2-driven SCF simulations on average perform the best (slightly better than the others) across the eight subregions with a NMB of 20% and a RMSE of 0.08. This, however, is due to the offset of two biased model factors (i.e., overestimated SCF from the biased SCF–SD relationship and largely underestimated SD).

3.5. Surface Albedo

Figure 10 shows the shortwave (0.3–5.0 μm) diffuse (white sky) surface albedo based on MODIS satellite observations and model simulations. The results and spatiotemporal patterns of direct (black sky) surface albedo (Figures S28–S34) are similar to those of diffuse albedo and will not be discussed in detail here. We find that model results driven by different precipitation datasets, with similar bias patterns and magnitudes (Figures 10 and S22), tend to significantly overestimate the surface albedo by up to 0.15 over the WUS mountains (except for AZ-NM) particularly during October to January, while the overestimates are smaller (less than ~0.05) during mid-winter to summer. This is partially due to the widespread overestimate in SCF across the WUS mountains (Figure 7), since snow albedo is typically much higher than that of other land surfaces.

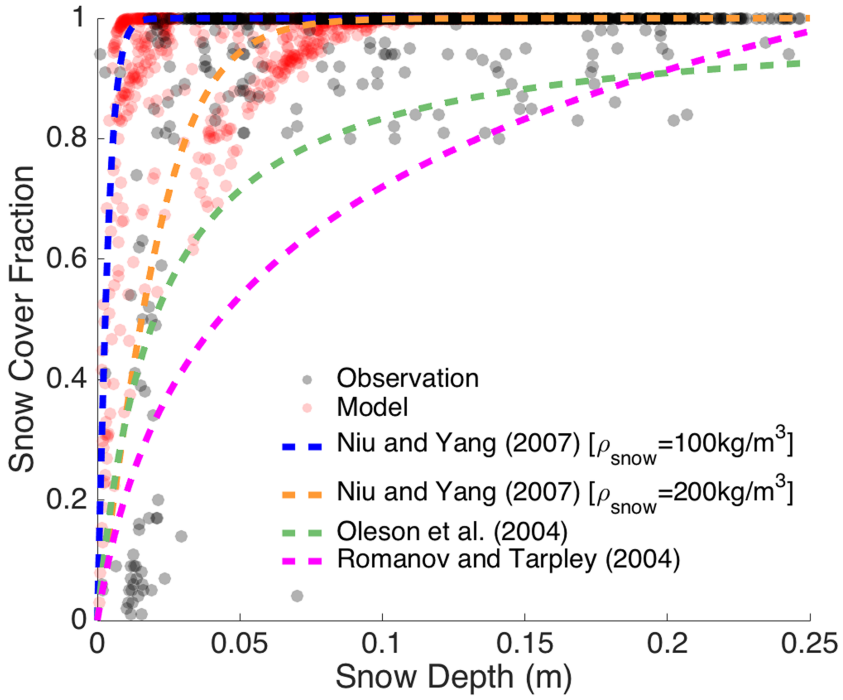

Figure 9. Snow cover fraction (SCF) as a function of snow depth (SD) on bare grounds (to avoid canopy interference) from spatiotemporally colocated observations (black dots) and model results driven by the WRF precipitation (red dots) over the western United States during Water Year 2013. The observed SCF and SD are from the daily MODIS/Terra satellite products and SNODAS datasets, respectively. Also shown are four widely used SCF–SD parameterizations (dashed lines) developed in previous studies, including the Niu and Yang (2007) relationship (used in Noah-MP) with snow densities of 100 and 200 kg m^{-3} (blue and orange), the Oleson et al. (2004) relationship (green; used in the Community Land Model), and the Romanov and Tarpley (2004) relationship (magenta). Note that the Niu and Yang (2007) parameterization with a snow density of 100 kg m^{-3} (black) is equivalent to the Yang et al. (1997) parameterization.

Minder et al. (2016) also found a large positive bias in surface albedo caused by the SCF overprediction in the central Rocky Mountains. Besides, the cold bias in surface temperature (see section 3.2 and Figure 6) could indirectly contribute to the positive surface albedo bias through changing surface energy balance and slowing snow melting, sublimation, and/or aging processes.

Moreover, Chen, Barlage, et al. (2014) showed that the snow albedo in Noah-MP is also biased high compared with MODIS and in situ measurements in the WUS mountains. The snow albedo overestimates may be related to inadequate model treatments of processes such as aerosol contamination in snow (He et al., 2018) and/or snow aging (Flanner et al., 2007) in Noah-MP. Further analyses reveal that the visible surface albedo in the model is underestimated during winter and spring (Figures S24 and S26), whereas the NIR albedo is substantially (by up to 0.2) overestimated throughout the year (Figures S25 and S27), which dominates and hence leads to the overestimates in shortwave albedo. This indicates that the modeled snow albedo is biased low and high at the visible and NIR bands, respectively. As a result, the missing aerosol contamination effect in Noah-MP is not likely the reason for the shortwave albedo overestimate, since impurities mainly reduce snow albedo at the visible band (He, Takano, Liou, Yang, et al., 2017). The positive NIR bias is partially due to the inaccurate model treatment that assumes the same visible and NIR snow albedo in the Canadian LAnd Surface Scheme, whereas the NIR snow albedo should be much smaller than the visible one in reality (Wiscombe & Warren, 1980). Too weak snow aging in the model may also contribute to the NIR positive bias, since aging processes result in larger snow grain sizes and thus smaller NIR snow albedo (Flanner et al., 2007; Wiscombe & Warren, 1980).

In addition, the biases in background (snow-free) albedo of different land surfaces and vegetation distribution could also contribute to the albedo overestimate in the WUS mountains (dominantly vegetated),

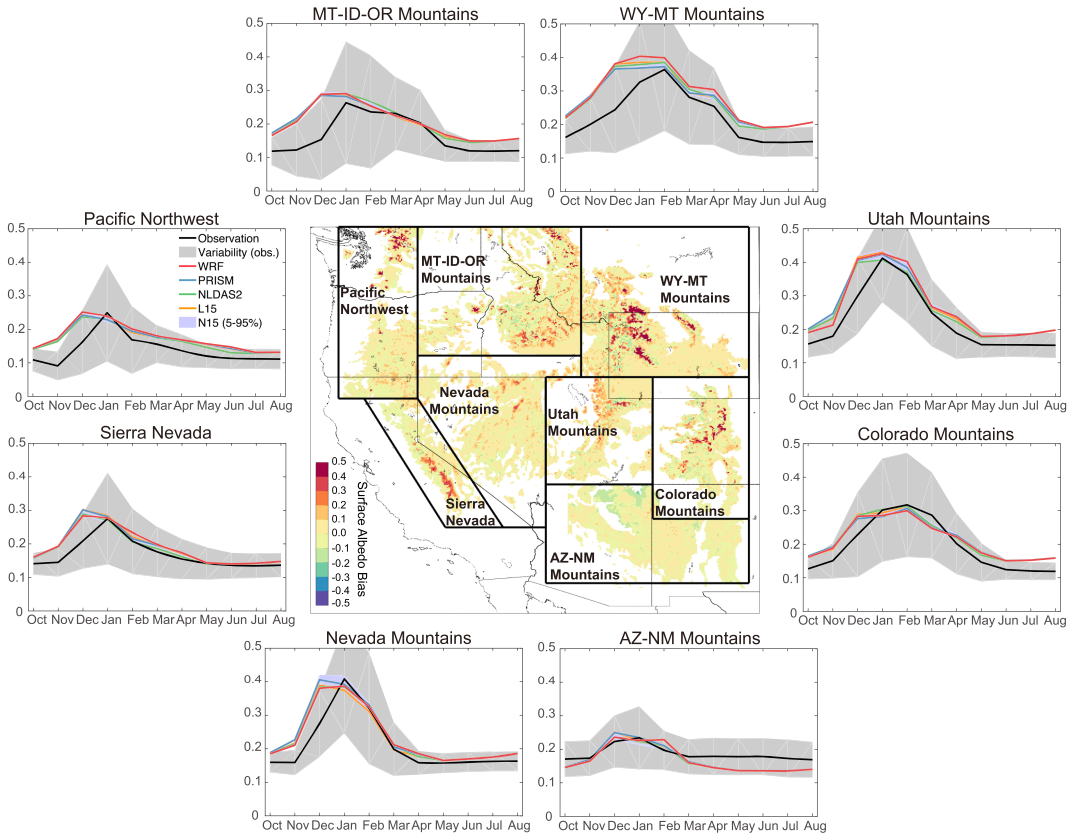

Figure 10. Same as Figure 6, but for shortwave (0.3–5.0 μm) diffuse (white sky) surface albedo. Observations are from the MODIS satellite products. Spatial distributions of model biases based on the other precipitation datasets are shown in Figure S22. The daily results are shown in Figure S23.

because the modeled surface albedo with negligible SCF during summer is still overestimated by up to ~0.05. For example, the albedo overestimates are much stronger at very high elevations (above ~3,000 m) than at low-elevation mountains (Figures 10 and S22), probably due to the uncertainties in vegetation distributions and interactions with snow in alpine areas. Previous studies (e.g., Chen, Barlage, et al., 2014; Essery et al., 2009; Minder et al., 2016) also suggested that the inadequate representation of canopy–snow interactions (e.g., vegetation shedding effects, canopy radiative transfer, and below-canopy turbulence), which is common in most LSMs, could lead to large biases in surface albedo and snowpack simulations particularly in alpine regions. This points towards the necessity of assessing and improving these model processes.

Overall, different precipitation datasets have rather limited impacts on surface albedo simulations in this study (Table S1), which consistently overestimate the shortwave albedo across the WUS mountains (except for AZ-NM). The NMB and RMSE of these simulations averaged over the eight subregions range from 12.4 to 14.4% and from 0.036 to 0.040, respectively.

4. Implication and Uncertainty

The preceding model assessments demonstrate that different precipitation datasets, with substantial cross-dataset variations over the WUS mountains, have important impacts on SWE and SD simulations but relatively small effects on SCF and surface albedo simulations in Noah-MP. Snowpack simulations driven by the WRF precipitation are as good as (or sometimes better than) those driven by the statistically interpolated precipitation (as summarized in Figures 11 and S35 and Table S1) in Water Year 2013. Thus, this study

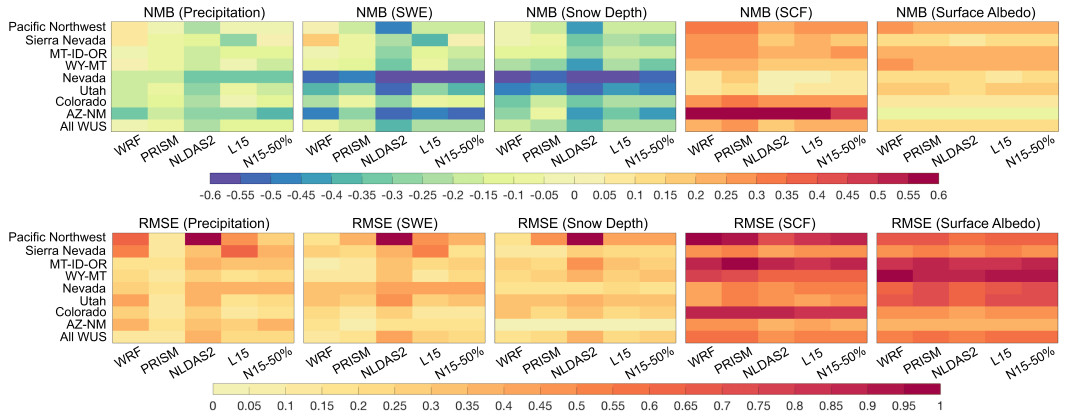

Figure 11. Top row: Normalized mean bias (NMB) of model simulations driven by different precipitation datasets (WRF, PRISM, NLDAS2, L15, and N15-50%) for precipitation, snow water equivalent (SWE), snow depth, snow cover fraction (SCF), and shortwave diffuse surface albedo in eight subregional mountains (Pacific Northwest, Sierra Nevada, MT-ID-OR, WY-MT, Nevada, Utah, Colorado, and AZ-NM) and the entire western US mountains (All WUS). Bottom row: Same as the top row, but for root mean square error (RMSE). Note that the RMSE here is normalized to a scale of 0~1 through dividing the original RMSE by the maximum RMSE for each variable (precipitation, SWE, snow depth, SCF, and surface albedo). The NMB and RMSE are from Table S1. The relative rank for the performance (NMB and RMSE) of different model simulations is shown in Figure S35.

justifies that convection-permitting modeling with proper configurations can provide adequately accurate precipitation for high-resolution snowpack simulations over the WUS mountains in a typical ENSO-neutral year. This further gives some confidence and reliability to the use of convection-permitting modeling in providing precipitation for high-resolution mountain snowpack modeling under future (changing) climate and particularly in high-elevated areas without sufficient in situ measurements.

In addition to the precipitation impact, this study also reveals several areas that need to be improved in snowpack–physics parameterizations in Noah-MP (see section 3 for details). For example, the largely overestimated SCF in all simulations suggests a biased SCF–SD relationship in the model and requires an imperative improvement. The positive snow and surface albedo biases in the model reflect inadequate treatments of several key processes, including canopy–snow interactions, snow albedo and aging parameterizations, and background albedo of different land surfaces. Many of these problems are common in LSMs (Chen, Barlage, et al., 2014; Essery et al., 2009). The albedo overestimates in the model could also feedback to the cold bias in surface temperature over the WUS mountains (Figure 6), which may further affect surface heat transport and energy balance, and hence snow melting and hydrological cycle (Chen, Liu, et al., 2014). These highlighted LSM deficiencies will be systematically assessed and improved in our future study.

Note, however, that several uncertainty factors may affect the analysis in this study. The observational datasets used in model evaluation could be associated with uncertainties. For example, the undercatch problem related to SNOTEL precipitation measurements may lead to an underestimate of the negative biases for those precipitation datasets in each subregion. This is probably the reason that a ~10% overestimate in the WRF precipitation could lead to the best agreement with SNOTEL SWE simulations, for example, over the Pacific Northwest mountains (Figures 3 and 4). Besides, different satellite and reanalysis products also have nontrivial uncertainties and differences. For SWE, the UA reanalysis product provides relatively accurate large-scale distributions compared with in situ SNOTEL measurements but still tends to underestimate over mountains in Water Year 2013 (Figure S4). For SD, previous studies (e.g., Clow et al., 2012) found that the SNODAS product performs satisfactorily in forested regions but poorly in alpine areas, because the assimilated observations are primarily from subalpine forested regions. The SNODAS data also tend to be higher than the UA SD reanalysis over the WUS mountains (section 3.3). For SCF, the three observational products (MODIS/Terra, MODIS/Aqua, and IMS) reveal noticeable differences, with the highest and lowest values for IMS and MODIS/Aqua, respectively (Figures S14–S21). Despite the uncertainties

and cross-product differences in observations, our analyses have shown that the spatiotemporal model bias patterns are valid and robust (see section 3).

Besides, the uncertainty in LSM processes could also influence the results in this study, particularly for snow-related processes including snow cover change, canopy–snow interactions, snow albedo evolution, and snow aging/metamorphism. For example, Chen, Barlage, et al. (2014) comprehensively compared six commonly used LSMs in terms of seasonal snowpack simulations over the Colorado Headwaters region. They found significant intermodel differences in snowmelt and sublimation efficiencies and systematical biases from all the models in snow albedo and snow–canopy interaction simulations. Among these models, Noah-MP performed relatively better in simulating many of snowpack properties. Chen, Liu, et al. (2014) pointed out that the differences in surface temperature and SWE simulations in the WUS mountains among different LSMs are largely attributed to distinct rain–snow partitioning, treatments of snow albedo and vegetation, and surface data. Gao et al. (2015) further showed that the uncertainty of soil heterogeneity representations in Noah-MP could also affect seasonal simulations of surface energy and water budgets over snow-covered Tibetan Plateau areas. Thus, it is necessary to quantify the model uncertainty by using an ensemble of different LSMs and/or an ensemble of different physics options for key model processes in Noah-MP, which will be investigated in future study.

Another uncertainty associated with Noah-MP is the model spatial resolution, which could be important to snowpack and hydroclimate simulations over mountains. Clark et al. (2011) showed that model resolution of sub-kilometer is needed for LSMs to capture the subgrid variability of snow depth and SWE over mountains particularly in melting seasons. Dutra et al. (2011) suggested that horizontal resolution of LSMs plays a crucial role in characterizing snow cover in Northern Hemispheric complex terrain. Kim and Park (2019) found that different LSMs (including Noah-MP) with horizontal resolution of 30 km tend to have large biases in snowpack simulations over Eurasian mountainous regions partially due to the relatively low spatial resolution. Thus, accurate snowpack simulations by LSMs require systematical understanding and assessment of model resolution effects.

In addition, the non-precipitation atmospheric forcing used in this study could also be associated with uncertainties. Liu et al. (2017) showed that the 4-km WRF dataset captures the spatiotemporal patterns of observed near-surface temperature with a correlation of >0.96 and a typical cold bias (within ±3 °C) over the WUS particularly in localized snow-topped mountains during cold seasons. Wang et al. (2018) found similar surface temperature bias patterns from 4-km WRF simulations over the interior WUS mountainous areas, with a spatial correlation of >0.96 and a RMSE of 1.5–1.8 °C. Nevertheless, these evaluations mainly focused on surface temperature among the non-precipitation forcing variables. Thus, a more systematic assessment of the other non-precipitation forcing variables such as downward radiation, which is sensitive to cloud microphysics in WRF (Liu et al., 2011), is necessary in future studies.

Another uncertainty source is the resampling of model results into observational grids/sites during evaluation and the resampling of precipitation forcing datasets into the 4-km model grids during simulations. We used the bilinear interpolation for the resampling process, which is widely used in previous studies (Henn et al., 2018; Jing et al., 2017; Liu et al., 2017). Further tests also showed that different interpolation methods (e.g., nearest neighbor) give similar results and do not change the main model bias patterns and conclusions in this study (not shown). We also note that the disaggregation of statistically interpolated datasets from daily frequency to hourly frequency (section 2.2) could introduce uncertainty in our analyses, which however may only have very limited effects since we focus on monthly to seasonal timescales in this study.

One limitation of this study is that we only focus on one typical ENSO-neutral water year (2013) as a case study, with the annual precipitation over the WUS in a near-normal state (close to the regional climatological mean value based on the NOAA precipitation database; see https://www.ncdc.noaa.gov/cag/). We note that multiyear analyses are necessary to investigate interannual variabilities (e.g., wet, dry, and normal years) and associated impacts on the results in this study, which will be included in a follow-up study. In fact, Liu et al. (2017) showed that the 13-year convection-permitting WRF simulations with proper configurations generally capture the seasonal and interannual variabilities of observed precipitation over the WUS, with as good performance as observation-based interpolated datasets such as the N15 data. Wang et al. (2018) further found that the 30-year high-resolution (4-km) WRF simulations with proper configurations

reproduce observed seasonal precipitation amounts and spatial distributions reasonably well over the interior WUS mountains, which are almost equivalent to a reanalysis.

We should note that this study focuses on evaluating and comparing the performance of convection-permitting WRF simulations and statistically interpolated observational datasets in terms of precipitation and snowpack simulations over the WUS mountains, where the WRF simulations show as good performance as (if not better than) the interpolated observational datasets with comparable biases (section 3). Thus, we conclude that convection-permitting modeling with proper configurations can provide relatively accurate precipitation for high-resolution snowpack simulations over the WUS mountains in Water Year 2013. However, this does not mean that the 4-km spatial resolution is good enough for WRF to produce accurate precipitation in complex terrain, since we do not assess the WRF performance with various spatial resolutions in this study, which is an important subject for investigation in our future study. In fact, Ikeda et al. (2010) evaluated the sensitivity of snowfall estimates in Colorado mountains to WRF model resolutions (i.e., 2, 6, 18, and 36 km). They found that using resolutions of 18-km and 36-km underestimates snowfall by 10–30% during cold seasons (November–April), whereas the 2-km and 6-km simulations show similar results and well capture the accumulated snowfall (biases <5%). Jing et al. (2017) further showed that using a WRF resolution (1.33 km) higher than 4 km resolves finer structures of terrain-relative winter precipitation in the interior WUS mountains and has slightly better agreement with SNOTEL measurements with a stronger correlation and a smaller RMSE.

Since this study uses the precipitation from only one WRF model realization in Water Year 2013, this may introduce uncertainty to our analysis and results due to atmospheric internal variability. Recently, Gowan et al. (2018) assessed precipitation forecasts from 10 member 3-km WRF ensemble simulations with different initial conditions during the 2016/17 cold season and found that the ensemble simulations generally capture the spatiotemporal distribution of precipitation over the WUS mountains but with insufficient spread. We note that it is important for future studies to conduct WRF ensemble simulations to quantify the impact of internal variability on precipitation and snowpack simulations over mountains.

5. Conclusions

This study focused on assessing the precipitation impact on high-resolution snowpack simulations over mountains through comparing precipitation estimates from convection-permitting modeling and observation-based statistical interpolation. We conducted 4-km snowpack simulations over the WUS mountains (divided into eight subregions) during a typical ENSO-neutral water year (2013) using the Noah-MP LSM driven by precipitation from the convection-permitting (4-km) WRF simulation and four widely used statistically interpolated datasets that are based on in situ measurements (i.e., PRISM, NLDAS2, L15, and N15). We systematically evaluated the precipitation and model simulations of SWE, SD, SCF, and surface albedo by comparing with a suite of observations.

We found substantial differences across the five precipitation datasets, particularly over mountains, with strong spatial heterogeneity. The cross-dataset differences are relatively large over the western and northern portions of the WUS mountainous areas (i.e., Pacific Northwest, Sierra Nevada, MT-ID-OR, and WY-MT), with the highest (lowest) precipitation from WRF (NLDAS2), whereas the datasets are more consistent in the interior WUS mountainous areas (i.e., Nevada, Utah, Colorado, and AZ-NM). Further comparisons with SNOTEL measurements showed that WRF well captures the observed precipitation over the Pacific Northwest, Sierra Nevada, MT-ID-OR, and WY-MT mountains with slight (~10%) overestimates in the former two subregions, while the statistically interpolated datasets tend to underestimate in the four subregions, with biases of <10% for PRISM and ~20% for NLDAS2. On the contrary, all the datasets consistently underestimate precipitation over the interior WUS mountainous areas, with the best (worst) performance from PRISM (NLDAS2) on average. This suggests the necessity of improving precipitation estimates in these mountains.

Simulations driven by the WRF precipitation generally capture the observed SWE pattern across SNOTEL sites over the western and northern portions of the WUS mountainous areas, while results driven by the statistically interpolated datasets all tend to underestimate SWE in the areas during winter and spring by up to ~20% (50%) for PRISM (NLDAS2) results. In the interior WUS mountainous areas, the SWE simulations driven by different precipitation datasets are systematically lower than observations by up to ~150 mm, though

with well-correlated temporal patterns. The PRISM-driven results are the closest to observations in the four subregions, while the WRF-driven results give an average performance. Further analysis revealed that averaged over the entire WUS mountains, the WRF-driven results perform the best in peak SWE values and as good as the PRISM-driven results in peak SWE dates.

Due to model underestimates of SWE, SD simulations are consistently biased low across the WUS mountains during December to May, although they capture the observed temporal variation. The resulting biases based on different precipitation datasets are similar in spatiotemporal patterns but largely differ in the magnitudes depending on subregions, with relatively strong differences in the western and northern portions of the WUS mountainous areas. The WRF-driven results show the smallest biases (up to ~0.2 m) in the areas, followed by the PRISM-driven results, while the NLDAS2-driven results have the largest biases (up to ~0.5). On the contrary, different simulations are more consistent in the interior WUS mountainous areas, with slightly better performance from the PRISM-driven results. Averaged over all the subregions, the WRF-driven (NLDAS2-driven) simulations show the minimum (maximum) biases.

In contrast to model underestimates of SWE and SD, the SCF simulations are consistently higher than observations across the WUS mountains, with similar bias patterns and magnitudes for different simulations. Model overestimates tend to be small when SCF reaches the peak in January, whereas they are particularly large during snow accumulation and/or ablation periods. The mean biases for different simulations are about 20–30% in the majority of the WUS mountains, with smaller values (<16%) over the Nevada and Utah mountains and larger values (up to 83%) over the AZ-NM mountains.

We found that model results based on different precipitation datasets, with similar bias patterns and magnitudes, tend to significantly overestimate the shortwave surface albedo across the WUS mountains (except for AZ-NM) particularly during October to January, due in part to the SCF overestimate. This is dominated by the strong positive bias at the NIR band, likely related to deficient model treatments of snow albedo and vegetation. The mean biases of different simulations range from 12.4 to 14.4% averaged over the WUS mountains.

Overall, different precipitation datasets, with substantial cross-dataset variations over the WUS mountains, have important impacts on SWE and SD simulations but relatively limited effects on SCF and surface albedo simulations, which may depend on the LSM. This study highlights that convection-permitting modeling with proper configurations can provide adequately accurate precipitation for high-resolution snowpack simulations over the WUS mountains in a typical ENSO-neutral year, which has added values particularly over regions without sufficient in situ measurements. This gives some confidence to the use of convection-permitting modeling in providing precipitation for high-resolution land surface modeling of mountain snowpack under future climate. However, this study also points towards necessary improvements for snowpack physics in the Noah-MP LSM, including the SCF parameterization, snow albedo and aging treatments, and canopy–snow interactions.

Author Contributions

CH and FC conceived the study. CH conducted model simulation, evaluation, and analysis. MB helped in preparing atmospheric forcing data. WT helped in data processing and analysis. AN provided precipitation data. KI and CL provided WRF data. FC, MB, CL, AN, WT, KI, and RR offered valuable discussions and comments in improving the study. CH prepared the manuscript with improvements from all the other coauthors.

Acknowledgments
The authors thank the three reviewers for their constructive comments that help to significantly improve this work. The authors thank Ben Livneh for developing and sharing the updated precipitation dataset and thank Zhe Zhang, Qian Cao, Dennis Lettenmaier, Guo-Yue Niu, and Margaret LeMone for helpful discussions. The authors thank the development teams for Noah-MP, NLDAS-2, PRISM, SNOTEL, SNODAS, UA, IMS, and MODIS products to make the datasets and model publicly available. The National Center for Atmospheric Research (NCAR) is sponsored by the National Science Foundation. C. He was supported by the NCAR Advanced Study Program Postdoctoral Fellowship. F. Chen and M. Barlage would like to acknowledge the support from NCAR Water System, USDA NIFA Grants 2015-67003-23460, NSF INFEWS Grant #1739705, and NOAA OAR Grant NA18OAR4590381. W. Tang thanks the NCAR Advanced Study Program's Graduate Visitor Program. Computing resources were provided by the Climate Simulation Laboratory at NCAR's Computational and Information Systems Laboratory (CISL), sponsored by the National Science Foundation and other agencies. The authors would like to acknowledge high performance computing support from Cheyenne (doi:10.5065/D6RX99HX) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation. Users can access the data used in this study via the website links provided in Tables 1 and S1 without any restrictions. Model data produced by this study can be accessed via the public repository (https://osf.io/gbw3r/). The authors declare no conflict of interest.

References

Bales, R. C., Molotch, N. P., Painter, T. H., Dettinger, M. D., Rice, R., & Dozier, J. (2006). Mountain hydrology of the western United States. *Water Resources Research*, *42*, W08432. https://doi.org/10.1029/2005WR004387

Barlage, M., Chen, F., Tewari, M., Ikeda, K., Gochis, D., Dudhia, J., et al. (2010). Noah land model modifications to improve snowpack prediction in the Colorado Rocky Mountains. *Journal of Geophysical Research*, *115*, D22101. https://doi.org/10.1029/2009JD013470

Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, *438*(7066), 303–309. https://doi.org/10.1038/nature04141

Berg, N., & Hall, A. (2017). Anthropogenic warming impacts on California snowpack during drought. *Geophysical Research Letters*, *44*, 2511–2518. https://doi.org/10.1002/2016GL072104

Betts, A., Desjardins, R., Worth, D., Wang, S., & Li, J. (2014). Coupling of winter climate transitions to snow and clouds over the Prairies. *Journal of Geophysical Research: Atmospheres*, *119*, 1118–1139. https://doi.org/10.1002/2013JD021168

Bohn, T. J., Livneh, B., Oyler, J. W., Running, S. W., Nijssen, B., & Lettenmaier, D. P. (2013). Global evaluation of MTCLIM and related algorithms for forcing of ecological and hydrological models. *Agricultural and Forest Meteorology*, *176*, 38–49. https://doi.org/10.1016/j.agrformet.2013.03.003

Broxton, P. D., Dawson, N., & Zeng, X. (2016). Linking snowfall and snow accumulation to generate spatial maps of SWE and snow depth. *Earth and Space Science*, *3*, 246–256. https://doi.org/10.1002/016EA000174

Carroll, T., Cline, D., Fall, G., Nilsson, A., Li, L., & Rost, A. (2001). NOHRSC operations and the simulation of snow cover properties for the conterminous U.S. Proc. 69thAnnualMeeting of the Western Snow Conf., Sun Valley, ID,Western Snow Conference, 14 pp. Retrieved from www.westernsnowconference.org/sites/westernsnowconference.org/PDFs/2001Carroll.pdf

Cayan, D. R., Das, T., Pierce, D. W., Barnett, T. P., Tyree, M., & Gershunov, A. (2010). Future dryness in the southwest US and the hydrology of the early 21st century drought. *Proceedings of the National Academy of Sciences*, *107*(50), 21,271–21,276. https://doi.org/10.1073/pnas.0912391107

Chen, F., Barlage, M., Tewari, M., Rasmussen, R., Jin, J., Lettenmaier, D., et al. (2014). Modeling seasonal snowpack evolution in the complex terrain and forested Colorado Headwaters region: A model intercomparison study. *Journal of Geophysical Research: Atmospheres*, *119*, 13,795–13,819. https://doi.org/10.1002/2014JD022167

Chen, F., Liu, C., Dudhia, J., & Chen, M. (2014). A sensitivity study of high resolution regional climate simulations to three land surface models over the western United States. *Journal of Geophysical Research: Atmospheres*, *119*, 7271–7291. https://doi.org/10.1002/2014JD021827

Chen, F., Manning, K. W., LeMone, M. A., Trier, S. B., Alfieri, J. G., Roberts, R., et al. (2007). Description and evaluation of the characteristics of the NCAR high-resolution Land data assimilation system during IHOP-02. *Journal of Applied Meteorology and Climatology*, *46*(6), 694–713. https://doi.org/10.1175/JAM2463.1

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

Clark, M. P., & Slater, A. G. (2006). Probabilistic quantitative precipitation estimation in complex terrain. *Journal of Hydrometeorology*, *7*, 3–22. https://doi.org/10.1175/JHM474.1

Clow, D. W., Nanus, L., Verdin, K. L., & Schmidt, J. (2012). Evaluation of SNODAS snow depth and snow water equivalent estimates for the Colorado Rocky Mountains, USA. *Hydrological Processes*, *26*(17), 2583–2591. https://doi.org/10.1002/hyp.9385

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

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., et al. (2008). Physiographically-sensitive mapping of temperature and precipitation across the conterminous United States. *International Journal of Climatology*, *28*(15), 2031–2064. https://doi.org/10.1002/joc.1688

Dee, D. P., Uppala, S. M., & Simmons, et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. https://doi.org/10.1002/qj.828

Dickinson, R. E. (1983). Land surface processes and climate-surface albedos and energy balance. In B. Saltzman (Ed.), *Theory of Climate, Adv. Geophys* (Vol. 25, pp. 305–353). San Diego, CA: Academic.

Dutra, E., Kotlarski, S., Viterbo, P., Balsamo, G., Miranda, P. M. A., Schär, C., et al. (2011). Snow cover sensitivity to horizontal resolution, parameterizations, and atmospheric forcing in a land surface model. *Journal of Geophysical Research*, *116*, D21109. https://doi.org/10.1029/2011JD016061

Ebert, E. E., Janowiak, J. E., & Kidd, C. (2007). Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bulletin of the American Meteorological Society*, *88*, 47–64. https://doi.org/10.1175/BAMS-88-1-47

Essery, R., Rutter, N., Pomeroy, J., Baxter, R., Stahli, M., Gustafsson, D., et al. (2009). SNPWMIP2, An evalutation of froest snow process simulations. *Bulletin of the American Meteorological Society*, *90*, 1120–1135.

Flanner, M. G., Shell, K. M., Barlage, M., Perovich, D. K., & Tschudi, M. A. (2011). Radiative forcing and albedo feedback from the Northern Hemisphere cryosphere between 1979 and 2008. *Nature Geoscience*, *4*(3), 151–155. https://doi.org/10.1038/ngeo1062

Flanner, M. G., Zender, C. S., Randerson, J. T., & Rasch, P. J. (2007). Present-day climate forcing and response from black carbon in snow. *Journal of Geophysical Research*, *112*, D11202. https://doi.org/10.1029/2006JD008003

Gao, Y., Li, K., Chen, F., Jiang, Y., & Lu, C. (2015). Assessing and improving Noah-MP land model simulations for the central Tibetan Plateau. *Journal of Geophysical Research: Atmospheres*, *120*, 9258–9278. https://doi.org/10.1002/2015JD023404

Garvert, M. F., Smull, B., & Mass, C. (2007). Multiscale mountain waves influencing a major orographic precipitation event. *Journal of the Atmospheric Sciences*, *64*(3), 711–737. https://doi.org/10.1175/JAS3876.1

Gergel, D. R., Nijssen, B., Abatzoglou, J. T., Lettenmaier, D. P., & Stumbaugh, M. R. (2017). Effects of climate change on snowpack and fire potential in the western USA. *Climatic Change*, *141*(2), 287–299. https://doi.org/10.1007/s10584-017-1899-y

Godsey, S. E., Kirchner, J. W., & Tague, C. L. (2014). Effects of changes in winter snowpacks on summer low flows: Case studies in the Sierra Nevada, California, USA. *Hydrological Processes*, *28*(19), 5048–5064. https://doi.org/10.1002/hyp.9943

Gowan, T. M., Steenburgh, W. J., & Schwartz, C. S. (2018). Validation of mountain precipitation forecasts from the convection-permitting NCAR Ensemble and operational forecast systems over the Western United States. *Weather and Forecasting*, *33*(3), 739–765. https://doi.org/10.1175/WAF-D-17-0144.1

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*(1), 262–281. https://doi.org/10.1175/2011JCLI4109.1

Hall, D. K., & Riggs, G. A. (2016). MODIS/Terra Snow Cover Daily L3 Global 0.05Deg CMG, Version 6. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/MODIS/MOD10C1.006, Last Access: January 3, 2019.

He, C., Liou, K. N., Takano, Y., Yang, P., Qi, L., & Chen, F. (2018). Impact of grain shape and multiple black carbon internal mixing on snow albedo: Parameterization and radiative effect analysis. *Journal of Geophysical Research: Atmospheres*, *123*, 1253–1268. https://doi.org/10.1002/2017JD027752

He, C., Takano, Y., & Liou, K.-N. (2017). Close packing effects on clean and dirty snow albedo and associated climatic implications. *Geophysical Research Letters*, *44*, 3719–3727. https://doi.org/10.1002/2017GL072916

He, C., Takano, Y., Liou, K.-N., Yang, P., Li, Q., & Chen, F. (2017). Impact of snow grain shape and black carbon-snow internal mixing on snow optical properties: Parameterizations for climate models. *Journal of Climate*, *30*(24), 10,019–10,036. https://doi.org/10.1175/JCLI-D-17-0300.1

Hedrick, A., Marshall, H. P., Winstral, A., Elder, K., Yueh, S., & Cline, D. (2015). Independent evaluation of the SNODAS snow depth product using regional-scale lidar-derived measurements. *The Cryosphere*, *9*(1), 13–23.

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.

Hong, S. Y., Noh, Y., & Dudhia, J. (2006). A new vertical diffusion package with an explicit treatment of entrainment processes. *Monthly Weather Review*, *134*(9), 2318–2341.

Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., et al. (2014). The global precipitation measurement mission. *Bulletin of the American Meteorological Society*, *95*(5), 701–722. https://doi.org/10.1175/BAMS-D-13-00164.1

Houze, R. A. (2012). Orographic effects on precipitating clouds. *Reviews of Geophysics*, *50*, RG1001. https://doi.org/10.1029/2011RG000365

Huning, L. S., & AghaKouchak, A. (2018). Mountain snowpack response to different levels of warming. *Proceedings of the National Academy of Sciences*, *115*(43), 10,932–10,937. https://doi.org/10.1073/pnas.1805953115

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., & Collins, W. D. (2008). Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research*, *113*, D13103. https://doi.org/10.1029/2008JD009944

Ikeda, K., Rasmussen, R., Liu, C., Gochis, D., Yates, D., Chen, F., et al. (2010). Simulation of seasonal snowfall over Colorado. *Atmospheric Research*, *97*, 462–477. https://doi.org/10.1016/j.atmosres.2010.04.010

Jimenez, P. A., Dudhia, J. J., Gonzalez-Rouco, F., Navarro, J., Montavez, J. P., & Garcia-Bustamante, E. (2012). A revised scheme for the WRF surface layer formulation. *Monthly Weather Review*, *140*, 898–918. https://doi.org/10.1175/MWR-D-11-00056.1

Jing, X., Geerts, B., Wang, Y., & Liu, C. (2017). Evaluating seasonal orographic precipitation in the interior western United States using gauge data, gridded precipitation estimates, and a regional climate simulation. *Journal of Hydrometeorology*, *18*, 2541–2558. https://doi.org/10.1175/JHM-D-17-0056.1

Jordan, R. (1991), A one-dimensional temperature model for a snow cover, Spec. Rep. 91–16, Cold Reg. Res. and Eng. Lab., U.S. Army Corps of Eng., Hanover, N. H.

Kim, D.-E., & Park, S. K. (2019). Uncertainty in predicting the Eurasian snow: Intercomparison of land surface models coupled to a regional climate model. *The Cryosphere Discussion*. https://doi.org/10.5194/tc-2019-15

Leung, L. R., & Qian, Y. (2003). The sensitivity of precipitation and snowpack simulations to model resolution via nesting in regions of complex terrain. *Journal of Hydrometeorology*, *4*, 1025–1043.

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

Liu, C., Ikeda, K., Rasmussen, R., Barlage, M., Newman, A. J., Prein, A. F., et al. (2017). Continental-scale convection-permitting modeling of the current and future climate of North America. *Climate Dynamics*, *49*(1-2), 71–95. https://doi.org/10.1007/s00382-016-3327-9

Liu, C., Ikeda, K., Thompson, G., Rasmussen, R., & Dudhia, J. (2011). High-resolution simulations of wintertime precipitation in the Colorado Headwaters region: Sensitivity to physics parameterizations. *Monthly Weather Review*, *139*, 3533–3553. https://doi.org/10.1175/MWR-D-11-00009.1

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*, 150042. https://doi.org/10.1038/sdata.2015.42

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*. https://doi.org/10.1175/BAMS-D-19-0001.1

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

Marshall, A. M., Abatzoglou, J. T., Link, T. E., & Tennant, C. J. (2019). Projected changes in interannual variability of peak snowpack amount and timing in the Western United States. *Geophysical Research Letters*, *46*, 8882–8892. https://doi.org/10.1029/2019GL083770

Meyer, J. D. D., Jin, J.-M., & Wang, S.-Y. (2012). Systematic patterns of the inconsistency between snow water equivalent and accumulated precipitation as reported by the snowpack telemetry network. *Journal of Hydrometeorology*, *13*, 1970–1976.

Minder, J. R., Letcher, T. W., & Liu, C. (2018). The character and causes of elevation-dependent warming in high-resolution simulations of Rocky Mountain climate change. *Journal of Climate*, *31*, 2093–2113. https://doi.org/10.1175/JCLI-D-17-0321.1

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

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

Mote, P. W., Li, S., Lettenmaier, D. P., Xiao, M., & Engel, R. (2018). Dramatic declines in snowpack in the western US. *Climate and Atmospheric Science*, *1*(1), 2.

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

Musselman, K. N., Clark, M. P., Liu, C., Ikeda, K., & Rasmussen, R. (2017). Slower snowmelt in a warmer world. *Nature Climate Change*, *7*(3), 214.

Musselman, K. N., Lehner, F., Ikeda, K., Clark, M. P., Prein, A. F., Liu, C., et al. (2018). Projected increases and shifts in rain-on-snow flood risk over Western North America. *Nature Climate Change*, *8*(9), 808–812. https://doi.org/10.1038/s41558-018-0236-4

National Ice Center (2008). updated daily. IMS Daily Northern Hemisphere Snow and Ice Analysis at 1 km, 4 km, and 24 km Resolutions, Version 1. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center. https://doi.org/10.7265/N52R3PMC. Last Access: January 3, 2019.

Newman, A. J., Clark, M. P., Craig, J., Nijssen, B., Wood, A., Gutmann, E., et al. (2015). Gridded ensemble precipitation and temperature estimates for the contiguous United States. *Journal of Hydrometeorology*, *16*, 2481–2500. https://doi.org/10.1175/JHM-D-15-0026.1

Niu, G.-Y., & Yang, Z.-L. (2004). The effects of canopy processes on snow surface energy and mass balances. *Journal of Geophysical Research*, *109*, D23111. https://doi.org/10.1029/2004JD004884

Niu, G.-Y., & Yang, Z.-L. (2007). An observation-based formulation of snow cover fraction and its evaluation over large North American river basins. *Journal of Geophysical Research*, *112*, D21101. https://doi.org/10.1029/2007JD008674

Niu, G. Y., Yang, Z. L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., et al. (2011). The community Noah land surface model with multi-parameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research*, *116*, D12109. https://doi.org/10.1029/2010JD015139

O'Brien, T. A., Collins, W. D., Kashinath, K., Rubel, O., Byna, S., Gu, J., et al. (2016). Resolution dependence of precipitation statistical fidelity in hindcast simulations. *Journal of Advances in Modeling Earth Systems*, *8*, 976–990. https://doi.org/10.1002/2016MS000671

Oleson, K. W., Dai, Y., Bonan, B., Bosilovichm, M., Dickinson, R., Dirmeyer, P., et al. (2004). Technical description of the Community Land Model (CLM), Tech. Rep. NCAR TN-461+STR, National Center for Atmospheric Research.

Pederson, G. T., Gray, S. T., Woodhouse, C. A., Betancourt, J. L., Fagre, D. B., Littell, J. S., et al. (2011). The unusual nature of recent snowpack declines in the North American Cordillera. *Science*, *333*(6040), 332–335. https://doi.org/10.1126/science.1201570

Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., et al. (2015). A review on regional convection-permitting climate modeling: demonstrations, prospects, and challenges. *Reviews of Geophysics*, *53*, 323–361. https://doi.org/10.1002/2014RG000475

Rasmussen, R., Ikeda, K., Liu, C., Gochis, D., Clark, M., Dai, A., et al. (2014). Climate change impacts on the water balance of the Colorado headwaters: High-resolution regional climate model simulations. *Journal of Hydrometeorology*, *15*(3), 1091–1116. https://doi.org/10.1175/JHM-D-13-0118.1

Rasmussen, R., Liu, C., Ikeda, K., Gochis, D., Yates, D., Chen, F., et al. (2011). High-resolution coupled climate runoff simulations of seasonal snowfall over Colorado: A process study of current and warmer climate. *Journal of Climate*, *24*(12), 3015–3048. https://doi.org/10.1175/2010JCLI3985.1

Rhoades, A. M., Jones, A. D., & Ullrich, P. A. (2018). The Changing Character of the California Sierra Nevada as a Natural Reservoir. *Geophysical Research Letters*, *45*, 13,008–13,019. https://doi.org/10.1029/2018GL080308

Rhoades, A. M., Ullrich, P. A., & Zarzycki, C. M. (2018). Projecting 21st Century Snowpack Trends in Western USA Mountains Using Variable-Resolution CESM. *Climate Dynamics*, *50*(1-2), 261–288. https://doi.org/10.1007/s00382-017-3606-0

Rhoades, A. M., Ullrich, P. A., Zarzycki, C. M., Johansen, H., Margulis, S. A., Morrison, H., et al. (2018). Sensitivity of mountain hydroclimate simulations in variable-resolution CESM to microphysics and horizontal resolution. *Journal of Advances in Modeling Earth Systems*, *10*, 1357–1380. https://doi.org/10.1029/2018MS001326

Risser, M. D., Paciorek, C. J., Wehner, M. F., O'Brien, T. A., & Collins, W. D. (2019). A probabilistic gridded product for daily precipitation extremes over the United States. *Climate Dynamics*, *53*(5-6), 2517–2538. https://doi.org/10.1007/s00382-019-04636-0

Romanov, P., & Tarpley, D. (2004). Estimation of snow depth over open prairie environments using GOES imager observations. *Hydrological Processes*, *18*, 1073–1087. https://doi.org/10.1002/hyp.5508

Rutz, J. J., Steenburgh, W. J., & Ralph, F. M. (2014). Climatological characteristics of atmospheric rivers and their inland penetration over the western United States. *Monthly Weather Review*, *142*(2), 905–921.

Scalzitti, J., Strong, C., & Kochanski, A. (2016). Climate change impact on the roles of temperature and precipitation in western U.S. snowpack variability. *Geophysical Research Letters*, *43*, 5361–5369. https://doi.org/10.1002/2016GL068798

Schaaf, C., & Wang, Z. (2015). MCD43C3 MODIS/Terra+Aqua BRDF/Albedo Albedo Daily L3 Global 0.05Deg CMG V006. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MCD43C3.006. Last Access: January 3, 2019.

Serreze, M. C., Clark, M. P., Armstrong, R. L., McGinnis, D. A., & Pulwarty, R. S. (1999). Characteristics of the western United States snowpack from snowpack telemetery (SNOTEL) data. *Water Resources Research*, *35*, 2145–2160. https://doi.org/10.1029/1999WR900090

Strachan, S., & Daly, C. (2017). Testing the daily PRISM air temperature model on semiarid mountain slopes. *Journal of Geophysical Research: Atmospheres*, *122*, 5697–5715. https://doi.org/10.1002/2016JD025920

Sun, F., Berg, N., Hall, A., Schwartz, M., & Walton, D. (2019). Understanding end-of-century snowpack changes over California's Sierra Nevada. *Geophysical Research Letters*, *46*, 933–943. https://doi.org/10.1029/2018GL080362

Thompson, G., & Eidhammer, T. (2014). A study of aerosol impacts on clouds and precipitation development in a large winter cyclone. *Journal of the Atmospheric Sciences*, *71*(10), 3636–3658. https://doi.org/10.1175/JAS-D-13-0305.1

Timmermans, B., Wehner, M., Cooley, D., O'Brien, T., & Krishnan, H. (2019). An evaluation of the consistency of extremes in gridded precipitation data sets. *Climate Dynamics*, *52*(11), 6651–6670.

Ullrich, P. A., Xu, Z., Rhoades, A. M., Dettinger, M. D., Mount, J. F., Jones, A. D., & Vahmani, P. (2018). California's drought of the future: A midcentury recreation of the exceptional conditions of 2012-2017. *Earth's Future*, *6*(11), 1568–1587. https://doi.org/10.1029/2018EF001007

Verseghy, D. L. (1991). CLASS-A Canadian land surface scheme for GCMS: I. Soil model. *International Journal of Climatology*, *11*, 111–133. https://doi.org/10.1002/joc.3370110202

Wang, Y., Geerts, B., & Liu, C. (2018). A 30-year convection-permitting regional climate simulation over the interior western United States. Part I: Validation. *International Journal of Climatology*, *38*(9), 3684–3704.

Westrick, K., Mass, C., & Colle, B. (1999). The limitations of the WSR-88D radar network for quantitative precipitation measurement over the coastal western United States. *Bulletin of the American Meteorological Society*, *80*, 2289–2298.

Wiscombe, W. J., & Warren, S. G. (1980). A model for the spectral albedo of snow: I. Pure snow. *Journal of the Atmospheric Sciences*, *37*, 2712–2733. https://doi.org/10.1175/1520-0469(1980)037%3C2712:AMFTSA%3E2.0.CO;2

Wrzesien, M. L., Durand, M. T., Pavelsky, T. M., Howat, I. M., Margulis, S. A., & Huning, L. S. (2017). Comparison of methods to estimate snow water equivalent at the mountain range scale: A case study of the California Sierra Nevada. *Journal of Hydrometeorology*, *18*, 1101–1119. https://doi.org/10.1175/JHM-D-16-0246.1

Xia, Y., Mitchell, K., Ek, M., Cosgrove, B., Sheffield, J., Luo, L., et al. (2012). Continental-scale water and energy flux analysis and validation for North American Land Data Assimilation System project phase 2 (NLDAS-2): 2. Validation of model-simulated streamflow. *Journal of Geophysical Research*, *117*, D03110. https://doi.org/10.1029/2011JD016051

Yang, D., Goodison, B. E., Metcalfe, J. R., Golubev, V. S., Bates, R., Pangburn, T., & Hanson, C. L. (1998). Accuracy of NWS 8″ standard nonrecording precipitation gauge: results and application of WMO intercomparison. *Journal of Atmospheric and Oceanic Technology*, *15*, 54–68.

Yang, Z.-L., Dickinson, R., Robock, A., & Vinnikov, K. Y. (1997). Validation of snow submodel of the biosphere-atmosphere transfer scheme with Russian snow cover and meteorological observational data. *Journal of Climate*, *10*, 353–373.