Evaluating Input Data and Rain Snow Separation Improvements to the National Water Model Simulation of Snow Water Equivalent

Irene Garousi-Nejad[1,1] and David Tarboton[2,2]

[1]Stanford University
[2]Utah State University

November 30, 2022

Abstract

We compared snowfall, and snow water equivalent (SWE) accumulation and ablation simulations from the WRF-Hydro model with the U.S. National Water Model (NWM) configuration against observations at a set of representative point locations from Snow Telemetry (SNOTEL) sites across the western U.S. We focused on the model's partitioning of precipitation between rain and snow and selected sites that span the variability of the percentage of rain on snow precipitation events. Our results show that the NWM generally under-estimates SWE and tends to melt snow earlier than observations in part due to errors in the precipitation and air temperature inputs. We reduced some of the observed and modeled discrepancies by using SNOTEL snow-adjusted precipitation and removing air temperature biases, based on observations. These input changes produced an average 59% improvement in the peak SWE. Modeled peak SWE was further improved using humidity-dependent rain-snow-separation. Both dew point and wet-bulb parameterizations were evaluated, with the dew-point parameterization giving better overall improvement, reducing the bias in SWE by 18% compared to the NWM air temperature-based scheme. This modification also improved melt timing with the number of site years having difference between modeled and observed date of half melt from peak SWE six or more days reduced by 6%. These SWE magnitude and timing improvements varied when analyzed for each rain-on-snow percentage class, with generally better results at sites where most precipitation events fall either as snow or as rain, and less improvement when there is a mix of snow and rain-on-snow events.

Evaluating Input Data and Rain Snow Separation Improvements to the National Water Model Simulation of Snow Water Equivalent

I. Garousi-Nejad[1] and D. G. Tarboton[1]

[1]Department of Civil and Environmental Engineering, Utah Water Research Laboratory, Utah State University, Logan, Utah 84322.

Corresponding author: Irene Garousi-Nejad (irene.garousi.nejad@gmail.com)

Key Points:

- The National Water Model (NWM), in general, under-estimates snow water equivalent due to both model errors and inputs errors.
- Using observed precipitation and bias-corrected air temperature improved the general downward bias in NWM snow water equivalent.
- NWM snow processes were further improved by using a dew-point based rain-snow separation scheme.

Abstract

We compared snowfall, and snow water equivalent (SWE) accumulation and ablation simulations from the WRF-Hydro model with the National Water Model (NWM) configuration against observations at a set of representative point locations from Snow Telemetry (SNOTEL) sites across the western U.S. We focused on the model's partitioning of precipitation between rain and snow and selected sites that span the variability of the percentage of rain on snow precipitation events. Our results show that the NWM generally under-estimates SWE and tends to melt snow earlier than observations in part due to errors in the precipitation and air temperature inputs. We reduced some of the observed and modeled discrepancies by using SNOTEL snow-adjusted precipitation and removing air temperature biases, based on observations. These input changes produced an average 59% improvement in the peak SWE. Modeled peak SWE was further improved using humidity-dependent rain-snow-separation. Both dew point and wet-bulb parameterizations were evaluated, with the dew-point parameterization giving better overall improvement, reducing the bias in SWE by 18% compared to the NWM air temperature-based scheme. This modification also improved melt timing with the number of site years having difference between modeled and observed date of half melt from peak SWE six or more days reduced by 6%. These SWE magnitude and timing improvements varied when analyzed for each rain-on-snow percentage class, with generally better results at sites where most precipitation events fall either as snow or as rain, and less improvement when there is a mix of snow and rain-on-snow events.

Plain Language Summary

In snow dominated regions, modeling the partitioning of input precipitation between rain and snow is important for flood prediction and water resources management. The National Water Model (NWM) includes equations to model this partitioning and the resultant snow accumulation and melt in national scale water forecasts. This paper compared NWM snow partitioning with observations at Snow Telemetry sites and found that the NWM generally under-estimates snow water equivalent (SWE) and tends to melt snow earlier than observations. This was due to both errors in the precipitation and air temperature inputs and inaccuracies in the precipitation partitioning. We identified that improving inputs of temperature and precipitation has the potential to produce 59% improvement in the modeling of peak SWE. We also evaluated alternative precipitation partitioning approaches based on dew point or wet bulb temperature, rather than simply air temperature, and found that the dew-point based approach that we evaluated reduced the bias in SWE by 18%. There were also improvements in the predicted melt timing that accrued from SWE magnitude being better modeled. The findings thus document the benefits for improved model inputs and better physically-based process representations and suggest these as opportunities for the operational forecasts to be improved.

1 Introduction

Snow models are a central component of hydrologic forecasting systems when snow and snowmelt are the dominant influence on the regional streamflow. Decades of model development, combined with advances in technology and software engineering, have gradually enabled snowmelt runoff models to evolve into large-scale, high-resolution, and physically-based distributed models such as the National Oceanic and Atmospheric Administration (NOAA) National Water Model (NWM) in the U.S. (https://water.noaa.gov/about/nwm). This evolution was driven in part by the need to shorten the time interval for streamflow forecasts; to accommodate the shift from simple temperature-index based to energy balance methods; and to enable predicting the effects of anthropogenic and environmental changes such as those caused by land-use change or climate change on large heterogeneous basins (DeWalle & Rango, 2008). The NWM is now part of NOAA's water resources information system that provides timely hydrologic forecasts and data to support and inform emergency services and water resources decisions (https://water.noaa.gov).

To provide accurate predictions of seasonal water supplies over the continental U.S. under future changing conditions, the NWM, operated by the National Water Center, uses an energy balance model (Noah-MP) to solve the surface energy and water balances based on first principles of conservation of energy and mass to calculate snowmelt (Gochis, Barlage, Cabell, Dugger, et al., 2020; Niu et al., 2011). In our previous work, we compared the Noah-MP models as implemented in the NWM version 2.0 retrospective simulations with snow observations at Snow Telemetry (SNOTEL) sites over the western U.S. and showed that the NWM generally underestimated snow water equivalent (SWE) early in the season and became progressively more biased later in the season compared to observations at SNOTEL sites, in part due to errors in inputs, notably precipitation and air temperature (Garousi-Nejad & Tarboton, 2022a). However, the discrepancies in model inputs were not the only sources of SWE differences. The SWE bias was persistent when the model precipitation input was relatively (statistically) close to the observed precipitation, suggesting that there were challenges in the current snow parameterization within the specific configuration of Noah-MP as implemented in the NWM version 2.0 retrospective configuration. We identified the current air temperature-dependent rain-snow-separation (RSS) parameterization within Noah-MP as a potential source of model error in SWE modeling, because this has been reported by other studies as a limitation of Noah-MP as used in the NWM (Chen et al., 2014; Liu et al., 2017; Wang et al., 2019). More generally, the accurate representation of RSS in hydrological models is important as the proportion of rainfall versus snowfall across mountainous regions changes, altering snowpack dynamics, streamflow timing and amount, and frequency of rain-on-snow events (Bales et al., 2006; Barnett et al., 2005; Gillies et al., 2012; Harpold et al., 2017; Knowles et al., 2006). Thus, research that evaluates the NWM performance and enhances model output accuracy through more realistic inputs and physics representations is essential. This motivated our focus on the NWM's partitioning of precipitation between rain and snow at sites selected to span the variability of precipitation events that were rain on snow present in the western U.S.

We addressed the following questions in this study:

- ***Question 1***. To what degree are discrepancies in NWM SWE and RSS predictions due to input errors and how much could they potentially be improved if inputs were better?
- ***Question 2***. How well does the NWM RSS (rainfall and snowfall separation) parameterization work in comparison to SNOTEL observations?
- ***Question 3***. Do any other RSS parameterization methods yield more accurate snowfall compared to SNOTEL observations?
- ***Question 4***. Does incorporating a statistically better RSS scheme into the NWM translate into appreciable improvements in modeling of SWE?
- ***Question 5***. How do improvements in modeled SWE vary over sites grouped according to the percentage of precipitation events that are rain-on-snow?

In what follows, we first review prior literature used in this work (Section 2). We then describe the data and model we used (Section 3) followed by the method and numerical experiment design developed to answer our research questions (Section 4). We then compare gridded model results from each scenario simulated with point-scale measurements across the western U.S. (Section 5). Following that, we discuss limitations and uncertainties associated with the data and model providing perspective on the results presented and identifying areas for input data improvement and model enhancements (Section 6). Finally, we summarize our conclusions (Section 7) and provide links to data we used and codes we developed.

2 Background

Seasonal mountain snowpack has key implications for mid-to high-latitude regions such as the western U.S., storing water in the winter when snow falls and then releasing it as runoff in spring and summer when the snow melts and contributes (up to about 70%) to the total runoff in these regions (Li et al., 2017). The recently published Intergovernmental Panel on Climate Change (IPCC) report indicates a 0.29 million km^2 per decade decline in April snow cover extent—commonly used as an indicator of water supply forecast for the following spring and summer season—in the Northern Hemisphere (Gulev et al., 2021). It is projected that seasonal snowpack decline will decrease water supplies for about 2 billion people this century (Mankin et al., 2015). In the western U.S., an average 30% decrease in areal extent of winter wet-day temperatures conducive to snowfall is projected (Klos et al., 2014). Given snowpack decline due to climate warming and its impact on water resources, accurate prediction of spring snowmelt will become increasingly important as the growing population demands more water and as operational agencies have to manage water under hydroclimate conditions outside of the historical record (Bhatti et al., 2016; Gergel et al., 2017; Mote, 2003; Mote et al., 2005).

Continued changes in the precipitation phase (rainfall, snowfall, or a mixture of both) are expected to alter snowpack dynamics, streamflow timing and amount, and frequency of rain-on-snow events; and thus present a new set of challenges for hydrologic modeling (Harpold et al., 2017; Musselman et al., 2018). RSS is one of the most sensitive parameterizations in simulating cold-region hydrological processes (Loth et al., 1993) and has a notable influence on the success of snowmelt models (Rutter et al., 2009). Despite advances in snowmelt modeling, most models rely on empirical algorithms based on air temperature to separate precipitation into rain and snow. For example, see the model comparison by Wen et al. (2013). These methods are empirical and ignore some of the physical processes involved in atmospheric formation of rain or snow where humidity and latent heat exchanges between a hydrometeor and the surrounding air play a role (Feiccabrino et al., 2015; Jennings et al., 2018). Such physical process representations warrant consideration if models are to improve their predictability by reducing their dependence on empirical parameterizations.

Inaccurate RSS may result in errors in SWE, snow depth, and snow cover duration at both point and basin scale (Harder & Pomeroy, 2014; Wang et al., 2019) because snow can be produced in air temperatures slightly above freezing if the wet-bulb temperature (the temperature to which air is cooled by evaporating water into the air at constant pressure until it is saturated) is below about -2 ^{o}C (Stull, 2011). Ultimately, these errors propagate into the hydrological response (runoff and streamflow) of the watershed and land-atmosphere energy exchanges (Jennings et al., 2018; Mizukami et al., 2013). Some studies suggest that using dew point temperature, wet-bulb temperature, or psychrometric energy balance based RSS schemes, which consider the impact of atmospheric humidity in the energy budget of falling hydrometeors, improves the modeling of precipitation phase and the accuracy of partitioning between rain and snow (Behrangi et al., 2018; Harder & Pomeroy, 2013; Marks et al., 2013).

While there has been significant prior work on RSS, our goal was to evaluate the NWM snow model performance across a set of SNOTEL sites that are representative of various precipitation regimes (dominantly rainfall or snowfall, or rain-on-snow) across the western U.S., and to identify where model biases can be removed by using a more physically accurate RSS method. The RSS methods that we used here include the air temperature-based method from Jordan (1991) currently used in the NWM, the air temperature-based method developed by the U.S. Army Corps of Engineers (1956) as used in the Utah Energy Balance (UEB) model (Tarboton & Luce, 1996), the dew point temperature-based method used in the SNOBAL model (Marks et al., 1999), and the wet-bulb temperature-based approach evaluated for the Variable Infiltration Capacity (Behrangi et al., 2018) and Noah-MP (Wang et al., 2019) models.

3 Data and Model

We used SNOTEL data, NWM input data, and an offline version of the WRF-Hydro model that serves as the basis for the NWM to evaluate different RSS parameterizations and their corresponding impact on the modeled SWE as detailed in the three subsections that follow.

3.1 SNOTEL Data

For more than 60 years, the automated SNOTEL network, currently consisting of 808 sites across the western U.S., has measured SWE using a pressure sensing snow pillow, precipitation (P) using a storage-type gage or tipping bucket, and air temperature (Ta) using a shielded thermistor sensor to monitor winter snow and inform spring and summer water supply forecasts. Our study used the daily snow-adjusted precipitation (start of the day) that accounts for uncertainty associated with snowfall measurements being subject to under-catch (Mote, 2003; Sun et al., 2019). We also used daily average air temperature and daily SWE (start of the day) at SNOTEL sites as a reference dataset to evaluate: (1) the snowfall fraction estimated from four different RSS parameterization methods, and (2) the accuracy of the NWM inputs (precipitation and air temperature) and outputs (SWE).

We recognize there are uncertainties associated with SNOTEL measurements that need to be considered in our analysis. However, SNOTEL provides the most comprehensive dataset we could obtain to explore our research questions because of its long, historically continuous records of P, Ta, and SWE across the western U.S. For our analysis, we focused on SNOTEL sites where complete daily data were available for water years 2008-2020. This led to a set of 683 SNOTEL sites. Even though it would have been technically possible to set up simulations and run WRF-Hydro for all 683 sites, it would have been computationally prohibitive, and we decided to focus on a representative set of them for this research. To select a representative subset of SNOTEL sites, we used a random sampling within rain-on-snow classes that led to a group of 33 sites that spanned site rain-on-snow variability, described later, and for which we set up simulations and ran WRF-Hydro.

3.2 National Water Model Input Data

The NWM surface physiographic and atmospheric meteorological inputs (1 km spatial resolution and hourly temporal resolution) were made available to us by the NCAR team (D.

Gochis and A. RafieeiNasab, personal communication, March 16, 2021) as a read only directory in the NCAR Cheyenne high-performance computer. The surface physiographic inputs included the model domain; initial conditions such as soil moisture, soil temperature, and snow states; geospatial inputs (such as topography, soil properties, land cover type, etc.) and parameter files (such as calibrated snowmelt factor used in calculation of the snow-covered area fraction). The meteorological inputs included the Analysis of Record for Calibration reanalysis dataset developed by NOAA National Weather Service (Kitzmiller et al., 2018; National Weather Service, Office of Water Prediction, 2021), hereafter referred to as AORC. AORC forcing data included incoming short- and longwave radiation, specific humidity, wind, air pressure, air temperature, and precipitation rate.

For each of the selected 33 SNOTEL sites we retrieved all required inputs for a four grid cell 2 km by 2 km area containing the SNOTEL site (Garousi-Nejad & Tarboton, 2022b). Then, we transferred data from Cheyenne to Expanse, an eXtreme Science and Engineering Discovery Environment (XSEDE) supercomputer (Towns et al., 2014) where we ran WRF-Hydro. The first water year (2008) was used for model spin up and, while the SNOTEL data extended to 2020, NWM forcing data was not available for 2020 at the time this work was done. Therefore, we used the period 2009-2019 for model comparisons.

3.3 WRF-Hydro National Water Model Configuration Code

The NWM is a physically-based, distributed model based on the WRF-Hydro modeling framework (Gochis, Barlage, Cabell, Dugger, et al., 2020) that provides operational hydrological forecasts at 1 km spatial and hourly temporal resolution for snow across the entire continental U.S. The NWM has evolved beginning from version 1.0 (August 2016) to the current version 2.1 (October 2021) with improved soil/snow physics, calibration, and data assimilation. The core of the NWM system is WRF-Hydro, developed by the National Center for Atmospheric Research (NCAR), which consists of different modules with different geospatial representation (e.g., grids in the land surface and terrain routing modules connected to stream reaches in the channel routing module) and resolution (e.g., 1 km in the land surface module versus 250 m in the terrain routing module) to simulate land and atmosphere energy/water fluxes and storages. Details about the NWM and WRF-Hydro are available in Gochis, Barlage, Cabell, Casali, et al. (2020). We obtained the Fortran source code from the WRF-Hydro GitHub webpage (https://github.com/NCAR/wrf_hydro_nwm_public/releases/tag/v5.1.1, version 5.1.1 corresponding to the NWM version 2.0 available at the time this work started (Gochis, Barlage, Cabell, Dugger, et al., 2020). Releases beyond this to date include WRF-Hydro version 5.1.2 and version 5.2.0, both available in GitHub(https://github.com/NCAR/wrf_hydro_nwm_public/releases), but to our understanding the rain and snow separation parameterization that we evaluated has not been changed in these releases.

In this study, we focused on the land surface module of the NWM, which is a particular configuration of the Noah-MP model (Niu et al., 2011), where all snow processes are simulated within a 1-dimensional vertical column over 1 km spatial resolution grid cells. The Noah-MP module uses up to three snow layers to solve the energy balance (Equation 1) and water balance (Equation 2) between the snowpack, atmosphere, and the ground surface. The snow state variables for each snow layer are the mass of liquid water, the mass of ice, layer thickness, and layer temperature.

$$\frac{dU}{dt} = Q_{sw} + Q_{lw} + Q_{lt} + Q_{sn} + Q_g + Q_p + Q_m \quad (1)$$

$$\frac{dSWE}{dt} = P_{snow} - M - E \quad (2)$$

where U is the snowpack internal sensible and latent heat storage, t is time, Q_{sw} is net shortwave radiation flux, Q_{lw} is net longwave radiation flux, Q_{lt} is convective latent heat of vaporization/sublimation flux, Q_{sn} is convective sensible heat flux, Q_g is conductive ground heat flux, Q_m is heat of fusion energy flux due to meltwater leaving the snowpack (which is solved for as a residual in Equation 1), P_{snow} is the snowfall (in terms of water depth) that reaches the ground after adjusting for canopy interception, M is the meltwater, and E is snow sublimation/frost (Shuttleworth, 2012).

4 Methods and Numerical Experiment Design

4.1 Input Data Evaluation

The first step in our work was to compare the NWM inputs (elevation, P, and T_a for water years 2009-2019) with observations at representative SNOTEL sites. Results showed biases in model inputs that needed to be considered in the analysis. There were discrepancies of up to approximately 250 m between model elevation and the elevation of SNOTEL sites (Figure 1a). This may be a contributor to differences observed in the daily mean air temperature comparison due to the lapse rate (Figure 1b).

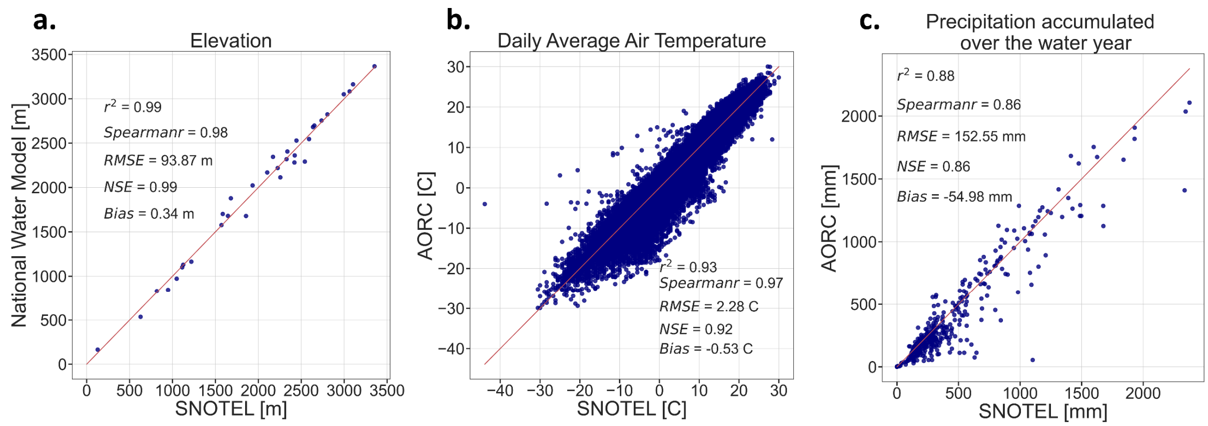

Figure 1. (a) NWM elevation inputs compared to SNOTEL site elevations (each point is a SNOTEL site), (b) AORC mean daily temperature compared to mean measurements at SNOTEL sites (each point is a day for a SNOTEL site during the 2009-2019 water years) excluding incorrect AORC air temperatures (see Figure 2), and (c) AORC annual precipitation compared to observations at SNOTEL sites (each point represents total precipitation during a water year at a SNOTEL site). Statistical metrics on graphs are coefficient of determination (r^2), Spearman's rank correlation (Spearmanr), root mean square error (RMSE), Nash Sutcliffe efficiency (NSE), and bias (Bias) for which equations are provided in Table 1.

For some years, we found artifacts in the air temperature inputs at three SNOTEL sites (Figure 2). After excluding these periods, we observed a negative bias (-0.53 ^{o}C) in AORC air temperatures compared to SNOTEL measurements (Figure 1b), meaning that T_a input to the

NWM is generally colder than observations. There were no artifacts in AORC precipitation for the period of our study; however, we observed a downward bias of about -55 mm (Figure 1c) when comparing the annual precipitation (accumulated from October 1 through September 30 for each water year at each representative SNOTEL site). These observations were the basis for designing our initial numerical experiments (scenarios), where we attempted to reduce biases in model inputs (details are provided in Scenario 2 and Scenario 3 in Section 4.5).

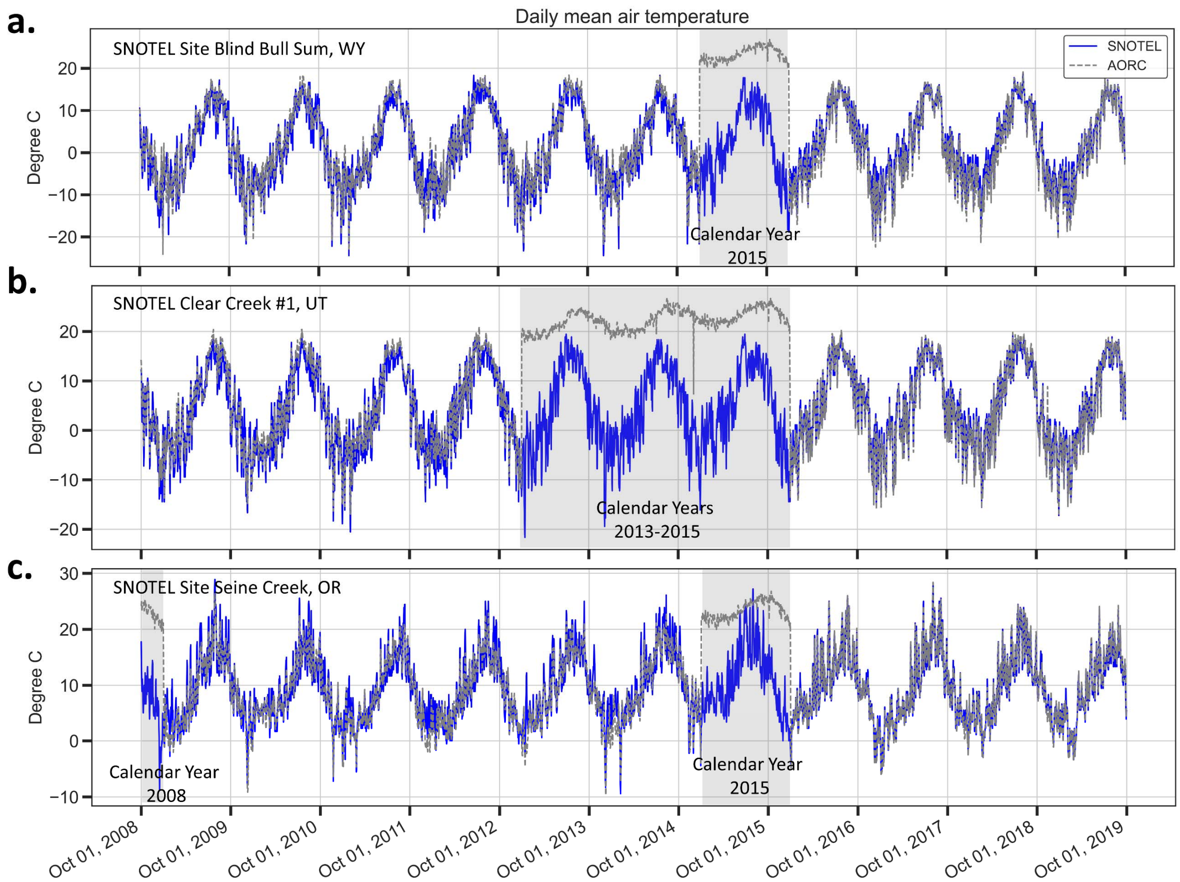

Figure 2. AORC and SNOTEL daily mean air temperature during 2009-2019 water years at (a) Blind Bull Sum SNOTEL site in Wyoming, (b) Clear Creek #1 SNOTEL site in Utah, and (c) Seine Creek SNOTEL site in Oregon with gray regions showing periods that AORC air temperature appear to be obviously incorrect. We considered these as artifacts and excluded these periods from our analysis.

4.2 Snow Rain Ratio

Evaluating simulated snowfall amounts from different RSS schemes is challenging due to the lack of reliable ground truth observations of the precipitation phase (Harpold et al., 2017). The Natural Resources Conservation Service (NRCS) reports a snow rain ratio (SNRR) for SNOTEL sites that estimates the fraction of precipitation that falls as snowfall calculated as the ratio of daily SWE increases to daily P for the same period. In theory, the SNRR should range from 0 to 1, with 1 indicating all precipitation falls as snowfall. We obtained daily SNRR values from NRCS Report Generator version 2 for 683 SNOTEL sites for water years 2008-2020 using a Jupyter Notebook script we developed (Garousi-Nejad & Tarboton, 2022b). We realized that this ratio was sometimes above 1 (100%) because it was calculated based on the daily P measurements which may be less than accumulated daily SWE. This may occur due to either precipitation measurement under-catch or processes that result in additional SWE being measured, such as snow drifting. The NRCS provides a snow-adjusted daily P estimate to account for this. We obtained this adjusted P and recalculated SNRR to get values within the range 0-1 (Algorithm 1). We used the computed SNRR values as a validation dataset to compare different rain/snow separation parameterizations. We acknowledge that there are uncertainties associated with this SNRR approach that may impact our analysis. However, this indicator was the best option available to us for evaluating RSS methods given the western-U.S.-wide dataset that we use in this study.

Algorithm 1. Snow rain ratio (SNRR) Calculation. P is the total precipitation and SWE is the snow water equivalent at the start of day. The index t and t+1 indicate the start and the end of the period (day).

```
If P_t > 0:

    // If there is an increase in SWE during the period,
    // compute SNRR
    If SWE_{t+1} - SWE_t > 0:
        SNRR_t = (SWE_{t+1} - SWE_t) / P_t
    else:
    // If there is a decrease in SWE during the period,
    // SNRR should be 0 due to the rain melting the snow
        SNRR_t = 0
else:
   // SNRR cannot be computed because there
   // is no precipitation to separate into rain and snow
    SNRR_t = nan
```

4.3 Representative SNOTEL Site Selection

We used the computed SNRR values to identify precipitation events that were rain-on-snow and classified sites based the percentage of rain-on-snow events they received to obtain a set to work with that spanned and is thus representative of the variability of rain-on-snow event percentages present across the western U.S. We designated precipitation events with SNRR >= 0.95 as snowfall and events with SNRR < 0.95 as rain-on-snow. We, thus, took rainfall or mixed rainfall and snowfall events for which SNRR < 0.95 as having a quantity of rain sufficient to be called rain-on-snow. We calculated the percentage of precipitation events that were rain-on-snow (ROS%) for each SNOTEL site over water years 2008-2020 using a script we developed (Garousi-Nejad & Tarboton, 2022b). For the 683 SNOTEL sites, ROS% values ranged between 30-100% (Figure 3a). We classified sites according to ROS% into seven groups each spanning a 10% class range. The largest number of sites fell in the 50-60% class, and the least frequent group (three sites) had ROS% between 90-100%.

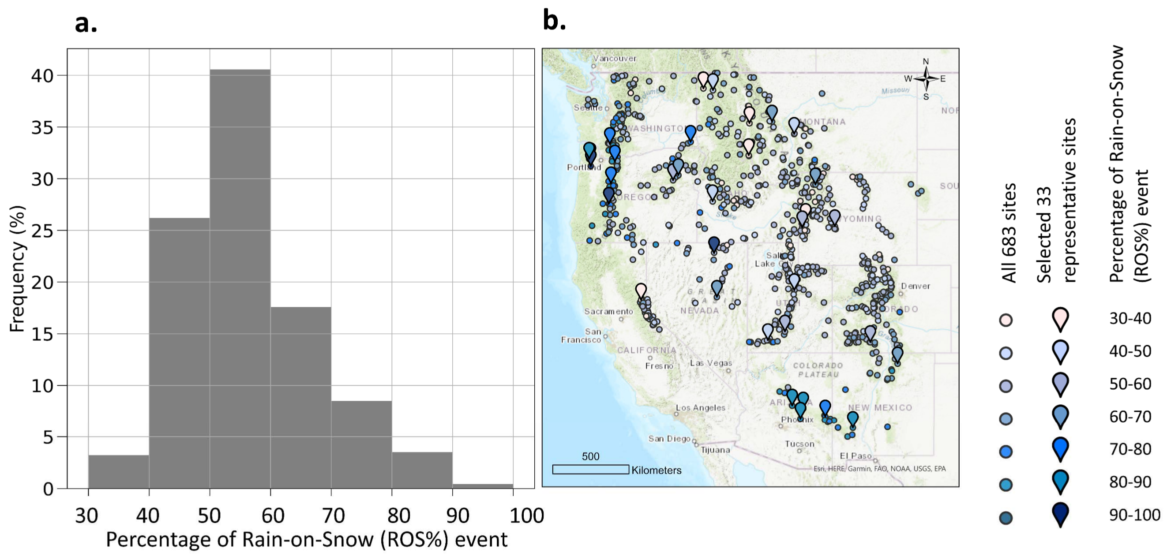

Figure 3. (a) Histogram of the percentage of historical Rain-on-Snow (ROS%) events inferred from the computed SNRR over SNOTEL sites (total of 683 sites) with data for 2008-2020 water years across the western U.S. (b) Location of representative SNOTEL sites selected based on the ROS%.

To select the representative set of SNOTEL sites to work with, we randomly selected five sites from each class with ROS% between 30-90% and selected all members within the 90-100% class because it contained only three SNOTEL sites using a script we developed (Garousi-Nejad & Tarboton, 2022b). This yielded a subset of 33 SNOTEL sites with different ROS% values spread across the western U.S. (Figure 3b). We obtained observed P, T_a, and SWE for these selected SNOTEL sites from NRCS Report Generator version 2 using Jupyter Notebook data retrieval scripts we developed (Garousi-Nejad & Tarboton, 2022b).

4.4 Evaluation of Rain-Snow-Separation (RSS) Parameterizations

We evaluated four different RSS schemes, including two air temperature-dependent and two humidity-dependent approaches, commonly used in hydrological models. The air temperature-based RSS schemes were from the U.S. Army Corps of Engineers, (U.S. Army Corps of Engineers, 1956; hereafter USCAE (1956)) as used in the UEB snow model (Tarboton & Luce, 1996), and Jordan (1991) as used in the current version of the NWM Noah-MP. The USACE (1956) T_a based method separates precipitation into rain and snow based on two temperature thresholds. All precipitation is rainfall if the air temperature is greater than or equal to 3 ^oC, snowfall if the air temperature is less than or equal to -1 ^oC, and varies linearly for air temperature between -1 and 3 (Algorithm 2). The Jordan (1991) T_a based method uses multiple thresholds (0.5, 2, and 2.5 ^oC) to separate precipitation into rain and snow (Algorithm 3). Both these methods only consider air temperature (Figure 4a, 4b).

Algorithm 2. Rain snow separation (RSS) scheme based on USACE (1956). T_a is air temperature in degree C and f_s is the fraction of snowfall.

```
If T_a >= 3:
    f_s = 0
else if T_a <= -1:
    f_s = 1
else:
    f_s = 1 - (T_a - (-1)) / (3 - (-1))
```

Algorithm 3. Rain snow separation (RSS) scheme based on Jordan (1991). T_a is air temperature in degree K, T_f is the freezing point in degree K, and f_s is the fraction of snowfall.

```
// Physical constants and parameters required
T_f = 273.16

If T_a >= T_f + 2.5:
    f_s = 0
else:
    f_s = 1
    if T_a <= T_f + 0.5:
        f_s = 1
    else if T_a <= T_f + 2:
        f_s = 1 - (-54.632 + 0.2 T_a)
    else:
        f_s = 0.6
```

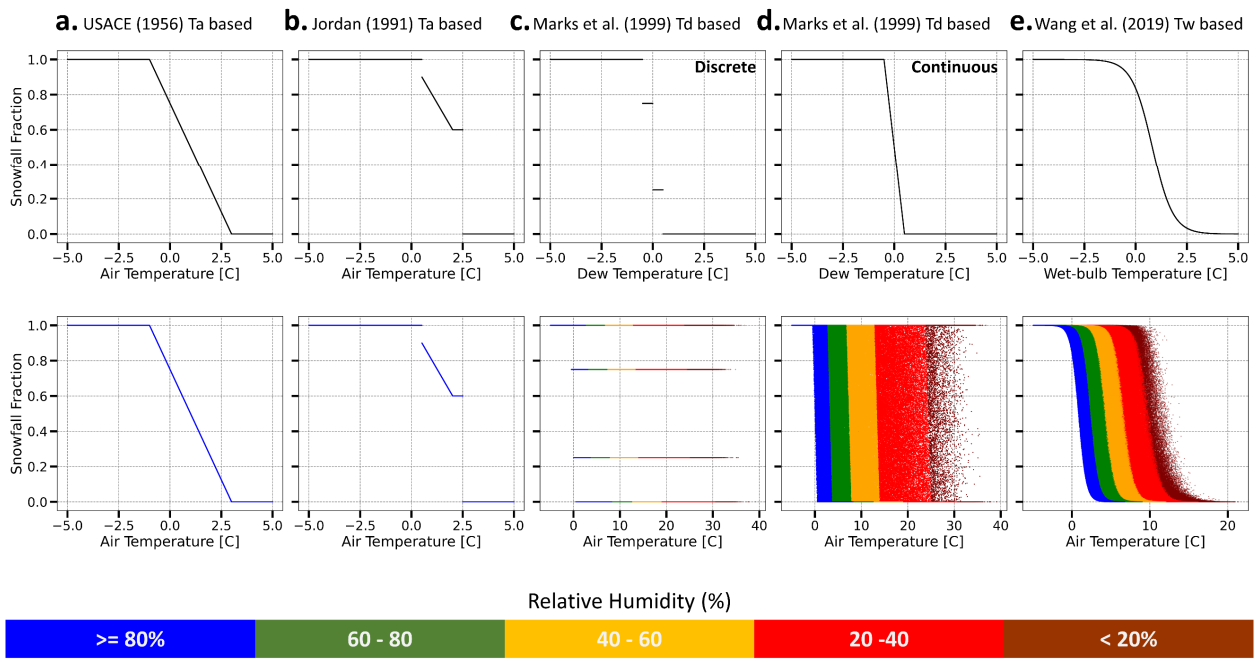

Figure 4. Snowfall fraction computed for the 33 SNOTEL sites using the observed precipitation and the NWM inputs (including air pressure, specific humidity, and bias-corrected air temperature) based on (a) USACE (1956), (b) Jordan (1991), (c) Marks et al. (1999): discrete version, (d) Marks et al. (1999): continuous version and (e) Wang et al. (2019) RSS methods. The plots on the top row show the relationship between snowfall fraction as a function of air temperature (T_a), dew point (T_d), or wet-bulb (T_w) temperature depending on the method. The plots on the bottom row illustrate the relationship between snowfall fraction and air temperature for all methods. The colors represent data with different relative humidity values.

The humidity-based RSS approaches were from the dew point temperature method (Marks et al., 1999) as used in the SNOBAL model and the wet-bulb temperature based method evaluated for Noah-MP (Wang et al., 2019). Dew point temperature (T_d), a measure of the vapor pressure of the air (Equation 3), is defined as the temperature to which air must cool at constant pressure for it to saturate, without any moisture addition/removal (Marks et al., 2013; Shuttleworth, 2012):

$$T_d = \frac{\ln(e) + 0.49299}{0.0707 - 0.00421\ln(e)} \qquad (3)$$

where e is the vapor pressure of the air in kPa and T_d is the dew point temperature in ^{o}C.

Marks et al. (1999) described a dew point based approach that uses discrete steps to partition precipitation into rain and snow (Figure 4c, Algorithm 4). The discrete stepped nature of the approach seemed limiting as there do not appear to be physical reasons for such step changes. We thus developed a continuous version of Marks et al.'s (1999) method to provide a smoother function of T_d (Figure 4d).

Algorithm 4. Rain snow separation (RSS) scheme based on Marks et al. (1991). e is the vapor pressure of the air in kPa, P_{air} is the air pressure in kPa, q is specific humidity kg/kg, T_d is dew point temperature in degree C, and f_s is the fraction of snowfall.

// Compute the vapor pressure of the air from
// Shuttleworth (2012) Equation 2.8
$e = (P_{air}\ q) / (0.622 + 0.378\ q)$

// Compute T_d from Shuttleworth (2012) Equation 2.21
$T_d = (\ln(e) + 0.49299) / (0.0707 - 0.00421 \ln(e))$

// Discrete version: compute snowfall fraction based on
// T_d from Marks et al. (1999) Table 1.
If $T_d < -0.5$:
    $f_s = 1$
else if $-0.5 <= T_d < 0$:
    $f_s = 0.75$
else if $0 <= T_d < 0.5$:
    $f_s = 0.25$
else:
    $f_s = 0$

// Continuous version: compute snowfall fraction using a
// continuous version of Marks et al. (1999) Table 1
If $T_d < -0.5$:
    $f_s = 1$
else if $-0.5 <= T_d < 0.5$:
    $f_s = 0.5 - T_d$
else:
    $f_s = 0$

Wet-bulb temperature (T_w) is defined as the temperature to which air is cooled by evaporating water into the air at constant pressure until it is saturated ($T_a \approx T_d \approx T_w$). According to thermodynamic laws, the air is thermally isolated in saturated environments. In other words, as the air cools to get to the saturation point, the heat (internal energy) removed from the air due to the cooling process must equal the latent heat required to evaporate water (from the hydrometeor surface in a precipitation event) to raise the specific humidity of the air to saturation (Shuttleworth, 2012). This can be mathematically represented as Equation (4) which can be reformulated as the wet-bulb equation (Equation 5):

$$\rho_a V (T_a - T_w) c_p = \rho_a \left[q_{sat}(T_w) - q \right] V \lambda \tag{4}$$

$$es_w(T_w) - e = \frac{c_p\ P_{air}}{0.622\ \lambda} (T_a - T_w) \tag{5}$$

where ρ_a is air density (kg/m^3), V is volume of air (m^3), T_a is (dry-bulb) air temperature (K), T_w is wet-bulb temperature (K), c_p is specific heat at constant pressure for air (1.04 kJ/kg K), $q_{sat}(T_w)$ is saturated specific humidity of air at T_w (kg/kg), q is specific humidity of air (kg/kg), λ is latent heat of vaporization (2.5 MJ/kg), $es_w(T_w)$ is the saturated vapor pressure of air at T_w (kPa), and P_{air} is air pressure (kPa). Equation (5) does not have an analytical inverse solution to calculate the wet-bulb temperature from air temperature and humidity (Stull, 2011), so was solved numerically using a Newton-Raphson scheme. We then used the sigmoid function of Wang et al. (2019) to calculate RSS (Algorithm 5).

Algorithm 5. Rain snow separation (RSS) scheme based on Wang et al. (2019). T_f is freezing point in degree K, cp is heat capacity of vaporization in j/kg, L_v is latent heat of vaporization in j/kg, NITER is number of iterations to iteratively solve the T_w equation, T_a is air temperature in degree K, P_{air} is air pressure in Pa, q is specific humidity in kg/kg, gamma is the psychrometric constant in Pa, e is the vapor pressure of the air in Pa, es_a is the saturated vapor pressure at T_a in Pa, RH is relative humidity, T_w is wet-bulb temperature in degree C, es_w is the saturated vapor pressure at T_w in Pa, and fs is the fraction of snowfall. Note that constant values are the same as used in the NWM Noah-MP code.

```
// Physical constants and parameters required
Tf = 273.16
cp = 1004.64
Lv = 2.5104E06
NITER = 20

Tc = Ta - Tf  // Kelvin to Celsius
gamma = (cp Pair) / (0.622 Lv)
e = (Pair q) / (0.622 + 0.378 q)
esa = 610.8 exp ((17.27 Tc) / (237.3 + Tc))
RH = e/es

if RH > 100:
    Tw = Tc
    esw = 610.8 exp ((17.27 Tw) / (237.3 + Tw))
else:
    Tw = Tc - 5                      // First guess for Tw to start the iterative method
    for i in range (1, NITER):       // Use Newton-Raphson method:
      esw = 610.8 exp ((17.27 Tw) / (237.3 + Tw))
      F = Tw - Tc + (1 / gamma) (esw – e)
      Fprim = 1 + (1 / gamma) (esw) [17.27 / (237.3 + Tw) - (17.27 Tw) / (237.3 + Tw)
         **2]
      Tw = Tw - F / Fprim            // Update Tw

      // Check the stopping criteria
      if ABS (F / Fprim) <= 0.01:
         break
    Tw = max ( -50, Tw)

// Compute fs using Wang et al. (2019) approach
A = 6.99*10**(-5)
B = 2
C = 3.97
fs = 1 / (1 +A exp (B (Tw + C)))
```

4.5 RSS Modeling Experimental Design

We developed a set of modeling scenarios to answer the research questions given earlier. For each of the 33 representative SNOTEL sites selected, we used the WRF-Hydro version 5.1.1 NWM configuration in the following scenarios:

1. **Base scenario with AORC inputs**. The hourly AORC forcing data was used to simulate snow processes from January 2008 to September 2019 (with the first nine months being set aside as model spin up) over 33 grid cells containing the representative SNOTEL sites. We call this scenario the base scenario as we kept all inputs and model settings the same as those used in the operational NWM version 2.0. The outputs that we evaluated are hourly snowfall (from the Jordan (1991) RSS scheme) and SWE values.

2. **Replacing AORC precipitation with observations from SNOTEL (Observed precipitation scenario)**. Scenario 2 was the same as the base scenario except for the input precipitation. In our preparation step (Section 3.3), we showed a downward bias for AORC precipitation compared to observations at SNOTEL sites. To isolate the effects of AORC precipitation biases on modeled snowfall and SWE, we used the SNOTEL observed precipitation as supplemental precipitation to run the model. This means that the model used all other AORC inputs, but the precipitation data were read from the additional forcing inputs. To generate supplemental precipitation input files, we followed the steps described in Gochis et al. (2020). We resampled observed daily precipitation into hourly precipitation by dividing the total daily precipitation from SNOTEL sites equally into 24 hours using scripts we developed (Garousi-Nejad & Tarboton, 2022b).

3. **Replacing AORC air temperature with bias corrected air temperature based on SNOTEL on top of the precipitation adjustments of Scenario 2 (Bias-corrected temperature scenario)**. Since we observed a negative bias in AORC air temperature compared to SNOTEL observations, we designed Scenario 3 to diminish the impact of errors in air temperature on the modeled snowfall and SWE. For each SNOTEL site we computed the average difference in daily temperature for the common data period (12 years) and used this difference to adjust the AORC hourly temperature inputs. This one difference value thus served as a bias correction offset for each representative SNOTEL site. The model physics settings were the same as in Scenarios 1 and 2, and precipitation was from SNOTEL observations (as prepared in Scenario 2).

4. **Inputs prepared for Scenario 3 but with USACE (1956) air temperature RSS modifications to the code**. In this scenario, we used inputs prepared for Scenario 3 to run the WRF-Hydro model modified to use the USACE (1956) air temperature-based RSS scheme (Algorithm 2). This was achieved by editing the rain snow separation code in the module_noahmplsm.F source code file and recompiling the model.

5. **Inputs prepared for Scenario 3 but with continuous dew point based RSS based on Marks et al. (1999)**. In this scenario, we used inputs prepared for Scenario 3 to run the WRF-Hydro model modified to implement the continuous version of the Marks et al. (1999) dew point based RSS method (Algorithm 4).

This was also achieved by editing the rain snow separation code in the module_noahmplsm.F source code file and recompiling the model.

6. **Inputs prepared for Scenario 3 but with Wang et al. (2019) wet-bulb based RSS**. In this scenario, we used inputs prepared for Scenario 3 and implemented the Wang et al. (2019) wet-bulb based RSS parametrization (Algorithm 5) in the NWM code as for scenarios 4 and 5.

4.6 Comparing Snow Accumulation and Melt

To assess the performance of the model, we first compared the computed snowfall amount from each RSS method and quantified the performance of each approach against observed RSS that was inferred from SNRR at SNOTEL sites through a set of statistical metrics, including Coefficient of Determination (r^2), Spearman's Rank Correlation (Spearmanr), Root Mean Square Error (RMSE), Nash Sutcliffe Efficiency (NSE), and Bias (Table 1). In addition to these statistical metrics, we used (1) SWE on observed peak date, (2) observed and modeled peak SWE, and (3) date of half melt from peak SWE metrics to compare the simulated SWE to observed SWE at SNOTEL sites (Garousi-Nejad & Tarboton, 2022b). First, we used the date on which peak SWE was observed to compare modeled SWE against observations. We refer to this comparison metric as a same-day comparison. Note that if there is a discrepancy in timing, model and observed peak SWE may be similar, while the model SWE on the observed peak date is different. To account for this the second metric compared observed and modeled peak SWE regardless of the dates when they occur. This is referred to as a different-day comparison in this study. This comparison may have limitations due to cumulative precipitation inputs being different up to the different dates. We did not report comparison of the Peak SWE timing because of variability associated with peak SWE time related to long periods where the SWE time series was flat near the peak. Instead, we chose the date of half melt from peak SWE as a metric to quantify the model's performance in terms of simulating the melt timing (Clow, 2010). This is the date (either modeled or observed) when half of the peak SWE has melted. To quantitively assess the difference between the modeled and observed half melt dates, we categorized the date differences into four groups—close, model early, model late, and far apart (Garousi-Nejad & Tarboton, 2022b). Close indicates that modeled and observed half melt dates are within 5 days of each other. Model early refers to the situation where modeled half melt dates are 6 to 19 days before observed, while model late means that modeled half melt dates are 6 to 19 days after observed. Lastly, far apart means that modeled an observed half melt dates are more than 20 days apart.

Table 1. Common statistical metrics used in this study to compare model inputs and outputs versus observations[†].

Name	Equation	Range	Description
Coefficient of determination (r^2)	$r^2 = \left(\frac{\sum_{t=1}^{N}(O_t - \overline{O_t})(M_t - \overline{M_t})}{\sqrt{\sum_{t=1}^{N}(O_t - \overline{O_t})^2 \sum_{t=1}^{N}(M_t - \overline{M_t})^2}}\right)^2$	-1 to 1 with 1 indicating a perfect positive linear relationship	Measures the linear relationship. Insensitive to proportional differences between modeled and observed data.
Spearman's rank correlation (Spearmanr)	$\text{Spearmanr} = 1 - \frac{6\sum_{t=1}^{N} d_t^2}{N(N^2 - 1)}$	-1 to 1 with 1 indicating a perfect positive correlation	Measures the strength of association between modeled and observed values.
Root mean squared error (RMSE)	$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{N}(O_t - M_t)^2}{N}}$	Depends on the variable with the best value of 0.	Measures how concentrated the data are around the line of best fit.
Nash Sutcliffe efficiency (NSE)	$\text{NSE} = 1 - \frac{\sum_{t=1}^{N}(O_t - M_t)^2}{\sum_{t=1}^{N}(O_t - \overline{O_t})^2}$	-infinity to 1 with 1 indicating observed and modeled data fits the 1:1 line	Determines the relative magnitude of the residual variance compared to observed values.
Bias	$\text{Bias} = \frac{\sum_{t=1}^{N}(M_t - O_t)}{N}$	Depends on the variable with the best value of 0.	Quantifies the average of the differences between modeled and observed values.

[†]M_t is model simulation, O_t is observation, t is time, N is the total number of simulations or observations, d_t is difference between observed and modeled rank, and the overbar indicates average.

5 Results

5.1 Changes in Snowfall

We compared the estimated annual snowfall magnitude from five different RSS methods with the observations inferred from SNRR from SNOTEL and found a persistent upward bias in snowfall from all methods (Figure 5). This is an average bias across all 33 sites and all years. USACE (1956) T_a based showed the smallest bias (about 6 mm) and Marks et al. (1999) T_d based (continuous version) had the most significant bias (about 45 mm). Results for Jordan (1991) T_a based (the current RSS scheme in the NWM Noah-MP) were slightly better than the dew point temperature-based (both discrete and continuous) methods (Figure 5b, 5c, and 5d). Among the two humidity-based methods, Wang et al. (2019) T_w based showed a smaller bias (more than 10 mm smaller), but its bias was still six times larger than USACE (1956) T_a based (Figure 5d and 5a).

The seasonal variations (11-year daily averages across selected SNOTEL sites) of accumulated snowfall from all methods indicated that more than 70% of the annual precipitation during February through May, independent of the RSS method, fell as snowfall averaged across the SNOTEL sites and water years (Figure 5f). Observations and USACE (1956) T_a based average accumulation matched well over the entire year. The other RSS methods tracked above observations and were all close together during the accumulation phase (October through May). Following May, Marks et al. (1999) T_d based (continuous version) diverged and produced more snowfall than other RSS methods and observations (50% more than observed in May). Also, Marks et al. (1999) T_d based was the only RSS method that showed 19% and 17% of precipitation falling as snowfall during July and September, respectively. This sets the Marks et al. (1999) T_d based method apart from other methods as the only one that estimated snowfall during warmer months (Figure 5f). Average air, wet-bulb, and dew point temperatures for each day across all site years indicated the general differences between these quantities that were inputs to the RSS methods (Figure 5g).

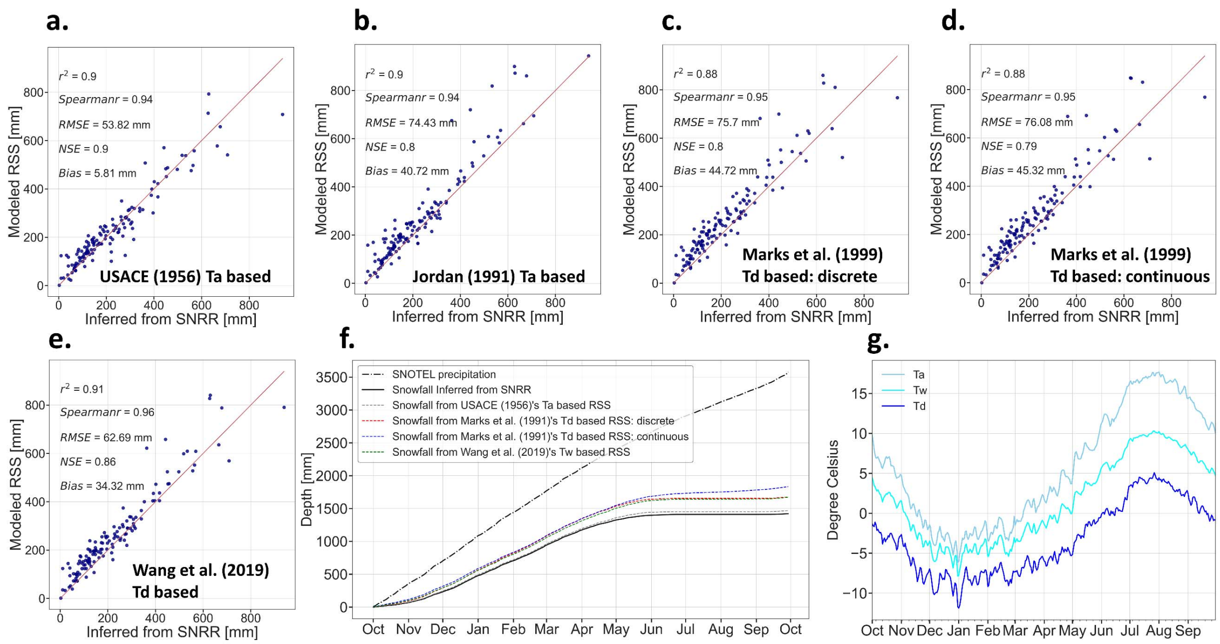

Figure 5. Analysis of annual snowfall estimated from different RSS schemes versus observations inferred from SNRR at SNOTEL sites for a period of 11 years (water years 2009-2019). (a) USACE (1956) air temperature-based RSS method versus SNRR, (b) Jordan (1991) air temperature-based RSS method (the current approach in the NWM version 2.0) versus SNRR, (c) Marks et al. (1999) dew point based (discrete version) RSS method versus SNRR, (d) Marks et al. (1999) dew point based (continuous version) RSS method versus SNRR, and (e) Wang et al. (2019) wet-bulb based RSS method versus SNRR. Each point in panels (a)-(e) represents a water year and a SNOTEL site. (f) The seasonal pattern of the long-term annual observed precipitation, observed snowfall inferred from SNRR, and modeled snowfall from all RSS schemes averaged across all sites and years. (g) Seasonal pattern of the long-term daily bias-corrected AORC air temperature (T_a) and computed wet-bulb (T_w) and dew point (T_d) temperatures using AORC data averaged across all sites and years.

5.2 Snow Water Equivalent on Observed Peak Date (Same-day Comparison)

The comparison between modeled and observed SWE on the date of observed peak SWE revealed a general downward bias in modeled SWE (Figure 6), suggesting that the NWM generally underestimated SWE on the date of observed peak SWE, independent of the model input errors (shown before in Figure 1) and model physics (specifically in terms of the different RSS methods as shown before in Figure 5). However, biases in modeled SWE were reduced when using observed precipitation instead of AORC precipitation, from -228 mm in the base scenario to -92 mm in the observed precipitation scenario (Figure 6b). This emphasizes the importance of using high-quality input forcing in the NWM. Even though we further reduced model input errors/biases by correcting the AORC air temperature biases, this did not improve SWE estimates (Figure 6c). Contrarily, it increased the downward bias in SWE. This should not be considered as a negative point as it is essential to have correct/accurate inputs, even though that may not necessarily translate into improvements in model outputs.

Even though our comparison of annual snowfall magnitude from different RSS methods (Figure 5) showed that USACE (1956) T_a based had the best agreement with observations, this agreement did not translate to the best same-day SWE comparison. Among the four RSS comparisons, when the best input estimates were used (Scenarios 3 to 6), USACE (1956) T_a based showed the largest negative bias (about -168 mm) and Marks et al. (1999) T_d based showed the least bias (about -111 mm) and best NSE and RMSE (Figure 6c, 6d, 6e, and 6f). Similar to the snowfall comparison, the modeled SWE from the current NWM RSS scheme (Jordan (1991) T_a based) and Wang et al. (2019) T_w based had almost statistically identical behavior when compared to SWE observations (Figure 6c versus 6f).

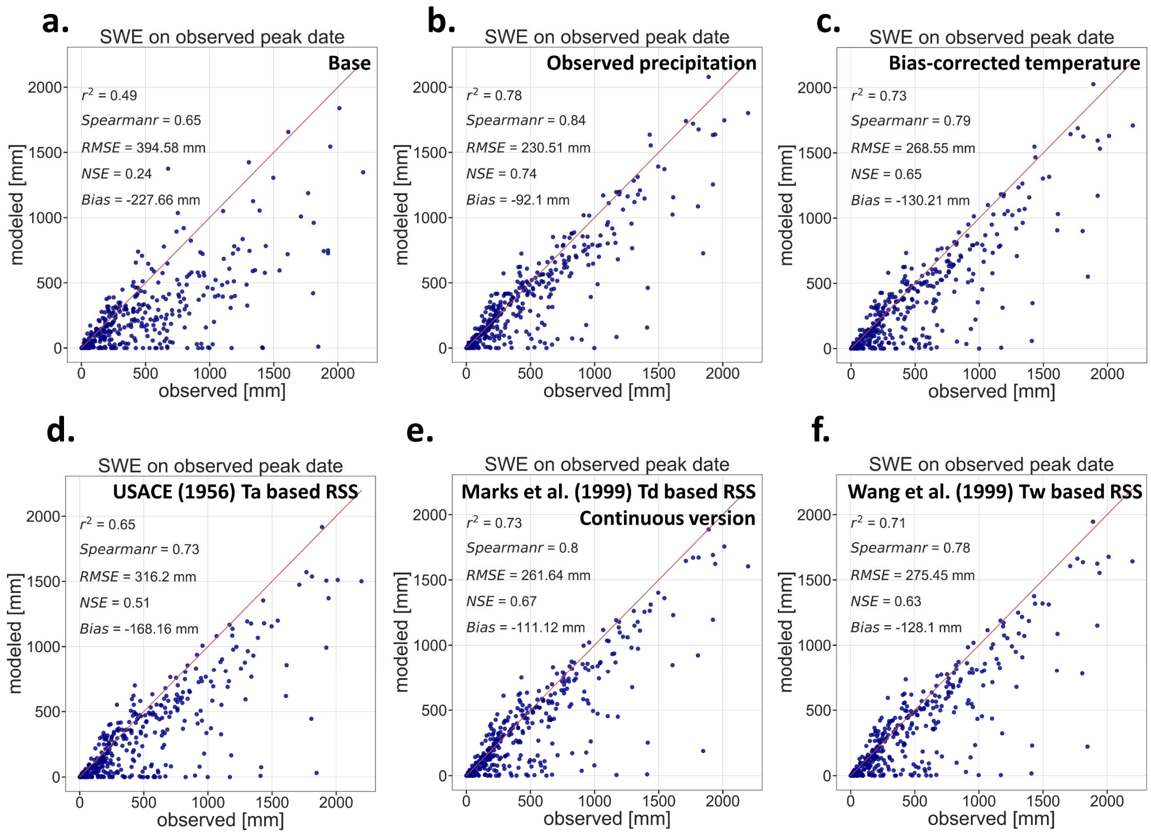

Figure 6. SWE Comparison on date of observed peak SWE. (a) NWM base scenario (Scenario 1) versus SNOTEL SWE, (b) NWM observed precipitation scenario (Scenario 2) versus SNOTEL SWE, (c) NWM bias-corrected temperature scenario (Scenario 3) versus SNOTEL SWE, (d) NWM using USACE (1956) air temperature (T_a) based RSS method (Scenario 4) versus SNOTEL SWE, (e) NWM using Marks et al. (1999) dew point (T_d) based (continuous version) RSS method (Scenario 5) versus SNOTEL SWE, (f) NWM using Wang et al. (2019) wet-bulb (T_w) based RSS method (Scenario 6) versus SNOTEL SWE. Each point on the graph represents a SNOTEL site and a water year.

5.3 Observed and Modeled Peak Snow Water Equivalent (Different-day Comparison)

Under-modeling of SWE was also evident in our comparison of observed and modeled peak SWE noting that the observed and modeled peak SWE do not necessarily occur on the exact same date (Figure 7). Among the four RSS schemes modeled (Scenarios 3 to 6) the dew point temperature-based scheme (Scenario 5) provided less biased modeled SWE similar to the same-day comparison. In general, these different day peak SWE comparisons had smaller error metrics than the comparisons presented above for the day of observed peak SWE.

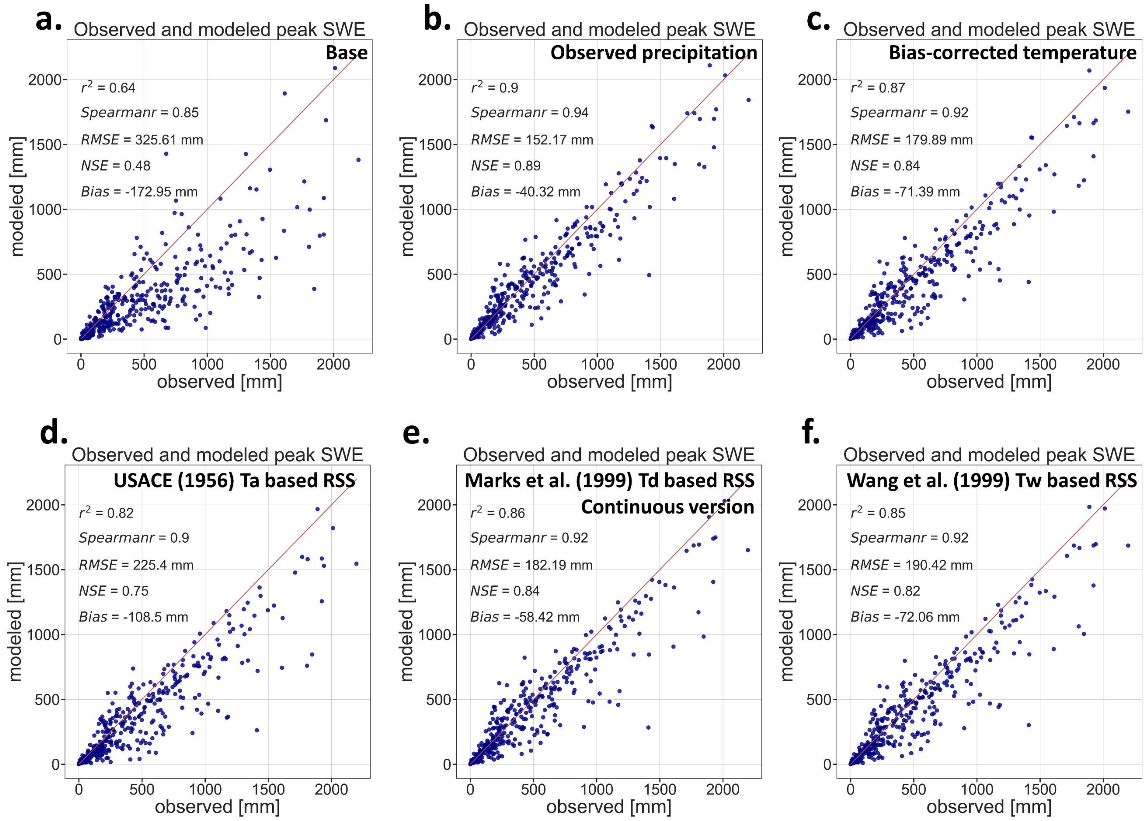

Figure 7. Observed and modeled peak SWE comparison (on the generally different dates they occur). (a) NWM base scenario (Scenario 1) versus SNOTEL SWE, (b) NWM observed precipitation scenario (Scenario 2) versus SNOTEL SWE, (c) NWM bias-corrected temperature scenario (Scenario 3) versus SNOTEL SWE, (d) NWM using USACE (1956) air temperature (T_a) based RSS method (Scenario 4) versus SNOTEL SWE, (e) NWM using Marks et al. (1999) dew point (T_d) based (continuous version) RSS method (Scenario 5) versus SNOTEL SWE, and (f) NWM using Wang et al. (2019) wet-bulb (T_w) based RSS method (Scenario 6) versus SNOTEL SWE. Each point on the graphs represents a SNOTEL site and a water year.

5.4 Seasonal Snow Water Equivalent

The seasonal pattern of SWE averaged across the representative SNOTEL sites indicated the general under-modeling of SWE relative to observations at SNOTEL sites in all scenarios, with USACE (1956) T_a based scheme (Scenario 3) being further apart from and Marks et al. (1999) T_d based scheme (Scenario 5) being the closest to the observations (Figure 8a). For the purpose of evaluating RSS options, we did not include results from scenarios that had inferior inputs (Scenarios 1 and 2) in this comparison. Furthermore, our results showed that discrepancies between seasonal patterns of SWE vary when analyzed for each ROS percentage class (Figure 8b-g). For SNOTEL sites with the smallest ROS% (30-40%, meaning that most precipitation events fall on average as snow), all RSS methods simulated almost identical SWE (Figure 8b). However, as ROS% increased, the impact of different RSS methods in modeling SWE became more evident in such a way that the T_d based RSS SWE simulations almost always stayed above the SWE from other RSS methods, meaning that it produced more SWE compared to other RSS methods. For the sites with ROS% between 80-100 (where rain-on-snow events are dominant), the T_d based RSS scheme simulated SWE was almost identical to observations during the accumulation period, October-March, while the other RSS methods underestimated SWE (Figure 8g). During the melt period all methods tended to melt the snow a bit slowly compared to observations, a difference likely due to model considerations other than RSS.

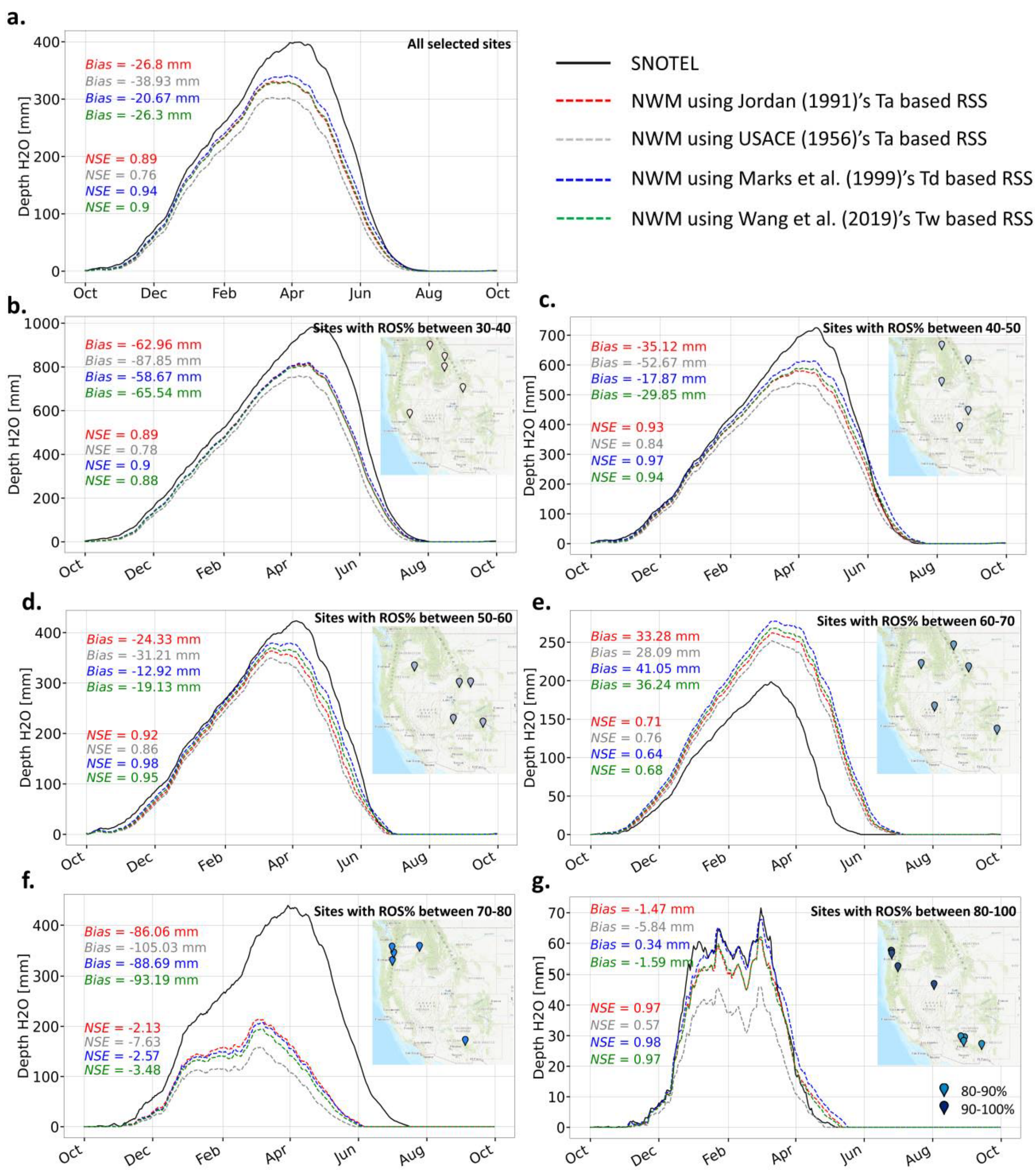

Figure 8. Observed and modeled SWE at the beginning of each date averaged across all years and (a) all selected SNOTEL sites, (b) sites with ROS% between 30-40%, (c) sites with ROS%

within 40-50%, (d) sites with ROS% within 50-60%, (e) sites with ROS% within 60-70%, (f) sites with ROS% within 70-80%, and (g) sites with ROS% within 80-100%.

5.5 Melt Timing Comparison (Half Melt from Peak Snow Water Equivalent Date)

Our comparison of the modeled half melt date (from scenarios that had valid inputs) with observations showed that the modeled half melt date was generally earlier than observations for more than 60% of the site-years (Table 2). When further classified depending on whether the differences between observed and modeled half melt dates from peak SWE were close, ahead, behind or far apart from observed melt dates, we observed that the NWM half melt date was off by 6 days or more for about 75% of site years (Figure 9a). This became even more noticeable when using the USACE (1956) T_a based RSS method (Figure 9b showing that about 79% of site-years deviated by 6 days or more from observations). Our results show that using humidity-based RSS methods improved the early melt issue in the NWM to some extent (Figure 9c and 9d), with the T_d based RSS method showing the most considerable degree of improvement compared to other RSS methods.

Table 2. Observed and modeled half melt dates comparison. Model half melt date is considered as early if it occurs one or more days before observations.

Scenarios that had observed precipitation and bias-corrected air temperature)	RSS scheme	Percentage of days with modeled half melt date earlier than observation across all sites and years
Scenario 3	Jordan (1991) T_a[†] based	67
Scenario 4	USACE (1956) T_a[†] based	72
Scenario 5	Marks et al. (1999) T_d[+] based	62
Scenario 6	Wang et al. (2019) T_w[*] based	65

[†]Air temperature
[+]Dew point temperature
[*]Wet-bulb temperature

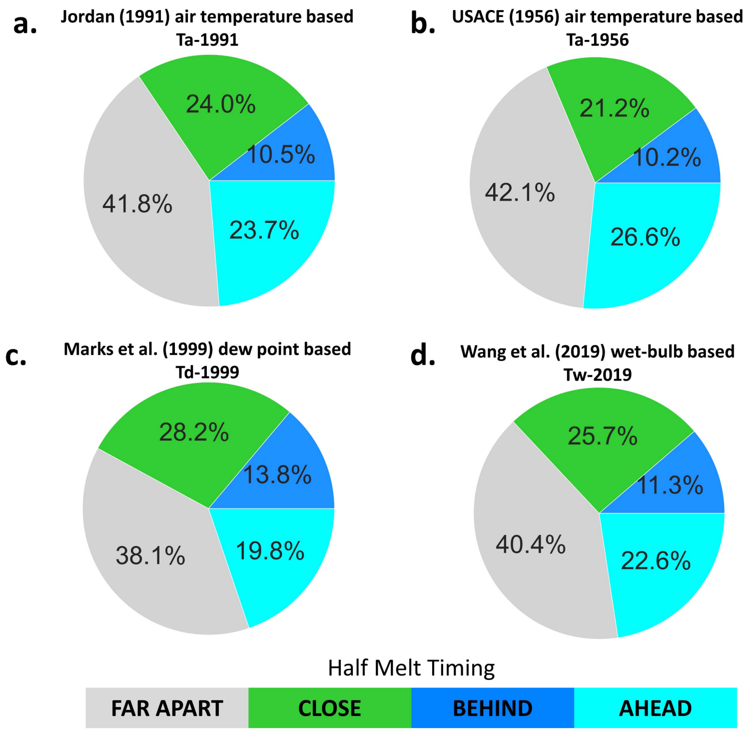

Figure 9. Analysis of melt timing based on classification of differences between observed and modeled dates of half melt from peak SWE. (a) NWM bias-corrected temperature scenario versus SNOTEL half melt dates, (b) NWM using USACE (1956) Ta based RSS method versus SNOTEL half melt dates, (c) NWM using Marks et al. (1999) T_d based RSS method versus SNOTEL half melt dates, and (d) NWM using Wang et al. (2019) T_w based RSS method versus SNOTEL half melt dates. In this figure, FAR APART: modeled and observed half melt dates are more than 20 days apart; CLOSE: modeled and observed half melt dates are within 5 days of each other; BEHIND: modeled half melt dates are 6 to 19 days after observed; and AHEAD: modeled half melt dates are 6 to 19 days before observed.

The NWM early melt issue inferred from the half melt date comparison between modeled results (Scenario 4 with Marks et al. (1999) T_d based method) and observations at selected SNOTEL sites during 11 years (the water year 2009-2019) was persistent across all sites but varied differently across ROS% classes (Figure 10). In this figure, the ROS% classes in the middle of the range, which represent sites with rain and snow mixes, as opposed to dominantly snow or dominantly rain, tended to have smaller percentages with close melt timing. For the sites where ROS% events were significantly high (>80%) or low (<40%), the modeled half melt date was close (off 6 days or less) more frequently (Figure 10a and 10f).

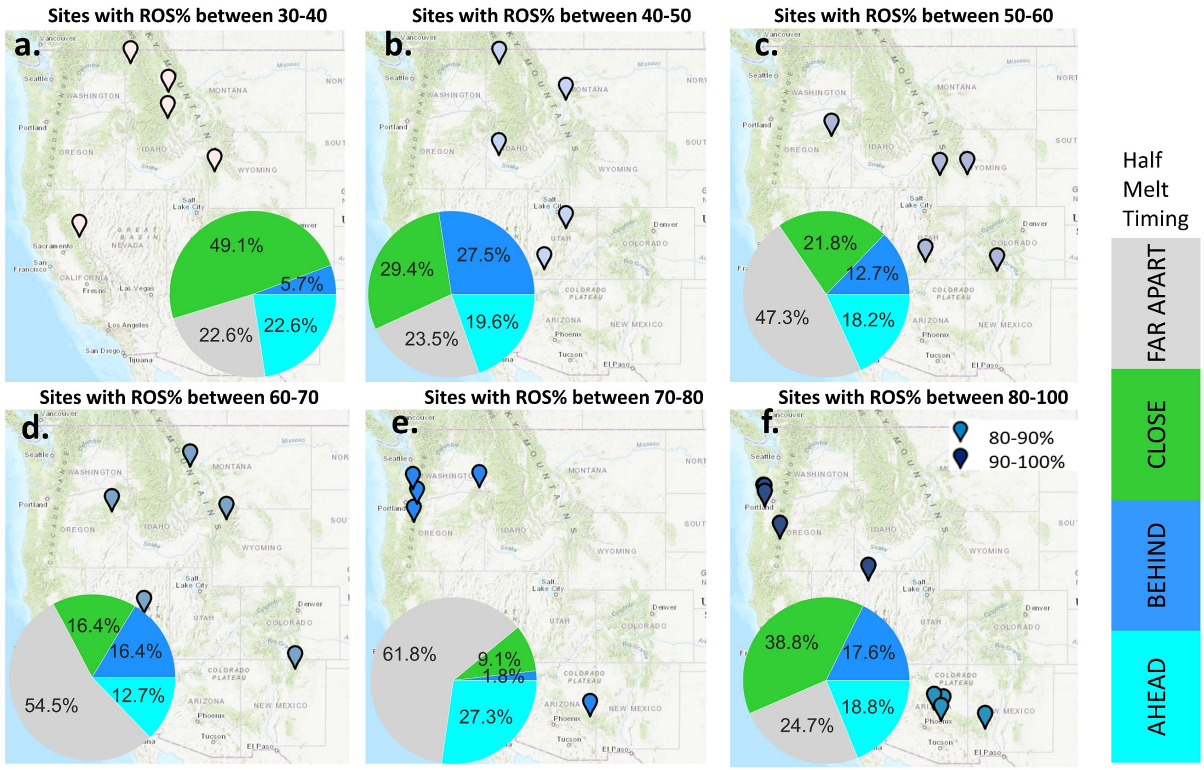

Figure 10. Analysis of melt timing from NWM using T_d based RSS scheme (the approach with the least bias and best NSE and RMSE in SWE comparisons) across different ROS% classes. (a) ROS% between 30 to 40%, (b) ROS% between 40 to 50%, (c) ROS% between 50 to 60%, (d) ROS% between 60 to 70%, (e) ROS% between 70 to 80%, and (f) ROS% between 80 to 100%.

6 Discussion, Perspective, and Future Work

In this study, our goal was to evaluate input data and three alternative RSS parameterizations to the NWM version 2.0 to find whether these improve SWE simulations. This section discusses findings for each of the research questions given in the introduction.

To what degree are discrepancies in NWM SWE and RSS predictions due to input errors and how much could they potentially be improved if inputs were better?

In this experiment, the most noticeable improvements in modeling SWE compared to the base scenario were achieved when we used observed precipitation from SNOTEL sites instead of the NWM AORC precipitation data (about 60% and 77% improvements in bias for same-day and different-day comparisons of peak SWE, respectively). Using better meteorological inputs to improve NWM performance has been reported by other studies (Lahmers et al., 2019; Viterbo et al., 2020). While stating that better inputs lead to better model performance is not new, this emphasizes the sensitivity to hydrometeorological input error, specifically precipitation and near-surface air temperature, in hydrological modeling predictions (Förster et al., 2014; Raleigh et al., 2015; Zehe et al., 2005).

Our model evaluation that quantifies how much the NWM performance in modeling SWE could improve by using more accurate meteorological inputs is important in considering where to invest time and effort in enhancing the NWM overall. We understand that model input improvements do not per se improve hydrologic process understanding; however, the ability to produce accurate hydrological forecasts is essential, and beyond forecast quality, the NWM does provide several outputs of hydrologic quantities, either not observed, or only observed in specialized field studies, and certainly not comprehensively across a continent. Examination of these outputs and their patterns across a continent does enhance process understanding. In addition, developing more accurate gridded precipitation products may reduce the need to make existing physical parameterizations more complex and add more uncertainties to the model due to new parameters (e.g., best fit coefficients in the Wang et al. (2019) T_w based approach).

How well does the NWM RSS (rainfall and snowfall separation) parameterization work in comparison to SNOTEL observations?

Our results showed that the NWM RSS (Jordan (1991) T_a based) performed statistically poorly (bias 41 mm, RMSE 74 mm) in separating precipitation into rain and snow compared to observed snowfall inferred from SNRR at 33 representative SNOTEL sites across the western U.S. Several challenges exist in this comparison, and each can be considered as a contributor to discrepancies observed. First, the spatial scale differences between SNOTEL and NWM datasets are a source of uncertainty in this analysis. As with all numerical models, the representation of sub-grid variability of snow processes may not be well parameterized when working with models such as the NWM that simulate snow processes across 1 km spatial resolution. Second, even though we used snow-adjusted precipitation from SNOTEL sites, there may still be systematic bias for SNOTEL precipitation due to under-catch (Mote, 2003; Sun et al., 2019). Third, even though we used observed precipitation from SNOTEL sites (instead of AORC precipitation that had downward bias) along with bias-corrected AORC air temperatures (corrected based on SNOTEL observations), there may still be uncertainties associated with other NWM AORC inputs, including specific humidity, in RSS calculations. Fourth, the method for inferring SNRR from SNOTEL measurements of precipitation and SWE has limitations. For example, rain that falls on a cold snowpack, freezes and adds to SWE mass will increase SWE and be interpreted to be snowfall. Other processes such as wind drifting or scouring of SWE at the SNOTEL site also introduce uncertainty. Lastly, while when SWE increases were more that P measurements they were used to infer and adjust for P under-catch, this does not adjust for under-catch of rainfall that may be present, even though it is commonly not thought to be as problematic as under-catch of snowfall (e.g., Meyer et al., 2012).

Do any other RSS parameterization methods yield more accurate snowfall compared to SNOTEL observations?

When considering other RSS alternatives from the literature, we observed that the dual-threshold air temperature-based method (USACE (1956) T_a based) yields noticeably better agreement between modeled and observed snowfall (bias 6 mm, RMSE 54 mm) compared to the other two humidity-based approaches (T_d based and T_w based). This may be interpreted as good, because it would be easier to apply a dual-threshold method with a linear decrease in between that takes only air temperature as the input to separate precipitation into rain and snow than T_d based or T_w based methods that determine the snowfall fraction using humidity information which potentially could add more errors if input data are not accurate. This finding is in line with the work of Feiccabrino et al. (2013) that reported on the superiority of the air temperature-based method over the dew point temperature approach based on data from 19 Swedish meteorological stations.

However, we should consider that this finding may be based on some assumptions that hinder us from concluding that USACE (1956) T_a based is the best among other methods tested in this study. Firstly, there are uncertainties associated with the NWM AORC data (even with our bias removal from precipitation and air temperature) we used as inputs to RSS methods and the reference data (SNRR) that we used to evaluate the performance of each RSS scheme. Secondly, even though air temperature-based RSS schemes are easy to use, they are empirically-based methods that have been developed based on historical data. Physically based methods are theoretically preferable for the simulation of processes under conditions that may differ from the historical conditions where empirical methods have been calibrated or optimized. We note that other studies report on the superiority of humidity-based approaches over air temperature-based ones in modeling both snowfall and SWE over mountainous regions (Ding et al., 2014; Marks et al., 2013; Wang et al., 2019). Further, as noted above, there are limitations associated with the SNOTEL inferred SNRR that may merit giving higher consideration to overall SWE simulation comparisons than snowfall ratio comparisons in assessing a RSS model. This is discussed below.

In this study, our results showed that snowfall estimates from Wang et al. (2019) T_w based scheme better agreed with observations inferred from SNRR at SNOTEL sites (Figure 5e: bias 34 mm, RMSE 63 mm) than those from Marks et al. (1999) T_d based scheme (Figure 5d: continuous version with bias 45 mm and RMSE 76 mm). This difference could be because T_w is more physically related to the precipitation phase as it considers the sensible and latent heat fluxes that determine the internal energy and temperature of a hydrometeor, and thus it is closer to the surface temperature of a falling hydrometeor than the air temperature (Wang et al., 2019). However, T_d only describes the cooling necessary for an unsaturated parcel of air to reach saturation over constant pressure, and it does not consider sensible and latent heat fluxes to the hydrometeor (Harder & Pomeroy, 2013). There may also be uncertainty related to best fit coefficients in the Wang et al. (2019) snowfall fraction equation that has been optimized to fit the observation-based relationship between snowfall probability and the T_w from Behrangi et al. (2018).

Does incorporating a statistically better RSS scheme into NWM translate into appreciable improvements in modeling of SWE?

Not only did incorporating a statistically better RSS scheme (Scenario 4 with USACE (1956) T_a based scheme) not translate into appreciable improvements in SWE estimates, but it turned out that this scheme was the least acceptable among the RSS alternatives evaluated when compared to SNOTEL SWE observations (evident in both same day and different day comparison of peak SWE).

When using observed precipitation and unbiased air temperature, our analysis showed that the humidity-dependent RSS schemes (dew point and wet-bulb temperature based) overcame the under-modeling of SWE to some extent. This is in line with previous work reporting on the impact of incorporating humidity into RSS processes on snowfall and snow mass compared to ground-based snow products (Behrangi et al., 2018; Jennings et al., 2018; Marks et al., 2013; Wang et al., 2019). In our study, while the Wang et al. (2019) T_w based RSS method showed better snowfall results than those from the Marks et al. (1999) T_d based RSS scheme, we found greater improvements in modeled SWE from the T_d based than T_w based RSS scheme (Figures 6 and 7). We give this finding that the T_d based RSS scheme performs better for direct comparisons against SNOTEL SWE observations greater credence than the USACE T_a based method performing best against inferred snowfall, due to the limitations associated with the SNOTEL SNRR separation method, and due to predictions of SWE being an ultimate target of this modeling. There was, however, remaining under-modeling of SWE which could be due to shortcomings associated with other meteorological inputs such as incoming solar and long-wave radiation which we did not study in this work and snow processes parameterizations in the NWM Noah-MP, such as the snow cover fraction calculations which have been reported to be problematic in modeling of SWE (Helbig et al., 2015; Magand et al., 2014; Wrzesien et al., 2015). These are open areas for future research to advance snow modeling in the NWM.

Collectively, our results showed that, on average, the NWM tended to melt snow early compared to observations at SNOTEL sites independent of the RSS scheme being used. However, the humidity-dependent approaches showed slightly better results. This observation that the modeling of melt timing was not significantly sensitive to the RSS scheme suggests that there is a need to investigate the overall energy balance and snow surface temperature calculations in the model.

How do improvements in modeled SWE vary over sites grouped according to the percentage of precipitation events that are rain on snow?

We observed that the degree of improvement in modeled SWE (in terms of both magnitude and melt timing) varied across ROS% classes. SWE was not well modeled for the ROS% classes in the middle rain dominated part of the range (60-80%), while at the lower end (predominantly snow) or higher end (predominantly rain) the model performed better. For these ROS% classes where the model performs better, Marks et al. (1999) T_d based separation gave the best improvements. A caveat of this analysis is that we characterized the representative SNOTEL sites based on the ROS% events metric that we computed based on the inferred precipitation phase from SNRR. We understand that this approach has limitations; however, without direct rainfall and snowfall measurements, which are rare across larger areas, it was the approach that was available to us.

7 Conclusions

Two key points emerge from this work. First, our comparison of the National Water Model (NWM) Noah-MP snow water equivalent (SWE) and SNOTEL snow water equivalent for representative sites and dates in the 2009-2019 water years reiterated that the accuracy of model inputs plays a key role in the accuracy of model outputs. Results showed that using observed precipitation and bias-corrected air temperature significantly improved the general downward bias in the NWM SWE magnitude and slightly improved early half melt timing of NWM compared to observations at representative SNOTEL sites across the western U.S. Second, our evaluation of three alternative RSS parameterizations in the NWM across a set of representative SNOTEL sites that spanned site rain-on-snow variability indicated that the negative bias in NWM SWE can be reduced, on average, by using RSS methods that incorporate specific humidity information in precipitation separation into rain and snow with consistent best estimates of the input data. Among the two humidity-based RSS schemes, the dew point temperature-based method was slightly better (smaller RMSE and Bias and larger NSE) than the wet-bulb temperature-based method at simulating peak SWE. Using the dew point temperature-based RSS also improved the modeling of melt timing slightly (early melt inferred from the half melt date comparison). Both SWE magnitude and timing varied across ROS% classes, with better results for the ROS% classes at the lower end (predominantly sow) or higher end (predominantly rain). These findings support the benefit of including physically based process representations in a model such as the NWM. Future work is required to assess the impact of improved SWE on streamflow.

Acknowledgments

This work was completed on the land of Eastern Shoshone Tribe, and was supported by the Utah Water Research Laboratory and National Science Foundation under collaborative grants OAC-1664061, OAC-1664018, and OAC-1664119. This work used compute allocation TG-EAR190007 from the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562 (Towns et al., 2014). We thank David Gochis at NCAR and Ed Clark at the National Water Center for facilitating access to the NWM inputs. Thanks to Mahidhar Tatineni at the San Diego Supercomputer Center who helped to optimize our computational work load on XSEDE. Thanks to Mahyar Aboutalebi for his help with computational simulation runs on XSEDE, and to Jeffery S. Horsburgh for his comments and suggestions.

Open Research

Codes developed for this research and the data we specifically used are publicly available in the HydroShare repository (Garousi-Nejad & Tarboton, 2022b).

The data and model sources that we drew from include:

- SNOTEL data accessed through the NRCS Report Generator v2: https://wcc.sc.egov.usda.gov/reportGenerator/
- WRF-Hydro version 5.1.1 source code was accessed in GitHub: https://github.com/NCAR/wrf_hydro_nwm_public/releases/tag/v5.1.1
- NWM physiographic and atmospheric meteorological inputs were made available to us by the NCAR team in the NCAR Cheyenne high-performance computer. The specific data we used from this source are in the HydroShare resource given above.

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: MOUNTAIN HYDROLOGY OF THE WESTERN US. *Water Resources Research*, *42*(8). https://doi.org/10.1029/2005WR004387

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

Behrangi, A., Yin, X., Rajagopal, S., Stampoulis, D., & Ye, H. (2018). On distinguishing snowfall from rainfall using near-surface atmospheric information: comparative analysis, uncertainties and hydrologic importance. *Quarterly Journal of the Royal Meteorological Society*, *144*(S1), 89–102. https://doi.org/10.1002/qj.3240

Bhatti, A. M., Koike, T., & Shrestha, M. (2016). Climate change impact assessment on mountain snow hydrology by water and energy budget-based distributed hydrological model. *Journal of Hydrology*, *543*, 523–541. https://doi.org/10.1016/j.jhydrol.2016.10.025

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: SENSITIVITY STUDY OF LSMS IN WRF. *Journal of Geophysical Research: Atmospheres*, *119*(12), 7271–7291. https://doi.org/10.1002/2014JD021827

Clow, D. W. (2010). Changes in the timing of snowmelt and streamflow in Colorado: A response to recent warming. *Journal of Climate*, *23*(9), 2293–2306. https://doi.org/10.1175/2009JCLI2951.1

DeWalle, D. R., & Rango, A. (2008). *Principles of Snow Hydrology*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511535673

Ding, B., Yang, K., Qin, J., Wang, L., Chen, Y., & He, X. (2014). The dependence of precipitation types on surface elevation and meteorological conditions and its parameterization. *Journal of Hydrology*, *513*, 154–163. https://doi.org/10.1016/j.jhydrol.2014.03.038

Feiccabrino, J., Gustafsson, D., & Lundberg, A. (2013). Surface-based precipitation phase determination methods in hydrological models. *Hydrology Research*, *44*(1), 44–57. https://doi.org/10.2166/nh.2012.158

Feiccabrino, J., Graff, W., Lundberg, A., Sandström, N., & Gustafsson, D. (2015). Meteorological Knowledge Useful for the Improvement of Snow Rain Separation in Surface Based Models. *Hydrology*, *2*(4), 266–288. https://doi.org/10.3390/hydrology2040266

Förster, K., Meon, G., Marke, T., & Strasser, U. (2014). Effect of meteorological forcing and snow model complexity on hydrological simulations in the Sieber catchment (Harz Mountains, Germany). *Hydrology and Earth System Sciences*, *18*(11), 4703–4720. https://doi.org/10.5194/hess-18-4703-2014

Garousi-Nejad, I., & Tarboton, D. G. (2022a). A comparison of National Water Model retrospective analysis snow outputs at snow telemetry sites across the Western United States. *Hydrological Processes*, *36*(1). https://doi.org/10.1002/hyp.14469

Garousi-Nejad, I., & Tarboton, D. G. (2022b). Data for Evaluating Input Data and Rain Snow Separation Improvements to the National Water Model Simulation of Snow Water Equivalent. HydroShare. Retrieved from http://www.hydroshare.org/resource/bdbecdef23b14848b5da46c4f465ec21

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

Gillies, R. R., Wang, S.-Y., & Booth, M. R. (2012). Observational and Synoptic Analyses of the Winter Precipitation Regime Change over Utah. *Journal of Climate*, *25*(13), 4679–4698. https://doi.org/10.1175/JCLI-D-11-00084.1

Gochis, D., Barlage, M., Cabell, R., Casali, M., Dugger, A., FitzGerald, K., et al. (2020). The WRF-Hydro® modeling system technical description, (Version 5.1.1). NCAR Technical Note. Retrieved from https://ral.ucar.edu/sites/default/files/public/WRFHydroV511TechnicalDescription.pdf

Gochis, D., Barlage, M., Cabell, R., Dugger, A., Fanfarillo, A., FitzGerald, K., et al. (2020). WRF-Hydro® v5.1.1 (Version v5.1.1). Zenodo. https://doi.org/10.5281/ZENODO.3625238

Gulev, S. K., Thorne, P. W., Ahn, J., Dentener, F. J., Domingues, C. M., Gerland, S., et al. (2021). *Changing State of the Climate System* (In Climate Change 2021: The Physical

Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change). Cambridge University Press. In Press.
Harder, P., & Pomeroy, J. (2013). Estimating precipitation phase using a psychrometric energy balance method: PRECIPITATION PHASE USING A PSYCHROMETRIC ENERGY BALANCE. *Hydrological Processes*, *27*(13), 1901–1914. https://doi.org/10.1002/hyp.9799
Harder, P., & Pomeroy, J. W. (2014). Hydrological model uncertainty due to precipitation-phase partitioning methods: HYDROLOGIC MODEL UNCERTAINTY OF PRECIPITATION-PHASE METHODS. *Hydrological Processes*, *28*(14), 4311–4327. https://doi.org/10.1002/hyp.10214
Harpold, A. A., Kaplan, M. L., Klos, P. Z., Link, T., McNamara, J. P., Rajagopal, S., et al. (2017). Rain or snow: hydrologic processes, observations, prediction, and research needs. *Hydrology and Earth System Sciences*, *21*(1), 1–22. https://doi.org/10.5194/hess-21-1-2017
Helbig, N., van Herwijnen, A., Magnusson, J., & Jonas, T. (2015). Fractional snow-covered area parameterization over complex topography. *Hydrology and Earth System Sciences*, *19*(3), 1339–1351. https://doi.org/10.5194/hess-19-1339-2015
Jennings, K. S., Winchell, T. S., Livneh, B., & Molotch, N. P. (2018). Spatial variation of the rain–snow temperature threshold across the Northern Hemisphere. *Nature Communications*, *9*(1), 1148. https://doi.org/10.1038/s41467-018-03629-7
Jordan, R. E. (1991). A One-dimensional temperature model for a snow cover : technical documentation for SNTHERM.89. Cold Regions Research and Engineering Laboratory (U.S.). Retrieved from http://hdl.handle.net/11681/11677
Kitzmiller, D. H., Wu, H., Zhang, Z., Patrick, N., & Tan, X. (2018). *The Analysis of Record for Calibration: A High-Resolution Precipitation and Surface Weather Dataset for the United States*. Presented at the American Geophysical Union, Fall Meeting, Washington, D.C. Retrieved from https://ui.adsabs.harvard.edu/abs/2018AGUFM.H41H..06K/abstract
Klos, P. Z., Link, T. E., & Abatzoglou, J. T. (2014). Extent of the rain-snow transition zone in the western U.S. under historic and projected climate: Climatic rain-snow transition zone. *Geophysical Research Letters*, *41*(13), 4560–4568. https://doi.org/10.1002/2014GL060500
Knowles, N., Dettinger, M. D., & Cayan, D. R. (2006). Trends in Snowfall versus Rainfall in the Western United States. *Journal of Climate*, *19*(18), 4545–4559. https://doi.org/10.1175/JCLI3850.1
Lahmers, T. M., Gupta, H., Castro, C. L., Gochis, D. J., Yates, D., Dugger, A., et al. (2019). Enhancing the Structure of the WRF-Hydro Hydrologic Model for Semiarid Environments. *Journal of Hydrometeorology*, *20*(4), 691–714. https://doi.org/10.1175/JHM-D-18-0064.1
Li, D., Wrzesien, M. L., Durand, M., Adam, J., & Lettenmaier, D. P. (2017). How much runoff originates as snow in the western United States, and how will that change in the future?: Western U.S. Snowmelt-Derived Runoff. *Geophysical Research Letters*, *44*(12), 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

Loth, B., Graf, H.-F., & Oberhuber, J. M. (1993). Snow cover model for global climate simulations. *Journal of Geophysical Research*, *98*(D6), 10451. https://doi.org/10.1029/93JD00324

Magand, C., Ducharne, A., Le Moine, N., & Gascoin, S. (2014). Introducing Hysteresis in Snow Depletion Curves to Improve the Water Budget of a Land Surface Model in an Alpine Catchment. *Journal of Hydrometeorology*, *15*(2), 631–649. https://doi.org/10.1175/JHM-D-13-091.1

Mankin, J. S., Viviroli, D., Singh, D., Hoekstra, A. Y., & Diffenbaugh, N. S. (2015). The potential for snow to supply human water demand in the present and future. *Environmental Research Letters*, *10*(11), 114016. https://doi.org/10.1088/1748-9326/10/11/114016

Marks, D., Domingo, J., Susong, D., Link, T., & Garen, D. (1999). A spatially distributed energy balance snowmelt model for application in mountain basins. *Hydrological Processes*, *13*(12–13), 1935–1959. https://doi.org/10.1002/(SICI)1099-1085(199909)13:12/13<1935::AID-HYP868>3.0.CO;2-C

Marks, D., Winstral, A., Reba, M., Pomeroy, J., & Kumar, M. (2013). An evaluation of methods for determining during-storm precipitation phase and the rain/snow transition elevation at the surface in a mountain basin. *Advances in Water Resources*, *55*, 98–110. https://doi.org/10.1016/j.advwatres.2012.11.012

Meyer, J. D. D., Jin, J., & 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*(6), 1970–1976. https://doi.org/10.1175/JHM-D-12-066.1

Mizukami, N., Koren, V., Smith, M., Kingsmill, D., Zhang, Z., Cosgrove, B., & Cui, Z. (2013). The Impact of Precipitation Type Discrimination on Hydrologic Simulation: Rain–Snow Partitioning Derived from HMT-West Radar-Detected Brightband Height versus Surface Temperature Data. *Journal of Hydrometeorology*, *14*(4), 1139–1158. https://doi.org/10.1175/JHM-D-12-035.1

Mote, P. W. (2003). Trends in snow water equivalent in the Pacific Northwest and their climatic causes: TRENDS IN SNOW WATER EQUIVALENT. *Geophysical Research Letters*, *30*(12). https://doi.org/10.1029/2003GL017258

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), 39–50. https://doi.org/10.1175/BAMS-86-1-39

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 Weather Service, Office of Water Prediction. (2021). Analysis of Record for Calibration: Version 1.1 Sources, Methods, and Verification. NOAA. Retrieved from https://hydrology.nws.noaa.gov/aorc-historic/Documents/AORC-Version1.1-SourcesMethodsandVerifications.pdf

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 multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research*, *116*(D12), D12109. https://doi.org/10.1029/2010JD015139

Raleigh, M. S., Lundquist, J. D., & Clark, M. P. (2015). Exploring the impact of forcing error characteristics on physically based snow simulations within a global sensitivity analysis framework. *Hydrology and Earth System Sciences*, *19*(7), 3153–3179. https://doi.org/10.5194/hess-19-3153-2015

Rutter, N., Essery, R., Pomeroy, J., Altimir, N., Andreadis, K., Baker, I., et al. (2009). Evaluation of forest snow processes models (SnowMIP2). *Journal of Geophysical Research*, *114*(D6), D06111. https://doi.org/10.1029/2008JD011063

Shuttleworth, W. J. (2012). *Terrestrial Hydrometeorology: Shuttleworth/Terrestrial Hydrometeorology*. Chichester, UK: John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119951933

Stull, R. (2011). Wet-Bulb Temperature from Relative Humidity and Air Temperature. *Journal of Applied Meteorology and Climatology*, *50*(11), 2267–2269. https://doi.org/10.1175/JAMC-D-11-0143.1

Sun, N., Yan, H., Wigmosta, M. S., Leung, L. R., Skaggs, R., & Hou, Z. (2019). Regional Snow Parameters Estimation for Large-Domain Hydrological Applications in the Western United States. *Journal of Geophysical Research: Atmospheres*, *124*(10), 5296–5313. https://doi.org/10.1029/2018JD030140

Tarboton, D. G., & Luce, C. H. (1996). Utah Energy Balance Snow Accumulation and Melt Model (UEB). Utah Water Research Laboratory and USDA Forest Service Intermountain Research Station. Retrieved from https://hydrology.usu.edu/dtarb/snow/snowreptext.pdf

Towns, J., Cockerill, T., Dahan, M., Foster, I., Gaither, K., Grimshaw, A., et al. (2014). XSEDE: Accelerating Scientific Discovery. *Computing in Science & Engineering*, *16*(5), 62–74. https://doi.org/10.1109/MCSE.2014.80

U.S. Army Corps of Engineers. (1956). Snow Hydrology, Summary report of the Snow Investigations. U.S. Army Corps of Engineers. Retrieved from https://usace.contentdm.oclc.org/digital/collection/p266001coll1/id/4172/

Viterbo, F., Mahoney, K., Read, L., Salas, F., Bates, B., Elliott, J., et al. (2020). A Multiscale, Hydrometeorological Forecast Evaluation of National Water Model Forecasts of the May 2018 Ellicott City, Maryland, Flood. *Journal of Hydrometeorology*, *21*(3), 475–499. https://doi.org/10.1175/JHM-D-19-0125.1

Wang, Y., Broxton, P., Fang, Y., Behrangi, A., Barlage, M., Zeng, X., & Niu, G. (2019). A Wet-Bulb Temperature-Based Rain-Snow Partitioning Scheme Improves Snowpack Prediction Over the Drier Western United States. *Geophysical Research Letters*, *46*(23), 13825–13835. https://doi.org/10.1029/2019GL085722

Wen, L., Nagabhatla, N., Lü, S., & Wang, S.-Y. (2013). Impact of rain snow threshold temperature on snow depth simulation in land surface and regional atmospheric models. *Advances in Atmospheric Sciences*, *30*(5), 1449–1460. https://doi.org/10.1007/s00376-012-2192-7

Wrzesien, M. L., Pavelsky, T. M., Kapnick, S. B., Durand, M. T., & Painter, T. H. (2015). Evaluation of snow cover fraction for regional climate simulations in the Sierra Nevada: EVALUATION OF SNOW COVER FOR REGIONAL SIMULATIONS IN THE SIERRA NEVADA. *International Journal of Climatology*, *35*(9), 2472–2484. https://doi.org/10.1002/joc.4136

Zehe, E., Becker, R., Bárdossy, A., & Plate, E. (2005). Uncertainty of simulated catchment runoff response in the presence of threshold processes: Role of initial soil moisture and precipitation. *Journal of Hydrology*, *315*(1–4), 183–202. https://doi.org/10.1016/j.jhydrol.2005.03.038