Water Resources Research

RESEARCH ARTICLE
10.1029/2023WR035303

The Role of Forcing and Parameterization in Improving Snow Simulation in the Upper Colorado River Basin Using the National Water Model

Yanjun Gan[1], Yu Zhang[1], Cezar Kongoli[2], and Ming Pan[3]

[1]Department of Civil Engineering, University of Texas at Arlington, Arlington, TX, USA, [2]Earth System Science Interdisciplinary Center, University of Maryland, College Park, College Park, MD, USA, [3]Center for Western Weather and Water Extremes, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA, USA

Abstract This study assesses snow water equivalent (SWE) simulation uncertainty in the National Water Model (NWM) due to forcing and model parameterization, using data from 46 Snow Telemetry (SNOTEL) sites in the Upper Colorado River Basin (UCRB). We evaluated the newly developed Analysis of Record for Calibration (AORC) forcing data for SWE simulation and examined the impact of bias correction applied to AORC precipitation and temperature. Additionally, we investigated the sensitivity of SWE simulations to choices of physical parameterization schemes through 72 ensemble experiments. Results showed that NWM driven by AORC forcings captured the overall temporal variation of SWE but underestimated its amount. Adjusting AORC precipitation with SNOTEL observations reduced SWE root-mean-square error (RMSE) by 66%, adjusting temperature trimmed it by 10%, and adjusting both decreased it by 69%. Among the physical processes, the snow/soil temperature time scheme (STC) demonstrated the highest sensitivity, followed by the surface exchange coefficient for heat (SFC), snow surface albedo (ALB), and rainfall and snowfall partitioning (SNF), while the lower boundary of soil temperature (TBOT) proved to be insensitive. Further optimization of the parameterization combination resulted in a 12% SWE RMSE reduction. When combined with the bias-corrected AORC precipitation and temperature, this optimization led to a remarkable 78% SWE RMSE reduction. Despite these enhancements, a persistent slow and late spring ablation suggests model deficiencies in snow ablation physics. The study emphasizes the critical need to enhance the accuracy of forcing data in mountainous regions and address model parameterization uncertainty through optimization efforts.

Plain Language Summary This study examines how accurately the National Water Model (NWM) predicts the amount of snow water equivalent (SWE) in the Upper Colorado River Basin (UCRB). Using data from 46 monitoring sites, we looked at how different factors affect the model's accuracy. We found that while the NWM can show changes in SWE over time, it consistently underestimates the actual amount when using certain forcing data. Adjusting this data with observations improved the model's accuracy significantly. We also identified specific model settings that strongly influence the results. Although optimizing these settings only improved accuracy by a small amount, combining them with adjusted forcing data led to a significant improvement. Despite these improvements, the model still struggles to predict late spring snowmelt accurately. This study emphasizes the importance of using accurate data and model settings for managing water resources effectively in mountainous areas.

Key Points:

- National Water Model driven by Analysis of Record for Calibration (AORC) captures SWE temporal variation but underestimates its amount
- Adjusting precipitation reduces SWE RMSE by 66%, adjusting temperature reduces it by 10%, and adjusting both reduces it by 69%
- Parameterization optimization reduces SWE RMSE by 12%, while it leads to a 78% reduction when combined with adjusted AORC

Supporting Information:

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

Correspondence to:

Y. Gan,
yanjun.gan@uta.edu

Citation:

Gan, Y., Zhang, Y., Kongoli, C., & Pan, M. (2024). The role of forcing and parameterization in improving snow simulation in the Upper Colorado River Basin using the National Water Model. *Water Resources Research*, *60*, e2023WR035303. https://doi.org/10.1029/2023WR035303

Received 8 MAY 2023
Accepted 4 JUL 2024

Author Contributions:

Conceptualization: Yanjun Gan, Yu Zhang, Cezar Kongoli
Formal analysis: Yanjun Gan
Funding acquisition: Yu Zhang
Investigation: Yu Zhang
Methodology: Yanjun Gan, Yu Zhang, Cezar Kongoli
Project administration: Yu Zhang
Resources: Yu Zhang
Supervision: Yu Zhang
Validation: Yanjun Gan
Visualization: Yanjun Gan
Writing – original draft: Yanjun Gan

1. Introduction

Snowmelt plays a crucial role in driving streamflow throughout the western United States (US; Li et al., 2017; Siirila Woodburn et al., 2021), accounting for 50%–80% of the region's annual runoff (Stewart et al., 2004). Therefore, accurate and timely prediction of snowpack evolution and snowmelt is essential for water resources management in this region (Hammond & Kampf, 2020; Trujillo & Molotch, 2014). Within the western US, snowpack prediction is particularly critical for the headwaters of the Upper Colorado River Basin (UCRB; Xiao & Lettenmaier, 2021), where runoff generated by snowmelt supplies to ~40 million people across five US states and Mexico and supports ~$1.4 trillion in annual economic activity and ~16 million jobs (James et al., 2014). However, this task is challenging since model-simulated snowpack often significantly deviates from in situ observations, showing discrepancies such as lower snow depth, delayed ablation, or other variations (Broxton et al., 2016; Mortimer et al., 2020).

Writing – review & editing: Yanjun Gan, Yu Zhang, Cezar Kongoli, Ming Pan

To date, a variety of hydrologic and land surface models (LSMs) with varying levels of complexity have been developed to simulate snow accumulation and ablation processes (Clark et al., 2011; Fisher & Koven, 2020). Intercomparison projects, including the Project for the Intercomparison of Land-Surface Parameterization Schemes (PILPS) Phase 2(d) (Slater et al., 2001) and Phase 2(e) (Bowling et al., 2003) and the Snow Model Intercomparison Project (SnowMIP) Phase 1 (Echevers et al., 2004) and Phase 2 (Rutter et al., 2009), have shown that (a) differences in model physics and parameterization schemes often result in large differences in simulated snowpack and snowmelt, and (b) simulated snowpack evolution is highly sensitive to the errors in the prescribed forcings.

Exploring the roles of model physics deficiencies and forcing errors in shaping errors in simulated snowpack and snowmelt has been an important theme in snow hydrology and land surface modeling (Günther et al., 2019; Kim et al., 2021). For instance, in a study by Chen et al. (2014), the performance of six LSMs in reproducing snow water equivalent (SWE) simulation within the Colorado River Headwaters region was evaluated. This study revealed a significant intermodal disparity arising from the treatment of snow albedo and its subsequent impacts on surface energy deficit, surface temperature, stability correction, and turbulent fluxes. Xiao et al. (2021) investigated the performance differences of four LSMs in reproducing observed SWE during the ablation season at 10 Snow Telemetry (SNOTEL; Serreze et al., 1999) stations. The study demonstrated that net radiation exhibited stronger correlations with observed melt rates than air temperature. The variations in net radiation were mainly attributed to differences in net shortwave radiation, which, in turn, were primarily caused by variations in ground surface albedo and vegetation shading effects among the models. Broxton et al. (2016) compared six reanalysis products and five Global Land Data Assimilation System (GLDAS; Rodell et al., 2004) products with three high-resolution SWE data sets. They found that all products underestimated SWE in the conterminous US (CONUS), especially in the western region, due primarily to excessive snow ablation occurring at near-freezing temperatures. In addition, forcing bias in precipitation and temperature also contributed to the underestimation of SWE, either worsening or alleviating the bias. Nevertheless, the specific contributions of these forcing biases have not been precisely quantified.

The past model intercomparison projects produced insights to key processes that regulate snow accumulation and ablation, but these alone are often too specific to inform optimal models or parameterization schemes (Terzago et al., 2020). For the latter purpose, investigators resort to unified modeling frameworks that implement multiple parameterization options for the same processes (Best et al., 2011; Clark et al., 2015; Niu et al., 2011). The control simulations using these options allow for testing hypotheses concerning the fidelity of a single, or a combination of parameterization schemes in reproducing specific processes. For example, Essery et al. (2013) used a single model with various options for snow-related processes to test 1701 scheme combinations of differing complexity and found a group of scheme combinations with good performance at an alpine site. You et al. (2020) applied the same methods as in Zhang et al. (2016) and Gan et al. (2019) to identify the most critical processes and optimal scheme combinations in the Noah-MP model (Niu et al., 2011) for snow simulations at eight sites with diverse snow climates. While these works shed light on the interplays among parameterization schemes and helped determine possible optimal combinations, they share the common flaw of not accounting for errors in the forcing data.

The impacts from forcing errors can be as large as, if not larger than, those stemming from inadequacy of model in representing snow processes, and their roles can be particularly prominent in mountain regions where meteorological variables are strongly modulated by terrain and where observation networks are sparse (Beniston et al., 2018; Raleigh et al., 2015; Y. Wang et al., 2020). For example, He et al. (2019) employed five different precipitation forcing data sets to evaluate precipitation uncertainty in snowpack simulation over the western US mountains. They found that precipitation significantly impacts simulated SWE and snow depth, while it has limited effects on snow cover fraction and surface albedo. M. Pan et al. (2003) found that errors in precipitation and air temperature are a main cause of low modeled SWE in the western US by comparing the first phase North American Land Data Assimilation System (NLDAS-1; Mitchell et al., 2004) forcings with SNOTEL measurements. Terzago et al. (2020) showed that a simple correction to the average temperature bias in the forcing data, based on site observations, can effectively remove the snowfall bias and lead to more accurate simulation of SWE. However, the same correction applied to precipitation data leads to a decrease in the snow model's performance. While traditional bias correction methods have been employed in previous research to rectify forcing errors, it is important to acknowledge the broader spectrum of efforts in the literature dedicated to bias correction using data assimilation (DA) techniques (Cluzet et al., 2022; Magnusson et al., 2014; Winstral et al., 2019). Some of these

more advanced DA techniques not only address errors in forcing data but also account for uncertainties in model parameterizations and physics.

The National Water Model (NWM; https://water.noaa.gov/about/nwm), which became operational for water prediction in 2016, uses Noah-MP as its core LSM and has a default set of parameterization schemes for various land surface processes. However, Gan et al. (2022) observed a pronounced negative bias in SWE simulation within the UCRB when driven by the North American Land Data Assimilation System project phase 2 (NLDAS-2; Xia et al., 2012) forcings. While forcing errors and model physics deficiencies both apparently played roles, we would like to explore the following questions: (a) How do errors in meteorological forcing data, particularly precipitation and air temperature, impact the precision of snowpack simulations? (b) What is the relative influence of model parameterization schemes on the accuracy of snowpack simulations? (c) Can the adoption of optimal parameterization schemes, combined with bias-corrected forcing data, enhance the predictability of snowpack evolution in the UCRB? (d) What insights can these investigations provide for enhancing snow models in complex mountainous regions?

In light of these questions, the present study conducts a series of experiments with the overarching aim to distinguish the specific roles played by the forcings errors and parameterization scheme limitations to the observed SWE bias in the UCRB. In these experiments, we employed both original and bias-corrected forcing data derived from the Analysis of Record for Calibration (AORC; Fall et al., 2023) to drive the NWM to quantify the individual impacts of forcings errors. Furthermore, we evaluated alternative parameterization schemes in conjunction with the bias-corrected forcing data through NWM simulations to identify specific deficiencies in land process representations and their consequent effects on simulated SWE. This research draws lessons from the diagnostic work by Broxton et al. (2016), but differs from the latter by (a) quantitatively assessing the isolated impacts of forcing errors and model physics limitations, and (b) identifying parameterization schemes that render the most physically realistic representation of snowpack evolution. The insights gained from this study may directly guide efforts to enhance the accuracy of forcings and improve physical process representations, not only for Noah-MP but also for other widely used, physics-based snow models. This, in turn, can significantly enhance their predictive capabilities for seasonal snowpack in complex mountainous terrains.

The rest of this paper is organized as follows. Section 2 describes the study area and data. Section 3 introduces the model and experimental setup. Results are described and discussed in Section 4. Finally, conclusions are summarized in Section 5.

2. Study Area and Data

2.1. Study Area

The UCRB (Figure 1) features a 277,000 km^2 drainage area, an elevation range of 1,056–4,328 m, and an average elevation of 2,146 m. According to the US Geological Survey (USGS) 24-category land cover classification (Loveland et al., 2000), about 57% of the land is shrubland, 22% evergreen needleleaf, 7% deciduous broadleaf forest, 7% grassland, and 4% barren/sparsely vegetated. 63% of the annual precipitation in the UCRB is snow, falling mainly from November–April, until melting starts in late April or early May (Kalra et al., 2017).

2.2. Static Data

We obtained the static data for the CONUS from the National Water Center (NWC) and subset it for the UCRB (870 × 603 1-km grid cells). The 1-km resolution land surface model grid and relevant geospatial data were processed from the geographical static data using the Weather Research and Forecasting (WRF) model Preprocessing System (WPS; W. Wang et al., 2017). The primary sources of static data include the following: (a) USGS 30-s Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010; Danielson & Gesch, 2011); (b) USGS 30-s 24-category land cover data (Loveland et al., 2000); (c) the National Environmental Satellite, Data, and Information Service (NESDIS) 0.15-degree monthly 5-year climatology surface albedo and 0.144-degree monthly 5-year climatology green vegetation fraction data (Gutman & Ignatov, 1998); (d) the State Soil Geographic (STATSGO)/Food and Agriculture Organization (FAO) 30-s 16-category soil texture data (FAO, 1991); and (e) the International Satellite Land-Surface Climatology Project Initiative I (ISLSCP-I) 1-degree annual mean 2-m air temperature (Sellers et al., 1996), which is used as bottom boundary layer conditions for soil models. In addition, the 250-m resolution routing-related data (e.g., elevation, flow direction, and

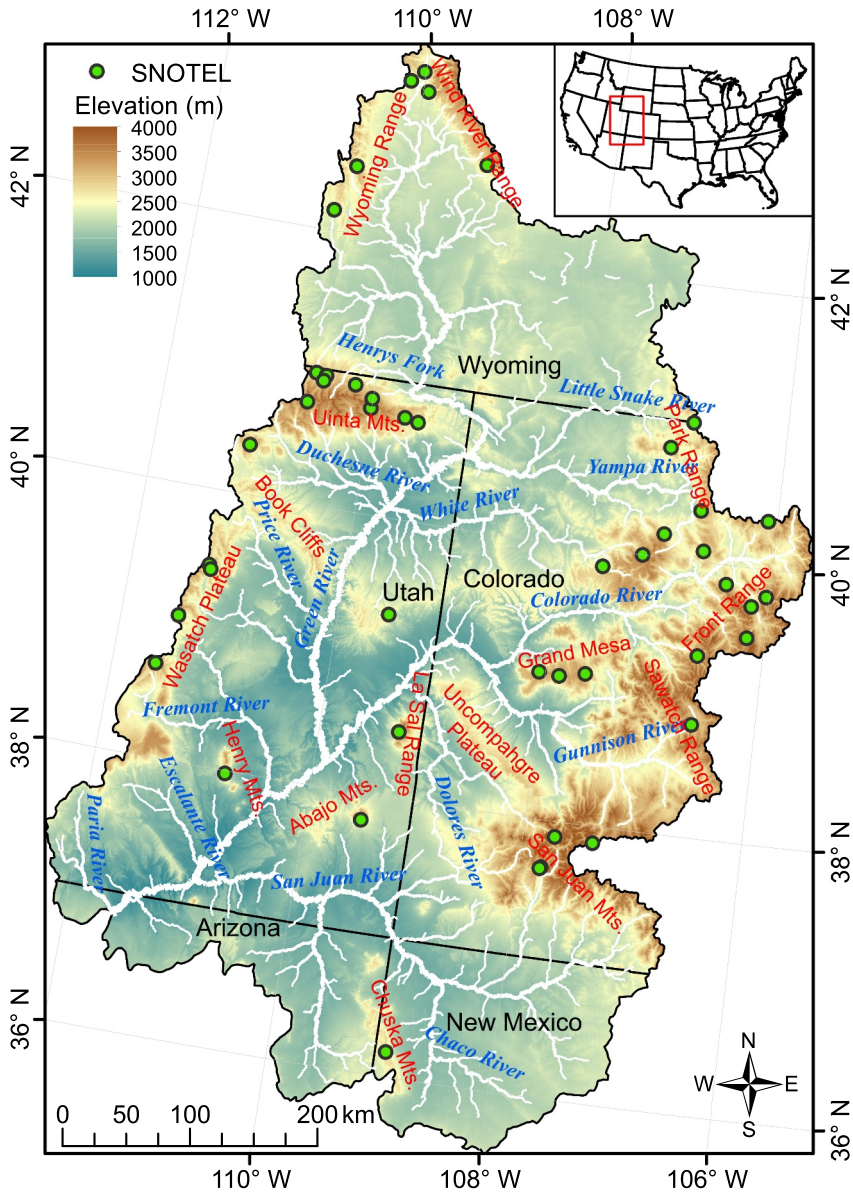

Figure 1. Spatial distributions of the selected SNOTEL snow survey sites in the Upper Colorado River Basin (UCRB). The upper right corner shows the location of the UCRB in the US.

stream order) were processed from the National Hydrography Data set Plus Version 2 (NHDPlusV2; McKay et al., 2012) using the WRF-Hydro GIS Pre-processor v5.1.1 (Sampson & Gochis, 2020).

2.3. Forcing Data

We evaluated the uncertainty of SWE simulation with the NWM using 1-km hourly AORC forcing data (Fall et al., 2023) for the water years (WYs) 2016–2019 (1 October 2015–30 September 2019). The AORC forcing data set was developed by National Oceanic and Atmospheric Administration (NOAA) National Weather Service (NWS) to ensure long-term consistency for land surface and hydrologic model calibrations. It covers the CONUS, southern Canada, and northern Mexico from 1979 to near present and is derived from various sources. For the period 2016–near present, UnRestricted Mesoscale Analysis (URMA; Raby et al., 2020), Next-Generation Weather Radar (NEXRAD) Stage IV (Lin & Mitchell, 2005), and NLDAS-2 (Xia et al., 2012) data sets were the main sources. Within the CONUS, the AORC products over recent years (after 2016) did not undergo bias correction due to unavailability of climatology for this period. The sources, methods, and verification of AORC

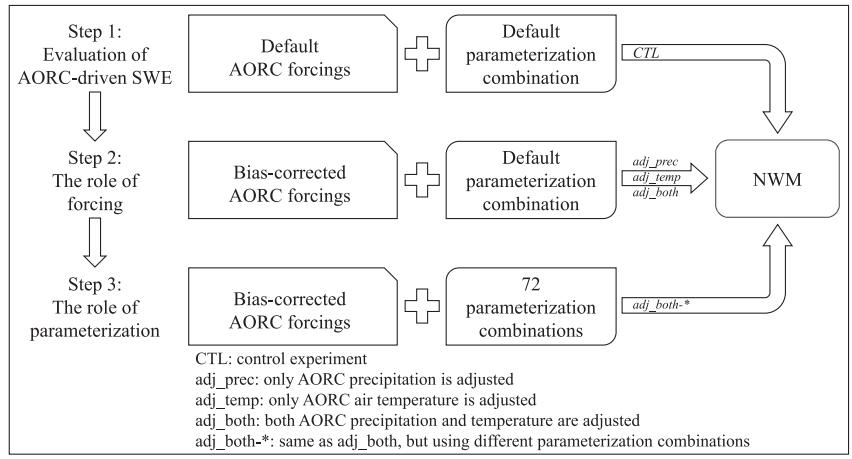

Figure 2. Flowchart illustrating the key steps in the research, highlighting the use of various forcing data and parameterization scheme combinations at each stage for running the NWM.

data are described in Office of Water Prediction (2021). The 1-km hourly AORC data was downloaded from Amazon Web Services (AWS; https://registry.opendata.aws/nwm-archive) and subset for the study area.

2.4. Observation Data

We compared SWE at point-scale using data from SNOTEL sites, as large-scale SWE observations in mountainous regions are unavailable (Dozier et al., 2016). The Natural Resources Conservation Service (NRCS) operates the SNOTEL program, which consists of over 800 mid- to high-elevation mountain stations in the western US. In operation since 1980, the program provides snowpack and basic meteorological data using automated measuring devices, such as pressure-sensing snow pillows, storage precipitation gauges, and air temperature sensors (Schaefer & Paetzold, 2000). To mitigate data uncertainty, we utilized bias-corrected and quality-controlled (BCQC) SNOTEL data developed at the Pacific Northwest National Laboratory (available at https://www.pnnl.gov/data-products), including daily SWE, precipitation, and mean air temperature. The BCQC procedure applied to the original SNOTEL data set involved several essential steps, such as the removal of outliers and erroneous values (e.g., negative SWE and precipitation), elimination of inconsistent SWE and precipitation values (e.g., peak SWE exceeds associated precipitation by more than 5%), and the correction of temperature and precipitation biases (e.g., addressing snowfall under-catch). Detailed information regarding this method can be found in Yan et al. (2018) and Sun et al. (2019).

We conducted a screening process by comparing the elevations of all SNOTEL sites with the equivalent grid cells. Sites with an absolute elevation difference greater than 50 m were excluded (63 out of 127), following the approach by M. Pan et al. (2003). Additionally, we removed sites where the total amount of missing data exceeded one-third of the validation period WYs 2017–2019 (18 out of 64). Figure 1 displays the spatial distributions of the final 46 SNOTEL sites in UCRB, with elevations ranging from 2,359 to 3,414 m. It's crucial to recognize the potential impact of scale differences when comparing model outputs with point-scale observations (Molotch & Bales, 2005; Trujillo & Lehning, 2015). We've taken measures to address the scale mismatch issue, including incorporating as many observation sites as possible and comparing the simulations with observations using multi-site averaging. Nevertheless, we advise readers to remain mindful of this scale-related limitation when interpreting the results.

3. Methodology

3.1. Experimental Design

Figure 2 presents the flowchart of our research, which outlines the three main steps involved in our study. An overview of the experimental design is described below.

Step 1: Evaluation of AORC-driven SWE. We evaluated the accuracy of NWM-simulated SWE driven by AORC forcings (CTL experiment) by comparing them to SNOTEL site observations. This helped

Table 1
Experiments With Different Adjusted Forcing Variables

Experiment ID	Experiment name	Adjusted forcing variables
1	CTL	None
2	adj_prec	Precipitation
3	adj_temp	Temperature
4	adj_both	Precipitation and temperature

determine the limitations of NWM and the forcings for SWE simulation. We focused on the uncertainties in precipitation and air temperature due to the continuous observations available from SNOTEL. We aggregated hourly AORC precipitation to daily and averaged hourly AORC air temperature to daily for comparison. We extracted daily SWE simulations, as well as daily precipitation and air temperature forcings for grid cells containing SNOTEL sites. The daily AORC precipitation and temperature forcings, AORC-driven SWE simulations, and corresponding SNOTEL observations are then averaged across all sites for comparison.

Step 2: The role of forcing. To improve the accuracy of SWE simulation, we corrected the biases in AORC precipitation and temperature forcings for grid cells with available site observations. We disaggregated observed daily precipitation into hourly data, using either the diurnal distribution of the forcing precipitation (when daily precipitation of the forcing is greater than 0) or a uniform distribution (when daily precipitation of the forcing equals 0). The newly created hourly precipitation data based on observations was then substituted for the hourly AORC forcing precipitation data. Hourly forcing temperature data was scaled to match the daily mean of observations. We conducted three experiments using different adjusted forcing variables in conjunction with the CTL experiment (Table 1) to assess the impact of precipitation and temperature on SWE simulation. These experiments used the same configurations, with only variations in precipitation and/or temperature inputs.

Step 3: The role of parameterization. We conducted 72 experiments to analyze how different parameterizations affect SWE simulation, focusing on five snow-related physical processes. Table 2 shows the parameterization schemes for these five physical processes, and a brief overview of the physical processes and their corresponding parameterization schemes is provided in Appendix A.

3.2. National Water Model Configuration

We utilized the NWM v2.1 for hydrologic simulation in UCRB. The model's core is WRF-Hydro (Gochis et al., 2020), which is a physically based distributed hydrologic model configured to use Noah-MP to simulate land surface processes. We used the default model parameterization schemes (Table S1 in Supporting Information S1) and parameters (available at https://water.noaa.gov/about/nwm) of the NWM v2.1 for our study. We set the land surface processes resolution to 1-km with an hourly timestep. The subgrid overland routing was executed on a 250-m grid mesh, with terrain and channel routing timesteps of 10 and 300 s, respectively. We used NCEP Final Operational Global Analysis data (National Centers for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce, 2000) as the initial conditions and spun up the model for 20 years (five cycles for WYs 2016–2019) to reach an equilibrium state for soil temperature, moisture, and groundwater.

Table 2
Noah-MP Parameterization Schemes Investigated in This Study

Physical process	Parameterization schemes
Surface exchange coefficient for heat (SFC)	1. Monin–Obukhov (Monin & Obukhov, 1954; default)
	2. Chen97 (Chen et al., 1997)
Snow surface albedo (ALB)	1. BATS (Yang et al., 1997; default)
	2. CLASS (Verseghy, 1991)
Rainfall and snowfall partitioning (SNF)	1. Jordan91 (Jordan, 1991; default)
	2. BATS (Dickinson et al., 1986)
	3. Noah (Chen et al., 1996)
Lower boundary of soil temperature (TBOT)	1. Zero-flux (Niu et al., 2011)
	2. Noah (H.-L. Pan & Mahrt, 1987; default)
Snow/soil temperature time scheme (STC)	1. Semi-implicit (Yang, Cai, et al., 2011; Yang, Niu, et al., 2011)
	2. Fully implicit (H.-L. Pan & Mahrt, 1987)
	3. Modified semi-implicit (Yang, Cai, et al., 2011; Yang, Niu, et al., 2011; default)

We conducted experiments for WYs 2016–2019, excluding results from the first WY for analysis to minimize initial condition uncertainty.

3.3. Evaluation Metrics

Model performances were evaluated against different sources of reference data using bias (BIAS), relative bias (RB), root-mean-square error (RMSE), Pearson correlation coefficient (CC; Pearson, 1895), and normalized standard deviation (NSD).

$$\mathrm{BIAS} = \frac{1}{N}\sum_{i=1}^{N}(S_i - O_i) \tag{1}$$

$$\mathrm{RB} = \frac{\frac{1}{N}\sum_{i=1}^{N}(S_i - O_i)}{\overline{O}} \times 100 \tag{2}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(S_i - O_i)^2} \tag{3}$$

$$\mathrm{CC} = \frac{\sum_{i=1}^{N}(O_i - \overline{O})(S_i - \overline{S})}{\sqrt{\sum_{i=1}^{N}(O_i - \overline{O})^2}\sqrt{\sum_{i=1}^{N}(S_i - \overline{S})^2}} \tag{4}$$

$$\mathrm{NSD} = \frac{\sqrt{\sum_{i=1}^{N}(S_i - \overline{S})^2}}{\sqrt{\sum_{i=1}^{N}(O_i - \overline{O})^2}} \tag{5}$$

where S_i and O_i are the simulated and observed values at time i, respectively; N is the number of time points; and $\overline{S}$ and $\overline{O}$ are the mean values of simulations and observations, respectively. The primary evaluation metric is RMSE, offering a more precise measure of the error between simulated and observed data. BIAS and RB are employed to assess whether the model exhibits positive or negative bias, while CC and NSD are utilized in the Taylor diagram (Taylor, 2001).

3.4. Analysis Methods

For most evaluations, we directly compare the evaluation metrics to assess the performance of different experiments. To objectively evaluate the distinctions among the large ensemble of model physical parameterization configurations, it's essential to establish a sensitivity index. Assume that there are m distinct physical processes (here $m = 5$), each with various parameterization schemes (2 or 3 for different processes). The mean value of the evaluation metric (RMSE in this paper) for each specific scheme j ($j = 1, 2, \cdots$) within a given process i ($i = 1, 2, \cdots, m$) can be represented as $\overline{Y}_j^{(i)}$. We defined an index to quantify the sensitivities of these physical processes as follows:

$$S_i = \frac{\Delta\overline{Y}^{(i)}}{\max\left\{\Delta\overline{Y}^{(1)}, \Delta\overline{Y}^{(2)}, \cdots, \Delta\overline{Y}^{(m)}\right\}} \tag{6}$$

where $\Delta\overline{Y}^{(i)} = \overline{Y}_{\max}^{(i)} - \overline{Y}_{\min}^{(i)}$ is the difference between the largest and the smallest mean values of the evaluation metric for the i th process.

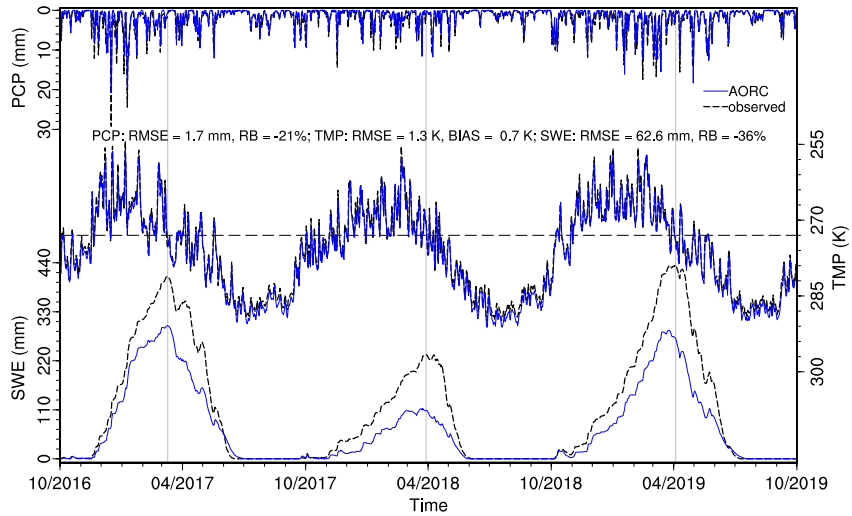

Figure 3. Comparison of mean daily AORC precipitation (PCP; top), air temperature (TMP; middle), and AORC-driven snow water equivalent (SWE; bottom) with SNOTEL observations. The horizontal dashed black line represents the freezing temperature (273.15 K), while the vertical gray lines segment the water years into snow accumulation and ablation periods based on observed peak SWE dates. The statistics illustrate the differences between AORC data and corresponding SNOTEL observations.

4. Results and Discussion

4.1. Evaluation of AORC-Driven SWE

We first assessed AORC precipitation and air temperature, as well as NWM-simulated SWE driven by AORC forcings (CTL experiment), against the SNOTEL observations. Figure 3 shows a comparison of mean daily precipitation, air temperature, and SWE between all SNOTEL sites and their corresponding grid cells. While NWM captures the overall temporal variation of SWE, it underestimates its amount by 36% (RMSE = 62.6 mm). This underestimation is partially attributed to a 21% precipitation underestimate (RMSE = 1.7 mm) and a slight air temperature overestimate of 0.7 K (RMSE = 1.3 K), despite which may fall within the range of instrument noise for some sites. An additional factor could be model deficiencies, evident in the notably higher peak SWE bias compared to the accumulated precipitation bias during the snow accumulation period (110.7 vs. 109.9 mm for WY 2017, 120.5 vs. 96.6 mm for WY 2018, and 146.2 vs. 138.7 mm for WY 2019). Furthermore, slower snow ablation is noted in WY 2017 (July 4 vs. June 18 for the snow ending date) and WY 2019 (July 18 vs. July 9 for the snow ending date), while faster snow ablation occurs in WY 2018 (May 29 vs. June 5 for the snow ending date) compared to the observations. These discrepancies may stem from model deficiencies in snow ablation physics, including aspects such as snow albedo and net radiation (Xiao et al., 2021), as well as errors in the forcing data, such as downward shortwave radiation. These findings align with studies by Chen et al. (2014) and He et al. (2019), which, despite employing different forcings and models, similarly attribute the SWE underestimation to forcing errors and model deficiencies.

Figure 4 presents the Taylor diagram for mean daily precipitation, air temperature, and SWE across different WYs. Among these three variables, AORC air temperature demonstrates the best agreement with the observations, only slightly overestimating them (with a normalized standard deviation of 1.02–1.03) and exhibiting a high correlation of 0.99 for all WYs. AORC precipitation, on the other hand, shows a relatively small negative bias (with a normalized standard deviation of 0.89–0.93) but the lowest correlation (0.75–0.84) compared to the other variables. While AORC-driven SWE maintains high correlations with the observations (0.97–0.99), it experiences the most significant negative bias (with a normalized standard deviation of 0.48–0.72) among the three variables. Additionally, WY 2018 exhibits the most significant negative bias in SWE compared to WYs 2017 and 2019, which could be attributed to the relatively higher bias of precipitation in this year as indicated by both the normalized standard deviation (0.89 vs. 0.90 and 0.93) and correlation (0.75 vs. 0.79 and 0.84).

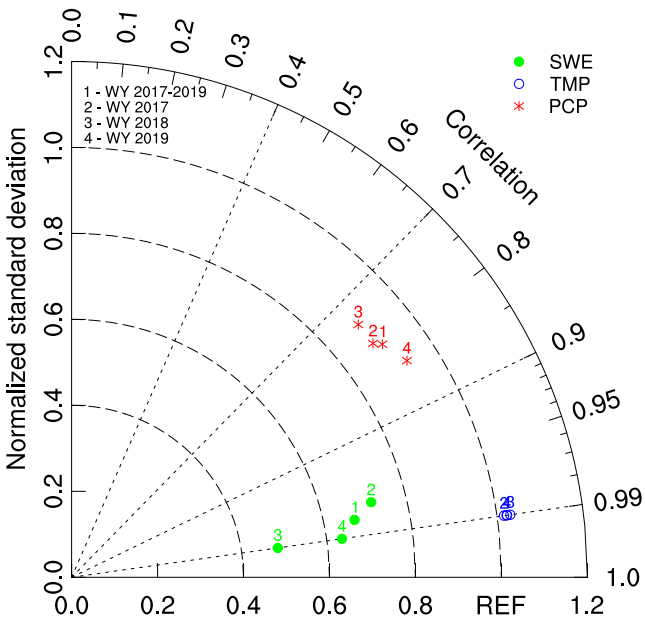

Figure 4. Taylor diagram showing the statistics between mean AORC data and corresponding SNOTEL observations for daily precipitation (PCP), air temperature (TMP), and snow water equivalent (SWE) of different water years. The radial distance from the origin to the points shows the normalized standard deviation and the azimuthal position of the points displays the correlation coefficient.

4.2. The Role of Forcing

We then examined the impact of bias correction applied to AORC precipitation and temperature on the simulated SWE. Figures S1 and S2 in Supporting Information S1 illustrate the comparison between the adjusted and default AORC forcing data against site observations for precipitation and air temperature, respectively. AORC precipitation displays the most substantial mean bias in March (−0.8 mm/d) and the least mean bias in June–August (−0.1 mm/d). Conversely, AORC air temperature shows its largest mean bias in August (1.2 K) and the least mean bias in March (0.1 K). The adjusted AORC forcings align seamlessly with the observations, with a correlation of 1 and bias of 0, in contrast to the default AORC forcings.

Figure 5 compares the mean daily SWE across all sites for four experiments (CTL, adj_prec, adj_temp, and adj_both) against observations. Notably, adjusting precipitation (adj_prec) results in a significant enhancement in SWE simulation. It minimizes RMSE by 66% (from 62.6 to 21.3 mm) and decreases bias by 87% (from −40.1 to −5.2 mm). Similarly, adjusting temperature (adj_temp) also has a positive impact on SWE simulation. It trims RMSE by 10% (from 62.6 to 56.6 mm) and lowers bias by 12% (from −40.1 to −35.5 mm). The adjustment of both precipitation and temperature (adj_both) exhibits remarkable improvements in SWE simulation. It reduces RMSE by 69% (from 62.6 to 19.2 mm) and decreases bias by 98% (from −40.1 to 0.8 mm). This implies that SWE simulations are more responsive to AORC precipitation adjustments compared to adjustments in air temperature.

4.3. The Role of Parameterization

4.3.1. Sensitivity Analysis of the Physical Processes

We investigated the impact of the parameterization schemes on SWE simulation by comparing 72 ensemble experiments driven by adjusted AORC forcings (both precipitation and temperature). Figure 6 shows the sensitivity indices of the five physical processes, calculated based on the RMSE of the simulated SWE. Among these processes, STC demonstrates the highest sensitivity, while TBOT is the least sensitive. Following in

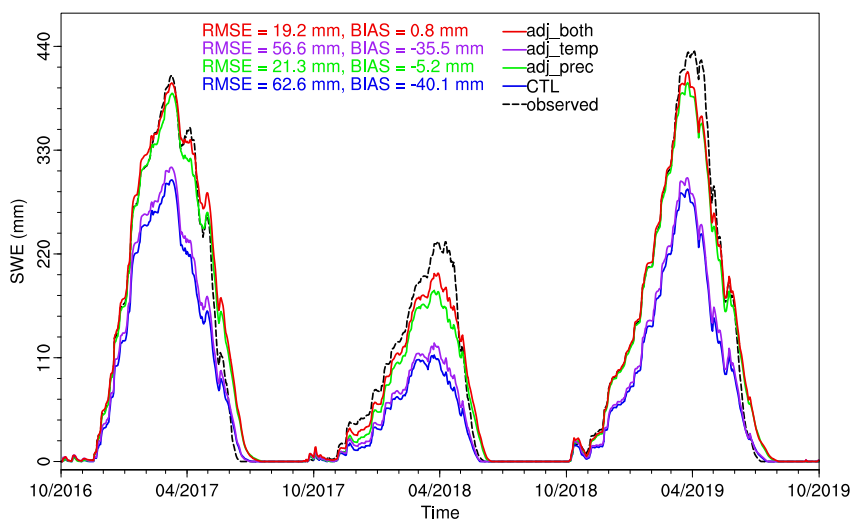

Figure 5. Mean daily SWE across all sites for four experiments with varying forcings. The CTL experiment used the default AORC forcings, the adj_prep experiment adjusted AORC precipitation, the adj_temp experiment adjusted AORC temperature, and the adj_both experiment adjusted both AORC precipitation and temperature for the simulations. All four experiments employed the default parameterization scheme combination of the NWM.

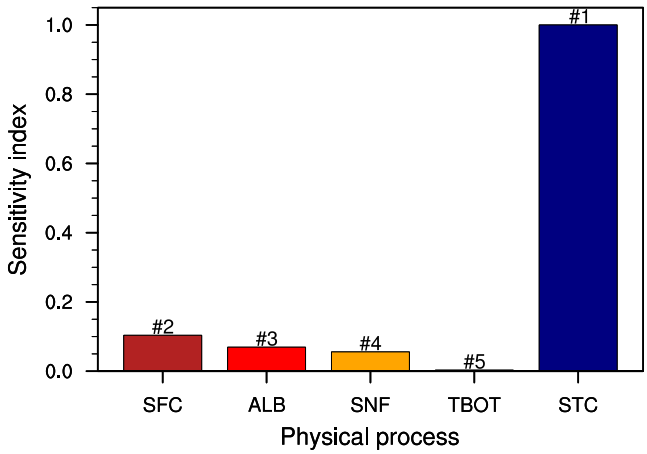

Figure 6. Sensitivity indices of the five physical processes. SFC = surface exchange coefficient for heat; ALB = snow surface albedo; SNF = rainfall and snowfall partitioning; TBOT = lower boundary of soil temperature; STC = snow/soil temperature time scheme.

sensitivity are SFC, ALB, and SNF, ranked as the second, third, and fourth most sensitive physical processes, respectively. While their differences are not substantial, all exhibit much lower sensitive compared to STC.

4.3.2. Combinatorial Optimization of the Parameterization Schemes

Figure 7 displays the distribution of RMSEs for all sites mean SWE time series of the ensemble experiments categorized by the five selected physical processes (left panel) and all sites mean SWE time series of the ensemble experiments classified by parameterization schemes of the identified sensitive physical process (right panel). By comparing the performance of different schemes, we can identify underperforming schemes and eliminate them, reducing uncertainty in subsequent analysis.

The first discarded scheme is STC-2, with a considerably higher mean RMSE (243.5 mm) compared to STC-1 (19.5 mm) and STC-3 (19.1 mm), as depicted in Figure 7a. STC-2 generates a more prolonged snow season with a higher and later SWE peak, significantly deviating from observations (Figure 7f). This finding aligns with previous studies by You et al. (2020) and Li et al. (2021), and is attributed to STC-2's (the fully implicit scheme) tendency to generate larger coefficients B in the thermal diffusion equation, resulting in smaller increments for the snow surface temperature, which leads to more extensive snow cover and delayed snow ablation.

The second discarded scheme is SNF-3, with a mean RMSE of 23.5 mm, higher than SNF-1 and SNF-2 (17.2 mm for both), as evident in Figure 7b. Additionally, its time series displays a greater negative bias compared to the other two schemes (Figure 7g). This is because SNF-3 (Noah) partitions less precipitation into snowfall due to its lower air temperature threshold. Despite outperforming SNF-3 in precipitation phase partitioning, SNF-1 and SNF-2 may still underestimate snowfall because precipitation is likely to fall as snow in warm and dry air conditions, while a low air temperature-based threshold may only represent cold and humid air conditions (Y.-H. Wang et al., 2019). Precipitation phase is influenced by various factors beyond just temperature, such as humidity, air pressure, and precipitation rates (Harpold et al., 2017). To improve the accuracy of snow simulation, incorporating humidity information into precipitation phase prediction models has been suggested, which could potentially increase the snow fraction (Jennings et al., 2018; Y.-H. Wang et al., 2019).

The third discarded scheme is SFC-1, with a mean RMSE of 18.4 mm compared to 15.9 mm for SFC-2 (Figure 7c). While SFC-1's temporal variation is comparable to SFC-2, it has a larger negative bias during the snow accumulation period and a larger positive bias during the late snow ablation period (Figure 7h), which aligns with Zhang et al. (2016) and Li et al. (2022). This is because SFC-2 (Chen97) produces a higher C_H than SFC-1 (M–O) due to the application of a different stability correction and the inclusion of planetary boundary layer effects (Yang, Cai, et al., 2011; Yang, Niu, et al., 2011), which results in a more efficient land surface ventilation and lower surface skin temperature (Niu et al., 2011). While Niu et al. (2011), Yang, Cai, et al. (2011), and Yang, Niu, et al. (2011) have reported that Chen97 may lead to greater sublimation, potentially resulting in less snow mass in certain river basins, we observed that in the case of UCRB, Chen97 generates more SWE due to its lower melting rate.

The fourth discarded scheme is ALB-1, which yields a larger mean RMSE than ALB-2 (17. 5 vs. 14.3 mm, as shown in Figure 7d). While ALB-1 (BATS) results in higher SWE than ALB-2 (CLASS), it prolongs the ablation period, leading to a greater bias at the end of the ablation season (Figure 7i). This is attributed to the BATS scheme producing a slightly higher snow surface albedo and, consequently, retaining more snow than the CLASS scheme, primarily due to its weaker snow aging effects (Niu et al., 2011). Molotch and Bales (2006) have noted that BATS tends to overestimate spring snow albedo due to an inclination to overestimate albedo for large grains, which are more prevalent during ablation periods. Although the simpler CLASS scheme can adequately simulate average albedo decay rates, it falls short in capturing the temporal variability in decay rates as influenced by time-change factors. This is because it assumes that optical or environmental factors affecting snow albedo remain constant over time (Abolafia-Rosenzweig et al., 2022). Optimizing BATS albedo parameters holds the potential to

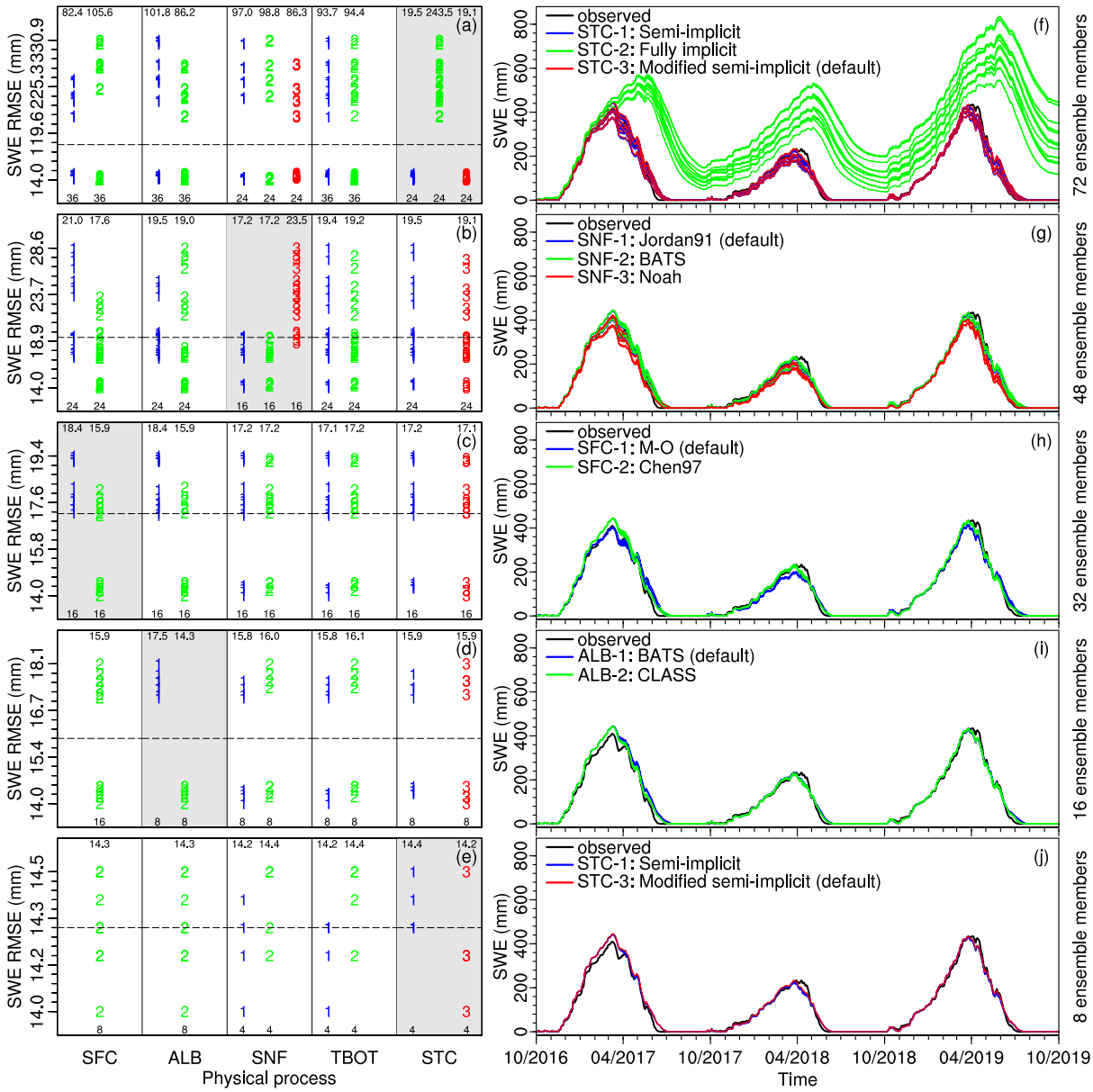

Figure 7. (a–e) Distribution of RMSEs for all sites mean daily SWE time series of the ensemble experiments categorized by the selected physical processes and (f–j) all sites mean daily SWE time series of the ensemble experiments classified by parameterization schemes of the identified sensitive physical process (highlighted with a gray background in the left panel). In each subfigure on the left panel, the horizontal dashed line represents the mean RMSE of all ensemble experiments, colored numbers denote different parameterization schemes, the numbers at the bottom show the counts of ensemble members for corresponding schemes, and the numbers at the top show the average RMSEs of all ensemble members with corresponding schemes. SFC = surface exchange coefficient for heat; ALB = snow surface albedo; SNF = rainfall and snowfall partitioning; TBOT = lower boundary of soil temperature; STC = snow/soil temperature time scheme.

enhance the simulation of albedo decay, as it accounts for complex factors like snow metamorphism and impurities that impact the evolution of snow albedo (Abolafia-Rosenzweig et al., 2022).

The final scheme that could be discarded is STC-1, with a slightly larger mean RMSE than STC-3 (14.4 vs. 14.2 mm, as seen in Figure 7e), although the difference in time series is not significant (Figure 7j). The difference between STC-1 and STC-3 lies in how they handle ground surface temperature for grid cells with snow but above freezing. STC-1 sets the whole grid cell to freezing temperature, while STC-3 only sets the snow-covered part to freezing temperature, producing more realistic ground surface temperature.

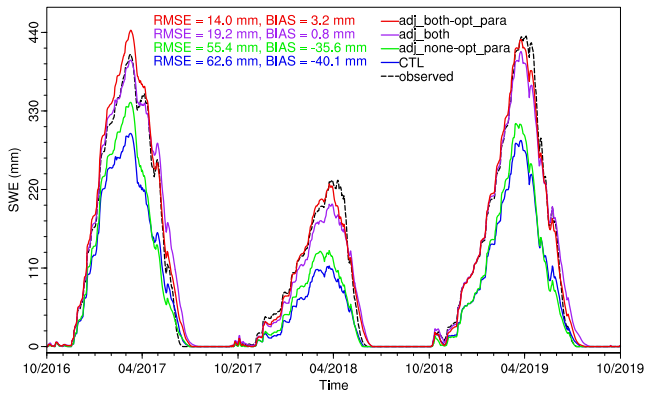

Figure 8. Mean daily SWE across all sites for four experiments with varying parameterization schemes and forcings. The CTL experiment employed the default AORC forcings and parameterization scheme combination, the adj_none-opt_para experiment utilized default AORC forcings and the optimized parameterization scheme combination, the adj_both experiment adjusted both AORC precipitation and temperature while using the default parameterization scheme combination, and the adj_both-opt_para experiment adjusted both AORC precipitation and temperature while utilizing the optimized parameterization scheme combination for simulations.

The optimal schemes identified for each process are SFC-2, ALB-2, SNF-1/SNF-2, TBOT-1/TBOT-2, and STC-3, respectively. This leads to four combinations that exhibit no significant difference in SWE simulation. The ultimate optimized scheme combination, yielding the smallest SWE bias, consists of SFC-2, ALB-2, SNF-1, TBOT-1, and STC-3.

4.3.3. Comparison Between the Optimized and Default Scheme Combinations

The comparison of mean daily SWE between the optimized and default scheme combinations is depicted in Figure 8. Two sets of comparisons, namely CTL versus adj_none-opt_para and adj_both versus adj_both-opt_para, were conducted to illustrate the impact of the optimized scheme combination. The first set employed the default AORC forcings, while the second set used the adjusted AORC forcings (adjusting both precipitation and temperature). The RMSEs for different experiments across different water years are given in Table S2 in Supporting Information S1.

The optimized scheme combination demonstrates improved performance, reducing the RMSE by 12% with the default AORC forcings (from 62.6 to 55.4 mm) and by 27% with the adjusted AORC forcings (from 19.2 to 14.0 mm) compared to the default scheme combination. In comparison to the CTL experiment, the optimized scheme combination coupled with adjusted AORC forcings (adj_both-opt_para) impressively reduces the RMSE by 78% (from 62.6 to 14.0 mm). Despite producing a larger positive bias between February and April 2017, the optimized scheme combination generally agrees better with the observations, not only during the accumulation period but also throughout the snow ablation period. The key differences between the default and optimized scheme combinations lie in the selection of SFC and ALB schemes. The optimized scheme combination utilizes SFC-2 and ALB-2 instead of SFC-1 and ALB-1 in the default scheme combination. This is because SFC-2 (Chen97) accumulates more snow during the accumulation period due to its lower melting rate (Figure 7h), and ALB-2 (CLASS) melts more snow during the ablation period due to its slightly lower snow surface albedo (Figure 7i) compared to their counterparts.

One notable observation from all experiments, whether with or without adjusted forcings and/or optimized parameterizations, is the model's manifestation of slow and late spring ablation, resulting in an extended snow season (Figure 8). This phenomenon aligns with findings from the SnowMIP2 project (Rutter et al., 2009) and is supported by Chen et al. (2014). The persistence of this behavior may be attributed to model deficiencies in snow ablation physics (He et al., 2019) and holds the potential for improvement through optimization of the snow albedo scheme (Abolafia-Rosenzweig et al., 2022).

5. Summary and Conclusions

This study aimed to assess the uncertainties of NWM in SWE simulation resulting from forcing and model uncertainties, using data from 46 SNOTEL sites within the middle elevation range of the UCRB. The influence of bias correction applied to AORC precipitation and air temperature was examined. The impact of parameterization selection on SWE simulation was also investigated by comparing 72 ensemble experiments of five physical processes. The key findings are summarized below:

1. NWM simulation driven by AORC forcings captured the overall temporal variation of SWE but underestimated its amount, yielding a mean RMSE of 62.6 mm/d.
2. Adjusting AORC precipitation with SNOTEL observations reduced SWE RMSE by 66%, adjusting the temperature trimmed it by 10%, and adjusting both decreased it by 69%.
3. Among the physical processes, snow/soil temperature time scheme (STC) demonstrated the highest sensitivity, followed by surface exchange coefficient for heat (SFC), snow surface albedo (ALB), and rainfall and snowfall partitioning (SNF), while the lower boundary of soil temperature (TBOT) proved to be insensitive.

4. Further optimization of parameterization scheme combination led to a 12% reduction in SWE RMSE. When combined with bias-corrected AORC precipitation and temperature, this optimization achieved a remarkable 78% reduction in SWE RMSE.

However, it is important to note that this study's scope is confined to the UCRB and may not be entirely generalizable to other regions. The development and implementation of high-resolution snow modeling frameworks, particularly for mountainous regions, remains a challenging but critical area of snow research. To further this field, there is a need to augment the quality of forcing data by incorporating more in situ observations, either directly or through climatological adjustments. This enhancement aims to improve the precision of high-resolution precipitation and temperature data in mountainous regions. Additionally, future research should focus on optimizing model structures and mitigating model parameterization uncertainties to bolster the accuracy of SWE simulations.

Appendix A: Physical Parameterization Schemes

A1. Surface Exchange Coefficient for Heat (SFC)

This process offers two methods for calculating heat exchange coefficients (C_H) between the land-surface and atmosphere: Monin–Obukhov (M–O) by Monin and Obukhov (1954) and Chen97 by Chen et al. (1997).

$$\text{M–O}: C_H = \frac{k^2}{\left[ln\left(\frac{z-d_0}{z_{0m}}\right) - \psi_m\left(\frac{z-d_0}{L}\right)\right]\left[ln\left(\frac{z-d_0}{z_{0h}}\right) - \psi_h\left(\frac{z-d_0}{L}\right)\right]} \tag{A1}$$

$$\text{Chen97}: C_H = \frac{k^2}{\left[ln\left(\frac{z}{z_{0m}}\right) - \psi_m\left(\frac{z}{L}\right) + \psi_m\left(\frac{z_{0m}}{L}\right)\right]\left[ln\left(\frac{z}{z_{0h}}\right) - \psi_h\left(\frac{z}{L}\right) + \psi_h\left(\frac{z_{0h}}{L}\right)\right]} \tag{A2}$$

where k is the von Kármán constant, L is the Monin–Obukhov length (m), z is the reference height (m), z_{0m} and z_{0h} are roughness lengths for momentum and heat, respectively (m), and d_0 is the zero-displacement height for the M–O scheme (m). Both methods take the same stability correction functions ψ_m and ψ_h for stable and unstable conditions, respectively. The difference between them is that M–O accounts for d_0 and assumes $z_{0m} = z_{0h}$, while Chen97 accounts for the difference between z_{0m} and z_{0h} but does not account for d_0.

A2. Snow Surface Albedo (ALB)

This process provides two methods for calculating snow surface albedo: BATS by Yang et al. (1997) and CLASS by Verseghy (1991).

BATS scheme calculates snow albedos for both visible and near-infrared bands under both direct and diffuse radiations. For diffuse radiation, snow albedos for visible and near-infrared bands are calculated by

$$\begin{cases} \alpha_s^{\text{dif,VIS}} = 0.95\left(1 - 0.2 f_{\text{age}}\right) \\ \alpha_s^{\text{dif,NIR}} = 0.65\left(1 - 0.5 f_{\text{age}}\right) \end{cases} \tag{A3}$$

where snow age function f_{age} is parameterized as

$$f_{\text{age}} = \frac{\tau^t}{\tau^t + 1} \tag{A4}$$

where t denotes current timestep and τ^t is defined as

$$\tau^t = \begin{cases} 0, & SWE^t \le 0 \\ \max\left(0, \left(\tau^{t-1} + \Delta\tau\right)(1 - \Delta s)\right), & SWE^t > 0 \end{cases} \tag{A5}$$

with

$$\Delta\tau = 10^{-6}\Delta t(r_1 + r_2 + r_3) \tag{A6}$$

and

$$\Delta s = \frac{\max\left(0, \mathrm{SWE}^{t} - \mathrm{SWE}^{t-1}\right)}{\mathrm{SWE}_c} \tag{A7}$$

where SWE^{t} and SWE^{t-1} are snow water equivalents of the current and previous timesteps, respectively; SWE_c is the critical value of new SWE to fully cover the old snow, which is set to 1 mm; Δt is the timestep; r_1 represents the effects of grain growth due to vapor diffusion; r_2 represents the additional effects of grain growth near or at the freezing of meltwater; and r_3 represents the effect of dirt and soot. The three parameters r_1, r_2, and r_3 are defined as

$$\begin{cases} r_1 = \exp\left(5000\left(\dfrac{1}{T_{frz}} - \dfrac{1}{T_g}\right)\right) \\ r_2 = \exp\left(\min\left(0, 10 \times 5000\left(\dfrac{1}{T_{frz}} - \dfrac{1}{T_g}\right)\right)\right) \\ r_3 = 0.3 \end{cases} \tag{A8}$$

where T_{frz} is the freezing temperature which is set to 273.16 K; and T_g is the ground temperature (K).

For the direct radiation, snow albedos for visible and near-infrared bands are calculated by

$$\begin{cases} \alpha_s^{\mathrm{dir,VIS}} = \alpha_s^{\mathrm{dif,VIS}} + 0.4 f_Z\left(1 - \alpha_s^{\mathrm{dif,VIS}}\right) \\ \alpha_s^{\mathrm{dir,NIR}} = \alpha_s^{\mathrm{dif,NIR}} + 0.4 f_Z\left(1 - \alpha_s^{\mathrm{dif,NIR}}\right) \end{cases} \tag{A9}$$

where f_Z is a factor accounts for the impact of solar zenith angle Z

$$f_Z = \max\left\{\frac{1.5}{1 + \cos Z} - 0.5, 0\right\} \tag{A10}$$

CLASS scheme assumes that the snow albedos for the visible and near-infrared bands under both direct and diffuse radiations are identical and can be calculated by

$$\alpha_s = \alpha_1 + \frac{\min(Q_s, \mathrm{SWE}_c/\Delta t)\,(0.84 - \alpha_1)}{\mathrm{SWE}_c/\Delta t} \tag{A11}$$

where Q_s is the snowfall rate (mm/s), SWE_c and Δt are stated before, and α_1 is parameterized as

$$\alpha_1 = 0.55 + \left(\alpha_s^{\mathrm{old}} - 0.55\right)\exp\left(-0.01\Delta t/3600\right) \tag{A12}$$

where α_s^{old} is the snow albedo from the previous timestep.

A3. Rainfall and Snowfall Partitioning (SNF)

This process has three methods to divide precipitation into rainfall and snowfall: Jordan91 (Jordan, 1991), BATS (Dickinson et al., 1986), and Noah (Chen et al., 1996). The proportion of snow (P_s) is calculated as follows:

$$\text{Jordan91}: P_s = \begin{cases} 0, & T_{sfc} > T_{frz} + 2.5 \\ 0.6, & T_{frz} + 2 < T_{sfc} \le T_{frz} + 2.5 \\ 1 - (-54.632 + 0.2T_{sfc}), & T_{frz} + 0.5 < T_{sfc} \le T_{frz} + 2 \\ 1, & T_{sfc} \le T_{frz} + 0.5 \end{cases} \tag{A13}$$

$$\text{BATS}: P_s = \begin{cases} 0, & T_{sfc} \ge T_{frz} + 2.2 \\ 1, & T_{sfc} < T_{frz} + 2.2 \end{cases} \tag{A14}$$

$$\text{Noah}: P_s = \begin{cases} 0, & T_{sfc} \ge T_{frz} \\ 1, & T_{sfc} < T_{frz} \end{cases} \tag{A15}$$

where T_{sfc} is the surface air temperature (K), and T_{frz} is the freezing temperature (K). A higher air temperature threshold leads to a higher snowfall rate.

A4. Lower Boundary of Soil Temperature (TBOT)

This process provides two methods for calculating heat flux from the bottom of soil column (F_b, W m^{-2}), which affects snow melting: zero-flux (Niu et al., 2011) and Noah (H.-L. Pan & Mahrt, 1987).

$$\text{Zero} - \text{flux}: F_b = 0 \tag{A16}$$

$$\text{Noah}: F_b = \frac{K_n(T_b - T_n)}{0.5(Z_{n-1} + Z_n) - Z_b} \tag{A17}$$

where n is the number of soil layers, K_n is the thermal conductivity for the n th soil layer (W m^{-1} K^{-1}), T_b is the bottom condition for soil temperature (K), T_n is the temperature for the n th soil layer (K), Z_{n-1} and Z_n are layer-bottom depths from snow surface for the (n^-1) th and n th soil layers (m), respectively, and Z_b is the bottom depth from the snow surface (m). The zero-flux scheme assumes that there is no heat flux from the bottom of the soil column, whereas the Noah scheme assumes that there is an existing heat flux at the 8 m soil depth.

A5. Snow/Soil Temperature Time Scheme (STC)

This process involves solving the following thermal diffusion equation using three numerical schemes: semi-implicit, fully implicit, and modified semi-implicit Crank-Nicolson methods.

$$C\frac{\partial T}{\partial t} = \frac{\partial}{\partial z}\left(K\frac{\partial T}{\partial z}\right) \tag{A18}$$

The thermal diffusion equation describes the temperature of snow/soil T (K) as a function of time t and snow/soil depth z (m), with C as the volumetric heat capacity (J m^{-3} K^{-1}) and K as the thermal conductivity (W m^{-1} K^{-1}).

The semi-implicit (or modified semi-implicit) and fully implicit schemes differ in their calculation of one of the coefficients (B) of the thermal diffusion equation for the surface snow layer.

$$\text{Semi} - \text{implicit/Modified semi} - \text{implicit}: B_m = \frac{K_m}{0.5 z_m z_{m+1} C_m} \tag{A19}$$

$$\text{Fully implicit}: B_m = \frac{K_m}{0.5 z_m z_{m+1} C_m} + \frac{K_m}{0.5 z_m z_m C_m} \tag{A20}$$

where m is the index for the surface snow layer (m could be −2, −1, and 0, which represents 3, 2, and 1 snow layer (s), respectively).

The semi-implicit and fully implicit methods handle the surface temperatures for vegetated ground ($T_{g,v}$, K) and bare ground ($T_{g,b}$, K) differently in comparison to the modified semi-implicit schemes. The semi-implicit and fully implicit methods set the ground temperatures to the freezing temperature (T_{frz}, K) if they are above it and the ground is covered by snow (snow depth >0.05 mm). The modified semi-implicit method uses a weighted average of the ground temperature and the freezing temperature.

$$\text{Semi}-\text{implicit/Fully implicit}: \begin{cases} T_{g,v} = T_{frz} \\ T_{g,b} = T_{frz} \end{cases} \tag{A21}$$

$$\text{Modified semi}-\text{implicit}: \begin{cases} T_{g,v} = (1 - F_s)\, T_{g,v} + F_s T_{frz} \\ T_{g,b} = (1 - F_s)\, T_{g,b} + F_s T_{frz} \end{cases} \tag{A22}$$

where F_s is the snow cover fraction.

Data Availability Statement

Software—The NWM used to generate the results is based on the WRF-Hydro (Gochis et al., 2020), which is accessible at https://ral.ucar.edu/projects/wrf_hydro. The WPS (W. Wang et al., 2017) was used to process the 1-km resolution land surface model static data and the WRF-Hydro GIS Pre-processor version 5.1.1 was used to process the 250-m resolution routing-related static data (Sampson & Gochis, 2020). Figures were created using ArcGIS version 10.7 (ESRI, 2019) and NCL version 6.4.0 (Meier-Fleischer et al., 2017). Data—The static data are available at https://water.noaa.gov/about/nwm. The AORC forcing data (Fall et al., 2023) are available at https://registry.opendata.aws/nwm-archive. The bias-corrected and quality-controlled SNOTEL observations (Yan et al., 2018) are available at https://www.pnnl.gov/data-products. The outputs of different experiments for the 46 SNOTEL sites are available in Gan (2024).

Acknowledgments
This work was partially supported by the National Oceanic and Atmospheric Administration (Grant NA18OAR4590410).

References

Abolafia-Rosenzweig, R., He, C., Skiles, S. M., Chen, F., & Gochis, D. (2022). Evaluation and optimization of snow albedo scheme in Noah-MP land surface model using in situ spectral observations in the Colorado Rockies. *Journal of Advances in Modeling Earth Systems*, *14*(10), e2022MS003141. https://doi.org/10.1029/2022MS003141

Beniston, M., Farinotti, D., Stoffel, M., Andreassen, L. M., Coppola, E., Eckert, N., et al. (2018). The European mountain cryosphere: A review of its current state, trends, and future challenges. *The Cryosphere*, *12*(2), 759–794. https://doi.org/10.5194/tc-12-759-2018

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Ménard, C. B., et al. (2011). The joint UK land environment simulator (JULES), model description—Part 1: Energy and water fluxes. *Geoscientific Model Development*, *4*(3), 677–699. https://doi.org/10.5194/gmd-4-677-2011

Bowling, L. C., Lettenmaier, D. P., Nijssen, B., Graham, L. P., Clark, D. B., El Maayar, M., et al. (2003). Simulation of high-latitude hydrological processes in the Torne–Kalix basin: PILPS Phase 2(e): 1: Experiment description and summary intercomparisons. *Global and Planetary Change*, *38*(1), 1–30. https://doi.org/10.1016/S0921-8181(03)00003-1

Broxton, P. D., Zeng, X., & Dawson, N. (2016). Why do global reanalyses and land data assimilation products underestimate snow water equivalent? *Journal of Hydrometeorology*, *17*(11), 2743–2761. https://doi.org/10.1175/JHM-D-16-0056.1

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*(24), 13795–13819. https://doi.org/10.1002/2014JD022167

Chen, F., Janjić, Z. I., & Mitchell, K. (1997). Impact of atmospheric surface-layer parameterizations in the new land-surface scheme of the NCEP mesoscale Eta model. *Boundary-Layer Meteorology*, *85*(3), 391–421. https://doi.org/10.1023/A:1000531001463

Chen, F., Mitchell, K., Schaake, J., Xue, Y., Pan, H., Koren, V., et al. (1996). Modeling of land surface evaporation by four schemes and comparison with FIFE observations. *Journal of Geophysical Research*, *101*(D3), 7251–7268. https://doi.org/10.1029/95JD02165

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*(7), W07539. https://doi.org/10.1029/2011WR010745

Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. A., et al. (2015). A unified approach for process-based hydrologic modeling: 1. Modeling concept. *Water Resources Research*, *51*(4), 2498–2514. https://doi.org/10.1002/2015WR017198

Cluzet, B., Lafaysse, M., Deschamps Berger, C., Vernay, M., & Dumont, M. (2022). Propagating information from snow observations with CrocO ensemble data assimilation system: A 10-years case study over a snow depth observation network. *The Cryosphere*, *16*(4), 1281–1298. https://doi.org/10.5194/tc-16-1281-2022

Danielson, J. J., & Gesch, D. B. (2011). *Global multi-resolution terrain elevation data 2010 (GMTED2010)*. U.S. Geological Survey.

Dickinson, R. E., Henderson-Sellers, A., Kennedy, P. J., & Wilson, M. F. (1986). *Biosphere-Atmosphere Transfer Scheme (BATS) for the NCAR community climate model*. NCAR. https://doi.org/10.5065/D6668B58

Dozier, J., Bair, E. H., & Davis, R. E. (2016). Estimating the spatial distribution of snow water equivalent in the world's mountains. *WIREs Water*, *3*(3), 461–474. https://doi.org/10.1002/wat2.1140

Echevers, P., Martin, E., Brown, R., Fierz, C., Lejeune, Y., Bazile, E., et al. (2004). Validation of the energy budget of an alpine snowpack simulated by several snow models (SnowMIP project). *Annals of Glaciology*, *38*(1), 150–158. https://doi.org/10.3189/172756404781814825

ESRI. (2019). ArcGIS desktop: Release 10.7 [Software]. Retrieved from https://www.esri.com/en-us/home

Essery, R., Morin, S., Lejeune, Y., & B Ménard, C. (2013). A comparison of 1701 snow models using observations from an alpine site. *Advances in Water Resources*, *55*, 131–148. https://doi.org/10.1016/j.advwatres.2012.07.013

Fall, G., Kitzmiller, D., Pavlovic, S., Zhang, Z., Patrick, N., St. Laurent, M., et al. (2023). The Office of Water Prediction's Analysis of Record for Calibration, version 1.1: Dataset description and precipitation evaluation [Dataset]. *Journal of the American Water Resources Association*, 1–27. https://doi.org/10.1111/1752-1688.13143

FAO. (1991). *The digitized soil map of the world*. FAO.

Fisher, R. A., & Koven, C. D. (2020). Perspectives on the future of land surface models and the challenges of representing complex terrestrial systems. *Journal of Advances in Modeling Earth Systems*, *12*(4), e2018MS001453. https://doi.org/10.1029/2018MS001453

Gan, Y. (2024). National Water Model simulated snow water equivalent data for the Upper Colorado River Basin SNOTEL sites [Dataset]. *Hydroshare*. Retrieved from http://www.hydroshare.org/resource/0bd94b2662cd4eee886420dba4b3d337

Gan, Y., Liang, X.-Z., Duan, Q., Chen, F., Li, J., & Zhang, Y. (2019). Assessment and reduction of the physical parameterization uncertainty for Noah-MP land surface model. *Water Resources Research*, *55*(7), 5518–5538. https://doi.org/10.1029/2019WR024814

Gan, Y., Zhang, Y., Liu, Y., Kongoli, C., & Grassotti, C. (2022). Assimilation of blended in situ-satellite snow water equivalent into the National Water Model for improving hydrologic simulation in two US river basins. *Science of the Total Environment*, *838*, 156567. https://doi.org/10.1016/j.scitotenv.2022.156567

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

Günther, D., Marke, T., Essery, R., & Strasser, U. (2019). Uncertainties in snowpack simulations—Assessing the impact of model structure, parameter choice, and forcing data error on point-scale energy balance snow model performance. *Water Resources Research*, *55*(4), 2779–2800. https://doi.org/10.1029/2018WR023403

Gutman, G., & Ignatov, A. (1998). The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models. *International Journal of Remote Sensing*, *19*(8), 1533–1543. https://doi.org/10.1080/014311698215333

Hammond, J. C., & Kampf, S. K. (2020). Subannual streamflow responses to rainfall and snowmelt inputs in snow-dominated watersheds of the western United States. *Water Resources Research*, *56*(4), e2019WR026132. https://doi.org/10.1029/2019WR026132

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

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*(23), 12631–12654. https://doi.org/10.1029/2019JD030823

James, T., Evans, A., Madly, E., & Kelly, C. (2014). *The economic importance of the Colorado River to the basin region*. Arizona State University.

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 Region Research and Engineering Laboratory, U.S. Army Corps of Engineers.

Kalra, A., Sagarika, S., Pathak, P., & Ahmad, S. (2017). Hydro-climatological changes in the Colorado River Basin over a century. *Hydrological Sciences Journal*, *62*(14), 2280–2296. https://doi.org/10.1080/02626667.2017.1372855

Kim, R. S., Kumar, S., Vuyovich, C., Houser, P., Lundquist, J., Mudryk, L., et al. (2021). Snow ensemble uncertainty project (SEUP): Quantification of snow water equivalent uncertainty across North America via ensemble land surface modeling. *The Cryosphere*, *15*(2), 771–791. https://doi.org/10.5194/tc-15-771-2021

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*(12), 6163–6172. https://doi.org/10.1002/2017GL073551

Li, Q., Yang, T., & Li, L. (2022). Quantitative assessment of the parameterization sensitivity of the WRF/Noah-MP model of snow dynamics in the Tianshan Mountains, Central Asia. *Atmospheric Research*, *277*, 106310. https://doi.org/10.1016/j.atmosres.2022.106310

Li, X., Wu, T., Wu, X., Chen, J., Zhu, X., Hu, G., et al. (2021). Assessing the simulated soil hydrothermal regime of the active layer from the Noah-MP land surface model (v1.1) in the permafrost regions of the Qinghai–Tibet Plateau. *Geoscientific Model Development*, *14*(3), 1753–1771. https://doi.org/10.5194/gmd-14-1753-2021

Lin, Y., & Mitchell, K. E. (2005). The NCEP stage II/IV hourly precipitation analyses: Development and applications. In *Paper presented at 19th Conference on Hydrology*.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L., & Merchant, J. W. (2000). Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*, *21*(6–7), 1303–1330. https://doi.org/10.1080/014311600210191

Magnusson, J., Gustafsson, D., Hüsler, F., & Jonas, T. (2014). Assimilation of point SWE data into a distributed snow cover model comparing two contrasting methods. *Water Resources Research*, *50*(10), 7816–7835. https://doi.org/10.1002/2014WR015302

McKay, L., Bondelid, T., Dewald, T., Johnston, J., Moore, R., & Rea, A. (2012). *NHDPlus version 2: User guide*. U.S. Environmental Protection Agency.

Meier-Fleischer, K., Böttinger, M., & Haley, M. (2017). The NCAR command language (version 6.4.0) [Software]. *UCAR/NCAR/CISL/VETS*. https://doi.org/10.5065/D6WD3XH5

Mitchell, K. E., Lohmann, D., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., et al. (2004). The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *Journal of Geophysical Research*, *109*(D7), D07S90. https://doi.org/10.1029/2003JD003823

Molotch, N. P., & Bales, R. C. (2005). Scaling snow observations from the point to the grid element: Implications for observation network design. *Water Resources Research*, *41*(11), W11421. https://doi.org/10.1029/2005WR004229

Molotch, N. P., & Bales, R. C. (2006). Comparison of ground-based and airborne snow surface albedo parameterizations in an alpine watershed: Impact on snowpack mass balance. *Water Resources Research*, *42*(5), W05410. https://doi.org/10.1029/2005WR004522

Monin, A. S., & Obukhov, A. M. (1954). Basic laws of turbulent mixing in the surface layer of the atmosphere. *Transactions of the Geophysics Institute of the Academy of Sciences of the USSR*, *24*(151), 163–187. (in Russian).

Mortimer, C., Mudryk, L., Derksen, C., Luojus, K., Brown, R., Kelly, R., & Tedesco, M. (2020). Evaluation of long-term Northern Hemisphere snow water equivalent products. *The Cryosphere*, *14*(5), 1579–1594. https://doi.org/10.5194/tc-14-1579-2020

National Centers for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce. (2000). *NCEP FNL operational model global tropospheric analyses, continuing from July 1999*. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. https://doi.org/10.5065/D6M043C6

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

Office of Water Prediction. (2021). *Analysis of Record for Calibration version 1.1—Sources, methods, and verification*. National Weather Service.

Pan, H.-L., & Mahrt, L. (1987). Interaction between soil hydrology and boundary-layer development. *Boundary-Layer Meteorology*, *38*(1), 185–202. https://doi.org/10.1007/BF00121563

Pan, M., Sheffield, J., Wood, E. F., Mitchell, K. E., Houser, P. R., Schaake, J. C., et al. (2003). Snow process modeling in the North American land data assimilation system (NLDAS): 2. Evaluation of model simulated snow water equivalent. *Journal of Geophysical Research*, *108*(D22), 8850. https://doi.org/10.1029/2003JD003994

Pearson, K. (1895). Note on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, *58*, 240–242. https://doi.org/10.1098/rspl.1895.0041

Raby, J. W., Cai, H., Dawson, L., & Dumais, R. (2020). *An evaluation of the Unrestricted Mesoscale Analysis as gridded observations for spatial model verification*. DEVCOM Army Research Laboratory.

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

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C., et al. (2004). The global land data assimilation system. *Bulletin of the American Meteorological Society*, *85*(3), 381–394. https://doi.org/10.1175/BAMS-85-3-381

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

Sampson, K., & Gochis, D. (2020). WRF hydro GIS pre-processing tools, version 5.1.1 documentation [Software]. *NCAR*. Retrieved from https://github.com/NCAR/wrf_hydro_arcgis_preprocessor

Schaefer, G. L., & Paetzold, R. F. (2000). SNOTEL (SNOwpack TELemetry) and SCAN (soil climate analysis network). In *Paper presented at Automated Weather Stations for Applications in Agriculture and Water Resources Management: Current Use and Future Perspectives*. NB.

Sellers, P. J., Meeson, B. W., Closs, J., Collatz, J., Corprew, F., Dazlich, D., et al. (1996). The ISLSCP initiative I global datasets: Surface boundary conditions and atmospheric forcings for land-atmosphere studies. *Bulletin of the American Meteorological Society*, *77*(9), 1987–2006. https://doi.org/10.1175/1520-0477(1996)077<1987:TIIIGD>2.0.CO;2

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 telemetry (SNOTEL) data. *Water Resources Research*, *35*(7), 2145–2160. https://doi.org/10.1029/1999WR900090

Siirila Woodburn, E. R., Rhoades, A. M., Hatchett, B. J., Huning, L. S., Szinai, J., Tague, C., et al. (2021). A low-to-no snow future and its impacts on water resources in the western United States. *Nature Reviews Earth and Environment*, *2*(11), 800–819. https://doi.org/10.1038/s43017-021-00219-y

Slater, A. G., Schlosser, C. A., Desborough, C. E., Pitman, A. J., Henderson Sellers, A., Robock, A., et al. (2001). The representation of snow in land surface schemes: Results from PILPS 2(d). *Journal of Hydrometeorology*, *2*(1), 7–25. https://doi.org/10.1175/1525-7541(2001)002<0007:TROSIL>2.0.CO;2

Stewart, I. T., Cayan, D. R., & Dettinger, M. D. (2004). Changes in snowmelt runoff timing in western North America under a 'business as usual' climate change scenario. *Climatic Change*, *62*(1), 217–232. https://doi.org/10.1023/B:CLIM.0000013702.22656.e8

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

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, *106*(D7), 7183–7192. https://doi.org/10.1029/2000JD900719

Terzago, S., Andreoli, V., Arduini, G., Balsamo, G., Campo, L., Cassardo, C., et al. (2020). Sensitivity of snow models to the accuracy of meteorological forcings in mountain environments. *Hydrology and Earth System Sciences*, *24*(8), 4061–4090. https://doi.org/10.5194/hess-24-4061-2020

Trujillo, E., & Lehning, M. (2015). Theoretical analysis of errors when estimating snow distribution through point measurements. *The Cryosphere*, *9*(3), 1249–1264. https://doi.org/10.5194/tc-9-1249-2015

Trujillo, E., & Molotch, N. P. (2014). Snowpack regimes of the Western United States. *Water Resources Research*, *50*(7), 5611–5623. https://doi.org/10.1002/2013WR014753

Verseghy, D. L. (1991). CLASS—A Canadian land surface scheme for GCMs. I. Soil model. *International Journal of Climatology*, *11*(2), 111–133. https://doi.org/10.1002/joc.3370110202

Wang, W., Bruyère, C., Duda, M., Dudhia, J., Gill, D., Kavulich, M., et al. (2017). ARW version 3 modeling system user's guide [Software]. *NCAR*. Retrieved from https://github.com/wrf-model/WRF

Wang, Y., Xie, Z., Jia, B., Wang, L., Li, R., Liu, B., et al. (2020). Sensitivity of snow simulations to different atmospheric forcing data sets in the land surface model CAS-LSM. *Journal of Geophysical Research: Atmospheres*, *125*(16), e2019JD032001. https://doi.org/10.1029/2019JD032001

Wang, Y.-H., Broxton, P., Fang, Y., Behrangi, A., Barlage, M., Zeng, X., & Niu, G.-Y. (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

Winstral, A., Magnusson, J., Schirmer, M., & Jonas, T. (2019). The bias-detecting ensemble: A new and efficient technique for dynamically incorporating observations into physics-based, multilayer snow models. *Water Resources Research*, *55*(1), 613–631. https://doi.org/10.1029/2018WR024521

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., et al. (2012). Continental-scale water and energy flux analysis and validation for the North American land data assimilation system project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *Journal of Geophysical Research*, *117*(D3), D03109. https://doi.org/10.1029/2011JD016048

Xiao, M., & Lettenmaier, D. P. (2021). Atmospheric rivers and snow accumulation in the Upper Colorado River basin. *Geophysical Research Letters*, *48*(16), e2021GL094265. https://doi.org/10.1029/2021GL094265

Xiao, M., Mahanama, S. P., Xue, Y., Chen, F., & Lettenmaier, D. P. (2021). Modeling snow ablation over the mountains of the western United States: Patterns and controlling factors. *Journal of Hydrometeorology*, *22*(2), 297–311. https://doi.org/10.1175/JHM-D-19-0198.1

Yan, H., Sun, N., Wigmosta, M., Skaggs, R., Hou, Z., & Leung, R. (2018). Next-generation intensity-duration-frequency curves for hydrologic design in snow-dominated environments [Dataset]. *Water Resources Research*, *54*(2), 1093–1108. https://doi.org/10.1002/2017WR021290

Yang, Z.-L., Cai, X., Zhang, G., Tavakoly, A. A., Jin, Q., Meyer, L. H., & Guan, X. (2011). *The community Noah land surface model with multi-parameterization options (Noah-MP)*. The University of Texas at Austin.

Yang, Z.-L., Dickinson, R. E., Robock, A., & Vinnikov, K. Y. (1997). Validation of the snow submodel of the Biosphere–Atmosphere Transfer Scheme with Russian snow cover and meteorological observational data. *Journal of Climate*, *10*(2), 353–373. https://doi.org/10.1175/1520-0442(1997)010<0353:VOTSSO>2.0.CO;2

Yang, Z.-L., Niu, G.-Y., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., et al. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins. *Journal of Geophysical Research*, *116*(D12), D12110. https://doi.org/10.1029/2010JD015140

You, Y., Huang, C., Gu, J., Li, H., Hao, X., & Hou, J. (2020). Assessing snow simulation performance of typical combination schemes within Noah-MP in northern Xinjiang, China. *Journal of Hydrology*, *581*, 124380. https://doi.org/10.1016/j.jhydrol.2019.124380

Zhang, G., Chen, F., & Gan, Y. (2016). Assessing uncertainties in the Noah-MP ensemble simulations of a cropland site during the Tibet Joint International Cooperation program (JICA) field campaign. *Journal of Geophysical Research: Atmospheres*, *121*(16), 9576–9596. https://doi.org/10.1002/2016JD024928