Basin-scale Evaluation of the Noah-MP Land Surface Model for Runoff and Snow Generation in the Missouri River Basin: Insights and Recommendations for Parameterization Scheme Selection

Eunsaem Cho[1,2], Eunsang Cho[3], Carrie Vuyovich[1], Bailing Li[1,2], Jennifer M. Jacobs[4,5]

[1]Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA
[2]Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA
[3]Ingram School of Engineering, Texas State University, San Marcos, TX, USA
[4]Department of Civil and Environmental Engineering, University of New Hampshire, Durham, NH, USA
[5]Earth Systems Research Center, Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, Durham, NH, USA

Correspondence to: Eunsang Cho (eunsang.cho@txstate.edu)

Abstract. Process-based land surface models, such as the Noah-Multiparameterization (Noah-MP) model, are widely used for large-scale hydrologic simulations because of their flexibility in selecting multiple parameterization schemes. However, limited guidance on choosing appropriate configurations constrains their reliability in representing runoff and snowmelt dynamics across diverse land-cover and snow conditions. This study evaluates the default and four alternative parameterization schemes in the Noah-MP land surface model, including Runoff and Groundwater (RUN), Surface Exchange Coefficient for Heat (SFC), Frozen Soil Permeability (INF), and Snow/Soil Temperature Time Scheme (STC), across 50 Hydro-Climate Data Networks (HCDNs) in the Missouri River Basin. Model performance was evaluated using USGS streamflow observations and snow water equivalent (SWE) estimates from the University of Arizona dataset for 2014 to 2023. Results showed that the alternative schemes improved runoff simulation compared to the default scheme through different approaches representing key hydrological and thermodynamic processes. Specifically, the RUN, SFC, INF, and STC experiments improved the Kling–Gupta Efficiency (KGE) by 0.19, 0.38, 0.44, and 0.14, respectively in representative subbasins, through enhanced groundwater dynamics, reduced evapotranspiration bias, improved rapid runoff response, and more accurate SWE evolution. SWE evaluation further indicates that the STC experiment most effectively reduced SWE bias in snow-dominated subbasins, though improvements in SWE did not universally translate into better runoff simulation. These results highlight the importance of basin-specific parameterization schemes within Noah-MP to improve hydrological prediction and water management across diverse hydroclimatic regions. The findings further indicate optimal parameterization schemes for different climates, land cover, and snow regimes.

1 Introduction

Snowmelt is one of the most important hydrological processes in snow-dominated regions, as it governs the timing and magnitude of water availability for ecosystems, agriculture, and human consumption (Barnett et al., 2005; Qin et al., 2020; Siirila-Woodburn et al., 2021; Cho et al., 2023). The gradual release of snowmelt sustains streamflow during

dry seasons, replenishes groundwater aquifers, and supports downstream water supplies (Huntington and Niswonger, 2012; Carroll et al., 2019). Conversely, rapid or excessive snowmelt can trigger flooding, infrastructure damage, and cascading hazards (Vormoor et al., 2016; Barnhart et al., 2020). Capturing these dynamics requires a robust understanding of hydrological processes—including snow accumulation, energy balance, infiltration, and runoff generation—because they determine how snowmelt is translated into streamflow.

Over the past several decades, snowmelt-related flooding has repeatedly affected communities across the United States and Canada, further demonstrating the widespread risks associated with changes in snowpack and runoff regimes (Graybeal et al., 2006; Stadnyk et al., 2016; Cho et al., 2021). The devastating 2019 Midwest Spring Flood exemplified snowmelt's societal impact, as rain-on-snow conditions breached more than 50 levees, unleashed unprecedented Missouri River Basin damages, and amassed nearly $12 billion in losses (NCEI, 2020; Francis and Melliger, 2021). Such events underscore the vulnerability of agricultural, urban, and ecological systems to changes in snowpack dynamics and runoff generation, highlighting the need for robust modeling frameworks capable of capturing both seasonal snowpack accumulation and the rapid hydrological transitions associated with snowmelt and rain-on-snow events (Barnett et al., 2005; Musselman et al., 2018).

Process-based land surface models (LSMs) have been widely employed for operational flood forecasting and water supply predictions (Alfieri et al., 2013; Wood and Lettenmaier, 2006; Cosgrove et al., 2024). LSMs simulate coupled exchanges of energy, water, and carbon between the land and atmosphere, enabling both climate diagnostics and hydrological forecasting (Overgaard et al., 2006; Abramowitz et al., 2008; Fisher and Koven, 2020). The Noah with multi-parameterization options (Noah-MP) model is one of the most widely adopted frameworks due to its role as the hydrologic core of the U.S. National Water Model and its unique design that allows multiple parameterization options for key processes including snow physics, runoff generation, frozen soil dynamics, and canopy water and energy exchange (Niu et al., 2011; Cuntz et al., 2016; He et al., 2023; Cosgrove et al., 2024). Theoretically, this flexibility enables Noah-MP users to select schemes that best represent local environmental conditions.

In practice, however, most applications of Noah-MP rely on its default physics or employ a single parameterization scheme across large domains (Ma et al., 2017; Li et al., 2022; Yin et al., 2023). Sensitivity studies have shown that snowpack accumulation, soil freeze-thaw dynamics, and runoff generation are highly dependent on parameterization choices (You et al., 2020; Cao et al., 2024), yet guidance for selecting among schemes remains limited. Prior evaluations have tended to focus on global- or continental- scale domains, emphasizing broad performance rather than basin-specific applicability (Gan et al., 2019; You et al., 2020). Moreover, while some studies have assessed SWE simulation (You et al., 2020; Yang et al., 2023) or runoff calibration (Su et al., 2024) independently, comprehensive analyses linking snow dynamics and runoff processes across diverse basin settings remain scarce. This knowledge gap constrains the ability of modelers to optimize Noah-MP for snow-dominated basins where both water supply forecasting and flood risk reduction depend on accurate snowmelt representation.

Within this context, we focus on systematically evaluating four alternative Noah-MP parameterization schemes, each governing one specific physical process: (1) runoff and groundwater, (2) surface exchange coefficients, (3) frozen soil permeability, and (4) snow/soil temperature dynamics. Each experiment modifies a single parameterization scheme while keeping the other three at their default settings, enabling an isolated assessment of each scheme's contribution to model performance. The objectives of this study are threefold: (1) to quantify how Noah-MP parameterization schemes influence runoff simulation in the Missouri River Basin; (2) to evaluate how basin characteristics—including climate class, land cover, and snow regime—influence the effectiveness of alternative parameterization schemes in runoff simulation; and (3) to assess how alternative schemes improve SWE and snowmelt-driven runoff representation relative to default model settings. To achieve these objectives, we introduce the Missouri River Basin as the target study area in Section 2. Section 3 outlines the datasets used for runoff and SWE comparisons, and Section 4 describes the model configuration and experimental methods. The results of parameterization experiments are presented in Section 5, followed by a discussion of region-specific sensitivities, implications for scheme selection, and broader modeling challenges in Section 6. Section 7 provides concluding remarks, highlighting recommendations for Noah-MP parameterization schemes and future directions for improving snow and runoff simulation in land surface models.

2 Study Area

This study focuses on the Missouri River Basin (USGS hydrologic unit code HUC-2 10) and Hydro-Climate Data Network (HCDN) located in Missouri River Basin (Fig. 1a). This basin was selected because it is highly vulnerable to spring snowmelt floods, as demonstrated by the devastating 2019 Midwest Spring Flood (Francis and Melliger, 2021; Shirzaei et al., 2021; Velasquez et al., 2023). In addition, the Missouri River Basin climate ranges from alpine tundra in the Rocky Mountains, through semi-arid steppe across the Great Plains, to humid subtropical conditions in the eastern lowlands (Qiao et al., 2014; Wise et al., 2018). In Fig. 1b, the digital elevation model (DEM) highlights the Rocky Mountain region in western part of the basin, which plays a dominant role in snow accumulation and snowmelt-driven runoff in the Missouri River Basin. Land cover type also varies, transitioning from high-elevation evergreen forests to wide grasslands and croplands (Jordan et al., 2012; Ahiablame et al., 2017). Snow regimes are diverse, with tundra, boreal, and montane forest snow in the Rockies, prairie snow in the Great Plains, and ephemeral snow in the southern lowlands (Sturm and Liston, 2021). The combination of varied climate and land cover and substantial snow make the Missouri River Basin ideal for evaluating Noah-MP parameterization schemes relevant to snowmelt-driven runoff.

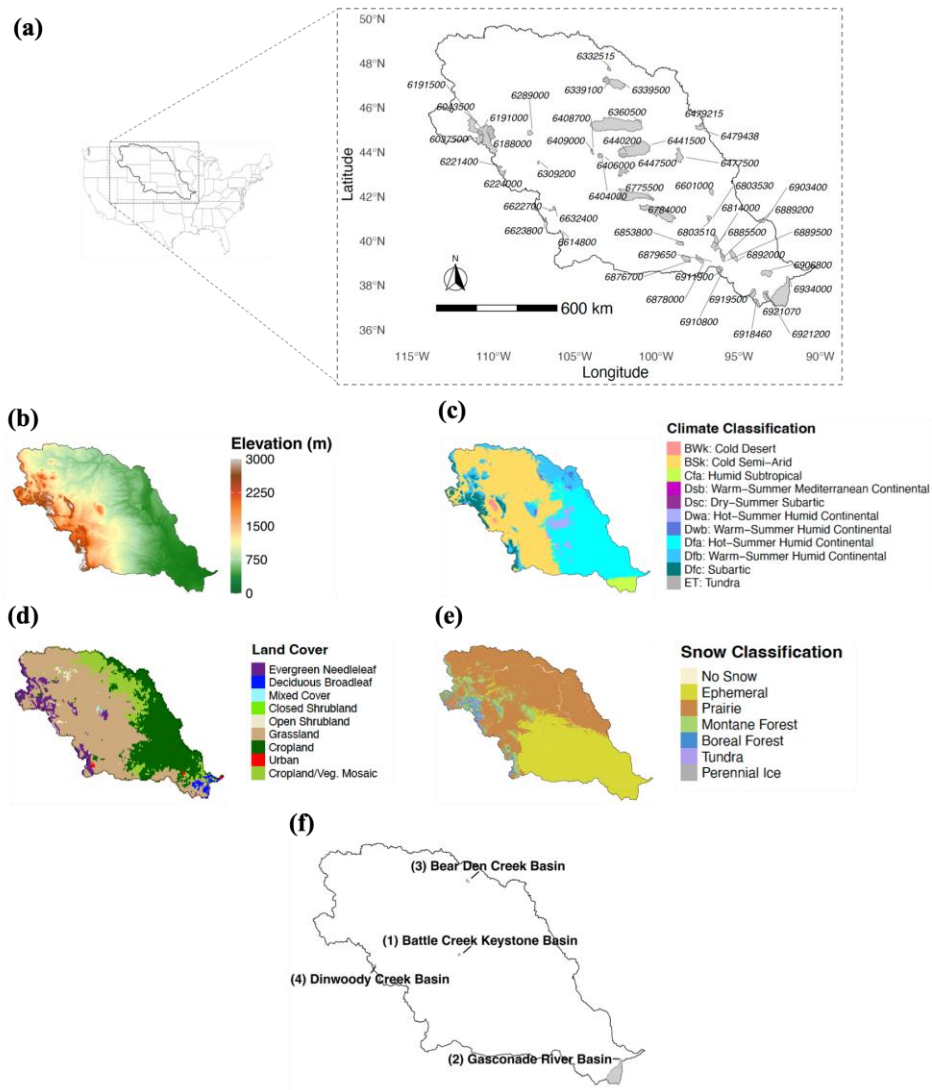

Figure 1 Overview of the Missouri River Basin showing (a) the 50 selected USGS Hydro-Climatic Data Networks (HCDNs) with at least 25 years of record and minimal dam influence, where numbers indicate site identifiers. Maps illustrating (b) the DEM, (c) climate classification, (d) land cover, and (e) snow classification. (f) Location of representative subbasins for time series analysis. Grey shading represents basin boundaries.

For runoff and SWE evaluation, we initially identified 75 subbasins with the USGS streamflow gauges in HCDN, which are characterized by minimal anthropogenic influence (USGS, 2012). Among these, 64 subbasins had at least

25 years of continuous observation, ensuring gauge reliability and operational consistency. Additionally, we excluded 14 subbasins containing dams with controlled outflows, as identified in the National Inventory of Dams (USACE, 2021), to avoid potential impacts of dam operations on natural streamflow patterns. In total, 50 subbasins were selected for analysis, and the upstream drainage area for each gauge was delineated using DEM of the Missouri River Basin (Fig. 1).

3 Data

In this study, land cover, climate and snow classification were used to characterize the basins. The land cover type for each subbasin was determined using the Modified IGBP MODIS vegetation dataset, which is based on the Terra-MODIS sensor and modified by NCEP (https://ral.ucar.edu/model/noah-multiparameterization-land-surface-model-noah-mp-lsm). Climate classification was assigned using the Köppen-Geiger climate classification map (Beck et al., 2018; https://www.gloh2o.org/koppen/) at a 1-km spatial resolution. The predominant snow class within each subbasin was identified using the Sturm snow classification map (Sturm and Liston, 2021; https://nsidc.org/data/nsidc-0768/versions/1), also available at 1-km resolution, which characterizes snow accumulation and ablation patterns based on long-term records of air temperature, precipitation, wind speed, and land-cover.

For model evaluation, we used streamflow observations from USGS HCDN gauges to evaluate simulated runoff and the University of Arizona (UA) gridded SWE dataset (Broxton et al., 2019) to evaluate SWE simulation. The observation-based UA SWE combines station-based SWE and snow depth observations from the Snowpack Telemetry (SNOTEL) network and National Weather Service Cooperative Observer Program (COOP) network with modeled SWE based on an empirical temperature index snow model. Daily gridded UA SWE values at 4 km from January 2014 to September 2023 were utilized in this study.

4 Model and Methods

4.1 Noah-MP

Noah-MP is a widely used, community-driven model designed to simulate the coupled energy, water, and carbon cycles at the land surface (Niu et al., 2011; Yang et al., 2011). Developed as an enhancement to the original Noah LSM, Noah-MP incorporates multiple parameterization options for vegetation processes, canopy radiative transfer, multi-layer snow physics, soil thermal and moisture dynamics, and groundwater interactions (Niu et al., 2011; Yang et al., 2011; He et al., 2023). The availability of multiple parameterization options within Noah-MP is particularly valuable for regional applications, because it allows researchers and practitioners to optimize model performance by selecting the most appropriate physical representations for specific geographical and climatological conditions (Hong et al., 2014; Ma et al., 2017; He et al., 2023).

In this study, Noah-MP 4.0.1 was forced with meteorological data from NLDAS-2. NLDAS-2 provides high-quality, observation-based atmospheric forcing at hourly temporal resolution and 0.125° (~12 km) spatial resolution, spanning

from 1979 to the present (Xia et al., ~~2012~~2012a). The forcing dataset includes hourly integrated precipitation, 2-meter air temperature and specific humidity, incoming shortwave and longwave radiation, 10-meter zonal and meridional wind speed, and surface pressure, making it well suited for driving land surface and hydrological models (Xia et al., ~~2012~~2012a; Xia et al., 2014; Nearing et al., 2016). NLDAS-2 has been extensively validated and is widely used in hydrological and climate studies across North America, supporting applications such as streamflow simulation, drought monitoring, and water resource management (Xia et al., ~~2012~~2012a; Hao et al., 2016; Xia et al., 2016; Pascolini-Campbell and Reager, 2024).

Initial conditions for all experiments were established through an experiment-specific spin-up procedure. Each parameterization experiment (default and all four alternatives) was independently initialized from January 1, 1980, and run through December 31, 2013, using NLDAS-2 forcing data with its respective parameterization scheme settings. The model states at the end of 2013 from individual spin-ups were used as initial conditions for the subsequent 2014–2023 analysis, ensuring that the model state entering the analysis period is fully consistent with each experiment's parameterization configuration.

4.2 Parameterization Schemes for Snow-dominated Basin

To assess Noah-MP's sensitivity to parameterization schemes, this study evaluates four alternative schemes governing runoff and groundwater generation, surface energy exchange, frozen soil permeability, and subsurface thermal dynamics (Table 1). These schemes were selected based on their perceived influence on snowmelt partitioning, groundwater interactions, and flood generation in prior sensitivity studies (Niu et al., 2011; You et al., 2020; Li et al., 2022; Yang et al., 2023). In the following, we briefly discuss these four schemes and their impact on snow and runoff simulations.

Table 1 Default and alternative parameterization schemes for the four Noah-MP experiments.

Experiment	Physical process	Default scheme	Alternative scheme
RUN	Runoff and Groundwater	TOPMODEL with groundwater storage (Niu et al., 2007)	TOPMODEL with an equilibrium water table (Niu et al., 2005)
SFC	Surface Exchange Coefficient for Heat	Monin–Obukhov (Brutsaert, 1982)	Chen97 (Chen et al., 1997)
INF	Frozen Soil Permeability	NY06 (Niu and Yang, 2006)	Koren99 (Koren et al., 1999)
STC	Snow/Soil Temperature Time Scheme	Semi-implicit (Niu et al., 2011)	Full implicit (Ek et al., 2003)

4.2.1 Runoff and Groundwater Process

For the runoff and groundwater simulations, we evaluated two options (1) TOPMODEL with groundwater storage and (2) TOPMODEL with the equilibrium water table. Both schemes compute surface and subsurface runoff as functions of the water table depth, but they differ in how the water table depth is determined.

In TOPMODEL with groundwater storage scheme (Niu et al., 2007), the water table depth is estimated using a linear reservoir located beneath the unsaturated soils. Groundwater storage in this unconfined aquifer is dynamically updated at each time step, reflecting the net gain or loss between groundwater recharge and discharge:

$$\frac{\partial \mathrm{GWS}}{\partial t} = Q_{recharge} - Q_{discharge} \quad (1)$$

where GWS represents groundwater storage changes; $Q_{recharge}$ is groundwater recharge, calculated using Darcy's law (Darcy, 1856); $Q_{discharge}$, groundwater discharge, represents subsurface runoff estimated based on the soil hydraulic characteristics and topographic properties. Changes in the groundwater table depth are then calculated by dividing groundwater storage changes with the specific yield, which is set at 0.2 globally in Noah-MP (Niu et al., 2007).

TOPMODEL with the equilibrium water table scheme estimates water table depths using a conceptual soil column, following the approach of Niu et al. (2005). The initial water table depth is set to be three times the default Noah-MP soil column depth (the original coarse soil column). The model also assumes a hypothetical, finer, and deeper soil column in addition to the original one. Total water deficit is calculated for the original soil column using maximum soil moisture and liquid water content, and for the hypothetical deeper soil column using the Clapp-Hornberger relation (Clapp and Hornberger, 1978), which incorporates saturated soil matrix potential and a soil-type exponent. The equilibrium water table depth is determined by iteratively adjusting the water table depth until the difference in total water deficit between the two approaches is less than 0.01 m.

While both schemes link runoff generation to water table depth, they differ fundamentally in their representation of groundwater dynamics and soil moisture coupling. In the TOPMODEL with groundwater storage scheme (Niu et al., 2007), groundwater recharge is calculated as a single vertical flux at the bottom of the unsaturated zone, governed by the hydraulic conductivity and matric potential of the lowest soil layer. As a result, groundwater storage dynamics are directly dependent on the moisture state of the deepest soil layer, making the scheme sensitive to conditions in the lower soil layers. In contrast, the TOPMODEL with the equilibrium water table scheme (Niu et al., 2005) determines the water table depth iteratively from the total water deficit of the entire soil profile, allowing the water table to reflect the integrated moisture conditions across all soil layers.

4.2.2 Surface Exchange Coefficient for Heat

For the process of the surface exchange coefficient for heat, we considered (1) the Monin-Obukhov (M-O) scheme (Brutsaert, 1982) and (2) the Chen97 scheme (Chen et al., 1997). The M-O scheme is based on Monin-Obukhov similarity theory, while Chen97 scheme (used in Noah version 3.0) differs in how surface roughness and displacement height are treated.

The M-O scheme uses the following equation for the surface exchange coefficient for heat (C_H):

$$C_H = \frac{\kappa^2}{\left[\ln\left(\frac{z-d_0}{z_{0m}}\right)-\psi_m\left(\frac{z-d_0}{L}\right)\right]\left[\ln\left(\frac{z-d_0}{z_{0h}}\right)-\psi_h\left(\frac{z-d_0}{L}\right)\right]} \quad (2)$$

where κ is the von Kármán constant; z is the reference height; d_0 is the zero-displacement height; z_{0m} and z_{0h} are the roughness lengths for momentum and heat (assumed to be the same in M-O scheme); L is the Monin-Obukhov length, and ψ_m and ψ_h are stability correction functions for momentum and heat, respectively.
The Chen97 scheme uses a slightly different approach to compute C_H:

$$C_H = \frac{\kappa^2}{\left[\ln\left(\frac{z}{z_{0m}}\right)-\psi_m\left(\frac{z}{L}\right)+\psi_m\left(\frac{z_{0m}}{L}\right)\right]\left[\ln\left(\frac{z}{z_{0h}}\right)-\psi_h\left(\frac{z}{L}\right)+\psi_h\left(\frac{z_{0h}}{L}\right)\right]} \quad (3)$$

This scheme explicitly distinguishes between the roughness lengths for momentum (z_{0m}) and heat (z_{0h}) where the latter is determined based on the Reynolds number using an empirical equation.

The key difference is that the M-O scheme can drive C_H toward near-zero under very stable (e.g., nighttime, weak-wind) conditions, whereas the Chen97 scheme relaxes this suppression and retains a small but finite exchange. As a result, Chen97 scheme often yields a higher C_H than the M-O scheme (Niu et al., 2011). A higher C_H enhances turbulent ventilation and sensible heat exchange, which cools the land surface and reduces the vapor pressure gradient, thereby suppressing ET, particularly under dry or energy-limited conditions.

4.2.3 Frozen Soil Permeability

For simulating soil water diffusivity under frozen conditions, we examined Niu and Yang (2006, NY06) scheme and Koren et al. (1999, Koren99) scheme. The approaches differ in their physical assumptions and mathematical equation for soil water diffusivity (λ).

In the NY06 scheme, soil water diffusivity, λ_{NY}, is calculated by

$$\lambda_{NY} = (1 - f_{imp,soil}) \times \lambda_{sat} \times \left[\max\left(0.01, \frac{\theta_{soil}}{\theta_{soil,max}}\right)\right]^{B_{exp}+2} \quad (4)$$

where λ_{sat} is the saturated soil water diffusivity, $f_{imp,soil}$ is the impermeable fraction of the soil layer, B_{exp} is a soil texture-dependent exponent, θ_{soil} is the actual soil moisture, and $\theta_{soil,max}$ is the maximum soil moisture. This equation produces a smooth reduction in diffusivity as soil freezing progresses, allowing water movement to remain possible even in partially frozen soils.

In the Koren99 scheme, soil water diffusivity, λ_{Kor}, is calculated as

$$\lambda_{Kor} = \lambda_{sat} \times \left[\max\left(0.01, \frac{\theta_{soil}}{\theta_{soil,max}}\right)\right]^{B_{exp}+2} \tag{5}$$

for unfrozen soils and for frozen soils, it is computed as

$$\lambda_{Kor} = \left[\frac{1}{1+(500\times W_{ice,soil})^3}\right] \times \lambda_{sat}\left[\max\left(0.01, \frac{\theta_{soil}}{\theta_{soil,max}}\right)\right]^{B_{exp}+2} + \left[1 - \frac{1}{1+(500\times W_{ice,soil})^3}\right] \times \lambda_{sat}\left[\min\left(\frac{0.05}{\theta_{soil,max}}, \max\left(0.01, \frac{\theta_{soil}}{\theta_{soil,max}}\right)\right)\right]^{B_{exp}+2} \tag{6}$$

where $W_{ice,soil}$ is the volumetric ice content. This nonlinear formulation causes a much sharper reduction in diffusivity as ice content increases, reflecting reduced water movement in frozen soils (Niu et al., 2011).

4.2.4 Snow/Soil Temperature Time Scheme

The semi-implicit (Niu et al., 2011) and full implicit (Ek et al., 2003) schemes are two approaches for simulating the evolution of snow and soil temperature in Noah-MP. Each scheme enforces that ground temperature beneath snow cannot exceed the freezing point, but they differ in how this constraint and dynamic processes are handled. In the semi-implicit scheme, the model checks during each timestep whether the snow depth exceeds 0.05 meters and if the ground temperature (T_g) is above the freezing point (T_{frz}). If both conditions are met, the ground temperature is immediately reset to the freezing point.

This adjustment is implicit because it enforces the phase change constraint within the timestep, ensuring that the ground cannot physically be warmer than the freezing point while snow is present. The process is dynamic because it is performed at every timestep, allowing the scheme to respond to evolving snow and temperature conditions as the simulation progresses. After the temperature is reset, net longwave radiation, sensible heat, latent and ground heat fluxes are recalculated using the updated ground temperature, ensuring that the energy balance reflects new physical states. The semi-implicit method thus maintains the phase change constraint and energy conservation dynamically, without requiring fully coupling all energy and mass fluxes at once.

The full implicit scheme does not recalculate fluxes but applies the freezing constraint separately to vegetated and bare soil fractions. If the snow depth is higher than 0.05 m and the ground temperature is above freezing point, both

the vegetated ground temperature ($T_{g,veg}$) and bare ground temperature ($T_{g,bare}$) are set to the freezing point. Here, bare ground temperature is used to calculate the surface temperature, which represents the skin temperature of the land surface and influences snowmelt and SWE accumulation through surface energy balance processes.

The semi-implicit and full implicit schemes both enforce the constraint that ground temperature beneath snow cannot exceed the freezing point, but they differ in how surface temperature is determined. In the semi-implicit scheme, surface temperature is consistently derived from flux–temperature relationships. In contrast, the full implicit scheme computes the grid-mean surface temperature as a weighted average of bare ground temperature and canopy temperature for the vegetated fraction. This structural difference may cause the semi-implicit scheme to produce more abrupt SWE changes due to the discrete flux recalculation, whereas the full implicit scheme can capture more spatially variable SWE evolution by explicitly representing heterogeneous surface fractions.

4.3 Performance Evaluation with Statistical Metrics

In this study, Kling-Gupta Efficiency (KGE) is employed as the primary performance metric to evaluate Noah-MP hydrological simulation. The KGE metric was developed to address limitations of traditional performance measures such as Nash-Sutcliffe Efficiency and has been widely adopted for calibration and evaluation of hydrological models (Gupta et al., 2009; Liu et al., 2022; Cinkus et al., 2023). KGE is calculated as:

$$KGE = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2} \quad (7)$$

where r represents the linear correlation coefficient between observed and simulated values, β is the ratio of simulated to observed means, and γ is the ratio of simulated to observed standard deviations. KGE ranges from negative infinity to 1, with values close to 1 indicating better model performance. Knoben et al. (2019) showed that KGE = -0.414 represents the performance of prediction using the observation mean.

5 Results

5.1 Effects of Parameterization Schemes on Water Balance components in the Missouri Basin

To evaluate the influence of parameterization schemes, a water balance analysis was conducted over the entire Missouri River Basin domain. All 0.125° resolution grid cells within the basin were considered. The analysis was performed using daily values from 2014 to 2023, with water years defined as October 1 through September 30 of the following year. For each water year, the following components were computed: total precipitation (P), total evapotranspiration (ET), annual maximum SWE (SWE), change in groundwater storage (ΔGW; groundwater storage at the final day minus the first day of the year), change in soil moisture storage (ΔSM; soil moisture at final day minus initial day), surface runoff (Q_s), subsurface runoff (Q_{sb}), and total runoff (Q_{total}; surface runoff + subsurface runoff). These variables were aggregated spatially across the basin and temporally averaged (sum for fluxes) over a 10-year period to represent long-term hydrologic behavior under each parameterization scheme.

Table 2 summarizes the 10-year mean values of each water balance component across the five parameterization schemes evaluated. Mean annual precipitation for the five experiments was 534.0 mm, of which 464–480 mm (87–90%) was lost through evapotranspiration and 64–80 mm (12–15%) was routed as total runoff. Variability among experiments was mainly expressed in the ET and runoff, with ET differences of up to 16 mm balanced by corresponding changes in runoff. For all experiments, the water-budget residual—defined as Residual = $P - ET - \Delta GW - \Delta SM - Q_{total}$ —was below 1%, indicating tight water-balance closure in the simulations.

Table 2 Basin-averaged, 10-year mean water balance components across five parameterization schemes for the entire Missouri River Basin. Components include precipitation (*P*), evapotranspiration (*ET*), snow water equivalent (*SWE*), change in groundwater storage (*ΔGW),* change in soil moisture (*ΔSM),* surface runoff (Q_s), subsurface runoff (Q_{sb}), and total runoff (Q_{total}).

Experiment	*P* (mm)	*ET* (mm)	*SWE (mm)*	*ΔGW (mm)*	*ΔSM* (mm)	Q_s (mm)	Q_{sb} (mm)	Q_{total} (mm)
Default	534.0	479.5	33.4	-0.3	-6.1	20.6	43.3	63.9
RUN	534.0	473.7	33.4	2.8	-5.6	26.8	40.7	67.5
SFC	534.0	463.6	36.9	0.2	-6.3	24.9	55.4	80.2
INF	534.0	476.9	33.4	-0.1	-6.0	27.9	38.8	66.7
STC	534.0	474.0	48.9	-0.3	-5.8	21.7	47.7	69.4

Table 2 demonstrates how key hydrological components differed by Noah-MP model parameterization scheme. The RUN experiment, which assumes an equilibrium water table, produced the largest increase in groundwater storage (ΔGW = +2.8 mm). The SFC experiment resulted in the lowest evapotranspiration (*ET* = 463.6 mm), because its alternative scheme limits latent heat flux. This reduction in ET contributed to the highest total runoff (Q_{total} = 80.2 mm), largely through increased subsurface runoff. In contrast, the INF experiment yielded the highest surface runoff (Q_s = 27.9 mm), likely reflecting reduced infiltration capacity under frozen soil conditions. The increased surface runoff was partially offset by lower subsurface flow, maintaining a total runoff value comparable to other schemes. The STC experiment produced higher surface, subsurface and total runoff than the default configuration, reflecting altered snow dynamics under its alternative snow/soil temperature time scheme. Notably, the STC experiment produced the highest average annual maximum SWE (48.9 mm) among all schemes, substantially exceeding the default value (33.4 mm). This difference reflects the altered snow accumulation and melt timing under the implicit snow/soil temperature time scheme, which delays snowmelt and prolongs snowpack retention. The enhanced snow storage under the STC scheme is consistent with the higher surface, subsurface, and total runoff relative to the default.

5.2 Evaluation of Runoff Simulation across Parameterization Schemes

Runoff simulation performance was evaluated by comparing monthly Noah-MP total runoff with observed streamflow at the 50 USGS HCDNs (Fig. 2). Overall, runoff simulation performed well in high-elevation area in west of the Rocky Mountains and near the basin outlet in the southeastern portion of the Missouri River Basin. However, sensitivity to parameterization varies by location. For example, the SFC experiment provided relatively higher KGE in lowland subbasins compared to the default experiment (Fig. 2c). The RUN experiment also shows improved

performance in lowland areas as well as in high-elevation subbasins in the western part of the domain (Fig. 2b). The INF experiment produced better simulation results in the northern part of the Missouri River Basin, where grassland dominates (Fig. 2d).

To highlight these differences, we generated ΔKGE maps that display the improvement in KGE for each experiment relative to the default (Fig. 3). The location of improvements differed by experiment. The RUN experiment demonstrated widespread performance improvements across most of the study domain (Fig. 3a). However, the effectiveness of this experiment is diminished in the subbasins located near the outlet of the Missouri River Basin, which is located at the bottom right of the domain. The SFC experiment performance improvements were concentrated in the lower subbasin regions (Fig. 3b). The INF experiment did not exhibit a clear spatial pattern of improvement, with relatively better performance observed in the northern part of Missouri River Basin (Fig. 3c). Finally, the STC experiment showed moderate improvements in subbasins near the outlet of the domain, while substantial improvements were limited (Fig. 3d). In terms of the fraction of subbasins showing improvement, the RUN experiment improved 58% of subbasins (29 of 50), the SFC experiment improved 64% (32 of 50), the INF experiment improved 64% (32 of 50), and the STC experiment improved 50% (25 of 50). "No Improvement" refers to subbasins where ΔKGE is zero or negative, encompassing both no change and performance degradation relative to the default. The absolute KGE values for all five schemes across each HCDN subbasin are provided in Fig. S1, allowing direct comparison of scheme performance at the individual subbasin level.

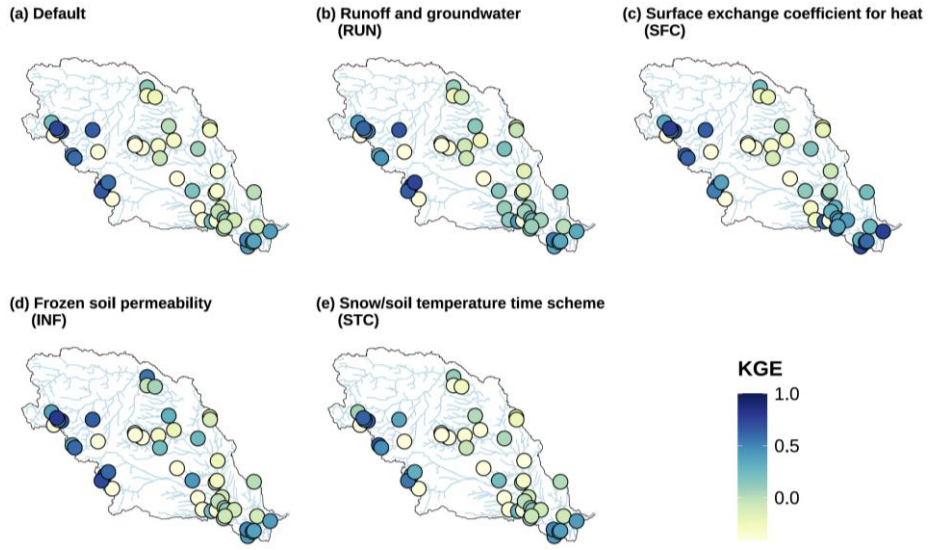

Figure 2 KGE of Noah-MP runoff relative to USGS streamflow at 50 HCDNs. Results are presented for (a) the default setting and with different parameterization scheme of (b) runoff and groundwater (RUN), (c) surface exchange coefficient for heat (SFC), (d) frozen soil permeability (INF) and (e) snow/soil temperature time scheme (STC).

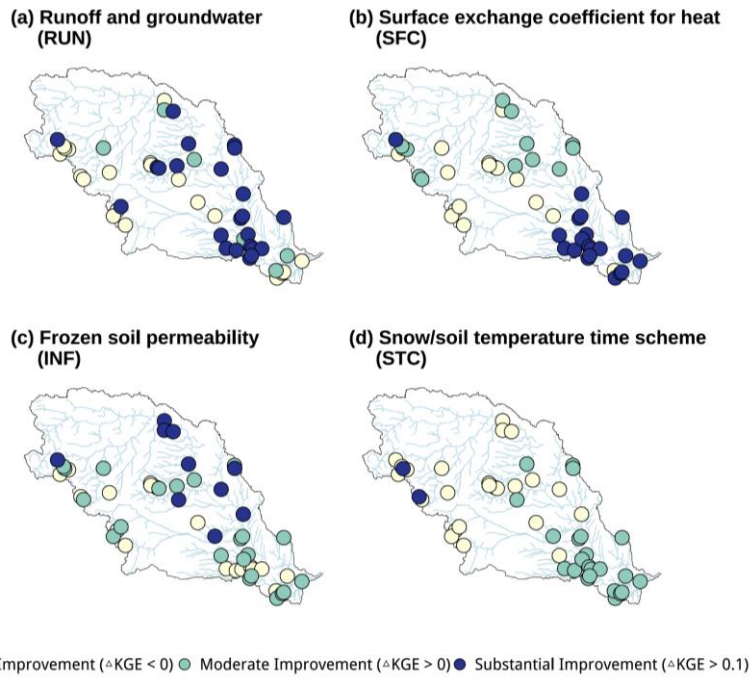

Figure 3 Changes in KGE at 50 HCDNs by the alternative parameterization scheme of (a) runoff and groundwater (RUN), (b) surface exchange coefficient for heat (SFC), (c) frozen soil permeability (INF) and (d) snow/soil temperature time scheme (STC). Subbasins are classified into three categories with different colors indicating No Improvement, Moderate Improvement, and Substantial Improvement.

To assess how alternative parameterization schemes affect hydrological processes, a comprehensive evaluation of hydrometeorological variables was conducted across multiple representative subbasins. Fig. 4 shows precipitation, evapotranspiration, groundwater storage, soil moisture, SWE and runoff time series for four selected subbasins. These subbasins were selected because they showed notable improvements in KGE values and exhibit process-specific response for each experiment. Their geographic locations are shown in Fig. 1f. The selected subbasins represent diverse hydroclimatic settings: Battle Creek Keystone Basin (warm-summer humid continental, evergreen needleleaf forest, montane forest snow), Gasconade River Basin (humid subtropical, grassland, ephemeral snow), Bear Den Creek Basin (warm-summer humid continental, vegetation mosaic, prairie snow), and Dinwoody Creek Basin (subarctic, shrubland, tundra snow). For comparison with USGS streamflow observations, simulated runoff at each 0.125° grid cell was aggregated by computing the area average of all grid cells within the basin boundary. Soil moisture reported throughout this study (e.g., Fig. 4) represents the full 2.0 m soil column, integrated across all four Noah-MP soil layers.

For the subbasins shown in Fig. 4, the time series comparisons demonstrate how the alternative schemes address limitations of the default setting. ~~The corresponding KGE values indicate improvements based on runoff comparisons with observed streamflow.~~ In Battle Creek Keystone (Fig. 4a), the default shows limited groundwater variability. The RUN experiment produces clear seasonal and interannual groundwater variability which in turn improved simulation of peak runoff in 2014, 2018, and 2023 and increases KGE by 0.19. The 2019 peak was still lower than observed but is more accurately represented than in the default scheme. The RUN experiment also shows that the improved groundwater storage dynamics influenced soil moisture variability, with more pronounced wet and dry cycles compared to the default. In the Gasconade River (Fig. 4b), the SFC experiment reduces evapotranspiration relative to the default and raises KGE by 0.38. It corrects the systematic underestimation of peak runoff from 2014–2023, including the largest flood event in 2017. Comparison with GLEAM evapotranspiration confirms that the default scheme systematically overestimates ET in this subbasin, while the SFC experiment produces more realistic ET estimates. The reduced evapotranspiration in the SFC experiment led to increases in soil moisture, contributing to wet conditions prior to flood events. In Bear Den Creek (Fig. 4c), the default fails to capture the rapid increases in surface runoff, while the INF experiment resolves this limitation. It improves KGE by 0.44 with better simulation of the 2019 and 2023 peaks. In Dinwoody Creek (Fig. 4d), the STC experiment enhances SWE simulation and produces a KGE improvement of 0.14. The alternative SWE aligns more closely with UA SWE and results in more realistic runoff, particularly during the 2015, 2019, and 2023 floods. The delayed snowmelt under the STC experiment produces a later but more sustained infiltration signal, resulting in distinct seasonal soil moisture patterns relative to the default. Collectively, these results show how each experiment resolves distinct weaknesses of the default and improves runoff representation across diverse hydroclimatic conditions.

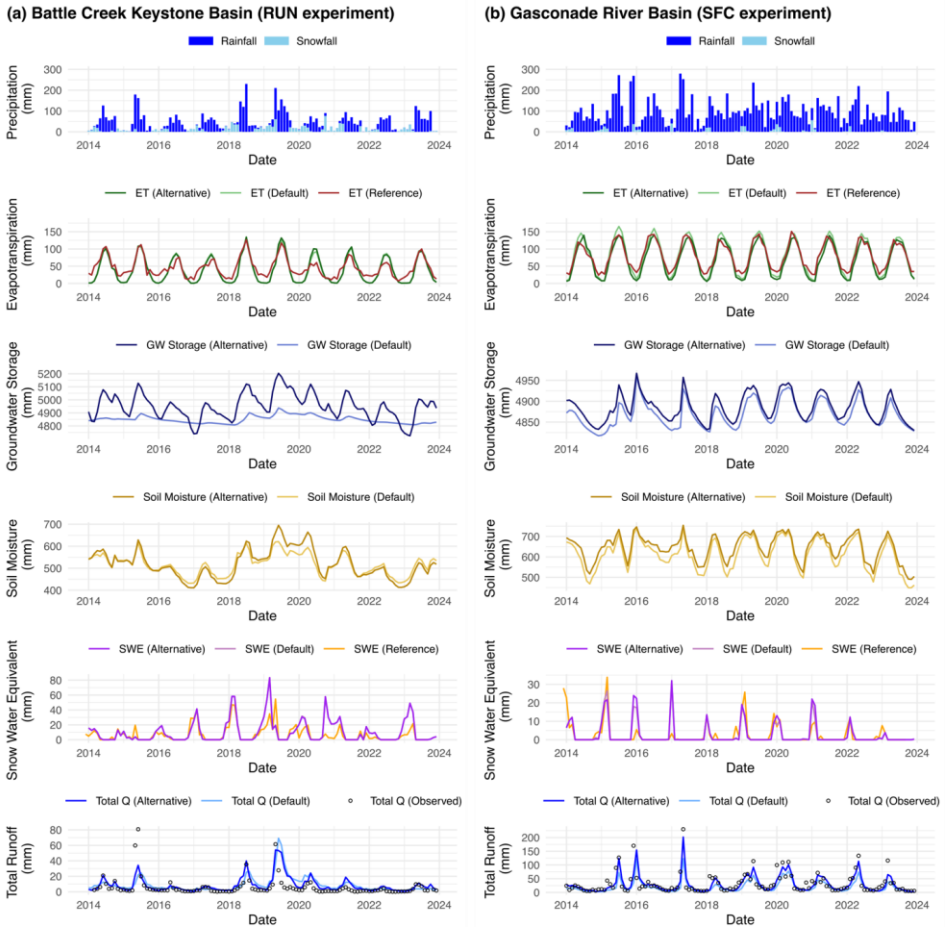

Figure 4 Time series of precipitation, evapotranspiration, groundwater storage, soil moisture, SWE and total runoff of subbasin with notable improvement of KGE by the alternative scheme of (a) runoff and groundwater (RUN), (b) surface exchange coefficient for heat (SFC), (c) frozen soil permeability (INF) and (d) snow/soil temperature time scheme (STC). USGS streamflow is used as observation, with UA SWE and GLEAM evapotranspiration as reference datasets. Default results are indicated by solid lines in lighter colors, and alternative results by solid lines in darker colors.

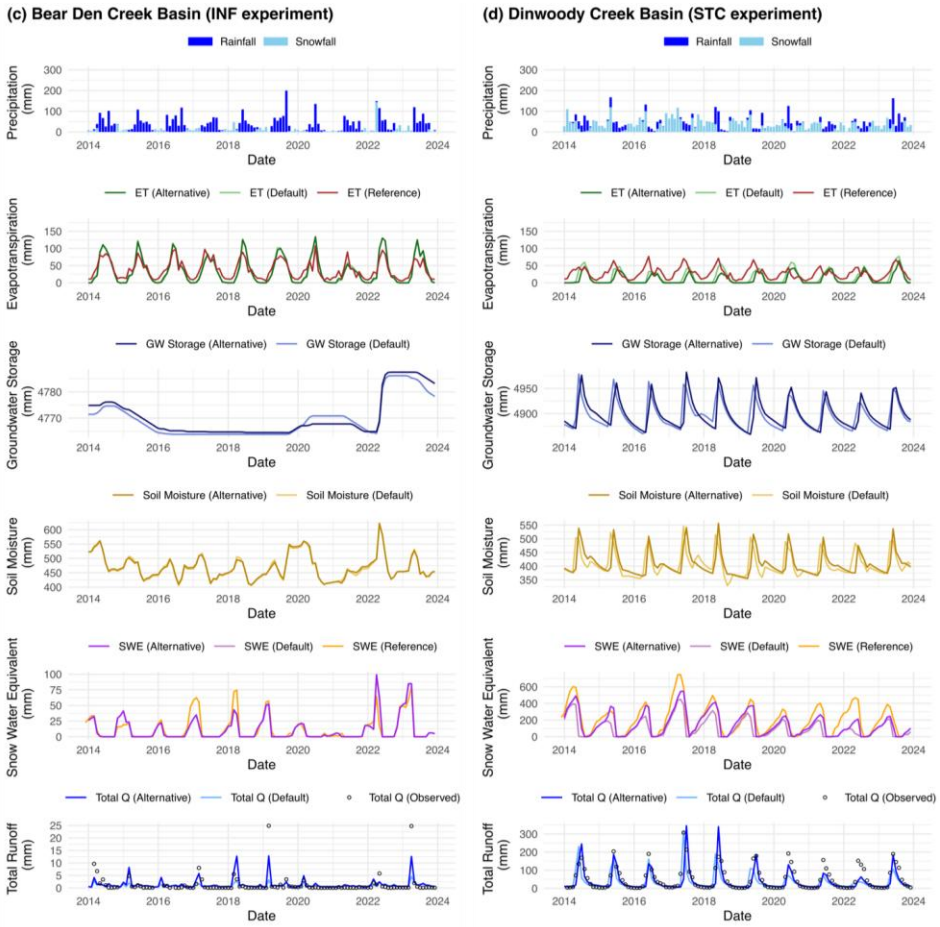

Figure 4 (continued)

5.3 Performance Improvement over Land Cover, Climate, and Snow Characteristics

Fig. 5 presents the KGE differences between the default and alternative parameterization schemes, illustrating how model performance varies with basin characteristics such as climate, land cover, and snow type. The RUN and SFC experiments showed broad effectiveness across multiple climate and land cover classes, with particularly notable improvements in Hot-Summer Humid Continental climates (0.18 and 0.28, respectively) and Cropland areas (0.22 and 0.24, respectively). The INF experiment demonstrated targeted improvements in Cold Semi-Arid climates (0.14), Vegetation Mosaic land cover (0.34), and Prairie Snow class (0.23), while the STC experiment provided only modest

improvements across most classifications. In contrast, the default scheme performed best in Subarctic climates, Evergreen Needleleaf Forest, and high-elevation snow classes including Montane Forest and Boreal Forest.

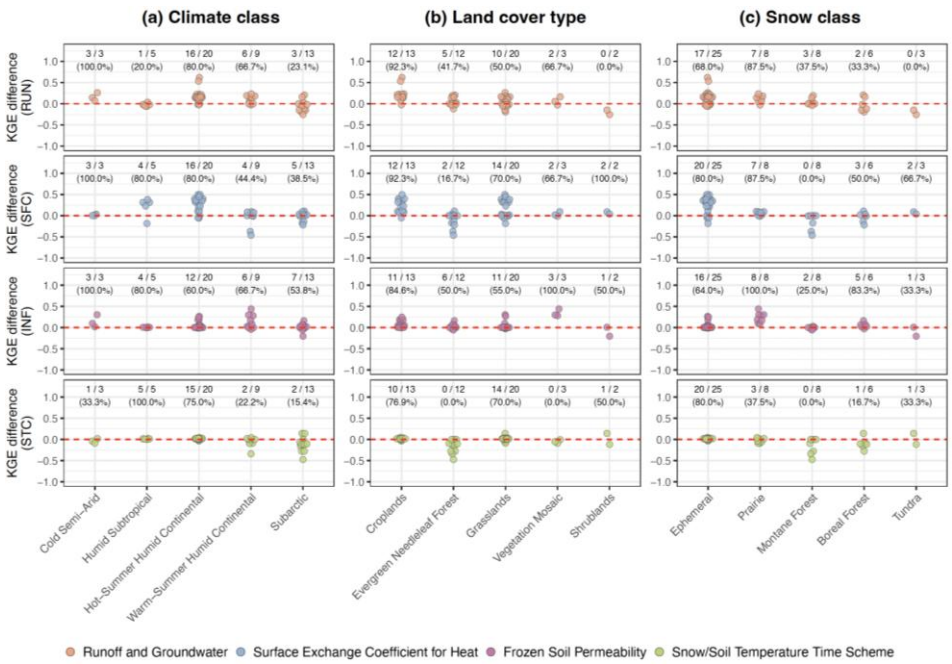

Figure 5 Dot plots showing changes in KGE across basin characteristics. Columns represent (a) climate classification, (b) land cover, and (c) snow classification. Rows represent the RUN experiment (top), SFC experiment (second), INF experiment (third), and STC experiment (bottom). Different experiments are distinguished by symbol colors.

5.4 Improvement of Snow Water Equivalent and Snowmelt-Driven Runoff Simulation

While previous results focused on overall runoff evaluation, this section examines the snow component of the Noah-MP model and assesses how alternative schemes impact SWE and snowmelt-driven runoff simulation. The analysis focused on the maximum SWE for each water year and total runoff from April to July (Q), as these metrics are important to snowmelt-driven runoff generation (Lapides et al., 2022; Gottlieb and Mankin, 2024). Relative bias was calculated as the difference between the average of simulated values and the average of observed values (simulated minus observed), divided by the average of observed values.

Fig. 6 shows how the relative bias of SWE varies among subbasins under different parameterization schemes. To examine whether these results relate to SWE variability, we calculated the mean annual maximum SWE for each HUC8 subbasins in Missouri River Basin over the 2014–2023 period (Fig. S2). Overall, the default, RUN, SFC, and

INF experiments produced similar spatial patterns, with systematic underestimation of SWE across most high-snowfall regions in the Rocky Mountains (Fig. S2). In contrast, the STC experiment effectively mitigated this underestimation issue. The number of subbasins with relative bias within ±10 percent (indicated by bold boundaries in Fig. 6) was notably higher for the STC experiment, with 14 subbasins, compared with the default, RUN, SFC, and INF experiments, which had four, four, eight, and six subbasins, respectively. This indicates that while most schemes perform comparably across the basin domain, the STC experiment provides an advantage in snow-dominated subbasins where annual SWE is high and snowmelt-driven runoff is the dominant source of streamflow.

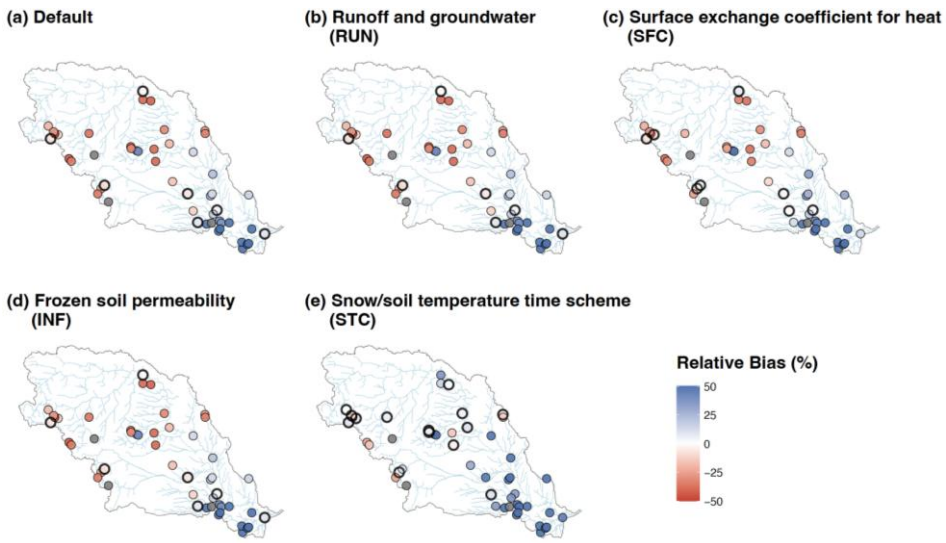

Figure 6 Relative bias of SWE in different HCDNs for (a) default setting and alternative scheme of (b) runoff and groundwater (RUN), (c) surface exchange coefficient for heat (SFC), (d) frozen soil permeability (INF) and (e) snow/soil temperature time scheme (STC). Subbasins with mean bias less than 10% are highlighted using bold boundaries.

Theoretically, improved SWE representation should lead to better runoff simulation. However, we found that SWE improvements do not universally translate into improved Apr–Jul runoff across all subbasins. To investigate this, we identified subbasins where both SWE and Apr–Jul runoff improved (blue circles) and those where SWE improved but runoff did not (red circles) under STC experiment. As shown in Fig. 7, both groups are spatially distributed across the Missouri River Basin without a clear regional pattern, suggesting that the translation of SWE improvements into runoff changes is mediated by multiple hydrological factors beyond SWE alone. This spatial distribution motivates a closer examination of the direction of runoff bias under the default scheme.

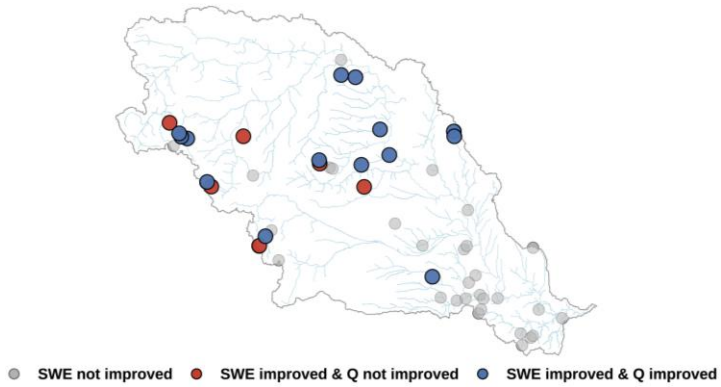

Figure 7 Spatial distribution of STC experiment outcomes across the 50 HCDN subbasins in the Missouri River Basin. Blue circles indicate subbasins where both peak SWE and Apr–Jul runoff were improved relative to the default scheme, red circles indicate subbasins where SWE was improved but runoff was not, and grey circles indicate subbasins where SWE was not improved. Major river flowlines are shown in light blue.

Fig. 8 shows the Apr–Jul runoff relative bias for the Default and STC experiment across SWE-improved subbasins, grouped by runoff improvement status. In subbasins where both SWE and runoff improved (Fig. 8a), the default scheme predominantly underestimated runoff, and the STC experiment effectively reduced this negative bias. In contrast, the subbasins where SWE improved but runoff did not, the default scheme already overestimated runoff in most cases, and the STC scheme further exacerbated this positive bias. These results suggest that the effectiveness of the STC scheme in improving runoff simulation depends critically on the direction of the pre-existing runoff bias in the default configuration, rather than on the degree of SWE improvement alone.

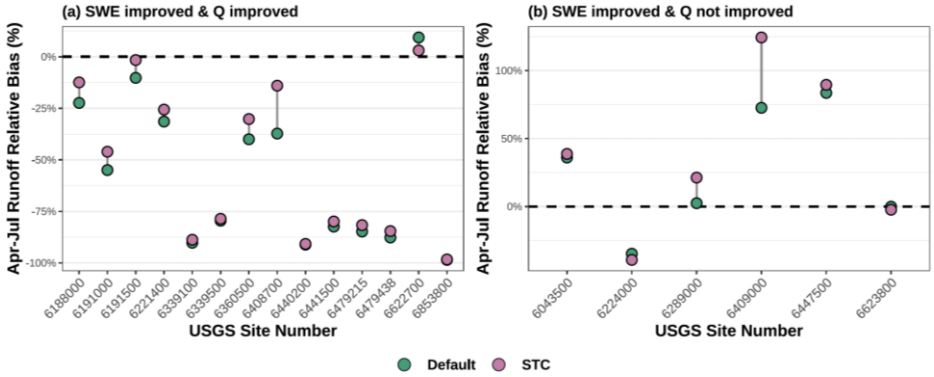

Figure 8 Apr–Jul runoff relative bias (%) for the Default (green) and STC (purple) schemes across SWE-improved subbasins, grouped by runoff improvement status: (a) subbasins where both SWE and Q were improved and (b) subbasins where SWE was

improved but Q was not. Points are connected by grey lines to highlight the direction and magnitude of change between the two schemes. The dashed line indicates zero bias.

6 Discussions

6.1 Influences of Alternative Parameterization Scheme on Runoff Generation

This study evaluated the performance of four alternative parameterization schemes in comparison with the default Noah-MP configuration to assess their impacts on runoff generation and other hydrological processes. The analysis revealed three limitations in the default parameterization scheme that were better addressed by the alternative schemes.

The first ~~limitation identified~~area for improvement was the ~~inadequate~~limited groundwater dynamics in the default groundwater scheme. Groundwater simulated by the default scheme does not always exhibit strong seasonality, with minimal variability throughout the year (Fig. 4a). In the default scheme, groundwater recharge is governed by the hydraulic conductivity and matric potential of the lowest soil layer. This makes groundwater storage dynamics sensitive to the moisture state of the deepest soil layer only. Overestimation of evapotranspiration depletes moisture in the lower layers, suppressing recharge and dampening groundwater storage variability (Li et al., 2021; Cai et al., 2014). Maina et al. (2025) addressed a related issue by integrating the ParFlow model with Noah-MP. In this study, the alternative TOPMODEL with equilibrium water table scheme overcame this limitation by determining the water table depth from the total water deficit of the entire soil profile.

The second ~~limitation~~area for improvement involved systematic overestimation of ET in vegetated areas, particularly those dominated by cropland. Previous research has demonstrated that the default Noah-MP crop parameterization ~~inadequately represents~~does not fully represent dynamic cropland processes, ~~leading~~which can contribute to ~~possible~~ overestimation of ET (Liu et al., 2016). Furthermore, Noah-MP's consistent overestimation of gross primary productivity in vegetated areas exacerbates this issue by elevating ET in both cropland and grassland environments (Ma et al., 2017). The alternative surface exchange coefficient scheme addresses this limitation by using a different function for calculating surface exchange coefficients for heat transfer. Unlike the default scheme, the alternative approach excludes zero-displacement height considerations and dynamically calculates thermal roughness length. This results in ET estimates that better align with observed values by enhancing the efficiency of heat and moisture exchange between the land surface and atmosphere.

A third ~~limitation is the potential underrepresentation of rapid~~area for improvement relates to winter-season surface runoff generation ~~that can occur during rain-on-snow (ROS) events.~~under frozen-soil conditions. The default Noah-MP configuration ~~inadequately captures~~does not fully capture the physical processes associated with ~~ROS events~~frozen and partially frozen soils, including ice lens formation, changes in snowpack thermal conductivity, and alterations in water retention capacity. The default frozen soil permeability scheme tends to predict more permeable frozen soils than typically observed in natural systems, particularly in conditions where ice layers inhibit infiltration (Koren et al., 1999; Niu and Yang, 2006). The alternative scheme overcomes this limitation through a nonlinear

Formatted: English (United States)

weighting function based on volumetric ice content, which reduces soil water diffusivity quickly as ice content increases. This reduced diffusivity effectively limits infiltration capacity under frozen soil conditions, causing a greater fraction of precipitation and snowmelt to be partitioned into surface runoff rather than subsurface flow. ~~Among the subbasins where INF improved overall performance, peak runoff estimates also improved at 28 sites, with an average improvement of 25.1% over the 10-year period, ranging from 1.0% to 71.0%.~~ This is particularly ~~important~~relevant for improving winter-season runoff estimates~~ during ROS events~~, which are often underestimated by the default scheme due to excessive infiltration ~~by the default scheme.~~ under frozen soil conditions. A more detailed, event-based evaluation of rain-on-snow processes specifically is left for future work.

It is important to note, however, that the identified ~~limitations of~~areas for improvement in the default scheme, and the ~~improvements~~gains offered by alternative schemes, reflect the specific hydroclimatic conditions of the Missouri River Basin. These findings may not be directly transferable to regions with substantially different climate regimes, land cover characteristics, or snow dynamics, as the performance of land surface parameterization schemes is known to be region-dependent (Gan et al., 2019; You et al., 2020; Li et al., 2022; Hu et al., 2023; Chang et al., 2020; Hamitouche et al., 2025). A more detailed discussion of these limitations is provided in Section 6.4.

6.2 Recommendations for Parameterization Scheme Selection

Two key criteria were used to recommend alternative parameterization schemes (Table 3). First, at least 50 percent of the subbasins within each environmental classification had to show improvement relative to the default to ensure sufficient representation. Second, the experiment needed to demonstrate an average KGE improvement of at least 0.1 within that classification. The distribution of the 50 HCDN subbasins across these classifications is summarized in Figure S3. It should be noted that classifications represented by relatively few subbasins warrant more cautious interpretation.

Table 3 Recommended Noah-MP parameterization schemes by environmental classification based on two criteria: at least 50% of subbasins within a class showing improvement relative to the default and an average KGE increase of at least 0.1.

Classification	Categories	Recommendation				
		Default	RUN	SFC	INF	STC
Climate	Cold Semi-Arid		✓		✓	
	Humid Subtropical			✓		
	Hot-Summer Humid Continental		✓	✓		
	Warm-Summer Humid Continental		✓		✓	
	Subarctic	✓				
Land cover	Cropland		✓	✓		
	Evergreen Needleleaf Forest	✓				
	Grassland			✓		
	Vegetation Mosaic				✓	
	Shrublands	✓				

	Ephemeral		✓	✓	
	Prairie		✓		✓
Snow	Montane Forest	✓			
	Boreal Forest	✓			
	Tundra	✓			

The recommendations summarized in Table 3 show that the RUN and SFC experiments are broadly effective across climate, land cover, and snow classifications. The INF experiment shows more limited but targeted improvements, mainly in Warm-Summer Humid Continental, Vegetation Mosaic and Prairie Snow classifications. In contrast, the STC experiment does not meet the criteria in any classification, reflecting its limited overall runoff improvement discussed in Section 5.4. The default is retained for Subarctic climates as well as Evergreen Needleleaf Forest, Shrublands, Montane Forest, Boreal Forest, and Tundra environments. Among these, the forest-related classifications (Evergreen Needleleaf, Montane, and Boreal Forest) may benefit from alternative canopy-related parameterizations such as interception or dynamic vegetation, which none of the four experiments targeted.

The effects of the recommended parameterization schemes in this study are consistent with previous research. For example, Hamitouche et al. (2025) showed that the TOPMODEL scheme with an equilibrium water table (RUN experiment) can improve model performance by reducing runoff underestimation and yielding a higher KGE than the default setting on a global scale. Chang et al. (2020) reported that the Chen97 surface exchange coefficient scheme (SFC experiment) enhances latent and sensible heat flux simulations, resulting in more accurate evapotranspiration estimates than the Monin-Obukhov scheme over a subtropical forest in southern China. Improvements related to frozen soil processes were also highlighted by Hu et al. (2023), who showed that Koren99's lower soil permeability (INF experiment) can better represent frozen soil hydrology by capturing stronger soil moisture variability and producing more realistic runoff responses in permafrost regions of the Tibetan Plateau.

6.3 The Role of Snow-Related Processes and Implications for Model Selection

An insight from this study is that improvements in overall hydrological performance do not always coincide with better representation of snow processes in Noah-MP. The STC experiment exemplifies this complexity, demonstrating limited improvements in overall runoff simulation (Fig. 3d) while achieving the most substantial enhancements in SWE representation (Fig. 6e). As demonstrated in Figs. 7 and 8, this disconnect is primarily governed by the direction of the pre-existing runoff bias in the default scheme, whereby SWE improvements translated into better runoff in most of the subbasins where the default configuration underestimated runoff. This finding suggests that the STC scheme's impact on runoff is conditional rather than universal, as its influence on snow accumulation and melt processes does not propagate consistently through the broader suite of hydrological processes that govern total runoff. Similarly, Sthapit et al. (2022) reported that although snowpack dynamics strongly influence runoff and soil moisture, mismatches and temporal lags in the simulated response indicate that improvements in snow representation do not always translate into better streamflow performance.

These findings highlight the importance of considering application needs when selecting a parameterization scheme. In snow-dominated basins where water supply forecasting, flood risk assessment, or ecological management depends on snowpack dynamics, the selection of parameterization schemes should prioritize snow process accuracy even when overall runoff metrics show modest improvements. In basins where snow plays a minor role, however, total runoff metrics may remain the most relevant criterion.

6.4 Limitations and Future Direction

While this study advances understanding of parameterization selection in the Noah-MP model, several limitations should be acknowledged regarding the generalizability of our findings. The current analysis is confined to the Missouri River Basin for the period from 2014 to 2023, which may limit the broader applicability of the recommended parameterization scheme. The literature consistently shows that the performance of land surface parameterization schemes varies considerably across different climatic and geographic settings (Gan et al., 2019; Li et al., 2022). For example, Hu et al. (2023) showed that the Koren99 frozen soil scheme (INF experiment) is particularly effective in permafrost regions of the Tibetan Plateau. Chang et al. (2020) found that the Chen97 surface exchange coefficient scheme (SFC experiment) improved ET simulations over subtropical forests in South China. Hamitouche et al. (2025) further documented substantial global variability in the performance of the TOPMODEL equilibrium water table scheme (RUN experiment). Taken together, these findings underscore that scheme recommendations derived from one region should be applied with caution elsewhere. Further validation across diverse geographical settings is therefore necessary to establish the robustness of the alternative parameterization approaches identified in this study.

This study focused on a subset of runoff and snow related parameters within the Noah-MP model. Future work should extend the analysis to include additional snow-related processes, such as snow thermal conductivity and snow cover fraction parameterizations. Snow thermal conductivity regulates the transfer of energy through the snowpack and thereby influences the timing of snowmelt, while snow cover fraction controls the representation of sub-grid snow heterogeneity and affects surface energy balance (Yang et al., 2023; Cao et al., 2024; Abolafia-Rosenzweig et al., 2025). Incorporating these processes would enable a more comprehensive assessment of Noah-MP sensitivity to snow dynamics.

An additional limitation of this study is that no river routing scheme was applied when aggregating grid-cell runoff to each basin. Instead, runoff was computed as an area average within each basin boundary. This approach follows established practice in continental-scale land surface model evaluation, where gridded runoff has been area-averaged and compared against USGS streamflow over 961 small basins ranging from 23 to 10,000 km² without explicit channel routing (Xia et al., 2012b). Routing effects on simulated streamflow are also known to diminish at monthly and annual timescales relative to sub-daily and daily resolutions, and to be more pronounced for larger contributing areas (Cortés-Salazar et al., 2023). Because this study evaluates runoff at a monthly resolution across relatively small, HUC-10-scale basins, the effect of omitting explicit channel routing is expected to be limited. Future studies could incorporate

explicit routing to better capture the timing of peak flows, particularly for larger or more topographically complex basins.

The Noah-MP model offers multiple physics parameterization schemes across various land surface processes, where interactions among different schemes can substantially affect model performance. Such interactions can be complex and non-linear: changes in ET, soil moisture, or snowmelt driven by one scheme may amplify or dampen the response of another, and compensating errors between schemes can be disrupted when any one of them is changed (Cuntz et al., 2016). Prior multi-physics ensemble studies have reported that certain scheme combinations produce emergent behaviors not predictable from individual scheme evaluations alone (Hong et al., 2014; Gan et al., 2019). Future investigations should implement methodologies that evaluate different combinations of physics schemes, thereby delivering a more comprehensive understanding of model uncertainty and overall performance.

As an alternative strategy to reduce the complexity arising from interactions among multiple physical processes, future studies could adopt a process-oriented evaluation framework that systematically compares multiple parameterization schemes within a single physical process. This approach would help isolate process-specific uncertainties while identifying parameterizations that provide the greatest performance improvements. For example, the current frozen-soil parameterization in Noah-MP does not fully account for preferential flow through soil macropores, which may contribute to biases in infiltration and runoff generation during frozen conditions. Agnihotri et al. (2023) evaluated combinations of supercooled liquid water and soil hydraulic property schemes in Noah-MP over the Mississippi River Basin and demonstrated that representing higher frozen-soil permeability to better account for preferential flow significantly improved spring streamflow simulations. In addition, Abolafia-Rosenzweig et al. (2026) tested 256 combinations of Noah-MP physics options across 790 SNOTEL stations in the western United States and identified an updated configuration that reduced both snow accumulation and ablation rate errors relative to the default. Such process-focused evaluations could be extended to other land surface processes to further improve the physical realism and predictive capability of Noah-MP.

In addition to these parameterization schemes, other improvements in model physics may help enhance model performance. This study revealed several key areas where process-based enhancements could address current model limitations, including the underestimation of evapotranspiration and SWE, along with the misrepresentation of rapid surface runoff generation and groundwater dynamics. Incorporating more sophisticated algorithms for snow processes, improving groundwater-surface water interactions, and refining the representation of rapid surface flow processes could substantially enhance the model's ability to simulate realistic hydrological responses.

7 Conclusions

We evaluated the impact of alternative Noah-MP parameterization schemes on runoff and snow generation across 50 USGS HCDN subbasins in the Missouri River Basin. By comparing the default and alternative schemes of runoff and groundwater processes, surface exchange coefficients, frozen soil permeability, and snow/soil temperature time

schemes, we identified basin-specific parameterization schemes that provide improved hydrological model performance in snow-dominated watersheds (Figs. 2 and 3). The effectiveness of each scheme varies by basin characteristics including climate, land cover, and snow regime, which supports the use of targeted parameterization schemes instead of relying on default physics (Fig. 5 and Table 3). The STC experiment, while offering limited improvement in overall runoff, produced the greatest increases in snow process representation by reducing SWE bias and improving snowmelt driven runoff dynamics (Figs. 6, 7 and 8).

While this study enhances understanding of parameterization selection in the Noah-MP model, several limitations should be acknowledged. The analysis was confined to the Missouri River Basin for the 2014 to 2023 period, which may limit the broader applicability of the recommended parameterization schemes. This study also focused on a subset of runoff- and snow-related parameters within the Noah-MP model. Future work should extend the analysis to include additional snow-related processes, such as snow thermal conductivity and snow cover fraction parameterizations. Furthermore, evaluating interactions among parameterization schemes will provide a more comprehensive understanding of model uncertainty and performance. A more comprehensive basin-scale analysis linking SWE, soil moisture, and runoff across all subbasins, particularly incorporating remotely sensed soil moisture products as reference data, is also identified as a direction for future research.

This study provides useful insights into basin-specific parameterization schemes for Noah-MP land surface model applications, with meaningful implications for operational hydrology. The results show that the default configuration remains effective in some environments, such as Subarctic climates and high elevation snow regimes. At the same time, there is potential to improve model performance in other settings by addressing the issues we identified, including overestimation of evapotranspiration, limited representation of frozen soil processes, underestimation of snow dynamics, and weak simulation of groundwater dynamics. In conclusion, these findings suggest that Noah-MP is well tuned in certain conditions, but targeted modifications can substantially enhance its reliability for water supply forecasting and flood risk management across diverse hydroclimatic regions.

Code Availability

All LIS configuration files, selected basin boundaries, R scripts, and RData files developed in this study are available at Zenodo (Cho et al., 2025; https://zenodo.org/records/17519636; last access: 4 November 2025). The LIS configuration files correspond to the Noah-MP simulations for the default setup and four alternative parameterization schemes (RUN, SFC, INF, and STC). The R scripts and RData files include all materials required to reproduce Figs. 1–8 and ensure full reproducibility of the analysis. The NASA LIS framework, including the Noah-MP land surface model, is freely available at https://github.com/NASA-LIS/LISF (Kumar et al., 2006, 2008; Peters-Lidard et al., 2007).

Data Availability

The NLDAS-2 meteorological forcing dataset used in this study is available from NASA's Land Data Assimilation Systems website (https://ldas.gsfc.nasa.gov/nldas/v2/forcing; Xia et al., ~~2012~~2012a). The basin shapefiles of HCDNs are available from the USGS ScienceBase Catalog (https://www.sciencebase.gov/catalog/item/631405bbd34e36012efa304a; Falcone, 2011). The DEM is available from the USGS National Map Download Platform (https://apps.nationalmap.gov/downloader/). Land-cover information is available from the Terra-MODIS sensor-based IGBP land classification map, modified by NCEP (https://ral.ucar.edu/model/noah-multiparameterization-land-surface-model-noah-mp-lsm). Climate classification data are available from the Köppen–Geiger climate classification map at 1 km spatial resolution (Beck et al., 2018; https://www.gloh2o.org/koppen/). The snow classification information is available from the Sturm snow classification map (Sturm and Liston, 2021; https://nsidc.org/data/nsidc-0768/versions/1). The observed daily streamflow data are available from the USGS National Water Information System (https://waterdata.usgs.gov/nwis). The UA daily 4 km SWE dataset (version 1) is publicly available at https://nsidc.org/data/nsidc-0719/versions/1 (Broxton et al., 2019). The GLEAM v4.2a evapotranspiration dataset is publicly available at https://www.gleam.eu (Koppa et al., 2022; Miralles et al., 2025).

Author contributions

EC1 (Eunsaem Cho) conceptualized the research, developed the methodology, and conducted the formal analysis and investigation, as well as prepared the original draft. EC2 (Eunsang Cho) conceptualized the research, contributed to methodology development, and supervised the research throughout the entire research process. CV contributed to the methodology, formal analysis, visualization, and supervised the research. BL and JMJ contributed to the methodology, formal analysis, and visualization, providing technical and scientific inputs. All authors reviewed and edited the paper.

Competing interests

The contact author has declared that none of the authors has any competing interests.

Acknowledgments

The authors further acknowledge the use of computing resources provided by the NASA Center for Climate Simulation.

Financial Support

This research was supported by NASA's Subseasonal-to-Seasonal Hydrometeorological Prediction program (Grant No. 80NSSC24K1278).

References

Abolafia-Rosenzweig, R., He, C., Ikeda, K., Liu, C., Lin, T.-S., Barlage, M., Xue, L., Rasouli, K., Zhang, Y., Newman, A., and Dudhia, J.: Revisiting snow water equivalent bias in the Noah-MP land surface model in the western U.S., Water Resour. Res., 62, e2025WR042586, https://doi.org/10.1029/2025WR042586, 2026.

Abolafia-Rosenzweig, R., He, C., Lin, T. S., Barlage, M., and Rittger, K.: Improved cross-scale snow cover simulations by developing a scale-aware ground snow cover fraction parameterization in the Noah-MP land surface model, J. Adv. Model. Earth Syst., 17, e2024MS004704, https://doi.org/10.1029/2024MS004704, 2025.

~~Abolafia-Rosenzweig, R., He, C., Ikeda, K., Liu, C., Lin, T. S., Barlage, M., Xue, L., Rasouli, K., Zhang, Y., Newman, A., and Dudhia, J.: Revisiting snow water equivalent bias in the Noah-MP land surface model in the western U.S., Water Resour. Res., 62, e2025WR042586, https://doi.org/10.1029/2025WR042586, 2026.~~

Abramowitz, G., Leuning, R., Clark, M., and Pitman, A.: Evaluating the performance of land surface models, J. Climate, 21, 5468–5481, https://doi.org/10.1175/2008JCLI2378.1, 2008.

Agnihotri, J., Behrangi, A., Tavakoly, A., Geheran, M., Farmani, M. A., and Niu, G.-Y.: Higher frozen soil permeability represented in a hydrological model improves spring streamflow prediction from river basin to continental scales, Water Resour. Res., 59, e2022WR033075, https://doi.org/10.1029/2022WR033075, 2023.

Ahiablame, L., Sheshukov, A. Y., Rahmani, V., and Moriasi, D.: Annual baseflow variations as influenced by climate variability and agricultural land use change in the Missouri River Basin, J. Hydrol., 551, 188–202, https://doi.org/10.1016/j.jhydrol.2017.05.055, 2017.

Alfieri, L., Burek, P., Dutra, E., Krzeminski, B., Muraro, D., Thielen, J., and Pappenberger, F.: GloFAS–global ensemble streamflow forecasting and flood early warning, Hydrol. Earth Syst. Sci., 17, 1161–1175, https://doi.org/10.5194/hess-17-1161-2013, 2013.

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

Barnhart, T. B., Tague, C. L., and Molotch, N. P.: The counteracting effects of snowmelt rate and timing on runoff, Water Resour. Res., 56, e2019WR026634, https://doi.org/10.1029/2019WR026634, 2020.

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., and Wood, E. F.: Present and future Köppen-Geiger climate classification maps at 1-km resolution, Sci. Data, 5, 180214, https://doi.org/10.1038/sdata.2018.214, 2018.

Broxton, P., Zeng, X., and Dawson, N.: Daily 4 km gridded SWE and snow depth from assimilated in-situ and modeled data over the conterminous US, version 1, NASA National Snow and Ice Data Center Distributed Active Archive Center (DAAC) [data set], https://doi.org/10.5067/0GGPB220EX6A, 2019.

Brutsaert, W.: Evaporation into the atmosphere: Theory, Reidel, Dordrecht, Holland, 299 pp., 1982.

Cai, X., Yang, Z. L., David, C. H., Niu, G. Y., and Rodell, M.: Hydrological evaluation of the Noah-MP land surface model for the Mississippi River Basin, J. Geophys. Res.-Atmos., 119, 23–38, https://doi.org/10.1002/2013JD020792, 2014.

Cao, X., Zhang, G., Chen, Y., Li, H., Li, J., Di, Y., and Cai, X.: Optimization of snow-related processes in Noah-MP land surface model over the mid-latitudes of Asian region, Atmos. Res., 311, 107711, https://doi.org/10.1016/j.atmosres.2024.107711, 2024.

Carroll, R. W., Deems, J. S., Niswonger, R., Schumer, R., and Williams, K. H.: The importance of interflow to groundwater recharge in a snowmelt-dominated headwater basin, Geophys. Res. Lett., 46, 5899–5908, https://doi.org/10.1029/2019GL082447, 2019.

Chang, M., Liao, W., Wang, X., Zhang, Q., Chen, W., Wu, Z., and Hu, Z.: An optimal ensemble of the Noah-MP land surface model for simulating surface heat fluxes over a typical subtropical forest in South China, Agr. Forest Meteorol., 281, 107815, https://doi.org/10.1016/j.agrformet.2019.107815, 2020.

Chen, F., Janjić, Z., and Mitchell, K.: Impact of atmospheric surface-layer parameterizations in the new land-surface scheme of the NCEP mesoscale Eta model, Bound.-Lay. Meteorol., 85, 391–421, https://doi.org/10.1023/A:1000531001463, 1997.

Cho, E., Cho, E., Carrie, V., Li, B., and Jacobs, J.: Data for: Basin-scale Evaluation of the Noah-MP Land Surface Model for Runoff and Snow Generation in the Missouri River Basin: Insights and Recommendations for Parameterization Scheme Selection, Zenodo [data set], https://zenodo.org/records/17519636, 2025.

Cho, E., Kwon, Y., Kumar, S. V., and Vuyovich, C. M.: Assimilation of airborne gamma observations provides utility for snow estimation in forested environments, Hydrol. Earth Syst. Sci., 27, 4039–4055, https://doi.org/10.5194/hess-27-4039-2023, 2023.

Cho, E., McCrary, R. R., and Jacobs, J. M.: Future changes in snowpack, snowmelt, and runoff potential extremes over North America, Geophys. Res. Lett., 48, e2021GL094985, https://doi.org/10.1029/2021GL094985, 2021.

Cinkus, G., Mazzilli, N., Jourde, H., Wunsch, A., Liesch, T., Ravbar, N., Chen, Z., and Goldscheider, N.: When best is the enemy of good–critical evaluation of performance criteria in hydrological models, Hydrol. Earth Syst. Sci., 27, 2397–2411, https://doi.org/10.5194/hess-27-2397-2023, 2023.

Clapp, R. B. and Hornberger, G. M.: Empirical equations for some soil hydraulic properties, Water Resour. Res., 14, 601–604, https://doi.org/10.1029/WR014i004p00601, 1978.

Cortés-Salazar, N., Vásquez, N., Mizukami, N., Mendoza, P. A., and Vargas, X.: To what extent does river routing matter in hydrological modeling?, Hydrol. Earth Syst. Sci., 27, 3505–3524, https://doi.org/10.5194/hess-27-3505-2023, 2023.

Cosgrove, B., Gochis, D., Flowers, T., Dugger, A., Ogden, F., Graziano, T., ... & Zhang, Y.: NOAA's National Water Model: Advancing operational hydrology through continental-scale modeling, JAWRA J. Am. Water Resour. As., 60, 247–272, https://doi.org/10.1111/1752-1688.13184, 2024.

Cuntz, M., Mai, J., Samaniego, L., Clark, M., Wulfmeyer, V., Branch, O., Attinger, S., and Thober, S.: The impact of standard and hard-coded parameters on the hydrologic fluxes in the Noah-MP land surface model, J. Geophys. Res.-Atmos., 121, 10676–10700, https://doi.org/10.1002/2016JD025097, 2016.

Darcy, H.: Les fontaines publiques de la ville de Dijon, Dalmont, Paris, France, 647 pp., 1856.

Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G., and Tarpley, J. D.: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, J. Geophys. Res.-Atmos., 108, 8851, https://doi.org/10.1029/2002JD003296, 2003.

Falcone, J. A.: GAGES-II: Geospatial attributes of gages for evaluating streamflow, US Geological Survey, Reston, Virginia, USA, https://doi.org/10.3133/70046617, 2011.

Fisher, R. A. and Koven, C. D.: Perspectives on the future of land surface models and the challenges of representing complex terrestrial systems, J. Adv. Model. Earth Syst., 12, e2018MS001453, https://doi.org/10.1029/2018MS001453, 2020.

Francis, B. and Melliger, J.: 2019 Missouri River Basin Historic Flood Event, J. Dam Saf., 18, 14–24, 2021.

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

Gottlieb, A. R. and Mankin, J. S.: Evidence of human influence on Northern Hemisphere snow loss, Nature, 625, 293–300, https://doi.org/10.1038/s41586-023-06794-y, 2024.

Graybeal, D. Y. and Leathers, D. J.: Snowmelt-related flood risk in Appalachia: First estimates from a historical snow climatology, J. Appl. Meteorol. Clim., 45, 178–193, https://doi.org/10.1175/JAM2330.1, 2006.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, J. Hydrol., 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Hamitouche, M., Fosser, G., Anav, A., He, C., and Lin, T. S.: Impact of runoff schemes on global flow discharge: a comprehensive analysis using the Noah-MP and CaMa-Flood models, Hydrol. Earth Syst. Sci., 29, 1221–1240, https://doi.org/10.5194/hess-29-1221-2025, 2025.

Hao, Z., Hao, F., Xia, Y., Singh, V. P., Hong, Y., Shen, X., and Ouyang, W.: A statistical method for categorical drought prediction based on NLDAS-2, J. Appl. Meteorol. Clim., 55, 1049–1061, https://doi.org/10.1175/JAMC-D-15-0200.1, 2016.

He, C., Valayamkunnath, P., Barlage, M., Chen, F., Gochis, D., Cabell, R., Schneider, D., Ikeda, K., Lin, T. S., Castro, C., Bruyere, C., Ahijevych, D., Rasmussen, R., Jin, J., Trahan, S., Neale, R., and Ek, M.: The community Noah-MP land surface modeling system technical description version 5.0, NCAR Technical Note NCAR/TN-575+STR, National Center for Atmospheric Research, Boulder, CO, USA, https://doi.org/10.5065/ew8g-yr95, 2023.

Hong, S., Yu, X., Park, S. K., Choi, Y. S., and Myoung, B.: Assessing optimal set of implemented physical parameterization schemes in a multi-physics land surface model using genetic algorithm, Geosci. Model Dev., 7, 2517–2529, https://doi.org/10.5194/gmd-7-2517-2014, 2014.

Hu, W., Ma, W., Yang, Z. L., Ma, Y., and Xie, Z.: Sensitivity analysis of the Noah-MP land surface model for soil hydrothermal simulations over the Tibetan Plateau, J. Adv. Model. Earth Syst., 15, e2022MS003136, https://doi.org/10.1029/2022MS003136, 2023.

Huntington, J. L. and Niswonger, R. G.: Role of surface-water and groundwater interactions on projected summertime streamflow in snow dominated regions: An integrated modeling approach, Water Resour. Res., 48, W11524, https://doi.org/10.1029/2012WR012319, 2012.

Jordan, Y. C., Ghulam, A., and Herrmann, R. B.: Floodplain ecosystem response to climate variability and land-cover and land-use change in Lower Missouri River basin, Landscape Ecol., 27, 843–857, https://doi.org/10.1007/s10980-012-9748-x, 2012.

Knoben, W. J., Freer, J. E., and Woods, R. A.: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores, Hydrol. Earth Syst. Sci., 23, 4323–4331, https://doi.org/10.5194/hess-23-4323-2019, 2019.

Koppa, A., Rains, D., Hulsman, P., Poyatos, R., and Miralles, D. G.: A deep learning-based hybrid model of global terrestrial evaporation, Nat. Commun., 13, 1912, https://doi.org/10.1038/s41467-022-29543-7, 2022.

Koren, V., Schaake, J., Mitchell, K., Duan, Q. Y., Chen, F., and Baker, J. M.: A parameterization of snowpack and frozen ground intended for NCEP weather and climate models, J. Geophys. Res.-Atmos., 104, 19569–19585, https://doi.org/10.1029/1999JD900232, 1999.

Kumar, S. V., Peters-Lidard, C. D., Tian, Y., Houser, P. R., Geiger, J., Olden, S., Lighty, L., Eastman, J. L., Doty, B., Dirmeyer, P., Adams, J., Mitchell, K., Wood, E. F., and Sheffield, J.: Land information system: An interoperable framework for high resolution land surface modeling, Environ. Modell. Softw., 21, 1402–1415, https://doi.org/10.1016/j.envsoft.2005.07.004, 2006.

Kumar, S. V., Reichle, R. H., Peters-Lidard, C. D., Koster, R. D., Zhan, X., Crow, W. T., Eylander, J. B., and Houser, P. R.: A land surface data assimilation framework using the land information system: Description and applications, Adv. Water Resour., 31, 1419–1432, https://doi.org/10.1016/j.advwatres.2008.01.013, 2008.

Lapides, D. A., Hahm, W. J., Rempe, D. M., Whiting, J., and Dralle, D. N.: Causes of missing snowmelt following drought, Geophys. Res. Lett., 49, e2022GL100505, https://doi.org/10.1029/2022GL100505, 2022.

Li, B., Rodell, M., Peters-Lidard, C., Erlingis, J., Kumar, S., and Mocko, D.: Groundwater recharge estimated by land surface models: An evaluation in the conterminous United States, J. Hydrometeorol., 22, 499–522, https://doi.org/10.1175/JHM-D-20-0130.1, 2021.

Li, J., Miao, C., Zhang, G., Fang, Y. H., Shangguan, W., and Niu, G. Y.: Global evaluation of the Noah-MP land surface model and suggestions for selecting parameterization schemes, J. Geophys. Res.-Atmos., 127, e2021JD035753, https://doi.org/10.1029/2021JD035753, 2022.

Liu, X., Chen, F., Barlage, M., Zhou, G., and Niyogi, D.: Noah-MP-Crop: Introducing dynamic crop growth in the Noah-MP land surface model, J. Geophys. Res.-Atmos., 121, 13953–13972, https://doi.org/10.1002/2016JD025597, 2016.

Liu, Y., Fernández-Ortega, J., Mudarra, M., and Hartmann, A.: Pitfalls and a feasible solution for using KGE as an informal likelihood function in MCMC methods: DREAM(ZS) as an example, Hydrol. Earth Syst. Sci., 26, 5341–5355, https://doi.org/10.5194/hess-26-5341-2022, 2022.

Ma, N., Niu, G. Y., Xia, Y., Cai, X., Zhang, Y., Ma, Y., and Fang, Y.: A systematic evaluation of Noah-MP in simulating land-atmosphere energy, water, and carbon exchanges over the continental United States, J. Geophys. Res.-Atmos., 122, 12245–12268, https://doi.org/10.1002/2017JD027597, 2017.

Maina, F. Z., Rosen, D., Abbaszadeh, P., Yang, C., Kumar, S. V., Rodell, M., and Maxwell, R.: Integrating the interconnections between groundwater and land surface processes through the coupled NASA land information system and ParFlow environment, J. Adv. Model. Earth Syst., 17, e2024MS004415, https://doi.org/10.1029/2024MS004415, 2025.

Miralles, D. G., Bonte, O., Koppa, A., Baez-Villanueva, O. M., Tronquo, E., Zhong, F., Beck, H. E., Hulsman, P., Dorigo, W., Verhoest, N. E. C., and Haghdoost, S.: GLEAM4: global land evaporation and soil moisture dataset at 0.1° resolution from 1980 to near present, Sci. Data, 12, 416, https://doi.org/10.1038/s41597-025-04610-y, 2025.

Musselman, K. N., Lehner, F., Ikeda, K., Clark, M. P., Prein, A. F., Liu, C., Barlage, M., and Rasmussen, R.: Projected increases and shifts in rain-on-snow flood risk over western North America, Nat. Clim. Change, 8, 808–812, https://doi.org/10.1038/s41558-018-0236-4, 2018.

NCEI (National Centers for Environmental Information): Extreme Wetness of 2019 Sets Records, National Oceanic and Atmospheric Administration, Silver Spring, Maryland, USA, available at: https://www.ncei.noaa.gov/news/extreme-wetness-2019-sets-records, 2020.

Nearing, G. S., Mocko, D. M., Peters-Lidard, C. D., Kumar, S. V., and Xia, Y.: Benchmarking NLDAS-2 soil moisture and evapotranspiration to separate uncertainty contributions, J. Hydrometeorol., 17, 745–759, https://doi.org/10.1175/JHM-D-15-0063.1, 2016.

Niu, G. Y. and Yang, Z. L.: Effects of frozen soil on snowmelt runoff and soil water storage at a continental scale, J. Hydrometeorol., 7, 937–952, https://doi.org/10.1175/JHM538.1, 2006.

Niu, G. Y., Yang, Z. L., Dickinson, R. E., and Gulden, L. E.: A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models, J. Geophys. Res.-Atmos., 110, D21106, https://doi.org/10.1029/2005JD006111, 2005.

Niu, G. Y., Yang, Z. L., Dickinson, R. E., Gulden, L. E., and Su, H.: Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data, J. Geophys. Res.-Atmos., 112, D07103, https://doi.org/10.1029/2006JD007522, 2007.

Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., and Xia, Y.: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements, J. Geophys. Res.-Atmos., 116, D12109, https://doi.org/10.1029/2010JD015139, 2011.

Overgaard, J., Rosbjerg, D., and Butts, M. B.: Land-surface modelling in hydrological perspective – a review, Biogeosciences, 3, 229–241, https://doi.org/10.5194/bg-3-229-2006, 2006.

Pascolini-Campbell, M. and Reager, J. T.: An Investigation of the Spatial and Temporal Characteristics of Extreme Dry and Wet Events across NLDAS-2 Models, J. Hydrometeorol., 25, 239–255, https://doi.org/10.1175/JHM-D-23-0038.1, 2024.

Peters-Lidard, C. D., Houser, P. R., Tian, Y., Kumar, S. V., Geiger, J., Olden, S., Lighty, L., Doty, B., Dirmeyer, P., Adams, J., Mitchell, K., Wood, E. F., and Sheffield, J.: High-performance Earth system modeling with NASA/GSFC's Land Information System, Innovations in Systems and Software Engineering, 3, 157–165, https://doi.org/10.1007/s11334-007-0028-x, 2007.

Qiao, L., Pan, Z., Herrmann, R. B., and Hong, Y.: Hydrological variability and uncertainty of lower Missouri River basin under changing climate, JAWRA J. Am. Water Resour. As., 50, 246–260, https://doi.org/10.1111/jawr.12126, 2014.

Qin, H., Thompson, M. R., Garcia, L. J., and Patel, S. K.: Extreme value snow water equivalent and snowmelt for hydrologic design, Water Resour. Res., 56, e2020WR028126, https://doi.org/10.1029/2020WR028126, 2020.

Shirzaei, M., Khoshmanesh, M., Ojha, C., Werth, S., Kerner, H., Carlson, G., Sherpa, S. F., Naghibi, F., and Lee, J. C.: Persistent impact of spring floods on crop loss in US Midwest, Weather and Climate Extremes, 34, 100392, https://doi.org/10.1016/j.wace.2021.100392, 2021.

Siirila-Woodburn, E. R., Rhoades, A. M., Hatchett, B. J., Huning, L. S., Szinai, J., Tague, C., Nico, P. S., Feldman, D. R., Jones, A. D., Collins, W. D., and Kaatz, L.: A low-to-no snow future and its impacts on water resources in the western United States, Nat. Rev. Earth Environ., 2, 800–819, https://doi.org/10.1038/s43017-021-00219-y, 2021.

Stadnyk, T., Dow, K., Wazney, L., and Blais, E. L.: The 2011 flood event in the Red River Basin: Causes, assessment and damages, Can. Water Resour. J., 41, 65–73, https://doi.org/10.1080/07011784.2015.1008048, 2016.

Sthapit, E., Lakhankar, T., Hughes, M., Khanbilvardi, R., Cifelli, R., Mahoney, K., Martin, R., Shrestha, R., and Rafieeinasab, A.: Evaluation of snow and streamflows using Noah-MP and WRF-hydro models in Aroostook River basin, Maine, Water, 14, 2145, https://doi.org/10.3390/w14142145, 2022.

Sturm, M. and Liston, G. E.: Revisiting the global seasonal snow classification: An updated dataset for earth system applications, J. Hydrometeorol., 22, 2917–2938, https://doi.org/10.1175/JHM-D-21-0070.1, 2021.

Su, L., Lettenmaier, D. P., Pan, M., and Bass, B.: Improving runoff simulation in the Western United States with Noah-MP and variable infiltration capacity, Hydrol. Earth Syst. Sci., 28, 3079–3097, https://doi.org/10.5194/hess-28-3079-2024, 2024.

USACE (US Army Corps of Engineers): National Inventory of Dams, US Army Corps of Engineers, Washington, D.C., USA, 2021.

USGS (US Geological Survey): USGS hydro-climatic data network 2009 (HCDN-2009), US Geological Survey, Reston, Virginia, USA, https://doi.org/10.3133/fs20123047, 2012.

Velasquez, N., Quintero, F., Koya, S. R., Roy, T., and Mantilla, R.: Snow-detonated floods: Assessment of the US midwest march 2019 event, J. Hydrol.-Reg. Stud., 47, 101387, https://doi.org/10.1016/j.ejrh.2023.101387, 2023.

Vormoor, K., Lawrence, D., Schlichting, L., Wilson, D., and Wong, W. K.: Evidence for changes in the magnitude and frequency of observed rainfall vs. snowmelt driven floods in Norway, J. Hydrol., 538, 33–48, https://doi.org/10.1016/j.jhydrol.2016.03.066, 2016.

Wise, E. K., Woodhouse, C. A., McCabe, G. J., Pederson, G. T., and St-Jacques, J. M.: Hydroclimatology of the Missouri River basin, J. Hydrometeorol., 19, 161–182, https://doi.org/10.1175/JHM-D-17-0155.1, 2018.

Wood, A. W. and Lettenmaier, D. P.: A test bed for new seasonal hydrologic forecasting approaches in the western United States, B. Am. Meteorol. Soc., 87, 1699–1712, https://doi.org/10.1175/BAMS-87-12-1699, 2006.

Xia, Y., Cosgrove, B. A., Mitchell, K. E., Peters-Lidard, C. D., Ek, M. B., Kumar, S., Mocko, D., Barlage, M., Livneh, B., Cai, X., and Wei, H.: Basin-scale assessment of the land surface energy budget in the National Centers for

Environmental Prediction operational and research NLDAS-2 systems, J. Geophys. Res.-Atmos., 121, 196–220, https://doi.org/10.1002/2015JD023889, 2016.

Xia, Y., Cosgrove, B. A., Mitchell, K~~., Ek, M., Sheffield, J., Cosgrove, B., Wood,~~. E., ~~Luo, L., Alonge~~Peters-Lidard, C~~., Wei, H., Meng, J., Livneh,~~. D., Ek, M B., ~~Lettenmaier, D., Koren, V., Duan, Q., Mo, K., Fan, Y., and~~ Kumar, S., Mocko, D., and Wei, H.: 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~~2. Validation of model ~~products~~-simulated streamflow, J. Geophys. Res.-Atmos., 117, ~~D03109~~D03110, https://doi.org/10.1029/~~2011JD016048, 2012~~2011JD016051, 2012b.

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, L., Alonge, C., Wei, H., Meng, J., Livneh, B., Lettenmaier, D., Koren, V., Duan, Q., Mo, K., Fan, Y., and Mocko, D.: 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, J. Geophys. Res.-Atmos., 117, D03109, https://doi.org/10.1029/2011JD016048, 2012a.

Xia, Y., Peters-Lidard, C. D., Huang, M., Wei, H., and Ek, M.: Improved NLDAS-2 Noah-simulated hydrometeorological products with an interim run, Hydrol. Process., 29, 780–792, https://doi.org/10.1002/hyp.10190, 2014.

Yang, K., Yang, J., Zhaoye, P., Zhang, F., Zhang, G., and Wang, C.: Optimization and validation of soil frozen–thawing parameterizations in Noah-MP, J. Geophys. Res.-Atmos., 128, e2022JD038217, https://doi.org/10.1029/2022JD038217, 2023.

Yang, Z.-L., Niu, G.-Y., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Longuevergne, L., Manning, K., Niyogi, D., Tewari, M., and Xia, Y.: The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins, J. Geophys. Res.-Atmos., 116, D12110, https://doi.org/10.1029/2010JD015140, 2011.

Yin, J., Zhan, X., Barlage, M., Kumar, S., Fox, A., Albergel, C., Mahfouf, J.-F., Hain, C., and Liu, J.: Assimilation of blended satellite soil moisture data products to further improve Noah-MP model skills, J. Hydrol., 621, 129596, https://doi.org/10.1016/j.jhydrol.2023.129596, 2023.

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