

Response to Comments by Editor

General Comments:

The manuscript has improved substantially, and most of the major concerns raised during the earlier review have been addressed. However, I am not able to accept the manuscript in its present form. A final revision is required to resolve several issues, including:

Response: We thank the editor for the constructive feedback and for the opportunity to further revise the manuscript. We have addressed each of the three issues raised below, with attention to revising the newly added ROS-related results to instead describe the improvement in terms of winter-season runoff more generally (Q1), ensuring consistent inclusion of the Abolafia-Rosenzweig et al. (2026) citation (Q2), and explicitly stating our runoff aggregation, routing, and soil-moisture depth methodology (Q3). We believe these revisions have strengthened the manuscript and thank the editor for the review throughout this process.

Q1. Rain-on-snow and peak-runoff analysis. The newly added analysis uses annual maximum monthly runoff and excludes years in which simulated and observed peak timing differs by more than one month. This does not identify rain-on-snow events and may condition the magnitude analysis on successful timing. Please either (a) reframe this as an exploratory analysis of annual monthly-peak runoff bias, remove the claim that it validates ROS or rapid flood response, report the number of retained and excluded years, and provide the underlying basin-by-year results; or (b) support the ROS claim with an event-based analysis at an appropriate temporal resolution. Please also reconcile the reported inconsistency range (71.0% versus 72.0%) and the mean improvement (25.1%), which is present in the response and marked-up manuscript but absent from the clean manuscript.

Response: This is an important clarification, and we agree that the previous analysis, based on monthly-peak runoff, could not identify discrete rain-on-snow events and mixed timing skill with magnitude skill. In response to this comment, we conducted an additional daily-resolution comparison to check whether the INF scheme's peak-runoff benefit held at a daily temporal resolution. For each of the 36 subbasins where the INF experiment improved overall KGE, we identified the observed peak daily streamflow within January–March for each water year (2014–2023), and compared the default and INF simulated runoff on that exact observed peak date. We found that peak-flow bias improved in 24 of the 36 subbasins (66.7%), with improvement ranging

from 0.3% to 93.0% and averaging 30.7%, which is broadly consistent with the direction of the original monthly-resolution result.

While this additional comparison provided useful context, we felt that it would not fully resolve the methodological concern raised by the reviewer. In particular, evaluating runoff on the observed daily peak date does not by itself identify whether the peak was associated with a rain-on-snow event. Thus, although the daily-resolution results supported the original analysis, we did not consider them sufficient to support a more specific interpretation in terms of ROS processes. A more direct assessment would require an event-based framework with explicit precipitation-phase and/or snowpack-based criteria at daily or sub-daily resolution. We therefore decided not to introduce the daily-resolution comparison as an additional quantitative result in the Discussion, but instead to revise the interpretation of the existing analysis to more appropriately reflect what it can support.

Accordingly, we removed the specific quantitative peak-runoff result from Section 6.1 and revised the framing of the discussion so that the benefit of the INF scheme is described more generally in terms of winter-season runoff, rather than being attributed specifically to ROS events. We also revised several statements in Section 6.1 to use more cautious wording and to avoid implying that the existing analysis provides direct validation of ROS-related peak-flow responses. We believe this more conservative interpretation better reflects the scope of the analysis while avoiding an overinterpretation of the results. The changed sentences are shown in the table below. A more direct event-based evaluation of ROS processes using precipitation-phase or snowpack-based criteria would be a valuable direction for future work.

Before	After
The first limitation identified was the inadequate groundwater dynamics in the default groundwater scheme.	The first area for improvement was the limited groundwater dynamics in the default groundwater scheme.
The second limitation involved systematic overestimation of ET in vegetated areas, particularly those dominated by cropland.	The second area for improvement involved systematic overestimation of ET in vegetated areas, particularly those dominated by cropland.
...the default Noah-MP crop parameterization inadequately represents	...the default Noah-MP crop parameterization does not fully represent dynamic cropland

dynamic cropland processes, leading to possible overestimation of ET (Liu et al., 2016).	processes, which can contribute to overestimation of ET (Liu et al., 2016).
A third limitation is the potential underrepresentation of rapid surface runoff generation that can occur during rain-on-snow (ROS) events.	A third area for improvement relates to winter-season surface runoff generation under frozen-soil conditions.
The default Noah-MP configuration inadequately captures the physical processes associated with ROS events, including ice lens formation, changes in snowpack thermal conductivity, and alterations in water retention capacity.	The default Noah-MP configuration does not fully capture the physical processes associated with frozen and partially frozen soils, including ice lens formation, changes in snowpack thermal conductivity, and alterations in water retention capacity.
This is particularly important for improving runoff estimates during ROS events, which are often underestimated due to excessive infiltration by the default scheme.	This is particularly relevant for improving winter-season runoff estimates, which are often underestimated by the default scheme due to excessive infiltration 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 the default scheme and the improvements offered by alternative schemes reflect the specific hydroclimatic conditions of the Missouri River Basin.	It is important to note, however, that the identified areas for improvement in the default scheme, and the gains offered by alternative schemes, reflect the specific hydroclimatic conditions of the Missouri River Basin.

Q2. The authors state in their response that they added Abolafia-Rosenzweig et al. (2026). This addition appears in the marked-up manuscript (ATC2) but is absent from the clean manuscript (version 3). The two submitted versions should be made consistent.

Response: We appreciate the editor pointing out this inconsistency. We have corrected this oversight, and the citation now appears in both the marked-up and clean versions of the manuscript, including in the reference list.

Q3. Please state explicitly how gridded runoff was aggregated to each gauge basin, whether river routing was applied, what soil-moisture depth is shown in Fig. 4.

Response: We appreciate the editor requesting these methodological clarifications, and we have revised the manuscript to state each of these points explicitly. 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, and no river routing scheme was applied. We have added a paragraph to Section 6.4 acknowledging the absence of explicit routing as a limitation of this approach, supported by two references from the literature. The soil moisture shown in Fig. 4 represents the full soil column depth (2.0 m), integrated across all four Noah-MP soil layers. We thank the editor for prompting these clarifications, which strengthen the methodological transparency of the manuscript.

Revised Paragraph (Page 13, Lines 370 – 373)

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.

Revised Paragraph (Page 23, Lines 585 – 594)

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., 2012). 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.

Added References

- 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.
- Xia, Y., Cosgrove, B. A., Mitchell, K. E., Peters-Lidard, C. D., Ek, M B., 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): 2. Validation of model-simulated streamflow, *J. Geophys. Res.-Atmos.*, 117, D03110, <https://doi.org/10.1029/2011JD016051>, 2012.