The First Reviewer

It's my pleasure to review this manuscript on NoahMP model. The study utilized the Noah-MP land surface model, implementing five different soil hydraulic parameterization schemes, including two soil hydraulic models, a dual-permeability infiltration scheme, and variations in surface ponding depth. By integrating soil moisture datasets from SMAP and ISMN, the study explored the effects of these schemes on Soil Moisture Memory (SMM). It was found that using the Van-Genuchten parameterization and dual-permeability infiltration scheme improved the simulation of SMM. Considering surface ponding extended soil moisture infiltration, thereby improving moisture conditions in the surface and root zone, leading to increased long-term memory and decreased short-term memory. Conversely, using macropores reduced SMM. In general, the manuscript is well written and fully within the scope of HESS. While I still have some concerns before the recommendation of its publication.

Major:

1. When designing the five different Noah-MP model scenarios, how did the authors determine the specific parameter values (such as the ponding depth threshold)? Were these parameter values selected based on previous literature, sensitivity analyses, or other considerations?

   R: In our study, we conducted sensitivity analyses for key parameters such as the ponding depth threshold, macropore fraction, and other relevant factors. The initial parameter values were determined based on testing various configurations to assess their influence on the model's performance in modeling surface soil moisture, runoff peaks, and baseflow. These tests allowed us to identify a range of values that produced realistic outcomes, but we acknowledge that the selected values are not final. Further parameterization scheme and calibration is necessary to optimize these parameters, and this will be part of our future work to refine the model's accuracy.

   The following is added to paper:
   "While we conducted sensitivity analyses on key parameters such as the ponding depth threshold and macropore fraction to identify ranges yielding realistic outcomes, we acknowledge that further calibration is necessary to optimize these parameters and will address this in future work to enhance the model's accuracy."

2. In Figure 4, the four VGM schemes exhibit greater variability in wet regions, unlike Figure 6 where their PDF distributions are generally consistent in both arid and wet regions. Ji et al. (2023) enhanced the synergistic effect of high-resolution soil dataset and model parameterizations. Here, this discrepancy may be also caused by differences in how each VGM scheme incorporates local soil moisture dynamics and interactions with vegetation, leading to varying responses across different hydrological regimes.

Reference: Ji, P., Yuan, X., & Jiao, Y. (2023). Synergistic effects of high-resolution factors for improving soil moisture simulations over China. Water Resources Research, 59, e2023WR035513.

R: The following is added to paper:
"The modeling results also indicate the long-term memory of the surface soil moisture is more sensitive to the four VGM schemes in the wet regions (Figure 4b) than the short-term memory (

Figure 6b). This can be attributed to the differences in how topsoil water responds to surface ponding and preferential flow as represented by the four VGM across different moisture regimes. Under higher soil moisture conditions right after a rainfall event, the persistence of soil moisture is mainly dominated by drainage of topsoil water to deeper soil, whereas at relatively lower soi moisture, the long-term memory is more controlled by persistent water inputs from surface ponded water and prolonged drainage by preferential flow. This also indicates that the effects infiltration of surface ponded water and preferential flow can last longer up to more than 10 days. Under dry conditions (Figure 4a and 6a), these hydrological processes become less important. However, the soil water retention curves as represented by the CH and VG schemes play a more important role under any conditions (Figure 4a and Figure 6a). Another possible reason could be the issue of time scale. Short-term memory has values up to 5 days, and given the SMAP revisit time of 3 days, generating values for intervals shorter than 3 days may challenge the validity of short-term memory as a reliable measurement for soil drainage, as demonstrated by McColl et al. (2019). Since we selected Noah-MP days corresponding to the SMAP revisit time, it is possible that the effects of different VG parameterizations were diminished by this sampling. We suggest that other measurements, such as streamflow and baseflow analysis, should be considered to better quantify the effect of soil hydraulics on soil drainage."

When comparing different model scenarios with SMAP and ISMN data, the study primarily focuses on the median of SMM. However, the shape and range of SMM distributions could also provide valuable information. Analyzing discrepancies between model-generated SMM distributions and observational data can offer a more comprehensive assessment of model performance.

R: We acknowledge that analyzing the shape and range of Soil Moisture Memory (SMM) distributions could provide valuable insights into the model's performance relative to observational data. However, the primary focus of this study is to assess the general trends and central tendencies of SMM by comparing the median values across different model scenarios. The use of median SMM serves as a robust metric that reduces the influence of outliers and extreme values, offering a clearer comparison of core model performance.

Hence the following is added to the paper:
"While analyzing the shape and range of Soil Moisture Memory (SMM) distributions could offer valuable insights, we focus on comparing median SMM values across model scenarios to assess general trends and central tendencies, thereby reducing the influence of outliers for a clearer evaluation of model performance."

Minor

1. Lines 399-447, The article extensively describes various parameterization schemes of the Noah-MP model, but could further discuss the underlying physical mechanisms of these schemes to help readers better understand their impact on SMM.

   R: Yes, we totally agree because this paper focuses on "why" not "accuracy." There is another paper in which we provide more details of the underlying physics, hence the following sentences is added to the paper:
   "A detailed description the underlying physical mechanisms of the schemes used in this

study could be found at Niu et al, (2024), also a brief description of equations and parameters is included in supporting material" We also revised the paper, in the result analysis, to add the underlying mechanisms to explain the modeling results, although we already provided explanations in the Discussion section.

2. Lines 466-467"The MF_CH experiment displays a spatial pattern that contrasts with the SMAP data, with a longer memory in the arid western regions but a shorter memory in the wet northeastern regions " What could be the reasons for this spatial distribution?

   R: We added an explanation to the paper for it:
   " This is likely caused by the faster drainage of topsoil water under the wetter conditions, whereas under the drier conditions, the spuriously stronger suction from the CH hydraulics sustain the surface soil moisture for a longer period."

3. While discussing the limitations of ISMN data, the paper mentions the issue of scale differences between point measurements and grid-scale data. I suggested to added some discussions on the high-resolution or hyper-resolution, which might be an efficiency way to solve the scale mismatch between observation and simulation.

   R: We added a paragraph updated to:
   "Another concern is the influence of ISMN spatial representation on SMM analysis. ISMN stations are point-based, and it is assumed that one point represents a 1/8-degree grid area. It is possible that the point measurements cannot fully capture the spatial variability within the Noah-MP grid cells, leading to discrepancies in the representation of values and spatial patterns. The limited number of stations may further amplify this issue. One potential solution to address the scale mismatch between point-based observations and grid-scale simulations is the use of high-resolution or hyper-resolution models. These models can provide finer spatial detail, allowing for a more direct comparison between observational data and model outputs, thereby improving the accuracy of the analysis and reducing scale-induced biases. Incorporating such approaches in future studies could help mitigate the limitations posed by the current scale differences."

4. Lines 712-713:" ... processes that influence SMM and to address the commonly observed overestimation/underestimation of long-term/short-term SMM in LSMs." I suggest to add some implications of these findings for future research on SMM.

R: Based on the Rahmati et al (2024) suggestions we added the following to the paper: "The findings from this study have important implications for future research on SMM. By identifying the specific parameterizations that lead to discrepancies in long-term and short-term SMM, future studies can focus on refining these parameters to reduce biases in LSMs. Moreover, while this study focuses on the effect of different parameterizations on the timescale of SMM, future research can analyze the impact of these parameterizations on the strength and legacy of SMM and assess whether the findings based on timescale align with those related to strength and legacy."

Citation: Rahmati, M., Amelung, W., Brogi, C., Dari, J., Flammini, A., Bogena, H., Brocca, L., Chen, H., Groh, J., Koster, R. D., McColl, K. A., Montzka, C., Moradi, S., Rahi, A., Sharghi S., F., and Vereecken, H.: Soil Moisture Memory: State-Of-The-Art and the Way Forward, Reviews of Geophysics, 62, e2023RG000828, 2024.