Response to Anonymous Referee #1

The manuscript evaluates seven runoff schemes in the Noah-MP land surface model for estimating global river discharge, comparing them to ERA5-Land data and streamflow observations. Results show varying accuracy, with TOPMODEL-based schemes underestimating runoff in some regions, while others like Schaake and BATS performed better. **Dynamic** VIC overestimated runoff globally. The study indicated that despite good performance, biases in high-flow extremes highlight the need for further model improvements. The study emphasizes improving hydrological models for accurate water resource management and climate adaptation.

The study addresses a quite interesting topic. The manuscript is well organized and neatly written with the appropriate scientific content. However, I have some suggestions and questions as follow.

Major comments:

1) The paper does not adequately address how the insights from these evaluations could be used to advance global hydrological modeling, particularly in the context of discharge simulation. While it provides a thorough assessment of the performance of seven runoff schemes within the Noah-MP Land Surface Model, its contribution to improving hydrological modeling remains unclear.

- ***Thank you for your valuable feedback. We agree with the reviewer, hence we have added a dedicated section "3.5 Implications for Global Discharge Simulation" to discuss how our findings contribute to improve hydrological modelling. See below.***

3.5 Implications for Global Discharge Simulation:

"This study provides a comprehensive analysis of how distinct Noah-MP runoff schemes impact discharge simulation in a global hydrological context, paving the way for enhanced modelling accuracy across different climate zones. By revealing the performance variability of different runoff schemes—such as Schaake, BATS, VIC, and XAJ—across cold, warm temperate, tropical, and arid regions, this research suggests that tailored scheme selection could improve discharge simulations for specific hydrological conditions. For instance, Schaake, BATS, and VIC exhibited reduced biases in warm temperate and tropical regions, while TOPMODEL-based schemes with groundwater performed notably better in arid areas, underscoring the need for strategic scheme selection based on regional climate and hydrological characteristics. This targeted approach to scheme selection can minimize bias and enhance model reliability in both regional and global discharge simulations, improving the accuracy of water resource management.

This study also addresses a crucial challenge in hydrological modelling: the significant biases in high-flow extremes by certain schemes. Given that accurate high-flow discharge predictions are essential for flood forecasting and disaster management, this finding suggests an urgent need for refining high-flow calibration. This enhancement is particularly relevant for global flood risk management, as it enables more reliable flood predictions that are vital for preparing for extreme weather events. By identifying critical parameters and proposing spatially variable adjustments, such as using data from sources like remote sensing products, this study sets a practical foundation for developing global calibration strategies that could yield more accurate discharge predictions (Beck et al., 2017). These strategies could be applied

universally across a range of climates, creating a more adaptable global model without extensive customization.

This research further advances current hydrological modelling by demonstrating the value of multi-model comparisons, which allow for a holistic approach to discharge simulation. Rather than depending on a single runoff scheme and potentially inheriting its limitations (Diks and Vrugt, 2010; Shoaib et al., 2018), a multi-scheme approach enables researchers to capture river discharge dynamics more comprehensively (Georgakakos et al., 2004; Huo et al., 2019). This approach aligns with a broader hydrological perspective that considers the interactions between multiple runoff dynamics, offering a pathway for more nuanced simulations that acknowledge the strengths and limitations of each scheme.

For coupled ocean-atmosphere regional models lacking complete river and discharge representations, integrating findings from this study could significantly improve their hydrological modules, particularly in complex regions like the Mediterranean where the freshwater flux from rivers remarkably affects the salinity near the coast close to river mouths (e.g. Reale et al., 2020). These refinements are expected to enhance the overall representation of the global water cycle within climate models, providing more realistic freshwater flux predictions and supporting more accurate climate projections.

Additionally, this study's analysis of seasonal and regional discharge cycles reveals new insights into the variability of discharge patterns across climates. This detailed understanding could facilitate the development of models better suited to capture seasonal dynamics in tropical and temperate regions, where runoff schemes like Schaake and ERA5-Land-driven simulations performed particularly well. By capturing the discharge seasonality more accurately, our findings have direct applications for both short- and long-term water resource planning (Pires and Martins, 2024), especially in regions facing pronounced seasonal changes in water availability.

Furthermore, the findings related to groundwater interactions underscore the importance of accurate groundwater dynamics in discharge simulation, especially in arid regions. The effectiveness of TOPMODEL-based schemes with groundwater dynamics in these areas suggests that future modelling efforts should prioritize improving groundwater parameterizations, particularly where groundwater plays a critical role in maintaining streamflow. This refinement could improve discharge simulations (Decharme and Colin, 2024) in water-scarce areas, supporting more efficient resource allocation and resilience against drought.

Finally, the implications of these findings extend to climate adaptation strategies, where reliable hydrological models are critical for anticipating shifts in water availability under changing climates. By advancing the accuracy of discharge simulations, particularly in high-flow and seasonal scenarios, this research provides a basis for better-informed adaptation planning, enabling decision-makers to prepare for anticipated changes in river flow and water availability. This study not only enhances current global hydrological modelling but also lays a foundation for more resilient water resource management, which is increasingly critical as climate variability challenges water availability worldwide."

2) Line 358-362: It is noted that the lags between peak runoff and peak discharge in large river basins, such as the Amazon, are attributed to the natural routing lag within the river network. Could these lags also be due to specific limitations within the CaMa-Flood global river routing model? Have you conducted any sensitivity analysis on the models to explore this?

- ***Thank you for raising this important point. We agree that the lags between peak runoff and peak discharge in large basins like the Amazon could potentially be influenced by specific limitations within the CaMa-Flood global river routing model, in addition to the natural routing lag within the river network.***

  Since our current analysis primarily focused on the sensitivity to runoff schemes, we did not conduct a detailed sensitivity analysis specifically targeting the routing parameterisation in CaMa-Flood. We acknowledge that this could provide valuable insights into the influence of model-specific limitations on the timing of peak discharge, particularly in large-scale river basins where routing dynamics are more complex. Besides, part of the analysis suggested by the reviewer for this specific basin has already been performed in other studies (e.g. Yamazaki et al, 2012; https://doi.org/10.1029/2012WR011869).

  Thus, as this aspect was outside the scope of the current study, we suggest that a sensitivity analysis on key parameters (such as river velocity, roughness coefficients, or floodplain dynamics) within the CaMa-Flood model could be an important direction for future research. Such an analysis would help isolate the contributions of model limitations from natural routing processes and provide a clearer understanding of the observed lags.

  We appreciate the reviewer's suggestion and will consider it in future studies to further refine the modelling of discharge timing in large river systems.

Minor comments:

1) Abbreviations are used in the abstract that may be unclear to readers who are not very familiar with the study.

   - ***Thank you for your feedback. We have revised the abstract to define all abbreviations upon first use, ensuring clarity for readers.***

2) Line 25: Rephrase this sentence to better highlight the importance of this study, e.g.: “These findings are critical for improving global hydrological models, which are essential for developing more reliable water resource management strategies and adapting to the growing challenges posed by climate change, such as shifts in water availability and extreme flood events."

   - ***We have revised the sentence incorporating the suggested phrasing.***

3) Line 39: Please use formal expression for the “On the flip side, ...”

   - ***Thank you for your suggestion. We have revised the expression to a more formal tone as recommended (“On the other hand, …”).***

4) Line 202: Is it correct that the ERA5-Land variables were regridded from 0.1° to 0.2°? If so, perhaps using a term like "regridding," "spatial aggregation," or "extrapolation" would be more appropriate.

- ***Thank you for your suggestion. We have replaced 'interpolated' with 'regridded', which is more correct.***

5) Line 241: In this study, ERA5-Land runoff is used as a benchmark for evaluating the runoff simulated by Noah-MP. While ideally, direct runoff observations would be used for this purpose, such data was not available, as you mentioned. To further strengthen your evaluation, and ensuring that the simulations align with real-world observations, it would be helpful to cite studies that have assessed ERA5-Land runoff against direct runoff measurements to demonstrate the reliability of ERA5-Land as a reference dataset.

- ***Thank you for your valuable suggestion. We were unable to find studies that directly assess ERA5-Land runoff against direct runoff measurements. However, we have included the available reference that assesses ERA5-Land runoff against P-E and simulated runoff highlighting strength and weakness of the dataset.***

6) Page 10, Figure 1: What was the reason behind selecting these specific river basins? Were they chosen based on their size as the largest river basins?

- ***The specific river basins were selected based on their status as large river basins globally, while also being constrained by the availability of consistent data.***

7) Line 393-394: I suggest using a more complete version of your statement something like this "According to the water balance equation, within a defined area over a specific period, the total inflows (such as precipitation) must equal the total outflows (including runoff and evapotranspiration), plus any change in storage (such as changes in soil moisture, groundwater, or surface water reservoirs)."

- ***We have revised the statement to provide a more complete explanation of the water balance equation as recommended.***

8) Page 18, Table 1: How do you interpret the potential reasons for why the performance metrics for equatorial, and warm temperate regions almost for all EXPs outperform those of other regions?

- ***Thank you for your insightful question. The better performance metrics in equatorial and warm temperate regions can be attributed to their wetter conditions, as most models were developed for humid and semi-humid regions, which supports their suitability in this study (as mentioned in lines 468-470 of the preprint).***
  Additionally, this may be due to effective land surface parameterization in these regions, which might not hold true for cold and arid areas. Lastly, the quality of ERA5-Land meteorological data may be limited in cold and arid regions, particularly in the absence of regional studies to confirm or refute this hypothesis.

References:

Beck, H. E., van Dijk, A. I. J. M., de Roo, A., Dutra, E., Fink, G., Orth, R., and Schellekens, J.: Global evaluation of runoff from 10 state-of-the-art hydrological models, Hydrol Earth Syst Sci, 21, 2881–2903, https://doi.org/10.5194/hess-21-2881-2017, 2017.

Decharme, B. and Colin, J.: Influence of Floodplains and Groundwater Dynamics on the Present-Day Climate simulated by the CNRM Model, EGUsphere, 2024, 1–33, https://doi.org/10.5194/egusphere-2024-3091, 2024.

Diks, C. G. H. and Vrugt, J. A.: Comparison of point forecast accuracy of model averaging methods in hydrologic applications, Stochastic Environmental Research and Risk Assessment, 24, 809–820, https://doi.org/10.1007/s00477-010-0378-z, 2010

Georgakakos, K. P., Seo, D.-J., Gupta, H., Schaake, J., and Butts, M. B.: Towards the characterization of streamflow simulation uncertainty through multimodel ensembles, J Hydrol (Amst), 298, 222–241, https://doi.org/https://doi.org/10.1016/j.jhydrol.2004.03.037, 2004.

Huo, W., Li, Z., Wang, J., Yao, C., Zhang, K., and Huang, Y.: Multiple hydrological models comparison and an improved Bayesian model averaging approach for ensemble prediction over semi-humid regions, Stochastic Environmental Research and Risk Assessment, 33, 217–238, https://doi.org/10.1007/s00477-018-1600-7, 2019.

Pires, C. and Martins, M. V: Enhancing Water Management: A Comparative Analysis of Time Series Prediction Models for Distributed Water Flow in Supply Networks, Water (Basel), 16, https://doi.org/10.3390/w16131827, 2024.

Reale, M., Giorgi, F., Solidoro, C., Di Biagio, V., Di Sante, F., Mariotti, L., Farneti, R., and Sannino, G.: The Regional Earth System Model RegCM-ES: Evaluation of the Mediterranean Climate and Marine Biogeochemistry, J Adv Model Earth Syst, 12, e2019MS001812, https://doi.org/https://doi.org/10.1029/2019MS001812, 2020.

Shoaib, M., Shamseldin, A. Y., Khan, S., Khan, M. M., Khan, Z. M., and Melville, B. W.: A wavelet based approach for combining the outputs of different rainfall–runoff models, Stochastic Environmental Research and Risk Assessment, 32, 155–168, https://doi.org/10.1007/s00477-016-1364-x, 2018.