General comments

This study presents the calibration and regionalisation of two land-surface hydrologic models in 263 catchments in the Western US. The results indicate that the median Kling-Gupta efficiency obtained in model calibration and regionalisation outperforms earlier/baseline study.

The study focuses on an interesting topic, but in its current form, it reads more like a technical report than a research paper. The Introduction nicely presents the context (i.e. why it is important to simulate water cycle components in the study region accurately). Still, the synthesis and formulation of the current research gaps need to be significantly improved. There is a large body of literature focusing on regional calibration of hydrologic models, transfer of model parameters at the regional scale, definition of the signatures used for similarity definition etc. It needs to be made clear how this study goes beyond the existing studies. The formulation of the research questions needs to be very precise and linked with presenting the research gaps. In its current form, it needs to be clarified whether the main goal is to propose and evaluate some methodological advance in model calibration and/or regionalisation or to present some new factual information about the study region (a case study analysis).

Thank you for your insightful comments. Streamflow forecasts are crucial for promoting sustainable water practices and building resilience to water-related challenges, and robust hydrological model simulations are at the core of streamflow forecasts. The calibration of parameters, although time-consuming and computationally expensive, is key to enhancing model performance. Our study acknowledges a gap in the availability of finely-tuned, high-resolution calibrated parameter sets for the Noah-MP and VIC models in the Western United States (WUS).

Addressing this gap, our study applies globally optimized calibration across 263 river basins in the WUS, at a fine resolution of 1/16 degree latitude-longitude. Utilizing the VIC and Noah-MP models, our approach extends beyond the existing scope of hydrological studies. We further regionalized these calibrated parameters to all 4816 HUC-10 basins in the WUS, developing high-resolution parameter sets. These sets are intended to bolster regional hydrological studies and climate change assessments, offering significant benefits for water resource management and environmental planning.

In response to your feedback, we will enhance our manuscript by elaborating on how our study fills an existing gap in globally calibrated hydrological model parameters at a fine spatial resolution and on the extensive spatial scope, which includes calibration at the above-mentioned 263 river basins in the WUS and the regionalization across 4816 WUS HUC-10 basins. Furthermore, our application of two widely used hydrological models, VIC and Noah-MP, introduces an additional layer of complexity and relevance. The use of these models allows us to address a broader range of uncertainties associated with hydrological modeling in varied climatic and geographic contexts.

The selection of the two models needs to be better justified. What are the differences in runoff generation between the models (and how is it linked with the regional variability of runoff generation in the study region)? It needs to be clarified why to use a 3hr simulation time step, when model inputs are daily. It is also not clear why to calibrate only selected soil-related parameters and how the selection is linked with the runoff generation processes and their variability in the study region. For example, are the snow accumulation and melt processes less important? Or are the snow-related model parameters already accurately calibrated? More importantly, the results and the differences between the two models need to be better linked with the main runoff generation processes (and their regional variability).

I missed the discussion of the results, which will link the new findings with previous studies. This can enhance the demonstration of the novel scientific contribution of the study.

Thank you for your comment. We answer your comments below:

(a) Why we selected these two models?

We chose to focus on the Variable Infiltration Capacity (VIC) model and the Noah-Multiparameterization (Noah-MP) LSM due to widespread previous application of these two models both in the U.S. and globally, as highlighted by Mendoza et al. (2015) and Tangdamrongsub (2023).

Our rationale for incorporating two distinct hydrological models lies in addressing the inherent variability and uncertainty in such simulations. By using two models, we aim to enhance the robustness of our study and better encompass structural uncertainties.

The VIC model is renowned for its widespread popularity and demonstrated success in simulating runoff on a global scale (e.g. 144 Adam et al 2003 & 2006; Livneh et al 2013; Schaperow et al 2021). Its established track record makes it an invaluable component of our analysis. The Noah-MP model is relatively newer, but is the hydrologic core of the National Water Model (NWM) which is being used increasingly domestically and internationally.

Further reinforcing our choice is a study by Cai et al. (2014), which evaluated the hydrologic performance of four LSMs in the contiguous United States using the North American Land Data Assimilation System (NLDAS) test bed. This study found that Noah-MP exhibited superior performance in soil moisture simulation and ranked highly in Total Water Storage (TWS) simulations. Conversely, the VIC model was distinguished for its excellence in streamflow simulations.

Our decision to utilize both the Noah-MP and VIC models is predicated on their proven effectiveness in simulating a wide range of hydrological processes. The unique runoff generation methodologies of each model are particularly pertinent for capturing the diverse hydrological characteristics of the WUS. This methodological diversity allows us to more comprehensively assess runoff generation mechanisms and their spatial variability within the region. We will revise the manuscript in section 2.2 Land Surface Models to address more on why we selected these two models.

(b) Differences in runoff generation between the models.

Noah-MP has four runoff physics options and after evaluation we decided that the free drainage exhibited the most substantial performance enhancement after calibration. As a result, we chose to continue using this option which is incorporated in the NWM. This runoff physics option is signified with infiltration-excess based surface runoff scheme and gravitational free-drainage subsurface runoff scheme [Schaake et al., 1996]. Noah-MP has four soil layers and each layer has a fixed depth (from top to bottom, 0.1m, 0.3m,0.6m,1.0m).

In VIC, each grid has up to three soil layers and the depth can be different for each grid cell. The infiltration into the top-most layers is controlled by variable infiltration capacity (VIC) parameterization (Liang et al., 1994). The flow is gravity-driven from upper layers to lower layers (Brooks and Corey, 1964). The function of the soil moisture in the third layer is linear below a soil moisture threshold and becomes nonlinear above that threshold. [Liang et al., 1994]. (could be seen as a combination of infiltration excess and saturation overland flow combination Liang and Xie (2001)). We will revise the manuscript to include these differences in the runoff generation between the models.

(c) How is it linked with the regional variability of runoff generation in the study region?

Both Noah-MP and VIC show good baseline performance along the Pacific Coast, in central to northern CA. Those areas have a high runoff ratio (specifically spring and annual runoff ratio) and high mean winter precipitation and mean annual max daily precipitation. These features might be favorable for runoff physics that have infiltration-excess mechanism, thus both VIC and Noah-MP perform well in these regions. VIC's baseline KGE generally is high in the inland northwest which has lower mean annual max daily precipitation and deeper groundwater table, VIC might be better to simulate these basins because it has varied soil moisture depth while Noah-MP has fixed soil moisture depth. Post-calibration improvements occurred for both models in most areas, especially in regions where the baseline KGE was low, such as southern CA and the southeastern part of the study region. We will revise our manuscript to include these above discussions.

(d) Why use a 3hr simulation time step, when model inputs are daily?

The choice of a 3-hour simulation time step, despite having daily model inputs, was intended to capture the diurnal cycle of energy balance and hydrological processes, which can be significant in regions with large variations in daily temperature and solar radiation. This finer temporal resolution aids in better representing the hydrological response and energy dynamics, especially in snow-dominated catchments. Besides that, we did an analysis on the timestep of Noah-MP and found that at least of 3-hour timestep is needed to generate robust simulations. Although VIC can be conducted at daily time scale, to make the two models comparable, we run VIC simulations at the same 3-hour timestep. We will revise our manuscript in the model section to address this comment.

(e) why calibrate only selected soil-related parameters and how the selection is linked with the runoff generation processes and their variability in the study region?

Our focus on calibrating soil-related parameters was based on their critical role in runoff generation. We aimed to address the key processes such as infiltration, soil moisture storage, and groundwater recharge, which are pivotal in the WUS's diverse hydroclimatic settings. The calibration of these parameters was prioritized to improve the representation of soil-water interactions, a major driver of runoff variability in the region. Concerning snow accumulation and melting processes, we acknowledge their importance. Before we conducted the calibration, we conducted snow simulation verification at 20 selected snotels across WUS and our assessment indicated that the existing parameterization for snow processes in both models was adequate for our study region. We will revise our manuscript in the calibration section to clarify this comment.

(f) Linking Results with Runoff Generation Processes and Previous Studies.

Analysis of Model Differences and Runoff Processes: We will enhance our results section to include a more detailed analysis of how each model's characteristics influence the simulation of runoff generation processes. This will include an examination of regional differences in runoff generation captured by each model, thereby providing insights into their respective strengths and limitations in representing the hydrological dynamics of the WUS.

Contextualizing Findings with Previous Research: To demonstrate the novel scientific contribution of our study, we will expand our discussion to more explicitly link our new findings with previous studies. This will include comparing our results with existing literature on hydrological modeling in the WUS, particularly focusing on how our approach contributes to a deeper understanding of regional runoff variability and model performance.

We appreciate the opportunity to refine our manuscript based on your feedback, and we believe these revisions will significantly enhance the clarity and impact of our research.

References:

Adam, J.C. and Lettenmaier, D.P.: Adjustment of global gridded precipitation for systematic bias, J. Geophys. Res., 108(D9), 1-14, doi:10.1029/2002JD002499, 2003.

Adam, J.C., Clark, E.A. , Lettenmaier, D.P. and Wood, E.F.: Correction of Global Precipitation Products for Orographic Effects, J. Clim., 19(1), 15-38, doi:10.1175/JCLI3604.1, 2006.

Brooks, R.H. and Corey, A.T. (1964) Hydraulic Properties of Porous Media. Hydrology Paper, Vol. 3, Colorado State University, Fort Collins.

Cai, X., Yang, Z.L., Xia, Y., Huang, M., Wei, H., Leung, L.R. and Ek, M.B., 2014. Assessment of simulated water balance from Noah, Noah-MP, CLM, and VIC over CONUS using the NLDAS test bed. Journal of Geophysical Research: Atmospheres, 119(24), pp.13-751.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges, 1994: A Simple hydrologically Based Model of Land Surface Water and Energy Fluxes for GSMs, J. Geophys. Res., 99(D7), 14,415-14,428.

Liang, X. and Xie, Z., 2001. A new surface runoff parameterization with subgrid-scale soil heterogeneity for land surface models. Advances in Water Resources, 24(9-10), pp.1173-1193.

Livneh B, Rosenberg, E.A. , Lin, C. , Nijssen, B. , Mishra, V. , Andreadis, K., Maurer, E.P. and Lettenmaier, D.P.: A long-term hydrologically based data set of land surface fluxes and states for the conterminous United States: Updates and extensions, Journal of Climate, doi:10.1175/JCLI-D-12-00508.1, 2013.

Mendoza, P.A., Clark, M.P., Mizukami, N., Newman, A.J., Barlage, M., Gutmann, E.D., Rasmussen, R.M., Rajagopalan, B., Brekke, L.D. and Arnold, J.R.: Effects of hydrologic model choice and calibration on the portrayal of climate change impacts. Journal of Hydrometeorology, 16(2), 762-780, 2015.

NOAA (National Oceanic and Atmospheric Administration): National Water Model: Improving NOAA's Water Prediction Services, 2016.

Schaperow J.R, Li, D., Margulis, S.A., Lettenmaier D.P. :A near-global, high resolution land surface parameter dataset for the variable infiltration capacity model. Scientific Data. Aug 11;8(1):216, 2021.

Tangdamrongsub, N.: Comparative Analysis of Global Terrestrial Water Storage Simulations: Assessing CABLE, Noah-MP, PCR-GLOBWB, and GLDAS Performances during the GRACE and GRACE-FO Era. Water, 15(13), p.2456, 2023.