Thank you for allowing me the opportunity to review your paper. From my understanding, these are the main points of your manuscript:

- You created SR models to created better lumped predictions by connecting LSTM predictions to routing modules.
- There is an option of including/excluding reservoirs.
- The model is trained exclusively on basins within the GL region.
- SR models perform better when performing Spatial validation, while a lumped LSTM does better at the temporal validation.

This manuscript is grammatically well-written and conveys a good message. I recommend the paper is accepted with minor revisions. I have listed Major and Minor comments below, but I believe the major comments are easy to address and do not require a change in the experiment. My biggest issue is with the novelty in the paper (see major comment 4).

Best,

Tadd Bindas

Major Comments:

1. There is a main paragraph explaining the importance of large-scale training of LSTMs and how individual catchment LSTMs are not suitable for comparison. However, it seems you're still limiting your data to only regional data from the Great Lakes region. While this is better than training an LSTM on one basin, it is best to give the model *all* of the data you can, including basins outside of the GL region.

   Have you trained your SR, or lumped, LSTM on all CONUS basins, or done any work regarding that?

2. Figure 1 is confusing as the numbering doesn't align with how the graph is read. It reads like it's circular, everything is somehow mapping to itself. I suggest redoing this figure as it would be easier for the reader to understand the SR workflow.

3. What loss function was used? While it makes sense to have a lot of your information in a supplemental paper, adding this information would be helpful to readers as they would know what information is used in training your lumped LSTM.

4. I'm unsure of the novelty here. After reading the paper, and the novelty statement, multiple times. I'm not convinced that what you're doing hasn't already been done. It appears you're only routing streamflow from a model created in Mai et al. (2022).

   I hate to share/recommend reading my own research as it can come off as scummy, but I have an accepted manuscript that has a very similar workflow what you're doing here (Training a LSTM that was created based on another paper's work, applying the LSTM to smaller catchments, routing those outputs, and comparing the routed flows against the

lumped predictions. My paper's novelty is that our Muskingum-Cunge routing is a differentiable model). The reason I'm sharing my work is because you may not be aware that I had used a very similar set up in my research. The paper has been under review for ~15 months and only available in preprint. See below for the Title and DOI.

Improving large-basin river routing using a differentiable Muskingum-Cunge model and physics-informed machine learning
10.22541/essoar.168500246.67971832/v1

To be fair to your work, I do think that your paper presents an angle that has not been done yet, but is not expressed in the novelty statement:
"The distinctive aspect of our research lies in the combination of components to form the SR model. The following subsections explain the specifics of the components we selected in this study to demonstrate the spatially recursive modelling approach."
I believe your work's novelty is both the scale (as my accepted work is done at the catchment scale, *not at the regional scale),* and my work does not have a reservoir module included in the routing. This will ensure your paper correctly identifies what is an expansion of domain knowledge.

Minor Comments:

1. Lines 107-108: This sentence reads a little weird. I suggest changing to: "Han et al. (2020) and Mizukami et al. (2016) include examples of hydrologic routing models."
2. Can you make the font bigger for figure 5 axis?