**Review EGUSPHERE-2026-2950**

**TITLE**

Technical note: Regional fine-tuning of LSTMs for improved streamflow predictions in ungauged catchments

**RECOMMENDATION**

Moderate-major revision

**REVIEWER**

John Quilty

**NOTE(S)**

I use the following notation to identify text of interest in the manuscript: L12-15 (lines 12-15).

**ASSESSMENT**

This paper proposes Regionalized Fine-Tuning (ReFT) as a method for fine-tuning a global LSTM to a target catchment, using donor catchments selected by inverse-distance weighting. ReFT uses a weighted loss function based on this inverse-distance weighting, selecting a fraction (5%) of the "closest" catchments as donors. The authors test their approach using 218 catchments from CAMELS-AUS against a globally trained LSTM and two other conceptual benchmarks. The authors consider two fine-tuning scenarios: all parameters and only those from the output head.

The paper would benefit from, among other things: coverage of recent literature on the same topic; additional elaboration to improve reproducibility; justification of fine-tuning strategies (all parameters vs. output head only). Further, some results require additional explanation (e.g., polarizing results of full parameter fine-tuning).

The authors discuss the spatial concentration of catchments in CAMELS-AUS and how the results may not show the generality of the proposed method to other scenarios where catchments are more spatially distributed (e.g., CAMELS-USA). This calls into question whether the adopted dataset is the right one for evaluating the proposed method. Given this is a technical note, if the editor agrees, this may not be seen as a major hindrance to the paper's suitability for this journal.

Overall, I think the paper has merit as a technical note and l look forward to seeing a revised version of the paper.

Best regards,

John Quilty

**MAJOR ITEMS**

1. The authors should ensure they include recent papers in hydrology on fine-tuning DL models, contrasting earlier methods with their approach. Why is the presented approach necessary, what problems with existing approaches does it overcome?

2. LSTM architecture: is this the right architecture for your dataset? I understand the adopted architecture is from earlier work and was selected by the original authors after much experimentation on CAMELS USA (531 catchments). However, that does not mean the architecture is necessarily the best suited for CAMELS-AUS (218 catchments). What experimentation was done to confirm the suitability of this architecture for CAMELS-AUS compared to potentially "lighter" architectures (fewer hidden states) or shorter sequences (e.g., 30, 180)?

3. What optimizer was used? How many epochs was the model trained for? More details on model development (global LSTM and fine-tuned models) would improve reproducibility.

4. Why use spatial separation alone instead of the "closeness" (e.g., cosine similarity or another appropriate distance metric) to a group of static features (which could include lat-lon-elevation)? For example, you could use the cosine similarity based on the global LSTM's embedding vector (hidden state) to estimate closeness with a target catchment and its donors (e.g., Jahangir et al., 2026; https://doi.org/10.1016/j.envsoft.2026.106978). This could be interesting to check as it may validate your distance-based approach.

5. How many epochs (or iterations) were used to train catchment-specific models using the ReFT loss? How did you check/evaluate overfitting?

6. Justify the fine-tuning configurations. Why these instead of other possibilities? Are there precedents that informed this selection? If so, they should be cited.

7. L115-116: Please describe the notion of nested catchments in more detail along with the issue of data leakage.

8. Please provide more details on the polarizing results when using ReFT to fine-tune all LSTM parameters.

9. As you discuss, the spatial concentration of catchments in CAMELS-AUS may limit the generality of the findings. Why not apply the method to CAMELS-USA instead?

**MINOR ITEMS**

1. Check verb tense (e.g., are vs. were)
2. L147: check period placement.