**Comments by JD in *blue*, our responses in black.**

*The concept of summation vs. unit hydrograph*

*By the title alone of their opinion paper, I admire the frankness of the authors admitting the shortcomings of the LSTM networks when applied to a SINGLE basin.*

**Response:** This is not a correct or meaningful interpretation of the message of this manuscript. The paper does not say that you should not use an LSTM for a single basin, the paper says that if you only want to model a single basin, you should still train the LSTM correctly (i.e., on data from multiple basins). This is not a shortcoming in the sense that we are not illustrating any situation where the model should not be used, we are providing insight into the proper way to use the model.

*In the LSTM, the form of the conversion function for the hidden gate h_t, is tanh, e. g., Kratzert et al. (2018, Equation 7). As observed previously by me, this is similar in shape to a summation or S-curve hydrograph in unit hydrograph theory, e.g., Lees (2022, CC1 and CC2 therein).*

*For the hidden gate, I suggest the authors consider taking one further step of using the first derivative of the conversion function. This is equivalent to using the form of an instantaneous unit hydrograph or impulse response function in convolution integral, e.g., Ding (1974). This, I believe, will inject some hydrologic realism into the LSTM.*

*The bottom half of Figure 4 for Buffalo Fork, Wyoming (USGS Gage 13011900) for the single LSTM basin model clearly indicates that an impulse response model having a distinct peak time and magnitude characteristic, e.g., Jeong and Kim (2023, Figure 2), would outperform the LSTM as now configured.*

*A reconfigured LSTM as suggested above may perform as well as, if not better than, the impulse response ones.*

**Response:** The purpose of this manuscript is not to investigate different modifications to the LSTM architecture but rather to make the community aware of the specific training setup that is required when training these (somewhat popular) models. The review has made this same (or similar) comment to previous LSTM publications in open review in HESS-D, and we suggest that if the reviewer is interested in this topic that they could try the experiment themselves. This type of experiment is out-of-scope for the current publication. This could be done, for example, starting from the LSTM implementation in https://github.com/neuralhydrology/neuralhydrology/blob/master/neuralhydrology/modelzoo/customlstm.py.