

Hydrol. Earth Syst. Sci. Discuss., referee comment RC1
<https://doi.org/10.5194/hess-2021-515-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on hess-2021-515

Ralf Loritz (Referee)

Referee comment on "Technical note: Data assimilation and autoregression for using near-real-time streamflow observations in long short-term memory networks" by Grey S. Nearing et al., Hydrol. Earth Syst. Sci. Discuss.,
<https://doi.org/10.5194/hess-2021-515-RC1>, 2021

Review of "Data assimilation and autoregression for using near-real-time streamflow observations in long short-term memory networks" by Nearing et al. (in discussion)

Summary and Recommendation

Nearing et al. test how near-real time streamflow observations can effectively be used in Long Short-Term Memory (LSTM) rainfall runoff models. They compare an autoregression (AR) approach with a data assimilation (DA) approach and test, additionally, how sensitive AR is to random gaps in the data. The manuscript (MS) is easy to follow, well-structured and suits the scope of HESS. I particular liked the comprehensive appendix in combination with a short MS. I think that this MS can be published after some minor revisions and provide only some smaller comments and questions below.

Sincerely

Ralf Loritz

Questions and comments:

- Reading the MS I would have like to see a couple of detailed results from three or four catchments where the AR or DA worked particular good or bad and what the

(hydrological and ML) reason for this might be (to underpin the discussion of Appendix F and G). For instance, what could be the reason that DA and AR reduces the predictive performance of a few of your models (Fig. F2)? You state that: "*We are unsure of the reason for this, but it warrants further exploration.*" (Line 352) maybe zooming into one of the catchments could help to give a better explanation.

- I find it a bit unrealistic how you added the missing data. I would assume that a broken gauging station is not working for a couple of days or maybe weeks in a row and wonder how this would alter your results (e.g. all streamflow data available for training but then two weeks or more only simulated data during testing with a closer focus on particular that period and not the entire testing period).
- Showing how the variance or the Shannon entropy changes of your simulations in addition to the median would be interesting (Fig.1 and Fig.3). If it remains constant, I would mention that the spread of the predictions is not affected by the data availability.

Personal comment: Three of the **seven** Co-Authors have presumably not contributed to this "technical note" as they are not mentioned in the author contribution section.