

Hydrol. Earth Syst. Sci. Discuss., referee comment RC1
<https://doi.org/10.5194/hess-2021-127-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on hess-2021-127

Anonymous Referee #1

Referee comment on "Benchmarking data-driven rainfall–runoff models in Great Britain: a comparison of long short-term memory (LSTM)-based models with four lumped conceptual models" by Thomas Lees et al., Hydrol. Earth Syst. Sci. Discuss.,
<https://doi.org/10.5194/hess-2021-127-RC1>, 2021

The manuscript outlines an application of LSTM-based runoff models, which were introduced in previous studies (Kratzert et al, 2018; 2019). In the present contribution, the focus of analysis is catchments in Great Britain. Similar to previous studies, the objective is to demonstrate the competitive ability of LSTM in rainfall-runoff simulations over traditional process-based models. The authors made considerable efforts to set up experiments and perform relevant analyses. Results are compared with four lumped conceptual models and show that the LSTM models outperform the traditional models as well when applied in Great Britain.

The manuscript is generally well written and organized, figures and tables support the results.

My main concern is the degree of innovation and scientific significance of this work compared to already published works. This is a critical aspect of the manuscript that should be improved.

A large section of the manuscript is dedicated to a discussion of the advantages reported in the previously developed LSTM model. This discussion focuses on predictive ability, without much methodological improvement and innovations in ideas, that in turn may impair the scientific importance of the research.

In recent years, LSTM models have been broadly assessed. Most of these studies indicate the generally better performance of LSTM models over lumped models. The results reported in this manuscript seem to confirm the previously reported conclusions. By comparison, the analysis in Sections 4.2 and 4.3 is limited, whereas IMHO this is the most insightful section of the paper which deserves additional in-depth discussion. I think the

authors should dedicate more space to discuss the implication of their findings.

Below are more detailed comments, questions, and suggestions that hopefully initiate a fruitful discussion and help improve the paper.

ABSTRACT: I would suggest mentioning the challenges in present LSTM applications for hydrological modeling and what is to be addressed, otherwise, it is difficult to tell the significance and necessity of the work.

INTRODUCTION: I do not think research gaps are well defined in the introduction. The research objectives should be motivated by the research gaps. The latter two of the three questions raised in the manuscript are related to overcoming limitations and model diagnosis, without indication of the explicit research gaps to be addressed. Are there some additional studies that investigate the correlation between LSTM model performance and catchment attributes?

The background should be more concise and emphasizes more about what is still to be investigated regarding the usage of LSTM models.

Furthermore, LSTM is but one of several machine learning frameworks used in rainfall-runoff modelling. Recent advances in evolutionary computation report theory guided and "hydrological informed" approaches that result in not only highly accurate but also readily interpretable models. See for example:

J Chadalawada, et al, 2020, Hydrologically Informed Machine Learning for Rainfall-Runoff Modeling: A Genetic Programming-Based Toolkit for Automatic Model Induction, Water Resources Research 56 (4), e2019WR026933

HMVV Herath, 2020, Hydrologically Informed Machine Learning for Rainfall-Runoff Modelling: Towards Distributed Modelling, Hydrology and Earth System Sciences Discussions, 1-42

Line 77: It seems only THREE research questions are being proposed.

METHODS:

Section 2.3: It is more suitable to use the term "layer" (e.g., LSTM layer and EA LSTM layer) when describing the specific layer structure.

Line 158: Please keep consistent notation using curly quotes or straight quotes throughout the manuscript.

Figure 2: In EA LSTM cell, is the input gate " i_t " or " i "? (see Equation 8)

Lines 203-206: The fully connected layer should be a part of the model architecture. It seems strange to introduce them in this subsection (model training).

Section 2.5.1: A brief description of the process-based models is required, especially what hydrological processes are included in the respective models because the discussion section involves the consideration of processes.

RESULTS AND DISCUSSION

Table 3, Figure 3, Figure 4, Figure 6, and Figure 7: All the results seem to merely be used to show the outperformance of LSTM models than other models in various cases. I think this part should be more concise if the result is not out of expectations, and more other implications should be discussed from the results.

Lines 496-503: The speculation of "connectivity" is interesting, while how the connectivity can be "learned" by LSTM models should be clarified, say whether the connectivity can be represented by hidden information within data or the model architecture (such as the memory of LSTM).

Lines 505-507: A simple strategy to examine the speculation is to train an LSTM model with/without `crop_perc` included for checking its role in improving the representation of hydrology in those catchments with a strong agricultural signal.