Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-247-RC2, 2018

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Rainfall-Runoff modelling using Long-Short-Term-Memory (LSTM) networks" *by* Frederik Kratzert et al.

Anonymous Referee #2

Received and published: 21 July 2018

General remarks

Artificial neural networks (ANN) enjoyed great popularity in the late 1990s and – as other data driven modeling techniques – are now part of the standard toolbox in rainfall-runoff modeling. Thus, it is surprising enough, that a limited number of studies can be found in the hydrologic literature which are applying the latest developments of the artificial intelligence research, such as e.g. deep learning.

This paper provides a first step into this direction and introduces Long-Short-Term-Memory (LSTM) networks for the task of rainfall-runoff modeling. In a comprehensive

comparative study the proposed method is applied to the CAMELS data set and is compared with the conceptional SAC-SMA model which was complemented by the Snow-17 routine. The study comprises 3 numerical experiments starting with the application to single catchments and ending with the test of potential applications for ungauged catchments using a regionalisation approach.

The paper is reasonably well written and a novel contribution for assessing the predictive performance of LSTM networks in rainfall-runoff modeling.This makes the study very interesting for scientists who did not use a LSTM networks before. Since it is a first application, the paper should describe more systematically the training procedure and characteristics of the LSTM network which in the present version turned out to be more art than science. In addition and although I am enthusiastic about the work, I think a balanced discussion of the new approach should also include limitations, especially in the "Summary and conclusion" chapter. I encourage the authors to make following major modifications as they prepare their manuscript for revision:

- Please check carefully the recent literature for applications of deep learning in water resources and discuss those, there are more than cited, e.g. ().
- I have concerns about the reproducibility of the performance of the LSTM network since the training is done by trial and error and it is not very systematically evaluated. But it is an important issue, because the number of free parameters od the LSTM network is huge and as I understand a gradient-based error back-propagation method is used for training. As a reference for the state of the art evaluation of data driven models I recommend () where a stochastic procedure, involving random sampling for training, cross-validation, and testing, is proposed.
- Finally, more information and discussion about limitations of the new approach would be helpful, e.g. the computational effort, extrapolation behavior, performance for extreme events (floods) etc.

Minor remarks

page 4,Eq. 1 U_f is not correct.

page 4 Give an equation for the calculations of the dense layer.

page 5, Fig. 2 Add bias b. Why c is capital letter?

page 5 Please give the reference on which the theory is based when starting with the description of the LSTM network – around Eq. 2.

page 6 l. 17 "For this study, we used a 2-layer LSTM network, with each layer having a cell/hidden state length of 20." First, I would split the theory and the setup of the LSTM for the numerical experiment. So move all the specific details to section 2.4. In addition, I would expect a table with all the specifications of the used LSTM including number of the parameters in W_c, W_f, W_i, W_o, U_c, U_f, U_i, U_o, b_c, b_f, b_i, b_o and hyperparameters. Second, I do not understand that the LSTM has a number of 365 inputs and the "hidden state length of 20". Please explain this!

page 6 I would skip section 2.1.1 or move this to the discussion since this is hypothetical and no mathematical equivalence is shown.

page 7 l. 10 Is the LSTM limited to MSE when backpropagation is used?

page 7 l. 19 spelling->"iteration"

page 11 Please give more information about the calibration od the SAC-SMA model and the computational effort.

page 13 Explain, why the LSTM network is better for the mean, but not for the median NSE (see Fig.6b). From my point of view, it is not surprising that the LSTM

network performance better for mean flows. So discuss in detail also the behavior for high flows.

page 15 "However, we want to highlight again that achieving the best model performance possible was not the aim of this study, rather testing the general ability of the LSTM in reproducing runoff processes."<-Since we already know that data driven techniques are able to reproduce runoff processes, the authors of the paper should be more ambitious and give some more details and discussion about advantages and disadvantages of the LSTM network.

page 21 I would skip Fig. 21 or would present a more detailed analysis of internal states and combine this with the hypothesis described in section 2.1.1.

References

Duo Zhang, Erlend Skullestad Hølland, Geir Lindholm, Harsha Ratnaweera, Hydraulic modeling and deep learning based flow forecasting for optimizing inter catchment wastewater transfer, Journal of Hydrology, 2017, DOI:10.1016/j.jhydrol.2017.11.029.

Elshorbagy, A., Corzo, G., Srinivasulu, S., and Solomatine, D. P.: Experimental investigation of the predictive capabilities of data driven modeling techniques in hydrology - Part 1: Concepts and methodology, Hydrol. Earth Syst. Sci., 14, 1931-1941, https://doi.org/10.5194/hess-14-1931-2010, 2010.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-247, 2018.