Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-552-RC2, 2020

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Groundwater Level Forecasting with Artificial Neural Networks: A Comparison of LSTM, CNN and NARX" *by* Andreas Wunsch et al.

Anonymous Referee #2

Received and published: 24 December 2020

The authors perform a comparison between NARX, CNN and LSTM on seq to val and seq to seq mode, over a set of wells. The authors investigate not only the performance of the models but also the computational effort required to calibrate them and the effect of the training length which are interesting and useful aspects. Another interesting and novel aspect is the combined approach to hyper parameters tuning and variable selection. The work is well written and exhaustive, making the work reproducible. The set of experiments was properly designed and explained. Therefore, I recommend only a few minor changes.

It could be interesting and useful to show a map of the study area with the wells loca-

tions and the mentioned surface water bodies.

In the introduction it could be useful to point out as a novelty aspect the approach to hyper parameters tuning and variables selection. Several other works use statistical methods to determine the input sequence length, which could have a hydrological meaning (Kisi et al. 2017; Hasda et al. 2020; Zanotti et al. 2019; Di Nunno 2020). In this case results (Tables S2 – S4) show a wide variability of the length of the input. This does not give any insight about the hydrological behaviour of the water bodies, but it could be useful in cases where the correlation is not linear.

In paragraph 2.5 could be better explained the relationship between the input delay, feedback delay and the additional GWt-1 data. Since NARX is autoregressive isn't it already considering previous groundwater levels? In table S2 you have ID GWLt-1: does it mean that you are feeding into the model more than one GWL (and the same for seq length when using GWLt-1)?

The performance of the models is well presented and discussed, but a discussion could be made also relatively to very poor performance on some wells: what could be the cause of the results of e.g. BW 781-304-2 or BW 138-019-0? Maybe it could be handy to add the length of the training set in figures S1-S68 or in their captions.

Fig. 4 and its relative discussion are in the results; it could be useful to mention in the materials and methods that you performed that analysis. Same for paragraph 4.4: it is an interesting analysis, and its methodology should be appropriately explained in the methods section and anticipated in the introduction.

Line 50-51 is not clear

Di Nunno, F., Granata, F., 2020. Groundwater level prediction in Apulia region (Southern Italy) using NARX neural network. Environ. Res. https://doi.org/10.1016/j.envres.2020.110062 Hasda, R., Rahaman, M.F., Jahan, C.S., Molla, K.I., Mazumder, Q.H., 2020. Climatic data analysis for groundwater level simulation in drought prone Barind Tract, Bangladesh: Modelling approach using artificial neural network. Groundw. Sustain. Dev. https://doi.org/10.1016/j.gsd.2020.100361 Kisi, O., Alizamir, M., Zounemat-Kermani, M., 2017. Modeling groundwater fluctuations by three different evolutionary neural network techniques using hydroclimatic data. Nat. hazards 87, 367–381. Zanotti, C., Rotiroti, M., Sterlacchini, S., Cappellini, G., Fumagalli, L., Stefania, G.A., Nannucci, M.S., Leoni, B., Bonomi, T., 2019. Choosing between linear and nonlinear models and avoiding overfitting for short and long term groundwater level forecasting in a linear system. J. Hydrol. 578, 124015.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-552, 2020.