Hydrol. Earth Syst. Sci. Discuss., 12, C4817–C4820, 2015
www.hydrol-earth-syst-sci-discuss.net/12/C4817/2015/

Hydrology and Earth System Sciences
Discussions

Open Access

HESSD

12, C4817–C4820, 2015

Interactive Comment

Interactive comment on "Technical Note: Application of artificial neural networks in groundwater table forecasting – a case study in Singapore swamp forest" *by* Y. Sun et al.

Y. Sun et al.

tmssy@nus.edu.sg

Received and published: 11 November 2015

Thanks for the comments on our paper.

It's fully acknowledged that application of ANN in hydrology research, more specifically in groundwater table modelling, has been a popular research topic. Two paragraphs are devoted to elaborating the most recent research and findings (please refer to lines 4 to 28 in page 9319 and lines 1 and 2 in page 9320).

Our study applied ANN to forecast the highly responsive groundwater table in a fresh water swamp forest, and the methodology in our paper is straightforward and easy to implement, which only uses the rainfall and reservoir levels as the ANN inputs and the lagged observed groundwater tables as the output. In this context, our paper is the first of this kind and is instructive for similar research in future. We think this paper meets the standard to be published as a 'technical note' in HESS.

In response to the 'comments', please find our reply embedded as follows:

1. The Authors say the used reservoir levels and rainfall as input to the ANN. It is not clear if the used lagged data or data at the step before the output.

R: The reservoir levels and rainfall are fed to the networks as input, while the output is the lagged observed groundwater tables after 1 day, 3 days and 7 days. We will revise accordingly in the manuscript (please refer to lines 9 to 12 in page 9324).

2. It is not clear how the Authors assumed the architecture of the network and how they chose the input.

R: The reservoir levels and rainfall are chosen as the inputs as they are the major water source and driving force for the regional groundwater (please refer to lines 4 to 7 in page 9320 and lines 9 to 12 in page 9324).

A single hidden layer MLP is selected due to the universal approximation theorem (please refer to lines 15 to 18 in page 9322), whereas the number of hidden neurons is determined by trial and error (please refer to lines 12 to 15 in page 9324).

3. Training data set seems to be too limited. I wonder if this may cause overfitting problems, as it seems to be.

R: An entire year's data are selected as the training data, with the objective to train the network in a more robust manner being exposed to the seasonal cycle (please refer to lines 20 to 23 in page 9324).

4. Looking at figures 3 and 5 as well as to table 1, it seems that there is an immediate decay of fitness, when the prediction is pushed at 3 and 7 days ahead. This may be

HESSD

12, C4817–C4820, 2015

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

related to overfitting problems or to a bad selections of the input.

R: Judging from Table 1, when move from 1 day to 3 and 7 days prediction, the performance of ANN does decay, but not to a drastic extent, e.g. at P1 from 5.4 to 8.2 and 9.9, at P3 from 5.2 to 6.6 and 8.6, etc. Therefore, the overfitting problems may not be dominating in our study case and we've also used cross validation data trying to avoid overfitting (please refer to lines 24 to 26 in page 9324 and lines 1 to 2 in page 9325).

5. The ANNs fail at reproducing peaks and dry periods, in particular for 3 and 7 days ahead prediction. Again, this seems to be related to an improper choice of the input or to a lack of information content of the input.

R: The peaks, especially at P4, are not perfectly captured because of the missing information of spillway discharge. The dry periods are not well predicted because such a drought condition does not exist in the training data. These reasons have been clearly stated in 'Results and discussion' (please refer to lines 8 to 26 in page 9325 and lines 1 and 2 in page 9326).

6. It is not clear if the Authors compared their model with a simpler one, i.e. linear models, like ARX or ARMAX. Maybe, these models may have similar performances with the proposed ANNs.

R: ANN is chosen mainly due to its ability in regression analysis and the usage of more accessible variables in mapping the input-output relationships. Its applicability has also been widely verified in many fields of hydrology research. Due to the complicated topography, geological characteristics and hydrological processes, the relationship between the input (reservoir level, rainfall) and the output (groundwater table) is not linear. Hence, linear regression model is not suitable to serve our study purpose (Please refer to lines 14 to 17 in page 9319).

In addition, for our study area – the NSSF, a numerical model has been setup (with Mike SHE) in order to simulate local hydrological conditions. ANN actually outperforms

HESSD
12, C4817–C4820, 2015

the numerical model in terms of forecasting accuracy at the piezometer locations but with a poorer spatial coverage. In consideration of the limited article length, either the linear regression model or the numerical model is further discussed in the paper.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 12, 9317, 2015.