Hydrology and Earth System Sciences Discussions

Interactive comment on "Real-time flood forecasting by employing artificial neural network based model with zoning matching approach" *by* M. Sulaiman et al.

Anonymous Referee #3

Received and published: 5 January 2012

The paper presents an attempt to use artificial neural networks as a flood forecasting tool. It proves, that the zone matching approach is essential for an adequate modelling of especially high water levels. Generally, the article could be a bit clearer and more focussed. The performance of the model is excellent. But the paper lacks clear descriptions of the conditions, which makes it challenging to assess the quality of the model. Generally, it would be desirable to extend the description of the conditions at the site, to contrast the target data with the input data and explain some assumptions. On the other hand it seems not to be necessary to present and compare such a big number of trained ANN using such a big number of performance measures.

Major Comments: 1.) The description of the study area should be more precise. Is it right, that the upstream gauge (input data) is Rantau Panjang (fig. 3) and the target time series are from Kota Tinggi? The catchment size of 2636 km^2 refers to any of these stations? For the evaluation of the model performance it would be appropriate to indicate both areas. The water levels at the target gauge result from both the discharge at the upstream gauge and the precipitation at / runoff from the area between the two gauges. The runoff from this area is certainly not independent from the runoff in the upper catchment. But however - the size of this 'unobserved' catchment (in the sense of not been represented by any input data) is quite interesting for a valuation of the model.

2.) Regarding the modelled system, it seems more adequate to deal with discharges. The riverbed geometry might change (within some 50 years!). And so might change the water level resulting from a certain discharge. What is the reason for modelling the water levels rather than the discharges?

3.) Is it necessary to use that many performance measures? The authors don't use all of them for interpreting the results and the model performance.

4.) To illustrate the issue and to get the modelling problem more imaginable, the referee would appreciate a plot of time series of a a single event (like fig. 10) showing both input and target data. As ANN are purely empirical models they are only valid within the range of the training data. And, unfortunately, extreme flood events are rare. To both guarantee a sufficient database and restrict the training/validation data to only flood events, the choice of the lower limit of the ZMA of 8000 mm seems very reasonable (regarding fig. 5). But, to my opinion, for an application of the model the distinction has to be made for the input data. In a word: The authors should refer more to the model inputs.

5.) It would be desirable to make clearer the choice of the input variables (time lag to the forecast value), as this is somehow the heart of the approach. Please, specify

explicitly the inputs of your 'best' ANN!

6.) For an application by the authorities, as suggested in the conclusions section, the authors should suggest a range of validity of the model.

Technical comments:

1.) Which software was used for the setup of the ANN?

2.) fig. 3: a bar plot would be more appropriate.

3.) 3.6 and fig. 6: Is the time step always 3h? Has the forecast time to be equal to the time step? If it is so, then why?

4.) 3.1 line 13-14: The neurons do not receive the weights from the adjacent layer! line 23: The bias does not stabilize the output between 0 and 1. The sigmoid transfer function does! next page, eq. 3: The water levels are not a function of the weights; the weights are parameters of this function. Further, it would be preferable to distinguish somehow between WL as a target value and WL from the upstream gauge. line 20: The weights aren't initialized randomly to speed the training process up!

5.) In the conclusions the authors refer to DNN. It doesn't get clear why this should be advantageous. But maybe it is not necessary to introduce this type of ANN in this chapter, if if was not applied in the paper.

—————————————————

C5610