Hydrol. Earth Syst. Sci. Discuss., 6, C144–C147, 2009
www.hydrol-earth-syst-sci-discuss.net/6/C144/2009/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Classification of hydro-meteorological conditions and multiple artificial neural networks for streamflow forecasting" *by* E. Toth

Anonymous Referee #2

Received and published: 25 March 2009

General comments

Toth has submitted an interesting paper that uses self organising maps (SOM) to cluster antecedent river and climatic conditions. The clustered data is then used for training of an ANN for the prediction of discharge data. The such trained ANN result in improved performance for predicting discharge during the validation period compared to a single ANN trained with the complete data set.

The work fits well in the scope of HESS. The work is well presented, the presented material is sufficient and overall the manuscript is well structured and clear to understand. The conclusions are based on the presented material. Title and abstract clearly

reflect and summarize the work. Related work is cited sufficiently, however, the new contribution of the work could be pointed out clearer. Please also make clearer that the reported improvement is due to the new approach (see also specific comment 1).

Specific comments

1. Increasing the complexity of the model often improves prediction results. From the study it is not entierly clear whether improved perfomance is due to the increased number of ANN (4 compared to 1 in the reference case) or whether the use of SOMs for data pre-clustering is the reason for the improved performance. Could you address the question whether you expect comparable results to be achieve using some other partitioning method in the discussion of your results?

2. Please point out more clearly how the method presented in manuscript differs from other rainfall-runoff modelling studies using SOMs (as reviewed by Kalteh (2008)). This will help to better understand the innovation resulting from the presented work.

3. p900L22ff Please also mention the purpose of the model for the Sieve River. If the goal is high flow prediction, then the presented study does not substantially improve predictions compared to the global model presented later on. The Nash Sutcliffe coefficient of efficiency is mainly influenced by high flow periods, and only minor improvements were achieved in terms of this measure.

4. The introduction reads as if the term "system theoretic" is interchangable with data driven? Wheater et al. (1993) seems to be the relevant reference, which unfortunatelly is not easily available to me. I do not agree with this interchangability of terms. It is my understanding, that system theory starts from a system understanding, defining system boundaries, system components and their interactions. Maybe you could clarify on your view of system theory.

HESSD

6, C144–C147, 2009

5. p902L9 Use of equations could make the following easier to understand: “In correspondence of a computational node, each one of the entering values is multiplied by a connection weight. Such products are then all summed with a neuronspecific parameter, called bias, used to scale the sum of products into a useful range. The computational node finally applies an activation function to the above sum producing the node output.”

6. p903L9 Please give a short description why the use of independent ANN with a single output node for each lead time is of advantage compared to having a single ANN with multiple output nodes, one for each lead time.

7. I was not able to find information about the units for Q and P. Please clarify this in a revised version.

8. P907L19 How does the result depend on the network geometry/number of nodes.

9. P908L17-25 Fig1 may not be representative for the entire simulation period. For example, it seems that nodes 4,5 and 9 form the class of the late recession period - whereas the text states that these nodes are less easily identifiable. Maybe a longer time series could be shown using a different representation of the class assignment?

10. Please give a graphic representation of Table 2 as this is much easier to read.

11. P910 L6ff. I like the use of multiple performance measures and the use of a benchmark model as suggested by Schaefli and Gupta (2007). To make your work more transparent, maybe you could highlight how you decided on which performance measures to use?

12. P911 L16-29 Please clarify why you don’t calculate a new SOM with fewer nodes instead of combining nodes in a somewhat subjective manner? I suggest to recalculate the SOM with only 4 nodes.

HESSD

6, C144–C147, 2009

13. multi-network and multinetwork are both used - please be consistent. I would suggest to use multi-network

14. Have a native speaker check for English. Examples of sentences, that are hard to understand are: p899 L11 where it is explicit the intention of modelling; p899 L21 In order to keep openly into account the fact; p911 L4-15 is hard to understand; p901 L18-21 is hard to understand;

Schaefli, B. Gupta, H. V. Do Nash values have value? Hydrological Processes, 2007 , 21 , 2075-2080

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 6, 897, 2009.