

Review of paper

'Classification of hydro-meteorological conditions and multiple artificial neural networks for streamflow forecasting'

By E. Toth

Dear Author, dear Editor,

I have reviewed the aforementioned work. My conclusions and comments are as follows:

1. Scope

The work is well within the scope of HESS.

2. Summary

With the proposed work, the author contributes to the field of short-term, real-time flood forecasting.

The approach is a data-driven, system-theoretic set of Artificial Neural Network (ANN) models. The set consists of one model per forecast lead time (one to six hours ahead) and hydro-meteorological situation (a distinction of nine and four situations was tested). The identification of the hydrometeorological situation is accomplished with a Self Organizing Map (SOM) fed with the same input as the forecast ANNs: streamflow observations of the last three hours and areal-averaged rainfall observations of the last two hours.

The major findings of the work are: a) SOMs allow classification of hydro-meteorological situations in accordance to intuition and basic hydrological process distinction, b) multi-model ANNs based on the aforementioned classification outperform global ANNs with respect to forecast error and, partly, efficiency, if the number of classes is adequate (not too many, here: four instead of nine).

3. Overall ranking

The work is ranked 'major revision'. This is due to aspects that require clarification and more detail, as described below.

4. General evaluation

Scientific significance

The use of SOMs for hydro-meteorological classification as basis for multi-model ANNs for flood-forecasting is not widespread yet and a promising technique. However, the new and original parts of the presented work need to be pointed out more clearly, especially in delineation to the work that has already been done in this field (e.g. Jain and Srinivasulu, 2006, referenced in the work).

Scientific quality

The scientific approach and the carrying out of the work is mainly clearly explained, self-contained and valid. Exceptions are:

- Working with areal-averaged rainfall in 830 km² catchment is a substantial loss of spatial information which influences shape and timing of the discharge hydrograph. The proposed model may perform better when rainfall input is provided in, say, 2 subregions (up- and downstream). The subdivision may be done applying a SOM on the individual raingauges.
- To a non-expert, the SOM technique is not completely clear from the explanation (P. 906 line 1 to p.907, line 12). Please explain in a more detailed, step-by-step manner.

- For the sake of bridging the gap between system-theoretic and 'classical' (i.e. conceptual or process-based) modelers, a comparison of model performance with a simple conceptual r-r model will be helpful. Thereby, the number of free parameters subject to optimization/calibration in each model should be addressed and compared.
- Minor exceptions: see 'specific comments'.

The major part of the conclusions is comprehensible and substantial: a) the suitability of SOMs for automated classification of hydro-meteorological conditions, b) the benefit of selecting an adequate number of classes. However, further reasoning is necessary to explain why the global ANN, with respect to efficiency, equals or outperforms the multi-model ANNs for short lead times (one to three hours).

Also, the conclusions should pick up and discuss findings of Jain and Srinivasulu, 2006 (referenced in the work) stating that decomposing a flow hydrograph based on physical concepts is better than using the SOM decomposition.

Presentation of the state of the art and related work in the relevant scientific fields is extensive and up to date.

Presentation quality

The work is structured in a logical and intuitive manner, text length is appropriate. Additional figures will enhance comprehensibility (see 'specific comments'), especially a graph of the one to six hour forecast hydrograph as combination of the six model outputs.

Title and abstract reflect the content of the work (minor remark: see 'specific comments')

5. Specific comments

Note: page and line numbers refer to the printer-friendly version of the work as published in HESSD (<http://www.hydrol-earth-syst-sci-discuss.net/6/897/2009/hessd-6-897-2009.pdf>).

- P. 898, line 12: name the information used: Streamflow and areal-averaged rainfall of the last three (two, respectively) hours.
- P. 902, line 8: include a short description of the hydrometeorological regime of the basin in general and of the calib/valid period (e.g. did singular floods occur, which can strongly influence the results etc.).
- P. 901, line 23 to p. 902, line 3: If you mention this, clearly state the connection to your work, i.e. explain what your approach is and how you overcome the existing deficiency (missing consideration of the influence of the calib dataset).
- P. 903, line 25: Add: 'see below' to clarify to reader that some of the interesting applications are presented in the next paragraph.
- P. 905, lines 5-10: As mentioned in the general evaluation, point out clearly the new and original parts of the presented work in comparison to the existing works presented in the section above. Delineate especially to work closely related to yours (e.g. Jain and Srinivasulu, 2006, referenced in this work).
- P. 905, lines 15-18: Maybe the point is clearer for the reader if the argumentation is turned around: 'It is important to underline that the combination of rainfall and runoff observations prior to the forecast contains valuable information about the systems' state of saturation and hence on its reaction to rainfall forcing in the forecast period'.
- P. 906, line 15: State that the Euclidean distance is the standard measure, otherwise explain why you have selected this specific measure.
- P. 908, line 17 / fig 1: Please add the textual explanation of the hydro-meteorological situations to figure 1. Especially in case of black/white printout, this increases understandability.

- P. 909 ff (chapter rainfall-runoff modeling): As the one to six hour forecast hydrograph in your work is based on six different models, it would be interesting to see a forecast hydrograph as combination of the six individual forecasts in order to asses its temporal autocorrelation and 'realistic looks'.
- P. 912, lines 14-15: replace '...only information available in the forecast instant' by '... most relevant information available in the forecast instant: streamflow and rainfall observations.'
- P. 912, line 18: replace with: '... appear penalized by the partly low class occupancy'.
- P. 912, line 27: As mentioned in the general evaluation, further reasoning is necessary to explain why the global ANN, with respect to efficiency, equals or outperforms the multi-model ANNs for short lead times (one to three hours).

Spelling / language:

- Replace 'numerousness' by 'occupancy'
- Replace 'precious' by 'valuable'
- Replace 'providing in input past inflows ...' by 'using past inflows as input'
- P. 899, line 12: explicitly
- P. 899, line20:In order to take into account ...
- P. 903, line 6: replace 'resulted' by 'was'
- P. 904, line 22: ... on past river flow only. However, information on the recent precipitation depths is valuable
- P. 910, line 15: [0,oo[
- P. 911, line 16: '... problem, the opportunity to was tested, so to ensure...'
- P. 911, line 25: recognized

Yours sincerely,
Uwe Ehret