**Hydrology and
Earth System
Sciences
Discussions**

Interactive comment on "Experimental investigation of the predictive capabilities of data driven modeling techniques in hydrology – Part 2: Application" *by* A. Elshorbagy et al.

Anonymous Referee #2

Received and published: 23 March 2010

General comments:

Improvement of good practice in model setup and application when dealing with data driven models (DDM) is of great importance in the domains of hydro- and environmental science where applications mostly suffer from a lack of observed data. In addition, the assessment of the generalizing capabilities, the predictability and the uncertainties of DDM is still intensively discussed in the scientific community and, as far as I know, no generic framework is available which can be used for this purpose. Thus, the 2-part paper provides a step towards a more elaborate setup, analysis and evaluation of

C3522

DDM. It is reasonably well written and a novel contribution for assessing the predictive performance of DDM techniques. In a comprehensive comparative study the proposed method is applied to six DDM techniques, which (accept M5 model trees) are available "out of the box" in different software packages from the internet. The impact of experience in model building with the one or the other DDM technique is not considered. This makes the study very interesting for scientists who did not use a certain DDM technique (for example artificial neural networks) before or who use it with standard parameter setups and/or training (calibration) procedures. For more advanced modellers the provided benchmark dataset (which should be available in the internet) combined with the proposed evaluation methods is a good opportunity to compare the performance of more "tuned" or sophisticated setups of various DDM techniques. The results of the DDM experiment are very carefully analysed with appropriate methods of statistics and conclusions from statistics are correct. However, I suggest to combine both manuscripts in a single, more condensed paper (for example section 4 of the 1st paper can be considerably shortened).

Specific comments: 2nd paper

2nd section: 1) Pages 7098/7099 – Why did you test several kernels for the SVM models and not different (also linear) activation functions in the case of ANN.

3rd section: 2) The results are very carefully and comprehensively analysed and discussed!

5th section: Page 7113 – line 26 I would also expect benefits if DDM techniques would be combined in ensembles e.g. using bayesian model averaging, see

Th.Wöhling and J. A. Vrugt: Combining multi-objective optimization and bayesian model averaging to calibrate forecast ensembles of soil hydraulic models. Water Resources Research, 44:W12432, 2008. doi: 10.1029/2008WR007154.

Technical corrections:

2nd paper

1st section: 1) Page 7097 – acronym AMI is used before defined

5th section: 2) Page 7114 – line 12 "pre-processing . . ."

figures section: Page 7138/7139 – ylabel->closing bracket