Hydrol. Earth Syst. Sci. Discuss., 6, C3514–C3516, 2010
www.hydrol-earth-syst-sci-discuss.net/6/C3514/2010/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Experimental investigation of the predictive capabilities of data driven modeling techniques in hydrology – Part 1: Concepts and methodology" *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 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:

1st paper

1st section: 1) Page 7057 – I think it is not necessary to introduce the term "soft computing techniques" if data driven modelling is used as a general term.

2nd section: 2) When studies about ANNs are discussed, it should be mentioned which architecture is used. Generally (in both papers), the reader may get the impression that the term ANN is synonym with feed forward neural networks (FFNN) (also see below).

3rd section: 3) Why those six DDM techniques were chosen and not others? In addition, I understand ANNs as a general class of architectures and training strategies which represents many data driven techniques – among them FFNN as a global approximation technique. How about local neural approximation techniques? 4) Page

7063 – Please provide also information about the error function which was used when applying each of the six DDM techniques. Also in the 2nd paper I found sparse information about it.

4th section: 5) The description of the various DDM techniques can be considerably shortened: e.g. ANN from Page 7066 – line 26 to the end of the paragraph, GP from Page 7068 – line 18 to the end of the paragraph, EPR from Page 7072 – line 14 to the end of the paragraph, SVM from Page 7073 – line 10 to Page 7075 – line 1, MT from Page 7075 – line 18 to the end of the paragraph. 6) A discussion of dynamic versus static regression strategies for EPR (Page 7072 – line 18 -21) would also be interesting in the context of the results of the conducted DDM experiments and can be included in the 2nd paper.

Technical corrections:

1st paper

1) Page 7057 – last line “technique”

2nd section: 2) Some acronyms like GP, SVM, EPR are used before defined 3) Page 7060 – line 11 “disaggregates . . . from the . . .”

3rd section: 4) Eq. 5 – “IPE_{ij}= . . .”

4th section: 5) Page 7073 – line 15 “where . . .” 6) Page 7074 – line 4 “Müller . . .” – u-umlaut

5th section: 7) Page 7080 – line 26 “subcatchment . . .” 8) Page 7081 – line 15 “(Berger, 1992) 9) Page 7081 – line 16 "More . . .“

figures section: 10) Page 7091 – add labels to x- and y- axis

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 6, 7055, 2009.