

## ***Interactive comment on “Calibration of hydrological models using flow-duration curves” by I. K. Westerberg et al.***

**J. Vrugt (Editor)**

jasper@uci.edu

Received and published: 11 April 2011

I would like to thank the three reviewers for their excellent comments and suggestions. I would like to accept the paper, but pending major revisions. In particular a correctly revised paper should

1. Better establish the connection with previous work, particularly on the idea of using Flow Duration Curves for model calibration and model evaluation. The review of Prof. Sivapalan contains important points and suggestions that should be forcefully addressed in a revision. Also, an earlier paper by Winsemius also uses a signature based approach to rainfall runoff modeling, but considers explicitly uncertainty in the original data used in the FDC. Actually, other Bayesian methods explicitly consider data

C5304

uncertainty as well. Data assimilation papers provide detail about this. Main point, the ideas in this paper cannot be presented as novel and new, and the paper should do a much better job in mentioning and discussing previous contributions. If this is not done, I suspect the paper will not reach its full potential.

2. Establish the limited information content of Flow Duration Curves for inferring rainfall - runoff model parameters that consider timing and onset of the hydrograph. In response to this comment, I have done a few calibrations myself (other model and data). and have noticed that the proposed procedure works relatively well for simple (low parameter) problems, but exhibits difficulty inferring the appropriate parameters for more complex models.

3. How would the presented results be if the discharge data are transformed prior to the analysis? Are the findings still the same? Indeed, the NSE has several shortcomings as it comes to low and intermediate flow (NSE emphasizes peakflow), and therefore a transformation of the data might yield completely different conclusions.

4. The NSE threshold used to differentiate between acceptable and non-acceptable behavior is subjective and strongly influences the outcome of the analysis. Moreover, major improvements have been made in the past decade with respect to model calibration, and very few studies rely on the NSE only to retrieve the appropriate parameters. A study by Yapo et al. (1996) uses multiple measures (BIAS, RMSE, etc). simultaneously. Therefore, NSE might be an insufficient benchmark; and other studies are available that have much better treated discharge error. Non parametric approaches are available that estimate the discharge error from the measurements itself. I also would like to refer to a study by Schaefli and Gupta that have clearly established the limitations of NS type of criteria.

I am looking forward to a manuscript that correctly incorporates the many suggestions of the reviewers and the main points considered herein.