

## ***Interactive comment on “Comparison of catchment grouping methods for flow duration curve estimation at ungauged sites in France” by E. Sauquet and C. Catalogne***

**Anonymous Referee #3**

Received and published: 25 May 2011

Thank you for the opportunity to review the paper titled “Comparison of catchment grouping methods for flow duration curve estimation at ungauged sites in France.” I think the paper is very relevant, within the scope of HESSD and provides a novel and simple approach to estimate the flow-duration curve at ungauged sites. The paper is also well-written and organized. I do have several specific comments and questions enumerated below which I hope will help to clarify the results. Many of the comments are suggestions that I do not feel should hold up final publication; however, I would like the authors to address comments (3), (6) and (7) before final publication.

C1772

1) p. 3242, lines 4-6: If a model is able to generate negative values, this would represent a serious limitation of the models presented to estimate streamflow at an ungauged location. Is there a more formal way that these methods could be constrained rather than by simply replacing negative values to 0.001?

2) I think the use of the concavity index is a clever and interesting way to think about categorizing regions of similarity between flow duration curves.

3) Please add more description to figure 9 in the caption and axis text and add the correlation values to the figure. It is very difficult to see how figure 9 support the statements made on p. 3250, lines 1-25 without more description on the figure itself.

4) Because different numbers of catchments were members of the groups, was some weighting applied to the regressions to handle the unequal sample sizes? Or do the authors feel the results are not sensitive to unequal sample sizes used to adjust the regression models?

5) In the discussion about canonical correlation analysis (CCA) on p. 3251-3252, it is interesting that CCA resulted in such poor results for this paper. I appreciate the authors looking into this issue but wonder if more could be said about why CCA performs so poorly. Is there something about flood data that makes CCA well-suited for application to that problem but not for flow duration curve estimation?

6) It is interesting that from figures 11 and 12, the visual grouping does not appear to provide a dramatic improvement in the regional regression model. Is this apparent marginal improvement due to the fact the top left graphs in figures 11 and 12 do not have the same scale for the y-axis? An alternative approach to evaluating the grouping would be to compare the results to a random grouping where groups were determined at random and regressions developed from those random groups.

7) p. 3251, lines 24-26: I think that the statement that regression tree performed the best is hard to conclude from figure 11 and 12. Add supporting evidence for this state-

C1773

ment. A mention is made of the errors in visual grouping being heteroscedastic but no supporting evidence is provided and, ultimately, I feel this argument is not strong enough to completely discredit the method if the cross-validation results show otherwise.

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 8, 3233, 2011.

C1774