

## ***Interactive comment on “Fuzzy committees of specialised rainfall-runoff models: further enhancements” by N. Kayastha et al.***

**Anonymous Referee #1**

Received and published: 1 April 2013

The paper presents an analysis and development of a theory that explores the improvement of predictive models by a soft combination of modules or models. The concept of fuzzy committee have been well conceptualized by some of the authors of this paper in their previous publications, however few papers have developed comparative examples with optimized models. Although it is a very interesting study I have a couple of concerns. The main conclusion is taken from calibration and not from the verification data set. There is also an important difference in performance between calibration and verification (Figure 3 and 4). I would suggest extending the verification results on only low flows or high flow regions, and to check where is the “gain or loss” of performance. With this, it should be possible to detect whether there is improvement or not of the fuzzy committee.

I should mention a number of important points that I believe should be addressed.

C588

1. The fact that RMSE is an error measure that squares errors and provides implicitly more weight to high values makes it not suitable to compare two different regimes at the same time. May be a normalized value could provide more information.
2. It is interesting to see that the high flows in Leaf catchment (fig. 3) have less RMSE than the low flows. I believe this is due to the RMSE used in the optimization of the models, that RMSE includes weights in its operation, the graph is misleading. The graph should show the normal RMSE to be able to provide information about the observed value and not the distorted reality balanced by the weights of hypothetical flow regimes determined by the arbitrary parameter Alpha.
3. On the other hand, it is well known that due to the random generation of some of the parameters the overall RMSE variability in the calibration imply performance values in verification. Therefore, most of the models might have RMSE lower than the committee error improvement presented in the paper, if so the conclusion is not really possible to be made out of such results (check table 3, low difference in values of errors). It is important to make either an analysis of the variability of the models used with each data set (verification) or either do a ten fold cross validation process.
4. The actual pareto front seems to show only calibration values, may be is better to show pareto graphs only with verification results; if the goal is to conclude something about performance. If the goal is to conclude on the calibration capabilities and its relation with verification samples, this should be identified in the same graphs. The Figure 3 and 4 does not separate the pareto local models used in calibration and verification.
5. It is important to provide the reader with a figure that allows him to visualize the hydrograph and highlights what is considered as low and what is high flow, according to the fuzzy parameters selected (Related to the conclusion Page 683 line 25). I believe that the difference in regimes is the most probable reason of the improvement difference.

Aside of this, it is possible to see a number of English mistakes that would be important to correct to improve the readability.

Page 677 line 19 to 22, please divide and explain better, is not clear. Page 680 line 14,

C589

what is viva versa? sentence is not clear Page 681 line 5, Check sentence. Page 681 line 10, why to use NSE coefficient if your targets have been built with RMSE weights. Did you check NSE per regime of the data? Page 681 line 24, check sentence. Page 683 line 17, this contradicts your 3rd conclusion.

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 10, 675, 2013.

C590