

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

E. Sauquet and C. Catalogne

eric.sauquet@cemagref.fr

Received and published: 23 June 2011

ABSTRACT. On line 2, it is reported that one of the aims of the study is the quantification of the uncertainty associated to the estimated FDC. In my opinion, this point has not been addressed in the paper. For instance, in the last figures the performances in terms of prediction errors are reported for each classification model, but there is not any evaluation of uncertainty in ungauged sites (e.g. with confidence bands).

==> Cross validation was performed to give an insight into uncertainty at ungauged sites. Due to discretization of the FDC uncertainty cannot be given out of the 15 per-

C2371

centiles. Nevertheless the distribution of the relative error (absolute value) is provided through boxplots (Fig. 10 to 13) and can be used thereafter to compute a p% confidence empirical interval at ungauged locations with upper and lower bounds at the $(1-p)/2\%$ and $(1+p)/2\%$ limit.

P3236, itemized list. On lines 5-7 the authors introduce the idea of a parsimonious model to represent dimensionless FDCs, and then list some advantages of this choice.

* lines 8-9. I do not understand the meaning and the usefulness of this sentence. In particular, it is not clear the double-reference to the long-term mean (once as 'qa' and then as 'index value'). I can try to guess an alternative sentence like 'the use of a parsimonious dimensionless FDC can provide an easy way to reconstruct the FDC, provided an estimate of the index flow'. Please, reformulate the sentence. * lines 10-11. The adoption of a model of FDC with few parameters certainly simplifies the regional analysis, during both the calibration and the application phases. However, I think that the 'steps of the regional procedure' represent a more general concept; e.g. the steps can be: the computation of hydrological parameters, the (possible) definition of homogeneous regions, the choice of a function relating hydrological variables to basin characteristics, ... The use of few parameters reduces only the effort of each 'regionalization step', but not the number of steps, so I would write 'It reduces the computational effort...' and I would add that 'a few parameters can be easier to interpret'.

==> We agree with you. The sentences have been reformulated to make these points clear.

* lines 12-15. The sentence states that the dimensionless FDC (through its shape parameters) provides a synthesis of the characteristic response of the catchment to rainfall. In general, however, the low-flow part of the FDC is not directly affected by rainfall as, in many cases, low-flows are mainly driven by the groundwater release and the geology of the basin (e.g. Smakhtin, 2001). This point needs to be reviewed.

==> We have modified the text, replacing rainfall by climate since the basin acts as a

C2372

filter to climatic inputs.

P3241, EQ (5). I have different concerns about the equation:

==> The equations and notations have been modified.

P3241-3242, L22-27. In this case study, FDCs have a large variety of shapes, as can be deduced from fig. 4. Due to this variability, some further information need to be reported to evaluate the fitting skills of the EOF method. In particular, I would report: * a plot of the shape function b_1 obtained for the case study; * a plot showing the variability of the empirical weights a_i ; * a plot showing how the modeled FDC changes as the parameter a_1 (and/or $\log(Q(i))$) changes, in order to represent how much 'flexible' the fitting is.

==> One figure (Fig. 4) has been added and briefly commented in the last version.

SECT 4 and 5. The authors defined fixed-region classifications (through the RT and CCA methods) based on two hydrological indices: IC and SR. I agree with them that IC and SR are useful because they have a physical meaning; however, a 'reference classification', directly based on the regional variables (a_1 and $\log(Q(i))$), should also be used in the regionalization and the results compared with the previous ones.

==> We have added justifications.

SECT 5 and 6. In general, it is not clear if any statistical test for homogeneity has been done to 'validate' each region. This point need to be discussed in the paper.

We have added some comments on this aspect.

Minor comments

P 3236, L2. The term 'identical' is a bit strong. I would prefer something like 'equivalent' or 'statistically identical'.

==> It has been changed.

C2373

P3237, L14. For the purpose of streamflow regionalization (prediction in ungauged basins), the classification must be done in a space based on some catchment characteristics (morphological, climatic, geological,...) which can be determined independently of the streamflow data. The catchments characteristics are also known as 'descriptors', so that this space is often called 'descriptors space'. The authors use the term 'hydrological space' to refer to the descriptors space (or space of auxiliary variables); this may be misleading because the descriptors space cannot contain variables related to the streamflow. In this context, I would change 'hydrological space' to 'descriptors space'. The same correction should be done on line 1 where the 'hydrological neighbourhood' is cited.

==> It has been changed.

P3238, L27. In the presented case study, only basins without significant human influence have been selected. Which are the criteria used to define the presence/absence of a significant anthropic effect?

==> More details are given in the revised version.

P3240, L20. Please, specify what is intended with 'standardized': zero mean and unit variance or just 'dimensionless' (divided by the mean value) like in P3236 L6?

==> Standardized is indeed not relevant.

P3242, L3. I would delete 'uncertainty'.

The term 'uncertainty' was deleted.

SECT 2. The number of suitable stream gauge used in the study is very large. Did the authors consider the effect of nested catchments?

We are aware when nested catchments are considered information on the upper part of the basins may be introduced several times in the regression formulas due to the over-lapping drainage area. This redundancy may bias the spatial analysis. There is

C2374

no objective manner to account for this kind of dependency. One option should have been to consider weighted regressions with weights defined by the proportion of non overlapping areas between gauging stations (even if there is no theoretical justification).

SECT 5.1.3. The indices i and p used in this section are in contrast with those used elsewhere in the paper. Other indices will avoid misunderstandings.

==> We have checked the notation and modified them to avoid confusions.

P3250, L3. Multivariate should be Multiple.

==> It was done.

P3250 EQ.(8)-(11). Please define the coefficients λ s.

==> λ s are the regression coefficients fitted on observations to each homogeneous group by the ordinary least squares method (using log transformed data to fit power-form models).

P3250 L19-21. The authors found one suitable model (regression) for each region. It is interesting to report: * the most significant descriptors selected; * if they considered completely different regressions for each region, or a unique model (i.e. the same set of descriptors) with coefficients changing from a region to another.

==> Results on regressions were commented in Sect 6.2.

Technical corrections

P3247, L7. 'Mahalanobis' instead of 'Mahanalobis'.

==> It was done.

P3251 L 2. Correct 'Sect. 4.5'.

==> It was done.

C2375

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

C2376

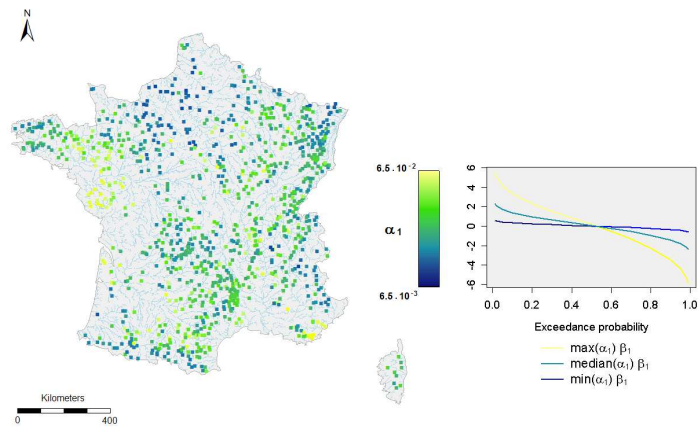


Fig. 1. Additional figure (Fig 4)

C2377