

Reply to interactive comment by Anonymous Referee #1 on “Understanding Hydrologic Variability across Europe through Catchment Classification”

Anna Kuentz, Berit Arheimer, Yeshewatesfa Hundecha, Thorsten Wagener

The authors would like to thank Referee #1 for his/her constructive comments and suggestions that will for sure help us improving the paper. Our replies to the different comments are written below (in blue font) after each comment/question (written in italics).

This paper uses established methods to classify regions with similar physiographic characteristics and with similar flow signatures to determine the best predictive relations at ungauged locations. For this reason, the manuscript reads more as a report rather than a novel contribution to the literature. For example, in lines 13-15, it seems the manuscript goals do not seem to be driven by scientific hypothesis but more by having a large set of data and wanting to develop/explore some relations which may (or may not) be useful at some later point. In this way, I think the motivation for the study seems weak as a scientific contribution. Despite this, I do believe that the novelty of the manuscript is in the application of these methods over such a large spatial domain. As hydrologic modeling efforts expand to cover continental scales, the ability to upscale existing approaches for model calibration across large ungauged regions becomes a limiting factor in these efforts. This point should be emphasized more in the manuscript to elevate the impact of the work beyond an application of existing methods to a larger region than had been tested previously.

The reviewer is completely right in his/her assumption that the scientific contribution of this paper is not in developing any new methods – but applying existing methods to learn more about nature, in this case hydrological controls across Europe. This is emphasized both in the Introduction and in the Discussion and Conclusion of the paper.

1. The selection of flow signatures needs more details as to how they were selected. Olden and Poff (2003) do not from my remembering of the paper - as the authors indicate in line 16 - provide 9 signatures. Their paper attempts to reduce redundancy in the 200+ statistics that have been used for hydro-ecologic classification but they do not provide a definitive reduced list. More details need to be provided as to why these signatures were selected, particularly because their usefulness in applications is not part of the analysis in the paper. This reads as quite an arbitrary choice.

Thank you for raising this point, the choice flow signatures was indeed made very carefully in this work. The reviewer is right in claiming that Olden and Poff (2003) do not provide a definitive reduced list of signatures, but they do suggest a way to select such a list when saying “One could reduce the population of indices to a minimum of nine, each of which exhibits the highest absolute loading for the first principal-component axes for each of the nine distinct components of the flow regime (Table III). This ensures that the majority of the variation is accounted for and that different facets of the flow regimes are adequately represented in subsequent analyses. Furthermore, given the particular ecological question being addressed, additional indices within each flow component could be selected (from the remaining significant principal components), which would not result in a substantial increase in redundancy”.

According to their suggestion we selected the nine indices with the highest absolute loading for the first principal-component axes for each of the nine distinct components of the flow regime, except for number of zero-flow days describing the component “duration of low events”, because this index was too specific to intermittent rivers (a very large majority of the rivers in our domain had a value of 0 for this index). For this component of the flow regime we selected instead the index with the highest absolute loading for the second principal-component axis.

2. In lines 12-13 (p. 3), the comment is made that this type of analysis has not been applied at the continental scale “including large rivers with human alteration...” Do the catchments examined here have human alteration? This is not noted in the methods? Does this bias your results?

When visually checking the hydrographs of each flow station, the catchments with obvious and very strong flow regulation were removed. Though, a part of the catchments used in the study still have various forms of human alteration. This has partly been taken into account with some indices like agricultural area, urban area or irrigated area. Unfortunately we haven't been able to find a good indicator of flow regulation available over the whole Europe but this would certainly be of interest if such an index became available. Nevertheless, impact from regulation was clearly identified in the hydrological interpretation of similarities between catchments in specific groups. This is part of the results (Table 3), which is discussed in Section 3.3.

3. In line 18 (p. 8), the statement is made that “identified gauging stations that should be further explored and filtered out...” Was this actually done?

Thank you for pointing out this imprecision. These stations were filtered out, but no further analysis was done on them yet. The sentence will be modified to make it clearer.

4. In Section 2.3, how were variables determined to be significant in the regressions? What diagnostics were used? How many variables were allowed to enter in each equation? It may be useful as an explanatory tool to see which variables are significant but to make predictions (which is the goal of this work), one needs to adhere to good statistical practices. How were these practices followed?

We agree with the reviewer on the importance of providing this information. The significance testing is described in section 3.1: significance of correlations was tested based on a t distribution with a threshold of 0.05. We agree that this information should be available in section 2.3 as well and will be added 1.13. The way the variables (and the number of variables) were selected for each regression is described in section 2.3 p. 8 l. 24-29 (stepwise regression based on the Bayesian Information Criterion). The built regression models were evaluated using statistical measures such as the coefficient of determination. These different steps constitute an established statistical procedure to build and evaluate regression models.

However, we want to point out here that, as stated in the introduction, the main goal doesn't lie in the prediction itself but in gaining better understanding in the hydrological patterns across the European continent. The regression models, like the classifications, are used as a tool to reach this better understanding by exploring the relationships between descriptors and signatures and highlighting the main controls of flow signatures in different types of European catchments.

5. I found myself questioning the value of Section 3.1. I do not think this offers any additional information beyond what can be determined from the CART and regression analyses. This section also contributes to the manuscript reading more as a report as this section seems to explain what could be characterized as exploratory data analysis that is completed before one settles on an approach and hypothesis to test. I also think that the manuscript is a bit laborious in its reading and removal of this section would help streamline the manuscript.

Thank you for this suggestion for improving the readability of the paper. This first part of analysis was performed to give a first overview of the links between descriptors and signatures and to closely study the catchment descriptors to decide whether it was reasonable to keep them for further analysis (13 of them were removed). We agree to move most of section 3.1 to the supplementary material and only state the main conclusions of this part of the study in the main text.

6. Section 3.2 seems to be missing a reference to how the classification was applied to the data. At the very least, reference Section 2.2 to describe how the classification was completed.

We agree with this suggestion, the way the classification was applied to the data is indeed described in section 2.2 and a reference to this section will be added in the result section.

7. I may have missed this but I think it is necessary to develop regression on the flow signatures using the entire dataset to compare to the regression results obtained for the classes. This analysis would determine the objective improvements provided by first classifying the data. If this analysis has been completed, please refer to this in the text when discussing the results.

The reviewer is right in raising the importance of comparing the regressions obtained for the classes with models calibrated using the entire dataset. This has been done in our study as described in section 2.3 (p. 8 l. 21-24). Both regressions using the entire dataset and regressions obtained for the classes are analyzed in the result section 3.2. As following the reviewer's suggestion, a reference to section 2.3 will be inserted in the result section (p. 12 l. 19).

8. There are two papers that I direct the authors to for potential citation. Singh et al. (2014) used CART to classify model parameter behavior across the United States and may be helpful to motivate some of other contexts in which CART has been utilized for model parameterization at ungauged locations. Oudin et al. (2010) ask almost the same question as this paper in how physiographic similarity is related to hydrologic similarity, although they answer this question using actual model results.

We thank the reviewer for bringing these interesting papers to our attention, they are indeed completely relevant in the context of our work and we will add a reference to them in the revised version.

Oudin, L., A. Kay, V. Andréassian, and C. Perrin (2010), Are seemingly physically similar catchments truly hydrologically similar? *Water Resour. Res.*, 46, W11558, doi:10.1029/2009WR008887.

Singh, R., S.A. Archfield, and T. Wagener, 2014, Identifying dominant controls on hydrologic model parameter transfer from gauged to ungauged basins - a comparative hydrology approach, *Journal of Hydrology*, doi: 10.1016/j.jhydrol.2014.06.030, 2014.