Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-123-RC1, 2016

Hydrology and Earth System Sciences
Open Access
EGU
Discussions

Interactive comment on "Delineation of homogenous regions using hydrological variables predicted by projection pursuit regression" *by* M. Durocher et al.

Anonymous Referee #1

Received and published: 20 May 2016

General comments

The article "Delineation of homogeneous regions using hydrological variables predicted by projection pursuit regression" by Durocher et al. describes an improvement of existing techniques for the regional estimation of flood quantiles. The topic is very relevant, but I found the manuscript not completely clear in some parts and with some methodological flaws. While readability (see comment n. 1) can be improved with a revision of the text, some methodological issues would require a complete reanalysis of the work. My main concerns (see comments n.2 and 3) are basically related to the use of a very complex procedure that is not justified by the results. This finding (i.e., that the procedure does not really produce improvements) would be a result itself, but the authors

Printer-friendly version

Discussion paper

seem to overlook it to support their initial hypotheses. For these reasons, I suggest a rejection of manuscript.

Major comments

1. While the aim of the work is very clear, I found quite intricate the description of the operational procedure. The point list on page 6, and in particular the step i) (which is the main focus of the paper) should be supported by a quantitative example to make the procedure easier to understand. For instance, the plots in figure 3 could be used in this part of the manuscript to better describe how the procedure works (and not only from page 11 to comment results). Moreover, step i) seems a kind of "preliminary" regionalization of the L-moments of the target site. Such L-moments are then used to support the delineation of the region. Why such preliminary estimates cannot be directly used in the prediction of flood quantiles? This point should be discussed by the authors, highlighting the possible differences with the direct estimation of the quantiles based on preliminary L-moments.

2. The whole procedure is rather complex, so I would expect the proposed method clearly outperforming the others. However, looking at figure 5, it seems that the residuals are scattered more or less homogeneously around the bisector, meaning that the "traditional" and the "new" methods performs, in average, the same. Figure 5 tells me that there is no significant difference between the methods, so I would use the simpler one. Of course, by computing an error metric over the whole set of residuals one may obtain slightly better performances of the RVN models (but this is not reported; a summary table would be appreciated). The authors state on page 12, line 25 onwards, that improvement is effective for sites with largest discrepancies. This seems not true except for two point in figure 5a and one point in figure 5b (all the points in the bottom-left corner of each panel). Figures 5c and 5d show points equally distributed around the bisector also in the bottom-left corner. Hence, is the complexity of the RVN model justified by a so small performance improvement?

3. On page 10, line 29, the nonlinear relationship between the (transformed) predictors and the (transformed) L-moment is mentioned and the authors say that it is shown in figure 2. This non-linear relationship would justify the use of a spline interpolator, but actually this is a questionable point. In fact, figure 2 tells a different story. Panel a clearly show a linear relationship (in the transformed variables; this is expected as often the mean value can be linearized with log transformations). In the b, c and d panels there is a much larger scattering, which does not allow to identify a clear complex pattern, even if all the plots show an increasing trend. Looking at the scatter plots I believe that most of the people would adopt a simple linear regression (said with 2 parameters) which is much more stable and robust. My personally idea is that the choice of the authors is not justified and that a linear model should be at least compared to the spline interpolator.

Minor comments

P7 L6 Please, give a more detailed description of “true neighborhood” meaning.

P11 L5-7 I found quite strange that the L-kurtosis performs much better that the L-skewness as in general the prediction ability deteriorates with increasing order of L-moments. The authors should investigate in more detail this issue.

Sections 2.1 and 2.2 are rather obvious but useful for the paper, so they could be placed in appendix and referred to in the main text. I would also add at the end of the “Multiple regression” section a note about weighted and generalized least squares.

Figure 5. Not clear which kind of information is provided by the “smooth fitting of the residuals”. Also in this case, the smooth fitting seems too complex tool which does not add any further information.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-123, 2016.