Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-123-AC1, 2016

Hydrology and Earth System Sciences
Open Access
Discussions

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

M. Durocher et al.

martin.durocher@uqtr.ca

Received and published: 1 June 2016

** see the attached PDF**

Major (1.1) Reviewer: "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)." The authors agree that the illustration provided by Figure 3 is useful to understand this step of the methodology. The Figure 1 was initially intended to provide a schematic illustration of this step (i). Figure 1 will be improved in a revised version of the manuscript in order to be more similar to Figure 3 in its

interpretation. (1.2) Reviewer: "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." In a situation where the target distribution is assumed to be known (i.e. the same family of distribution is assumed everywhere) the preliminary prediction could be used indeed to deduce the parameter of the target distribution from the L-moments. However, in the framework the distribution is unknown and the L-moments cannot be used directly to estimate the flood quantile without specifying a family of distributions, which is achieved here by using the information of relevant neighborhoods. The comment of the reviewer also applied only to the utilization of the index-flood model inside the neighborhood. In a revised version of the manuscript, the authors agree to add a comparison between quantiles estimated by the parameter of the PPR method and the regional estimates. (2.1) Reviewer: "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). "

Notice that two error metrics are actually used in the present study: RRMSE and NHS. The authors agree that an additional table will help to summarize more directly the information provided in the text and in Figure 4 and 6 for the calibrated model. The Table below will be included and discussed in a revised version of the manuscript.

Table 1: Model RRMSE NHS AHM ARE Index-Flood ROI 46.2 86.5 72.8 57.3 CCA 45.4 86.2 41.7 42.9 RVN-LM 45.0 87.1 14.5 36.9 RVN-HYB 40.1 86.2 16.5 43.1 Regression-based ROI 44.9 86.9 72.8 64.7 CCA 45.2 86.4 40.7 26.7 RVN-LM 42.6 87.3 14.2 34.3 RVN-HYB 42.3 86.2 21.2 51.1 Best criteria in bold

(2.2) Reviewer: "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?"

The authors want to make the precision that the comment "improvement is effective for sites with largest discrepancies" applied only to the relative residuals as this comment is made while describing Figure 5a and 5b. To make it clearer, the terms "largest relative discrepancies" will be used in a revised manuscript. The additional complexity of the RVN method in respect of the traditional ROI method concern only the preliminary step (i), where the missing hydrological information is substituted by predicted values. The question is thus if the addition of the preliminary step is justified. Notice that RRMSE and NHS are cross-validated criteria and hence they are criteria that ÂńpenalizedÂż excessive complexity in the prediction. Consequently a better RRMSE implies that a model truly brings additional information on the quantiles. The overall improvement in terms of RRMSE for the RVN-HYB method in respect of the ROI and the CCA method is respectively 6.1% and 5.3% (see table in major comment 2.1 or Figure 4). Although of moderate amplitude, the authors believe that these improvements are nonetheless substantial. In particular, the authors agree with the reviewer that the red line in the bottom-left corner of Figure 5a and 5b are mostly influenced by few points. Nevertheless, in Figure 5a, the 4 largest relative discrepancies are better predicted by RVN-HYB and the two best relative improvement are of 77.2% and 68.5%. In Figure 5, small changes in average can be difficult to assess by the naked eye. The red lines are smooth curves fitted between the residuals of the two models is there to indicate which models in average have locally lower residuals. It can be seen that for the residuals above approximately 0.2 (i.e. 20% of the observed values) the red line is distinctively under the bisector (upper right corner) and this is true for several points. This means that the sites that are overestimated are in average less overestimated by the RVN-HYB method. Actually, 8 out of the 9 residuals located at the most right are

better predicted by the RVN-HYB model. Similar behavior is also noticeable in upper-left corner of Figure 5b. The present case study is an example of a region where some sites are problematic for likely any models. In practice, the residuals are not known, consequently we don't know if the target sites of interest will be "problematic" or not. Globally what the results and Figure 5a indicate is that, in a certain way, the RVN-HYB model is more robust, because for the sites that are well predicted by most models, including ROI, RVN-HYB will perform in average similarly. However if the target site is predicted less accurately, the RVN-HYB model will, in average, be better in terms of relative errors. Consequently, the overall gain may seem of moderate in magnitude, but for some problematic station the gain is more substantial. In a revised version of the manuscript the meaning of Figure 5 and the red line will be better explained and the arguments above will be added.

(3) Reviewer: "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." The authors understand the concern of the reviewer, but want to add the remark that the CCA model is actually the model that linearly predict the reference variables. Therefore, in the case of the index-flood model, the impact of assuming nonlinearity is measured by comparing the CCA and the RVN-LM model. These points could be made clearer in the revised version of the manuscript. Nevertheless the authors understand the concern of the reviewer in the

used of the PPR in Figure 2. In that case, the NSH for the two regressions method are provided in the table below. Table 2: NHS(%) L1(log) LCV(log) LSK(log) LKT(log) PPR 91.5 33.3 6.7 55.7 linear 90.9 28.2 7.8 48.1

For L1 and LSK the difference between the NHS criteria is about 1 %, which the authors agree is very similar. However, the gains for the LCV is of 5.1% and of 7.6% for the LKT are substantial. Notice that the NHS is a cross-validation criteria and hence this result does not represent a form of overfitting in favor of the more complex PPR approach as the prediction are obtained without the use of the predicted sites. Unnecessary complexity is then ÂńpenalizedÂż by such criteria. The fitting of LCV and LKT are cases of mild nonlinearity, but such mild nonlinearity can be adjusted here because 151 is a reasonable number of sites. The authors agree that the situation would be different if for instance only 20 sites where available. Therefore, in the present case study, it is false to say that linear models are more stable and robust than PPR as the performance is assessed by cross-validation. In a revised version of the manuscript the term "nonlinearity" will be changed for "mild nonlinearity" to be more precise. The interpretation of the criteria NHS as cross-validation criteria will be underlined and the particularity of having a sufficient number of sites to use PPR will be made clearer. Minor (1) Reviewer: "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." The authors, agree with the reviewer that L-skewness is expected to be, in general, better predicted than the L-kurtosis. The data have been investigated for data manipulation errors and nothing have been found. It appears to be a legitimate exception to the rule. (2) Reviewer: "P7 L6 Please, give a more detailed description of "true neighborhood" meaning." To clarify the term "true neighborhood" the following lines would be added to a revised manuscript around P7-L5: "If the hydrological variables were known at the target location, the distance would be available and the neighborhood that truly regroups the most hydrologically similar sites to the target location could be formed. However in practice this true neighborhood is unknown. Using instead an estimate

has the effect of falsely including that are less hydrologically similar than the sites that would be included in the true neighborhood." (3) 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. As briefly discussed in the major comment (2.2), at the proximity of a point say (X,X) in Figure 5, several points maybe under or above the bisector, which makes seeing small improvement for a method difficult. The smooth fitting provided by the red lines indicates specifically the average between the residuals of two models in the proximity of a point X. The authors agree with the reviewer that parsimony is important. The smoothing splines procedure used here as a visual guide is widely available in most numerical softwares and the "degree of complexity" is controlled by the generalized cross-validation criteria (GCV), which is a widely used criteria to avoid overfitting. Hence, with the information from 151 sites, the authors believe that the smooth fitting is not "too complex", but "as complex" as the residuals suggest it. More details on the interpretation of the red lines will be provided in a revised version of the manuscript.

Please also note the supplement to this comment:
http://www.hydrol-earth-syst-sci-discuss.net/hess-2016-123/hess-2016-123-AC1-supplement.pdf

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