

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

T. Gado (Referee)

tamergado@hotmail.com

Received and published: 22 May 2016

General Comments:

The present manuscript investigates the utilization of hydrological information in Regional Frequency Analysis (RFA) in order to improve homogenous properties of neighborhoods and then improve regional flood estimation.

In general, the paper is well organized and the contribution of the study is relevant. Nevertheless, throughout the manuscript, the authors support the idea of using the estimation of hydrological variables, instead of site characteristics, to delineate homogeneous regions. Yet, the estimation of hydrological variables is based on subjective

C1

selections of site characteristics and subject to model errors. Thus, using such estimations to improve homogenous properties of neighborhoods in RFA is questionable. Consequently, the improvements in the results are insignificant as mentioned by the referee #1. Moreover, the manuscript is not clear in some parts especially in the methodology and the results. In my opinion, the article may need major revisions before publication. Major and specific comments are shown below.

Major Comments:

1. The literature is not complete and does not state what other researchers have done in order to improve the flood estimation at ungauged sites. So, the authors should improve the description of the existing literature on the topic investigated. In particular, the manuscript should elaborate a little bit better on the evolution of the ROI method as the study focuses on the neighborhood approach for homogenous region delineation.
2. The methodology is blurred, difficult to follow, and contains some odd judgements. For instance, LSK was maintained because it is associated to better predictive performance, however, it is poorly predicted by the site characteristics (P11 L5 – 12). Also, the authors did not show some details such as the additional translation which necessary to avoid numerical difficulties of LSK and LKT due to negative values (P10 L19).
3. Although the authors introduced a complicated methodology, they did not make enough efforts to clarify the description of the results; such as confusing explanation of Fig. 4 (P12 L8 – 17) (e.g., why 80 sites? in P12 L12), and unclear Fig. 5 and its explanation (P12 L26 – 33). Furthermore, the presentation of the results of the regression-based model needs improvements to be clearer (P12 L35 – P13 L10). I recommend using the simple Q-Q plot to assess the compared methods regarding the estimation of regional flood quantile. Also, the results should contain numerical tables to quantitatively clarify the differences between the considered methods. The authors can find a close example for the presentation of such results in the reference Gado and Nguyen (2016). Finally, comparing the results of the index flood and the regression

C2

methods would be valuable here.

Specific comments:

1. P1 L16 – 17. Which properties does the hydrological information in Regional Frequency Analysis enforce for a group of gauged stations? I suggest to add “desired properties”.
2. P1 L18. Ungauged sites can be defined by site characteristics in the neighborhood delineation methods (e.g., ROI). Therefore, there is no a challenge for using neighborhoods in RFA regarding the unavailable hydrological information at ungauged sites.
3. P1 L23. The regional frequency analysis can be applied for flood or extreme rainfall or any other extreme events. Hence, it should be stated that the case study is for regional flood estimation.
4. P2 L21. “the distance between hydrological variables”. The distance is between locations not variables.
5. P3 L4. “as an estimation model”
6. P9 L4. Please, define NH in equation 14.
7. P9 L14 – 15. How can a regression model fitted on two different neighborhoods, for the same target location, obtain identical values?
8. P10 L2. I don’t believe that 15 years of data are enough to get statistically reliable results, why did authors choose 15 years as the minimum time series used in the study.
9. P10 L3. I think you should have at least a map showing the locations of the selected stations in the case study (Quebec).
10. P10 L7. Using the maximum likelihood for parameter estimation with small time series (e.g., 15 years) may cause convergence problems, I would recommend using L-moments instead.

C3

11. P11 L16. What does HYB denote for in “RVN-HYB”?
12. P11 L19. “One of the objectives of RFA is to identify a proper family of distributions from regional information” This is not an objective of the RFA. I suggest to write one of the main steps.
13. P11 L23. Please, define here the $Q(r)$ as the regional quantile.

References:

Gado, T. A., and Nguyen, V.T.V., 2016. Comparison of homogenous region delineation approaches for regional flood frequency analysis at ungauged sites. *J. of Hydrol. Eng.*, 21(3), Doi: 10.1061/(ASCE)HE.1943-5584.0001312.

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

C4