Hydrology and
Earth System
Sciences

Open Access

Discussions

EGU

Interactive comment on "Heterogeneity measures in hydrological studies: review and new developments" *by* A. I. Requena et al.

Anonymous Referee #2

Received and published: 13 July 2016

GENERAL COMMENTS

I really enjoyed reading this paper. I appreciate the efforts taken by the authors in contributing to the field of Regional Frequency analysis (RFA) by proposing a new index to assess heterogeneity of a region. There is clarity in describing the assumptions/drawbacks associated with the existing heterogeneity measures. Also, it is nicely stated why there is a need for new heterogeneity measures in RFA. The criteria that are defined to compare various heterogeneity measures with a new measure look adequate.

SPECIFIC COMMENTS

1. In the introduction section, literature review regarding regional hydrological frequency analysis is not complete.

2. Page 3, Line 32, "...focused on the delineation step". What do you mean by "..focussed on delineation step"? How Gini Index (proposed in this paper) account for the delineation step?

3. Page 4, line 28-29; what is the reason behind considering a linear relationship for L-CV and why gamma/2 was used?

4. Page 10, line 25-26; Gini index considered in this study is a function L-CV . Can GI be expressed in terms of L-skewness coefficient on the similar line as it was expressed in terms of L-CV? If yes, then will there be any change in the overall results from this study? I am asking this question because L-skewness would be more uncertain for observed data (due to involvement of higher (third) moment) which can influence the heterogeneity measure H2 or V2 (e.g variation in (V2) H2 heterogeneity measure would be more as compared to (V1) H1 heterogeneity measure; even visible in Figures 1, 2 and 3). As the indices are compared based on hypothetical regions (using Monte Carlo simulation), the effect on L-skewness coefficient may not be noticeable. But in practice, there can be uncertainty associated with the estimation of L-skewness coefficient.

5. What is the range GI index?

6. Page 16, line 3-10 describes use of GI in assessing the delineation methods. The methodology discussed is not clear. Different delineation methods have different criteria to select the optimal delineation/regionalization solution (e.g. AIC or BIC, Davies-Bouldin index in the case of hard clustering algorithms, Xie-Beni in the case of fuzzy clustering algorithm). Are you trying to say that we can rank the delineation methods based on GI index value to arrive at best delineation method for regionalization which then can be used to perform regional frequency analysis? If yes, then I think this kind of approach is even possible with any other Index (e.g., H1, H2). Authors have mentioned on Page 3, line 32 that the proposed index (e.g., GI) can alleviate the drawback associated with delineation method as compared to the conventional heterogeneity indices (H1, H2 etc).

7. In Figure 7, impact of heterogeneity due to addition of discordant sites to homogeneous regions is presented. A similar observation can be concluded from Figures 1 where the indices are assessed with increase in the heterogeneity percent (gamma value). In the case of Figure 1, mean of conventional heterogeneity measures (H1, H2, AD) tend to diverge with increase in the degree of heterogeneity while in Figure 7, Heterogeneity measures tend to converge (coming close) with increase in the number of discordant sites (i.e. increase in the degree of heterogeneity). Kindly clarify this point. However, in the case of GI, in both figures1 and 7, convergence is visible. Why is this happening?

8. The results obtained from the Monte Carlo simulation study are encouraging. Only thing missing from the paper is to perform analysis using observed data which would help to strengthen the conclusions drawn from the study. Analysis based on observed data would provide more clarity on the aforementioned concerns especially on Specific comment 4.

TECHNICAL CORRECTIONS

1. First line of Abstract reads "Regional frequency analysis is needed to estimate hydrological quantiles at ungauged sites or to improve estimates at sites with short record lengths, by transferring information from gauged sites." I think the regional transfer of information is possible with hydrologically similar sites only. Hence, "...hydrologically similar gauged sites" instead of "...gauged sites" looks more appropriate.