Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-136-AC3, 2016

Hydrology and Earth System Sciences Discussions
Open Access

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

A. I. Requena et al.

ana.requena@ete.inrs.ca

Received and published: 29 August 2016

Authors' reply to Anonymous Referee #2

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.

Reply: The authors thank the reviewer for the thorough revision as well as for the constructive comments provided.

SPECIFIC COMMENTS

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

Reply: The literature review on regional hydrological frequency analysis included in the introduction of this study is not intended to be exhaustive. Its aim is to provide a general idea of the diversity of tools and methods existing for performing regional hydrological frequency analysis, with the aim of underlining the need for heterogeneity measures for the evaluation of the corresponding regional heterogeneity. In order to guide the reader in the search of additional references, "see Ali et al., 2012" will be changed to "see Ali et al., 2012 and references herein" (page 2 line 6). Also, the sentence "For further references on regional flood frequency analysis, please see Ouarda (2013), Salinas et al. (2013) and references herein" will be added to page 2 line 10.

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?

Reply: The authors apologise for the lack of clarity. The intention was to highlight that the use of a heterogeneity measure will allow direct comparison of the heterogeneity of regions delineated by different methods. Please, see reply to comment 6 for additional explanations about the use of the Gini Index for comparing the heterogeneity of regions delineated by different delineation methods. For clarity, "... and focused on the delineation step" will be replaced by the aforementioned explanation "Furthermore, the use of a heterogeneity measure will allow direct comparison of the heterogeneity of regions delineated by different methods".

3. Page 4, line 28-29; what is the reason behind considering a linear relationship for

L-CV and why gamma/2 was used?

Reply: This linear relation and the use of gamma/2 is a common and plausible way of simulating a varying at-site L-CV over a region. It has been used in other studies, such as Hosking and Wallis (1997) and Wright et al. (2015). In this regard, the following sentence will be added to page 4 line 29: "Note that this relation is commonly used in other studies (e.g. Hosking and Wallis, 1997; Wright et al., 2015) as a plausible way of simulating varying conditions over a region".

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.

Reply: The GI has been expressed in terms of the L-CV due to the fact that the L-CV has been recognised as a good surrogate of the difference between at-site flood distributions. It also has a physical meaning, as it is related to the slope of the frequency curve. In theory, the GI could be expressed in terms of the L-skewness as well. However, as mentioned by the reviewer, the use of the L-skewness would imply high uncertainty when applying it to real data. Note that the length of observed data is usually short to obtain an appropriate estimation of the L-skewness in practice, making the use of a lower L-moment such as the L-CV preferable. Also note that when traditionally assessing the homogeneity of a region, the use of the H1 statistic (only based on L-CV) has been generally recognised as more powerful than others such as H2 (based on L-CV and on L-skewness), and that to our knowledge, statistics only based

Printer-friendly version

Discussion paper

on L-skewness are not usually considered.

5. What is the range GI index?

Reply: This can be found in Sect. 3.3, after Eq. (10): "Theoretically, GI ranges from zero to one. The former is obtained when all the x_i values are equal, and the latter is given when all but one value equals zero (in an infinite population)".

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).

Reply: The authors apologise for the lack of clarity. The GI is intended to be used in assessing delineation methods by comparing the heterogeneity of the regions already identified as the "best regions" by each delineation method. For instance, each one of two different delineation methods (methodA and methodB) delineates two different sub-regions (sub-regionA1 and sub-regionA2; sub-regionB1 and sub-regionB2). The GI will then be applied on the at-site L-CV obtained from the flood data of each sub-region (GIA1, GIA2, GIB1, GIB2); and the GI for each delineation method will be obtained as the GI average of the corresponding sub-regions (GIA, GIB). Then, the best delineation method will be the one with the smallest GI (i.e., min(GIA,GIB)), as it obtains the less heterogeneous sub-regions. For clarity, this example will be properly incorporated into the Discussion section. Note that the GI is identified as the best measure for ranking regional heterogeneity in this study, as it is the one with the best behaviour over the four steps of the assessment procedure.

Regarding the comment of the reviewer concerning the text in page 3 line 32, the authors also apologise for the lack of clarity. Such a drawback was not related to conventional heterogeneity indices, but to procedures that imply the need of performing additional steps in the regional analysis for comparing delineations methods. Please, see reply to specific comment 2 for further explanations.

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?

Reply: It is important to highlight that, although both graphics have in common an increase in the regional heterogeneity, their characteristics and displayed information are different. In the case of Fig. 1, heterogeneous regions with increasing heterogeneity (from 0% to 100%) are generated by considering two different regional L-CVs and their heterogeneity values are compared. As a result, measures such as H1 and H2 "diverge" with increasing heterogeneity, as these measures are not originally thought to compare several (heterogeneous) regions. In the case of Fig. 7 the situation is different. A given homogeneous region is considered, and "homogenous" sites are exchanged by "discordant" sites. Hence, in this case, progressive changes in the heterogeneity of a given region are assessed (not between different regions). As a result, the heterogeneity of the region is more difficult to be ranked for all measures when there is an increase in the number of discordant sites and in the difference between L-

CVs (of homogeneous region and discordant sites). For clarity, the following sentence will be added to the Sect. 4.4, after the existing description of Fig. 7: "Note that unlike Fig. 1, where the heterogeneity value of two kinds of regions with the same degree of heterogeneity but different regional L-CV may be compared; in Fig. 7 progressive changes in the heterogeneity of a single homogeneous region are assessed."

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.

Reply: The authors agree with the reviewer on the need to extend the present study to observed data. However, the present manuscript is itself a very dense paper, and such application would imply the delineation of homogenous regions for a given case study by using different delineation methods, which may be considered as a whole new study. The authors would also like to highlight that many studies over the literature compared measures by using simulations without including a real case study (e.g. Genest et al. 2009). Furthermore, in a real case study, the performance of the measures cannot be evaluated because of the absence of the reference heterogeneity value (unlike in simulations). The authors believe that including a case study may not provide an added value to the paper but could have a negative impact in terms of clarity and readability, while making the manuscript even longer. In this regard, the following sentence will be added to the last paragraph of Conclusions: "In this study, the performance of heterogeneity measures is evaluated through a simulation-based procedure and recommendations are given. Further research may focus on their application to a variety of case studies in order to analyse practical aspects."

TECHNICAL CORRECTIONS

1. First line of Abstract reads "Regional frequency analysis is needed to estimate hy-

drological 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.

Reply: The suggestion of the reviewer will be taken into account when moving the aforementioned sentence from the abstract to the Introduction as suggested by Referee#3. The authors thank the reviewer for this comment.

References:

Genest, C., Rémillard, B. and Beaudoin, D.: Goodness-of-fit tests for copulas: A review and a power study, Insurance: Mathematics and economics, 44(2), 199-213, 2009.

Hosking, J.R.M. and Wallis, J.R.: Regional frequency analysis: an approach based on L-moments, Cambridge University Press, 240 pages, 1997.

Ouarda, T.B.M.J.: Hydrological Frequency Analysis, Regional. Encyclopedia of Environmetrics, John Wiley & Sons, Ltd, 2013.

Salinas, J., Laaha, G., Rogger, M., Parajka, J., Viglione, A., Sivapalan, M., et al.: Comparative assessment of predictions in ungauged basins–Part 2: Flood and low flow studies, Hydrology and Earth System Sciences, 17, 2637-2652, 2013.

Wright, M.J., Houck, M.H. and Ferreira, C.M.: Discriminatory Power of Heterogeneity Statistics with Respect to Error of Precipitation Quantile Estimation, J. Hydrol. Eng., 04015011, 2015.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-136, 2016.