Review 28-05-2019

A WATERSHED CLASSIFICATION APPROACH THAT LOOKS BEYOND HYDROLOGY: APPLICATION TO A SEMI-ARID, AGRICULTURAL REGION IN CANADA

Jared D. Wolfe, Kevin R. Shook, Chris Spence, Colin J. Whitfield

I have read the authors' response to reviewers and their revised manuscript with interest. To me the revisions seem thorough and the manuscript has become much clearer as a result. I commend the addition of a sensitivity analysis of the clustering approaches in section 4.4. The editor has requested special attention to original comment 27, in which I ask the authors to evaluate the usefulness of their classification with independent data. This is addressed in point 6 in this review. My comments use the line numbers in the author response document (hess-2018-625-author_response-version1).

Response to reviewers

1. L430 (comment 5): the authors provide more detail about the accuracy of HydroSHEDs and state that "[…] the dataset provided an objective delineation over the region of interest and was sufficient for purpose of the current study." This argument would gain in strength if the authors can add how they came to this conclusion.
2. L474 (comment 8): the authors provide additional context for choosing the Thornthwaite PET method (which I think is justified) and also state a disadvantage of the method. It might be helpful to also include the practical impact of this disadvantage, because I don't quite understand.
3. L484 (comment 9): the authors provide reasons for not using any metrics related to snow in their study but acknowledge that this might be important. Is this mentioned anywhere in the manuscript? For example as a study limitation or an opportunity for further work.
4. L532 (comment 14): the authors provide a statement about the accuracy of the findings of Spence and Saso (2005). Is this accuracy dependent on the number of observations used? (Spence and Saso (2005) seem to use n = 34, compared to n = 11 in this paper). Addition: I see the authors have clarified this on L1267.
5. L555 (comment 17): the changed text in this response refers to Table 3, but the text in the manuscript refers to Table S3 (L1179).
6. L656 (comment 27): the authors provide evaluation against another data source (section 4.3) and include a sensitivity analysis of their clustering approach (section 4.4). The also point out (as reviewer #1 has mentioned) the rough correspondence between their clusters and the current understanding of eco-regions (section 5.1.2). The authors also comment that further evaluation is difficult due to the lack of data sources (e.g. L678).
   The authors state that (L928) "… those areas that are climatically and physio-geographically similar, and thus might be expected to respond in a hydrologically coherent manner to climate and land management changes." This is the critical assumption that underlies this clustering exercise. As I understand the manuscript, section 4.3 does not as much evaluate the entire classification, but only a part of it (wetland density). Further demonstrating that the defined clusters indeed respond in a coherent manner will add much more weight to this paper. However, if there is no data available than that is clearly not an option.
   If this is the case, the authors might want to further highlight the novelty of their work compared to the current understanding of eco-hydrology on the Prairies (e.g. the need for fuzzy treatment of watershed similarity as evidenced by section 4.4; the increased granularity possible with an approach such as the authors use, …).

7. L733: the authors provide more detail about how they scaled variables during the PCA. I'm bringing this point up again in relation to the text on L1326: "Climate and elevation gradients are likely responsible for the west to east watershed clustering pattern." I wonder to what extent this is forced by the data preparation, where these variables are log-transformed but not normalized. Is it possible that the log-transformed range of climate and elevation variables spans a wider range than that of the other variables? For example, if log-transformed mean precipitation has range [0,3] (assuming P = 1 to 800mm) it would span three times the range of a fractional variable with range [0,1]. This might skew the clustering procedure towards treating P and elevation as more distinctive attributes for each cluster. I don't believe this is necessarily a bad thing, for example if there are reasons to believe that P and elevation are relatively important. However, the authors also comment that "[…] if one is particularly interested in such variables, one should consider strategies to weight their importance." Is it possible that some form of weighting has already happened in the current manuscript as a result of only log-transforming the variables?

Investigation of the log-transformed ranges of each variable might indicate this. This would be a relatively low-effort check compared to re-doing the full clustering analysis with differently prepared data. If found relevant, this might be added to the discussion in L1595-1602.

Comments on revised manuscript

8. L877: "regime" > "regimes"?
9. L1033: I think the term "wet climate cycles" might still be confusing. Would "wet climate periods" be a suitable alternative?
10. L1208: it might be helpful to the reader to briefly summarize why wetland area distributions are simulated, if observations are also available (in the GSW data set). Am I correct in saying the GSW only gives the maximum wetland area, and the GPD simulation gives estimates of the full distribution of wetland sizes?
11. L1224: "4175 set" > "set of 4175 watersheds"
12. L1299: "TPC3" > "PC3"
13. L1425: "[…] less than 1." > "[…] less than 1 km^{-2}"
14. L1652: This paragraph might be better placed directly after (or as part of) the paragraph that starts on L1610.
15. L1683: Is the reference to Wagener et al, 2007 correct? I don't believe that paper talks about the relation between management practices and classification approaches. Perhaps this should be Wagner et al, 2007 (which I haven't read but its title suggests it as being more likely)?
16. L2127, Figure 6b: the number of points make this plot difficult to read. x-axis could be changed to cover the width of the page. Possibly cut of the y-axis at 10 for additional clarity.
17. Figure S1, c: text in the centre overlaps and is unreadable.