A novel algorithmic framework for identifying changing streamflow regimes: Application to Canadian natural streams (1966-2010)

Point-to-point Reply to Anonymous Reviewer 1

Masoud Zaerpour, Shadi Hatami, Javad Sadri, and Ali Nazemi

We greatly acknowledge the time and effort dedicated to evaluation of our manuscript by the Anonymous Reviewer 1 (AR1). We have reflected deeply on the thoughtful comments given by AR1, and now are ready to provide our point-to-point responses as of the following. In this response letter, comments provided by the AR1 are numbered and listed in the same order received and are in *Italic* fonts. Our responses are given in Plain text. Please note that some of the comments received discuss multiple issues. We attempt to address each issue individually by separating them from one another.

Part I. General comments

1. The analyses presented are interesting, and may be useful in establishing the changes in the regimes of Canadian rivers. Unfortunately, the results are undermined by their presentation. With some revisions, I believe that the paper can make a good contribution.

Response: We highly appreciate the positive and extremely constructive review comments given by the AR1. We are grateful that AR1 has found our work interesting and with practical merits.

2. The writing needs quite a bit of editing. There are too many grammatical mistakes to list here, and the writing is often unclear.

Response: Many thanks for your comment. We do acknowledge your concern. We have rigorously edited our paper to avoid grammatical mistakes. The revised version of our manuscript is now much improved. We also plan to ask a professional English editor to review our revised manuscript prior to submission to the journal.

3. There are missing articles in many of the sentences, such as the first one: Page 1, Line 28 "Natural streamflow characteristics have been critical consideration". This is sentence is missing the article "a" before "critical", or needs to make "consideration" a plural

Response: Many thanks for your comment. This is taken care of in the revised paper.

4. In many sentences there are disagreements in number, i.e. between singular and plurals: Page 2, Line 34 "some others determines"

Response: Many thanks for your comment. This is taken care of in the revised paper.

5. As was stated above, the writing is often unclear, as in the caption of Figure 9: "Figure 9. Mapping shifts in natural streamflow throughout Canada during 1966 to 2010. Rates of shift among various regime types in each stream are shown by shades of grey that quantifies how much decline in the giver regimes shown in the x-axes in each panel can result into incline in the receiver regime type corresponding with the column in which the panel is located. Columns filled with diagonal lines show the identical regime types with the receiving regimes identified in the column where the panel is located."

Response: Many thanks for your comment. This is now taken care of. Please note that in the quest for more clarity, we have revised Figure 9 with a set of six maps. Please see the **Figure R5 below** in response to the comment #9.

Part II. Specific comments

6. There appears to be only one gauging station within the Canadian Prairies. This is disappointing as the hydrology of the region is very important and has seen many effects of changes in climate. There are several RHBN stations within the prairies, according to this website https://www.canada.ca/en/environment-climate-

change/services/wateroverview/quantity/monitoring/survey/data-products-services/reference hydrometric-basin-network.html.

Whitfield et al. (2020) grouped responses of streams into 3 clusters in the Prairies and adjacent areas, using a very different clustering methodology. I assume that there were no other prairie stations which met the authors' criteria. However, it would be good to have this explained. Would the use of a slightly different analysis period have allowed the inclusion of more prairie streams?

Whitfield, P.H., Shook, K.R., Pomeroy, J.W., 2020. Spatial patterns of temporal changes in Canadian Prairie streamflow using an alternative trend assessment approach. Journal of Hydrology 582, 124541. <u>https://doi.org/10.1016/j.jhydrol.2020.124541</u>

Response: Many thanks for your comment. In the first step of our study, we did a rigorous analysis to accommodate as many RHBN stations as possible with the longest common period. In fact, as noticed also by Whitfield et al. (2020), the Prairie region does not include many RHBN stations with long-term and continuous data record. Unfortunately, we could find only two stations in the prairie region during 1966 to 2010 that met our data criteria, i.e., having a continuous daily record with less than ONE month worth of missing data in a typical hydrologic year. To address your comment, we altered our data period to 1976-2010 and repeated our search for new stations in the Prairie region and accordingly our clustering analysis. This new effort has resulted into consideration of NINE new stations in the Prairie region. We compared the result of our new analysis with the results we obtained with previously selected RHBN stations in the Prairie region, namely, Waterton River near Waterton Park (05AD003) and Belly River near Mountain View (05AD005) during 1976-2010. Figure R1 below summarizes our findings. Left panel shows the clustering results related to the new NINE stations, i.e. S1 to S9 in comparison with the two previously selected stations, i.e. 05AD003 and 05AD005. The right panel shows the analysis of trends in anomalies of decadal memberships, in which the stations in each ecozone are sorted from the lowest to highest elevations from the top to the bottom. This analysis shows that the three stations in the southwest of prairie are mostly of type C2 (i.e., fastresponding/summer peak regime; previously named as Naivo-glacial – see our response to your 10th comment below). The station S8 in the northwest and S1 in the east of prairie are mostly belong to C1 (i.e., slowresponding summer regime – previously named as Glacial). The rest of stations are mostly belong to the C3 and C4. The analysis of trends in anomalies of memberships shows mainly decreasing trends in belongingness to C1 and C2 clusters and increasing trend for C5 and C6. This analysis will be included in our revision as a part of our Discussion section.



Figure R1. Validation of the proposed clustering approach in the Canadian Prairies using nine new and two previouslyselected RHBN stations during 1976 to 2010. The color bar in the left map show the degrees of membership to each cluster. The right panel show the trends in the degree of membership in the six clusters at the considered 11 stations. Positive and negative trends are shown with red and blue colors, respectively. Sharp colors show significant cases. New stations S1 to S9 are sorted from the lowest to the highest elevations from the top to the bottom.

7. Although hydrologists are used to working with river basins, grouping the stations by basin is not always useful. As shown in Table 2, Canadian river basins are very large. Wong et al. (2017) identify 15 ecozones in Canada, many of which are spanned by single basins. For example, the Nelson River system spans the Montane Cordillera, Prairies, Boreal Plains, Canadian Shield and the Hudson Plain. Stations in differing ecozones would not be expected to behave in similar ways, given that their elevations, geologies, topographies, vegetations and climate forcings are very different, even if they are within the same basin.

Wong, Jefferson Razavi, Saman Bonsal, Barrie Wheater, Howard Asong, Zilefac Elvis. (2017). Intercomparison of daily precipitation products for large-scale hydro- climatic applications over Canada. Hydrology and Earth System Sciences. 21. 2163- 2185. 10.5194/hess-21-2163-2017.

Furthermore, many ecozones are split among several basins. The Montane Cordillera stations are divided among the Nelson, Peace-Athabasca, and Fraser basins. These stations would be expected to show some similarities, although local conditions would also apply. It would be very useful to have the ecozones superimposed on the maps. It would also be useful to take the ecozones into account when grouping the analyses.

Response: Many thanks for this very insightful comment. We have done a major revision in our manuscript to address your comment and a similar comment raised by Anonymous Reviewer 2 (AR2) regarding the appropriateness of the basin/sub-basin system for integration and discussion of our results. In the revised manuscript that will be provided upon the approval of the editor, we will present and discuss our results entirely at the ecozone scale except only in positioning against earlier studies (currently Section 5.1 in the submitted manuscript). We should also mention that our revised manuscript has significantly improved and the result can be much better interpreted by moving to the ecozone scale. To just demonstrate here how

mappings look like with consideration of ecozones, below in **Figure R2** and **Table R1** (these are counterparts of Figure 4 and Table 1 in the submitted manuscript, respectively), we show how the considered RHBN stations are situated and distributed within the 15 Canadian ecozones. We have also shown in **Figure R3** the results of our clustering analysis at the ecozone scale. This is equivalent to the results we provided in Figure 6 in the submitted manuscript but this time at the ecozone scale. As it can be seen, the results are much more interpretable at ecozone scale. We highly appreciate AR1's constructive comment.



Figure R2. The distribution of the selected 106 RHBN streamflow stations within the Canadian ecozones.

Table R1. List of Canadian ecozones with at least one RHBN station in this study, along with their abbreviations and the number of RHBN stations considered within each ecozone.

Abbreviation	Ecozones	# of stations	Abbreviation	Ecozones	# of stations
EZ2	Northern Arctic	1	EZ8	Mixedwood Plains	5
EZ3	Southern Arctic	1	EZ9	Boreal Plains	6
EZ4	Taiga Plains	1	EZ10	Prairies	2
EZ5	Taiga Shield	4	EZ12	Boreal Cordillera	7
EZ6	Boreal Shield	25	EZ13	Pacific Maritime	9
EZ7	Atlantic Maritime	26	EZ14	Montane Cordillera	19



Figure R3. The distribution of the identified regime types across Canadian ecozones. Sizes of circles are proportional to the membership degrees that quantify the association of the streams to regime types C1 to C6. Only streamflow stations with degrees of membership of 0.1 or higher are shown in each panel. The red stars are the archetype stations related to each regime type.

8. Line 338 "...the Arctic has the least diversity in the streamflow regime. All considered 12 streams are associated with large degrees to glacial regime, out of which five and six streams show increasing and decreasing trends in the membership, respectively."

The fact that half of the streams in the basin change in each direction is confusing. Does this imply that the changes are not a result of climate shifts, but rather of short- duration weather trends? Or is it that the streams are in different climatic zones?

Response: Many thanks for this additional and very constructive comment. After performing the analysis of trends – based on your previous comment – at the ecozone-scale, it becomes clear that different directions of trend in one drainage basin correspond mainly to different ecozones – see **Figure R4**. Specific to your

comment, Figure R4 shows that all gauging stations in the Arctic basin that are located in the Taiga Shield show declining degrees of membership to cluster C1 (named previously Glacial, in the revised version named as slow response/summer peak regime). In contrast in the Boreal Plains, degrees of membership to C1 increase. While the ecozone classification, to a great extent, determines how the streamflow changes inside each drainage basin, we can see further variations in the direction of trends in membership values inside each correspond to elevation – see **Figure R5**. Figure R5 is the same as Figure 8 in the submitted manuscript, but rows are rearranged based on ecozones, and are sorted in each ecozone by elevation. Please note that all our discussions and framing in the revised paper is based on ecozone not the basin/sub-basin system, except in Section 5.1.



Figure R4. The trends in the degrees of membership to each regime type in 106 considered RHBN streams over the period of 1966 to 2010. Colors show the direction of Sen's slopes in the anomalies of memberships. Positive and negative trends are shown with red and blue colors, respectively. Bigger sizes of triangles show significant trend cases (*p*-value ≤ 0.05).



Figure R5 The evolution in degrees of membership to each regime type in 106 considered RHBN streams along with the corresponding Sen's slopes. For each stream, shades of grey show decadal memberships over the period of 1966 to 2010. The color bar shows the direction and significance of the Sen's slope of the trend in the anomalies of memberships. Positive and negative trends are shown with red and blue colors, respectively. Sharper colors show significant trends. The RHBN stations at each ecozone are sorted from lowest to the highest elevations from the top to the bottom.

9. It would be very useful to have a map, or maybe more than one map, of the sites showing their changes in regime type. This would allow the reader to see if the changes are spatially related. Again, it would be very useful to have the ecozones superimposed.

Response: Many thanks for your comment. As mentioned in our responses to your 7th and 8th comments, we majorly revised the presentation and discussion of our result at the ecozone scale. To address your concern, both Figure 8 (see <u>Figure R4</u> above) and Figure 9 (see Figure R6 below) are now presented as a sets of six maps, in which ecozones are the spatial units for performing the analysis and discussing our results.



Figure R6. Mapping shifts in natural streamflow throughout Canada during 1966 to 2010. In each panel, colors show the clusters from which transition happen to the reference cluster. Rates of shift in regime types for each stream are shown by the size of each circle.

For better communication of the forms of transitions in streamflow regimes, we also add a set of Sankey diagrams, showing how the streamflow regime in our considered RHBN streams transform from one regime type (left side in each panel) to another type (right side in each panel). Again, the gauging stations are grouped based on ecozones. Colors are related to the six regime types. Width of arrows are proportional to rates of shift. Stations within each ecozone are sorted from the lowest to the highest elevation from the top to the bottom – see **Figure R7** below.



Figure R7. Sankey diagrams showing various transitions in streamflow regimes throughout Canadian ecozones during 1966 to 2010. Each panel is related to one target cluster located in the right side of the panel. Streams at the left side of each panel are grouped in ecozones and are sorted in each ecozone from the lowest to highest elevations from the top to the bottom. Colors show the six regime types. The width of arrows are proportional to the rate of shift.

10. The "glacial" type is problematic. Looking at Figure 6, at least 16 of the "glacial" basins cannot include any glaciers at all, as they are not in mountains. No doubt many of the mountain basins do not contain glaciers, either. The same issue is true of the "niveo-glacial" type. I understand that the authors are using the term "glacial" to refer to the shape of the cluster's annual hydrograph, but the term is confusing. Worse, the authors are grouping together streams with very different causes for their behaviours.

Response: This is another great comment from AR1 that we highly acknowledge. Yes! Our initial idea behind the naming of our clusters was to point at the shape of the annual hydrograph rather than the place where streams are initiated from. To address your very relevant comment and to avoid any potential confusion, we changed the naming of our regime types to exclusively point at the shape without any referral to the source of the streamflow. Our new naming system is based on two key characteristics, namely the timing of the peak flow (i.e. cold season peak, freshet peak and summer peak) and the form of hydrologic response (i.e. fast *vs.* slow). The form of hydrologic response can be identified by streamflow variability, with high variable annual streamflows being identified as fast response. Using this new system, our previously termed "glacial" regime will be now "slow response/summer peak", comprising streams with very strong seasonality and high discharge in summer and relatively less variation throughout decadal timeframes compared to the C2, which was previously termed as "nivo-galacial" regime is now named as

"slow response/freshet peak" regime and includes streams in which the annual streamflow volume is mainly contributed by the spring snowmelt, showing high discharge in spring with relatively less variation in the hydrograph compared C4, previously "nivo-pluvial" shape of to named as regime. C4 is now named as "fast response/freshet peak" regime, showing more variation in the shapes of hydrographs. C5 regime, previously termed "pluvio-nival regime", is now named as "slow response/cold season peak" regime, which comprises streams with weak seasonality with slightly more discharge in fall and winter seasons. C6, previously termed as "pluvial" is now named "fast response/cold season peak" regime, showing higher variation in shapes of decadal hydrographs. In our revised version, we will keep some of the discussion on the potential sources of the streamflow but we will avoid terming the regime type by names that can resemble the source of streamflow. We believe applying this comment has considerably improve the clarity in our revised version.

11. The source of the archetypal "glacial" stream, Kazan River above Kazan Falls, is in northern Saskatchewan, where there are no glaciers. Looking at Figure 7, the main difference between the "glacial" and "nivo-glacial" types would appear to be that the former has a shallower recession limb. According toen.wikipedia.org/wiki/Kazan_River:

"The river headwaters are in northern Saskatchewan[7] at Kasba Lake... Along its course the river flows through several lakes, including Ennadai Lake and Yathkyed Lake."

So the cause of the shallow recession limb is almost certainly storage within the lakes in the basin.

Response: Many thanks for your comment. You are absolutely right! We believe the new naming we suggested above can avoid this problem as C1 regime, i.e., "slow response/summer peak" regime can refer to the flows initiated from both glaciers and/or lakes. In addition, we separated the glacier-fed stations from lake-dominated. Please see **Figure R8** below in response to your 13th comments.

12. Line 245: "Architype (sic) streams are those streams that have the highest association to the identified regime types and can represent the characteristics of a given regime better than other members of the cluster."

As the Kazan River is controlled by lakes, it would be very difficult to transition to another cluster type. I see that many of the "glacial" and "niveo-glacial" streams lie within the Canadian Shield. Are many of these also dominated by lakes?

Response: 14 out of 16 stations located in the Canadian Shield which are pointed out by AR1 in comment #10 are lake-dominated. Please see **Figure R8** and the discussion below for further information on how glacier-fed streams are distinguished from lake-dominated ones.

13. Where an unglaciated stream transitions between "glacial" and "niveo-glacial" types, as in the Hudson Bay and Arctic Seaboard basins, it cannot represent a change in the glacial contribution. It is therefore important to separate those basins containing glaciers, from those which do not.

Response: Many thanks for your comment. You are right. This confusion should be greatly resolved now by the new naming. When we see a transition from C1 to C2, it refers to increases in variability of the decadal

streamflow. To specifically address your comment, we have used additional data, i.e. GLIMS and NSIDC Glacier Database (https://doi.org/10.7265/N5V98602), to separate streams with different types of origin. In particular, please see **Figure R8** below in which for "slow response/summer peak" regime type (i.e. C1), glacier-fed streams are distinguished from lake-dominated by blue circles and red squares, respectively. The discussion on the source of streams are included in the revised manuscript.



Figure R8. Streams with "slow response/summer peak" regime type (C1) with different streamflow origins. Glacier-fed and non-glaciated streams are separated from one another with blue circles and red squares, respectively.

14. Where there are glaciers, and the stream transitions from "niveo-glacial" to "glacial", would this imply an increase in glacial contributions? If so, would this be justified by what we understand about glacial hydrology?

Response: Many thanks for this extremely thoughtful question. We have not directly explored the role of glacial contribution by analyzing glacial coverage/storage data; however we believe that transition from C2 to C1, i.e., "fast response/summer peak" to "slow response/summer peak" regimes (previously termed as "niveo-glacial" and "glacial", respectively) means more streamflow contribution from glacial retreat rather than the annual snow pack. This means that response of the annual snow pack is lessened and more glacial retreat take place from glacial storage. We believe that this can make a potential hypothesis, which should be investigated in more details using glacial coverage data in an independent study, clearly remaining beyond the scope of this study. Having said that, we added this hypothesis to the discussion in the revised paper and suggested this as a potential area for future exploration.

15. Line 244 "Figure 5 summarizes the results, showing c = 6 as the optimal number of clusters.". I think this needs to be explained in more detail. Why is 6 the optimum number of clusters? I can see that the indices become quite flat around c = 6, but what do the indices mean, i.e. are small index values better (this is not explained)? If so, why not use c = 7, as it looks to be slightly better for the Separation Index and the Xie and Beni Index? Is there a reason why it is advisable to use fewer clusters?

Response: Many thanks for this comment. We used the elbow method for determining the optimal number of clusters (e.g., Zhao et al., 2008; Satopaa et al., 2011; Kuentz et al., 2017). In fact, we chose c=6 as the cutoff point, because while the separation indices does continue to decrease slightly, the added complexity cannot be justified. Please note that adding a new cluster in this case means identifying 30 new parameters for the cluster center. When c=6, a significant change in the slopes of all three validity indices take place from steep to shallow (i.e., an elbow) and the curves become flat, indicating that adding more cluster has only insignificant marginal values. We also noted that adding to the number of clusters leads to ambiguity in our results as there would be less significant difference between regime types.

Zhao, Q., Hautamaki, V., and Fränti, P.: Knee point detection in BIC for detecting the number of clusters. In International conference on advanced concepts for intelligent vision systems (pp. 664-673). Springer, Berlin, Heidelberg, https://doi.org/10.1007/978-3-540-88458-3_60, 2008.

Satopaa, V., Albrecht, J., Irwin, D., and Raghavan, B.: Finding a" kneedle" in a haystack: Detecting knee points in system behavior. In 2011 31st international conference on distributed computing systems workshops (pp. 166-171). IEEE, https://10.1109/ICDCSW.2011.20, 2011.

Kuentz, A., Arheimer, B., Hundecha, Y., and Wagener, T.: Understanding hydrologic variability across Europe through catchment classification, Hydrol. Earth Syst. Sci., 21, 2863–2879, https://doi.org/10.5194/hess-21-2863-2017, 2017.

16. Figure 7 is useful to demonstrate the differences among the clusters. It would be extremely useful to see similar plots indicating the cluster transitions. For example, what does it look like when the streams transition from "glacial" to "niveo-glacial", or vice-versa? Because so many climate signals are used, it is not easy to see how the changes in the hydrograph relate to the transition from one cluster to another.

Response: Many thanks for this comment. To address your concern and to better demonstrate transitions between regime types, we revised our visualization approach in different figures in the revised version to explicitly illustrate how changes in regime types correspond to changes in streamflow characteristics. As an example in **Figure R9** below, we provide three examples for types of transitions between clusters along with their attribution to changes in streamflow features. The transitions from the earliest (1966 to 1975) to the latest (2001 to 2010) episodes are shown with grey and pink envelops in the left panels. The right panels show the rate of shift as well as the attribution of regime shift to the key changes in streamflow characteristics.



Figure R9. Three samples for transitions between regime types along with their corresponding changes in streamflow characteristics. Panels in left show shape of hydrographs in the earliest (1966 to 1975) and the latest (2001 to 2010) decadal episodes using grey and pink envelopes, respectively. Panels in right show shifts in the flow regimes attributed to the changes in streamflow characteristics. In each right panel, left envelopes show the transition from multiple regime types to one clusters but one path is the dominant and is visualized through the thickest grey envelope, proportional to the rate of shift. Right envelopes show the association of the identified transition to changes in streamflow characteristics using \mathbb{R}^2 .

Part III. Technical comments

17. How were the calculations performed? I assume that some software was used. It should be credited and described. If possible, the software should be made available for others to test and use.

Response: Many thanks for your comment. We have coded up everything in MATLAB. The indicators of hydrological alteration are described at Ritcher et al., (1996). The procedure of fuzzy clustering method is based on the algorithm proposed by Bezdek (1981) and it is carried out in MATLAB using *fcm* function. The formulation of three validity indices of Xie-Beni index, separation index, and partition index can be found in Xie and Beni (1991) and Bensaid et al., (1996). The implementation of these three indices are also in MATLAB software. All other procedures, i.e. trend analysis, attribution etc. are also coded in MATLAB.

We would be very happy to provide our in-house MATLAB codes developed throughout this paper with our

revised manuscript. We wait for further guidance by the editor and the journal for providing our computational codes.

Bensaid, A.M., Hall, L.O., Bezdek, J.C., Clarke, L.P., Silbiger, M.L., Arrington, J.A. and Murtagh, R.F.: Validity-guided (re) 635 clustering with applications to image segmentation. IEEE Transactions on fuzzy systems, 4(2), pp.112-123, https://doi.org/10.1109/91.493905, 1996.

Bezdek, J. C.: Pattern Recognition With Fuzzy Objective Function Algorithms, Plenum, New York, https://doi.org/10.1007/978-1-4757-0450-1_3, 1981.

Richter, B.D., Baumgartner, J.V., Powell, J. and Braun, D.P.: A method for assessing hydrologic alteration within ecosystems. Conservation biology, 10(4), pp.1163-1174, https://doi.org/10.1046/j.1523-1739.1996.10041163.x, 1996.

Xie, X.L., Beni, G.A.: Validity measure for fuzzy clustering. IEEE Trans. PAMI 3(8), 841–846, https://doi.ieeecomputersociety.org/10.1109/34.85677, 1991.

18. I believe that "architype" is a misspelling of "archetype"

Response: Many thanks for your comment. We fixed this typo.

19. Figure 5: The x axis labels are misspelled – "Numer" should be "Number".

Response: Many thanks for your comment. We fixed this typo.

20. Figure 7: In the interests of space, it would be a good idea to omit the periods in the x-axis label month names, and also the x-axis title "Month". The caption refers to the "expected" annual hydrograph. What does this mean? Are these the mean (or median) weekly values? The y-axis label is in "mm/week-1", i.e. in mm x week. Obviously this is incorrect.

Response: Many thanks for your comment. In Figure 7, the solid black and red lines are the mean of grey (1966-1975) and pink (2001-2010) envelops, respectively. These indicate the expected annual hydrograph in each decadal period, which is calculated, as you noted, by the mean of weekly values. That is why the caption refers to the "expected" annual hydrograph. As you suggested, the title of x-axis is removed and months are shown with one letter. The y-axis label is revised as ""mm week⁻¹". Below, please see the revised Figure 7.



Figure 7 (revised). Alterations in the decadal streamflow regimes at the archetype streams through time. The envelopes of 10-year annual hydrographs for the earliest (1966 to 1975) and the latest (2001 to 2010) episodes are shown with grey and pink colors, respectively. The expected annual hydrographs during the earliest and the latest 10-year periods are shown in solid black and red lines. The change in the membership degree of each archetype stream is shown within parentheses.