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

**Point-to-point Reply to Anonymous Reviewer 2**

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 2 (AR2). We have reflected deeply on AR2's thoughtful comments, and are ready now to provide our point-to-point responses as of the following. In this response letter, comments provided by the AR2 are numbered and listed in the same order received and are shown in *Italic*. Our responses are given in Plain texts. Please note that some of the comments received discuss multiple issues. We attempt to address each of these issues individually by separating them from one another.

1. *The paper presented by Zaerpour and colleagues proposes a new framework for identifying shifts in streamflow regimes with a subsequent application to Canadian streamflow. While the methodical challenge and the chosen case study are clearly interesting, the convoluted nature of both the methodology and the presentation of the result prevent me from recommending the paper for publication in its present form. Below, I summarize my main points that need to be clarified.*

**Response:** Many thanks for your thoughtful and positive review. We are grateful for receiving your comments and have done all of our responsible efforts to address them in the best way possible. In the revised manuscript, we explained our methodology in more details and clarified points raised by you and Anonymous Reviewer 1 (AR1). Since we shifted the discussion now from drainage basins into ecozones, according to a series of comments we received from you as well as AR1, the presentation of our results and framing our discussions have been majorly changed. Furthermore, we added a section to address your legitimate concern regarding the stability of the clustering results to chosen timeframes. We believe our revised manuscript, carefully evolved based on your comments and AR1's, is much improved.

**Part I. Methodology**

2. *Comment on clustering and change detection: The authors propose to use a fuzzy clustering algorithm to first identify groups of stations (or degree of membership of each station to each group) within a matrix of Indicators of Hydrologic Alterations (IHA) for a given – short – time window. Subsequently the degree of membership if each station to each class is computed for further timeframes. Overall, I can follow this approach and on first reading the description of the methodology makes sense (note a minor issue mentioned below). However, I have several questions regarding the choice of this particular approach and the stability of the analysis:*

**Response:** We appreciate AR2's attention to the methodological aspect of our study. We have added more detailed explanations in the methodology section to clarify the detailed questions you pointed in the following. We have also majorly extended on Section 5.2 and included a brand new analysis for the stability of the results to the choice of the timeframe. Our key findings in this regard are shared below.

3. *How stable is the estimate? The results of the analysis are crucially dependent on the identification of clusters in the first period. However, the streamflow climatologies (or IHAs) used for estimating these clusters are only computed using a small fraction of the available data, which is likely to yield unstable estimates. As a consequence investigating shifts in these clusters may be confounded by estimation errors. For example, I wonder if the authors would reach the same conclusions if clusters would have been identified using another period. Since the paper does not report on the stability of the estimate, it is hard to evaluate whether the overall conclusions are affected by the arbitrary choice of the first time window for identifying clusters (e.g. why not use the last window or one in the centre). Approaches for combatting this issue could be to (a) use all time windows for identifying the clusters and subsequently assessing how the degree of membership of each station to each cluster changes over time or (b) repeating the analysis with clusters identified for each time window and report the associated spread.*

**Response:** Many thanks for your thoughtful comment on the issue of stability. We did look into the uncertainty in our results, although from a different angel, and investigated the stability of the result to the length of timeframe in Sect. 5.2, particularly through **Figure S1** in the supplement. Figure S1 clearly shows that the centers of the clusters do not change significantly by the length of timeframe. While the uncertainty of our results due to the length of the timeframe was investigated thoroughly (please see the rest of discussion in Sect. 5.2 as well as the materials provided in the supplement), your comment have opened up a new way of looking at the stability of our results. By performing your suggested sensitivity analysis, see **Figure R1** below, it is now clear that cluster centers do not significantly change by altering the decadal timeframe in which the clustering is made. In Figure R1, we used all possible decadal timeframes throughout the study period to recalculate the cluster centers. The dots show the centers of clusters scaled into two dimensions using Multidimensional Scaling (MDS; Cox and Cox, 2008). Black crosses show the centers of the first decadal timeframe. The result shows that distinctions between clusters are clearly maintained, despite changing the decadal timeframe chosen for identifying the cluster centers.

Nonetheless, we feel necessary to mention that the choice of the first decadal timeframe is not arbitrary in our work, as we do a formal trend analysis on the membership values. Obviously in this context and for the purpose of understanding the evolution in the streamflow regime using the trend analysis, we should start from the first timeframe and finish in the last one and go systematically throughout all other possible timeframes in between. Our moving window methodology is particularly designed to address this.

Cox M. and Cox T.: Multidimensional Scaling. In: Handbook of Data Visualization. Springer Handbooks Comp.Statistics. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-33037-0_14, 2008.

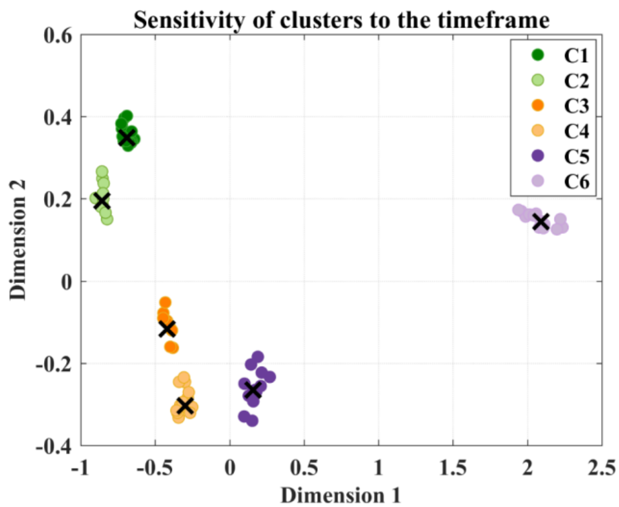


**Figure R1.** The stability of the cluster centers to the choice of decadal timeframe for clustering. The dots show the two dimensional scaling of the cluster centers based on the relative distance of cluster centers from one another. Black crosses show the centers identified by choosing the first decadal timeframe.

4. *What is the benefit over a classical EOF/PCA analysis? Technically the analysis has distinct similarities to applications of dimension reduction methods such as Principal Component Analysis (PCA)/Empirical Orthogonal Function (EOF) analysis or Multidimensional Scaling (MDS) to spatially distributed time series. Of course, these methods do not evolve around the idea of "clusters" but identify modes of similar variability, but the strategy to first identify a membership matrix (analogue to "leading EOF patterns") which are then projected onto individual stations. In the EOF/PCA world, an analogue approach would yield a filtered time series at each station in which then again could be used to assess regime shifts without the need of developing a new (and somewhat convoluted methodology).*

**Response:** Many thanks for your comment. As you noted, the key conceptual difference in our methodology is the consideration of intersected clusters to describe regime types. Through the use of our clustering-based approach, regime types can be identified using empirical data in a fully self-organizing manner and we are able to measure how degrees of belongingness to each cluster change through trend analysis. The beauty of our methodology is in its absolute transparency and the fact that transitions in both regime types (clusters) and their corresponding hydrograph shapes and/or IHAs are fully traceable and can be explicitly linked to one another. Although approaches such as PCA/EOF are informative, they cannot provide such an opportunity. For example, components in PCA analysis are combinations of several IHAs, whereas we directly link the changes in clusters to individual IHAs and measure the strength of such links using rates of shift and the coefficient of determination. Unlike the EOF/PCA approach, which assesses the shifts in each individual stations, combining the clustering algorithm with the moving-window technique provides an opportunity to identify the shifts in the flow regime across the whole country in a fully integrated way using a set of specific reference points, i.e., cluster centers in the first time episode. We believe these strengths and added benefits justify the use of our proposed methodology although it might not seem as straightforward as the use of a method such as PCA. We discussed these points in the revised manuscript.

5. *Why use a combination of Kendall's tau and R2 for the attribution work? To me it appears to be a bit convoluted to use two very different metrics (Kendall's tau and R2) for the attribution work. While Kendall's tau operates on ranks and is thus less sensitive to non-linarites or outliers, R2 is in essence a linear metric. As an alternative single metric I could e.g. imagine to rely on Spearmans rank-correlation coefficient together with a simple test of significance thereof.*

**Response:** Many thanks for your comment. We did not combine these two metrics nor the concept behind the Kendall's tau dependence and coefficient of determination. In fact, we use these two metrics for two different purposes: On the one hand by using Kendall's tau, we identify the sign and significance of dependencies between changes in membership degrees and changes in streamflow characteristics. On the other hand, we use coefficient of determination, $R^2$, to quantify how much of the variability in a given set of membership degrees can be described by changes in a specific streamflow characteristics. By using these two measures together, we not only provide a formal approach to assess the dependencies between changes in membership degrees and streamflow characteristics (through the use of Kendall's tau with formal *p*-value), but also we can facilitate quantitative communication of the impact of changes in a specific streamflow characteristic on transition from one regime type to another. We added this clarification in our revised manuscript.

6. *Why not just look at trends in time series of monthly means and timing indicators? The analysis revolves around what the authors refer to as Indicators of Hydrologic Alterations (IHA). These are essentially a: The annual mean. b: the mean of each month and c: the timing of low/high flows. While reading I wondered if it would not have been sufficient to simply show maps of the trends of each of these metrics to arrive at the same conclusions?*

**Response:** Many thanks for your comment. There are two issues here. First, while we indeed looked into the expected values of annual and monthly mean flows as well as timings of low/high flows, we also looked into the variability of IHAs as well. While the majority of current literature limit the analysis of change to expected values of IHAs, there are strong evidences, particularly in Canada that changes in the variability of streamflow characteristics can be as important as changes in the expected values – please see Table 4 in our submitted manuscript for some of the already established findings.

Second, while looking at the individual trends in mean and variability of IHAs can be informative, it has certain limitations. First and foremost, looking solely at the trends does not provide any information on how the combination of streamflow characteristics can lead into formation of specific regime type. In addition, while it would be very difficult to look into trends in 30 different characteristics, we only look into the trends in six membership values and we know that an increasing trend in membership of one regime type will inevitably translate to decreasing trends in membership values for at least another cluster. Finally, through the use of our proposed approach, it would be possible to formally relate the changes in the streamflow characteristics to changes in regime types. None of these would have been possible by looking at the simultaneous trends in the individual streamflow characteristics only.

7. *Why stratify the analysis along large drainage basins. While I acknowledge the tradition of stratifying the analysis of streamflow data along drainage basins I wonder if this is the ideal choice in this particular instance. The regime classes identified by the authors essentially reflect different climatological regions (e.g. colder, snow dominated vs. warmer, rainfall dominated). Continental-scale drainage basins typically cover large climatic gradients and apart from the case where stations are hydrologically connected (how many of them are?), we would not expect a-priory the drainage basin would have much explanatory power on the climatology. Alternatively, I could imagine an assessment of changes in the underlying climate drivers (e.g. temperature, precipitation) would contribute to a deeper understanding of the associated changes.*

**Response:** Many thanks for your insightful comment. A similar comment was also given by AR1 regarding the suitability of the basin/sub-basin system in discussing/framing our results. Accordingly, we majorly revised our manuscript and considered ecozones (Wiken, 1986; Lespinas et al., 2015) as units in which we frame our results, as oppose to the drainage basins/sub-basins we used in our initial submission. This effort has significantly improved the presentation and interpretation of our results. We have also considered climate regions provided by Environment and Climate Change Canada, but we figured that ecozones provide the most suitable unit for discussing our results, as ecozone not only considers climatic factors but also geology, soil characteristics, vegetation, topography, etc. (Wong et al., 2017). Just as an example, **Figures R2 and R3** as well as **Table R1** below show the distributions of the considered RHBN streams as well as the result of our clustering analysis presented at ecozone scales.

Regarding the hydrological connectivity of the considered RHBN streams, we did a rigorous analysis over all selected stations to figure out if the stations are connected. For this purpose, we used the HYDAT data along with the National Hydrographic Network (NHN) provided by Natural Resources Canada to determine the streamflow network and the flow direction at each sub-basin. Accordingly, we realized that there is only one pair of stations (i.e., 01AD002 and 01AD003) located in Saint John- St. Croix sub-basin that are hydrologically connected. To keep the consistency of our analysis, we will exclude 01AD003 from our revised manuscript, so that we can come up with 105 streams that are hydrologically independent from one another.

Lespinas, F., Fortin, V., Roy, G., Rasmussen, P., & Stadnyk, T.: Performance evaluation of the Canadian precipitation analysis (CaPA). Journal of Hydrometeorology, 16(5), 2045-2064, https://doi.org/10.1175/JHM-D-14-0191.1, 2015.

Wiken, E.B.: Terrestrial Ecozones of Canada. Ecological Land Classification, Series No. 19. Environment Canada. Hull, Quebec. pp. 26, 1986.

Wong, J. S., Razavi, S., Bonsal, B. R., Wheater, H. S., and Asong, Z. E.: Inter-comparison of daily precipitation products for large-scale hydro-climatic applications over Canada, Hydrol. Earth Syst. Sci., 21, 2163–2185, https://doi.org/10.5194/hess-21-2163-2017, 2017.

**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 |

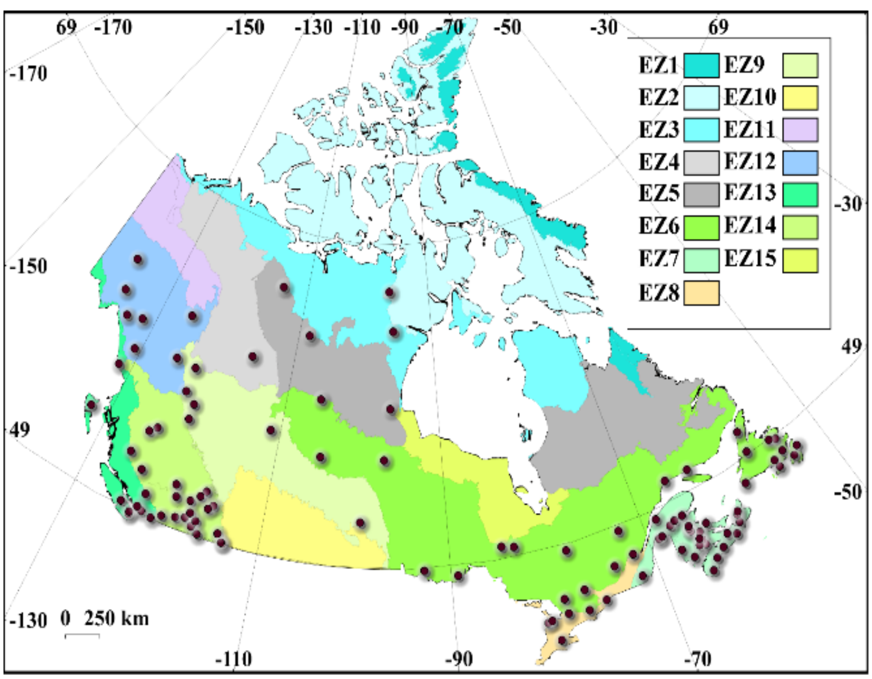


**Figure R2.** The distribution of the selected 106 RHBN streamflow stations within the Canadian ecozones.

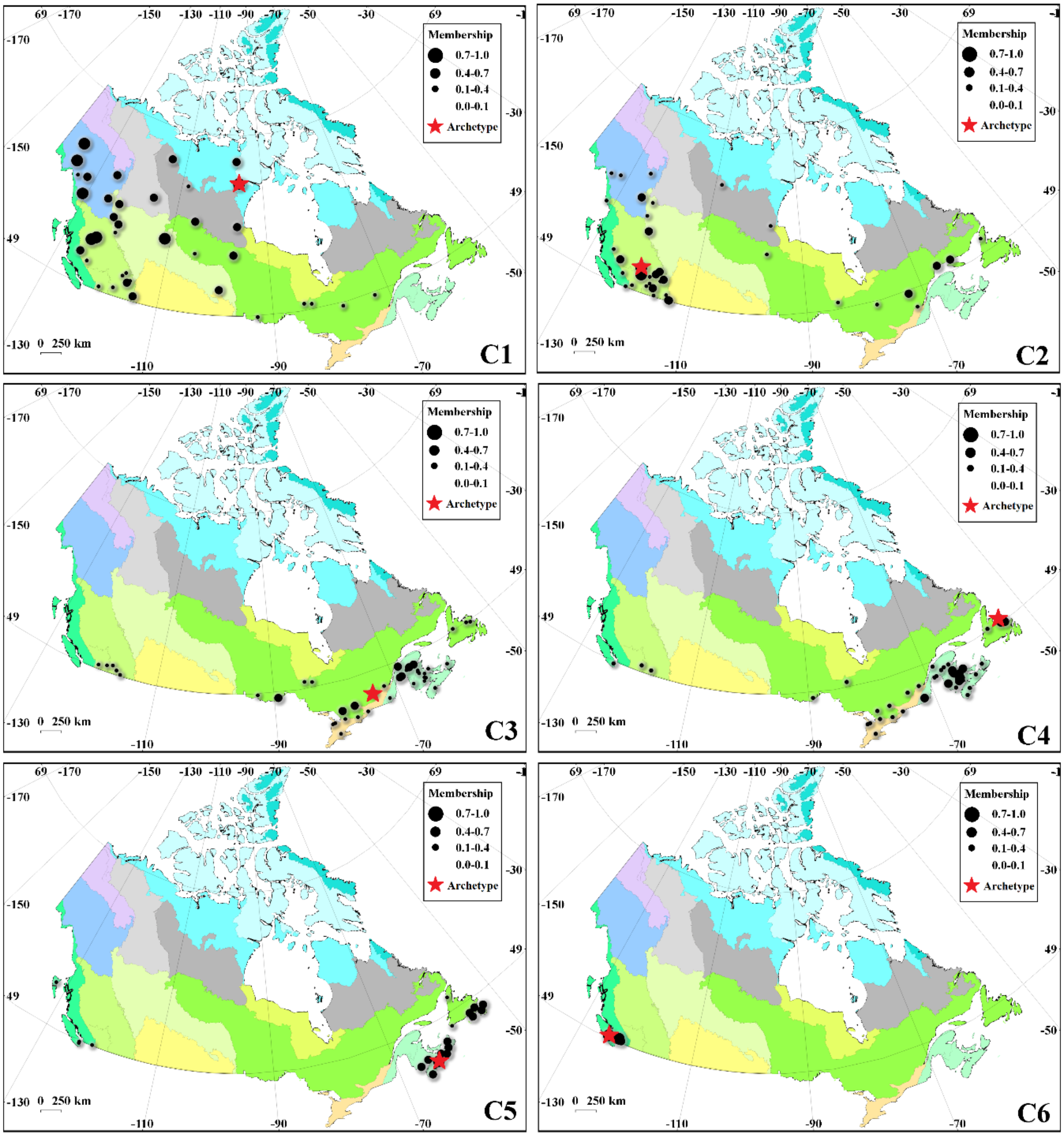


**Figure R3**. The distribution of the identified regime types across Canadian ecozones. Sizes of circles are proportional to the membership degrees that quantify the association of the streams to regime types C1 to C6. Only streamflow stations with degrees of membership of 0.1 or higher are shown in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

## Part II. Presentation

8. *Overall, I found the paper too long, a bit convoluted and therefore cumbersome to read. Some reasons:*

**Response:** Many thanks for your constructive comment. We have rigorously revised our manuscript to address your comment and a similar comment raised by AR1 regarding the presentation of our results. Now that we extended on the methodological aspect of our study, strategically shifted the point of discussion from basins to ecozones (and accordingly changed our figures) and focused more on key take-home messages, we believe that our results are easier to follow.

9. *The key selling point advertised in the title (i.e. the methodology) is featured in about 10% (4 of 39 pages) of the article and the properties of the methodology (e.g. stability or relation to alternative techniques) are neither assessed and nor discussed.*

**Response:** Many thanks for your comment. While we wanted to be very clear in what we do methodologically and how we do it, we wanted to also be concise and convey the majority of the methodological details through referencing. In the revised version, we indeed extended on explaining our methodology and our text in methodology has increased by around 20%. First and foremost, we have added the analysis of stability formally as part of our methodology. In addition, we have discussed the pros and cons of the proposed fuzzy-based framework compared to the alternative approaches e.g., EOF/PCA and/or other clustering methodologies in the Rationale section. Finally, we extended on the analysis related to validation indices and explained the elbow/knee method briefly, with which the optimal number of clusters are found.

10. *The description of the results is very convoluted and I found it difficult to extract the key message upon first reading. For example, I would value if the results description would focus on overarching patterns/conclusions instead of a diligent, but lengthy description of details.*

**Response:** Many thanks for this very constructive comment. As the unit of our discussion in now shifted from drainage basins to ecozones, key messages as well as the general patterns are more obvious in our revised manuscript. Please note that we have changed several figures related to the analysis of trends in memberships, shift in regime types as well as the attribution of regime shift to the alteration in streamflow characteristics. Accordingly, we have shortened the description of the results and moved the diligent details into the supplement.

11. *I found figures 8,9,10,11 quite hard to assess on first reading. Would it be possible to summarize these results e.g. in sets of 6 maps (one for each cluster).*

**Response:** Many thanks for your comment. To address you concern, we change the Figures 8 and 9 and show the changes in degrees of membership as well as shift in streamflow regimes in a set of 6 maps. Please see the **Figures R4 and R5** below. Regarding Figures 10 and 11, we removed the unnecessary panels, combine the two figures and improve the readability of the figures. Please see **Figure R6** below. We have also considered a couple of new Sankey diagrams, showing how streamflow regimes in our considered RHBN streams transform from one regime to another and how this change in regime types can be attributed to the change in the streamflow characteristics. Please see **Figures R7 and R8** below. For additional changes made in the presentation of our results, we invite you to also review our responses to the comments given by AR1.

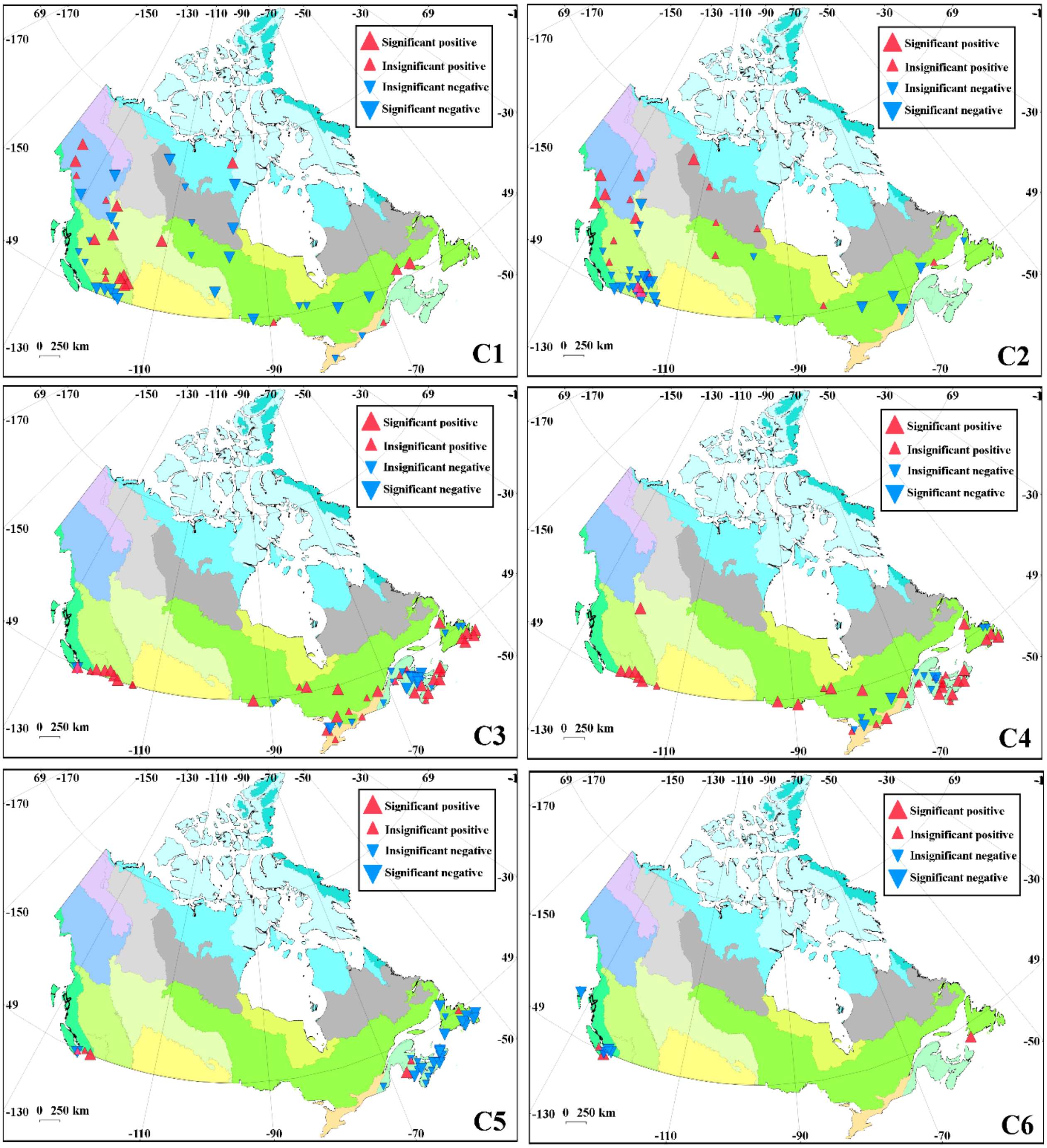


**Figure R4.** The trends in the degrees of membership to each regime type in 106 considered RHBN streams over the period of 1966 to 2010. Colors show the direction of Sen's slopes in the anomalies of memberships. Positive and negative trends are shown with red and blue colors, respectively. Bigger sizes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

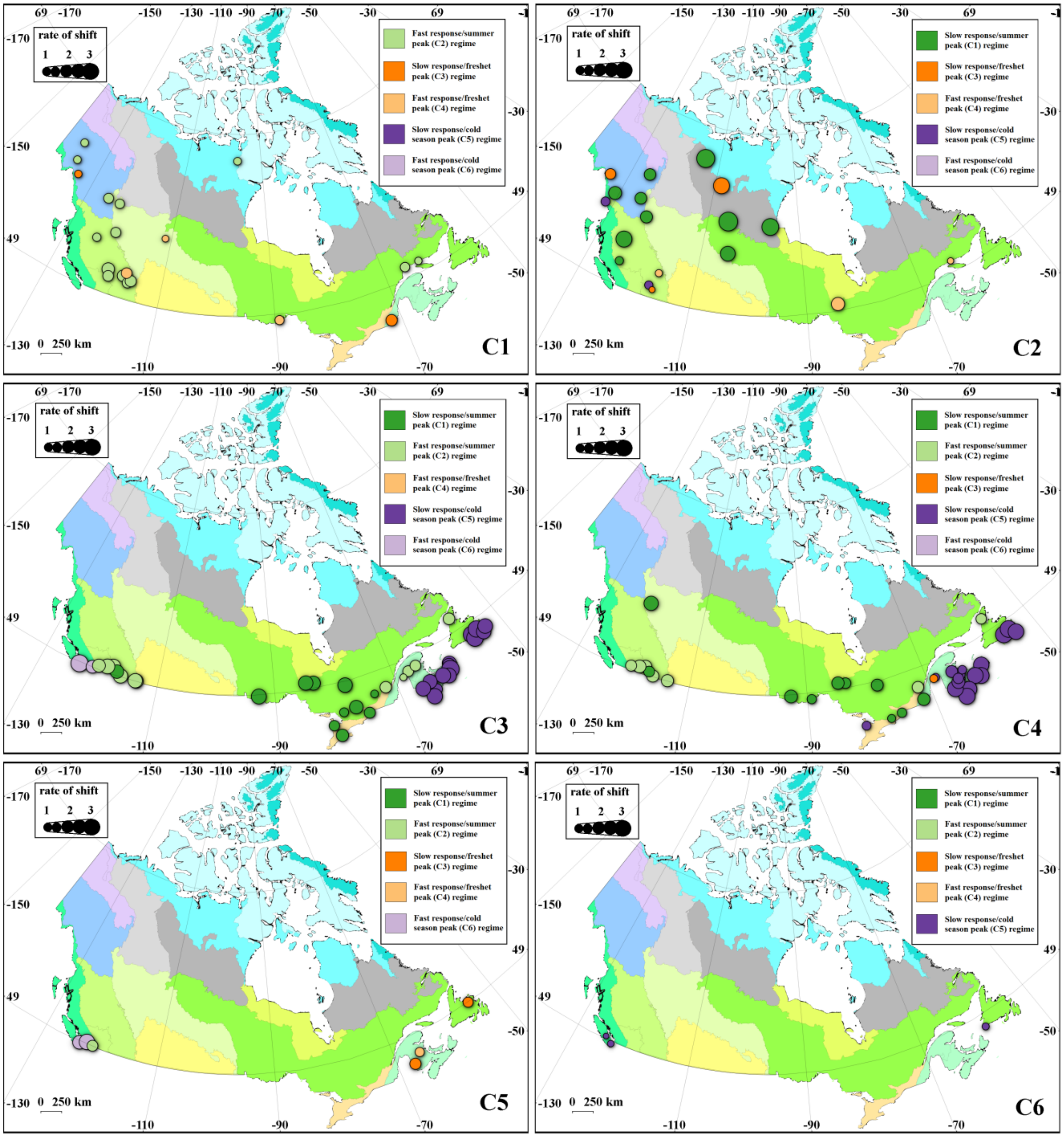

**Figure R5.** Mapping shifts in natural streamflow throughout Canada during 1966 to 2010. In each panel, colors show the clusters from which transition happen to the reference cluster. Rates of shift in regime types for each stream are shown by the size of each circle.

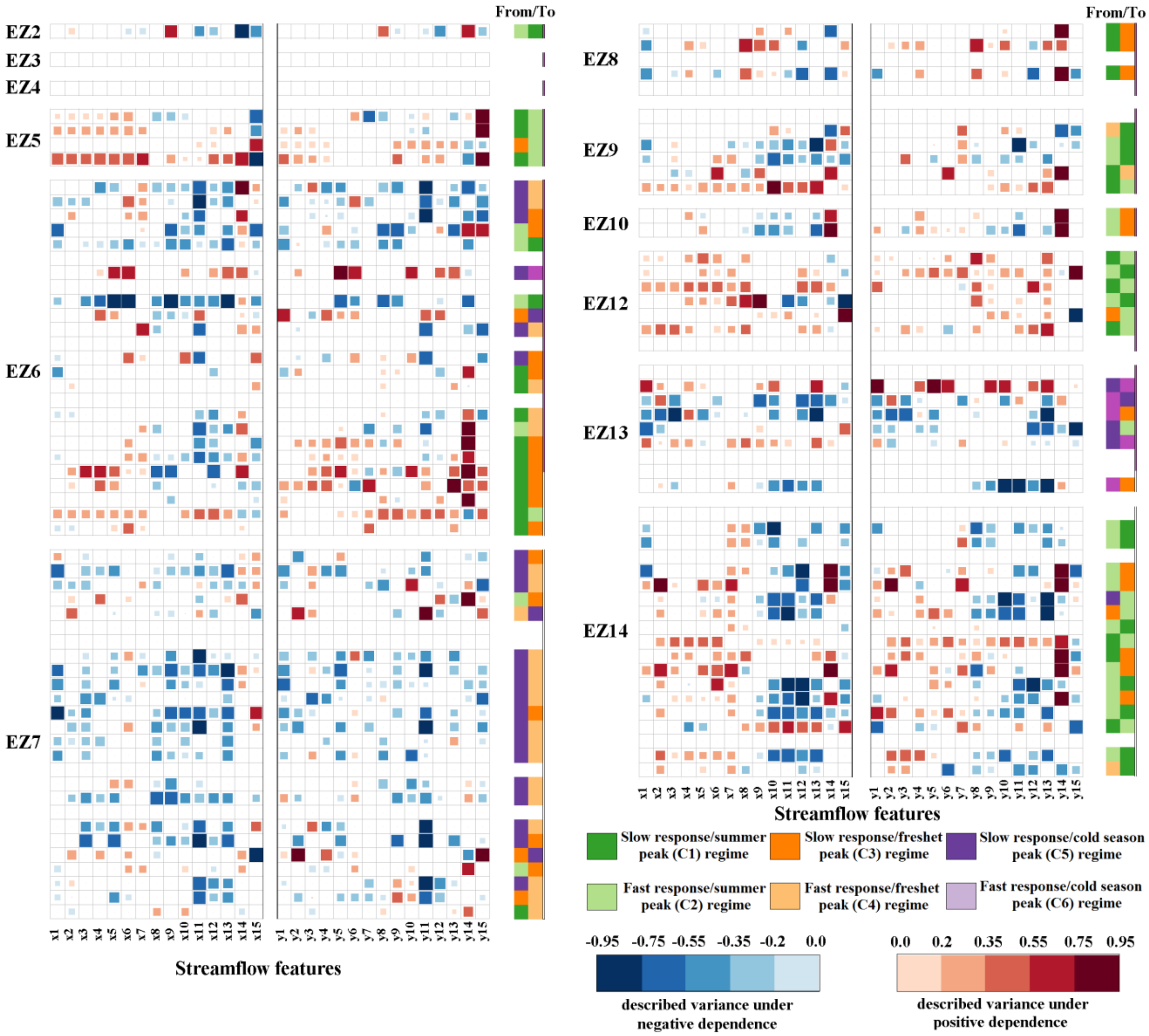


**Figure R6.** The alterations in regime types for 106 RHBN streams attributed to the first and second moments of the 15 IHA considered. Shades of red and blue show the positive and negative dependencies between changes in streamflow features and the degrees of membership, respectively. Blue and red color saturations show coefficients of determination between changes in the streamflow features and the degrees of membership. The dominant regime shift at each stream is identified by the color scheme described in the legend above blue and red color bars. Streams at the left side of each panel are grouped in ecozones and are sorted from the lowest to highest elevations from the top to the bottom in each ecozone.

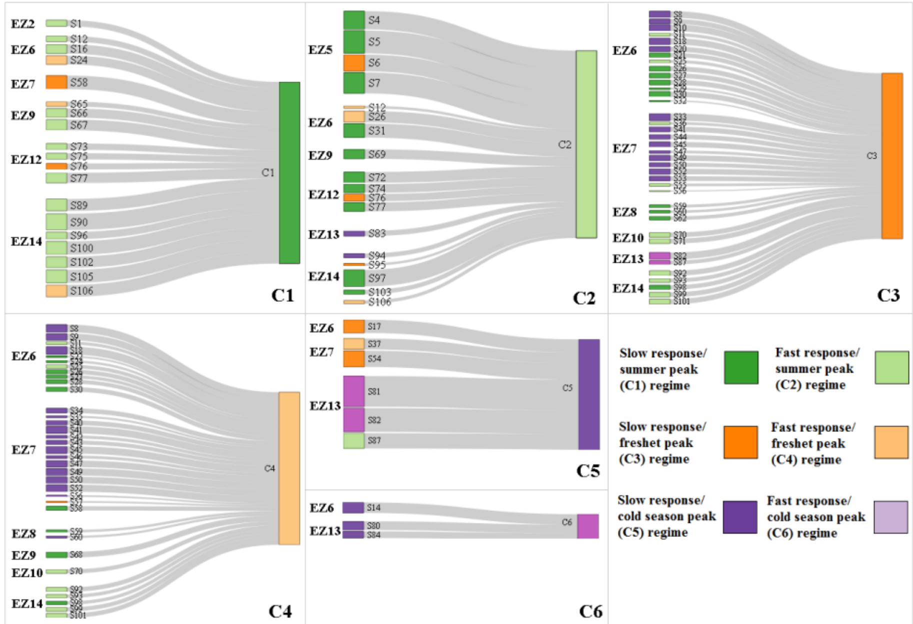


**Figure R7.** Sankey diagrams showing various transitions in streamflow regimes throughout Canadian ecozones during 1966 to 2010. Each panel is related to one target cluster located in the right side of the panel. Streams at the left side of each panel are grouped in ecozones and are sorted in each 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.

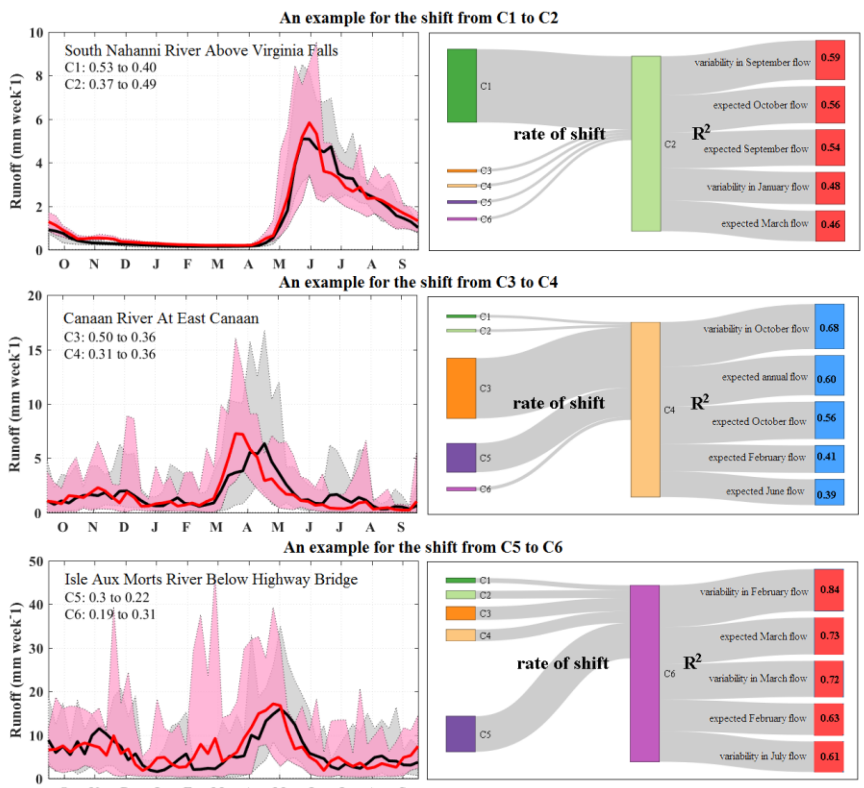


**Figure R8.** Three samples for transitions between regime types along with their corresponding changes in streamflow characteristics. Panels in left show shape of hydrographs in the earliest (1966 to 1975) and the latest (2001 to 2010) decadal episodes using grey and pink envelopes, respectively. Panels 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. Left envelopes show the association of the identified transition to changes in streamflow characteristics using $R^2$.

**Part III. Minor issues**

12. *There is a significant number of grammatical mistakes (i.e. missing articles) in the paper that need to be resolved by the authors.*

**Response:** Many thanks for your comment. We rigorously edited our paper to avoid grammatical mistakes. We will also seek for a professional native-speaking editorial service prior to submission of our revised manuscript to the journal.

13. *Text following equation 2c: V (matrix of centroids) is described twice. This indicates that the methods section might have been written in a sloppy manner, raising the question if everything is correct. I did not have the time to check all the indices etc. in detail.*

**Response:** Thank you very much for your thorough review. We cleared this issue in the text. We also double-checked all formulas and their description in the methodology to avoid such mistakes. All other typos and inconsistency issues in the formulas and notations are also taken care of.

14. *How is the "timing of annual low/high flow" defined? Is it the day of year of the smallest/largest value? If yes: how is the discontinuity between day 365 and day 1 handled?*

**Response:** Many thanks for your comment. The timing is defined on the weekly scale (week 1 to 52). We noted that going finer into daily scale can raise issues, particularly in smaller catchments, in which our results become very sensitive to abrupt weather events such as warm spells or rain-on-snow events (Dery et al., 2009).

Déry, S. J., Stahl, K., Moore, R. D., Whitfield, P. H., Menounos, B., and Burford, J. E.: Detection of runoff timing changes in pluvial, nival, and glacial rivers of western Canada. Water Resources Research, 45(4), https://doi.org/10.1029/2008WR006975, 2009.