Response to Reviewer #1:

We would like to thank the reviewer for their productive comments on our manuscript. Please find our point by point answer as follow:

This manuscript presents an index to describe warm, heavy rain events during winter in southern Ontario. This index is then used to evaluate the occurrence of such events with different large scale circulation patterns and projections for the future. The paper presents a thorough analysis and I appreciate the detailed discussion included in Section 4. Some concerns are listed below.

My main concern is that the authors have not provided sufficient justification for the need to create a new index. The introduction focuses on rain on snow events, but this new index does not consider snow cover. This disconnect needs to be more fully explained. Why is the proposed index better than those used in previous literature?

The main goal of this study was to understand how the frequency of winter weather extreme events (temperature and precipitation), simulated by CRCM5-LE, is modulated by large scale atmospheric circulation. Studying such events are mostly relevant if they have societal implications. A strong shift in high flows occurrence from spring to winter was observed recently in southern Ontario and is expected to continue in the future. Therefore, we decided to define temperature and precipitation thresholds that may explain the generation of high flows in several watersheds in Ontario. Defining an index based also on snow would have been interesting but is not in the scope of this study. A major originality of the study was the calculation of future weather regimes for each member of CanESM2-LE to investigate how the variability of atmospheric circulation will impact the winter weather extremes. The weather regime of a given day impacts directly local temperature and precipitation conditions and investigating also snowmelt adds a level of complexity. Indeed, snowmelt of a given day depends also on the atmospheric conditions occurring weeks before the extreme events (major snowfalls following by cold conditions keeping the snow on the ground). Therefore, weather regimes of these days would also need to be investigated. The need of studying the sequence of weather regimes occurring prior to a high flow event in future studies was discussed at the end of section 4.4.

Moreover, when using snowmelt in the index (With Rain on snow index (ROS) for example) some questions are arising. ROS index does not take into consideration the rain only events while it can have a significant impact on high flows. The occurrence of ROS events is decreasing in the Great Lakes region because of an increase in days without snow on the ground (Jeong and Sushama, 2018), but this doesn't lead to a decrease in high flows. Our index takes into consideration rain events even with the absence of snow on the ground, conditions that are expecting to become more frequent in the future. The proposed index is not meant to be better than ROS but is adapted to the study of weather extreme events simulated by CRCM5-LE and how these events are impacted by large scale atmospheric circulation. As stated in the

discussion, ROS and our index can be studied together to understand the future evolution of different hydrometeorological extreme events (Rain only, rain on snow, snowmelt).

PRMS hydrological model was previously set up in this region (Champagne et al., 2019) which gave us the opportunity to discuss the shortcoming of this index to explain high flows events. We used PRMS to investigate how the future evolution of high flows is corelated to the future evolution of weather extreme events. But the objective was not to create an index using snow data from PRMS output. Nevertheless, to strengthen the discussion around snowmelt, we propose to add a figure showing the evolution of snowmelt between 1961-1990 and 2026-2055 corresponding to each weather pattern. This will feed the discussion on the need to study the impact of the sequences of weather regimes on snowmelt and high flows.

Considering this is a special issue on large ensembles, I think there should be more discussion about how this work is taking advantage of the large ensemble used here.

The main objective of this work was to assess how the internal variability of climate has an impact on the variability of local meteorological extreme events in southern Ontario. To investigate these extreme events in a small region such as southern Ontario, high resolution simulations are required. A regional large ensemble such as CRCM5-LE has the double advantage of simulating the local climate and that each member of CRCM5-LE can be related to large-scale atmospheric circulation from its corresponding CanESM2-LE member. Statistical stochastic methods to downscale GCM data could also represent well the local climate but cannot be related to a corresponding large-scale atmospheric circulation. Few sentences explaining the double advantage of a regional large ensemble will be added in the method section.

Specific comments:

Line 52: Is the "precipitation" here referring to rain or snow or both?

Precipitation is referring to both rain and snow. A mention ''Rain and snow'' will be added to improve the clarity of the sentence.

Line 84: Was only the future period bias corrected?

The bias correction was applied for the entire period (past included). The mention ''Future'' will be removed to avoid confusions.

Line 87: Is precipitation frequency also modeled with a gamma distribution? This makes more sense for intensity.

The method developed by Ines and Hansen (2006) involves two steps in the bias correction of precipitation. For each month, the first step consists on truncating the distribution of the modelled frequency of daily precipitation in order to match the observed distribution of precipitation frequency. The second step used the truncated distribution of precipitation

intensity into a gamma distribution fitted to the observed intensity distribution. For clarity, these explanations on the bias correction method will be added to the manuscript.

Line 138-141: What effect would the trends within these two periods have?

The Z500 anomalies were calculated and normalized for two distinguished period (62 years in the past and 62 years in the future) to avoid the low frequency variability (Here positive trend in Z500) to be disproportionate compare to the high frequency variability. However, this method does not remove all the low frequency Z500 trends. Following this comment, we investigated the change in occurrence of regimes within these two periods and results show a large increase in occurrence of regime HP within the 2026-2055 period. Therefore, the regimes HP are occurring mostly at a period when the conditions are warmer (At the end of the 2026-2055 period), which overestimates HP average temperature and number of warm events. To avoid this artefact, we decided to slightly change our original method and normalize the anomalies in period of 30 years before calculating the regimes. This method minimizes the impact of low frequency on each regime average climate while keeping sufficient periods length (30 years) for the calculations of the anomalies. The number of warm events corresponding to regimes HP in the future period is clearly decreased with this new method. The results and discussion will be modified according to the new figures made with the new method. We will also modify the figure 9 *(DJF change in number of precipitation and warm events between the historical (1961-1990) and the future period (2026-2055) for the 50 members CRCM5-LE average)* **and turned it into change in number of extreme events per occurrence of weather regimes. Indeed we want to remove the impact of the 50-members average change in occurrence of weather regimes on extreme events and study only the variability between members. The goal of this figure is to rather investigate the stability between the past and the future regarding the number of weather extreme events occurring for each weather regime. The impact of change in occurrence of weather regimes on weather extreme events is anyway investigated through the variability between the members of CanESM2-LE/CRCM5-LE (Figure 11 and 12, Table 1 and 2).**

Section2.4/Section 3.1: Were the top regimes determined separately for the model and reanalysis? If so, then I think the good agreement warrants more commentary. If not, can you describe more clearly how they are related?

A principal component analysis (PCA) was first applied to the daily 20thCR Z500 anomalies calculated between 1956 and 2012 for each grid of the domain. The principal components explaining 80% of the spatial variance were identified and their eigenvectors were used to transform the original time series into the principal components time series. This daily principal components time series was then classified using the k-means algorithm. The k-means algorithm identifies iteratively classes centroids that maximize the interclass variance and minimize the intraclass variance. To classify the daily Z500 for each member of CanESM2, the same regimes identified with 20thCR were used. First, the same eigenvectors identified in the PCA using 20thCR Z500 were used to calculate the principal components time series for each

member of CanESM2. Then, the k-means classes centroids identified with 20thCR were used to classify the principal components time series for each member of CanESM2 large ensemble. This more accurate description of the CanESM2 regimes identifications will be added to the manuscript.

Section 3.2: Is the modelled and bias-corrected data being compared against the dataset used for the bias correction target? If so, can you add a brief discussion on the implications?

The modelled and bias corrected data have not been explicitly compared in the study. The extreme events identified from the raw data from CRCM5-LE will be compared to the extremes calculated with the observations. A new figure (Figure R1) will be include in supplementary materials to show the need in applying bias correction. This figure show that the difference between simulations and observations is much higher when using the raw data (Figure R1) compared to the bias corrected data (Figure 4 main manuscript).

Section 3.3: What does the model bias in the South regime mean for these projections?

The model bias between simulated Z500 and observed Z500 in the regime South in the past is likely to remain similar in the future because the same simulated dataset (CanESM2-LE) is used for the future period. The number of events for a regime South calculated with CanESM2 will likely be overestimated compare to a pattern that would look like the 20thCR regime South. However what is relevant in this study is to understand why the atmospheric circulation of the regime South using CanESM2 anomalies produces weather extreme events and how the change in number of regime South occurrence will modulate the number of extreme events. We will add a new column to the figure 3 showing the future evolution of Z500 corresponding to each regime. A discussion on the impact of a change in weather regimes Z500 anomalies will be added to the discussion.

How robust are the streamflow projections when they do not account for changes in snowfall/snowpack or associated feedbacks?

A new figure showing the evolution of snowmelt corresponding to each extreme weather events will be added to the manuscript. This figure will provide more discussion on the difference between weather extreme events and high flows.

Line 229: Do you mean the magnitude of the correlations here?

We meant the magnitude of the correlations. This will be added in the new version of the manuscript.

Line 237: Can you clarify what is meant by the combination of weather regimes? The value of using these combinations instead of just changes in individual regimes is not clear.

A combination of two weather regimes has been done by summing the seasonal occurrence of these two weather regimes. If for a given winter the regime HP occurs 20 times and the regime South 15 times, there combination will be 35. The goal of using these combinations is to

identify the impact of a combination of weather patterns occurring the same winter on weather extremes and high-flows. The weather patterns are linked to each other because are a discretization of a continuous process (Atmospheric circulation). As stated in the discussion, a given pattern recurrently succeed to the same patterns because the systems (cyclones and anticyclones) are following a general direction (West to east). Therefore, the weather regimes are not independent from each other. A combination of weather regimes occurrence shows the impact of a simulteanously large occurrence of two weather regimes on hydrometeorological extremes. It is particularly valuable to understand their impact on high-flows because atmospheric conditions days before the events are also relevant to explain the generation of high flows events (i.e. Formation of Snowpack). A clearer explanation on what is a combination of weather regimes and its purpose will be added to the new version of the manuscript.

Figure 1: I found it very difficult to orient myself in this figure. It is hard to discern the coastlines of the Great Lakes from the contours in the large scale figure and the legend covers the land separating two lakes. It would be very helpful to label the lakes and key places here, especially since so much of the later discussion is very specific about local geography.

This figure will be replaced by a new figure with a modified large-scale figure, a clearer separation between Lake Erie and Ontario and the name of the Lakes (Figure R2).

Figure 2: Why is the scale of the horizontal axis nonlinear? With the current spacing and connecting line, this time series is misleading.

The scale of horizontal axis is not linear because represent different high flow events. The connecting lines representing the average will be removed in the new figure as they are confusing and do not give any valuable information (Figure R3).

Figures 4-7: The figure label ("simulations minus observations") does not agree with the caption ("CanGRD minus CRCM5-LE")

The words simulations and observations will be added to the legend.

Figure 8: Are the ensemble values pooled or averaged?

The ensemble values are pooled. This information will be added to the manuscript.

Line 462: This paper should be cited as Jeong and Sushama

This modification will be done

References:

Champagne, O., Arain, M. A. and Coulibaly, P.: Atmospheric circulation amplifies shift of winter streamflow in Southern Ontario, Journal of Hydrology, 124051, doi:10.1016/j.jhydrol.2019.124051, 2019.

Ines, A. V. M. and Hansen, J. W.: Bias correction of daily GCM rainfall for crop simulation studies, Agricultural and Forest Meteorology, 138(1–4), 44–53, doi:10.1016/j.agrformet.2006.03.009, 2006.

Jeong, D. and Sushama, L.: Rain-on-snow events over North America based on two Canadian regional climate models, Climate Dynamics, 50(1–2), 303–316, doi:10.1007/s00382-017-3609-x, 2018.