Response to Reviewer 2:

Comments/Text of reviewer posted in **black;** our answers are posted in blue.

RC2:
This manuscript examines long-term trends and variability in the seasonal and annual temperature, precipitation and runoff in the Svartberget catchment, within the Krycklan catchment, located in the northern Sweden's boreal zone. This manuscript compliments many of the past studies done within this catchment, adding a much-needed analysis on the long-term effect of changes in temperature and precipitation on the runoff.

Overall, the manuscript is well written and well structured. I believe that the current manuscript requires additional analyses and information to be publishable as a journal article in Hydrology and Earth System Sciences. Provided below are a list of questions, comments, and suggestions towards an improved version of the manuscript.

Response: We thank the reviewer for the thorough review and positive feedback on this research. We appreciate the detailed comments and suggestions for improvement of the message and analysis. We will address each point as detailed below.

General Comments:

1a) Line 85: With the hypothesis that warmer winters will result in higher runoff during the winter, exhausting summer baseflow, how much more runoff is expected to occur during your defined winter period from snowmelt? If temperatures remain below freezing during the winter, then I suspect an increase in winter baseflow would be from the previous autumn season.

Response: In response to "how much more runoff is expected to occur during your defined winter period from snowmelt", we can see that using the accumulated freeze degree days (number of days less than 0; AFDD) between the period 1992-2022 on average was -1579^{o}C with an average winter runoff of 0.13 mm day $^{-1}$. In the colder years (such as 1994 and 2003) with an AFDD of below -2000 ^{o}C we observe the lowest Q (below 0.1 mm day $^{-1}$). Contrary during warmer years with AFDD above -1250 ^{o}C (such as 2015) we observed significantly higher runoff (commonly above 0.2 mm day $^{-1}$) where significantly higher runoff was observed over time (r^2= 0.25, p<0.01 Fig 3B). We will add this to the results to improve the novelty of this research.

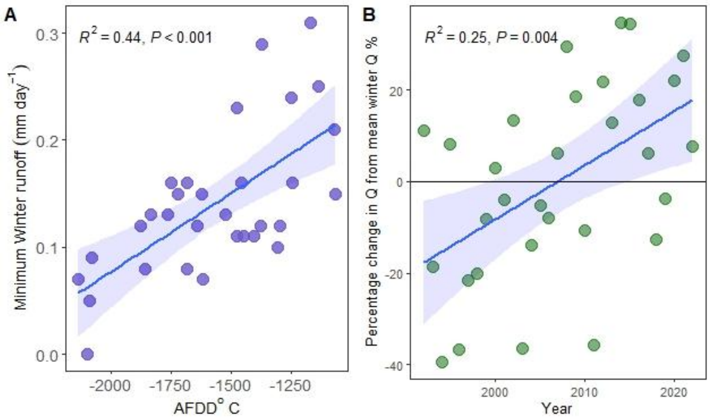


Fig 1 Change in winter runoff in relation to Accumulated freeze degree days over time

1b) Would the changes in the timing of snowmelt most likely affect the spring season thus creating more low flow events in the summer?

Response: Analysis of the initiation day of melt during the spring (day that runoff increased above baseflow) showed that snowmelt occurs earlier over time (as shown using the Mann Kendal trend test (tau =-0.29 p=0.02 Fig 4A)). Using regression analysis to detect the relation between winter extreme indices and spring initiation day of melt, we note that winters with a higher number of frost days correlated with earlier spring melt, which supports that less water is available in the catchment for summer low flow (Fig 4B). However, since we focused only on minimum runoff indices and extreme climate indices, including other runoff indices is beyond the scope of the manuscript.

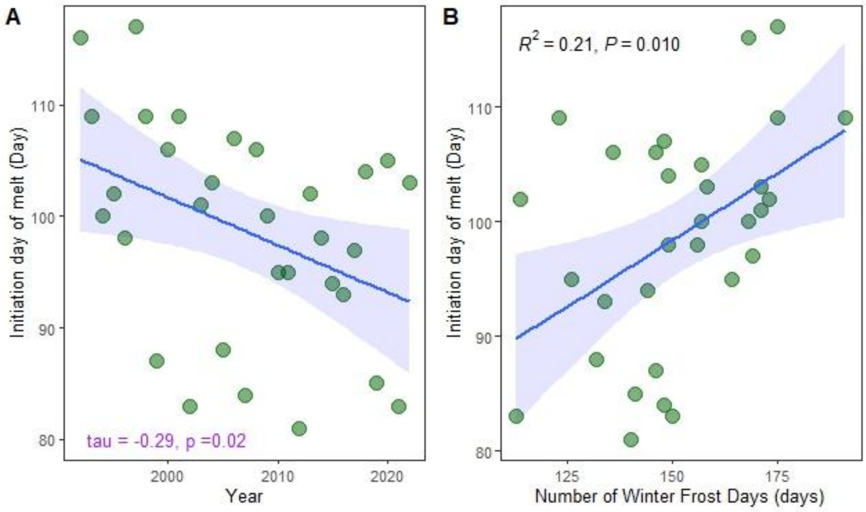


Fig 2 Analysis of changes in the timing of spring melt (A) and its relation to winter climate (B)

2) Line 126: Why was the autumn season not used in this study? Late autumn discharge, enhanced by autumn rainfall and early snowmelt sessions, will influence the total winter flow. An analysis of the autumn season should not be omitted in this study.

Response: Following the suggestion of reviewer 1 (and this comment), we have tested the effect of autumn precipitation and runoff conditions on winter runoff (see reviewer 1 comment 2) using three techniques. The results did not show any significant relation between Autumn variables and winter runoff. However, we do appreciate this question in testing the validity of the winter runoff driven by winter snowmelt rather than stored water from previous autumn events. This information will be added to the manuscript.

3) Are there any strong connections of climate oscillations (e.g. AO or NAO) to the climate variables or climate indices that could then influence runoff in this catchment?

Response: While previous studies have shown the connection between North Atlantic Oscillation (NOA; Ulen et al 2019) and AMOC (Schenk et al 2018) to winter climate (winter temperature, precipitation, snow accumulation, frozen soil) in South-western Sweden and along the coast of Norway, no such effects have to the best of our knowledge been observed related to the climate in Northern Sweden. However, using the NAO index (https://www.ncei.noaa.gov/access/monitoring/nao/) to test the correlation with seasonal hydro-climatic variables from the Krycklan catchment, we found significant correlations with four winter and four summer temperature extreme indices. Although we did not find any direct connection to runoff or precipitation variables during the seasons, these new results do suggest that NAO can have an important role in regulating local temperatures in the Krycklan catchment, which in turn can drive processes regulating winter runoff. These new results will be incorporated into the discussion section to support the effect of climate change on NAO and local climate in Northern Sweden and the figure will be added to the supplementary figures.

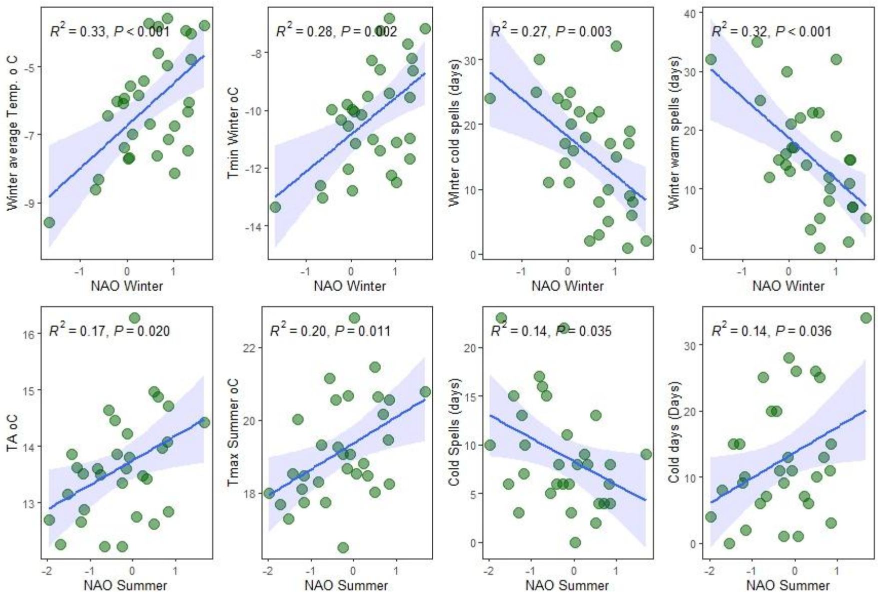


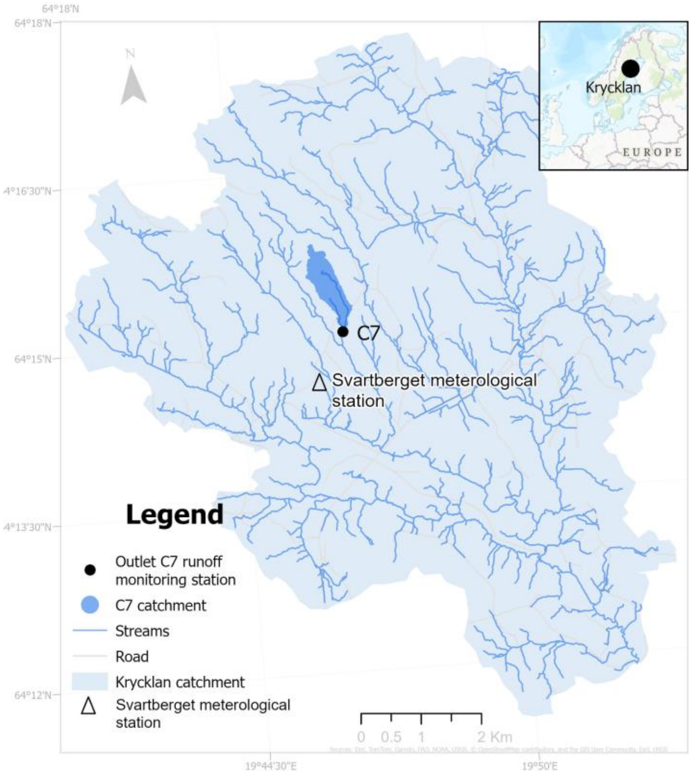
Fig 3 Regression of winter and summer variables with the North Atlantic Oscillation Index (NAO) during the winter and summer

4) In order to detect trends and significant changes, I am wondering why a more suitable Mann-Kendall test and Sen's slope was not employed on the time series?

Response: Indeed the non-parametric Mann-Kendall test was first performed for detecting trends in the dataset during the initial analysis (see also comment by Reviewer 1). However, we did not include the results in the submitted version of the manuscript but have done so now following this comment. It should be noted that the results did not change by including the Mann Kendal test other than by improving the prediction of trends in some cases. All figures will be updated with the tau values and a table has been added to supplementary information containing all trend values.

5) Any map that you could provide for the readers of the Krycklan and Svartberget catchment, including location of sampling sites, met tower(s), and hydrometric stations?

Response: The map will be added following this and the other reviewer's suggestion



R Fig 4 Map showing the location of the study site in the Krycklan catchment in relation to Northern Countries

6) Have you considered using a change-point detection on the time series to detect any regime shifts?

Response: Following this comment, we checked for change points using r statistical software change point package for the Runoff, Average Winter Temperature, Precipitation and Evapotranspiration using the local climate data (1982-2022). No change point was detected in the average temperatures, however, we did identify change points in Q, P and ET. These change points did not coincide, and instead occurred at different times for each variable (Q in 1998, P in 1994 and ET in 2005) and therefore could not provide any additional information that could help in understanding the connection between winter climate and seasonal runoff.

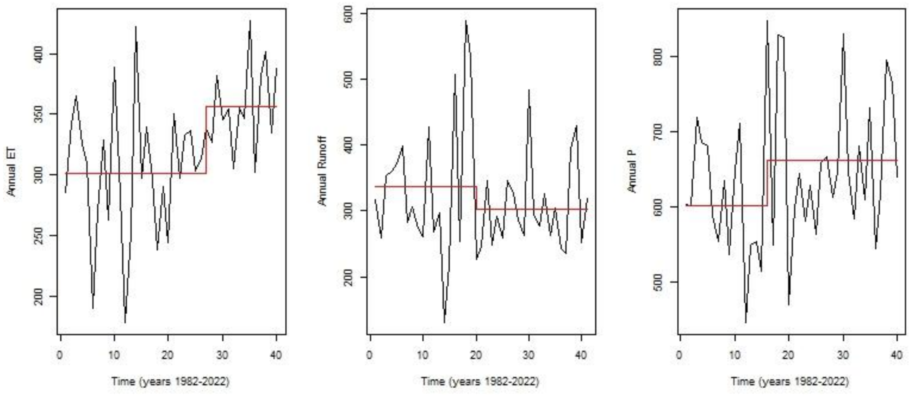


Fig 5 Change point detection analysis using R studio strucchange package for the Svartberget annual data on Evapotranspiration (ET), Annual Runoff and Annual precipitation

Specific Comments:

1) Page 1, Title: In the manuscript, you focus on the other seasons, not just winter. I suggest changing the title to reflect your analysis.

Response: We appreciate the comment but would like to emphasize that we are interested in showing the effects of winter conditions on successive seasons and therefore would keep our original title.

2) Line 105: What is the difference between the two different data sets? Was the shorter period data set only missing daily maximum and minimum temperatures? Why not use the 1982-2022 data set to detect long-term trends and "extreme climate change"?

Response: Indeed the difference in the dataset was the maximum and minimum temperatures. Since the extreme indices used Max and Min to define each index, we were not able to calculate indices using the longer dataset. We will make this clearer in the next version of the manuscript in the method description section.

3) Line 110: What is the difference between the two stations? Was any analysis conducted to see the similarities between the two locations that are 150 km apart?

Response: The data from the Stensele and Svartberget synchronise with an $r^2=0.92$, root mean square error (RMSE 0.05) for the overlapping period 1982-2004 using the daily average time series. The systematic bias in the dataset lies in the extremes where maximums are higher in the Svartberget than the Stensele dataset while minimums are lower in the Stensele than the Svartberget dataset. However, these differences are small and do not affect the general trends in the dataset used in this study. Based on these results, we expect that the trends observed in the long-term Stensele dataset can be used to reflect similar trends in Svartberget. Thank you for the comment, we have included this in the supplementary information section to support the data analysis.

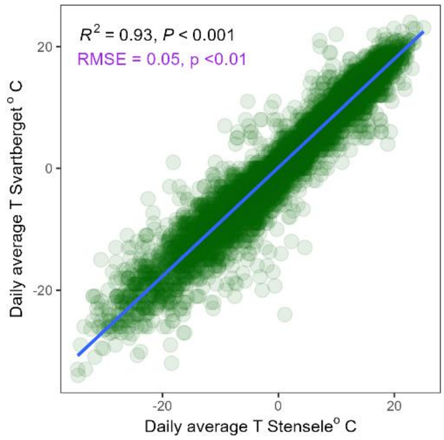


Fig 6 The correlation between daily mean temperature from the Stensele meteorological station and the Svartberget meteorological station between 1982-2004 showing the r^2 and RMSE (root mean square error) between the two datasets

4) Line 116: What percentage of the data set was gap-filled and the quality of the data? The reference, Karimi et al. (2022), does not mention this information for the 1982-2022 time period.

Response: Gap filling was done for 4% of the dataset as described in Karlsen et al 2016. Thank you for the comment; we will add this to the method description to improve the clarity of the dataset limitations.

5) Line 144: Do you avoid possible inhomogeneity as described in Zhang et al. (2005)? Zhang, X., Hegerl, G., Zwiers, F.W. and Kenyon, J. (2005) Avoiding Inhomogeneity in Percentile-Based Indices of Temperature Extremes. Journal of Climate, v18, 1641-1651. DOI:10.1175/JCLI3366.1.

Response: To address inhomogeneity in the dataset, we used the bootstrap technique to test if the trends in climate indices would vary depending on the window use ie 3 days, 5 days, 7 days) based on the description of the climate index (https://etccdi.pacificclimate.org/list_27_indices.shtml). Since these data were aggregated by seasons and not by weeks or months, the windows did not cause the results to vary so we used the definition as published to be consistent.

6) Line 144: For Table S1, could you provide more details on how these are calculated?

Response: We will add a link to the 27 extreme climate indices that describe in detail how each index was determined and will add a short description to Table 1 in the manuscript to improve the clarity of how the extremes were determined.

7) Line 146: For the maximum 1-day precipitation total, what if the precipitation event starts at night and ends in the morning? How do you account for this?

Response: We used a 24-hour technique where the day starts at midnight 00:00 and ends at 23:59. Within this period, all recorded precipitation was summed to reflect the daily total. If a precipitation event started before 00:00 in the previous evening, it was recorded in the previous day's daily sum. We assume that even if we missed an extreme in precipitation because of how a day was defined in maximum 1-day precipitation, we expect that this would be reflected in other indices such as maximum 5-day precipitation, simple precipitation index etc.

8) Line 175: Define SOI as the Seasonal Origin Index here.

Response: We will add the definition of SOI to this section based on this suggestion

9) Line 191: The change in average daily temperatures from two different years is misleading and is best to just stick to the slope value of 2.2. Similar comment for the select years on Page 8, unless you are mentioning the upper and lower boundary of the variability.

Response: We agree and will use the long-term trend values instead of the changes in the years

We now suggest formulating this as " In the last four decades, the increase in temperature has accelerated by 2.1 °C where long-term average daily temperature trend changed from 1°C in the 1980 to 3.1 °C in 2022".

Page 8

We now suggest formulating this such as "A closer look at the last 30 years of variability in temperatures in the Krycklan catchment (1992-2022) showed annual average temperatures ranging from 0.5 to 5 °C across the years increasing from 1.5 °C in 1992 to 3.1 °C by 2022 (Fig 1B) when looking at the long-term trend line."

10) Figure 1B: What does the grey shaded area represent on these scatter plots?

Response: The grey shaded area represents the standard error in the dataset. This info will be added to the caption.

11) Figure 1C: What does the blue shaded areas represent? Are the numbers 1-12 suppose to represent each month? The caption for Figure 1 in general is very confusing and jumps between Figure A, B and C.

Response: We will restructure the caption to improve clarity as follows

We now suggest formulating this as "Figure 1 Trend in long-term temperature in the Krycklan catchment showing the relation to the much longer-term time series from the SMHI-200 Stensele data (1891-2004) (A), trends in average daily air temperature (Avg), maximum (Max) daily air temperature and minimum (Min) daily temperatures, and variability in daily temperatures in the months across the year (C), across 30 years from 1992-2022. In panel A, the Red symbols indicate the average within a 30-year period from 2022 backwards while purple and orange error bars represent the standard deviation in the Stensele and Svartberget datasets, respectively. The grey shaded area in panel B represents the standard error in the datasets and the blue jitter dots in panel C represent daily average temperatures in Svartberget."

12) Line 207: What is the difference between these minimum temperatures and those in the previous sentence?

Response: In line 206, we show the annual average temperatures, which ranged between 0.5-5 º C and in line 207, we show the annual minimum temperature, which ranged -4.1 to -0.4 º C. We will omit the term "average" in line 207 to improve the clarity of this sentence.

13) Line 208: Was this "coldest month" always the same month?

Response: Looking back at the dataset, we see that across the years, the coldest months were Jan and Feb. We will change this sentence to reflect this.

We now suggest formulating this such as: "The annual average minimum temperature ranged from -4.1 to -0.4 ºC with the coldest average monthly temperatures occurring in either January or February, ranging from -20.6 to -7.6 ºC (Fig. 1C)."

14) Figure 2: This should be placed after the following paragraph. What are the legend labels in Figure 2A? This is the first time evapotranspiration has been mentioned. I would include a quick comment in the Methods section on how it was calculated. How much confidence is there in these values and the increasing general trend?

Response: Good suggestion, we will add a short description of ET in the methods and have reworded the cation for improved clarity. Based on the Mann Kendal analysis which showed a positive tau and p-value <0.02, we can show high confidence in the ET trends.

We now suggest formulating this as: "Figure 2 Long-term trends in annual Precipitation (P) and runoff (Q) from the Krycklan catchment across 40 years showing the differences in the percentage of precipitation (P) from runoff (Q) that represents evapotranspiration (ET) (panel A). The variability as measured by the standard deviation in runoff (Q) and precipitation (P) during the time period is shown in panel B. The shaded areas represent the standard error in the datasets. "

15) Line 228: What about the variability in precipitation and its increasing trend?

Response: We will add a sentence to explain the variability in P.

We now suggest formulating this as: "It should be noted that the variability in runoff has decreased by 0.5 mm day^{-1} from 1.6 mm day^{-1} to 1.1 mm day^{-1} while precipitation increased from 3.5 to 4.2 mm day^{-1} during the period 1982-2022 (Fig. 2B, Table S2, Fig. S2)."

16) Line 233: Is the same method as described in the Methods section? I would delete this sentence or at least make it more clear for Figure S3.

Response: This sentence will be omitted to improve clarity based on the reviewer's suggestion

17) Line 235: What is the difference between "daily averages" and "average winter temperatures" here?

Response: This sentence will be omitted to improve clarity

18) Figure 3: What is the "blue box" described in the caption? What do the shaded parts of the trend lines represent? In panels C and D, it looks like both the trend and variability, with the variability lines hardly noticeable.

Response: The caption will be reworded and the error bars omitted to improve the consistency of the figure.

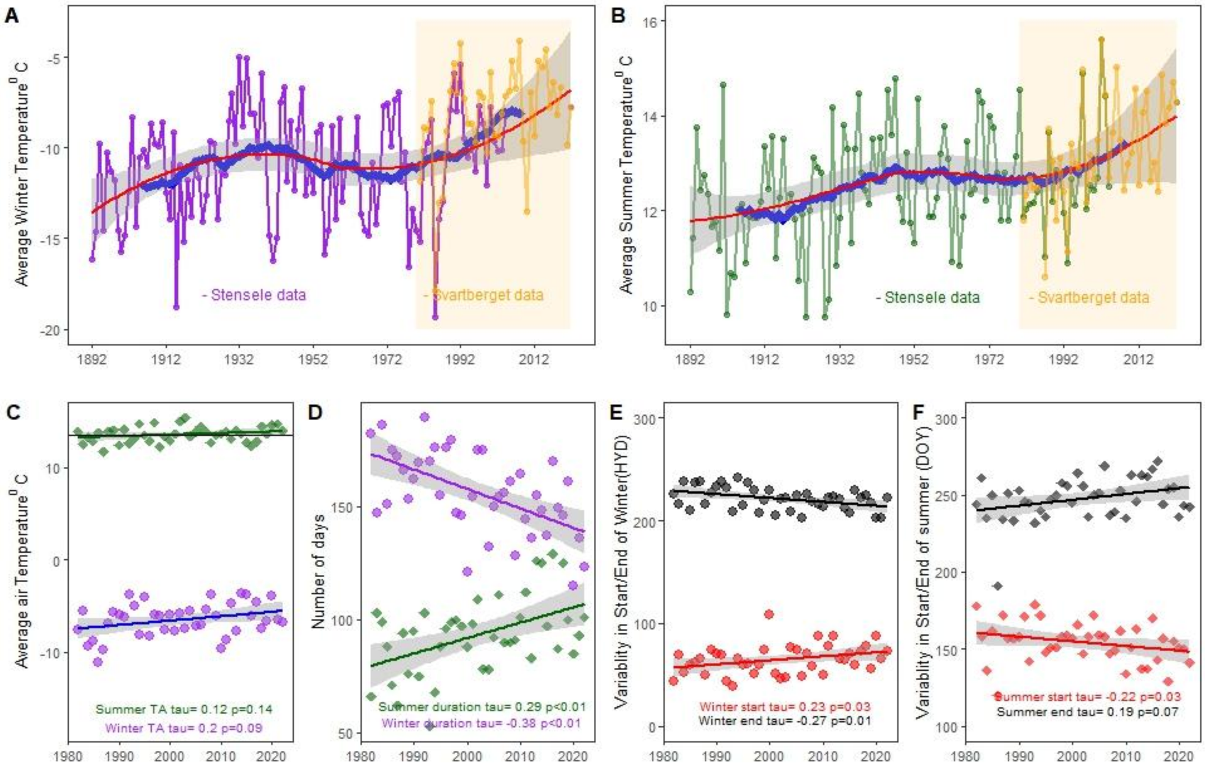


Fig 7

We now suggest formulating this as: "Figure 3 Variability in seasonal temperature across 40 years in the Krycklan catchment (orange box) showing changes in temperature during the winter (A) and Summer (B) in relation to the SMHI Stensele longer-term dataset (1891-2004). The variability in annual air temperature (C) and duration (D) is represented by purple symbols for the winter period and green symbols for the summer period. Changes in the start (red) and end (black) of the winter (circle) and summer (diamonds) are depicted in panels E and F, respectively. Standard errors in the datasets (A, B, D, E, F) are shown as grey shading."

19) Line 254: "MinTmin" was shown not to be significant in Table S1.

Response: Thank you for the observation, we have omitted the error.

20) Line 255: Based on Table S1, you have some indices with significant changes, but no significant trends in any of the spring indices or seasonal runoff variables? Is this correct? The same question for Figure 4.

Response: Using the Mann Kendal trend test, we could not detect any significant trends or changes in the spring indices. We have changed the analysis to the Mann-Kendel trend test to make this clearer. In Figure 4 (manuscript), we observe increases in Q during the winter as shown by the positive trend test (tau=0.3 $p<0.01$). During the summer although we see a decrease across the year, these changes were not significant at $p<0.05$.

21) Line 262: Are these trends described here different than those listed in Table S1? Seems like you are using two different data sets to examine similar trends. Table S1 is 1992-2022 and this analysis uses 1982-2022.

Response: While we do look at the long-term trends to gain a more robust long-term perspective on how runoff is changing across time as shown in manuscript fig 4, we have been limited to truncating the dataset from 1992-2022 for the modelling analysis based on the availability of extreme climate indices data. Based on this comment, we have added both results to Table S1 and added a sentence to reflect this difference.

We now suggest formulating this as: "Looking at the time series between 1992-2022, we observe similar increases in winter trends tau= 0.3 $p<0.01$). Winter runoff varied between 0.001-0.05 mm with total season runoff during the winter varying between 18 mm and 144 mm. On average runoff ranged between 0.1 mm day^{-1} and 1.1 mm day^{-1} with the lowest runoff recorded in 1996 and the highest runoff occurring in 2007 (Fig. 4). During the summer, the minimum trends in runoff were also decreasing (r^2=-0.05, p=0.14) with average daily runoff varying from 0.13 mm to 2.45 mm between 1982 and 2022. Trends between 1992- 2022 were also consistent with this (tau= -0.08 p=0.14). Daily minimums varied from 0.01 mm day^{-1} to 0.42 mm day^{-1} while maximum runoff during the summers varied from 0.8 mm day^{-1} to 21.9 mm day^{-1}. The driest years were recorded in 2006 (0.03 mm day^{-1}) and the wettest years were recorded in 1993 (0.52 mm day^{-1}) (Fig. 4)."

22) Line 278: Is "MaxTmax" a winter variable or a summer variable?

Response: Thank you for the comment, we will change all indices to reflect S for summer and W for winter to avoid further misinterpretation

23) Figure 6: Any discussion on Figure 6A? Are the numbers on the x-axis representing months?

Response: We will add a citation to where Fig 6A will be explained and the months' label will be added to the x-axis

24) Line 320: The SOI values show an increased contribution from winter precipitation to streamflow. Any reason as to why this does not show up in your analysis of the data sets?

Response: In our analysis, the contribution from winter precipitation to stream flow is reflected in increasing negative values across time. This is indeed shown in Figure 6B where we see SOI shifting from 0.2 to -0.6 during the winter period.

25) Line 321: Were there more mid-winter melting events found in the data sets? Would the shoulder seasons (spring and autumn) be the most affected by the warming rather than winter?

Response: In this analysis, we focused mainly on the lowest flow across the years as an indication of mid-winter melt. Another indicator that can be used is snow depth or soil frost, which also indicates warming. In terms of the effect on the shoulder seasons, we have tested the effects of autumn runoff on mid-winter melt (reviewer 1 comment 2) which did not show any significant correlation with the runoff during the winter. Additionally, for spring, we have found earlier initiation days of melt (not included in this study) in relation to winter frost days (see reviewer 1 comment 1b) indicating that indeed warmer winters are also affecting spring flood dynamics.

26) Line 355: As mentioned previously in the discussion, there was an increased contribution from winter precipitation with the SOI analysis. Figure S4 showed a decrease in total winter precipitation and an increase in total summer precipitation. Any comment on the different results? How would this influence the probability of droughts in this boreal catchment?

Response: Increased contribution of precipitation to stream runoff during the winter is an indication of increased loss to the catchment through runoff rather than stored in snowpack. This, together with reduced winter precipitation can further induce low groundwater recharge for aquifers (similar to other studies by Nygren et al, 2020, Jasechko et al., 2017, Klove et al 2017). Even though precipitation is increasing during the summer (although not significantly), this may not be enough to recharge the storage or the dried-out landscapes which are water deficit. This is reflected in the low correlation between minimum summer runoff and total precipitation where years with the highest precipitation are not the year with the highest baseflow.

27) Line 372: "Shrinking snowpacks" As mentioned in previous comment, one analysis shows a greater contribution and the other, a non-significant decrease. Did you find that the snowpacks were shrinking?

Response: In this analysis, we have focused on the effects of extreme winter indices, which infers impacts on snowpack indirectly. With the SOI analysis, we showed that the isotopic signature in the stream water during the winter reflects more of the winter precipitation signal across time. With the extreme indices analysis, we have not found any significant increase in contribution from winter precipitation or in the number of precipitation days in the winter. However, the extreme event analysis showed warmer winter temperatures across time, alluding to an increased winter melt of the snowpack. This coincides with the isotopic analysis results.

1) Table S1: What is the difference between Tmin, Tmax and MaxTmax and MinTmin?

Response: Temperature is recorded on a 10-minute interval in the Krycklan catchment where hourly maximum, minimum and averages are calculated. To identify the coldest recorded temperature of the day we used the MinTmin, which selects the lowest hourly recording of each day while Tmin calculates the average minimum temperature of that day. Similarly, MaxTmax identifies the highest hourly temperature recorded in a 24hr period while Tmax reflects the average maximum temperature in a day. We will make this clearer in the description of the methods to improve clarity.

2) Table S1: Which 27 climate change indices were developed by the World Meteorological Organization out of the 33 variables listed?

Response: Great comment, we will separate the 27 indices to improve clarity

3) Table S1: Should the growing season length be under the summer category instead of winter?

Response: We will move the growing season length to summer. Thank you for the suggestion

4) Figure S1: Are the dotted lines the average daily temperature in blue? It is difficult to distinguish between the three data sets in this figure.

Response: We will remake this figure and hope that it is now clearer

We now suggest the following figure:

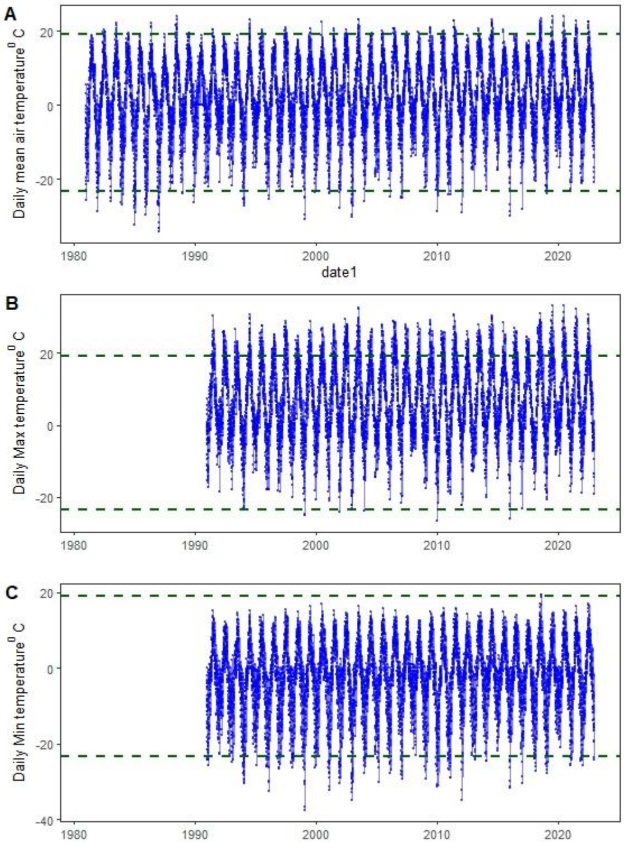


Fig 8 Daily air temperature showing (A) Hourly averages, (B) Hourly Maximum temperature and (C) Hourly Minimum temperatures

5) Figure S2: For panel C and D, what does the pink shade represent?

Response: We will improve the description of this caption.

We now suggest formulating this such as: "Variability in daily precipitation (A) and runoff (B) and the variability in annual values for precipitation (C) and runoff (D) in the Svartberget catchment. The pink shade represents the standard error in the dataset"

6) Figure S3: If this is average daily temperature for each year, where is the "isolation of the winter period" in each figure?

Response: The winter period would be the area where the average temperature (blue line) falls below the zero line (black). We will add this to the caption and hope that it improves the clarity of this figure.

7) Figure S5: What does the blue shade in the background of each figure represent?

Response: We will add a description to this caption

We now suggest formulating this such as: "The standard error in each dataset is represented by the blue shades."