Response To Reviewer #1

The paper by Croghan and co-authors entitled "Seasonal and interannual Dissolved Organic Carbon transport process dynamics in a subarctic headwater catchment revealed by high-resolution measurements" presents a comprehensive data set of high-frequency flow and DOC data (and associated hydroclimatic variables) over four years to assess the DOC dynamics in an arctic catchment (without permafrost). The authors utilize this data to explore high-frequency metrics and other indices. They use random forest models to assess the drivers of these metrics, and find some interesting and some confounding patterns.

This paper is well structured and clear - with a thorough analysis that doesn't stretch too far beyond the data in hand and seeks to advance our understanding of the system. The data set is valuable and unique, and the process insights strong. There are a number points I would ask the authors to consider, and also explore a newly published paper that has a similar analysis in a subarctic environment with at times similar and at times different results that likely was not spotted before submission. Some of the authors interpretations may (or may not) be informed by this paper:

Shatilla, N. J., Tang, W., & Carey, S. K. (2023). Multi-year high-frequency sampling provides new runoff and biogeochemical insights in a discontinuous permafrost watershed. Hydrological Processes, 37(5), e14898.

>> We thank you for your positive evaluation of our paper and constructive feedback. The suggested paper was indeed not spotted before submission but has been very insightful. Specifically, we appreciated your supporting observations that CDOM-Q events usually had negative hysteresis index values, outside of Spring and Fall, which was very much in line with what we found.

Further, we also found the system differences in seasonal DOC patterns interesting. In this suggested paper DOC/CDOM peaks in freshet and rapidly declines, while for our site, DOC remains elevated after freshet, though naturally some differences are expected due to a lack of permafrost in our study catchment.

We have now used this paper as a supporting reference throughout this manuscript, thank you.

Other comments:

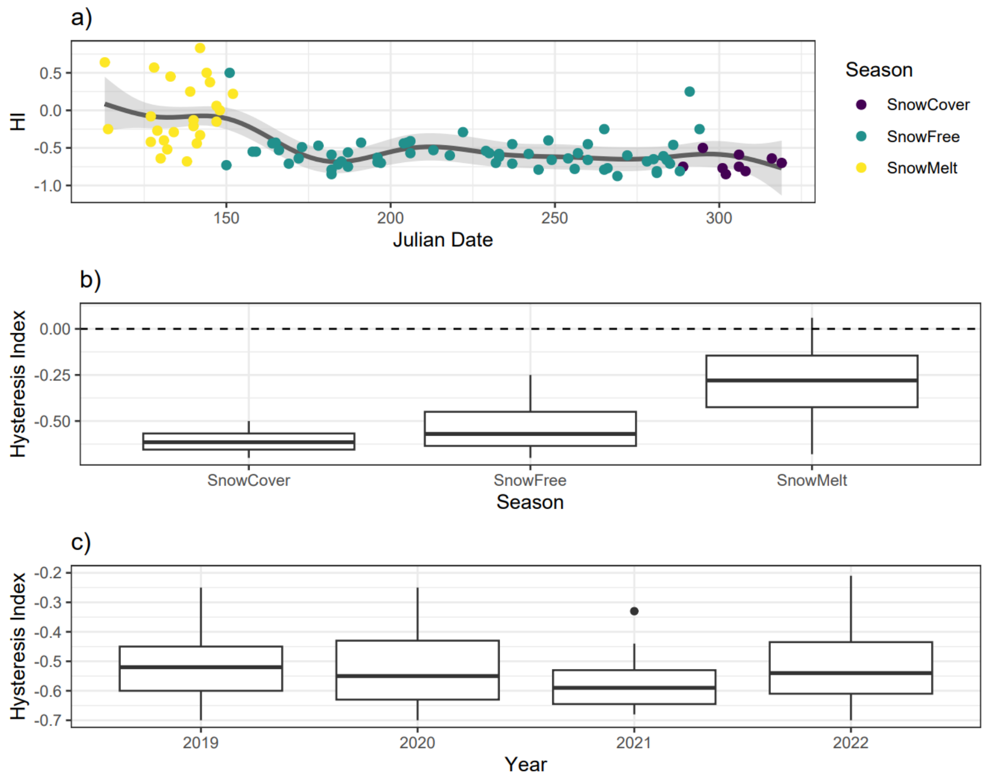
~Equation 1. Is there a reference for this equation?

>> The equation was derived through internal laboratory calibrations of the instrument thus there is no reference. We have made this clear in the revised manuscript.

L205: “FDOM required corrections for turbidity effects (Downing et al., 2012). Corrections for turbidity were undertaken using the following equation which was derived using internal lab calibration of the instrument”.

~Line 237. Only single peaks were used for analysis. How many events were single vs multiple peaks in the entire data set? Did the authors examine multiple peaks or was this simply 'too messy'?

>> The motivation was to compare between but certainly there was merit in examining multiple peaks. We've subsequently included an additional 37 peaks (from multi peak events) to see how multiple peaks would change the dataset, but there is no overall difference in results. The snow melt season still has the highest average HI values and in snowmelt and snow cover the values are still consistently very negative in comparison. Thus, the single peaks vs multiple peaks did not make a difference and there will be no need to change any of the text (outside of edits to the statistical test values), though we have included the multiple peaks data (shown in the figure below) in the revised manuscript.



~I would like the authors to consider including summary hydroclimatic information that is easier to 'digest' than the time series. This could be in the SI, but - and yearly temperatures and precipitation (as well as rain v snow) may be informative to the reader as they interpret the data. At several times during the discussion I was trying to evaluate warm vs cold seasons or wet vs dry and it was difficult from the time series alone.

>> This is a good idea thanks, and we agree it will be useful to feature in the main manuscript to better inform the data. We have adapted table one in the revised manuscript, so it summarises by season and year and now includes precipitation.

~If I am interpreting this right, DOC is strongly influenced by water temperature (and all the other factors as well) in the winter (Table 2). I'm struggling to interpret the very large values of node purity (compared with the summer months) and the % variance explained (which is less in the winter). This likely is due to my lack of knowledge with the random forest analysis although it is quite commonly applied. I'm more used to seeing the %Var ttributed to each predictor variable.

>> The node purity tells us the relative contribution of each variable to a models performance but should not be compared between models. Thus, though water temperature has a higher value for node purity in the winter months, this does not mean it is a better predictor than water temperature in the summer months.

~Perhaps a small thing, but air and water temperature are highly related, does this affect the random forest models at all (co-linearity issues)

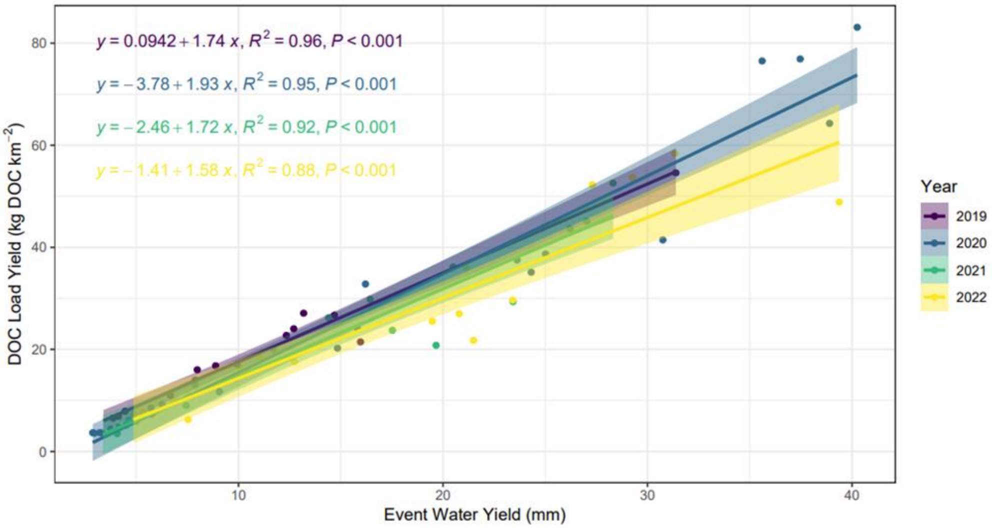
>> By using RandomForests instead of linear regressions, the models shouldn't be impacted by co-linearity issues. At any rate the motivation for including both air temperature and water temperature in the model was that for the snow cover and snowmelt models, the data wasn't strongly correlated thus they potentially represented different processes.

~ Line 426. How as the event water determined? Via isotopes? Please simply state the method.

>> Indeed, event water was determined using isotope separation, we have clarified this in the revised manuscript.

~For Figure 6, how much of the inter-annual differences is driven by the differences in range of event water yield? Some years have very high flows, whereas others are much more modest. I'm wondering if the 'differences' are simply driven by the end-members. Can a test/consideration be done for values within the same range (say < 50 mm). Largely I'm wondering if doing this constrains the process drivers to more common events, etc., and whether the inter-annual variability in load vs event water yield holds. I'm struggling a bit here and in the interpretation to figure out 'why' this would be. Obviously different climate conditions would be the first place to look, but it remains somewhat unresolved.

>> This is a helpful observation. We re-ran the analysis (figure shown below) and you were correct that the large events accounted for the variation between years and by removing these the differences between years were non-significant. Thus, we have edited the discussion in the revised version, and feature this revised analysis without extreme events in the supplementary version to show that it was indeed the case that the large events drove the differences between years. Note that we also removed 2018 from the year to year analysis as data collection only began in September 2018, thus it was perhaps misleading to compare to other years.



L567: " however when extreme events were removed from the analysis (Supplementary Figure 2), differences between the years disappeared. Thus, the relationship between DOC yield and event water yield appears to be consistent between years, with differences between years driven by differences in the extent of annual extreme events"

~Line 469-470. While the forest is mostly conifer, is there much of an understory or mixed forest input of leaf litter? Do you think this affects your interpretation of the DOC sources?

>> The catchment, particularly in the upstream area has an extensive mixture of shrubbery and other non-coniferous vegetation thus this certainly a seasonally variable source of DOC in the catchment. Furthermore, in the peatlands there is extensive dying back of vegetation in the Autumn months which is an ample potential supply of organic matter. With the present analysis we can only speculate exactly to how it contributes to the DOC we observe, but we note that in Figure 1 DOC concentration seems to increase from September onwards which provides some support to the hypothesis that decaying organic matter (detritus) acts as a source of DOC.

~At the end, you mention stable isotopes of water as an area of future focus (line 537). Do you think that looking at DOM quality would be helpful in advancing our interpretation of DOC data?

>> Certainly, and this is something we are planning to do in the future. Determining the seasonal variability in composition and quality would be extremely beneficial to obtaining better understanding of the system. Using a combination of isotope analysis with compositional analysis is something we are particularly interested in a pursuing going forward to better understand how DOM composition varies between different water sources in the catchments and how subsequently this reflects on the DOC data. We have added to the revised manuscript to reflect these thoughts.

L614: "To further enhance our understanding, future research should focus on better understanding how DOM compositional changes impact DOC fate in headwater catchments, and establishing causal relationships between transport metrics, in-stream processing and empirical indicators of sources and transport pathways. A promising avenue for further research involves integrating high-resolution stable water isotope monitoring and spatially distributed hydrological modelling with in-situ DOC monitoring of quantity and quality (e.g. combined fluorescence and absorbance measurements)

Response To Reviewer #2

This paper used high frequency in situ measurements of DOC and discharge from 2018 to 2022 to examine the seasonal variation in concentration-discharge (C-Q) relationship. Overall, this paper presented a novel data set with comprehensive analyses of the C-Q relationship, including the C-Q slopes, the hysteresis patterns, and DOC load yield. The paper is mostly well written and clearly presented although some clarification of concepts and interpretation of results may need improvements . I hope the comments below help the authors improve the manuscript.

>> Thanks for your helpful evaluation and constructive feedback. We have used your comments to make improvements to the revised manuscript.

In the introduction, the authors used the term "transport process" as a general term for what their research focuses on. This term, however, has a very broad meaning and thus it may not be clear to readers what exactly the authors will investigate. I thus suggest the authors clarify what they mean by transport process. Although the scope and meaning of this term becomes clear in the methods, it would help readers understand the paper better if the term can be explicitly explained in the introduction.

>> We have added clarification to the opening sentence of the second paragraph of the introduction to better clarify what we mean by transport processes:

L81: "DOC transport processes (referring to the mobilization of DOC from catchment sources to the stream through differing flow paths) in the Arctic exhibit pronounced seasonality and are highly susceptible to change (Bowering et al., 2023; Csank et al., 2019; Shatilla and Carey, 2019)."

The introduction section does not have a clear statement on the current knowledge gap and how this paper will address that gap. It is not clear what scientific question this paper tries to address. The analysis essentially computed almost all metrics one can do to a concentration-discharge relationship. I think a more clearly defined research question/hypothesis would make the paper's motivation more clear.

>> We attempted suggesting where research was lacking in the final sentence of each preceding paragraph, but we agree the paper will benefit from a clearer overall statement of the research gap at the start of the final paragraph in the introduction. In the revision we have highlight in the revised manuscript our research gap. Chiefly, that there is a need for multi-year high resolution datasets to understand seasonal and annual variation, particularly within the Arctic where seasons are experiencing rapid changes.

L142: “The rapid evolution of controlling DOC processes due to climate change necessitates a pressing need to document transport processes in understudied high-latitude headwater catchments (Shatilla et al., 2023). The scarcity of multi-year, high-frequency datasets in these high-latitude catchments has impeded our understanding of seasonal and inter-annual DOC dynamics. As the underlying drivers of DOC transport processes are undergoing substantial changes, there is a need to understand baseline levels of variability (Shatilla and Carey, 2019; Shogren et al., 2021). This is particularly essential for assessing the dynamic evolution of the Arctic carbon and water cycles, underscoring the need for a concerted effort to address these knowledge gaps. (Laudon et al., 2017; Marttila et al., 2021; Pedron et al., 2023).“

Furthermore, we have added two hypotheses for our two research questions in our introduction to better inform the papers motivations:

L161: “We hypothesized that:

H1 At the intra annual scale, DOC transport processes would significantly differ between snow melt, snow free, and snow cover seasons;

H2 At the inter annual scale, the metrics of DOC transport processes would significantly differ between years with the most different hydrometeorological conditions.”

The authors suggest that the variation in C-Q slope between years were generally smaller than seasonal variation (e.g., line 321). Is this really the case? If we look at figure 3(a), it appears that, at least in November, the among-year variation in C-Q slope is larger than month-to month variation within the same year. I thus think the statement that month-to-month variation is larger is not fully supported by the data.

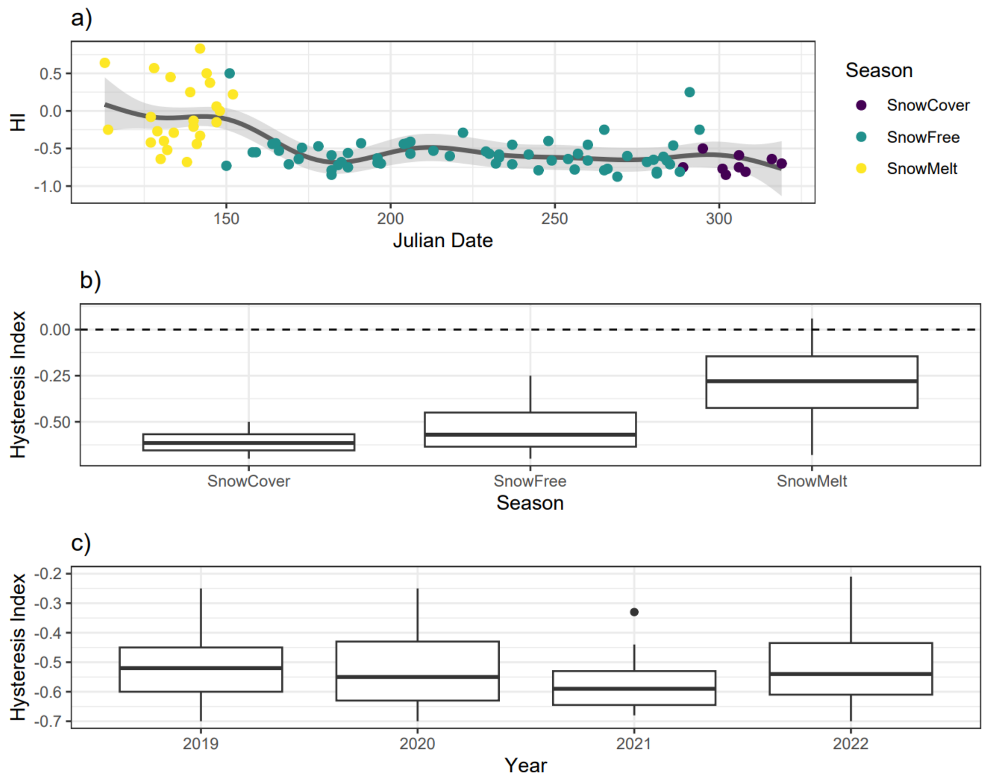
>> We agree it would be more accurate to say that in most months the month-to-month variation is greater than the year to year variation. The differences in April and November were driven by the fact that they vary year to year between being either in complete snow cover for the entire month (meaning very little variation in flow range), to sometimes having large events, which leads to much bigger between year to year differences than other months. Thus we have clarified this in the revised paper:

L542: “However, exceptions were noted in the shoulder months of April and November, where variations in inter-annual transport metric values stemmed from the fact that, in certain years, minimal flow variation occurred due to the catchment remaining in frozen conditions, whereas in other years, large events occurred during these months.”

The authors only used single peak data for analysis of event C-Q relationship. How many events have multiple peaks? If multi-peak events were not excluded, would the results remain the same? Some additional analysis including the multi-peak event would be helpful.

>> We initially used single peak events for consistency across the analysis, but we would agree it is a good idea to examine multiple peak events. We subsequently calculated multiple peak events (by calculating HI for each peak) which we show in the figure below. This added 37 new events to the analysis. The new analysis does not change any interpretation of the results. The

snowmelt season still has the highest average (and most variable) HI, the snowfree and snowcover season are still consistently very negative in their HI values. Hence, we have included this new figure to the revised manuscript, but it did not necessitate further changes to the text (beyond minor changes to the statistical test values) or interpretation thereafter.



Below, I listed detailed line-to-line comments.

Line 187: please provide citation information for equation 1.

>> The equation is derived from our own calibrations of the instrument thus there is no reference, though we have made this clear in the revised manuscript.

L205: "FDOM required corrections for turbidity effects (Downing et al., 2012). Corrections for turbidity were undertaken using the following equation which was derived using internal lab calibration of the instrument".

Line 209-214: Could you please provide numeric criteria you used to define snow cover, snow melt and snow free season. From figure 2, snow depth fluctuates even in what you classified as snow cover season. Wouldn't that cause some period of time to be classified as snow melt season within the currently defined snow cover season?

>> We gave the following criteria. Snow free is the period from which the snow measurement at meteorological station is 0 cm, to the period when the permanent snow cover of the year begins (defined as the point beyond which snowfall never falls to 0 again for the rest of the year. The snow cover season refers to the period where snow depth is greater > 0 till the following spring snowmelt. The snow melt season is defined as the point in spring where a decrease in snow depth occurred alongside flow increase. The logic for classifying snowmelt

in this way was that snow depth often fluctuates for reasons that are not due to melting (for example through snowpack consolidation), so a pure snow-based definition of snowmelt would not be helpful. We did not classify events in early snow cover season as “snowmelt season” as in this study we are interested in seasonal differences, and October/November have very different seasonal characteristics to the snowmelt in spring snowmelt season, which was borne out in the analysis (for example in the hysteresis data).

In the revised manuscript, we have made clear the distinction that the snowmelt season is in specific reference to the spring snowmelt season, and better clarify the numerical definition:

L230: “The snowmelt season was defined as the period starting from the onset of snowmelt, indicated by a decline in snow depth with a concurrent increase in flow, until snow cover at the Kenttärova site reached 0 cm (Fig. 1). The spring snowmelt season was classified using both snow depth and flow as snow depth alone varies for reasons not due to melting (e.g. snowpack consolidation). The snow cover season referred to the period when permanent snow cover occurred (i.e. the point of the year snow depth was > 0 cm till the spring snowmelt). . The snow free season referred to the period between the snowmelt and snow cover seasons where snow depth was 0 cm.”

Line 225: is it each season or each month? From the figure, it seems C-Q analysis was done to data in each month.

>> We have changed this to say “each month” in the revised manuscript.

Line 237: A brief explanation on how HI index was calculated could be helpful.

>> Indeed, we have added a brief description of how the HI index is calculated to better inform the reader in the revised manuscript.

L264: “Briefly, the HI is calculated by subtracting the falling limb standardised DOC value from the rising limb standardised DOC value at each 20th flow percentile across the loop. The HI is then calculated as the average HI of the loop for each event. “

Line 248: the term “dynamics” is a bit vague here. Could you please be more specific and explicit about the meaning of this term here?

>> We have changed this line to be clear that the yield analysis will reflect changes in transport limitation and source activation:

L281 “Differences in the linear regression relationships signify variations in transport limitation and source activation among seasons and years.”

Table 1: Is the unit of flow wrong? Shouldn’t it be L s-1 km-2?

>> Yes, we have changed this in the revised manuscript.

Figure 3: how was coefficient of variation calculated? What is the standard deviation used in the calculation of CV? If I understand correctly, the C-Q slope here is derived by linear

regression using data in each month. Is the standard deviation used in CV calculation obtained from the standard error of slope from the regression? If so, the CV calculated here is not meaningful because the SE of slope from regression shows the uncertainty of estimation, not true variation in slope over time within a month.

>> The coefficient of variation was calculated for the DOC data, not for the C-Q slope data. Thus it shows the CV for DOC by month. The reasoning for including it was to see how much the DOC data varied month to month in comparison to the C-Q data. We have added to the methods to make it clear that the CV refers to the DOC data, and our justification for including it:

L259: “The coefficient of variation (CV) for monthly DOC was also calculated alongside monthly C-Q slopes to identify the amount of variation in DOC relative to changes in C-Q slope.”

Figure 6: similar to my comment to table 1, shouldn’t the unit of DOC load yield be kg DOC km-2?

>> Yes, we have changed this in the revised manuscript.

Line 364: please give exact p-value, not just a range.

>> The individual p-values have now been added for P values greater than 0.001. Below this we use the < sign for brevity.

Line 375: please give the degrees of freedom of the F test statistic.

>> We have added this to the manuscript.

Line 382: what’s tested here is that the slope is statistically different from zero assuming there is a linear relationship. Whether the relationship is linear or not is not tested. Thus, the term “remained strongly linear” is a bit misleading.

>> We have changed the sentence from “remained strongly linear” to “were strongly linear”. While we test slopes, we do also feature individual linear regressions on the graphs to show that in all cases the relationships were linear.

Line 420-432: the explanation here may need further consideration. The trend in C-Q slope seen using data within each month (figure 3) is not evident when analyzing event C-Q behavior. Thus to say that the patterns seen in figure 3 suggests limited source does not reconcile with what is shown in figure 4(a), particular considering that ~60% of flow occurs in events.

>> Figure 3 is all flows, which includes event flow, therefore across the entire range of flow the catchment is more limited in source (or less transported limited) in the snowmelt period compared to other months. In Figure 4 we compare only event flows. I think the important distinction to add would be that at the kind of flows that constitute event flow, there is a tendency for all events to become much more similar (ie very weakly transport limited), likely due to increased depletion of sources at higher flows (as shown by the lower C-Q slopes for all months for events compared to the entirety of flow). Thus, we will discuss the difference

with event flow, although keep the assertion that snowmelt was more source limited across the entire flow-range.

L485: "When exclusively examining event flows rather than encompassing all flow conditions (Fig. 4a), no significant disparity was observed between snowmelt conditions and other months. In contrast, C-Q slope values exhibited a general reduction across all months, when compared to slope values for all flow conditions. This reduction may suggest that transport limitation reduces, and increased source depletion occurs relatively quickly after flow increases beyond baseflow conditions across all seasons."