Reviewer 1

General comments

Reviewer comment: *The topic of DOC mobilization and fate is well within the scope of HESS, but the manuscript needs to be improved substantially before publication is possible.*

Reply: Thanks for your positive and constructive comments. Below, we provide a point-by-point response to all your general and specific comments.

Reviewer comment: *Some explanations of the model structure and setup are unclear. The paragraph explaining the strategy to account for uncertainty in the simulations (L.232-244) is on the one hand hard to follow and some decisions seem arbitrary. On the other hand, simulation uncertainties are not discussed at all in the manuscript, even though they are shown in the main results figure 3.*

Reply: We have revised Sections 2.3 and 2.4 to better clarify the model structure and setup. In particular, we have changed how we handle uncertainty in the model and we hope that the clarity of our explanation of this approach in Section 2.4 is improved. In addition, we further address sources of uncertainty in the Discussion section.

Reviewer comment: *The model performance is rated by using four different likelihood measures and pre-defined thresholds for behavioural models. When the model performance is discussed, it seems arbitrary which likelihood measure is highlighted. The discussion would be stronger if the model performance evaluation would be more systematic.*

Reply: That is a fair point. In the new version of the manuscript, we focused on two metrics: correlation and PBIAS. We have reported the performance metrics more systematically and have visualized the metrics in figure 5 and 6.

Reviewer comment: *Further, the influence of DRIPs on DOC concentrations is only visible (in the sampling data) if the gain in streamflow is > 40% (Figure S1). Otherwise, the lake outflow dominates, which was also stated in the discussion. Thus, events A, B, F, G and H are the most important for the testing of the hypothesis. These events should be discussed in more detail. For example, 'diff_nobio', which does not consider UCAs, often performs better or similar good compared to the UCA-version. Is the UCA-concept really helpful or are the stream concentrations rather driven by the groundwater concentrations and not the inflow volumes? If the UCA-concept is helpful, in which hydrological situations? After a structured analysis and discussion, more specific conclusions should be drawn.*

Reply: We agree and we have emphasized these events in the results and discussion.

Specific comments

Reviewer comment: *In my opinion, the title does not correctly represent the content. In the manuscript, the model itself is rather used as a tool to analyse processes. The title, however, suggests that the main focus is on model structure and testing.*

Reply: Following both reviewer suggestion, we change the title to "Groundwater flow paths drive longitudinal patters of stream DOC concentrations in boreal landscapes".

Reviewer comment: *L.16. Change 'concentration' to 'concentrations'*

Reply: Done (line 16).

Reviewer comment: *L.29. The word 'Importantly' seems inappropriate here.*

Reply: Change to "Further" (line 29).

Reviewer comment: L.55. Change 'By contrast' to 'In contrast'

Reply: This sentence no longer exists.

Reviewer comment: L.58. Remove 'a'

Reply: This sentence no longer exists.

Reviewer comment: L.79. Remove 'what' and 'is'

Reply: This sentence is rephrased (line 70).

Reviewer comment: L.89. Remove 'than'

Reply: Done (line 78).

Reviewer comment: *L.89. The term 'passive pipe' might be misleading. Pipe flow is a specific type of hydraulic flow, which is not addressed here at all. Probably better refer to 'non-reactive' or 'recalcitrant'.*

Reply: The term "passive pipe" is commonly used in stream biogeochemistry to refer to those streams that do not produce or retain organic matter; thereby "passive" transporting it to downstream ecosystems (see Casas-Ruiz et al. 2017 and references therein). Yet, to avoid confusions, we now refer to "streams that do not take up the supplied DOC" (line 78).

Reviewer comment: *L.136. Change 'talweg' to 'thalweg'.*

Reply: Done (line 118).

Reviewer comment: *L.141. What kind of filters were used for filtering (filter material)? How long was the time between sampling and lab-analysis? Is there a reference explaining DOC sampling and analysis in more detail?*

Reply: We filtered all water samples with 0.45 µm MCE syringe filters (Millipore®). We kept all samples refrigerated at 4ºC before DOC analyses (< 7 days after filtering). DOC analysis consisted of acidification of the sample to remove all inorganic carbon, followed by combustion using a Shimadzu TOC-VCPH (analytical error: 3%; Laudon et al. 2011). The analysis was repeated at least three times per sample resulting in a DOC concentration in mg/l and a percent standard deviation. Following the reviewer suggestion, we added all this information in the methods (lines 120-125).

Reviewer comment: *L.165. How were the reaches classified into DRIP and non-DRIP zones?*

Reply: We classified reaches based on if they comprise a DRIP or not. DRIP were located previously from topographic and isotopic measurements. Therefore, 5 reaches were classified as "DRIP zones" and 20 reaches as "non-DRIP zones". We clarified this issue in the methods section lines 89-95.

Reviewer comment: *L.166. What do you mean by '…was weighted based on the mean…'? Did you want to say '…was derived by using the average of all…'?*

Reply: Yes, we meant that. We change the sentence accordingly (lines 146).

Reviewer comment: *L.168. Change 'upslope contributing are' to 'UCA'. It was already expanded in the introduction.*

Reply: This part has been rephrased (line 144)

Reviewer comment: *L.217. Remove 'yielded'*

Reply: Done.

Reviewer comment: *L.232. In this paragraph, the method for uncertainty estimation is explained and the uncertainty bounds are shown in figure 3. However, the uncertainty of the model results is not discussed at all in the results and the discussion sections. Please add* it.

Reply: We have added discussion the uncertainty in section 4.2, line 323.

Reviewer comment: *L.233-234. '…we compared the range of uncertainty in our simulations to the range of uncertainty in the observations.' Where did you do that? I can't find it in the manuscript.*

Reply: We have removed this sentence**.**

Reviewer comment: *L.234 The number of model runs (100 for each scenario) appears really low, especially for such a simple and fast model (in terms of computational time). Usually, several thousands of runs are performed for uncertainty estimation. How did you check that 100 runs are enough for stable uncertainty bounds?*

Reply: Following the reviewer suggestion, we now run each model 10,000 times.

Reviewer comment: *L.236. Why did you choose to use 66% uncertainty bounds? More common are 80% (10^{th} - 90^{th} quantile) or 90% (5^{th}-95^{th} quantile).*

Reply: We changed the uncertainty bounds to min and max.

Reviewer comment: *L.246. I understood that 100 model runs were performed for each scenario. Which one was evaluated with likelihood measures? All? Only the best? Which values are shown in table 2?*

Reply: We estimated the likelihood metrics R and PBIAS for each run (i.e., n = 10,000); and then presented the boxplots.

Reviewer comment: *L.252-254. The description of R^2 is - at least – unusual. I would recommend using a more straight-forward description. In this context, the most important thing might be, that R^2 shows the ability of the model to simulate the dynamics of the measured time series, but not the absolute value.*

Reply: The reviewer is right. We changed the sentence accordingly (line 196).

Reviewer comment: *L.265. The reference to 'Fig. 3F, K-M' should rather be 'Fig. 3E, K-M'*

Reply: Thanks for noticing! This sentence is no longer in the manuscript.

Reviewer comment: *L.277. change 'NSE<0' to 'NSE was <0'*

Reply: This sentence is no longer in the manuscript.

Reviewer comment: *L.295-296. Probably it is worth mentioning that R^2 generally has problems with evaluating uniform values such as D and I.*

Reply: We have changed the metrics to PBIAS and R

Reviewer comment: *L.315. Please re-format the table. Floating point numbers and scenario identifiers are in two lines, which makes them hard to read.*

Reply: The reviewer is right. We changed the table to a figure

Reviewer comment: *L.319-328. This paragraph reads like a conclusion/summary. At this point, however, the results were not yet discussed. The paragraph should be removed.*

Reply: Following the reviewer suggestion, we removed this whole paragraph.

Reviewer comment: L.337. change 'influence stream' to 'influence on stream'

Reply: Done (line 267).

Reviewer comment: L.341. Remove 'of'

Reply: This sentence is no longer in the manuscript.

Reviewer comment: *L.347 The mentioned UCA_BIO is not the best model under the conditions of 3E (Neff=0.5; R^2=0.84). DIFF_BIO performs much better with Neff=0.81 and R^2=0.86.*

Reply: Thank you for noticing. The metrics have been changed.

***Reviewer comment*:** L.364. Change 'that' to 'than'

Reply: Changed, line 296

Reviewer comment: *L.364-365. Change '... spatial variability in groundwater ...' to ''... spatial variability in groundwater concentrations…'*

Reply: Changed, line 297

Reviewer comment: *L.402. '…demonstrated that UCA can be useful to identify "reactive" reaches…' Actually, the model results don't show a clear signal towards UCA explaining the variability better. Often, the Diff-version is similar good or even better.*

Reply: this sentence is no longer in the manuscript

Reviewer comment: *L.419ff. In the conclusions, UCAs are not mentioned at all. Since they play an important role in the whole manuscript, please add your conclusions about this concept.*

Reply: The UCAs are in a way mentioned because DRIPs represent them.

Reviewer comment: *L.425. 'occasionally' is imprecise language. Please, be more specific.*

Reply: Removed

Reviewer comment: *L.429-431. 'This study …. (Mineau et al., 2016)'. This is not a conclusion of this study. Delete or shift the sentence.*

Reply: Removed

Reviewer comment: Figure S1. I would recommend adding figure S1 to the main text. I think the different hydrological conditions, even though they are provided in the text, are much easier to understand with Figure S1.

Reply: The figure is added to the main text

REVIEWER 2

General comments

Reviewer comment: *This is a potentially interesting paper but one that needs some attention before publication. There are several aspects of the paper that rely on results from previous papers. Some, particularly the locations of DRIPs, need more explanation here. Other aspects (such as the discussion of uncertainties) also need more detail.*

Reply: Thanks for your detailed and constructive comments. Following the reviewer suggestion, we now provide a detailed explanation on how DRIPs were identified (lines 94- 94), as well as included DRIPs in Fig. 1 to better show where they are. Further, we also provide a detailed discussion of uncertainties (lines 323-328). Please, see the response to your specific comments on this regard).

Reviewer comment: *I also do not get a clear sense of how this paper helps address important problems or knowledge gaps. A clearer explanation in the Introduction of where those gaps are and the Conclusion of the broader aspects would help here. It is an important topic, but the importance needs to be better explored.*

Reply: That is a fair point. We rewrote the introduction and conclusions to highlight the important knowledge gaps it addresses. Specifically, the introduction now highlights that accounting for spatial patters of stream DOC concentrations is valuable for understanding in-stream C retention along rivers (Bernal et al., 2018) and catchment-integrated evasion of C (Wallin et al., 2013), as well as for assessing and managing the brownification of large water bodies and coastal ecosystems (Kritzberg et al., 2020). Yet the main drivers controlling the spatial variation in DOC concentration remain unclear, partly because processes occurring at various scales interact to influence the export of DOC to downstream aquatic ecosystems (Laudon and Sponseller 2018). In boreal landscapes, the organization of groundwater flow paths can regulate spatial patterns of stream DOC concentration by (i) supplying terrestrial DOC to streams and (ii) promoting in-stream C cycling. The role of DRIPs as DOC suppliers and in-stream C cycle promoters likely depends on catchment hydrology. Since the entire stream network relies on the inputs and fate of DOC in headwaters (Raymond et al., 2016), it is important to understand where and when DRIPs hydrologically connect to headwaters, as well as their broader effects on stream C cycling and overall catchment C exports. This is especially relevant in boreal regions, which store a large fraction of the Earth's organic soil C and particularly susceptible to climate change (Bradshaw and Warkentin 2015).

Reviewer comment: *The paper is generally clear but has minor grammatical errors throughout and occasional odd phrasing. I have sympathy with having to write in a second language, but checking the English would help communicate the research better.*

Reply: Thank you. Following the reviewer advice, we checked the English in the new version of the manuscript.

Specific comments

Reviewer comment: *Not sure that this reflects the paper's contents well. There is no explanation of parsimony in the paper (ie what are the alternative 'non-parsimonious' approaches?) If parsimony is important it needs to be reflected in the Introduction and Discussion.*

Reply: Following both reviewer suggestion, we change the title to "Groundwater flow paths drive longitudinal patters of stream DOC concentrations in boreal landscapes".

Reviewer comment: The Abstract is a reasonable summary of the paper, but there is a lot of focus on what was done rather than what was concluded. Also as with the rest of the paper try to add some indication of importance in the statements at the start and the concluding statements at the end of the abstract.

Reply: we have revised the abstract

Reviewer comment: The introduction is reasonably comprehensive but suffers from some odd terminology. More importantly, it does not convey a sense that the paper is addressing an important general question. What is it that we really don't know that this paper will help understand? DOC (and C in general) in rivers is important but there have been many studies over recent years. If you explain what the key gaps in our understanding are and how this study helps address those, then the paper will have more impact with the scientific community. Your aims should also refer back to those broader themes.

Reply: we have rewritten the introduction to cover these topics.

Reviewer comment: *L35. "Running waters" is odd terminology if you specifically mean rivers.*

Reply: We changed it to "streams and rivers".

Reviewer comment: *L39-43. Point not clearly made. There are obviously differences in sources of water, retention times, and flowpaths in headland streams but HOW do those result in differences in DOC? Terms such as 'dynamics' are a bit vague, better to explain exactly what you mean (sources, loads, reactions etc).*

Reply: We have avoided the use of dynamics and rewritten the introduction.

Reviewer comment: L52. *What is a 'spiralling framework'?*

Reply: The nutrient spiraling framework (Webster 1975) indicate that, in streams and rivers, carbon and nutrients do not cycle in place, but are transported downstream as they complete a cycle. While this framework is widely used in stream biogeochemistry, we decided to use the term "in-stream C cycle" or "in-stream DOC processing" throughout the manuscript.

Reviewer comment: *L61-66. Is this universally true? Reactions in groundwater likely reduce DOC loads and, in many catchments, the deeper groundwater seems to have lower DOC concentrations than surface runoff. In which case groundwater inflow may dilute surface water DOC concentrations. Perhaps be more circumspect here?*

Reply: The reviewer is right; deep groundwater generally has lower DOC concentrations than streams (e.g. Tiwari et al. 2017). In boreal headwaters, the contribution of deep groundwater to stream flow is minimal (Tiwari et al. 2017), and most water enters to streams through near-surface groundwater flow paths that convey substantial fluxes of water from large upslope contributing areas to narrow sections of the stream, referred as discrete riparian input points (DRIPs). These near-surface groundwater flow paths influence riparian soils through effects on plant community, redox conditions and organic matter accumulation (Grabs et al. 2012; Kuglerová et al. 2015), leading to greater concentrations of labile DOC in DRIPs than in more diffuse groundwater flow paths (Ploum et al., 2020; Demars et al., 2021). Following the reviewer suggestion, we added this information in the introduction (lines 45-49).

Reviewer comment: L81-90. *Your aims are fine as far as they go, but I don't get a good idea of how these fit in to a broader understanding of the science.*

Reply: Following the reviewer suggestion, we rewrote the introduction to better state the novelty of the study (*see* our response to your general comment). Further, we now also indicate that our aim was to assess the relevance of DRIPs as primary drivers of spatial patterns of stream DOC concentrations along boreal headwater streams (line 71). Resolving these land-stream interactions is critical for better integrating headwater streams within the broader carbon cycle, as well as for understanding how they affect carbon transport and processes in downstream ecosystems.

Reviewer comment: *L89. Not sure the 'passive pipe' argument is valid. It is generally accepted that streams are dynamic environments, particularly around reactions involving compounds such as DOC. This makes it seem that you are addressing a problem that may not exist.*

Reply: Yes, it is accepted that streams are dynamic environments affecting carbon cycle (e.g. Mineau et al. 2016). However, the timing of inputs and fate of DOC in headwater streams remain largely unknown. For instance, while most DOC might be transported downstream during peak flows (pulse-shunt concept; Raymond et al. 2016), changes in hydrological flow paths can also bring bioreactive DOC facilitating its biogeochemical process (hot spot concept, McClain et al. 2003). By referring to these concepts throughout the manuscript we clarified this distinction.

Reviewer comment: *Given the importance of understanding the groundwater inflows to the stream and the origins of the groundwater, more details are needed here. It seems that some of these are in other papers but better explanation would make this paper more convincing. Specifically:*

Reply: Following both reviewers suggestion, we included more information about DRIPs in lines 86-94.

Reviewer comment: *L112-114. It is not clear what you mean by "route 60% of the upslope contributing areas to 5% of the stream length". I presume that your DRIPs occur along 5% of the stream, but where does the 60% area come from? This is probably in the Leach et al. paper but more details are needed here.*

Reply: We meant that despite DRIPs discharge into a very small fraction of the stream (only 5% of its total length), they collectively drain 60% of the groundwater inflows discharging along the C5-C6 stream segment (line 93).

Reviewer comment: *What exactly is the groundwater? Are you sampling only shallow riparian zone waters or a mix of those and deeper groundwater? Explain how deep the bores are and what units they are screened in.*

Reply: The wells had a mean depth of 95 cm (σ = 37 cm) and were fully screened. Therefore, we assumed that the water sampled was a weighted average of the phreatic aquifer, down to the depth of the well. Yet, it is important to note that 90% of the DOC is transporter thought near-surface flow paths (top 50 cm) in this catchment (Ledesma et al. 2015). We added this information in the methods (lines 115-116) sections.

Reviewer comment: *How are the DRIPs identified and do you have diffuse discharge between the DRIPs?*

Reply: Following the advice of both reviewers, we now state how DRIPs were identified (lines 86-94). We extracted DRIPs and their upslope contributing area from a digital elevation model with a resolution of 2 x 2 m, assuming that there is no diffuse flow between DRIPs.

Reviewer comment: *Some descriptions (e.g., shallow and near-surface) could be more specific (i.e., are the groundwater levels above the land surface?)*

Reply: There are occurrences of saturated overland flow - mostly at the DRIPs. For example, Leach et al 2017 mentions: "The stream reach has no tributaries although saturation excess overland flow is sometimes visible in select areas along the riparian zone during wet conditions." We have removed the terms shallow and near-surface throughout the manuscript.

Reviewer comment: *L98. What underlies the regolith and soils? Are there deeper aquifers that can contribute groundwater and if so do we know anything about that groundwater?*

Reply: As most headwater streams in Krycklan, deep groundwater inflows are minimal in this stream segment (Tiwari et al. 2017). Indeed, a previous study showed that 90% of the DOC is transporter thought near-surface flow paths (top 50 cm) in this catchment (Ledesma et al. 2015). We clarified this issue in the Study Area section (line 95)

Reviewer comment: *L100-110. This description is difficult to envisage. You should illustrate the variations in river flow and groundwater responses on a Figure. Most of those details are in Fig. S1. Given that you have only 3 figures in the paper (and only 1 is multipart) you could easily fit Fig. S1 into the main text (and also Fig. S2 if you wanted) – figures in supplements rarely get looked at.*

Reply: Following the reviewer suggestion, we included Fig S1 in the main manuscript (new Fig 2).

Reviewer comment: *Figure 1. What are the contours (I presume elevation)? Given the topic of the paper, can you show groundwater flow paths – you refer to these in the caption (L120) so presumably you know something about them.*

Reply: That is a good point! We changed Fig 1 to show the location of preferential groundwater flow paths.

Reviewer comment: *L123-127. That is not really a drought. Droughts imply that the catchment as a whole dries up, which impacts more than just river flows (groundwater levels, soil moisture are also impacted). While, this analysis is useful, 'artificial low flows' or something similar would be better.*

Reply: Following the reviewer suggestion, we now use the term "artificial drought" to refer to the lake damming experiment (line 108). Yet, it is important to notice that, during the course of the artificial drought, the strength of DRIP-stream hydrological connections declined, generating a patchy distribution of lateral DOC inputs similar to those occurring under natural droughts (Gómez-Gener et al., 2020). Hence, we consider that our results could also be applied to natural drought conditions as well. We added this information in the manuscript (lines 109-110).

Reviewer comment: *L132-135. Not clear how the DRIPs were identified from the wells (which is what you imply). Or do you mean that you located the wells following identification of DRIPs? Again, I think that these details are in other papers but would be better summarised in section 2.1.*

Reply: Yes, we first identified the location of DRIPs and then installed the wells. We clarified this procedure in sections 2.1 and 2.2 (lines 89-95).

Reviewer comment: *L136. "thalweg"*

Reply: Done (line 118).

Reviewer comment: *L137. "evacuated" not "vacuumed".*

Reply: Done (line 119).

Reviewer comment: *L146. Report the typical or range of analytical uncertainties here.*

Reply: The analytical error for DOC concentrations was 2% (line 124)

Reviewer comment: *L165-169. As noted above, the classification of reaches and location of DRIPs needs more justification. It is important to the study. It seems that much of this comes from the geometry of the catchment, specifically the UCA., with the details only in other papers. It may be correct, but as presented here it is not very convincing.*

Reply: Following the reviewer advice, we now explain in section 2.1 and 2.2 the procedure to identify the location of DRIPs and the subsequent classification of reaches (lines 89-95).

Reviewer comment: *L176-177. The interpretation that all increase in streamflow is due to groundwater inflows is also important and needs more detail. From Figure 1 there looks not to be any significant tributaries, but what about overland flow, small rivulets, draining pools etc (especially at high flows). You should be more specific in ruling those out if they do not exist.*

Reply: The stream segment did not have any tributary, and not rivulets nor draining pools were observed during the study period. We added this information in section 2.1 (lines 89-95).

Reviewer comment: *L198. see comment earlier about 'passive pipe' and perhaps come up with a better term.*

Reply: Following the reviewer advice, we now state that one scenario considered that all terrestrial DOC inputs were transported to downstream ecosystems (i.e. no in-stream DOC uptake, pulse-shunt concept) (line 167)

Reviewer comment: *L199-201. Not very clear what you mean by 'upstream', do you mean above the lake?*

Reply: Yes, we meant the upstream lake. We changed it accordingly (line 169).

Reviewer comment: *L232-255 Do you present the uncertainties in the paper (it is not clear that you do)? There should be more discussion in the results or discussion sections.*

Reply: We have revised the explanation of how we treat uncertainty in modelled stream DOC. The model uncertainty is represented by the range in modelled values when comparing the 10000 model runs. We also discuss further issues around model uncertainty in the discussion section.

Reviewer comment: *L234-239. Is there any reason for only using 100 model runs? I would have thought that the model is not computationally limited and more runs would be possible. If increasing the number of runs resulted in the same rand and distribution of outcomes than you should state that. I agree that studies should not do redundant analyses, but knowing where redundancy occurs is useful.*

Reply: Following the advice of both reviewers, we not run each scenario 10,000 times.

Reviewer comment: *L259-265. This is really a summary of the results and would make more sense after you have described the data from the different studies. Suggest moving to the end of this section or to the beginning of the Discussion.*

Reply: The result section has been fully revised and this part no longer exists

Reviewer comment: *L283. See comments earlier about calling this a 'drought'*

Reply: We have rephrased the term drought and refer to artificial drought

Reviewer comment: *Table 2. Difficult to read (final version needs line breaks fixing)*

Reply: We have replaced the table with two figures

Reviewer comment: *Sections 3.1 to 3.3. These are detailed but a bit dense to read. The level of detail also tails off a bit. In 3.1, you report the DOC concentrations but not in the other sections. Given the level of detail here, it is difficult to have to keep looking at Figure 3 to see the changes. Adding the values in here would help.*

Reply: The discussion has been fully revised

Reviewer comment: *I think that a better discussion of the uncertainties (which are the most important, does that differ between the different conditions is needed). All I understand to this point is that the uncertainties were used to generate the predictions in Figure 3, but not what the main uncertainties are.*

Reply: We have discussed the uncertainties in the newly revised discussion under section 4.2

Reviewer comment: *L325-329. Using the model names introduced above would help to key this discussion into the previous sections and Figure 3.*

Reply: We incorporated model names.

Reviewer comment: L347-349. Do you mean that DIFF BIO is the better model (not UCA BIO)? The NSE and R2 in the table would indicate that was the case.

Reply: We have revised this part and changed the metrics, so NSE and R2 are no longer considered.

Reviewer comment: L355. ‘drought’ and ‘pipe’ comments again.

Reply: we have avoided the terms

Reviewer comment: L361-368. This is written as if the catchment is really considered to have been in drought (rather than just the flows being artificially low). Are these findings what you would expect if the catchment was relatively wet and just the flows were impacted?

Reply: This part of the discussion has been rewritten.

Reviewer comment: As with the Introduction, there is no broad overview in the conclusions. As it is the Conclusions are understated. A final paragraph or two explaining the broader relevance of this study or outlining what you think you have done which is novel is needed. Be more specific how you contribute to the “greater goal” of understanding these processes. That would make the study more interesting and impactful to a wider audience. Perhaps speculation on what this means for DOC if flow regimes change would be interesting and useful here.

Reply: We have rewritten the conclusion