REVIEWER 1

Comments by the reviewer are in black text

Our responses are in blue text

Where applicable, proposed changes to the revised manuscript are in blue and italics

GENERAL COMMENTS:

The manuscript deals with an important and debated problem, the permafrost carbon feedback. More specifically, it uses a numerical simulation approach for assessing solute transport in a permafrost catchment, in order to study the impacts of climate change on DOC export from the soil of the considered watershed. The mechanistic modelling of dissolved species in permaftrost affected areas is at the forefront of the current research in cryohydrogeology, and the used tool is a state of the art one. Thus this work is both novel and of great interest for cryospheric sciences.

Meanwhile, critical weaknesses affect its reliability. First of all, the study focuses on a specific site, while the claimed aim of the work is to draw general conclusions about lateral transfers in permafrost regions. This is contradictory, and if the authors want to produce a general study then the problems of upscaling and transposing to different biogeoclimatic permafrost contexts must be carefully dealt with. I have also a number of technical concerns about the set up of the performed numerical simulations, for instance the poorly justified pseudo-3D approach, the absence of any convergence study, and some problems in the definition of boundary conditions. Finally the discussion of the obtained simulation results is rather weak and should be strengthened. All these points are detailed in the specific comments below.

Overall I think that the manuscript cannot be published in its present form. A thorough effort is needed for making it more solid and meaningful, including possibly new computations and results depending on the answers to some of my concerns (e.g.: numerical convergence, relevance of top and bottom boundary conditions). So I recommend that a major revision of this work should be undertaken prior to reconsider whether or not it should be published in The Cryosphere.

We thank the reviewer for their thoughtful review. In response, we plan to make several modifications to enhance the manuscript. Our primary focus will be on providing a thorough justification of the meshing strategy, which was not adequately addressed in the original version. Additionally, we will clarify the applicability of our results to a broader pan-Arctic context, improve the clarity of the figures, and include additional descriptions where necessary. We are confident that these revisions will significantly strengthen the manuscript.

SPECIFIC COMMENTS:

1. - Title and l 8-9: The title refers to ‘a permafrost catchment’, while abstract draws conclusions for ‘permafrost regions’. There is no *a priori* reason that conclusions made from the study of a peculiar catchment should be relevant for permafrost regions as a whole.

We agree that the conclusion that our results apply on a larger scale may be misleading as it is currently written. While it is not possible to generalize our results for the entire northern hemisphere permafrost region, the results from this study can be extended to other permafrost hillslope catchments with similar climatic, hydrological and geological conditions. This site considered in our study is representative of gravely footslopes with a gentle inclination (5-15°, Hamm and Frampton 2021). We will adjust the phrasing in the abstract and will improve the discussion on how representative these results may be in a broader context of similar permafrost hillslope regions.

2. - l 17: Why considering mineralization of organic matter to CO_2 only, and not also to CH_4? May be due to the peculiarity of the study site. This should be explained.

We included this information in the methods section, rather than in the abstract. To better highlight our decision, we will add more explanation as to why we do not consider mineralization into CH4. The main reason is that the study site is a hillslope and compared to wetlands and peatlands is unlikely to contribute significantly to overall permafrost CH4 release. We will add a reference to this statement in the methods section.

3. - l 27-41: Other questions than PCF maybe related to lateral transfer of dissolved chemical species from the active layer under climate change (e.g.: impacts on river / ocean ecosystems).

We will add a sentence on potential impacts on river ecosystems and the ocean, such as rusting rivers, mercury release, and ocean acidification.

4. - l 71-72: The assumption that DOC may be considered as a non reactive tracer should be discussed here, with relevant references.

A passive tracer is used to numerically represent advective solute transport in the simulations; the conceptualisation used is that a given quantity of DOC is transported by the advective flow field but can undergo mineralisation which is dependent on soil thermal and saturation dynamics. Although this does not account for the full complexity of carbon reactions it embodies an important advancement in coupling soil and groundwater dynamics with potential DOC transport and is a novel contribution of our study. The implications are mentioned in the introduction but further discussed in the discussion section. We refrain from adding it to the introduction section.

5. - l 80-82: Most likely this behavior will be site dependent! For instance not the same for a tundra hill slope with low evapotranspiration or a boreal forest hill slope with high evapotranspiration, depending on the slope itself, etc.

The emphasis on the site-specific hypothesis is missing. We will add it.

6. - l 111: Is Qs sink/source term accounting for ionic exclusion? To say, does the used model take into account the specificity of solute transport with freeze/thaw?

Yes, exclusion of solute from the ice is included, although it is enforced by the nonlinear solver and the formulation, and not in the Qs term. The solute in ATS is represented as a moles of solute per moles of liquid water. This way, when the water in the soil freezes, the solute remains in the unfrozen water fraction of the soil and solute concentration increases to honor mass conservation, which is enforced by the nonlinear solver. Upon melting of the pore ice, the opposite occurs. We propose to add the following text in the methods section to clarify this:

"Note that solute is excluded from the ice phase by the formulation which represents solute as moles of solute per moles of liquid water and by the nonlinear solver which enforces solute mass conservation. Thus freezing of pore water results in an increase of solute concentration and the opposite is true for melting of pore ice."

7. - l 113: Why neglecting dispersivity? The errors associated with this simplifying assumption should be discussed.

This study deliberately focuses on lateral transport by advection. Although it would be interesting to also consider dispersion, it is highly substrate-dependent and can not easily be quantified with field observations. Hence, any assumptions of dispersion coefficients would add to parameterisation uncertainty and obfuscate the advection-specific results discovered here. For example, the observation of vertical tracer movement by wetting/drying and cryosuction could have been misinterpreted as dispersivity had it been included. We will add an explanation to the text better explaining what the consequences of this decision are.

8. - l 118-119, « path of highest flow accumulation »: What about lateral transfers from the slopes to the thalweg? For studying the drainage of active layer waters and the associated solute fluxes the choice of focusing on the thalweg line does not seems obvious to me. This should be discussed.

We will discuss these choices more thoroughly in an expanded description of the meshing process (please see next comment)

9. - l 121, « pseudo-3D approach »: I cannot understand the usefulness of varying y-axis width of cells along the x direction. According to Fig.2, there is only one mesh along y-axis at every x position. So this seems to me a 2D mesh (as said in l 169). The Figure 3 and the associated explanations are hard to follow and should be thoroughly improved (e.g.: a large blue area appear on Fig.2, but only green and red area are mentioned in the text.) Whatever the pseudo-3D approach really is, the statement made by the authors that it « allows us to account for thermal and hydrological balances across the entire catchment area without the need for a complex and computationally intensive full 3D mesh »

should be justified. How does the y-axis width varies along the thalweg? Has any comparative study between results of a full 3D approach and results of this pseudo-3D approach been done? It should have been, prior to use the pseudo 3D approach, or at least if it is not practicable due to computation time the approximation should be discussed, as well as the associated errors.

We acknowledge that the justifications as to why we chose the particular meshing approach were not adequately explained. We will revise that section to better provide the justification. The key point here is that the use of a variable-width representative hillslope is a well-established modeling approach in catchment and hillslope hydrology. In this approach, the sides of the hillslope are determined from topography by gradient descent, thus aligning the sides with the surface flow and allowing for no-flow boundary conditions on the sides. The fact that the width is now variable results directly from the 3D topography and preserves flow convergence/divergence, which is well-established to be important. We propose to change the corresponding paragraph in the revised manuscript is as follows:

"*To capture the main topographical characteristics of the Endalen sub-catchment without the computational expense of a fully three-dimensional model, we created a variable width representative-hillslope mesh (Fig. 2) extending 1040 m along the direction of highest flow accumulation from the catchment divide (376m relative elevation) to a groundwater spring close to the valley bottom (0m relative elevation). The sides of the delineated hillslope are aligned with the topographic gradient and thus the surface flow, which results in variable-width mesh elements. Such hillslope-based approaches to approximating a three-dimensional landscape with a lower-dimensional model have a long history in catchment hydrology (e.g. Fan and Bras, 1998; Troch et al. 2003; Hazenberg et al. 2015) and have been successfully compared to three-dimensional models (e.g. Hazenberg et al. 2015; Paniconi et al. 2003). Importantly, the plan shape and profile curvature of the catchment, the dominant topographic controls on flow (Dunne and Black, 1970; Anderson and Burt, 1978; Freeze, 1971), are preserved by the variable width mesh.*"

We will also expand on the explanation of the mesh by moving mesh-specific information to the Appendix to be able to elaborate more thoroughly without increasing the already lengthy main text significantly. We will add details to the representation of the catchment hydrology, the boundary conditions, and the importance of accounting for all the water in the form of precipitation within the catchment rather than only accounting for a short transect and the precipitation falling onto its surface. We will also add a sentence on the complexity of the processes represented in the model and how this does not allow for a feasible representation in 3D due to computational costs.

10. - l 126-127, « main area of interest »: Why focusing on such a tiny area of 20 m length, while simulations are done on a 1040 m large domain? This important choice should justified and discussed.

As described in the text, the upper parts of the transect are mainly in place to provide the adequate amount of water to the lower parts of the slope, where organic matter can be expected to exist in the permafrost. We will add to this explanation by highlighting that the rocky nature of the upslope part leads to very low carbon abundance and that carbon transport in this section is not of interest. Only in the lower parts of the slope, where the inclination is not as steep anymore, carbon may be present. We will further add that the computational costs for these simulations were very high and due to the mesh discretization, a longer transect with a higher lateral resolution to observe transport would significantly increase these costs.

11. - l 129-130, « By directing precipitation from the upper slope to the lower areas, we ensure realistic hydrological conditions with flow accumulation towards the valley bottom. »: What is meant here? The physical equations solved by the numerical model do ensure that gravity exerts a vertical descendant driving force on water, so that water flows from top to bottom when gravity dominates. This sentence seems pointless, I recommend to delete it.

This will be part of the reworked mesh-section described in comment #9 that will be added to the Appendix. The sentence will either be deleted or rephrased for clarity.

12. - l 130-131,« This division of the mesh allows for accurate modeling of the thermal-hydrological processes in the catchment. » : Such a statement should be justified.

This will be part of the reworked mesh-section described in comment #9 that will be added to the Appendix.

13. - l 132, « Each column in this mesh area [...] varies in width in the y-direction. »: Why and how? See above the point on the pseudo-3D mesh approach.

This will be part of the reworked mesh-section described in comment #9 that will be added to the Appendix. We will add information on how the mesh has been created and that its shape has been chosen to represent the approximate catchment area derived from the digital elevation model.

14. - l 132 – 141: Here mesh cells dimensions are described, but not justified. In a modelling study relying on PDE spatio-temporal discretizations (e.g.: finite differences method, finite volume methods, finite elements methods, etc), it is mandatory to assess the truncation errors by dedicated convergence studies, designed for the simulation case under concern. The results of such a convergence study should be given here, by means of a upper bound of the truncation errors for the outputs of interest. Only in this way one can be sure that the variations discussed in the numerical results are not simply due to truncation errors. Please include here the results of such convergence study for the case under consideration.

Were the goals of this study to make a precise prediction of an effect of interest or to demonstrate a new discretization method or other algorithmic improvement, it would absolutely be necessary to demonstrate mesh convergence. However, the goals of this study are to improve our qualitative understanding of how likely previously frozen carbon is to be exported or mineralized and it is clear the broad stroke conclusions will not be sensitive to details of mesh refinement. The mesh used is already highly resolved, especially in the region of interest. Outside the region of interest coarser mesh elements are adequate as those regions are only acting to establish correct flow boundary conditions on the highly refined region. Based on community experience with these types of models and our own experience with dozens of permafrost hillslope simulations, we are confident the mesh is adequately refined for the purpose of the study. To confirm this, we will prepare an alternative mesh with additional refinement and study the effects of mesh resolution on our results. We will add results of those simulations to the Appendix and add a sentence on mesh refinement and sensitivity to the manuscript accordingly.

It's also important to note that mesh size does not affect mass and energy conservation. ATS allows for user defined convergence criteria and error tolerance and enforces conservation of both mass and energy to those tolerances. The relative error tolerance in the present simulations is set to 1e-6 for all variables.

15. - l 134-137: I do not understand why a buffer zone is needed. This should be explained.

This will be part of the reworked mesh-section described in comment #9 that will be added to the Appendix. The buffer zone is used to avoid effects of the outlet boundary condition; please see response to comment #17 below.

16. - l 142-146: All of these sentences look like unjustified statements. Moreover what is exactly stated is not completely clear. For instance what means « preserving the subsurface volume representation of the catchment » and « a natural equilibrium without artificially imposed boundary conditions »? This paragraph should be either deleted or rewritten in a clearer and justified way.

This will be part of the reworked mesh-section described in comment #9 that will be added to the Appendix. The confusing paragraph will be removed and the text will be clarified.

17. - l 147-148: « The vertical sides of the model are assigned zero-flux boundaries for water » this is questionable, especially for the outlet vertical boundary. Are there any field observation that can be used to justify this choice? This should be discussed.

It is important to note that the outlet boundary is not closed to flow. As noted in the original manuscript, water expressing on the surface is removed. Thus, by closing the subsurface on the outlet, water flow is forced to the surface where it is removed. This can create an artifact in the immediate vicinity of the outlet, which is why we use a buffer region and do not analyze results in that region. By previous experience (Jan and Painter, 2021; Hamm and Frampton, 2021) and physical arguments we know this boundary condition produces physically correct behavior away from the immediate vicinity of the boundary. We will improve the description in the main text accordingly.

18. - l 151, « in line with borehole observation in Svalbard »: This is important, since it is likely the reason why a 40 m thickness as been chosen for the modeling domain. Please add a figure with the mentioned soil temperature profile evolution, as well as a discussion for explaining in which way these data where used for choosing not only the bottom thermal boundary conditions but also the domain thickness.

We propose to add a plot with trumpet curves showing that the seasonal surface temperature signal does not penetrate the ground deep enough to affect the bottom boundary condition. The maximum annual temperature variations at depth -15 m are 0.1°C, indicating this is a reasonable approximation of the depth of zero annual amplitude (DZAA). Also, see our reply to comment #23 on why we are confident that a 40 m deep domain is sufficient.

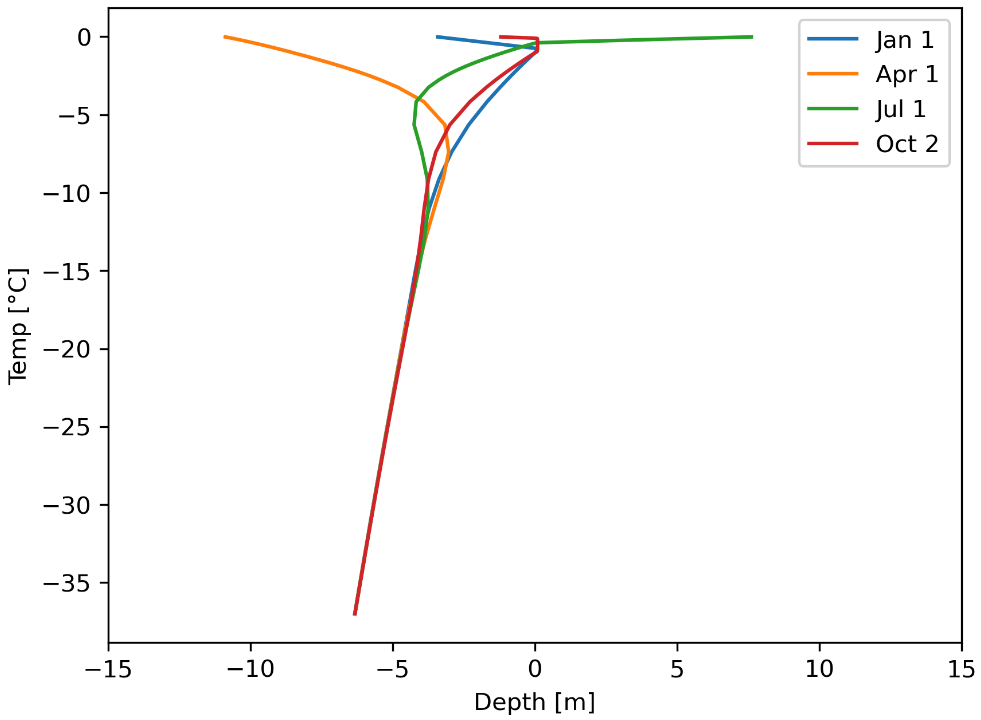

Fig. 1: Trumpet curve plot indicating that the 0 annual amplitude depth is located at ~15 meters depth.

19. - l 159-165: I have serious concerns about the chosen methodology for building ‘present day-like’ precipitation. Why not using a day-of-the-year average like for other forcings? This is not justified. Stating that « the resulting rainfall distribution resembles the variability of natural rainfall throughout the year» is not enough for making the arbitrary, artificial precipitation forcing data set relevant. The use of arbitrary precipitation data may impair the interpretability of the simulation results, so either the ‘resemblence’ between the artificial data set and the observation

should be quantitatively demonstrated, either the observation data set itself should be used as forcing data.

This approach was previously presented in Magnússon et al., 2022. For precipitation, a day-of-year-average is not representative of the physical system because it will always create very small values for each day in a predominantly dry place like the Adventdalen area, as opposed to many days with no precipitation and a few days with significant precipitation. These small precipitation values would lead to an unrealistically high fraction being evaporated due to the surface energy balance and the overall windy conditions in the valley. Additionally, there would be no days with absolute 0 precipitation, which is not realistic. As described in Magnússon et al. 2022, the daily rainfall amounts are not arbitrary, but represent the frequency-intensity distribution of the natural rainfall variability. Details can be found in the Supplementary Methods VI in Magnússon et al. 2022 and in the included repository with an annotated script that was used to create this precipitation time-series.

20. - l 165-166, « Soil physical properties are defined to resemble highly conductive material. »: This should be justified. Why not moderately conductive material?

We will add a justification to this statement. We assume a highly conductive material in the model because observations in the field site suggest that most of the subsurface material is very gravely or mossy. There are no field observations that describe the subsurface composition in detail.

21. - l 169-170, 'to establish a water table at target depth »: I do not understand. What is the 'target depth'? Are there observational evidences of a water table 'at target depth'?

The spinup process needs to establish an ice saturated subsurface below the active layer. A multistep procedure for doing this with ATS was developed more than a decade ago and is the standard spinup procedure for ATS simulations of continuous permafrost. The first step in that procedure is to establish a hydrostatic water column with the water table close to the surface. We then freeze that water in place with an open boundary condition on top to allow water to be pushed out by the liquid-to-ice volume expansion. That establishes an ice table as close to the surface as possible. Subsequent spinup steps with seasonally varying top boundary conditions then remove any excess ice/water and place the system in the desired ice-saturated state in cyclic equilibrium with the atmosphere. The final step enables the active layer to develop together with a seasonally perched water table at a depth consistent with present day conditions. We propose to replace the confusing sentence with the following text:

"*First, a single column model extending to the full depth of the final 2D mesh is used to establish an ice-saturated subsurface with an ice table near the surface. This was accomplished by freezing an initially hydrostatic and isothermal water column from below keeping an open top boundary to allow the volume expansion of the phase change to push out excess water.*"

22. - l 173-174: Why 10 years? Please provide the criterium used for this choice.

10 annual cycles for the second spinup step is a commonly used length for this initialization step in a 2D problem (e.g. Gao and Coon 2022, Jafarov et al. 2022) to reach an annual steady/cyclic state. In the plot below, it can be seen how the mean annual difference in surface temperature at the surface of the first mesh column in the area of interest between 2 adjacent years during the spinup decreases and reaches a minimal temperature variation after 5 years. We propose to include this plot in the Appendix and include the following sentence.

"We confirmed that the temperature is in a cyclic steady state at the end of the 10 year spinup period (see Appendix)."

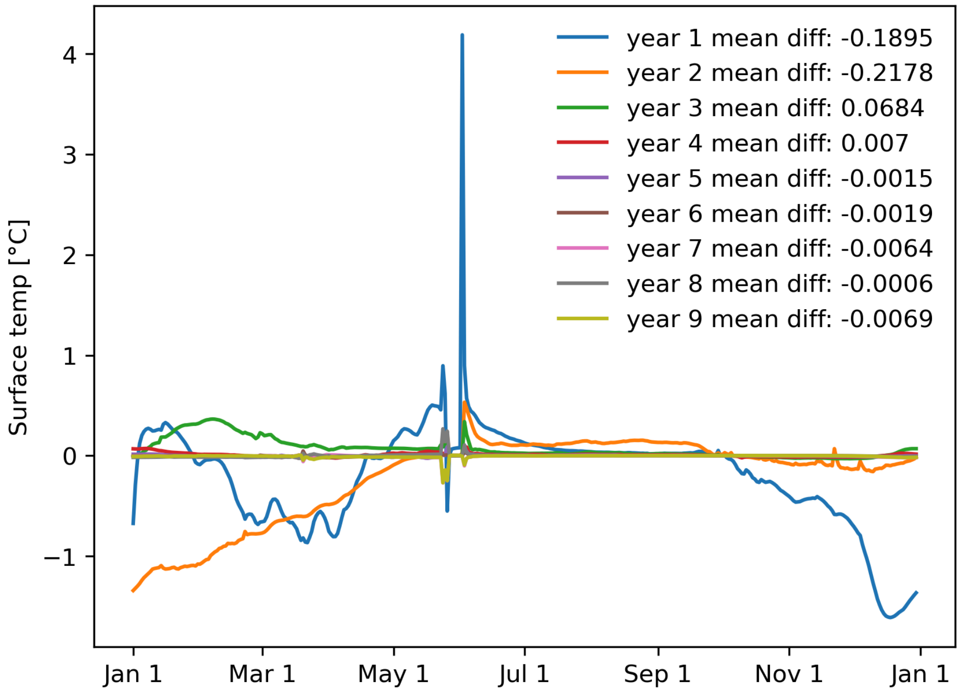

Fig 2: Plot of difference in surface temperature at the surface of the first mesh column in the area of interest. Mean differences are denoted as mean annual differences between the current year and the prior year. Despite some variability in the temperature difference, the mean difference in the last year is below 0.01°C with an absolute maximum daily difference of 0.2°C. The system is therefore considered to be in a cyclic steady state.

23. - l 176-178: The targets of this study are simulations under climate change, while the use of a thermal boundary conditions of constant temperature equal to present day temperature at the bottom of the domain (l 150-151) is not compatible with the simulation of climate change scenarios. Or at least, it implies the assumption that the temperature at 40 m depth is not impacted by surface temperature variations at the considered time scale of 50 years. This is a major concern for the validity of the produced simulation results. In order to demonstrate that this strong approximation does not impair the discussed results, the time variations of temperature in depth, close to the bottom boundary, must

be shown. If they are not negligible, then the simulations should be re-ran with an appropriate bottom boundary condition (e.g.: geothermal heat flux).

A geothermal heat flux boundary condition us problematic for these simulations because at 40 m depth the temperature profile is likely to contain a memory of past climates. Rather than deal with spinup in a reconstructed historical climate and all the uncertainties that this would create, it is better to use deep borehole temperature in regions where it is available.We used two simplified 1D column simulations with the same surface forcing as the full 2D transect, and with 1) constant temperature boundary condition and 2) geothermal heat flux boundary condition (which is set to 0) to show that the bottom boundary condition does not significantly affect the temperature in the active layer. We propose to include the following figure in the Appendix and add the following text to the body of the manuscript:

"*We confirmed with 1D simulations (see Appendix) that the treatment of the lower boundary has negligible effect on the near-surface region of interest up to year 40 and then only minor effects from year 40-50*."

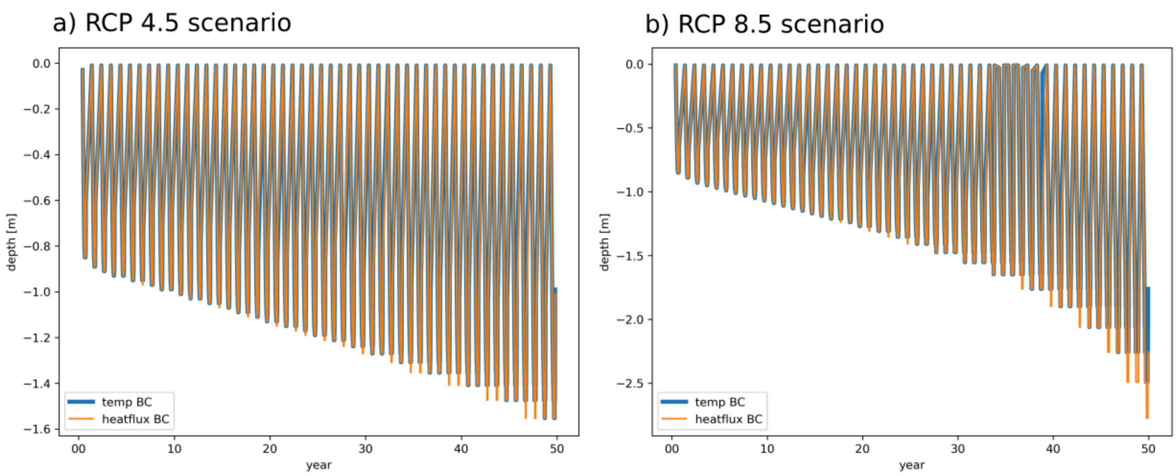

Fig. 3: Thaw depth progression in simplified 1D column models using the same forcing as the full 2D transect in the main text for the (a) RCP 4.5 and (b) RCP 8.5 scenario, with each a constant temperature boundary condition (blue) and a heat flux boundary condition (orange) prescribed at the bottom of the domain.The simulations show that the choice of bottom boundary condition has little effect on the processes in the active layer throughout years 1-40 and only small effects (never more than 1 mesh cell difference between the two versions) in the years 40-50. Hence, we are confident that the bottom boundary condition is adequate and the domain depth is deep enough.

24. - l 184-187: Additional explanations should be added here: why this rate and total amount of injection, why this moment?

We will add more information on the injection rate. The moment of injection is, as described, meant to represent a day at which the domain is at fully frozen conditions before the onset of thaw. The rate of injection is determined by the target length of injection (1 daily cycle) and the total amount of tracer injected is arbitrary since the transport is conservative. We set it to 100 mol simply to represent 100 units of solute.

25. - l 204-205, «air temperature (T avg), which is the only variable that changes over time in the respective scenarios. »: Why? Precipitation does also change along time in IPCC climate change scenario. This choice should be discussed.

Precipitation changes are difficult to predict and model predictions for high latitudes often diverge. Given that precipitation changes are not the focus of the study, we believe it is reasonable to adopt a stylized scenario that neglects precipitation changes to focus on temperature effects. We will add the following sentence to explain our motivation for this scenario:

"*This scenario is stylized and only aims to capture the effects of air temperature warming without considering complex, difficult-to-predict, and highly uncertain changes in precipitation*."

26. - Figure 4: Why the BTCs are multimodal? For instance, 2 modes for TOL carbon in Fig4.b and 2 modes for buried carbon in Fig4.c?

The release of tracer in the model is determined by the thawing rate of the active layer, which in turn is determined by air temperature changes. In our model, we use site-specific average air temperature and precipitation, which can vary from day to day. In a more idealized setting (e.g., annual temperature representation in the shape of a sine-curve or constant precipitation/infiltration rates) may cause a more homogenized image. In our case, daily variations are causing non-linear thaw rates and hence non-linear BTCs as tracer gets mobilized according to the thaw rates and groundwater flow through infiltration. We will extend the description of Figure 4 accordingly.

27. - l 271-272: Is this temporal partitioning between surface and subsurface transport in agreement with field observation? At least has the ponded water before mid-June has been observed in the field? Whenever it is possible numerical results should be discussed at the light of field observations.

We will add a reference to field observations when possible in the main text (for example for the subsurface material as well as the ponded water observations argued in this comment). The catchment is indeed characterized by inundated conditions in the beginning of the thaw season (mid-June) due to snowmelt and shallow thaw depths.

28. - l 279-281: In Figure 4.d the peak of the 13 August seems higher than the one of the 14 June, although the opposite is stated in the text. This bimodality should be better explained.

The bimodality is caused by two main modes of transport at two different times of the active layer development. The first peak in the subsurface at 20m can be explained by the initial surface overland flow which causes the spike in Fig 4b as it transports solute mass with it, which then infiltrates into the subsurface and produces a subsurface BTC peak. Later, tracer that has not been transported on the surface initially, will be transported in the subsurface by slower groundwater flow. However, this happens later in the season, leading to the second peak. We agree that the "less pronounced" peak is misleading and will delete it, while adding some more explanation about the bimodality to the description of the figure.

29. - Figure 8: Vertical peak on 30th of August, dual peaks in the 1st of September … I think these features are strange. May be due to convergence problems? The convergence study must be done for assessing it.

These features are not caused by convergence issues, rather by the fact that we are representing a realistic topography instead of an idealized flat transect. Changes in topography combined with the undulating permafrost table can lead to vertical and sub-vertical groundwater flow and cause the observed vertical increase in solute mass. Further, we mask values smaller than 0.05 mol for better visualization, which does not imply that there is no tracer in the remaining cells, but the quantity is below the defined threshold. This has been described briefly in the text (Line 315). We will extend this explanation and adjust the figure legend to reflect the lower solute mass limit.

30. - l 323-324: « A substantial fraction of the initially injected tracer mass (~ 40%) moves vertically (both upwards and downwards) within the same mesh column in which it was injected (see Fig. 9a and d and Fig. 10). » What phenomena are responsible for this vertical redistribution ? Diffusion, freeze/thaw cycles related effects? This should be explicated.

This vertical movement is a combination of drying/wetting and cryosuction. This is explained using Figure 10. The upwards movement (cryosuction) is so far only discussed in the discussion section, but we will also add clarifications on this to the results section.

31. - Figure 10.h: Numerical instability ? Should be corrected, or explained if it is not an artifact.

The blue cell downslope of the injection point in Fig. 10h is not an artifact nor numerical instability, but rather indicates that some tracer has moved downslope already. It only appears in this specific time instance plot due to the high temporal (daily) variability in output. In the original figure, we masked values smaller than 1 to improve visibility. Since this is not essential to the interpretation of the results and may confuse the reader, we will increase the threshold to 1.5, which will further improve clarity by removing the distracting downslope tracer movement. We will include the information about the masking in the figure caption.

32. - l 346 : « This observation highlights the importance of mesh resolution in lateral transport simulations. ». I fully agree. A convergence study must be done.

Convergence issues are not causing this observation. The mesh resolution in the present-day climate simulations is highest (1 cm per cell) in the contemporary active layer and decreases slowly with depth. However, energy- and mass conservation apply to these cells in the same way that it does in the contemporary active layer cells. The mesh resolution is therefore adequate.

33. - l 352, « vertical mobilization »: once again the involved mechanism must be explicated.

This will now be mentioned earlier on (see comment #30) and is also part of the discussion.

34. - l 407-411 : « We partly address this by representing a converging slope model setup, where the cell width in transverse direction varies depending on the distance in longitudinal direction. This way, the surface area of the catchment is preserved, and it is possible to accurately represent water and energy balances as well as infiltration and evaporation rates throughout the catchment. This approach has previously been applied by Gao and Coon (2022). » I do not think that using a one cell-thick discretization in the transversal direction may allow to simulate the effect of the watershed geometry, either convergent or divergent (using the terminology of Gao and Coon 2022). Nor in the present manuscript or in Gao and Coon 2022 are presented arguments for supporting the validity/usefulness of such a 'pseudo-3D' meshing methodology. A proper comparative study should be done for this, between results obtained with« pseudo-3D » meshes and with full 3D meshes. Of course with only one cell no lateral fluxes may be computed.The only interest I would see would be to weight the inward fluxes through the top cell face according to the area of the cell, but then why not simply apply a spatial weighting on the incoming fluxes prescribed by the boundary conditions? Including this in the meshing seems to me inappropriate and confusing. Anyway in this case the methodology used for computing the cells widths must be explicated.

See our reply to comment #9 and the proposed changes to the text. We aligned the mesh with the topographic gradient and thus the flow; by construction there is no tangential lateral flow, only in the vertical and along the topographic gradient. Flow-aligned meshes are a standard approach for reducing the computational complexity of hillslope simulations, as noted in our reply to comment #9.

35. - l 465, « Under the simulated environmental and soil physical conditions in this study, »: This precision hold true not only for this point, but for all the listed conclusions. The writing of the manuscript should better reflects the fact that this is a numerical study of transport in a specific site, with no possibility of automatic generalization for permafrost regions as a whole.

This very important point will be discussed more thoroughly in the discussion. While the results of this study may apply for foothill systems, they may unfold very differently for permafrost locations in e.g., plains with micro-topography.

TECHNICAL CORRECTIONS :

36. - l 25: Missing) at the end of the line.

Noted

37. - l 29, l30, and elsewhere: I would recommend to systematically use ‘organic carbon’ instead of just ‘carbon’ for naming the C part of the organic matter stored in permafrost.

We prefer to keep using the word “carbon” as it is mentioned nearly 180 times in the current manuscript. We will clarify in the introduction that we refer to organic carbon throughout the manuscript.

38. - l 32: Missing s at ‘question’.

Noted

39. - l71-84: Part of this paragraph should be in the Methods section (e.g.: choice of distinguishing four carbon pools and using labelled tracer for identifying them).

This final paragraph in the introduction section summarizes our approach to address the questions that arise from the literature summary in the rest of the introduction and hence is essential to the formulation of the hypothesis. A thorough explanation of the carbon pools and labels follows in the methods section.

40. - l 151, «bottom horizontal boundary »: According to Figure 3, the bottom boundary is not horizontal.

We will remove “horizontal”

41. - l 251-260: These information should be included in the Methods section, along with a figure for quantitative locations of the injection points and of the measurement points within the modelling domain.

We agree that a zoomed-in figure with the injection points would be beneficial for the interpretation of the results and will prepare such an illustration. We prefer to keep the introductory sentences in the results section as they currently are to preserve the flow of reading and introduce the way in which the results are presented.

42. - l 258, « explicit », « continuous in space » : Odd vocabulary. A BTC represents the temporal evolution (not exactly continuous, since there is a time discretization) of concentration at a given location in space, while a plume vizualisation represents the spatial distribution of concentration at a given moment.

We will clarify this sentence.

43. - Legend of table 2: The use of different concentration thresholds for TOL and buried carbon should be mentioned in the body of the text, in the Methods section.

We will add this to section 2.5 “Model forcing” under “Present-day weather conditions”.

44. - l 273: « when all runoff is occurring in the subsurface »: Then it is not run off, but groundwater flow.

We prefer to keep our original phrasing as it encompasses both percolation and groundwater flow, as well as highlighting the surface-subsurface nature of runoff and flow interactions. Subsurface runoff is a defined hydrological term (e.g., Pilgrim et al. 1978)

45. - l 285: Do not refer specifically to a Figure in Supplementary material (Fig. S3) in the body of the text. Instead, refer to the supplementary material as a whole.

Supplementary information will be moved to the Appendix

46. - l 286 – 287 : « However, given the specific solute transport patterns in this model, » What is meant here ? Unclear.

We will reword this.

47. - l 299, « vertically transported upward »: Oddly said. Strictly speaking, the topographical effect mentioned here does not generate ascendant flow.

In this part of the domain, the local topography combined with the undulating permafrost table, which acts as an impermeable boundary for flow, causes the active layer to be locally saturated with liquid water at the time of tracer arrival, which in turn enables the observed vertical/sub-vertical tracer movement. The manuscript text will be revised as follows:

“*A small fraction of buried carbon tracer (tracer mass < 0.005 mol) experiences surface transport (also visible in Fig. 8) because it gets transported upwards by groundwater upwelling. This is caused by a combination of the terrain unevenness and the undulating impermeable permafrost table, causing local downslope water accumulation and a fully saturated active layer.*”

48. - l 302: Do not refer specifically to a Figure in Supplementary material (Fig. S4) in the body of the text.

Supplementary information will be moved to the Appendix

49. - Caption of Fig. 6, « Note that the tracer mass is restricted to the uppermost subsurface cell in this snapshot and is difficult to visualize in this illustration » : True, then this figure has to be improved. May be that plotting the two variables

separately would be an option? Besides, Fig. 6 is mentioned only very briefly once in the text. Either it should be deleted or more extensively commented.

We will add an inset to the figure, zooming into the area where there is tracer and will also change the color scheme so that it becomes more visible (see a first draft of these changes in figure below).

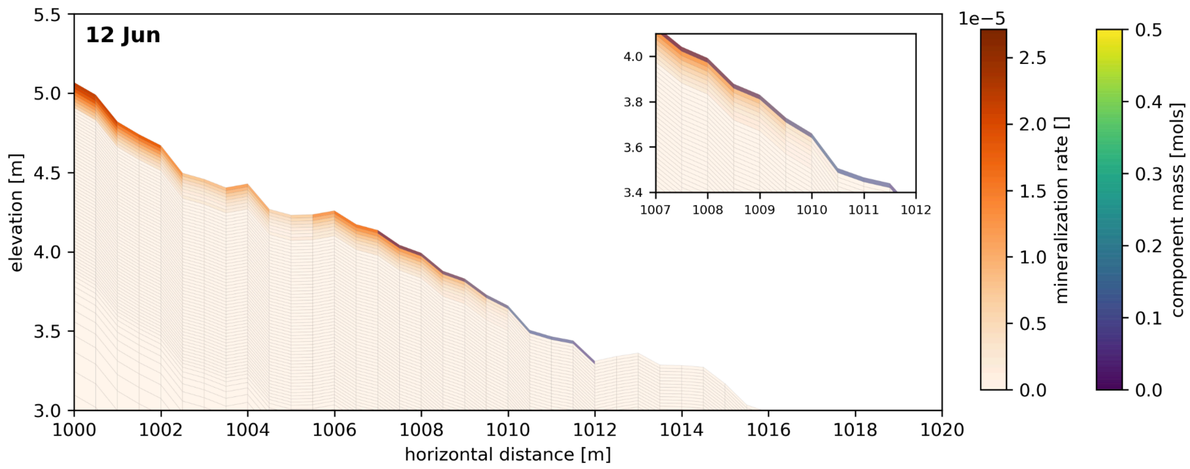

Fig 4.: New proposed visualization of Figure 6 in the main text

50. - l 315, « some tracer is moved upwards »: see comment on l 299.

Noted

51. - l 317: 12 % does not obviously look negligible.

Will be changed to “small”

52. - Figure 7: I can’t see anything regarding the solute mass. This Figure should be thoroughly reworked so that it becomes informative. Besides, what is the « ponded depth » ?

We will add the information about the ponded depth and also the thaw depth into the figure caption. Apologies for the oversight. However, we argue that the tracer mass is as visible as possible while still showing the entire transect of interest. For better visualization, we will decrease the opaqueness of the mesh and the thaw depth line.

53. - Figure 9: The annual cycles/peaks should be discussed. « The visual increase in mass above 100 mol in the injection columns in (a) and (d) is not a physical phenomenon but a result of aggregating and rounding errors during post-processing of the model output. » Then the post-processing should be improved.

Please see response to comment #54 below.

54. - l 350-355: If the post-processing method does not allow to conclude, then it should be improved.

The issue here is that a limited precision is used in output files causing very slight round-off errors. A limited precision is used due the large quantity of output data. To improve this, an increased precision in the output files could be implemented, but then the simulations would need to be re-run, which is not practical as the runtime is over 3 months wall clock time. As this is a minor issue which has no bearing on the results of the study we prefer to keep the data processing as is, but propose to improve the figures (Fig 9 and Fig 11) by truncating to the 100 molar upper limit where applicable. This correction should help clarify the presentation of results without distracting from cumbersome explanations of data post-processing. The explanation of data post-processing will be provided in the Appendix.

55. - Caption of Figure 11: « The visual increase in mass above 100 mol in the injection column in (b) is not a physical phenomenon but a result of aggregating and rounding errors during post-processing of the model output. » Once again the post-processing method should be improved.

Please see response to comment #54 above.

56. - l 414-420: this should be part of the introduction, not the discussion.

This part places the work in the context of what has been done on permafrost carbon transport modeling previously and how it can be improved in the future. We therefore think it is important to mention this in the discussion to highlight future development of this work and prefer to keep it as is.

57. - l 420-425: this should be part of the Methods section, not of the Discussion.

This section is discussing limitations of using a non-reactive representation of the tracer and we feel it is therefore important to elaborate on in the discussion section. We wish to refrain from discussing alternatives in the methods section.

58. - l 432-447: This should be part of a perspective section, not of the Discussion.

We feel it is useful and customary to include perspectives such as this in the discussion section. A dedicated section for this part would likely hinder the flow of the discussion. We therefore prefer to keep this part in the discussion.

59. - Section 4 Discussion: Given the parts that should not be included in the discussion (see the three comments above), the Discussion section is rather short (a bit more than one page), and lack of in depth analyses of the produced results. I recommend to strengthen this section, including for instance elements for linking the simulated behavior and the site specificity.

We will extend the discussion by a discussion about the site specificity, its role in a wider context in the Arctic, and add a discussion point on wetness conditions. We believe this will strengthen the discussion.

60. - l 452, « mechanical transport »: Sounds weird. Passive transport would be more relevant.

Will remove “mechanical”

References

Gao, B. and Coon, E.T., 2022. Evaluating simplifications of subsurface process representations for field-scale permafrost hydrology models. *The Cryosphere*, *16*(10), pp.4141-4162.

Hamm, A. and Frampton, A., 2021. Impact of lateral groundwater flow on hydrothermal conditions of the active layer in a high-Arctic hillslope setting. *The Cryosphere*, *15*(10), pp.4853-4871.

Jan, A. and Painter, S.L., 2020. Permafrost thermal conditions are sensitive to shifts in snow timing. *Environmental Research Letters*, *15*(8), p.084026.

Magnússon, R.Í., Hamm, A., Karsanaev, S.V., Limpens, J., Kleijn, D., Frampton, A., Maximov, T.C. and Heijmans, M.M., 2022. Extremely wet summer events enhance permafrost thaw for multiple years in Siberian tundra. *Nature Communications*, *13*(1), p.1556.

Pilgrim, D.H., Huff, D.D. and Steele, T.D., 1978. A field evaluation of subsurface and surface runoff: II. Runoff processes. *Journal of Hydrology*, *38*(3-4), pp.319-341.