Letter to Editor

Dear Elin and co-workers,

Thank you again for the strong improvement of your manuscript. I have received and studied the reports of two referees. Moreover, I quickly went through your manuscript myself. Referee 1 raised a series of concerns to sharpen the structure and to strengthen your conclusions with respect to insights about the system – not the model. Referee 2 is more positive about the reached state. He/she points to the lack of structural detail in the model setup description.

I strongly agree with both referees and would like to ask you for another round of major revisions along the suggested lines. In my view, your study has a strong data-related focus, which is somehow difficult to trace and obscured when referring to your MikeSHE model and the added corrections. Moreover, I find it particularly difficult to draw the lines to catchment functioning beyond small vs. large or GW dominated vs. OF dominated. I have full confidence that you can and will turn your manuscript into a nice paper when adhering to and working clearly towards the key messages about landscape functioning.

I hope you find the reviews (incl. mine) helpful for this. If you have further questions, please contact me.

All the best. Many cheers.

Conrad

Reply: Dear Editor, we are grateful for your thorough work, both reading and discussing the manuscript, as well as your constructive criticism. It is obvious that you gave our manuscript a lot of time and effort, which we much appreciate, and we hope that our answers and changes are adequate and satisfactory. We received commented on the structure, especially regarding the method section about the Mike SHE model as well as the discussion section, both from you and reviewers. We hope that the updated version of the manuscript will remedy these concerns. The major changes in the method section include an extended description of the Mike SHE model, and an edited version of the section regarding the stream chemistry and seasonality. Our answer is built up as follows: a list of major changes, a reply to the editor and thereafter, a reply to reviewer #1 and #2 respectively. All questions and comments are stated unbridged, with our replies below each question in brown. Blue text was taken from the old version of the manuscript and red text was taken from the new version of the manuscript. Please note that the abbreviation for geometric mean of the travel time distribution is now shortened to MTT_{geo}.

List of major changes

The major changes to this version of the manuscript includes:

- A more in-depth description of the Mike SHE flow model in the method section
- A new figure describing Mike SHE
- An improved structure for the method section and the discussion section
- Updated figures and tables from input by Editor and reviewers
- Table 5 have been simplified and additional information was moved to appendix
- New abbreviation geometric mean of the travel time distribution (MTT_{geo})

Reply to Editors review

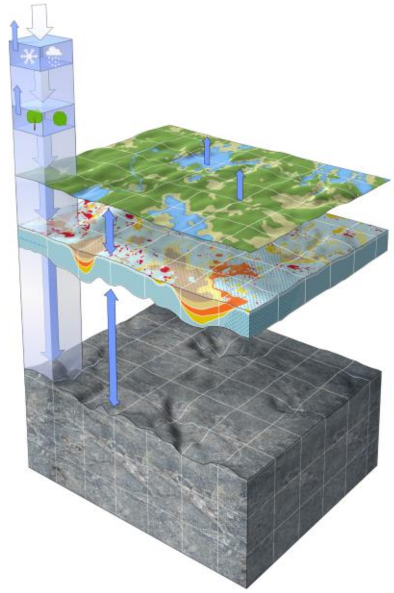
Here is my review:

1. L115-124: After studying your methods and consulting your 2018 paper, I am not quite sure how you used MikeSHE and if your application can really hold as physical 3D GW, 1D vadose zone and surface water model. If I am not mistaken, MikeSHE allows to use different approaches to calculate the hydrological fluxes through the different model units – which are not necessarily "physically" defined.

Reply: Reviewer #1 and #2 also commented on this part of the method section, and we agree that a more in-depth description of Mike SHE could be useful. We, therefore, extended the flow model section according to the text example below (there is an extended version of this text in the manuscript). We also added a figure explaining the Mike SHE model. We hope that these additions will give sufficient information regarding the Mike SHE model setup.

Part of new method section added to the manuscript:

We applied the Mike SHE/Mike-11 hydrological modelling tools to quantify travel times in a pre-calculated 3D transient flow field. The simulated terrestrial hydrological system for the Krycklan catchment includes: the saturated and unsaturated flow, ET, snowmelt, overland flow, and streamflow processes. The fully distributed 3D modelling tool use topography, soil properties, and time-varying climate inputs to calculate the water fluxes throughout a catchment (Rahim et al., 2012; Sishodia et al., 2017; Wang et al., 2012; Wijesekara et al., 2014). The ET processes include canopy interception, open surface evaporation, root uptake, sublimation, and soil evaporation from the unsaturated zone based on a methodology developed by Kristensen and Jensen (1975). Flow in the saturated zone (SZ) is calculated in 3D by the Darcy equation. The flow in the unsaturated zone (UZ) is calculated in vertical 1D using the Richards equation, and overland flow (OL) is calculated using a horizontal 2D diffusive wave approximation in the Saint-Venant equations (Fig. 3)….



(New) Figure 3: Schematic Mike SHE model set up. Precipitation falls on the ground as rain or snow. Evapotranspiration (ET) processes include canopy interception, open surface evaporation, root uptake, and soil evaporation from the unsaturated zone (UZ). The overland flow (OL), saturated zone (SZ), and UZ interact depending on the saturation level. The SZ is divided into ten calculation layers (CL), while the UZ has a much finer description. Streamflow is modelled through Mike 11 and is not restricted to the Mike SHE resolution. The figure is used on the courtesy of SKB. Figure illustrator: LAJ.

2. Moreover you superimpose mixing assumptions (eq. 3 and 4), which tries to compensate for some model limitations. L123f is full of further claims, which I find difficult to recover in your manuscript. I suggest to really stick to the approach using data from 14 sub-basins and calculating seasonal gMTT from this data and the model and evaluation of this gMTT findings.

L123: The complexity of the model allows for an in-depth investigation of advective travel times by non-reactive particle tracking simulations in a transient flow field.

Reply: We hope that the extended version of the Mike-SHE model setup, is sufficient to allow this claim. Regarding the question of the specific equation see answer below:

Equation 1-4: We agree with later comments by the editor regarding the necessity to explicit write out these equations (equation 1 to 4). To simplify the manuscript the text has been modified so that these equations instead are written in text format in the method section.

Equation 3: Adjusted cation concentration = Observed cation concentration/ (1-fraction of mire coverage)

While we agree that Eq. 4 is an attempt to compensate for some limitations of the particle tracking in Mike SHE, we must emphasise that Eq. 3 is not. The purpose of using base cations and oxygen isotopes in the manuscript was to enable a comparison of the modelled transit times to something observable in order to strengthen the credibility of the performance of the model. However, it requires a fundamental understanding of how they work and, in many cases, also some adjustments in order to allow a fair and meaningful comparison. The underlying assumption concerning the base cations, as illustrated in Fig. 2, is that the kinetically controlled weathering of base cations from the mineral soil could act as a proxy for the contact time between the minerals and the groundwater, i.e., the groundwater transit time. The assumption is, however, only valid when the groundwater is in contact with mineral soils, which complicates things because mires are abundant in some of the investigated sub-catchments. Because no minerals are present in the peat the base cation concentration cannot be expected to increase during the time the groundwater spends there. Hence, a direct comparison between base cations and groundwater age in a mire-dominated would be misleading so this needs to be addressed somehow to allow a fair comparison.

The fact that mire-dominated sub-catchments have lower concentrations of base cations (and other weathering-derived elements) as a result of no weathering in the peat is well demonstrated and we refer particularly to Lidman et al. (2014), which is cited in the manuscript, for further details and discussions on this matter. All data presented in the manuscript come from the same catchments as were investigated in this study, so we believe that there is strong underlying evidence for Eq. 3. Again, this is not to correct or compensate the modelling results, but to adjust the observed cation concentrations in certain sub-catchments so that the more accurately reflect the transit time, thereby allowing some sort of validation of the model.

Furthermore, the effects of this adjustment should not be exaggerated, and this may be helpful to illustrate (Figure Y below). Because most investigated sub-catchments only contain a minor fraction of mires the correction will only have a major impact on a few sub-catchments, e.g. C4 and C5. Hence, the correlations and relationships we present would be valid without this adjustment, but they would be a bit weaker and, in our opinion at least, somewhat unfair because the underlying conceptual model would not make full sense, since it would assume weathering of base cations in peat.

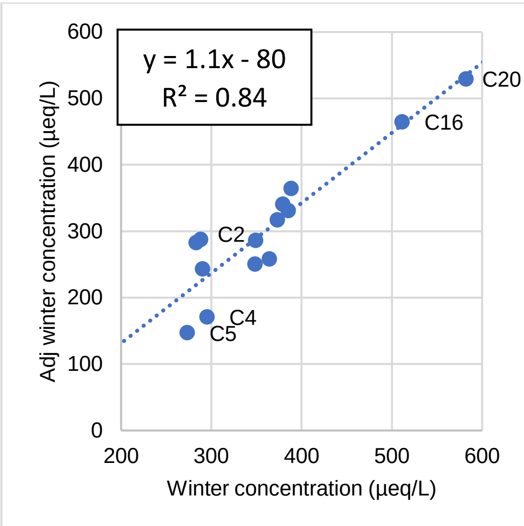


Figure Y: Winter Stream concentration before (x-axis) and after implementing equation 3 (y-axis)

Equation 4: MTT_{geo}=groundwater MTT_{geo}*(1- fraction overland flow).

There is a limitation with particle tracking in Mike-SHE, since it only functions in the saturated zone. This is due to that the particle tracking is based on saturated flow using the advection-dispersion ekvation based on the Darcy velocity. This correction can however be seen at the soil contact time, which is needed to compare the MTT_{geo} to stream BC concentrations. The stream BC concentration is affected by the water-soil contact time and is diluted by water taking the overland route to stream. To account for the OL water, especially in spring, the MTT_{geo} was reduced using the OL fraction as scaling factor.

However, for most catchments, almost all water comes to the streams as groundwater flow (for example, the full-scale catchment receives approximately 10% water as overland flow). The overland flow is also mainly linked to the spring snowmelt and to sub-catchments with mires. It has no impact on the winter gMTT (water only from groundwater) and very little impact on summer and annual MTT_{geo}. Please see the Fiq. X below, which compares the annaul adjusted MTT_{geo} and unadjusted MTT_{geo}. Although we mainly discuss the adjusted values, we will state the unadjusted values in Appendix.

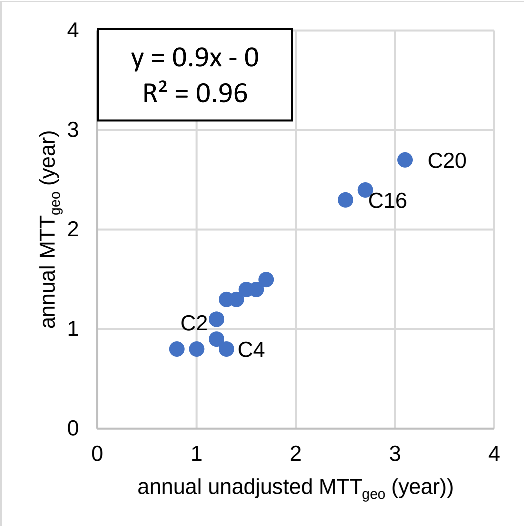


Figure X: Annual unadjusted MTT_{geo} (x-axis) and after implementing equation 4 → annual MTT_{geo} (y-axis)

3. Sec. 2.2: Despite being in favour for the reached brevity, the essential points are difficult to discern. Even after consulting your 2018 paper, I cannot really see, if the numerical discretisation and setup is as presented in Bosson et al. (2008) table 3-1 with a more fine vertical discretisation near the surface to enable a calculation of unsaturated flow (see Vogel and Ippisch, 2008, 10.2136/vzj2006.0182). L174 is rather confusing - especially when consulting Tab 2 with the greatest depth of surface soils to 7 m (out of 100 m) and the rare occasions of GW tables deeper than 10 m (Fig 3b).

Reply: The unsaturated zone and saturated zone have different discretization, with the unsaturated zone having a discretization from centimeters to meters while the saturated zone has a discretization from a meter to a few meters. Since the particle tracking occurs in the saturated zone, we believed that the most important information was the properties of the saturated zone. However, for clarity purposes, we added more information regarding the complete Mike SHE model according to the answer to question 1.

4. Also the actually used "physical" concepts remain unclear to me. In L180 you refer to depth-dependent drainage functions with given hydraulic conductivity values (Tab 2), which I could not fully understand from a quick scan of your 2018 paper and Bosson et al. (2008). I am under the impression, that this is some sort of Darcy flux with unit gradient.? Is this correct? How is infiltration calculated? Please do not get me wrong: I have no problem with your model selection and the used approaches. Moreover I like the brevity. However, I would suggest to clarify which "numerical engine" (to use the term from Graham and Butts (2005)), which concepts and which discretisation is used. Maybe tab 2 could also come as a figure for easier understanding and to get the model out of the way for the more interesting insights?

Reply: The reason behind the exclusion of a more descriptive part regarding the flow model was that we did not want the flow model to be the focus point for this study. However, we understand from the reviewers and the editor that a more explicit description of the Mike SHE model is needed. We therefore extended the part about the flow model in the method section. As the editor suggested, we also added a figure explaining the Mike-SHE model (see answer to question 1).

5. I got confused by the brief announcement of the particle tracking approach in Sec 2.2 (L190ff.) but the actual subsection 2.4 with the more detailed description. Also giving tab 3 without explanation how porosity affects the particle tracking added to my confusion. I suggest to tackle all particle tracking in sec. 2.4.

Reply: We agree that this part of section 2.2 should be moved to section 2.4 for better clarity. We also clarified in the text which equation is used for the particle tracking:

The advection-dispersion equation governs the transportation of particles for a porous medium. The Darcy velocity is divided by the porosity to calculate the groundwater velocity. Therefore, the only complementary input data needed to run the particle was porosity values (Table 4).

6. Fig 3 is insightful, however I find it difficult to understand the low number of particles. If you get about 20 particles per year and most particles are released rather quickly, can your really robustly sample the tails? I would also welcome a word about the idea behind only "releasing" particles in the saturated zone (and not simply adding markers to all incoming water).

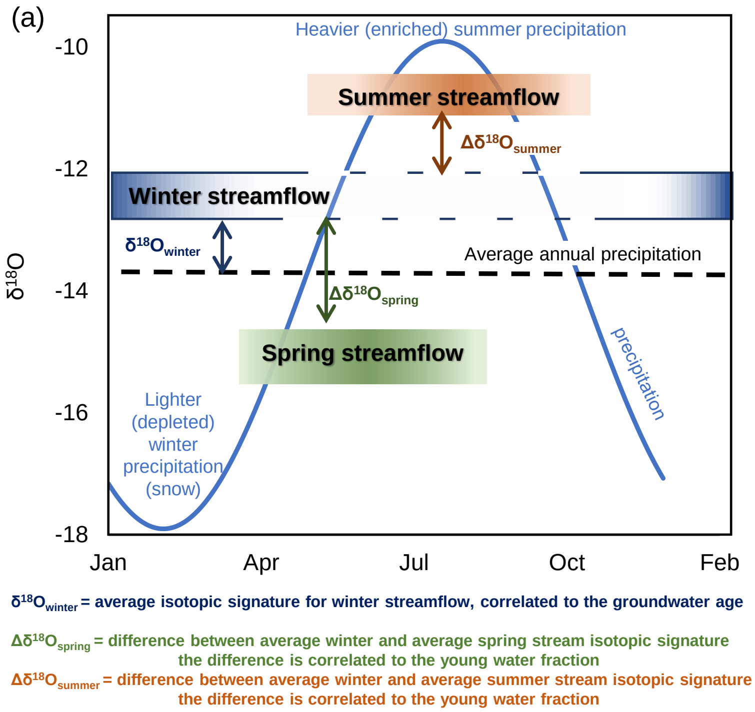
Reply: The word "release" might be a somewhat misleading. Particles are added through the groundwater recharge. However, the number of particles released must be limited to run the model. With 0.5 particles per 10 mm recharge per cell, we are still able to capture the main recharge inputs, such as the spring snow melt (approximately 200 mm → 10 particles) and the larger precipitation events during the summer and autumn (150 mm → 8 particles). Note, that 20 refers to the number of particles per cell. Due to the model size of the catchment, there is in total more than ½ million particles injected in the model over a year, which is very computational heavy.

7. Sec. 2.3: I am slightly confused by the title and the following description of the system's seasonality. Maybe I went through the manuscript too quickly. I suggest to include a statement about the references (δ18O and BC) and the respective concepts (Fig 2) incl. the site seasonality before going into detail. L125-135 does not really serve this.

Reply: We changed the title to: "2.2 Linking seasonal base cation concentration and isotopic signature to stream water age". We also agree that a better statement about the seasonality could be useful in this section. Therefore, we changed the first paragraph to:

This study was focused on three seasons in Krycklan, winter, spring, and summer. We defined the winter to occur from late November to late March because it is characterized by an extensive and permanent snow cover with little groundwater recharge. We assumed that the winter stream composition reflects the chemistry of deeper groundwater (Fig. 2). Similarly, we defined spring by the rapid transition in hydrology and biogeochemistry in April-May. During snowmelt, ca 50 % of the annual precipitation leaves the system in a short period of time, diluting baseflow with new input of water. Finally, we defined the summer season as the period between July and September when the hydrology is characterised by rain, high ET and relatively little runoff. June and October were excluded because, hydrologically, they are transition months between the three distinct seasons. This is because snowmelt can still influence runoff in June, and winter conditions (snowfall, soil frost, etc.) can sometimes begin to establish in October.

Figure 2: We received some additional suggestions from reviewer #1, which we also added. The main changes are, ∝ have been removed, MTT_{geo} was removed and changed to mean soil contact time, and the line for precipitation was removed for clarity. We also added a description of $\delta^{18}O$ and BC directly in the figure, instead of mainly having it in the figure caption.



8. With respect to the employed "correction terms" (eq. 1,3,4) I cannot really see how much this influences your results. At the moment it simply resides in the realm of post processing of the MikeSHE model output and the data. However, it comes with a series of assumptions.

Reply:

Equation 1: This method is standard when working with water isotopes (unless you specifically wish to trace the lake water). Incomplete evaporation from a water surface, e.g. a lake, will alter the isotopic composition of the water so this needs to be accounted for when looking at the isotopic signal from lake-influenced catchments (Peralta-Tapia et al., 2015).

Peralta-Tapia, A., Sponseller, R. A., Ågren, A., Tetzlaff, D., Soulsby, C., and Laudon, H.: Scale-dependent groundwater contributions influence patterns of winter baseflow stream chemistry in boreal catchments. Journal of Geophysical Research: Biogeosciences, 120(5), 847–858. https://doi.org/10.1002/2014JG002878, 2015.

Equation 3 and 4 – Please see answer to question 2

9. For linear regressions (like eq 1) it might be sufficient to name the scaling factor and intercept and focus on what is behind the regression? (Is there seasonality in the signal… Shouldn't there be a weather dependency of this relationship (depending on actual evaporation)?)

Reply: The use of water isotopes in catchment hydrology is based on the seasonal fractionation, which is shown in Fig. 2. The signal is indeed strongly affected by the weather and in particular the temperature at which the phase transitions occur, e.g. water vapor to rain and rain to snow. This causes a distinct seasonal variation in the input signal to the system.

Further fractionation may occur also in the catchments when there are phase transitions. This is the case, for instance, when there is evaporation of inception on tree, but since this typically ends in complete evaporation of the water drops it does not alter the input signal to the system. However, if there is incomplete evaporation, i.e. parts of the evaporating water remain in the systems, this can cause a substantial alternation of the isotopic signature of the water. This is a problem mainly in lakes, where there is continuous evaporation of water from the lake surface, particularly during the summers. This is enough to significantly change the isotopic signature of the lake water. In all applications of water isotopes in catchment hydrology it is therefore customary to adjust the isotopic signature for the effects of evaporation from the lake (unless the purpose is to trace lake water), (Peralta-Tapia et al., 2015).

Peralta-Tapia, A., Sponseller, R. A., Ågren, A., Tetzlaff, D., Soulsby, C., and Laudon, H.: Scale-dependent groundwater contributions influence patterns of winter baseflow stream chemistry in boreal catchments. Journal of Geophysical Research: Biogeosciences, 120(5), 847–858. https://doi.org/10.1002/2014JG002878, 2015.

10. Eq 3 in my eyes says that the flux in a cell is averaged over all surface fractions but that the concentrations of non-mire surface fractions are diluted by the same flux of non-loaded water from mires. I see its practical meaning but I am not quite sure about its theoretical underpinning and if this has effect on the results.

Reply: The conceptual idea behind Eq. 3 is that weathering occurs in mineral soils but not in peat. Hence, base cations will be released to the groundwater, causing higher concentrations, as long as the groundwater remains in the mineral soils. Any period that the groundwater may spend in a mire would, however, not lead to higher concentrations because there are no minerals present that can release base cations. Hence, if a sub-catchment has large portion of mires where the groundwater may spend substantial amounts of time, the transit time of the groundwater in the system cannot be expected to be reflected in the base cation concentration. Hence, this undermines the assumption we need to make in order to test the modelled transit times against real data. Therefore, some sort of correction is needed.

While this is a conceptual idea, it was developed based on actual data from the very same sub-catchments as in this study. This has been presented in detail in above all Lidman et al. (2014), which is cited in the manuscript. This paper includes a wide range of weathering-derived elements, which all showed the same pattern of lower concentrations from mire-dominated catchments. Depending on how strongly different elements bind to organic matter there may also be a substantial accumulation in the peat, but for the base cations this effect was minor. For example, if a mire-dominated catchment (like C4 in the manuscript) contains ca. 40% mires, the Na concentration in that stream would be 40% lower than in a 100% forested stream. This was statistically valid across a wide range of elements. That is the empirical underpinning of Eq. 3. Hence, we believe that it is firmly justified.

The effects of the correction on the results are shown in Fig. Y. As can be seen, the overall pattern is still the same because most sub-catchments do not contain so much mires, but above all the mire-dominated streams C4 and C5 deviate and display conspicuously low base cation concentrations given the modelled transit times. One possible interpretation of this figure would be that the model fails to accurately calculate the transit times of the water in the mires, but that would then rely on the erroneous assumption that there is widespread weathering of base cations in mires. Hence, it would be a misrepresentation of the results, and therefore we feel that this correction is relevant and justified.

11. Eq 4 is a similar scaling approach. This appears to have strong implications. After thinking about it I am not really sure what the propose of MikeSHE remains if groundwater flow is indeed calculated as "drainage function" based on conductivity and when overland flow is actually imposed as surface fraction. I assume that your

<u>Reply</u>:

The drain constant: The uppermost saturated calculation layer's depth, 2.5 m, was calibrated by studying the calculation layer thickness's influence on the groundwater table's location and dynamics. Above this depth, the ET processes were shown to have an impact on the groundwater table dynamics. If the groundwater table falls below the uppermost calculation layer, a more simplistic method, not taking capillary rise and all ET-processes into account, is applied. The top calculation layer's thickness results in all soils shallower than 2.5 meters being averaged into one soil type. Kh is averaged using the thickness of each soil layer. Vertical flows are more dependent on the lowest Kv. Therefore, the harmonic weighted mean value is used to calculate the new Kv instead. In MIKE SHE, a drain function can be applied to account for small ditches and drains in the landscape not captured by the model grid or the 1D river flow network. It has also been used to account for higher hydraulic conductivity in the uppermost part of the first calculation layer. The function is activated whenever the groundwater reaches 0.5 m below the ground surface, above which higher K-values have been observed.

Equation 4: As discussed above, Eq. 4 is an attempt to look at the implications of the limitations imposed by the particle tracking module in Mike SHE. Since the particle tracking is limited to the groundwater, the modelled transit times will of course be representative for the groundwater, but in cases where there is considerable overland flow, the age of the groundwater will not be the same as the age of the stream water. This is important above all during the spring flood and, in particular mire-dominated systems. We see this as a way to further test the performance of the model and to address the impact of overland flow on the overall transit time of the stream water.

We also wish to emphasize that Eq. 4 is not used to substitute the modelled groundwater transit times – these are still presented in the manuscript. Instead, it offers an additional metric, which deepens the understanding of the hydrology of these systems. Otherwise, the limitations of the particle tracking would only leave us with groundwater age, but for many applications a quantification of the stream water age is also valuable. As shown in Fig. X above (question 2), the difference between the two is quite small.

12. Eq. 1-4 in general: I am not really convinced by the chosen notation with a mix of variables and names and everything. First of all, the equations are rather simple. Hence I could imagine that they are actually not needed. The calculation and averaging of Δδ18O is also quite straight forward (eq 2). Generally speaking, the average of an average is equal to the average of the whole set as long no weights are added. Hence I am not sure if this can be presented more easily?

Reply: We agree that the chosen notation was unnecessarily complicated and inconsistent. This has changed in the revised manuscript. The equations are of course relatively simple, but for the sake of clarity we believe that it might be worth to explicitly write them out.

We also realize that the delta-delta notation ($\Delta\delta^{18}O$) may seem a bit confusing. Lower case delta of course refers to the standard delta notation ($\delta^{18}O$) that is used for many isotope systems. What we then look in the testing of the model is the change in $\delta^{18}O$, and change is often denoted by upper case delta.

13. Sec 2.4: I find it very difficult to understand why particles are only traced in the GW and not form the entry into the system. I suspect this to be a limitation of the model. However, this forces the correction with eq. 4.?

Reply: Yes, this is a limitation in the model. Please see answer to question 2, regarding equation 4.

14. The second post processing step comes from the geometric vs. arithmetic MTT calculation. Obviously, this has dramatic impact. Using the median or a log(time) conversion could be other approaches. In the later course you refer to percentiles below 1 year… First of all, I would expect a little more reference to the literature about these steps. I suggest to clarify the used methods also with respect to your forthcoming comparison of Δδ18O and BC vs. gMTT. Maybe a more conceptual description about the theoretical similarities about the respective approaches would be helpful to prepare and clarify your findings? I would also appreciate to be able to discern what of your results can be attributed to the model or the decisions in your post processing.

Reply: It is indeed a complex question how to reduce a distribution to a single number in the most meaningful way. Both geometric and arithmetic means as well as median values and log time conversions are common choices (Destouni et al., 2001; Kaandorp et al., 2018; Massoudieh et al., 2012, 2017; Unlu etal., 2004) , which all have their strengths and weaknesses. None of them is, however, able to fully represent the underlying distribution, but the way we see it they are all valid metrics, which are ubiquitous in the literature. Figures A, B and C below show different metrics of the distribution compared to the old groundwater fraction (groundwater older than 5 years). Since the BC concentration is thought to be well correlated to the age of the groundwater, we needed a stable measurement that can account for the tail, without giving the tail to great weight. The main reason we favour geometric mean over arithmetic mean is that the latter is highly sensitive to the tail of the distribution, and this is where one might suspect that there is considerable uncertainty in the model. Whether the oldest particles reach ages of say 100, 1000 or 5000 years, will have profound impact on the arithmetic mean, but less so on the geometric mean (Taagepera, 2008; Unlu etal., 2004; Zhang etal., 1996). For example, the a_MTT goes up to 20-30 years (Figure A, below), while the g_MTT only goes up to 3-4 years (Figure C, below), much closer to the median. However, compared to the median (Figure B, below), the geometric mean still considers the length of the tail. The median of the distributions becomes almost equal for most sub-catchments, making it hard to say anything about the difference of the distributions. However, the geometric mean, because it takes the length of the tail into account, gives a spread to the travel times, which can be directly linked to the tail. All metrics are however stated in Table 5 and Appendix.

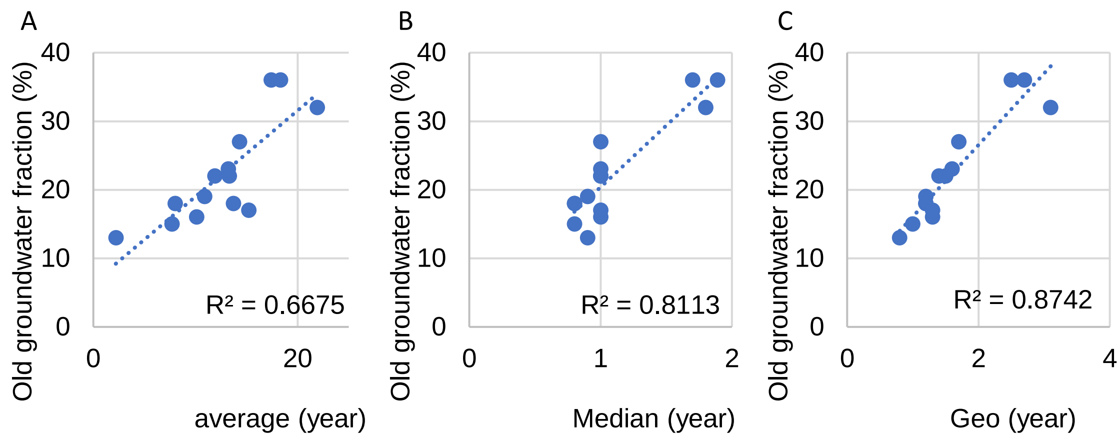


Figure A, B and C: The old groundwater fraction (>5years) compared to distribution metrics of arithmetic mean (average), median, and geometric mean (geo).

The geometric mean is mainly recommended to use as a when the distribution is significantly skewed, i.e., when the standard deviation (SD) is more than half of the arithmetic mean (Taagepera, 2008). This is the case for the travel time distribution. However, as for the isotopic signal and the base cation (BC) concentration, the SD is much smaller than the arithmetic mean (Table 2). As for the comparisons to BC and oxygen isotopes, the choice of metric is not crucial because they are all strongly correlated to one another. We realise that we should have emphasised this better in the manuscript. We agree it is vital to make sure that the findings ultimately reflect an underlying pattern and are not merely an artefact from decisions in the post processing of the results. We hope that this is clear in the revised manuscript.

References:

Destouni, G., Simic, E. and Graham, W.: On the applicability of analytical methods for estimating solute travel time statistics in nonuniform groundwater flow, Water Resour. Res., 37(9), 2303–2308, doi:10.1029/2001WR000348, 2001.

Kaandorp, V. P., de Louw, P. G. B., van der Velde, Y. and Broers, H. P.: Transient Groundwater Travel Time Distributions and Age-Ranked Storage-Discharge Relationships of Three Lowland Catchments, Water Resour. Res., 54(7), 4519–4536, doi:https://doi.org/10.1029/2017WR022461, 2018.

Massoudieh, A., Sharifi, S. and Solomon, D. K.: Bayesian evaluation of groundwater age distribution using radioactive tracers and anthropogenic chemicals, Water Resour. Res., 48(9), doi:https://doi.org/10.1029/2012WR011815, 2012.

Massoudieh, A., Dentz, M. and Alikhani, J.: A spatial Markov model for the evolution of the joint distribution of groundwater age, arrival time, and velocity in heterogeneous media, Water Resour. Res., 53(7), 5495–5515, doi:https://doi.org/10.1002/2017WR020578, 2017.

Taagepera, R.: Making Social Sciences More Scientific: The Need for Predictive Models. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199534661.001.0001, 2008.

Unlu, E. and Faller, J.F. (2001), Geometric mean vs. Arithmetic mean in extrusion residence time studies. Polym Eng Sci, 41: 743-751. https://doi.org/10.1002/pen.10770

Zhang, C., Zhang, S. A robust-symmetric mean: A new way of mean calculation for environmental data. GeoJournal 40, 209–212 (1996). https://doi.org/10.1007/BF00222547

15. As a minor point I was wondering if 1 particle per 20 mm GW recharge is really sufficient? That means that one year is represented by ±18 bins only…

Reply: Please see answer to question 6.

Results: I will go through the remainder rather quickly.

Fig 4: Where is the gMTT, MTT, median, percentiles etc.? How does this compare to estimates from data (before turning to the much more aggregated Fig. 6 and 7)? I think a brief summary of the nice water dynamics performance from your 2018 paper would be worth mentioning. (This could also come as staring point in your methods.)

Reply: We decided that the figure became confusing to showcase all these values in it. However, we agree that showcasing the values for at least one sub-catchment could be useful for a visual representation (see figure below). Most values can however be found in table 5 and Appendix Table A1. We compare our results with earlier findings in the discussion section. Here we have other studies made for C7. Otherwise, we are restricted to the other stream chemistry results shown in figure 6 and 7. We also added a brief summery if the results for the 2018 paper in the method section.

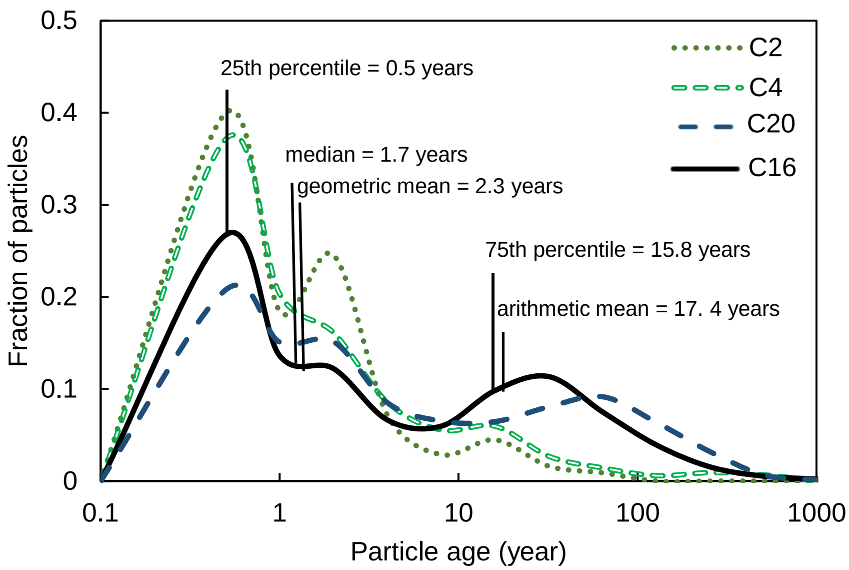


Figure 5: Examples of particle tracking results. The figure shows the age of particles reaching sub-catchment outlets. The solid line showcases the statistics for C16, including the 25th percentile, the median, the geometric mean, the arithmetic mean, and the 75th percentile (Appendix, A1). Moreover, the figure shows three other example distributions, including C2 (small till and forest-dominated catchment), C4 (small mire dominated catchment), C20 (small silt dominated catchment).

Tab 5 and Fig 5: I find it not easy to look through these many results with the few guidance offered. Likewise I found it difficult to reconnect your conclusions to the presented graphs and numbers.

Reply:

Table 5: Reviewer #1 also commented on Table 5, and we found that the explanatory text below the table was missing. We therefore placed this text above the table. We also made a shortened version of Table 5 with the most important information, and put the other information in Appendix, Table A1. We believe that A1 provides more in-depth information about the travel time distribution but this might not be necessary for the major results and discussion of this manuscript.

Figure 5: Fig. 5 contains plenty of information and they may therefore require some effort when looking through, but we agree that we as authors should provide the reader with sufficient guidance to make this as easy as possible. The inclusion of this figure was based on previous comments on the manuscript, where reviewers requested a more thorough presentation of the modelled distributions. To make figure 5 easier to understand, we added indicators for each season, starting with spring and ending with winter. We also added a descriptive text to the y-axis of the smaller chart, more descriptive titles for each sub-catchment and a line at 50 % so that the smaller charts more easily can be compared with each other. The figure text was also updated accordingly. See figure below:

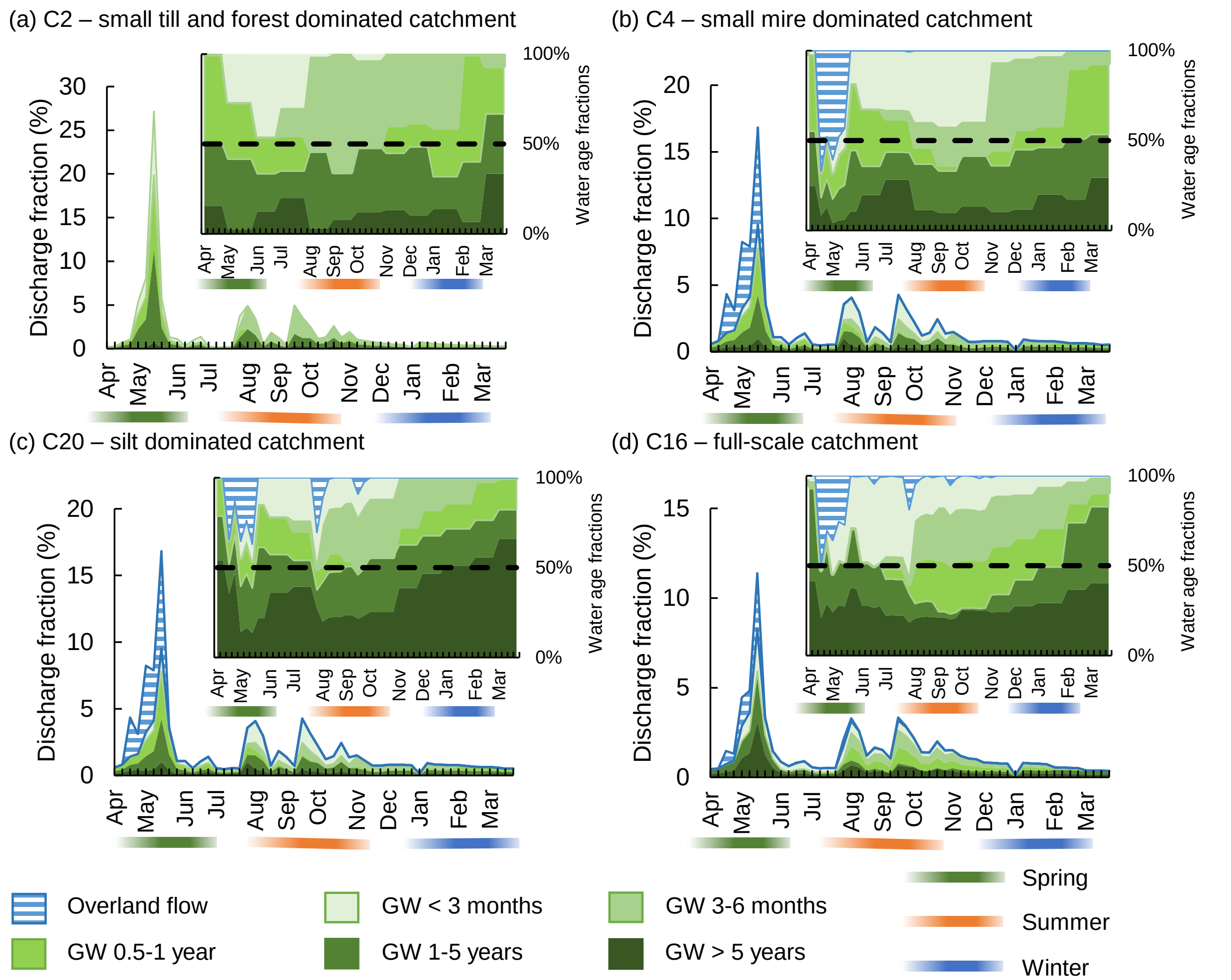


Figure 6: Seasonal fraction of discharge to streams. The figure shows the proportion of annual stream discharge arriving as groundwater and overland flow. Four sub-catchments are exemplified, including (a) the small till and forest dominated C2, (b) the small mire dominated C4, (c) the silt dominated C20 and (d) the full-scale Krycklan catchment C16 with mixed mires and forests (extended version in Appendix Fig. A1). The figure showcases the water age fraction discharging to the streams. The fractions are both shown as part of the total annual discharge as well as the water composition. The bands below the months highlight the three investigated seasons, spring, summer, and winter.

(L344 how can I see deep and long flow paths?)

Reply: We chose to remove this sentence. It was supposed to say that they are more connected to sub-catchments with deeper soil. But it seemed to work better as a discussion point.

Fig 5: What is "discharge fraction" in the main panels referring to? (percentage of annual discharge?) Why is the overall shape of the curves of each GW fraction very much similar although the flow paths should be very different? Would it be an option to avoid the C# but to refer to real descriptors (like whole catchment, small forest subbasin)?

Reply: Yes, it is percentage of annual discharge. Note that therefore the actual discharge from the small sub-catchment of C2 is much smaller than the discharge from the entire catchment.

Although the main shape is similar for all sub-catchments, the age of the groundwater as well as the fraction of overland flow is very different. We also decided to showcase the annual discharge on a weekly basis, because the daily discharge graphs were hard to distinguish. The monthly discretization also makes the shapes similar because daily differences are not showcased. We like to keep the C# so that the study can be compared easily to other studies in Krycklan. These denotations are already well established in the literature from Krycklan, which is quite extensive. However, we have added the suggested descriptors since we agree that these would clarify the examples.

Fig 6: Why is not always gMTT used on x-axis? Would it be an option to give focus to some sub-basins, which I can find in each plot? Eg. C20 is in b, d, f but not in a, c, e. Also the others from Fig 5 are not to be found in all plots. Maybe this could also strengthen your story, when you can adhere to some examples motivated throughout the manuscript?

Reply: We agree that we could give a better focus to some sub-catchments. We have changed these figures accordingly and showcase C2, C4, C16 and C20. Note that we don't have isotopic data for C20, and we therefore cannot showcase it in all plots. The same is shown in Fig. 5. We explained in the method section that the ^{18}O is more related to the young water fraction rather than to the actual age of the water. Therefore, we showcase MTT_{geo} in some plots and the young water fraction in some plots.

Discussion: Again just quickly.

I follow referee 1 that the focus should really be the data and the characteristics. When it come to the model, I do not see that there is sufficient basis about the employed concepts and numerics to attribute found behaviour to model specifics. Which are the most valuable and/or most delicate properties which will be affected by changes in snow melt and evaporation? What can we learn for other catchments with far less data than Krycklan can offer?

Reply: We agree with the editor and reviewer #1, that the discussion section could use some structural changes to make it clearer what the most important parts of the study has to offer. The first section is now dedicated to the limitations of the model. The second and third parts are dedicated to model and stream chemistry and their consistent picture of how the boreal landscape functions. The last section discusses the causes of intra-annual variations of MTT_{geo} and young water fraction, and which catchment characteristics these can be linked to. Finding characteristics that increases or decreases discharging travel times or young water fraction are important to understand, especially for boreal catchments. Future land-use or climate changes may cause soil types and landscape types to respond differently in the future and therefore, to have a well-established baseline understanding of these areas is important for future predictions.

You state that soil hydraulic conductivity is essential. However, this appears to be somehow one of the few site parameters. I am inclined to claim this importance a model artefact - especially given the rather broad definition (in orders of magnitude).

Reply: One uncertainty is related to the calibration of the flow model, especially regarding depth and soil properties (Jutebring Sterte et al., 2018). However, the flow model was calibrated and evaluated based on daily discharge measurements for the 14-sub-catchments, groundwater measurements, and all other available data from the Krycklan site. We believe that the well-represented groundwater levels and the correlations to stream discharge amount and peak flows gave a credible model. It should be noted that the study is still based on a model, therefore, we also used chemistry based indications, such as ^{18}O and BC, to ensure that the results are not related to a model artefact but instead related to the catchment functioning.

Following Darcy's law, which describes the theoretical understanding of the physics that underlies most geohydrological models, the hydraulic conductivity, soil porosity, and hydraulic gradient are the three parameters needed to model groundwater travel times, and of these the conductivity generally varies the most (orders of magnitude). We do not see this strong dependence on hydraulic conductivity as a model artefact, but as an expected outcome based on our physical understanding of groundwater flow. Interestingly, in previous studies for the site, this effect from soil properties on travel times wasovershadowed by the effect of the sizes of the different sub-catchments. One thing that the present study in combination with the new data from the C20 catchment has demonstrated is that small catchments can have long MTTs, but it all seem to depend on the soil type.

We also want to emphasize that the modelled travel times have a strong agreement with observable data, i.e., stream chemistry observations linked to travel times such as isotopic data and BC (without being calibrated to relate to these factors). Combined this indicates that the results are not just an artefact but is a result of a well calibrated physically based flow model whose results also can be linked to stream chemistry. Our study is also in agreement with previous studies which have shown that the main part of the groundwater discharged from the small till catchments is younger groundwater (Bishop, 1991; Kolbe et al., 2020; Laudon et al. 2011). Groundwater investigations have shown a much older age (i.e. several decade) already at a few meters depth below the water table compared to shallower groundwater, and a greater variability of the shallow groundwater isotopic signal. This demonstrates that the main part of water being discharged in the till dominated areas is relatively young groundwater. This has been linked to the decreasing hydraulic conductivity in the till soil, especially in the upper part of the till soil.

Bishop, K., Seibert, J., Nyberg, L., and Rodhe, A.: Water storage in a till catchment. II: Implications of transmissivity feedback for flow paths and turnover times. Hydrological Processes, 25(25), 3950–3959. https://doi.org/10.1002/hyp.8355, 2011.

Laudon, H., Berggren, M., Ågren, A. et al.: Patterns and Dynamics of Dissolved Organic Carbon (DOC) in Boreal Streams: The Role of Processes, Connectivity, and Scaling. Ecosystems 14, 880–893. https://doi.org/10.1007/s10021-011-9452-8, 2011.

Kolbe, T, Marçais, J, de Dreuzy, J-R, Labasque, T, Bishop, K.: Lagged rejuvenation of groundwater indicates internal flow structures and hydrological connectivity. Hydrological Processes; 34: 2176– 2189. https://doi.org/10.1002/hyp.13753, 2020.

.

You tackle the role of mires throughout the manuscript but it remains rather implicit. Maybe this could be worked out more clearly based on a hypothesis? Or is it really a side topic?

Reply: Wee agree that the fraction of mires could be stated earlier as an important factor to investigate. Therefore, we added a section to the introduction that states:

We focused especially on the catchment characteristics that have been suggested to be important factors for regulating stream chemistry of the Krycklan sub-catchments, including the areal coverage of mires, catchment size, soil properties, and seasonal changes in groundwater recharge (Karlsen et al., 2016; Klaminder et al., 2011; Laudon et al., 2007; Peralta-Tapia et al., 2015; Tiwari et al., 2017).

Conclusions

You state "Yet, the combination of stable water isotopes, stream water chemistry, and particle tracking provided a consistent picture of how the boreal landscape functions hydrologically and what processes and factors are of importance." I am not sure if these threads have been really brought together yet.

Reply: We like to thank the editor for this comment. We believe that this comment also highlights some of the issues with the structure of the discussion section. Therefore, we have re-written this section and re-structured it to strengthen the conclusion section. I.e., that the previous isotopic studies and our study link the young water fraction and variations of young water fraction to the fraction of mires, the mire frost, and the high hydraulic conductivity of the shallow part of the till soils. In turn the BC concentration can be linked to the discharging water age, which in turn can be linked to the soil properties and in part to the catchment size.

Moreover, the role of snow packs is not really worked out in your study (or I have missed it). Since it is also in the title I suggest to really give it more attention.

Reply: The main impact of the snowpack is the reduced groundwater recharge in winter, and the spring flood due to snowmelt in spring. We agree that the snowpack was given too much attention in the title, so we therefore changed the title accordingly.

Minor things I spotted:

L96f. Verb missing?

Snow-dominated landscapes have gained increasing attention due to their importance as water resources (Barnett et al., 2005) and their vulnerability to climate change the last decades (Tremblay et al., 2011; Aubin et al., 2018).

Reply: We appreciate that this error was noticed, and we have changed this sentence to:

Snow-dominated landscapes have received increasing attention in the last decades due to their importance as water resources (Barnett et al., 2005) and their vulnerability to climate change (Tremblay et al., 2011; Aubin et al., 2018).

Tab1 caption: Please clarify "Soil proportion based on the soil"

Reply: We have corrected this text according to:

Table 1: Sub-catchment characteristics. The list includes all 14 monitored sub-catchments in Krycklan, called C1 to C20, including the entire Krycklan catchment, C16. Different branches of the stream network are illustrated in

distinct colours in Fig. 1. The table includes the sub-catchment area, average elevation, and average slope. Further descriptions of these characteristics can be found in Karlsen et al. (2016). The table also includes soil proportion based on the soil map (1:100,000) from the Swedish Geological Survey (2014).

Reply to reviewer #1

General comments:

The paper has definitely seen some improvements during the last revision. The authors did a thorough job of considering and answering the reviewers questions and concerns. In particular, the introduction is more complete and the methods section contains more important details.

The addition of investigating the seasonal differences has definitely made the paper more interesting and impactful. Stressing the specific local settings (with a long, snowy, low-recharge winter period, mire abundance and regional soil types) and directing the investigation and discussion towards these factors moves the manuscript in the direction of presenting novel results and insights.

The presentation style and structure has, however, suffered from incorporating the new research. Often the manuscript is not that clearly written and the individual sentences are not connected in order to show relationships and tell a coherent story. Regularly there is a lack of precision in the wording that makes it hard to follow the discussion and arrive at an otherwise obvious conclusion.

The authors decided to move away from investigating the entire travel time distributions and concentrate on mean travel times and young water fractions alone. I find this to be a missed opportunity since the use of the distributed model allows for a more in-depth analysis of the catchment system dynamics (and travel time distributions contain a lot more information on that).

The discussion section is mainly filled with an evaluation of the catchment model (2.5 pages) while the new results concerning travel times only occupy 1 page. I would like to see a significant improvement of the way the discussion and conclusion is presented in terms of structure and writing style.

Reply: We would like to thank Reviewer #1 for his/her review of the manuscript. The constructive criticism concerns and suggestions have helped improve the new version of the manuscript even further. We agree that the introduction is at a higher level this time around and are pleased that the reviewer noticed the effort put into this section. We also appreciate that the reviewer thinks that the manuscript became more interesting due to the new focus on seasonal travel times. After reading the comments from the two reviewers and the editor we, however, agree that the text needs to be re-worked for better clarity. Here we used the specific comments to improve the language and clarity of the text. We especially addressed the concerns regarding the discussion and conclusion, with more focus on the actual results and we addressed the structure and writing style. Regarding the comment on the manuscript moving away from the entire travel time distribution is slightly confusing to us. The first iteration of the manuscript focused on the mean travel time of winter base flow. The distribution of the travel times was shown, but this is still the case in Figure 5 (new manuscript). We also believe that Figure 6 (new manuscript) gives a better representation of the travel time distribution for each stream and how it varies within a year, which the previous version of this study did not have.

Our explanations and responses to all Reviewer #1 comments and questions are listed below. All the comments have been included unabridged in this document, together with our response following each comment in brown. Blue sections are text directly imported from the old version of the manuscript, and red was taken from the new version. Please note that the abbreviation of geometric mean of the travel time distribution and the young water fraction is now shorten to MTT_{geo} and young water fraction respectively.

Specific comments:

Key points

Line 18: '…while THE FRACTION of mires…'

Reply: We have changed this to "areal coverage of mires"

Abstract

Line 29: Why do you write 'the following snowmelt'? Following what exactly?

In this study, a particle tracking model approach in Mike SHE was used to investigate the travel time, and pathway of water in 14 partly nested, long-term monitored boreal sub-catchments characterized by long and snow rich winters with little groundwater recharge and highly dynamic hydrology during the following snowmelt.

Reply: This sentence was changed to:

In this study, a particle tracking model approach in Mike SHE was used to investigate the travel time, and pathway of water in 14 partly nested, long-term monitored boreal sub-catchments. These sub-catchments are characterized by long and snow-rich winters with little groundwater recharge and highly dynamic runoff during spring snowmelt.

Line 30: Better describe what potentially causes the seasonality in MTTs (changes in precipitation, temperature, vegetation?). 'Seasonality' causing seasonal differences is not a good explanation…

The seasonality caused considerable variation in travel times between different seasons and landscape types, with winter mean travel times ranging from 1.2-7.7 years and spring mean travel times ranging from 0.5-1.9 years.

Reply: This sentence was changed to:

The variations were found to be related to the distribution of different landscape types and their different response to seasonal changes. Winter MTTgeo ranged from 1.2-7.7 years, while spring MTTgeo varied between 0.5-1.9 years.

Line 32: '…(YOUNGER than three months)...'

Reply: We have changed this to <3 months

Line 34: This is the only time in the entire manuscript that you mention the word 'synchronicity'. This should be explained in the results, discussion and/or conclusion section.

Reply: We appreciate that reviewer #1 noticed this, and we have addressed it in the discussion section.

Line 36: What about the mires affected the young water fraction? Was it the fraction of mires?

Mires were found to affect the young water fractions of the stream contribution

Reply: This sentence was changed to:

The areal coverage of mires was found to be especially important for the contribution of young water in spring …

Line 37: So catchments with mires cannot also be forest-dominated?

Mires were found to affect the young water fractions of the stream contribution (r=0.96, P<0.0001) by introducing larger differences in the mean travel times between the seasons compared to forest dominated sub-catchments.

Reply: Yes, that is true. Sub-catchments can have a minor mire areas but still be forest dominated. What we were trying to say here is that the more mires there are in a catchment, the more we can see a difference between seasonal travel times. We tried to clarify this by changing the sentence to:

The areal coverage of mires was found to be especially important for the contribution of young water in spring (r=0.96, P<0.0001). The main factor for this was attributed to extensive soil frost in mires, causing considerable overland flow during the snowmelt period…

Introduction

The introduction has improved a lot. Previous work is now covered more completely and the authors have obviously read and inserted the suggested additional references appropriately. Also the focus and objective of the study has been shifted towards exploring TTDs and their controls in different seasons which adds significantly more relevance to the results.

Line 105-106: A strange sentence: '…water that have spent…', '…sub-subsurface…'.

These specific conditions provide unique opportunities to study the source of water that have spent the longest time in the sub-subsurface environment.

Reply: This sentence was changed to:

Landscapes with long-lasting snow cover that often melts rapidly in the spring creates both opportunities and challenges for determining stream water age and pathways.

Method

Line 152: '…has been included…'?

The catchment is divided into 14 nested sub-catchments and has been included multi-disciplinary biogeochemical and hydrological research for more than 30 years (Laudon and Sponseller, 2018).

Reply: This sentence was changed to:

For more than 30 years, multi-disciplinary biogeochemical and hydrological studies have been conducted in Krycklan (e.g., Laudon and Sponseller, 2018). Streamflow is monitored in 14 nested sub-catchments, called C1 to C20, with the longest continuously monitored time-series starting at the beginning of the 1980s.

Line 183: '…gave a more precise location…'? What do you mean by this? Better spatial resolution in general? A more precise location of certain features? And what happened to 'the threshold level of the lake outlet of C5'?

Reply: Both the reviewers and the editor were confused by the Mike-SHE model section. We have therefore updated this section for more in-depth information regarding the model tool itself. Regarding the question at hand, streams are modelled using a 1D module for river hydraulics called Mike 11. This module also uses high-order dynamic wave formulations of the Saint-Venant equations. The Mike11 model is not restricted to the grid size of the Mike-SHE model. Instead, Mike 11 can be much more complex to allow for a more precise calculation of water levels and flow rates. If the cross section of the stream is smaller than the grid size the M11 cross section is placed between two adjacent grid cells in MIKE SHE, which is the case in the present study, the small forest streams in Krycklan are much smaller than the applied grid size of the MIKE SHE model. Due to this coupling between the MIKE SHE and the MIKE 11 model we could give the outlet of the sub-catchment a more precise location as well as the lake threshold level.

Line 193: What happens to particles that leave the saturated zone but re-enter it later (re-infiltration)? Is that somehow accounted for? If not, in which way would that influence your results?

Reply: This is a source of uncertainty in the particle tracking model. Particles that reach the unsaturated zone or the overland waters are removed from the calculations and are not used to calculate the mean travel times for the stream flow. If this part of water was accounted for it would generally increase the mean travel times for most sub-catchments. However, all recharge areas over an annual cycle, are included as sources of particles in the model. Which means that areas where water are discharging under certain conditions might turn into recharge during another time of the year. This means that we include most of the flow paths that annually contribute water to the streams.

Line 228: I would prefer a more unambiguous description of the isotopic value. Instead of 'more depleted' I would write 'more negative' because the signal could be more depleted in the heavy but also more depleted in the lighter isotope.

Reply: We agree that "more depleted" is an ambiguous term unless it is specified which isotope is depleted. It seems to be widely used when describing isotopic fraction, but it would clearer to use another description. Therefore, we changed this to "much lighter (depleted) signal".

Line 230: In the equation you write 'spring/summer', above and below you only use 'spring'.

Reply: We have corrected this in the new revision of the manuscript.

Line 273-274: Does this include snowmelt isotope measurements? If not, a huge amount of water input to the groundwater (with very different isotopic composition) would be neglected.

Reply: This data includes all precipitation, i.e. both rain and snow. We added a statement about this in the method section regarding the chemistry data.

Line 290: What about the time it takes to reach the water table? What deviations from the 'real', 'complete' MTT do you expect when using this simplification?

Reply: We addressed this uncertainty in the discussion section:

Another uncertainty related to the particle tracking model in Mike SHE is related to the travel time from the point of infiltration through the unsaturated soil horizons to the saturated groundwater. Due to technical limitations, this travel time cannot be accounted for in the particle tracking calculations. Particles are placed at the groundwater table proportionally to the groundwater recharge (Fig. 4). Therefore, the main fraction of particles introduced to the model occurs at high infiltration rates when the groundwater level is close to the soil surface. Under these conditions, the water has, in most cases, spent a relatively short time in the unsaturated zone. However, some particles are also introduced when the groundwater level is lower, such as early snowmelt or following extended dry periods. Under such conditions, the model uncertainty increase. In this context, the smallest potential uncertainty occurs in mires that seldom experience a groundwater table below one meter below the soil surface. The uncertainty becomes somewhat larger in the till areas where the unsaturated zone on average is above 1 m but can extend down to 3 m below the ground during low flow. C14 and the lower part of C16 are exceptions to these relatively shallow saturated conditions as a deep esker traverses the sub-catchments resulting in a groundwater level up to 10 m below the soil surface (Fig. 1). Accounting for the travel time from infiltration to recharge could impact the results and provide, especially for C14 and C16, longer MTT than if the groundwater level were at the same level throughout the whole catchment. This limitation primarily affects catchments with the longest MTTs and, therefore, does not seriously question the general pattern observed. The distance from the ground surface to the groundwater table is for most model cells much shorter than the distance to the nearest stream so most of the transit time should be related to the groundwater flow rather than to percolation. Although water, especially during dry conditions, no doubt can spend considerable time in the unsaturated zone, it must also be acknowledged that this water volume is small compared to the groundwater inventory in the saturated zone so its impact on the average MTTs should be relatively small.

Line 297: Can a standard deviation be 'high'. I'm not sure – but I'm not a native speaker.

Reply: We are not native speakers either, but to our knowledge, a high standard deviation indicates that the data is more spread out, while a low standard deviation means that the data is more grouped around the mean. Our impression is that "high" and "low" frequently are the words used to describe the standard deviation.

Line 298: The median would also be an appropriate measure that has been used frequently.

Reply: We agree that median also can be used to find the bulk of a population. However, it only states the middle value regardless of the size of the values to the right and left of the middle value, whereas the geometric mean will account for the size of the values. We discussed this in an answer to the Editor (see answer to Editor question 14) and extended the section regarding the use of the geometric mean in the method section.

Line 301: I am confused because I thought you wanted to look at MTTs for different seasons. What do you mean by '…when the median was stabilized for all sub-catchments'? Do you mean that the seasonal change in the median TT was stable for all sub-catchments?

To identify the minimum particle tracking time needed for robust travel time estimates, we compared median travel times for varying lengths of particle tracking. We assumed that the calculation was run for enough time when the median of the travel time was stabilized for all sub-catchments. The median stabilized after 500 years of simulation time, but in the end, we let the particle tracking run in total 1000 years to ensure that the results were stable for all parts of the catchments. Thereafter, we used all particles that reached a stream or lake to calculate mean travel times for each sub-catchment.

Reply: We agree that this part was confusing. We changed most of this section, so we cannot state the exact change, but this is the most relevant change in response to this question:

Particle tracking was used to assess groundwater travel times from groundwater recharge to stream runoff for each sub-catchment. The model was run for 1000 years to capture the travel times of all discharging groundwater for each sub-catchment. One year of simulated flow results from Jutebring Sterte et al. (2018) was cycled 1000 times to extend the particle tracking simulation. The year 2010 was selected, as the water balance was close to the long-term annual averages observed for the Krycklan catchment. All particles were released at the top of the transient groundwater table the first year. Numerical constraints restricted the number of particles released to 0.5 particles/10 mm modelled groundwater recharge, which corresponds to a total of approximately 0.6 million particles. This number of particles was assumed to be enough to capture the timing of recharge patterns (Fig. 4).

Line 306: Don't forget the time that water needs to percolate to the water table. This could be significant in certain situations (especially when it's dry).

Reply: We addressed this concern at the comment of line 290. Unfortunately, this is a limitation in the Mike SHE.

Results

Line 338-340: A bit confusing. So now it reads as if first you determined the mean travel time of all catchments then you determined the groundwater mean travel time and then you determined the groundwater geometric mean travel time? It would be easier to understand if you connected the sentences better…

The particle tracking model in Mike SHE was used to establish mean travel time in the 14 sub-catchments. The time from groundwater recharge until the groundwater reached a stream was used as an estimation of groundwater travel time. The geometric mean (gMTT) was used to describe the central tendency of travel times because of the skewed distribution (Table 5, Fig. 4).

Reply: We like to thank the reviewer for this comment. We tried to clarify this section and have changed it to:

The particle tracking results were used to establish travel time distributions and MTT of stream water runoff of the 14 sub-catchments in Krycklan. Since the travel time distributions were significantly skewed, we assumed that the geometric mean of the travel time distributions provided the best representation of MTT (Table 5, Fig. 5). However, all metrics are stated in the Appendix, Table A1. The annual MTTgeo for all sub-catchments ranged from 0.8 to 3.1 years (Table 5). Most groundwater discharging to a stream had a travel time of less than one year (34% to 54%). The longest stream MTTs were connected to silt dominated catchments such as C16 and C20. We used some sub-catchments for result representation, but all results are provided in Table 5 and Appendix A1. The displayed sub-catchments were: C2 (small till and forest dominated catchment), C4 (small mire dominated catchment), C20 (small silt dominated catchment), and C16 (the full-scale Krycklan catchment).

Line 350: Enhance or dilute? These are not antonyms. What do you mean exactly? Also it is strange to mention that they are affecting stream solutes in 'different ways' without giving any more details…

On an annual basis, a fraction of the water reached a stream as overland flow, which may enhance or dilute various stream solutes in different ways.

Reply: What we meant to say by "enhance or dilute" was the the concentration of a certain solute might *either* increase *or* decrease depending on the difference in concentration between the stream water and the overland flow. For clarification we change this sentence to:

On an annual basis, a fraction of water reached the streams as overland flow. A major part of the overland flow occurred during the snowmelt in spring, especially in sub-catchments with mires such as C4 (Fig. 6).

Line 353: You are using age and gMTT interchangeably. Please stick to one term for clarity. Also, I think you want to say that the 'discharging' groundwater is youngest. There could be very old groundwater in the system that does not discharge to the stream during that period of time.

Line 354: Try to be clearer in your use of language. It is not the groundwater that 'increases'. It is either the age of the groundwater or the gMTT of the groundwater that increases.

Line 355: Again, more clarity is required. Travel times are not 'old', travel times are 'short' or 'long'. Discharging water can be 'old' or 'young'.

Line 353 to 355: On an annual basis, a fraction of the water reached a stream as overland flow, which may enhance or dilute various stream solutes 350 in different ways. The major part of the overland flow occurs during the snowmelt in April to May, especially in sub-catchments with mires such as C4 (Fig. 5). Each site has the oldest age during the winter season (1.2-7.7 years) and the youngest age in spring and summer (0.5-1.9 years). The input of new water is also reflected in the seasonal groundwater gMTT. The groundwater is youngest in connection with the snowmelt during late spring, then increases during the summer period with little groundwater recharge (Jun-Jul). The oldest groundwater travel times occur during the winter, before the beginning of snowmelt in late March 355 or early April.

Reply: These comments regarding Line 353-355 are connected to the same section. Therefore, we will address them at the same time. We agree that this section could be re-written for clarity. We agree with the suggestions by the reviewer and have used them to change the section to:

On an annual basis, a fraction of water reached the streams as overland flow. A major part of the overland flow occurred during the snowmelt in spring, especially in sub-catchments with mires such as C4 (Fig. 6). Both the fraction of young water reaching the streams and the MTTgeo displayed strong seasonal trends. The longest seasonal MTTgeo, 1.2-7.7 years, and the smallest young water fraction were found during the winter season. In winter, the fraction of older water successively increased until the spring snowmelt began in early April. Conversely, the smallest fraction of old discharging water and short MTTgeo, 0.5-1.9 years, were connected to events of larger groundwater recharge, such as the spring snowmelt and heavy summer rains.

Line 361: Why do you write 'despite a larger young water fraction'? Since this larger fraction is not considered in the particle tracking, it should not have a big influence on the groundwater gMTT. Also, why do you write 'even more'? The decrease in groundwater gMTT for C4 is smaller than for C2.

Mire sub-catchments have the youngest mean travel time during spring snowmelt. However, as exemplified by the similar-sized sub-catchments of C2 (forest) and C4 (mire), groundwater is not renewed to the same extent in mires as in forested sub-catchment (Table 5). The groundwater gMTT of C2 was reduced from 1.2 years to 0.7 years from winter to spring. In C4, groundwater gMTT 360 was reduced from 1.5 years to 1.2 years, despite a larger young water fraction. The overall gMTT of C4 decreased even more, from 1.5 years to 0.7 years. A more pronounced seasonality in mean travel times also occurs for catchments with a larger proportion
of mires combined with low conductive soils (LCS). For example, C20 had an overall gMTT that reduced from 7.7 years to 1.9 years from winter to spring, while the overall gMTT of the similar-sized till sub-catchment C6 only changed from 2.8 to 0.6 years.

Reply: We tried to refine this section and clarify which numbers in table 5 we refer to and why. Table 5 was also shortened to the values we discuss most. However, all other information is still available in appendix. We changed this section to:

In spring, mire sub-catchments had the shortest MTTgeo. However, as exemplified by the similar-sized C2 and C4 sub-catchments, groundwater was not renewed to the same extent in mire dominated systems due to a larger fraction surface runoff (Fig. 6). Mire dominated sub-catchments (like C4) displayed stronger seasonal variations in MTTgeo, with shorter MTTgeo than till dominated sub-catchments (like C2) in spring and longer MTTgeo than C2 in winter (Table 5). In C4, the MTTgeo reduced from 1.5 to 0.7 years from winter to spring, while the corresponding change in C2 was 1.2 to 0.7 years. The seasonality of MTTgeo was even more pronounced for catchments with a larger areal coverage of mires combined with a larger areal coverage of silt. For example, C20 had an MTTgeo that reduced from 7.7 years to 1.9 years from winter to spring, compared (Table 5).

Line 390: 'strong statistical significance' or 'significant statistical correlation'?

The overall gMTT always had a strong statistical significance to the BC concentration (Fig. 6 b, d, and e), generally agreeing with our conceptual model (Fig. 2b).

Reply: We have clarified this sentence to:

The MTTgeo had a significant correlation to the BC concentration during all seasons (Fig. 7 b, d, and e), again agreeing with our conceptual model (Fig. 2b).

Discussion

I would wish for a careful revision of the style and structure used throughout the discussion. Often there is a lack of cohesion between the individual statements and it is not explained why statement a) would support statement b). One example of this can be found at lines 471-473.

According to the conceptual model (Fig. 2), older baseflow water should result in an isotopic signature closer to the precipitation average. There was a strong negative correlation between groundwater age and the streams isotopic signature during baseflow (Fig. 6a), suggesting that the model produces credible water age patterns for the winter season. The larger sub-catchments, including C14, C15, and C16, are close to the long-term precipitation average, which limits the ability to estimate the travel times using isotopes. Water older than 4-5 years is argued not to be accurately quantifiable using isotopes due to amplitude loss (Kirchner., 2016). These theoretical considerations strengthen our results of a winter MTT of 4-6 years for the larger sub-catchments and provided new insights into travel times for these systems.

Reply: We have tried to restructure the discussion to get a better and easier to follow structure and style. Regarding the example, we changed this paragraph to:

Following the conceptual model (Fig. 2), patterns in stream isotopic signatures can be explained by seasonal changes in travel times. The modelling results show that all sub-catchments discharged the oldest water in winter, somewhat younger water in summer, and water of the youngest age in spring. When winter arrived, the main precipitation was snow resulting in that groundwater recharge effectively ceased, which caused an increasing proportion of old groundwater discharging into the streams (Fig. 6). In agreement with our conceptual model (Fig. 2), a strong negative correlation between winter MTTgeo and the isotopic stream signatures during winter baseflow was observed (Fig. 7a). At an average water age older than four years, it can be expected that the groundwater has reached full mixing. Hence, older water can no longer be accurately quantified using water isotopes only due to amplitude loss (Kirchner., 2016). These theoretical considerations strengthen the results of a winter MTTgeo between four and six years for the larger sub-catchments as their stream isotopic signatures were close to the long-term precipitation average and, therefore, should have reached complete mixing.

Line 430: '…for the MEAN water age…'.

Reply: We changed this sentence accordingly.

Line 437: I would add the acronym of the model for better recognition (STARR or something like that).

Reply: We changed this sentence accordingly.

Line 450-459: I understand your reasoning. However, especially in catchments where the soils have low conductivities, the difference between groundwater TTs and overall TTs could be quite large. Can you at least try to give some estimations on how long the percolation through the different soils to the groundwater table would take?

Reply: The hydraulic conductivity in the unsaturated zone varies with level of saturation, which in turn varies every time step. This makes it difficult to estimate travel times for the unsaturated zone. We therefore focused on the groundwater recharge and saturated zone flow and not unsaturated flow. Given that the distance between the ground surface and the groundwater table is so short compared to the average distance from a cell in the model to the nearest stream we would expect that most of the transit time in the soil is spent in the saturated zone. However, there are no doubt water that also can spend considerable time in the unsaturated zone, but when looking at the average transit time we must also acknowledge that the volume of water in the unsaturated zone is small compared to the groundwater compartment. Therefore, we would envisage that this limitation in the particle tracking in most cases should not have dramatic consequences for the modelled transit times. We also added a discussion part about this in the discussion section 4.1.

Line 472: How does the negative correlation suggest that the model produces credible patterns?

According to the conceptual model (Fig. 2), older baseflow water should result in an isotopic signature closer to the precipitation average. There was a strong negative correlation between groundwater age and the streams isotopic signature during baseflow (Fig. 6a), suggesting that the model produces credible water age patterns for the winter season.

Reply: Please see the first answer for the discussion section

Line 484: 'less' is not the correct descriptor for a difference.

The larger young water fraction was generally found in mire dominated sub-catchments, such as C4 and C5. In contrast, equally sized sub-catchments without mires, such as C1 and C2, had a less $\Delta\delta O_{18spring}$ and hence smaller young water fraction.

Reply: We agree that this is not a correct descriptor. However, the specific sentence was removed.

Line 489: 'biogeochemical consequences'?

Those earlier results showed a large overland flow component in wetland catchments because of frozen conditions during spring flood with biogeochemical consequences during snowmelt.

Reply: This section was changed and biogeochemical consequences, is no longer used in the text.

Line 491: In general, try to be more precise. Instead of writing: 'Inter-annual precipitation and evapotranspiration variability likely caused the relationship to be less evident…', better write 'A larger inter-annual precipitation and evapotranspiration variability likely caused the relationship to be less evident…'. This also applies to the variable names – one time you write 'the summer ΔδO18', another time you write 'ΔδO18summer'. Be more consistent.

In summer, the conceptual model predicted that ΔδO18summer should also be correlated to the young water fraction, but with the opposite sign, due to the enriched summer rain (Fig. 2). Inter-annual precipitation and evapotranspiration variability likely caused the relationship to be less evident compared to the spring flood results as the snowmelt conditions are more consistent from year to year. However, although less strong than compared to the spring ΔδO18, there was still a strong connection between the average summer ΔδO18 and the modelled young water fraction (Fig. 6e).

Reply: We agree to these changes. We have applied them to the section in question and checked the naming of $\Delta\delta^{18}O_{summer}$ I the entire manuscript.

Similar to the conditions in spring, the conceptual model predicted that the difference in stream isotopic signature between winter baseflow and summer flow, Δδ18Osummer, should be correlated to the young water fraction in summer, but with the opposite sign, due to isotopically heavier summer rains (Fig. 2). A larger inter-annual variation in precipitation and high ET likely caused the relationship to be less evident compared to the spring results as the snowmelt conditions are more consistent from year to year. However, although less evident than compared to the Δδ18Ospring, there was still a significant correlation between the average Δδ18Osummer and the modelled young water fraction (Fig. 7e).

Line 525: Again, more precision in how you describe your results is necessary: It is 'the fraction of low conductive soils' and not just 'low conductive soils'.

Reply: We have checked the manuscript for this error and have changed it accordingly.

Line 550: They don't 'receive', they 'produce'. Also, what exactly is the 'annual age of water'?

Sub-catchments with mires receive the highest young water fraction during the spring snowmelt; however, the annual age of water is not as strongly connected to that landscape feature (Table 6).

Reply: We have checked the manuscript for the word receive and changed it to produces whenever it seemed appropriate. We agree with the specific line change, however, the specific sentence was removed in the new version.

Line 555: A 'similar annual age'? Please be more precise.

Reply: This specific sentence was removed in the new revision of the manuscript. However, it was meant to be the mean annual age expressed as MTT_{geo}, Table 6.

Line 558: Does it displace the older water or is it just added to the older water (and mixed)? One would cause the stream water to first become very old and then very young, the other one would cause the stream water to become somewhat younger on average.

Line 559: What you describe is not called 'transmissivity feedback'. Transmissivity feedback describes the process that discharge from hillslopes increases over-proportionally when the storage in the hillslope increases - due to the fact that soil layers closer to the surface have higher hydraulic conductivity. Sloppy and makes me question the general understanding of the processes taking place.

Reply: Line 558-559: We understand that this section needed to be restructured, since the message was lost. As regards the "transmissivity feedback" we believe we fully agree with the reviewer on what it means. We regret that we expressed ourselves in a manner that led the reviewer to think otherwise and hope that this matter will be clearer in the revised manuscript. We hope the new section is clearer:

Earlier studies have demonstrated that fluxes of old groundwater are more stable throughout the year than younger groundwater showing a more variable temporal pattern (Rinaldo et al., 2011; van der Velde et al., 2015; Kaandrop et al., 2018). In our system, such a pattern can mechanistically be linked to till soils dominating most sub-catchments, where the groundwater response to precipitation events can be described by transmissivity feedback processes (Bishop, 1991), caused by the fact that the hydraulic conductivity increases exponentially towards the soil surface. When water infiltrates the ground, the water table rises and activates more conductive soil layers, resulting in rapid increases in the lateral flow. This implies that much of the water transport in till soil occurs relatively close to the surface, while the groundwater in deeper layers is more stagnant, which further explains the relatively short and consistent MTTs of till soils. Measurements of chlorofluorocarbons (CFCs) further support that deeper groundwater water transport in till soils in Krycklan is slow. Not far below the groundwater table CFCs have indicated that the groundwater can be several decades old, suggesting that most of the groundwater transport occurs close to the surface (Kolbe et al., 2020). Consistent with this explanation, silt dominated areas, that have more consistent hydrological conductive with soil depth, had much longer MTTs than comparatively sized sub-catchments underlain by till soils (Fig 6, Fig. 8 and Appendix Fig. A1).

Figures:

Figure 1: Catchment topography is shown in panel (d) – not (e).

Reply: Thank you for noticing this error. The error was corrected.

Figure 2: (a) The use of the symbol ∝ is incorrect here. It means 'proportional to'. The difference in the isotopic signature can only be proportional to the fraction of young water. Also I think you are missing a Δ before baseflow MTT. (b) You haven't explained yet what gMTT is. Also, there should not be a line for the precipitation BC concentration (it is not connected to any MTT), you can instead mark it on the y-axis.

Reply: We have changed this figure according to both the editor and reviewer 1# suggestions. The new figure (a and b) is displayed below. The main changes are, ∝ have been removed, MTTgeo was removed and changed to groundwater age, and the line for precipitation was removed for clarity. We also added a description of $\delta^{18}O$ and BC directly in the figure, instead of mainly having it in the figure caption

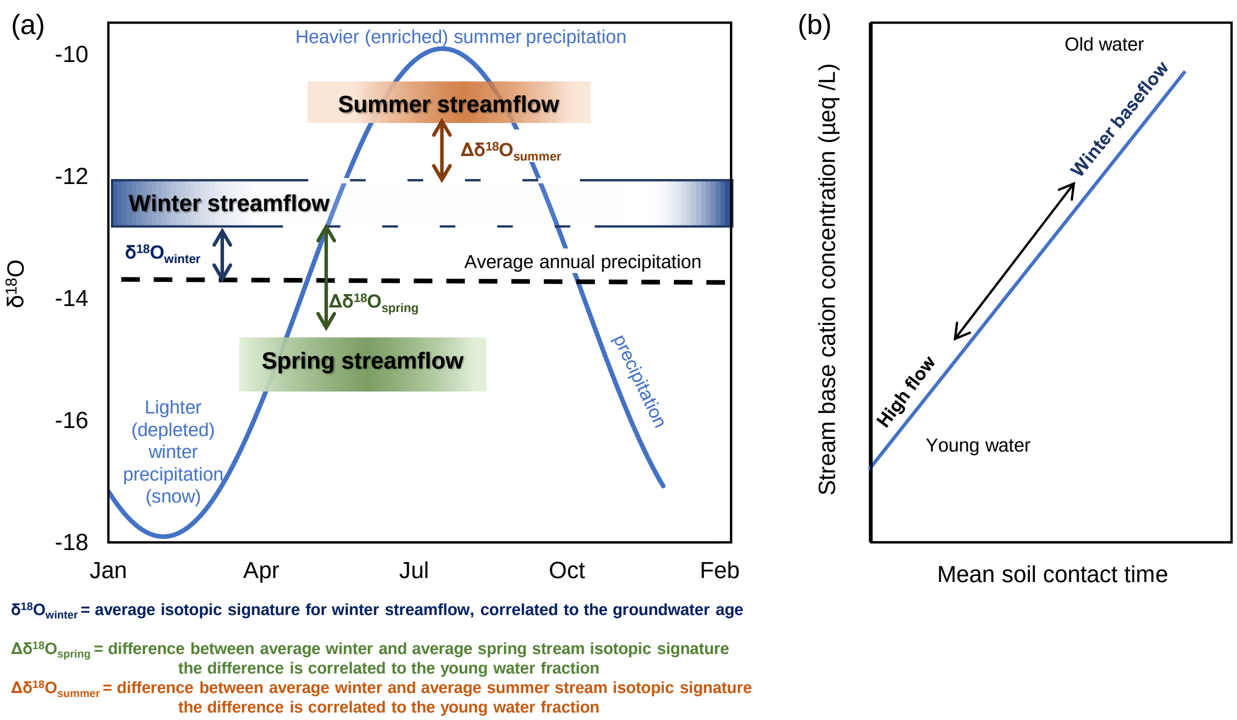


Figure 3: Better write '(a) Steps of the particle tracking'. In panel (b) write 'Average depth to the groundwater table'. I don't understand what you want to say here: 'Note that the top vertical layering of the saturated zone was set to 2.5 m below the ground surface, and the thickness thereafter follows the soil layers (thickness increasing with depth).' Can you add another panel to the figure to show what you mean?

Reply: We changed the figure and caption according to the reviewer's suggestions. The text in question was, however, removed from the caption and better described in the method text.

Figure 4: The figure shows the 'age' of the particles reaching the outlet, not the 'timing' as you write in the caption.

Reply: We agree that "timing" should be changed to "age"

Tables

Table 1: What are the numbers in parentheses in the 'Sorted sediments' column?

Reply: Thank you for noticing this error. Some text has disappeared in an earlier iteration of the manuscript. We have added this yet again to the table text:

Table 1: Sub-catchment characteristics. The list includes all 14 monitored sub-catchments in Krycklan, called C1 to C20, including the entire Krycklan catchment, C16. Different branches of the stream network are illustrated in distinct colours in Fig. 1. The table includes the sub-catchment area, average elevation, and average slope. Further descriptions of these characteristics can be found in Karlsen et al. (2016). The table also includes soil proportion based on the soil map (1:100,000) from the Swedish Geological Survey (2014).

Table 2: One row just reads: 'Soil frost'. What does that mean?

Reply: This part of the table was removed and written as text in the method section for clarity. We applied soil frost on the mires since it was an important part of the hydrological functioning of the catchment, which mean reduced horizontal flow and infiltration at freezing temperatures and at the start of the snowmelt.

Table 5: Is aMTT the arithmetic mean of the 'particle tracking (groundwater gMTT)' or of the 'particle tracking adjusted for overland flow according to Eq. (4)'? Please clarify.

Reply: We apologize. The explanation for these was stated below the table. We found that this information easily was missed when looking at the table. Therefore, we moved it to above the table. We also made the table smaller to just include the most important parts discussed in the manuscript. However, all data are still available in the Appendix.

Table 6: I guess something went wrong in the formatting of the caption 'according to a p>0.05, b a p<0.05, and c p<0.01'.

Reply: Thank you for noticing this error. It has been corrected.

An idea: Since the upper two rows of each table are identical, why not make one bigger table out of the four smaller tables? Something similar to the attached table (or transposed).

Reply: Wee agree that this table could be smaller, and we like the suggestion from the reviewer. We used some of the ideas and remade the table:

Table 6: Correlation matrix – young water fraction (YWF), geometric mean travel time (MTT_{geo}), and catchment characteristics. The catchment characteristics include the log catchment size, km^2 (Log C.-size), the areal coverage of mires, and the areal coverage of silt. The table includes annual (grey), winter (blue), spring (green), and summer (orange) results. Darker colours show $|r| > 0.5$ with the connected p-value according to [a] $p<0.05$ and [b] $p>0.05$.

	Winter season					Summer season				
	Log C.-size	Mire (%)	Silt (%)	MTT_{geo}	YWF	Log C.-size	Mire (%)	Silt (%)	MTT_{geo}	YWF
Log C.-size	1					1			0.91 [a]	
Mire (%)		1	0.92 [a]		0.64 [a]		1	0.80 [a]	-0.50 [b]	0.68 [a]
Silt (%)	0.58 [a]		1		0.58 [a]	0.58 [a]		1		0.58 [a]
MTT_{geo}	0.63 [b]	-0.51 [b]	0.90 [a]	1		0.55 [a]	-0.55 [a]	0.92 [a]	1	
YWF (%)		0.96 [a]		-0.53 [b]	1		0.95 [a]		-0.52 [b]	1
	Annual					Spring season				

Appendix

Line 835-837: What is ‘ai’?

Reply: We added this text for clarification: whereas: ai= data set values, n=number of values.

Line 841: ‘is Krycklan’?

Reply: This has been corrected to: “in Krycklan”

Supplements

The supplement was supposed to be deleted, I guess. It is still available with the manuscript though. Don’t forget to delete it, otherwise readers will be confused.

Reply: Yes, we chose to remove the supplements and have all necessary information in the manuscript and the Appendix. We thank the reviewer for pointing it out and we contacted the editorial support team about this issue. They in turn told us not to upload a new version of the supplements. Hopefully, this issue will therefore be resolved for the next version of the manuscript.

Reply to reviewer #2

General comments:

I would like to commend the authors on the extensive revisions they undertook in response to the two reviews in the first round. The most significant concern raised by myself, and the other reviewer, was that the organization around the catchment size hypothesis lacked novelty. I commend the authors for taking this comment to heart and significantly revising the text to re-focus their analysis on the aspects of their study that are more novel: (1) their unique study location and (2) the ability to explore seasonal differences in MTTs within this setting. I think this change along with the expanded literature discussion greatly improved the manuscript. And I agree with the authors that this new focus brings much more novelty to the study.

Reply: We thank reviewer #2 for his/her time and effort to read, comment, critique and discuss our manuscript. We are delighted to read that reviewer #2 agrees with most changes of the manuscript compared to its earlier state, especially with regards to the new study focus. The remaining concerns from reviewer #2 has been commented on below (brown). Text from the new manuscript was marked red. We hope that the new revised manuscript and the answers to all reviewers and editor concerns and remaining critiques will be sufficient for publication.

I also appreciate the addition of Figure3. This makes the approach much more clear and addresses some of my concerns about the confusion of the methodology.

Reply: Thank you for the comment, and we agree that the Figure 3 (old manuscript) made the method section clearer.

My only remaining concern is in the description of the MIKE-SHE model itself. While I realize that most of this is published in other locations, I was hoping that the authors would provide a better description of how the model is actually working -- ie. how physical processes and interactions are represented. Section 2.2 was modified but is still mostly about the parameterization of the layers. I think this section should be updated to provide a more comprehensive overview of the model for those who are not familiar.

Reply: We agree that a description of a more in-depth description of Mike SHE could be useful. This was also addressed by reviewer 1# and the editor. The section about the Mike SHE model was therefore re-written in the updated version of the manuscript.

Also a more detailed comment on this section: This description of the upper most soil layer thickness (line 175) is confusing and doesn't match with the table where it looks like this is only true in the case of till? Sandy sediments appear to have a thickness of 0.8 for the top layer? Can you clarify this?

Reply: We agree that this section was a bit confusing. Therefore, we simplified the table and changed the text in the method section of Mike SHE. We hope that these changes will make it clearer. As stated in the new version of the manuscript:

For the Krycklan model, the horizontal grid was set to 50*50 m. Vertically, the model is divided into ten calculation layers (CL) and extends to a depth of 100 m below ground. The CLs of the SZ vary with depth and are thinner closer to the soil surface; the first CLs extend to 2.5, 3, 4, and 5 meters below the ground surface, with the soil properties and depth extension following the stratigraphy (Table 3). The ET and UZ processes are only fully active in the uppermost SZ-CL, and here the ET and UZ are calculated at a finer resolution, leading to a detailed calculation of the groundwater table level. If the groundwater table falls below the first SZ-CL, a more simplistic method, not taking capillary rise and all ET-processes into account, was applied. The depth of the first SZ-CL was set to 2.5 m and was calibrated using the influence of the CL thickness on groundwater table level, UZ, and ET dynamics…