Revision notes of the manuscript

Kindly note that in the revision notes, comments from the reviewers are marked with "Comment", while our responses are marked with "Response".

Responses to the comments from reviewers:

Reviewer 2

Comments: Main comments
This research addressed the question of SRP release to surface water during baseflow conditions in different seasons (Winter and Summer). The research combines measurements of SRP in stream water sections, in streambed sediments, sediment pore water and groundwater, as well as tracers' tests and Rn measurements to localize and quantify GW inflow. This is important to better understand P delivery and release to surface water and the respective role of the different P sources.

I think that it is a manuscript with great potential but with a lot of information that need to be more organised to have a clearer story. There are quite a lot of grammatical issues to correct throughout the manuscript, a thorough proofread is needed. There are some details in the material and methods section to correct/add (measurement methods, uncertainty…). The discussion comes also very late and is too short, without many references to other studies. I would separate the results and discussion sections and would develop the later by discussing YOUR results and the processes that may be involved.
Response: We thank reviewer 2 for the positive comment. Regarding the concerns, please see our detailed point-to-point responses below. We will consider a thorough proofreading.

Comments: Technical comments
I would improve the resolution of Figure 1 and the y axes (colours when the scales of the 2 y axes are different, labels).
Response: We will choose a higher resolution for Figure 1 and use different colours for the two y-axes in Figure 2

Comment: I would add one figure with all the sampling/measurements points, that would make things much clearer.
Response: We will prepare a catchment figure which includes all sampling points within the study site

Comments: Check the references to figures and tables (all need to be cited in text and be the right one).
Response: We will check all references and correct them accordingly in the revised manuscript

Comment: No abbreviations ("SRP") or numbers ("52%") at the beginning of a sentence.
Response: We will consider this in the revised manuscript.

Comment: Superscript and subscripts to check.
Response: We will consider this in the revised manuscript.

Comment: There is a lack of consistency in the terminology (Summer/Winter, September/January).
Response: We will consider this in the revised manuscript.

Comments: Specific comments
Line 20: I think there is a mistake in the email address.
Abstract:
Line 30: SRP-fraction for DP? It is unclear.
Introduction:
Line 50: I would try to improve the transition between the temporal variability of SRP concentrations and the different P sources, it is too quick as it is now.
Lines 56-59: I would suggest separating in two sentences, it is a long sentence that may be difficult to follow and understand.
Lines 63-66: I would also here separate the sentence in two.
Lines 68-69: The sentence about temperature-dependent processes and its link with the redox conditions discussed above are hard to understand, I would clarify this. The transition to temperature-dependant processes would need to be improved.
Lines 71-75: The first part of this paragraph is hard to follow and understand due to grammatical errors, lack of clarity and organisation.
"In situations, …": I would not use this, but would go directly to the point.
"fed by baseflows": I would say "fed by groundwater" or "during baseflow conditions".
"Data suggest…": which data? Reference needed here.
Response: We will consider all this technical notes all through the revised manuscript.

Comments: Line 79: I am missing here a paragraph about geogenic sources of P, as it has nicely been done for the other P sources in the above paragraphs.
Response: We will add a short new paragraph about geogenic sources of P.

Comments: Line 88: I am not convinced using the expression "headwater baseflows", maybe instead "in headwaters during baseflow conditions?".
Response: We will consider this suggestion accordingly.

Comments: Material & Methods
Lines 110 and 111: I would not use "e.g." when referring to conductivity values, I would give a range or a mean value instead.
Line 112: I think "circa" is commonly used before dates, so I would delete it and just keep "below 0.4 m".
Response: We will consider this through the revised manuscript.

Comments: Line 113: Which "detailed topographic characteristics"? I would give more details here.
Response: These topographic characteristics are slope and slope position, valley bottom and exposition. We will add these to the revised manuscript.

Comments: Line 121: DPS is Degree of P Saturation.
Response: This is correct. We will clarify this in the revised manuscript

Comments: Line 124: I would specify "declines of WSP in the topsoil…".
Response: We agree with this suggestion and consider this in the revised manuscript.

Comments: Lines 135: We go from precipitation/Q (Figure 2) to GWL (Figure 4) without mentioning air temperature (Figure 3). Either the data presented in Figure 3 should be presented in the text or Figure 3 should be deleted.
Response: We agree with the reviewer. We will combine Figure 2 and 3 and will refer to the new figure 2 in the text.

Comments: Lines 148 and 149: Change "between X-X" to "between X and X" as before.
Response: We agree with this suggestion and consider this in the revised manuscript.

Comments: Lines 146-150: To which period(s) (e.g. 2010-2020?) do these values refer to?
Response: Yes, the water age analysis of Yang et al. 2021 is relate to the period of 2010-2020 but the earlier water quality investigations are related to the years from 1999 to 2010. This is also true for the cited study of Dupas et al. (2017). We will add this information to the revised manuscript.

Comments: Line 152: I would write "January 2019 during a period of…" instead of "January 2019 with..".
Response: We do not see the need to change this sentence and would therefore keep it as it is.

Comments: Line 153-155: It reads like dilution tests and 222Rn measurements were also used to characterise stream water, groundwater and sediments, which is not true. I would correct this; the sentence can then be used to organise the section.
Here we have: 1) in-stream tracer dilution tests and Rn222 measurements to analyse lateral inflows and 2) what measurements? to characterise stream water, groundwater and sediments properties.
Response: We will follow this comments to better clarify the measurement and their purpose as follow: These campaigns comprised in-stream salt tracer dilution tests and 222Rn measurements in order to analyze lateral inflows to the stream, and water quality measurements to characterize stream water, riparian groundwater and stream sediment properties.

Comments: I would slightly change the headings of the section 2.2. to improve its organisation:
2.2.1. Lateral inflows to the stream
2.2.1.1. Water balance of stream sections measured by tracer dilution tests
2.2.1.2. Groundwater discharge investigated by Radon measurements
2.2.2. Stream water, groundwater and sediments chemistry
Response: This is an excellent suggestion and makes the purpose of the measurements much clearer. We will consider this suggestion in the revised manuscript.

Comments: Lines 169-191: This is a very good section; the methodology is clear.
Response: Many thanks for this comment.

Comments: Lines 203-204: How do these 6 locations relate to the 6 locations used for the tracer dilution tests? Are they different? If so, why? How far are they from each other?
Response: The Rn and tracer dilution test locations are similar. We will adjust the text and also refer to the new figure of the sampling locations here. The locations can be considered to be essentially identical.

Comments: Lines 208-210: I would develop on the method on how to get from Rn data to the localisation of groundwater discharge and its quantification. How is the rate of radon degassing determined?
Response: Schubert et al. (2020), sect. 2.2.2, discusses comprehensively how the degassing rate was determined at the Schäferbach. In order to keep our manuscript as concise as possible, we prefer not to repeat this lengthy discussion but rather to cite the Schubert et al. paper. A few explanatory details and a related remark will be added to the text of our manuscript in sect. 2.2.2.

Comment: Lines 212-228: I would improve the organisation of this section, maybe follow: field instrumentation-field sampling-field measurement-lab analysis? Which method did you use for iron analysis?
Response: We would like to keep the structure of the presented measurements because this is in line with the structure of the presentation of the complete set of field measurements which is organized according to the purpose and not the method.

Response: Dissolved Fe concentrations were measured by ICP-OES (Perkin Elmer 7300 DV). The limit for determination was 0.01 mg L^{-1}. We include this information in the revised manuscript.

Comment: Line 215: Please specify which methods you used (and with references) for P analysis, this is important. Include the method detection limit or uncertainty.
Response: Total (TP) and reactive phosphorus (SRP) were measured using the ammonium molybdate spectrometric method (DIN EN ISO 6878, 2004). Detection limit is 0.005 mgPL^{-1}. We will add this information to the manuscript.

Comment: Lines 229-235: This part of the section is more organised and easier to understand.
Response: Thanks for the positive comment.

Comment: Results & Discussion
Line 240: Change "in…" to "during the two summer campaigns...".
Lines 241-242: I would not use "highly" and "very".
Line 241: "Constant" instead of "consistent"?
Lines 241-243: This is a long sentence; I would separate it in two sentences.
Lines 243-244: I would rather use a factor for comparison instead of a concentration.
Line 250: Should it be summer instead of autumn?
Line 256: I would specify in the heading "along the study reach of the stream".
Line 267: "proportion" not "proportionate".
Line 268: Do not start a sentence with number.
Line 270: The term "neutral" cannot be used here, I would rewrite the sentence.
Response: We agree with this technical suggestions and will consider them in the revised manuscript.

Comments: Lines 279-288: I would gather this part into a first paragraph presenting the results: longitudinal patterns and concentrations.
Lines 288-319: Then, in a second paragraph, I would discuss why we see these patterns and concentrations by bringing in the info on groundwater discharge localisation and rate. I think that would improve the organisation of this section.
Response: The reviewer might be right. However, we thought about the structure of sect. 3.3 quite intensely, too, and came up with the structure presented in the manuscript. First, we describe the recorded ^{222}Rn patterns qualitatively, and then compare and discuss the quantitative differences of winter and summer result. Subsequently we describe the resulting groundwater discharge rates, again first qualitatively then quantitatively. We believe that this is the most appropriate way to describe the situation. Hence, we would like to stick to the structure of the section as it is.

Comment: Lines 296-309: More information in the methods section on the FINIFLUX model would help to better understand the uncertainties related to the modelled results.
Response: The FINIFLUX model is described in all detail in Frei, S. and B.S. Gilfedder (2015). The uncertainties of the model results, which are mainly correlated to the uncertainties of the applied degassing coefficients, are discussed in Schubert et al. (2020). In order to keep our manuscript as concise as possible, we prefer not to repeat this lengthy introduction into the model but rather to cite the Frei, S. paper and B.S. Gilfedder as well as the Schubert at al. paper for the more interested reader. Still, a few explanatory sentences on the uncertainty of the degassing coefficient have been added to the manuscript in sect. 3.3.

Comment: Line 312: I think you already said that before (lines 287-288?), in a different way. I would avoid repeating the results.
Response: We agree with the reviewer

Comment: Line 321: Why are you investigating sources of SRP only in September 2020? Explain why, it is not clear for me.
Response: -14C and pore water measurements were motivated by the finding of stream reach balances obtained in summer 2019. Furthermore we had financial restrictions for the use of the financially always costly 14C measurements. We therefore selected the most extreme and ecologically relevant period for SRP transport from land to water during end of summer only.

Comment: Line 322-324: It is hard to locate these observation points, a figure showing all sampling/measurement points would help a lot.
Response: We will consider this point and revise the Figure with sampling points "technical comments"

Comment: Line 356: "are at work", please rewrite this sentence.
Response: We agree with the reviewer and revise the sentence in the revised manuscript.

Comment: Lines 351-364: There are some good things here but there is no references at all to support your points, and the discussion about the underlying processes is almost absent. The discussion needs to be developed.
Response: We will modify the sentence in line 355-356 to read and add: „Sediment pore water, … exhibited high concentrations of DP, Fe and NH4+ but low concentrations of NO3- …, suggesting that reductive conditions predominated. This indicates that sediment was a source of P to the stream, but that transport was quantitatively limited and probably dominated by diffusive processes." We further add more discussion and two references to line 360: „The young radiocarbon age of the stream DOC suggests that its source lies in shallow sediments rich in organic material such as riparian soils. In contrast, the presence of aged DOC in groundwater, here with a radiocarbon age of more than 2000 years is typical of low flow periods (Schiff et al. 1997, Tittel et al. 2022). The groundwater itself need not be that old, the DOC may have been dissolved and mobilized recently by hydrolysis of old organic carbon from the soil."

Comment: Lines 374-386: It is too much focused on presenting the data, and not enough on discussing them. I would discuss briefly how your results compare with long-term data ("our results are consistent with...") but discuss more about YOUR data and the processes explaining what you observed.
Line 374-375: I would rewrite this sentence, some grammatical issues there.
Line 385: Any reference to support the suggested dilution pattern?
Response: We see the point raised here and will add more linkages to the seasonal campaigns. The dilution pattern is an observation from the data here. We would add more discussion and also references to comparable dilution patterns in Germany and elsewhere after this section as the lines 374-386 are a description of results and not a discussion which follows from line 394.

Comment: Line 394: I feel like the REAL discussion starts here, so very late. I think separating the results and discussion would be beneficial.
Response: We disagree here. Yes, we state pure results in terms of concentration and fluxes. But later on (3.4 & 3.5) we do both together. Separating the results from discussion here would need a lot of repetition. However, we can make all this more clear in the chapter title: 1. Observed discharge and stream SRP concentrations. 2. Observed longitudinal water and SRP fluxes. With words like "assessing" and "integrating" in the last chapters we make pretty clear that we also discuss the results.

Comment: Line 414-426: I really like this part where you discuss your findings, related them to land use and soil type. Some references are missing when you refer to other studies in the same catchment.
Response: Many thanks for this positive comment and we will add related references to the discussion in the revised manuscript

Comment: Line 418: Please rewrite the beginning of the sentence, it does not seem right.
Response: We will change the beginning of the sentence to make the context more clearer

Comment: Figures/Tables:
Figure 1: This figure is hard to read, the resolution needs to be improved. In the legend, I would not use “soil types” since it does not refer to WRB soil types. Maybe hillslope or topographical position?
Response: We will revise the figure accordingly as already considered in the response of Reviewer 1.

Comments: Figures 2/3/4: I would gather the three figures and use panels. Units of the y axes should be in square brackets.
Response: We would like to show only a combination of figure 2 and 3 and put Figure 4 in the supplement. Regarding Figure 4 please see the next response.

Comments: Figure 4: I would show only the same period as in Figures 2 and 3 in the text, a longer time series can be shown in the Supplement maybe. Is there an issue with where the vertical lines are located? It does not look to be the same dates as in Figures 2 and 3.
Response: There was a mistake in the timing in Figure 4. We will correct that and put Figure 4 in the supplement but referring to the groundwater levels of the campaigns in the text as follows: "The groundwater levels [m below surface] in the sampling period ranged between 0.5 and 1.1 m. Relating to the 3 sampling campaigns groundwater levels were at 0.65 m (Jan. 2019), 1.0 m (Sep. 2019) and 0.85 m (Sep. 2020) below surface."

Comments: Figure 5: Maybe consider colouring the y axes (blue/red) so it is easier to see that they have different scales.
Response: We will adjust the color of the axes as suggested.

Comments: Figure 6: In the caption I would use “during the three sampling campaigns” and not “in”. Should it be “SRP net flux” instead of “SRP net” in the y axis title? Maybe consider colouring the y axes (blue/red) so it is easier to see that they have different scales between January and September. I would also change the order of the panels since Q net is calculated from Q and SRP net flux is calculated from SRP flux: Q, Q net, SRP flux, SRP net flux.
Response: We agree and will adjust the figure captions, order and axes. See also our response to reviewer 1 on the X-axis.

Comments: Figure 7: I would add a sentence explaining the dashed line in the caption (even though it is already in the text), so the reader does not have to look for it in the text. In the y axes titles, you use here winter/summer but in Figures 5 and 6 you use January/September, I would stick to one of them and not mix the two, be consistent. I would also change the colours to red and blue to be consistent with the other figures and I would colour the y axes (blue/red).
Response: We color-code the y-axes in Fig. 7, as suggested. Furthermore, the information “January 2019” and “September 2020” will be added to the figure captions of Figure 7 and 8. Even though the meaning of the dashed line (Fig. 7) is explained in the text above Fig. 7, we add the information to the figure captions of Figure 7 and 8, too, as suggested by the reviewer.

Comment: Figure 8: What is the uncertainty of the modelled Rn concentrations?
Response: The uncertainly of the result is mainly a function of the assumed radon degassing rate. The related FINIFLUX parameters have been chosen carefully and are reasonable. Still, a detailed

discussion of the FINIFLUX model (including error propagation related to the degassing coefficient) would be beyond the scope of the study presented here. Tor the more interested reader the FINIFLUX model is described in detail in Frei, S. and B.S. Gilfedder (2015). Uncertainty and error propagation related to the FINIFLUX degassing coefficient is discussed in Schubert at al. (2020). Still, a few explanatory sentences on the uncertainty of the degassing coefficient will be added to the manuscript in sect. 3.3.

Comment: Figure 9 : Why are there only 3 points here? And not all the measured concentrations points? Is it the average of each campaign? I am surely missing something here, I would clarify.
Response: Because all historical sampling data have been carried out at the outlet of the headwater catchment in Figure 9 we used only the samplings at the catchment outlet of the three campaigns. Due to nutrient processing within the stream reach we did not use upstream nutrient measurements from our campaigns for this comparison.

Comment: Table 2: "nd" refers to "not determined" or "not detected"? If it refers to not detected I would say "< MDL" instead. These MDL need to be given in the method section.
Response: n.d. actually means not determined, but I have mixed this with a non-analyzed value from another sample. The correct value is 0.027 mg L-1 , which we enter in the table. Thanks for the hint.

Comment: Table 3: Where do these data come from? Any reference?
Response: The origin of the data is described in the text above. Surface and drainage water samples are taken from Ollesch (2008) while groundwater data came from an internal campaign. We can add this information to the caption.

New References

Schiff SL, Aravena R, Trumbore SE, Hinton MJ, Elgood R, Dillon PJ (1997) Export of DOC from forested catchments on the Precambrian Shield of Central Ontario: Clues from 13C and 14C. Biogeochemistry 36:43–65.

Tittel, J., Musolff, A., Rinke, K., Büttner, O. 2022. Anthropogenic transformation disconnects a lowland river from contemporary carbon stores in its catchment. Ecosystems 25: 618-632.