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:

General comments

I commend the authors for their revision of this manuscript. The study is valuable for the literature on the interaction between hydrology and P biogeochemistry, and deserves to be communicated as effectively as possible.
The manuscript is much clearer throughout. Here are a few remaining points to be ultra-clear about:

Comment:
In the discussion on Rn degassing, its impacts on GW measurements, etc. – shouldn't the estimate of k in this study be reported? Pointing out that prior work (e.g. Raymond et al. 2012) uses much bigger streams is fair, but hopefully the ranges for k on L334-335 are still plausible. What k did FINIFLUX estimate? Note also that L234 calls it k_Rn but k is used elsewhere.

Response:
In the revised manuscript we describe the challenges related to the quantification of radon degassing (i.e., the k value) based on empirical equations in more detail. The k values from our FINIFLUX model calculations are now reported in an additional paragraph. As described in detail in Frei and Gilfedder (2015) and Schubert et al. (2020) (both cited in the "methods" section of the paper). FINIFLUX is using the empirical equation from O'Connor and Dobbins (1958) (O'Connor, D.J., Dobbins, W.E.: Mechanism of reaeration in natural streams, Transactions of the American Society of Civil Engineers, 123, 641-666, 1958.). As suggested by the reviewer, the k values resulting from FINIFLUX are compared to k values resulting from alternative empirical equations (Raymond et al., 2012). The differences in the results are discussed.
"k_Rn" in L234 was changed to "k".

Comment:
The new text added around the summer radon / GW modeling results (L344-358) certainly raises questions about the FINIFLUX results (primarily, how the sensitivity analyses give lower discharges than what was calibrated). However the authors discuss the challenges involved. (This could perhaps serve as a good case study for future development of FINIFLUX.) Just some food for thought: could uncertainties in the stream discharge measurements (via salt injections) also be involved? Is it possible that some sections were losing water to GW?

Response:
We did not investigate uncertainties in discharge associated with salt injections but if these uncertainties would be available they could in principle be considered using the FINIFLUX

model. As shown in Fig. 4 especially the January campaign shows clear losing conditions in the most downstream section.

Comment:
The GW P fluxes compared to stream P flux (L412-414) is a nice addition to the text. I would caution that the relatively wide range in uncertainty for the GW P fluxes leaves room for interpretation. Assuming the true GW P flux is contained within that bracket, what if the summer GW P flux was 4 mg s-1? Or 64 mg s-1? (Roughly 4-fold less and 4-fold more, respectively, of the summer stream P flux.)

Response:
The range of calculated GW-P flux is based on the variability of SRP concentrations in individual groundwater wells. This range represents the maximum range if we assume that the GW P flux can be calculated using the lowest or highest concentration in a well. Because the GW P flux is most likely a mixture of different GW SRP concentrations, GW P fluxes calculated with concentrations near the mean GW SRP concentrations are much more likely. Therefore, in our view, the reported maximum ranges of GW P fluxes represent rather theoretical maximum ranges. The P fluxes at the catchment outlet agree quite well with those based on mean GW SRP concentrations.

Comment:
The Discussion could use some revision for efficiency and clarity. Some examples: the sentence beginning on L423 reads like a repeat of L417-418. The first half of section 4.2 (L451-488) could be condensed/simplified. I don't follow how the "findings confirm… that SRP concentration in the saturated zone is controlled by sorption equilibrium under oxidizing conditions in the upper groundwater" (L517-518).

Response:
We agree with the comment and shortened the discussion accordingly and took the sentence from L423 out. For setting our results in a broader context of earlier investigations we think that it is necessary to present those findings in the suggested detail and we therefore would like to keep this discussion mostly as it is. Nevertheless we shortened this sections slightly (L486).
We rewrote the sentence (L517-518) as follows: Our findings suggest that SRP concentrations in the saturated zone is controlled by oxidizing conditions in the upper groundwater, and that SRP losses through seepage are largely buffered.

Comment:
L538 suggests sediment P fluxes were quantified; I don't believe they were, but rather suggested to likely be low. Please clarify.

Response:
We rewrote the sentence as follows:
Because of the limited quantitative P exchange between stream sediments and stream water, stream sediments that may have originated from agricultural soils eroded into the stream did not contribute significantly to the SRP loads exported from the stream.

Comment:

I suggest reading through the manuscript once more with fresh eyes for small fixes. Below are a few things I noticed.

DOC needs to be defined in the abstract.
Suggest defining P on L42.
L208: ‘was’ instead of ‘were’
L216: “…stream is provided in”
L336: ‘ran’ instead ‘run’
L380: DOP
L421: “…SRP-concentrations, TDP, …” – odd wording
L438: remove the “that”
L505: van Dael, not van Deal
L515: Clarify the 0.05 mg/L reference (mean/median?)
Conclusions should now be Section 5
Some references are repeated.

Response:

All done with the exception of L42 and the manuscript has been checked again.