Instytut Geografii i Przestrzennego Zagospodarowania
Polska Akademia Nauk
Zakład Badań Geośrodowiska

Institute of Geography and Spatial Organization
Polish Academy of Sciences
Department of Geoenvironmental Research

31-018 Kraków, ul. św. Jana 22; tel./fax +48 12 422-40-85; e-mail: zbg@zg.pan.krakow.pl

Kraków, 23.04.2026

Dear Editor,

Thank you for your decision, which allows us to revise the manuscript '**The effect of beaver dams on the physico-chemical properties of stream water in the Carpathians (Poland and Slovakia)**' for publication in Biogeosciences. We greatly value all the comments we have received, particularly those that have highlighted the weaknesses of our previous statistical approach. All sections of the manuscript have been improved, following your helpful suggestions. Please see our responses and detailed explanations of the subsequent modifications (marked in red) below:

General comments:

More specifically, you describe usage of U-tests or Kruskal-Wallis Anovas. Both tests accept only one factor as structuring your dataset. You either use pond age (3 levels), pond type (2) or season (4? or 2?) as factors in your analyses. You do this by pooling data across all levels of factors not targeted in any single test. This strategy is in partly just statistically weak, in partly, however, also plainly wrong. Your study design is, in fact, a two-way factorial design (with two treatment or fixed factors pond age and pond type) and one additional within-subject factor that codes for seasons. The latter is because the SAME ponds were repeatedly sampled, in earlier times this would have been called a "repeated measures ANOVA", nowadays such designs are handled by random factors. To account for the repeated measuring in the same ponds is necessary because otherwise by pooling across the 4 seasons you pseudoreplicate your dataset. The sample size for your two treatment factors must be 9 and not 9*4. When you ignore the repeated measures nature of your dataset, you run a U-test assuming a sample size of 9*4, i.e. you actually work with strongly inflated degrees of freedom. So, while ignoring pond type in a 1-way analysis targeting pond age may just mean a statistically weak strategy (the variance caused by pond type will increase apparent measurement noise), the ignorance of seasons being a within-subjects factor is statistically incorrect. Proper treatment of seasons will also make an analysis more powerful as the seasonal effect can be identified given the paired nature of your dataset. To illustrate, the problem I describe here is the same issue as you try to solve by using Dc to get rid of catchment-caused differences in upstream-pond-downstream data. You could also think of a factor coding for location along any river (with 3 levels up/pond/down). This factor is a within-subject factor requiring a random factor river in a proper statistical design. I am fine with your solution of using Dc, however, these thoughts just for illustration of the problem. Please adapt your statistics. Describe clearly why you decided to run a 1-way design instead of a two-factorial design and that this means pooling of data across levels of factors not targeted in any particular test. Much cleaner would be to actually run a 2-way analysis. This is not possible using ordinal scaled data (i.e. using a test like Kruskal-Wallis), but will require more sophisticated GLM-type models. Going this way would also allow you to include the seasons factor in a correct way. From my experience, STATISTICA should be able to handle this.

Answer: We sincerely thank you for this important and constructive comment. We agree that our previous statistical approach, based on separate one-factor non-parametric tests, did not adequately reflect the repeated-measures structure of the dataset and could lead to pseudoreplication. In response to this concern, we completely revised the statistical analysis. We abandoned the previous 1-way testing strategy and re-analysed the data using a Linear Mixed Models (LMM). In the revised analyses, pond age, pond type, season, and sampling location were treated as fixed effects, whereas stream identity was included as a random effect. This allowed us to account for repeated measurements collected within the same stream and thus to better represent the dependence structure of the dataset. For factors showing a significant overall effect, least-squares means (LS means) were examined, and pairwise comparisons were performed to identify specific differences between factor levels. The Materials and Methods section has been changed accordingly, and the Results section and relevant tables have been updated to present the outcomes of the LMM analyses. We are grateful for this comment, as it substantially improved the statistical treatment of our dataset and the overall quality of the manuscript.

I also urge you to invest a bit into briefer text expressions. The paper still has some potential for shortening without compromising content.

Answer: We appreciate this suggestion. In response, we revised the manuscript to make the text more concise and precise. Several passages were shortened or rephrased, and unnecessary wording was removed, while preserving the full scientific content of the paper.

Minor line-by-line comments:

13: better “pond types differing in flooding extent”, so you already establish the word TYPE here.

Answer: We appreciate the suggestion. While we have introduced the term "type" in the Methods section of the abstract (as suggested in your subsequent comment), we would prefer to keep the original wording in this specific line to emphasize that flooding extent was our primary variable of interest. However, we have ensured that the transition to "pond types" later in the text is clear and consistent.

17: types instead of categories

Answer: We applied this change.

17/25/27: please make sure to always use the same words when addressing the two levels of the factor pond type to avoid confusion.

Answer: We have standardized the terminology used to describe the two levels of the 'pond type' factor throughout the manuscript

18: make clear that age has 3 CATEGORIES

Answer: This sentence has been changed. We have clarified that there are three age categories.

21: processes instead of activity. It makes no sense to phrase “more statistically significant”. In a statistical sense the word “significant” cannot be intensified. You may use other qualifieers like “marked”, “pronounced”, etc.

Answer: The sentence has been modified.

26: when addressing a process like denitrification (or any other) just name the process, don´t call it “the denitrification process”. So here, this has to be “…which suggests more intense denitrification”. This applies to many instances throughout the MS, also for other processes.

Answer: We have removed the word 'process' when referring to named biogeochemical or physical processes. This correction has been applied throughout the manuscript.

31: Delete this last sentence, lame and meaningless. Close with an implication of your findings or some statement on the importance of your results even if limited.

Answer: The last sentence has been removed. We have focused only on the most important outcomes of the study in the revised abstract.

34: Castor needs italics

Answer: It has been changed.

45: the population statement is valid for where? Global?

Answer: We have modified this sentence. It now indicates that the population increase occurred generally in Europe and Asia. We also used the second part of the sentence to introduce the study area, and link it to the rest of the paragraph, particularly after removal of the sentence indicated in the following comment, which was previously meant to introduce the study area.

48-49: Cannot see relevance of this sentence at this position, maybe move to (end of) next paragraph.

Answer: We have removed this sentence entirely

76-77: source and sink for what? Changes and transitions = redundant, reword. Cool to bring up self-purification, but you may want to also point to gases CO2, CH4, N2O, which build up as mineralization products in beaver ponds.

Answer: We have revised the sentence to clarify that ponds act as sources or sinks specifically for nutrients. To avoid redundancy, we simplified the phrasing by focusing on seasonal variability instead of using 'changes and transitions' together. Additionally, we have included a statement regarding the role of beaver ponds as

hotspots for greenhouse gas emissions (CO_2, CH_4, and N_2O), providing a more balanced view of their impact on the ecosystem.

82: delete “complex”.

Answer: We deleted ‘complex’

83: bother to give an example for these contradictory patterns?

Answer: We have added an example of the effect of beaver dams on downstream phosphate concentrations. We have also updated the reference in the general sentence. Initially, we selected the works of Majerova et al. (2020) and Weber et al.(2017) because they illustrated different temperature patterns, but we have now changed these references to meta-analyses, which e.g. highlights the effects of beaver dams on the given element (Ecke et al., 2017).

91: find briefer expressions, here is a good example. The whole phrase “especially….could” can be replaced by “may”.

Answer: We have simplified the sentence as suggested, replacing the wordy phrase with 'may'.

160ff and Fig 1b): I cannot see the added benefit of this figure. No chance to see how the valley actually drives a particular pond type. If you want to make this useful, then show cross-sections at a much smaller scale, so that readers can see in-channel vs. overflowing ponds.

Answer: Following improvements to our methodological approach, the pond type was deemed to be insignificant. We have therefore removed this part of the figure, and the sentence referring to it, to eliminate distracting elements and focus the manuscript on the age factor.

164: replace “the archival” by “archived”.

Answer: As suggested, we have replaced the word 'archival' with 'archived' in the description of the orthophotos.

Fig 2: Decide whether you want to use the word sequence or cascade in concert with ponds, but not both. In general use this less often, also in abstract, just confusing.

Answer: We have changed the wording in the figure to ‘sequence’ to align with the text. We have also corrected some instances where ‘sequence of multiple ponds’ was previously used instead of ‘pond sequence’.

191+196: make the order in which these ponds pop up in the text agree with the figure.

Answer: We matched both orders by changing the order in which the ponds appear in two sentences, and the order in which they appear in Figure 2c.

210: sulphate and sulfide.

Answer: We have verified the nomenclature. As noted in the manuscript, sodium sulphite (Na_2SO_3) was indeed used for the zero-oxygen verification, as it is the correct reagent for removing dissolved oxygen. In the manuscript instead of the full name we used a chemical formula.

215-225: The detection limits are fairly high in comparison to the used standards. In essence, of your 6 standards, only the highest one was relevant, all others were below detection. This meanse 1-point calibrations for all parameters. Please clarify. Use L-1 instead of dm-3.

Answer: Thank you very much for drawing our attention to this sentence. Such calibrations (0, 1, 10, 50, 100 and 1000 $\mu g \cdot dm^{-3}$) were performed for trace element measurements taken from samples collected at the same study sites. These are not analysed in this paper, so an inclusion of this part was our oversight. For ion chromatography, multi-point calibration curves were established based on the specific concentration ranges of the standards. For anions, a 6-point calibration curve was used: 0 (blank), 0.2, 0.5, 1.0, 2.0, and 5.0 $mg \cdot L^{-1}$. For cations, the ranges were adjusted to the standard mix composition: Na^+: 0, 0.4, 1.0, 2.0, 4.0, 10.0 $mg \cdot L^{-1}$; NH_4^+: 0, 0.5, 1.25, 2.5, 5.0. 12.5 $mg \cdot L^{-1}$; K^+ and Ca^{2+}: 0, 1.0, 2.5, 5.0, 10.0 to 25 $mg \cdot L^{-1}$. These ranges ensured that the calibration points remained significantly above the detection limits (0.4 $mg \cdot L^{-1}$ for Ca and 0.1 $mg \cdot L^{-1}$ for other ions), providing reliable multi-point quantification for all parameters. We also found that the order of the information in this paragraph was potentially confusing, so we restructured it.

234-235 and eq2: The MEAN absolute difference should be denoted with a horizontal bar on top.

Answer: We corrected this.

238-244: This is not understandable to me.

Answer: We have tried to make this paragraph more clear by rephrasing and improving the consistency in the use of *Dc* and $|\overline{Dc}|$.

250: There was no research question for this particular analysis. At least provide a motivation for doing this here at this location, not necessarily in the intro.

Answer: We have modified this section by adding brief reasoning to these analyses.

253-254: Not understandable to me. If a result is significant, then it is significant and it does not matter whether P is 0.04 or 0.0001 – you cannot make an extra choice step. Also, you here bring up a potential confounding between the factors age and type. This confounding should be made clear in the design section. It may qualify as reasoning for 1-way instead of 2-way designs.

Answer: Following all the comments regarding the statistical approach we gave up the previous design. We have now used only p>0.05 as the significance threshold, which is stated in the descriptions of each method (both LMM and correlation). Moreover, while presenting additional plots, we have focused on the only one factor, which turned out to be significant, namely pond age, alongside the seasonal variation.

We have now stated clearly in the Methods that there was some overlap between the two categorisations, as most of the youngest study sites were characterised by in-channel pond type and all the oldest ponds were overflowing.

254-259: This section is hard to understand. Make clear how many comparisons you assumed for your Bonferroni correction. The logic behind what the last sentence here describes is really obscure to me.

Answer: We have removed this section because it is no longer applicable following the changes to the methodological approach.

262: what are “baseline conditions”.

Answer: We meant reference (upstream) conditions here. We have now clearly defined it as ‘water upstream to beaver ponds’

266-267: P-values for these correlations?

Answer: We specified in the Methods that the $p > 0.05$ was considered statistically significant and only correlations $r > 0.6$ were considered high. Only the results which met those criteria were mentioned in this paragraph. However, to avoid lack of clarity, while revising the Methods section, we specified the *p* threshold for each method separately (both LMM and correlations). We replaced ‘high’ with ‘strong’ in the indicated section to match the description used in the Methods.

Fig 3: use a color gradient that reflects the rank order of pond age levels, e.g. red/orange/yellow.

Answer: We agree that the red-orange-yellow seems to be reasonable as it generally represents increase/decrease (e.g. in age) very well. However, including blue alongside the warm colors makes this figure more color-blind-friendly. That is why we would prefer to keep the current set of colors. Unfortunately, due to the size of the plots and number of elements presented on the figure, we could not resolve this issue differently, by change in signature shapes. Nevertheless, we have tried to achieve the ‘warming’ effect for the increase in age by reversing the order in which the colors are assigned.

286: Does this statement refer to absolute Dc values or MEAN absolute Dc values?

Answer: We have shortened this sentence and used only one variant of the change (*Dc*), without referring to $|\overline{Dc}|$, to avoid such confusion.

Fig 4: The asterisks in the panel headers should sshow “seasonal difference” – what is this exactly referring to?

Answer: We have removed this part of the analysis and focused on the seasonal differences obtained from the LMM instead, since these seem to be more reliable as they take into account more factors simultaneously.

Table 3: Note that Kruskal-Wallis and U-tests cannot deal with dependence in the data set (repeated measures). Make clear how many repeated data usages you actually accounted for in the Bonferroni correction. As far as I can see, the same data was used many times, especially upstream data of spring and summer. It´s not trivial to here decide about an appropriate Bonferroni correction.

Answer: Due to statistical issues related to repeated measurements at the same study sites, we refined our approach by using a method that accounts for the dependence between samples taken from the same river. Consequently, this table has been replaced with the results of LMM.

309: do not intensify “significant”.

Answer: In this case ‘more’ was supposed to indicate number and not magnitude. However, we appreciate the comment, since we previously used such a phrase with the indicated intention in the abstract, which indeed qualifies for an error. We avoided using this phrase in the revised manuscript.

319-321: clearly describe confounding (“overlap”) of factors much earlier in the MS (methods). I cannot follow the reasoning of this sentence: How can identifying a pH effect be driven by confounding?

Answer: We have described the overlap at the beginning of the revised section 2.3. The paragraph indicated in the comment has been removed as a result of a change in the methodological approach.

Fig 7: This plot (also in Appendix) must be done differently, it´s confusing, hard to read and also wrong. Why not computing averages from the two profiles? There is no justification to show single data points. Also the plot implies pairing of measurements 1-2-3 with 9-10-11 but that is simply not correct. 1-2-3 may also be opposed to 12-13-14, right?

Answer: Thank you for this suggestion. We agree that the previous figure design unnecessarily suggested pairing of only two profiles at a time. As the profiles were located close to each other and the samples were taken from the same levels of the water column, there is indeed no need to show them separately. We have presented the average values from the two profiles in the revised figures.

354: Please clean up the “deposition” argument. Certainly beaver ponds will be depositional habitats where PARTICULATE organic matter will accumulate by sinking to the bottom or getting filtered by dam and hyporheic interfacts. However you measure TOC from a water sample, likely without having disturbed the bottom, so this can only include small suspended particles.

Answer: Yes, we agree with your comment. Our findings regarding higher TOC concentrations in the pond water than in the streams above do not indicate deposition of organic matter in the ponds, but rather that the ponds' water is enriched with organic matter. We have revised this statement of the text.

359: downwelling “under” (?) the dam?

Answer: We have corrected this sentence.

365: delete “a specific”

Answer: We removed this word.

366: delete “a relatively”

Answer: We removed this word.

373-375: Please revisit nitrogen-transforming processes. Standard denitrification does not produce NH4+, this is a different process called DNRA = dissimilatory nitrate reduction to ammonia. The current text ignores textbook knowledge.

Answer: Thank you for this important comment. We agree that standard denitrification does not lead to NH_4^+ production and that this process should be distinguished from DNRA (dissimilatory nitrate reduction to ammonium). In response, we have revised the manuscript to clarify the interpretation of nitrogen-transforming processes and to explicitly include DNRA as a potential pathway contributing to NH_4^+ formation under reducing conditions in beaver ponds. The manuscript (Discussion section) has been corrected accordingly.

370-378: When coming up with denitrification, it would be good to also bring up N2 and N2O as gaseous products. This creates a link to gaseous mineralization products to be explained in the intro.

Answer: We took this suggestion into account. We have added information about the gaseous products of denitrification to the text.

386-387: I don´t understand the Margolis statement. A reduction of SO4 will increase pH but hardly ANC itself.

Answer: The decrease in NO_3^- and SO_4^{2-} concentrations is one of the main contributors to the acid-neutralising capacity (ANC) of beaver ponds, as it results in an increased HCO_3^- ratio within anions (anion-cation balance). However, this would indeed require a broader explanation. That is why we removed that sentence.

380-395: You may want to actually include increase CO2 (mineralization product) as a driver of pH changes. Outgassing in the pond will be slower than up- or downstream.

Answer: Thank you very much for that great suggestion. We agree that CO_2 produced during the mineralisation of organic matter can lower the pH. However, we also believe that CO_2 may be released into the atmosphere due to elevated water temperatures in the pond, even in the absence of turbulent flow. Finally, we added a statement to the Discussion section about how CO_2 production may affect the pH of water.

404-406: “over” the top of the dams?

Answer: We have revised this sentence as suggested.

420: consider CO2 emissions.

Answer: In accordance with your suggestion, we have incorporated the potential role of CO_2 emissions into the interpretation of the observed pH increase below the beaver dams. The relevant text has been revised accordingly.

429: statement only makes sense for ice-cover – were the ponds covered by ice?

Answer: This sentence has been removed.

438: delete “one of”.

Answer: We removed ‘one of’ as suggested.

442: Meaning of “can be partly related to” is unclear to me.

Answer: We have modified this sentence for clarity.

447: logic of “can occur despite” unclear to me.

Answer: We modified this sentence to be more clear.

Fig 8 Use “season” instead of period in headline, get rid of “geomorphic”.

Answer: We have removed the pond type completely from the figure since an improved statistical approach did not indicate its significant effect. In the revised figure, we have reduced the information presented to only the patterns associated with pond age. In the case of a seasonal effect, we have stated the seasons precisely (spring and summer).

470: “significance… lower” is not a statistically meaningful statement.

Answer: This sentence has been removed entirely following modifications to the statistical approach and its results.

497-505: Upstream is your reference, so I would call that the “control”. So, what you have is a control-impact design.

Answer: Thank you. We modified this paragraph.

508: Any evidence shown for this statement?

Answer: We have revised parts of the manuscript related to this issue. We clarified the description in the Methods: ‘Moreover, the correlations between the upstream values and *Dc* were analysed, as this could indicate potential unreliability of the obtained percentage changes.’ and in the Results: ‘There were no strong correlations between the percentage change (*Dc*) in the physico-chemical parameters and the parameters above beaver ponds (Table A2)’ since those statements were previously imprecise. We have also supplemented the last sentence with additional data presented in the appendix.

515: why “most of”? there were only 2, right?

Answer: This statement was meant to refer to the profiles obtained for most of the parameters. We have corrected this sentence.

524-526: Unclear statement. Do you refer to the confounding issue?

Answer: We have replaced the 'length of beaver occupancy' with 'age of beaver pond sequences' and rephrased this sentence to make it more clear.

528. Last sentence of this paragraph no clear to me.

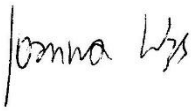
Answer: We have modified this paragraph to make it more clear and focused.

Table A1: abs(Dc) was described as being a MEAN, so how can there be a SD for i? What is shown in this table seems to not agree with the description in Methods.

Answer: Thank you for pointing out the inconsistency in the use of |Dc| symbol. We applied the use of $|\bar{Dc}|$ (with horizontal bar) in the revised manuscript, to indicate an average value (the one used in the equation). In this case, we referred to 'plain' |Dc| (just absolute value of *Dc*). We have made adjustments to the table caption and its header.

On behalf of all authors of the submitted manuscript, sincerely

Joanna Wąs



corresponding author: joanna.was@zg.pan.krakow.pl