21-06-2022

# Response to Decision on manuscript hess-2021-401

Hydrology and riparian forests drive carbon and nitrogen supply and DOC:$NO_3^-$ stoichiometry along a headwater Mediterranean stream

**Editor comments to the authors**

*Dear authors,*

*Two referees have evaluated your revised manuscript. One found that the small number of events included in your analysis is a major weakness but still recommended acceptance. Another reviewer recommended major revisions, and they questioned the exclusion of the medium-sized storms (that is leading to a small number of events considered in the end). Given that both reviewers are raising the same concern, here, I am returning your manuscript for further revision. One of the reviewers suggested that you either need to better articulate why your current approach is novel or adjust your analysis/sample size. I agree with these suggestions, which will only make your manuscript stronger.*

*With best wishes,*

*Genevieve*

**[Reply]**:

Dear Editor,

Thank you once more for handling our manuscript and returning it for further revisions. First, we would like to thank Rémi Dupas for the time evaluating our manuscript and for the recommendation to accept it this time around. Given that there is still one remaining concern, which was common in the former review round and it is still shared by both reviewers, we realized that it was necessary to include the medium storm events more explicitly in our analyses and conceptual model in order to strengthen our study. In short, we now have (i) included a second PLS regression model with all (large and medium) storm events, (ii) added the medium storms in Table 2, (iii) added a third panel in Figure 8 describing the medium storms in our conceptual model of DOC and $NO_3^-$ mobilization during storm events, and (iv) revised materials and methods, results, and discussion accordingly. All in all, we are satisfied with the revisions, which, as you anticipated, have made our manuscript stronger.

Please, find below our detailed response to reviewer #2, including descriptions of changes made to the previous version of the manuscript. Note that the text in *italics refers to literal comments by the reviewer*, the text in blue contains literal quotations from the revised version of the manuscript, and line numbers refer to those in the revised, clean version.

With thanks in advance for your effort handling our manuscript,

José L. J. Ledesma and co-authors

**Reviewer #2**

*Major issue:*

*I'm struggling with the exclusion of the medium-sized storms from the PLSR. And I think the exclusion masks the actual novelty of their (small) dataset. First, the authors classify storm size based on the combination of precipitation totals and GW table response, with an emphasis on the latter, as storms in the total precipitation range of 67-98 mm were classified as either medium or large based on whether they produced a large change in the GW table. The authors then proceed with their PLSR analysis with the 5 storms that produced a substantial GW table response AND include change in GW table as a predictor in their PLSR models, which seems like circular logic and tautological. Further, their subsequent conceptual model, based on this analysis, does not build on our current understanding of how DOC and NO3- are transported from riparian areas to streams during large precipitation events during dry vs. wet watershed conditions. It seems like the authors are aware of this because they site a few key papers, which all suggest that you will see large DOC and NO3- transport if there is enough precipitation when the system is transport-limited (buildup of solutes during dry conditions) vs. when the system is more source-limited during wetter conditions.*

*I think the authors are missing an opportunity in their analysis, though I think it might be difficult with the number of events they have (though they proceeded with 5 events previously). The actual novelty in their dataset is about the interaction between precipitation totals and antecedent watershed dryness/wetness, especially in that 67-98 mm event precipitation range. Specifically, what is the threshold in antecedent watershed dryness/wetness as quantified using P-7/30 when storms actually produce a GW table response and export more DOC and NO3-? We don't typically classify storm size based on the GW table response, we measure and quantify storm size based on precipitation totals. And this is how we make predictions about how storm size influences stream biogeochemistry.*

*I wish I had the analytical solution for including the interaction between total precipitation and P-7/30 and/or finding the threshold in P-7/30 that produces at GW table response, but I don't. Maybe the authors can include interactions in the PLSR? Regardless, for this portion of the manuscript to build on our understanding of DOC and NO3- export during storms, the authors need to include the data from medium-sized storms and adjust their conceptual model to reflect this. There is novelty there, the authors just need to think a little more about this final analysis.*

**[Reply]**: We thank the reviewer for their detailed evaluation of our manuscript and for their perseverance in requesting further clarifications/interpretations in the relation to the storm event analysis. First, we must note that in our replies to the reviewers' comments that were posted in the online Interactive Discussion we included a PLS regression model that contained all (large and medium) storm events. We used that figure to illustrate the different nature of large *versus* medium storm events in relation to DOC and $NO_3^-$ mobilization and to justify the exclusion of the medium storm events from the analyses and conceptual model. However, in order to keep it simple, we did not include this figure in our former response letter, which would have helped to clarify our case. At any rate, we now realize that this PLS regression analysis and subsequent interpretations not only should be included in the response letter, but also in the manuscript. Therefore, the present revised version includes results from this PLS regression analysis and further information of both large and medium storm events characterized in this study.

<u>Reply to the general concern and description of changes made in the manuscript</u>

In general, the essence of the reviewer comment lies on the fact that we excluded the medium storm events from our PLS regression analysis and from further interpretation in our conceptual model. We have explicitly addressed this concern by including in the manuscript a PLS regression model where all 16 storm events (5 large and 11 medium) act as observations, the nine hydroclimatic descriptors as predictors, and stream DOC and $NO_3^-$ concentrations as response variables. The inclusion of this new PLS regression model have, in fact, helped to make our case stronger in a number of ways. First, in the PLS ordination biplot we observed two clear clusters falling in opposite sides across the first component. These two clusters precisely correspond to large and medium storm events, respectively. Moreover, the hydroclimatic descriptors considered important in this model (VIP>1) were precipitation amount (*P*), groundwater table range ($\Delta Gw$), and average groundwater table ($Gw_{avg}$). These results provide further detail supporting our former classification based on storm size and groundwater table response and the idea that large and medium storm events display a distinct mechanism of DOC and $NO_3^-$ mobilization. Additionally, these results highlight that storm size and associated groundwater table responses provide a first order control (in terms of relevance) on the mobilization of DOC and $NO_3^-$ from the riparian zone to the streams during storm events.

Furthermore, the nine hydroclimatic descriptors showed small variability among the medium storm events, but varied widely among the five large storm events. For this reason, we performed a second PLS regression analysis using only the subset of large storm events (i.e., our original PLS regression model) in order to gain further insights into the hydroclimatic characteristics that most effectively mobilize DOC and $NO_3^-$ during these larger events. Indeed, the results from this second PLS indicate that during large storm events the mobilization of DOC and $NO_3^-$ depends on antecedent soil moisture conditions and the relationship between riparian groundwater tables and stream flow (important descriptors in this second model were *P-30*, $Q_{avg}$, *slope*, and $Q_{avg}/Gw_{avg}$), as explained in the previous versions of the manuscript. In light of these results, we conclude that antecedent soil moisture conditions and the relationship between groundwater tables and stream flow play a minor role during medium storm events in the mobilization of DOC and $NO_3^-$ from the riparian zone, but are essential for understanding such mobilization during large storm events.

We have integrated all these results in the revised version of the manuscript. We have added a new panel in Figure 6 (now Figure 5) showing the PLS ordination biplot of the whole dataset (large + medium storms) together with the former PLS ordination biplot including only the five large storm events. Former Table A1 is now Table 2, where hydroclimatic descriptors and chemical characteristics are shown for both large and medium storm events. Figure 8 conceptualizing the mobilization of DOC and $NO_3^-$ from the riparian zone to the stream during storm events now includes an additional panel for the conceptualization of medium storms.

We have simplified the material and methods to clarify that we use PLS regression models that consider the whole storm events dataset (Lines 183-184: "We then related those descriptors with the observed stream DOC and $NO_3^-$ concentrations using partial least square (PLS) regression models").

We have now extensively described the results of the new PLS regression model in the results section, as well as presented the classification between large and medium storm events in this part using the results from the new PLS to support it (Lines 289-296: "A two-component PLS regression model using these 16 storm events as observations and the nine hydroclimatic descriptors as predictors explained 72% (i.e., $R^2Y$ = 0.72) of the variation in average stream DOC concentration and 55% (i.e., $R^2Y$ = 0.55) of the variation in average stream $NO_3^-$ concentration. The $Q^2Y$ values, representing the ability of the model to predict new data, were 0.49 for DOC and 0.25 for $NO_3^-$, both below the 0.50 threshold for good models. Three out of the nine predictors reached VIP>1, indicating that they were important for explaining the variability in the response variables (Fig. A2a). Specifically, higher precipitation amount (higher $P$), larger groundwater table range (higher $\Delta Gw$), and shallower average groundwater table (lower $Gw_{avg}$) related to higher stream concentrations of both DOC and $NO_3^-$, which fell close together in the PLS ordination biplot (Fig. 5a)"; Lines 297-310: "Remarkably, we observed two clusters of observations falling in two opposite sides across the first component of the PLS biplot: a dense cluster of eleven storm events located in the right side, close to $Gw_{avg}$ and opposite $P$ and $\Delta Gw$, and a more scattered cluster of five storm events located in the left side, closer to $P$ and $\Delta Gw$ and far from $Gw_{avg}$ (Fig. 5a). Indeed, there were large hydroclimatic differences between these two subsets of events. The eleven storm events clustered in the right side were characterized by (i) lower $P$, (ii) lower average stream flow ($Q_{avg}$), (iii) deeper $Gw_{avg}$, and (iv) smaller $\Delta Gw$ (i.e., smaller thickness of the riparian layer that becomes hydrologically activated during the event) compared to the five storm events located in left side (Wilcoxon rank-sum test, p<0.01 in all cases; Fig. 6a-d). Given the large difference in $P$ and associated $\Delta Gw$ between the two subsets of events, we classified them as 'medium storms' ($P$ = 29 to 98 mm, $\Delta Gw$ = 2 to 9 cm, n = 11) and 'large storms' ($P$ = 67 to 174 mm, $\Delta Gw$ = 21 to 55 cm, n = 5). Additionally, both stream DOC and stream $NO_3^-$ concentrations were significantly lower during the medium storm events than during the large storm events (Wilcoxon rank-sum test, p<0.01 in both cases; Fig. 6e-f). As a result, stream DOC concentration increased (on average) with respect to the average concentration observed during base flow conditions by 133% during large storms and only by 20% during medium storms. Likewise, stream $NO_3^-$ concentration increased 110% during large storms and only 21% during medium storms compared to average base flow conditions"; Lines 324-326: "Given their variability in hydroclimatic descriptors and in stream DOC and $NO_3^-$ concentrations, we performed a second PLS regression analysis using only the subset of five large storms in order to gain further insights into the hydroclimatic characteristics that most effectively mobilize DOC and $NO_3^-$ during these larger events").

Finally, we have extensively included the medium storm events in our interpretations and in the context of our conceptual model in the discussion section (Lines 448-461: "Results from the first PLS regression model including all storm events showed that precipitation amounts and associated hydrological responses clearly distinguish two subsets of events. Large storms, characterized by higher precipitation amounts, produced higher stream flows and considerable groundwater table elevations, whereas medium storms produced only moderate responses in stream flow and limited groundwater table elevations (Table 2; Fig. 6). These differences in climatic and hydrological characteristics between large and medium storm events led to marked differences in the resulting stream chemistry, with both DOC and $NO_3^-$ concentrations being significantly higher during the large storm events. Hence, the first PLS showed that higher precipitation (higher *P*), larger groundwater table elevations (higher $\Delta Gw$), and shallower groundwater tables (lower $Gw_{avg}$) related to higher stream DOC and $NO_3^-$ concentrations (Fig. 5; Fig. A2). Mechanistically, chemical differences between the two types of events are likely explained by the hydrological activation of a thicker and shallower riparian layer during large storm events that leads to the mobilization of relatively larger amounts of DOC and $NO_3^-$ stored in the riparian soil compared to the amounts mobilized from the deeper and narrower layers that are hydrologically activated during medium storm events (Fig. 8a). These results suggest that large and medium storm events display distinct mechanisms of DOC and $NO_3^-$ mobilization from the riparian zone and that large storms are responsible for a disproportionally larger supply of these solutes"; Lines 462-463: "Thus, storm size and associated groundwater table responses provide a first order control (in terms of relevance) on the mobilization of DOC and $NO_3^-$ from the riparian zone in the Font del Regàs catchment"; Lines 516-525: "Our conceptual model of riparian DOC and $NO_3^-$ mobilization during large storm events based on data from a sub-humid Mediterranean catchment is similar to previous conceptualizations presented for temperate sites. The model highlights that solute mobilization depends on antecedent soil moisture conditions and the relationship between riparian groundwater table and stream flow (Werner et al., 2019; Beiter et al., 2020). Thus, the mechanisms proposed here for large storm events can be representative of other forest headwaters located in similar climatic settings. In addition, our analyses indicate that in Mediterranean forest headwaters, 'medium storms' show limited groundwater table responses and therefore solute mobilization is more restricted and less dependent on antecedent soil moisture conditions or the relationship between riparian groundwater table and stream flow. This distinctive feature of Mediterranean catchments compared to temperate sites might relate to their overall higher temperatures and evapotranspiration rates and, thus, could extent geographically in the future as climate becomes warmer in temperate areas (Spinoni et al., 2018)").

<u>Groundwater table responses and antecedent conditions in the 67-98 mm event precipitation range</u>

The reviewer correctly points out that "*storms in the total precipitation range of 67-98 mm were classified as either medium or large based on whether they produced a large change in the GW table*", and requests us to investigate the threshold for antecedent conditions that lead to significant groundwater table responses in that precipitation range. While this question is interesting, it is both difficult to resolve with the data at hand and somehow outside of the scope of the present work.

Nevertheless, we have addressed this question qualitatively in the discussion section. Our results suggest that seasonality is the key factor determining whether a storm event in the given precipitation range will result in either large or small groundwater table responses. We have also indicated that future studies should further investigate this interesting topic (Lines 463-479: "[…] in our dataset of 16 storms, we observed an overlap in the 67 to 98 mm precipitation range in which two events (*6 and *7) showed small groundwater table elevations (4 and 9 cm, respectively) and were classified as medium storms, and two other events (#1 and #5) showed significant groundwater table elevations (36 and 21 cm, respectively) and were classified as large storms. Therefore, there is an apparent threshold range in which storm size (defined only by precipitation amount) can lead to either substantial or limited groundwater table responses and, thereby, to relatively more or less DOC and $NO_3^-$ mobilization. While the sample size (n = 4) is too small to clearly determine what processes or pre-conditions define whether a storm in the 67 to 98 mm range will lead to significant groundwater table responses, we noted that medium storm events *6 and *7 took place in July-August and October, respectively, whereas large storm events #1 and #5 took place in March and November, respectively. We argue that seasonality and, in particular, the differences in evapotranspiration between the vegetative and the dormant seasons, might provide a plausible explanation in this context. Mediterranean catchments such as Font del Regàs experience long periods of high evapotranspiration that lead to catchment drying and hydrological disconnection in late spring-summer and a subsequent re-wetting in autumn-late autumn (Medici et al., 2008). In this sense, the time of the year when a storm occurs might determine the magnitude of the riparian groundwater table response and the consequent DOC and $NO_3^-$ mobilization for intermediate-medium to large storm sizes in Mediterranean catchments. Future studies should specifically investigate this ambiguous catchment response to similar precipitation inputs in light of the results presented here (e.g. Beiter et al., 2020)").

<u>Concluding remark</u>

Overall, we think that the extended changes described help to highlight and strengthen the novel aspects of our work. We trust that the extended explanations provided above and the exhaustive revisions included in the manuscript offer (i) a more rigorous justification of the classification between large and medium storm events and of the implementation of the PLS analysis that does not appear circular or redundant, (ii) a better explanation of how storm size operates in this context, (iii) a more complete conceptual framework that can be clearly linked to our current understanding of how DOC and $NO_3^-$ are mobilized from riparian zones to streams during storm events, and (iv) clear support from appropriate literature describing the physical mechanisms that we believe to be relevant in explaining our results and conceptual model.

*Minor comments:*

*Lines 274-276: I'm confused about what the numbers in the parentheses indicate here (eg., 33%). Does the 33% indicate that the ratios were optimal 33% of all storm flow time? Please clarify what larger whole these percentages are a part of.*

**[Reply]**: Yes, this is correct; the 33% refers to all storm flow time. The other percentages given in these lines are to be interpreted in the same way. To avoid confusion, we have made this clarification explicit in the text.

*Lines 218-282: Same issue as above. Does 11% mean 11% of the time during baseflow there were N-limited conditions?*

**[Reply]**: Yes. We have now made it explicit.