Manuscript: Disentangling Scatter in Long-Term Concentration-Discharge Relationships: the Role of Event Types

Response to Anonymous Reviewer 1

We thank the reviewer for a positive evaluation of our manuscript and a comprehensive review. Below we provide our point-by-point replies (black color) to the reviewer comments (blue color). New or modified text in the revised manuscript is presented in italics.

General Comments:

1. This study presents a unique, spatially and temporally extensive dataset of nitrate C-Q relationships across 184 German catchments of varying size and land cover/land use from 2000-2015. The authors found that the degree of catchment hydrologic connectivity (and the closely related factor of runoff event type) strongly regulates the pattern of catchment nitrate export. Divergence of event-scale nitrate C-Q responses from the more generalized long-term response were attributed to a combination of catchment topographic properties and event type. The study dataset is impressive, providing a long-term view of catchment nitrate responses across gradients of event type, topography, and land use. The paper is very well written, making it both easy and enjoyable to read; there are only a few places in the manuscript where grammatical clarifications are needed. The statistical analyses presented in the paper are well-presented and wholly appropriate for the research questions that are being asked. Overall, this paper represents a meaningful addition to the existing C-Q literature and I recommend it for publication with only a few minor revisions. My main concern with the paper is that the potential influence of C-Q hysteresis on the observed patterns is not addressed anywhere in the paper. For example, if a particular catchment (or a particular event type) is characterized by a strong hysteresis signal, then the observed C-Q pattern (i.e., dilution or enrichment) would be highly influenced by the timing of sample collection. If a strong sampling bias exists where samples are more frequently collected on the rising limb relative to the falling limb (or vice versa), then the observed catchment or event C-Q signal might be confounded by the presence of hysteresis processes. I would not expect this to be an issue for the long-term C-Q pattern, but it may be an issue for the event-scale patterns and this would, in turn, cause a problem for the interpretation of the "Δres50" term presented in the paper. Given the low frequency of sample collection in this dataset (biweekly to monthly), it seems likely that a particular runoff event would be represented in the dataset by a sample collected either on the rising limb OR the falling limb but not both. It would be fascinating to see an additional analysis of this extensive dataset that incorporates the potential influence of sample timing on the hydrograph, but this is likely beyond the scope of this paper in its current form. However, I do think the authors need to include at least some discussion of the potential influence of this potential "hysteresis effect bias" associated with low-frequency event sampling.

We appreciate the reviewer's comments and the ideas on the possible effect of hysteresis at the event scale. We agree that this effect requires additional attention in the manuscript.

We quantified the proportion of samples taken in the rising limb, falling limb and near to the discharge peak (near-to-peak) of the event hydrographs. The rising limb starts at the beginning of the runoff event and finishs one day before the day of the peak discharge. The falling limb starts one day after the day of the peak discharge and finishs at the end of the runoff event. The beginning and the end of the runoff events are obtained from the runoff event detection method explained in detail in the original manuscript (Lines 132-135). We defined near-to-peak as samples collected from one day before to one day after the day of the peak discharge. We allowed some overlap between near-to-peak and other two groups to use a larger number of samples than considering samples collected on the day of the peak of discharge only. Of the total samples taken during runoff event types 34%

correspond to the rising limb, 55% to the falling limb and 30% to near-to-peak (11% of the samples were collected during the day of the peak discharge). This information will be shown in Figure S6a in the revised manuscript. In addition, we quantified the deviations of the long-term C-Q relationship (Δres50) for samples taken during the rising limb, falling limb and near-to-peak. We computed the deviations for these three groups of samples following the same bootstrapping procedure shown in the Method section (Lines 165-172) of the original manuscript.

The new results provided in Figure S6b show that the deviations from the long-term C-Q relationships for different event types are very similar for all three cases (samples taken during falling limb, rising limb or near the event peak) and resemble the deviations that we have previously observed for all collected samples (Figure 5 in the main manuscript). This suggests that the relative time of sampling during an event does not affect deviations from the long-term C-Q relationships that we detected for different event types.

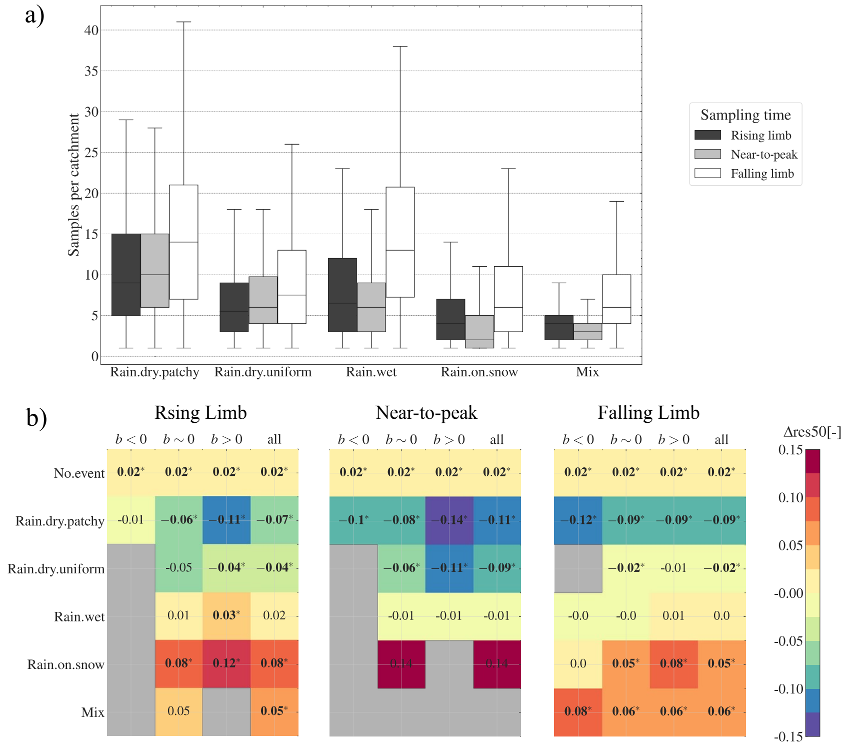

*Figure S6. a) Number of samples per catchment per event type corresponding to the samples taken during the rising limb, falling limb, or near to the peak (i.e., samples taken from one day before to one day after the peak of the hydrograph). b) Median deviations of nitrate concentrations from the long-term C-Q relationships (Δres50) for samples taken during the rising limb, falling limb, and near to the peak. Deviations are computed analogously as for Fig. 5 in the main manuscript. The three first columns of the heatmap correspond to one of the long-term export patterns (i.e., dilution (slope b<0), neutral (slope b~0), and enrichment (slope b>0)), and the fourth column corresponds to all study catchments. Bold font and * indicate significant differences (Kruskal-Wallis test, p<0.05) between median deviations across catchments for each event type and median deviation across catchments of all nitrate samples. At least 5 catchments with sufficient data (more than 10 samples per event type) are required to evaluate the significance of the deviations. Gray squares indicate cases where this requirement is not met.*

We will insert the following description in the revised manuscript in the Method section.

L181: *"Low frequency datasets such as the one used in our study might contain samples collected during different phases of the event hydrograph (e.g., falling or rising limb). This might hamper the interpretability of the results due to possible bias in observed nitrate concentration linked to the time of sampling and the hysteresis effect revealed in high-frequency observations (e.g., Lloyd et al., 2016;*

Vaughan et al., 2017). In fact, Pohle et al. (2021) showed systematic differences in nitrate concentration between samples collected during rising and falling limbs for numerous catchments in Scotland. To understand the potential effect of the hysteresis on the deviations from long-term C-Q (∆res50) we repeat the bootstrapping procedure described above considering samples collected during the rising limb, falling limb and near the event peak (near-to-peak). The rising limb of a runoff event starts at the beginning of the event and finishes one day before the day of the peak discharge. The falling limb starts one day after the day of the peak discharge and finishes at the end of the runoff event. In addition, we defined near-to-peak as samples collected from one day before to one day after the day of the peak discharge. Of the total samples taken during runoff event types 34% correspond to the rising limb, 55% to the falling limb and 30% to near-to-peak. Notice that definition of near-to-peak samples allows some overlap with the other two groups of samples to use a more balanced number of samples than considering samples collected on the day of the peak of discharge only (11% of the samples were collected during the day of the peak discharge). "

We will add the following lines in the Result section.

L236: *"The time of sampling in runoff events did not interfere with our main results (Fig. S6b). Although data limitations for a few groups of samples (gray tiles in Fig. S6b), we could reproduce the analysis for most of the cases. We found that similarly to our results using all the samples (Fig. 5b) values of ∆res50 for samples taken during the rising limb, near to the peak and falling limb, are positive for Rain.on.snow and Mix events, negative for Rain.dry.patchy and Rain.dry.uniform, and intermediate for Rain.wet events."*

We will add the following text discussing the results of the additional experiment to the Discussion section of the original manuscript.

L443: *Although the presence of the event-scale hysteresis effect might considerably affect nitrate concentration during rising and falling limbs of the event hydrograph in some catchments (Pohle et al, 2021) we found a similar direction of deviations from the long-term C-Q relationships when we considered samples taken during rising limb, falling limb and near to the peak (Fig S6b). Hence, our results suggest that the variability potentially added by the presence of hysteresis patterns is lower than the deviations observed for different event types from the long-term C-Q relationship. Increasing availability of high-frequency datasets coupled with new statistical modeling approaches might be used in the future to evaluate hysteresis-related effects in the existing long-term C-Q datasets to further disentangle inter- and intra-event variability of nitrate dynamics at larger scales."*

Specific Comments:

2. Introduction: Very well-written and cited, providing a concise but informative review of the relevant C-Q literature. However, the Introduction focuses heavily (almost exclusively) on the hydrologic drivers of observed C-Q patterns, with little mention of the role of biogeochemical drivers. Particularly in the case of nitrate, biogeochemical drivers—emphasizing the "bio" aspect-- can also influence C-Q patterns. Because this paper focuses solely on nitrate concentrations, I think it is worth mentioning the potential role of biogeochemical processes as drivers of the observed C-Q patterns (this might fit well in the paragraph starting on L56 or after). For example, seasonality of microbial processes might influence soil nitrate concentrations and affect the observed patterns of C-Q especially during seasonal events (e.g., rain-on-snow). Similarly, one might expect the "C" side of the nitrate C-Q relationship to be strongly influenced by the timing of nitrogen fertilizer applications in agricultural catchments. In each of these two examples, the biogeochemical drivers exert as much (of not more) control on the C-Q relationship as the hydrologic drivers. Salli Thompson's 2011 paper "Relative dominance of hydrologic versus biogeochemical factors on solute export across impact gradients" might be a useful paper to consider here.

We thank the reviewer for the point made here on the role of biogeochemical processes. We agree that this point was not addressed sufficiently in the Introduction. We revised accordingly and we will add the following lines to the introduction section.

L46: "*Biogeochemical processes that affect nutrient cycles in soil and water might add variability to long-term C-Q relationships. The effectiveness of the denitrification process, which removes nitrate from the soil, depends on periodic environmental factors such as temperature and soil moisture and the availability of electron donors (Korom et al., 2012; Ortmeyer et al., 2021). Instream removal processes are also more efficient during low flows and higher temperatures, adding more variability to the low-flow portion of the long-term C-Q relationships (Dehaspe et al., 2021; Moatar et al., 2017). Moreover, availability of nitrate sources is balanced by fertilizer application and mineralization of organic nitrogen compounds and hence varies in time adding temporal variability to C-Q relationships. The time of fertilizer application is often unknown, and the mineralization processes depend on chemical soil conditions and environmental factors (e.g., soil moisture and temperature) that mediate communities of microorganisms (Curtin et al., 2012; Guntiñas et al., 2012). On the other hand, average residence times of nitrate in agricultural catchments can last for decades, producing a legacy in soil (Meter et al., 2016; Puckett et al., 2011; Tesoriero et al., 2013; Vervloet et al., 2018) that can buffer the periodic effect of biogeochemical processes reducing the variability in the concentration of nitrate (Basu et al., 2011; Bieroza et al., 2018; Thompson et al., 2011).*

L57: *"Disparate patterns of the event C-Q relationships in a catchment over time are mainly attributed to varying dominant flow sources (e.g., groundwater, shallow subsurface flow), antecedent wetness conditions (Inamdar et al., 2006; Knapp et al., 2020; Vaughan et al., 2017), time of fertilizer application (Bowes et al., 2015; Dupas et al., 2016; Outram et al., 2016), biogeochemical cycling (Heathwaite and Bieroza, 2021) and runoff event characteristics or types (Butturini et al., 2006; Bauwe et al., 2015; Chen et al., 2020; Knapp et al., 2020)."*

3. Methods: If it is possible with your dataset to quantify the proportion of rising limb and falling limb samples, it would be good to include that quantitative information in the Methods section. If the proportions of the two are widely unbalanced, then the potential influence of that sampling bias on your results should be discussed in the Discussion. If it is not possible to determine the rising- or falling-limb status of samples in your dataset, then a brief acknowledgement of the implications of this should still be included in the Methods.

Thanks for the suggestion, we agree that this is important to show. We now include this information in the Methods section and add Figure S6. Please refer to Comment 1 for more details.

4. L134: What is meant here by "precipitation attribution"? Does this mean precipitation classification as rain or snow? Otherwise, I'm not sure what precipitation would be attributed to.

By the term precipitation attribution we mean that the method links runoff events with the corresponding inducing rainfall and/or snowmelt event. We clarify this in the manuscript:

L134: "The method includes baseflow separation, precipitation event attribution *(i.e., corresponding inducing rainfall and/or snowmelt events are linked to runoff events)* and an iterative procedure to adjust site-specific thresholds for the refinement of multi-peak events."

5. L243-247: For these correlation analyses, how did you account for the potential interaction between catchment topographic characteristics and land use? For example, one would expect at least some of the flatter catchments to also be used for agriculture (indeed, Figure S4 seems to indicate this). Thus, a simple correlation between median catchment slope and nitrate C-Q response is not straightforward if it does not somehow control for potential biases due to land use effects on nitrate availability.

We agree with the reviewer that there could be considerable intercorrelation between the characteristics as we acknowledge in the original manuscript at L256 for the fraction of forest. We will additionally highlight this point in the results and Discussion section:

L250: "*Specifically, flatter catchments (low median topographic slope) with greater soil depths that are mostly located in the Northern Germany and Alpine Foreland tend to exhibit more positive residuals for Rain.wet, Rain.on.snow and Mix events, and more negative residuals for Rain.dry.patchy events and samples taken during no event conditions (Fig. 5a). Catchments with these characteristics often also have a higher fraction of agricultural land cover (Fig. S4); however the latter feature shows less significant correlations with nitrate C-Q deviations*."

L395: "*We acknowledge that catchment characteristics might be highly correlated (Fig. S4). Flatter catchments often exhibit higher fractions of agriculture, therefore more diffuse source availability. Although the correlation of the fraction of agriculture and C-Q deviations during Rain.on.snow events was less significant than topographic descriptors, a potential increment of diffuse sources in flatter catchments might also enhance the mechanism of nitrate bypassing the buffer capacity of catchments during Rain.on.snow events generating higher C-Q deviations."*

6. L253: "Instead, we *observed* strong..."? It seems like a word is missing here…

Revised as suggested.

7. L264-266: It would provide useful context here to also provide the ranges around these median values, not only the medians themselves.

We appreciate the reviewer's insightful suggestion. We will add the ranges for the coefficient of variation across event types. We report the coefficient of variation of median runoff coefficient across catchments since mean values of Rain.dry.patchy and Rain.dry.uniform events are very close to zero and obscure computation of their coefficient of variations.

L264: "*Event runoff coefficients exhibit a larger variability across event types than across catchments for most of the catchments. Catchment median event runoff coefficients exhibit a coefficient of variation of 41% across catchments. Nevertheless, median runoff coeffients of event types exhibit coefficients of variation in different catchments from 12% to 118%, with a median value of 67% across catchment."*

L276-296: These two paragraphs are basically invoking the same hydrologic driver for the observed C-Q patterns: catchment wetness status associated with a given event type. But catchment wetness also changes *during* events, and this is where the need to consider potential hysteresis effects / sampling biases becomes important. I am not sure where a discussion of this issue fits in best in the Discussion section, but it should be included somewhere.

Thanks for the insightful comment. We include this concern in the new paragraphs added to the discussion section in Response #1.

L358: The word "wetness" is not needed here.

We agree with the comment and we will remove the word.

L373: "… controls of the variability *of* C-Q …" I think another "of" needs to be added here.

Revised as suggested.

L414: The word "prompt" does not make sense here. I'm not sure what you're trying to convey with that word, but "prompt" doesn't work. Do you mean "prone"?

The comment is right. Corrected.

L432: Do you mean “increase” instead of “increment”?

Corrected.

L457-458: I generally agree, but it’s also important to consider that the Δres50 metric uses INDIVIDUAL grab sample deviations from the long-term C-Q pattern, whereas event-scale metrics like runoff coefficient integrate hydrologic conditions across an entire event. So accounting for potential biases due to the timing of sample collection and hysteresis become important to consider.

We thank the reviewer for pointing this out. The analysis suggested by the reviewer showed that the relative timing of the sample during events does not affect the long-term average deviations for different event types. Please also refer to our response to Comment 1.

References:

Basu, N. B., Thompson, S. E., & Rao, P. S. C. (2011). Hydrologic and biogeochemical functioning of intensively managed catchments: A synthesis of top-down analyses. *Water Resources Research*, *47*(10). https://doi.org/10.1029/2011WR010800

Bauwe, A., Tiemeyer, B., Kahle, P., & Lennartz, B. (2015). Classifying hydrological events to quantify their impact on nitrate leaching across three spatial scales. *Journal of Hydrology*, *531*, 589–601. https://doi.org/10.1016/j.jhydrol.2015.10.069

Bieroza, M. Z., Heathwaite, A. L., Bechmann, M., Kyllmar, K., & Jordan, P. (2018). The concentration-discharge slope as a tool for water quality management. *Science of The Total Environment*, *630*, 738–749. https://doi.org/10.1016/j.scitotenv.2018.02.256

Bowes, M. J., Jarvie, H. P., Halliday, S. J., Skeffington, R. A., Wade, A. J., Loewenthal, M., Gozzard, E., Newman, J. R., & Palmer-Felgate, E. J. (2015). Characterising phosphorus and nitrate inputs to a rural river using high-frequency concentration-flow relationships. *The Science of the Total Environment*, *511*, 608–620. https://doi.org/10.1016/j.scitotenv.2014.12.086

Butturini, A., Gallart, F., Latron, J., Vazquez, E., & Sabater, F. (2006). Cross-site Comparison of Variability of DOC and Nitrate c–q Hysteresis during the Autumn–winter Period in Three Mediterranean Headwater Streams: A Synthetic Approach. *Biogeochemistry*, *77*(3), 327–349. https://doi.org/10.1007/s10533-005-0711-7

Chen, X., Parajka, J., Széles, B., Valent, P., Viglione, A., & Blöschl, G. (2020). Impact of Climate and Geology on Event Runoff Characteristics at the Regional Scale. *Water*, *12*(12), 3457. https://doi.org/10.3390/w12123457

Curtin, D., Beare, M. H., & Hernandez-Ramirez, G. (2012). Temperature and Moisture Effects on Microbial Biomass and Soil Organic Matter Mineralization. *Soil Science Society of America Journal*, *76*(6), 2055–2067. https://doi.org/10.2136/sssaj2012.0011

Dehaspe, J., Sarrazin, F., Kumar, R., Fleckenstein, J., & Musolff, A. (2021). Bending of the concentration discharge relationship can inform about in-stream nitrate removal. *Hydrology and Earth System Sciences*, *25*, 6437–6463. https://doi.org/10.5194/hess-25-6437-2021

Dupas, R., Jomaa, S., Musolff, A., Borchardt, D., & Rode, M. (2016). Disentangling the influence of hydroclimatic patterns and agricultural management on river nitrate dynamics from sub-hourly to decadal time scales. *Science of The Total Environment*, *571*, 791–800. https://doi.org/10.1016/j.scitotenv.2016.07.053

Guntiñas, M. E., Leirós, M. C., Trasar-Cepeda, C., & Gil-Sotres, F. (2012). Effects of moisture and temperature on net soil nitrogen mineralization: A laboratory study. *European Journal of Soil Biology*, *48*, 73–80. https://doi.org/10.1016/j.ejsobi.2011.07.015

Inamdar, S. P., O’Leary, N., Mitchell, M. J., & Riley, J. T. (2006). The impact of storm events on solute exports from a glaciated forested watershed in western New York, USA. *Hydrological Processes*, *20*(16), 3423–3439. https://doi.org/10.1002/hyp.6141

Knapp, J. L. A., von Freyberg, J., Studer, B., Kiewiet, L., & Kirchner, J. W. (2020). Concentration–discharge relationships vary among hydrological events, reflecting differences in event characteristics. *Hydrology and Earth System Sciences*, *24*(5), 2561–2576. https://doi.org/10.5194/hess-24-2561-2020

Korom, S. F., Schuh, W. M., Tesfay, T., & Spencer, E. J. (2012). Aquifer denitrification and in situ mesocosms: Modeling electron donor contributions and measuring rates. *Journal of Hydrology*, *432–433*, 112–126. https://doi.org/10.1016/j.jhydrol.2012.02.023

Lloyd, C. E. M., Freer, J. E., Johnes, P. J., & Collins, A. L. (2016). Using hysteresis analysis of high-resolution water quality monitoring data, including uncertainty, to infer controls on nutrient and sediment transfer in catchments. *Science of The Total Environment*, *543*, 388–404. https://doi.org/10.1016/j.scitotenv.2015.11.028

Meter, K. J. V., Basu, N. B., Veenstra, J. J., & Burras, C. L. (2016). The nitrogen legacy: Emerging evidence of nitrogen accumulation in anthropogenic landscapes. *Environmental Research Letters*, *11*(3), 035014. https://doi.org/10.1088/1748-9326/11/3/035014

Moatar, F., Abbott, B. W., Minaudo, C., Curie, F., & Pinay, G. (2017). Elemental properties, hydrology, and biology interact to shape concentration-discharge curves for carbon, nutrients, sediment, and major ions. *Water Resources Research*, *53*(2), 1270–1287. https://doi.org/10.1002/2016WR019635

Ortmeyer, F., Begerow, D., Guerreiro, M. A., Wohnlich, S., & Banning, A. (2021). Comparison of Denitrification Induced by Various Organic Substances—Reaction Rates, Microbiology, and Temperature Effect. *Water Resources Research*, *57*(11). https://doi.org/10.1029/2021WR029793

Outram, F. N., Cooper, R. J., Sünnenberg, G., Hiscock, K. M., & Lovett, A. A. (2016). Antecedent conditions, hydrological connectivity and anthropogenic inputs: Factors affecting nitrate and phosphorus transfers to agricultural headwater streams. *Science of The Total Environment*, *545–546*, 184–199. https://doi.org/10.1016/j.scitotenv.2015.12.025

Pohle, I., Baggaley, N., Palarea-Albaladejo, J., Stutter, M., & Glendell, M. (2021). A Framework for Assessing Concentration-Discharge Catchment Behavior From Low-Frequency Water Quality Data. *Water Resources Research*, *57*(9), e2021WR029692. https://doi.org/10.1029/2021WR029692

Puckett, L. J., Tesoriero, A. J., & Dubrovsky, N. M. (2011). Nitrogen Contamination of Surficial Aquifers—A Growing Legacy. *Environmental Science & Technology*, *45*(3), 839–844. https://doi.org/10.1021/es1038358

Tesoriero, A. J., Duff, J. H., Saad, D. A., Spahr, N. E., & Wolock, D. M. (2013). Vulnerability of Streams to Legacy Nitrate Sources. *Environmental Science & Technology*, *47*(8), 3623–3629. https://doi.org/10.1021/es305026x

Thompson, S. E., Basu, N. B., Lascurain Jr., J., Aubeneau, A., & Rao, P. S. C. (2011). Relative dominance of hydrologic versus biogeochemical factors on solute export across impact gradients. *Water Resources Research*, *47*(10). https://doi.org/10.1029/2010WR009605

Vaughan, M. C. H., Bowden, W. B., Shanley, J. B., Vermilyea, A., Sleeper, R., Gold, A. J., Pradhanang, S. M., Inamdar, S. P., Levia, D. F., Andres, A. S., Birgand, F., & Schroth, A. W. (2017). High-frequency dissolved organic carbon and nitrate measurements reveal differences in storm hysteresis and loading in relation to land cover and seasonality. *Water Resources Research*, *53*(7), 5345–5363. https://doi.org/10.1002/2017WR020491

Vervloet, L. S. C., Binning, P. J., Børgesen, C. D., & Højberg, A. L. (2018). Delay in catchment nitrogen load to streams following restrictions on fertilizer application. *Science of The Total Environment*, *627*, 1154–1166. https://doi.org/10.1016/j.scitotenv.2018.01.255

Winter, C., Tarasova, L., Lutz, S., Musolff, A., Kumar, R., & Fleckenstein, J. (2022). *Explaining the Variability in High-Frequency Nitrate Export Patterns Using Long-Term Hydrological Event Classification*. https://doi.org/10.1002/essoar.10507676.1