Response to Review

Reviewer #1:

This manuscript combines data-driven analyses and modeling approaches to understand how catchment and drought characteristics impact annual N export, transit times, and overall catchment N retention across a set of nested sub-catchments in Germany. The manuscript is very well-written and easy to read, and the authors present their ideas and results in a clear and compelling way. The authors combine of data-driven C-Q analyses with catchment modeling which is a unique approach.

Overall, this paper has the potential to be a strong contribution to the literature, but additional information/considerations are needed. Below are some general suggestions, as well as line by line comments.

We thank the reviewer for this positive evaluation and constructive feedback. In the following, we provide a point-by-point response to the reviewers' suggestions. We are confident that we could appropriately address all points raised, which helped to significantly improve the quality of our manuscript.

General comments:

This paper explores how antecedent drought impacts N export and retention. There has been a fair bit of work on this in the US – in particular Davis et al. 2014 (Journal of Environmental Quality) and Loecke et al. 2017 (Biogeochemistry). The idea of "weather whiplash" may be a useful reference and motivation for this work.

Thank you for this suggestion and the provision of interesting literature. We have now considered ideas from the suggested papers in the Introduction and also in the Discussion of our work. For example, we now write: "For example, Loecke et al. (2017) reported that the shift from very dry to wet conditions resulted in an increase in flow-weighted nitrate export compared to previous years, and Davis et al. (2014) suggested that the antecedent soil moisture conditions play an important role in the response of nitrate export during runoff events."

The manuscript relies on a previously calibrated model (mHM-SAS, Nguyen et al 2022) for a fair bit of the analysis. While it is beyond the scope of this manuscript to include/replicate all the calibration information in the original modeling paper, the authors need to provide enough detail for readers of this paper to understand the analysis presented in this paper. For example, SAS functions were not introduced until line 202, and there is no description of what they are or how they are defined, and how SAS functions relate to nitrate generally. The a/b ratio (line 292) was never defined or explained, despite it being a central figure in the paper. It seems important to have at least some content explaining the model/SAS functions (and how they relate to nitrate cycling) in the Introduction, and then substantially more context in the methods, perhaps even results. For example, the authors mention briefly the relatively low NSE efficiency for the lower basin model in the methods section (line 204) but never expand in results or Discussion.

Thank you for this helpful comment. In our revised manuscript, we now introduce SAS functions in the Introduction and add a general description of the mHM-SAS model with a focus on StorAge Selection (SAS) functions, the a/b ratio, and their relation to nitrate transport in the Method section from line 189 onwards. Specifically, we added that SAS functions describe the selective removal of water from a subsurface storage with different water ages and nitrate concentrations, which allows for a nitrate transport formulation based on time-variant transit time distributions (TTDs). The SAS function in mHM-SAS is described using a beta function, *beta(a,b)*, with *a* and *b* being two fitted parameters that vary in time depending on the antecedent wetness condition (see Supplement, eq. 6). The a/b ratio implies the selection schemes for discharge, e.g., preference for young water (a/b ratio < 1) or old water (a/b ratio > 1; as documented in the Supplement or in Nguyen et al., 2022). As nitrate is transported by water, the distribution of water ages in discharge (calculated from the SAS function) controls the export of nitrate.

One central component of the analysis is the catchment vs soil N retention capacity. The catchment N retention metric is using a simple mass balance approach. However the soil N retention capacity is a simulated parameter from the mHM-SAS model. After digging back into the original Nguyen 2021 manuscript (which appears to be missing from the reference list), it doesn't seem like this parameter has been validated in any way. Were there any measurements of root zone leachate made to validate this particular parameter estimate? At the very least, there is likely uncertainty in this parameter estimate, but this is not propagated into the Nret-soil metrics. There are numerous fates of N in the soil, and this seems like it has the potential to be highly uncertain. It would be helpful if the authors could provide more nuance here, as this is potentially an over-simplification.

Thank you for this remark. We added all relevant references for the mHM-SAS model in the revised version (in the model description). We agree that there is uncertainty in simulated N in the soil leachates resulting from the mHM-SAS model, which has not yet been incorporated into the soil N retention capacity. Unfortunately, we do not have measurements of nitrate concentrations in the soil leachates available for direct validation. Therefore, to address this point, we calculated the propagation of the parametric model uncertainty in our soil N retention metric. In the revised version of our manuscript, we added this model uncertainty in the form of error bars to the soil N retention capacity (Figure S4). Furthermore, we will shift the focus from a direct comparison of catchment N vs. soil N retention more toward catchment N retention capacity, which was calculated from measured nitrate loads in the stream and is, therefore, the more robust variable.

Finally, the forest change scenarios seemed like side component of the manuscript that wasn't fully explained, explored, or integrated into the main narrative of the manuscript. The scenarios used were extreme (100% vegetation loss?), and overall it came across as somewhat of a diversion from the main story line which was so focused on drought. I don't think this is a necessary part of the manuscript.

Thank you for this comment and the direct suggestion on how to solve it. We followed this suggestion and took the scenarios out of the manuscript and the SI, reducing it to a few sentences in the Discussion on forest dieback in the upper part of the catchment, which might affect future nitrate export.

Line by line comments:

Line 100: is the agriculture in the lower catchment irrigated? If so, how would this impact the findings discussed throughout the rest of the manuscript?

Thank you for this question. We agree that irrigation could potentially change drought impacts on N cycling. However, irrigation is not (yet) a common practice in the study area (EEA, 2018). Still, irrigation might become more common in the future, and we therefore added information on its potential impact in the Discussion of our work (line 395 onwards).

Specifically, we added that irrigation would increase the soil moisture content and, therefore, might buffer drought impacts, such as the decrease in plant uptake and denitrification. Moreover, irrigation might lead to the mobilization of nitrogen as nitrate, which would otherwise be retained in the upper soil until rewetting in autumn. Hence, crop irrigation might counterbalance the reduction of N retention in the catchment soils and the accumulation of N during summer. However, irrigation would also increase the pressure on the available water resources and might enhance greenhouse gas emissions from agricultural soils (Sapkota et al., 2020).

Line 115: I think this description could be clarified/simplified. eg - 1 year = 1 drying/wetting cycle. the year starts in May. Calling it a drying-rewetting cycle instead of year is clunky.

By calling it a drying-wetting cycle, we wanted to avoid confusion for readers that might have skipped the method part. However, we see that the term “drying-wetting cycle” sounds clunky. In the revised manuscript, we therefore refer to “year” instead of “drying-wetting cycle” and emphasize that it starts in May throughout the figures of our manuscript.

Line 160: This is results, not methods.

We have shifted that part to the result section (from line 230 onwards).

Line 195: I think this section needs to be expanded to provide more context about the modeling approach here. For example – what are SAS functions? Why is the NSE so low for the lower catchment?

As explained above, we fully agree and have therefore added more information on SAS functions, the a/b ratio, and sub-catchment differences in the NSE efficiencies (from line 189 onwards).

Line 202: have the authors considered whether/how a model calibrated under moderate climate conditions will perform under extreme climate conditions? Might there be specific hydrologic processes that are more sensitive to drought and therefore do not simulate well during a period with different climate conditions than the calibration period?

The mHM-SAS model is a deterministic model with a strong physical basis, with the hydrological concept from mHM, soil nitrogen dynamics from the HYPE model, and subsurface transport from the SAS concept. These concepts have been demonstrated to be able to work under different climatic conditions (Lindner et al., 2010; Samaniego et al., 2010; Benettin and Bertuzzo, 2018). Therefore, we believe that the model can also work under a wide range of hydro-meteorological conditions. However, we agree with the reviewer that a model calibrated under moderate climate conditions might not perform well under extreme climate conditions, which is not the case in this study. Here, we demonstrated that the model performance is acceptable (at least for instream nitrate concentrations at the three gauging stations) in both calibration and validation periods (Nguyen et al., 2022). For other studies, this might not be true; for example, the model performance under extreme conditions might deteriorate compared to that under normal conditions.

We further agree with the reviewer that there might be "specific hydrologic processes that are more sensitive to drought and therefore do not simulate well during a period with different climate conditions than the calibration period". In this study, the model was calibrated using integrated signals (streamflow and nitrate concentrations at gauging stations), and we do not have data about the internal processes for a quantitative evaluation of the model performance under normal and drought conditions. However, the internal processes under normal and drought conditions (simulated by the model) are based on our physical understanding of natural processes, also under very dry conditions. For example, TTs were longer during low flow conditions, and denitrification rates decreased in very dry soils. To account for this, we added a short form of this explanation in the section from line 184 onwards.

Line 220: This sensitivity analysis is good - but do you think that N inputs might systematically have changed during the drought? How might applying the long-term N inputs to the drought period impact your conclusions?

This is indeed an important point to consider. In our study site, N input did not significantly change because fertilization (the main source of N input) is mainly done in spring when it is not yet clear to the farmers whether the summer will be dry. Atmospheric deposition was slightly reduced due to the decrease in rainfall. However, atmospheric deposition makes up only a small fraction of N input to the Selke catchment, and thus its reduction can be considered of minor importance. We made these points clearer in the revised manuscript (line 215).

Line 231: instead – word choice?

We replaced "instead" with "In contrast" to point out the difference between $N_{ret\text{-}soil}$ and N_{ret}.

Line 273: More information on model fits needed (found it in line 209, but it could use a brief description in the results)

We added information on the model fit in line 279 of the revised manuscript.

Line 292: a/b ratio has note been defined or described; substantially more context is needed here.

We added this information from line 189 onwards.

Line 295: It is surprising to that the median TTs are dominated by old water in the lower catchment, which is dominated by agriculture. Typically agricultural land use encourages rapid movement of water through the surface/subsurface (via tile drainage, etc), which might lead to younger water fractions. Can you provide more Discussion as to why the median TTs are so long here, particularly given the land use? It also would be helpful to have more information in the site description about irrigation practices in the agricultural portion of the watershed, as this might impact TTs and flushing of N.

Thank you for this important remark. As elaborated before, agriculture in the catchment is not irrigated, but this might change in the future. We fully agree that tile drainage in agricultural areas leads to short transit times. However, in the lower Selke catchment, tile drainage is extremely sparse (map of the potential need for irrigation in the province of Saxony-Anhalt, in German: https://lhw.sachsen-anhalt.de/fileadmin/Bibliothek/Politik_und_Verwaltung/Landesbetriebe/LHW/neu_PDF/5.0_GLD/Dokumente_GLD/Wasserhaushalt_Bio/Dr%C3%A4nagekarte_ST_2011.pdf). This can be explained by the relatively deep groundwater levels and therefore rarely any wetted areas that need to be drained for agricultural purposes (Musolff et al., 2015). Additionally, the long transit times in this lowland part of the catchment are in agreement with results from previous studies (Winter et al., 2021; Nguyen et al., 2022). For example, our estimated long TTs are supported by the fact that after the reunification of Germany in 1898/1990, fertilizer levels drastically decreased across Eastern Germany, while it took around 12 years downstream until in-stream nitrate levels in this part of the catchment started to decrease (Winter et al., 2021). In contrast, in the upstream part of the catchment, the decrease in nitrate concentrations instream could be observed already in the following year, in agreement with the relatively short TTs reported here. The long TTs in the downstream part of the catchment can be explained by the thick soils (mainly Chernozems), the flat topography, and especially by the deep sedimentary aquifers. We added that information in the discussion part of our manuscript (from line 421 onwards).

Line 324: this applies to throughout the manuscript, but the language of upper Selke could be clarified by in parentheses including the site names (SH, MD). It’s hard to remember which sites are upper vs lower based on the initials (which is what is labeled on all the figures).

That is good to know – thank you. We modified our wording by adding site names in parentheses throughout the manuscript.

Line 336-337: This is an incomplete sentence

We corrected the sentence so that it now reads, “We found an increased intra-annual variability of nitrate export, with lower concentrations during the dry periods (July-October) and exceptionally high concentrations during the subsequent wet periods (January-April).” (lines 335-337).

Line 337: Were the concentrations during the drought exceptionally high? They appear to plot on top of the existing historical data, albeit on the high end – but this makes it sound like the concentrations were much higher than what was previously observed.

Nitrate concentrations during the dry summer periods were relatively low. However, with rewetting in autumn, nitrate concentrations increased to the highest values ever measured at the upstream part of the catchment, and they were among the highest values measured at the downstream gauge. Therefore, we argue that nitrate concentrations are exceptionally high. This might not become apparent from the C-Q plot in Figure 2 due to the log-transformed y-axis; however, it is visualized more clearly in Figure 1, where measured high-frequency nitrate concentrations are depicted.

Line 455: this is a really interesting point! Very cool!

Thank you! That is very nice to hear.

Line 343: Overall this appears to be true, but it also warrants some more nuanced discussion/context – similarly high concentrations were observed at other times in the historical record. Were those droughts as well?

Previous dry years include 2003, 2006, and 2016 (Figure 1b), which also show relatively high nitrate concentrations during the subsequent high-flow season but are not as pronounced as in 2018 and 2019 (Figure 1c). The significant shift in the C-Q slope for the years 2018 and 2019 compared to previous years (Figure 2 a-c) shows that nitrate concentrations were higher than in previous years, given the same discharge. Still, we agree that the sentence in is somewhat an oversimplification and have therefor replaced it by relating our results to the ones from Loecke et al. (2017) and Davis et al. (2014).

Line 358: Can you provide more explanation here? Why is the potential for denitrification in groundwater lower in the downstream sub-catchment? Why would the capacity for in-stream retention be greater in this area (if denitrification is reduced)? Why would those two microbial processes respond differently to the same drivers? Reduced flow velocity would like enhance both?

Thank you for this very important remark. The potential for denitrification in the groundwater differs between the upstream and the downstream part of the catchment, likely due to differences in the geology (Table 1). Hannappel et al. (2018) looked at evidence for groundwater denitrification in that area and found such evidence only in the upstream sub-catchment. This can be explained by a lack of electron donors in the aquifer which are needed for denitrification (Rivett et al., 2008). Nevertheless, this does only apply to denitrification in the groundwater, not in the soils or in the stream. We see now that this was not sufficiently clear in the manuscript and reframed it accordingly. Microbial processes in the catchment soils are affected by more environmental factors than the sub-catchment specific differences in geology and we expect denitrification in all parts of the catchments, except under very dry conditions as in the summer period of 2018 due to the low soil moisture content.

Regarding in-stream processing: A study by Rode et al. (2016) analyzed N assimilatory uptake in the Selke River and detected strong differences between upstream and downstream. These differences were explained by the different light availability (less light available in the forested upstream area, where the river is more often under the shade of trees) and changes in flow velocity because the upstream area is located in the Harz Mountains with steeper slopes, whereas the downstream part of the catchment has a relatively flat topography (Table 1). In terms of denitrification, Trauth et al. (2018) found the intense interaction of surface water and groundwater in the hyporheic zone and the riparian aquifer to facilitate more denitrification, especially under summer low flow conditions. The upper Selke lacks sand- and gravel bars and an extensive riparian aquifer as a hot-spot of denitrification.

We added a short form of these arguments to our Discussion from line 361 onwards.

Line 380: this is also a general comment – It seems like there is a disconnect between the upper and lower catchment behavior throughout the manuscript. How much of this might be driven by lower model fit metrics/uncertainty in the model outputs for that spatial scale specifically? Also does the land use change factor in here? How about irrigation in the lower agricultural catchment during the drought?

That is a very good point. Differences between the upstream and the downstream part of the catchment are a crucial factor throughout our manuscript and have already been analyzed in previous studies. Certainly, the differences in land use do play an important role, especially because of the differences in land use and related N input via fertilization. As explained above, irrigation is not a driving factor for the observed differences, but this might change in the future. Additionally, changes in the topography from the Harz Mountains to the lowland play an important role as well as the geology, with relatively impervious bedrock upstream in the Harz Mountains compared to deep sedimentary aquifers downstream that allow for longer and deeper flow paths. Moreover, annual precipitation in the elevated upstream area is higher (Table 1), which allows for a faster transit of water through the sub-catchment. Our findings of relatively fast TTs upstream and longer TTs downstream are in agreement with these geological and meteorological settings, and they are in agreement with previous studies that, for example, analyzed the time it took for a strong decrease in fertilizer application to become apparent in decreasing nitrate concentrations in the stream (Winter et al., 2021). Therefore, we argue that this is not an artifact that was driven by the differences in the model uncertainty downstream but rather that it can be physically explained and was confirmed by other studies that did not use the same model. We agree that there is model uncertainty, and even more so in the downstream part of the catchment, which is even larger and therefore integrates more influencing factors. Nevertheless, also our combination of data-driven analysis and process-based modeling showed that the differences between upstream and downstream are quite consistent. We have further elaborated on these sub-catchment-specific differences in section 4.2 (Sub-catchment-specific contributions to the integral nitrate response to drought).

Figure 1: This figure is very data dense and could potentially be split apart into panel a/b and a separate figure for panel c. Further, the flipped double axis is difficult to read – typically discharge is not shown as a hanging plot (usually this is used for precipitation). One suggestion

is to pull the hydrographs for all 3 onto a single panel of this figure. And finally, panel C seems like it should be in the results section – it is showing the calibration results of the model.

Thank you for these suggestions. We followed them in the way that we i) separated panels A and B from C, ii) placed discharge in C in a separate combined panel, and iii) shifted it to the result section.

Figure 2: The color of the slope lines for panel a-c is confusing, because the color scheme flips (grey line corresponds to orange points, red line corresponds to grey points). Also, the Discussion starting around line 335 is drawing attention to differences in these C-Q plots by wet vs dry season. The colors here are differentiated by year, and obscure the point about wet vs dry season (within a year). In contrast, there isn't a discussion of how the C-Q slope changes between 2018 and 2019. So you could consider shifting the colors to show the wet/dry season comparison and not accentuate year. This might fit the Discussion better.

We noticed a typo in the legend of this plot which explains the confusion. We have now corrected this so that grey dots correspond to grey lines and reddish dots correspond to the dark red line. Generally, the C-Q slope goes from low Q to high Q, representing the low-flow vs. high-flow season. Therefore, we would argue that a separation into seasons would not add much new information because this is already incorporated in the "Q" of the C-Q slope on the x-axis. Our main message here is that the C-Q slope during the drought, which is strongly controlled by intra-annual differences in Q, differs from previous years, meaning that nitrate concentrations differ from previous years in their relation to Q; thus, similar Q conditions would show different nitrate concentrations.

Figure 3: suggest labeling y axis in panels d-f in same units used in written text/results section (years, not days).

Good point. We have done that.

Reviewer #2:

Referee report on: Droughts can reduce the nitrogen retention capacity of catchments

By Carolin Winter, Tam V. Nguyen, Andreas Musolff, Stefanie R. Lutz, Michael Rode, Rohini Kumar, Jan H. Fleckenstein

I agree with the first referee that this is a nicely written manuscript with clear hypotheses and a good concept to analyse altered N dynamics under drought conditions within a mesoscale catchment. In particular, I like the strategy to combine a data-driven and model-based analysis. Although I am no native speaker, I think that language issues are largely settled, which is why a reviewer can concentrate on the most important issues: science and novelty behind the presented analysis. So I congratulate all authors for this work.

We would like to thank the reviewer for his/her positive and encouraging feedback. This is very nice to hear. We would further like to thank the reviewer for the constructive remarks and suggestions, which we address point by point below.

This said, I still have some major points that I think should be addressed prior to publication in egushere.

1. Information about N-Inputs

You state that exported N loads have generally decreased (e.g. L 268). What about the input? I propose to present the entire N-balance for the different subcatchments instead of only four components in Table 2. Moreover, a diagram would be more intuitive than a table.

Thank you for this suggestion. We agree that it is beneficial to the reader to have the entire N balance, including N input, presented as a figure. Therefore we included N input and presented all N fluxes in the form of a stacked bar-plot. In this new Figure 4, we display N fluxes leaving the system (denitrification, plant uptake, and leachates) as negative values and N fluxes entering the system (N input, amt. deposition, and mineralization) as positive values. Furthermore, we added error bars to visualize the inter-annual variability in the long-term reference period (1997-2017). Regarding trends in N input: In the Selke catchment, N inputs have stabilized since around 1995 (Winter et al., 2021). Therefore long-term N input to mHM-SAS does not show any trend. We have added this information in line 215.

2. Stability of catchment properties

This is a general drawback: If you compare N-retention–discharge relations during the recent drought to longterm values, you assume that catchment properties stayed the same. Can you really do this? There might be landuse changes (forest diebacks after storm events, or due to preceding droughts, e.g. 2003) and there is for sure a rising trend in temperatures. This trend per se alters the N-cycle, prolongs vegetation periods and presumably intensifies processes. I would like to see this point in the discussion section.

We thank the reviewer for this important remark. We agree that catchment properties are not necessarily stable in time, especially not in the face of climate change. In the following, we discuss in more detail which characteristics might have changed, which characteristics are relatively stable, and the potential implications for N retention:

Land cover change: We used the CORINE land cover map of the Copernicus Programme (https://land.copernicus.eu/pan-european/corine-land-cover) to check for land use changes from 1990 over 2000 until 2018. Overall, land use changes were minor. We detected a slight increase in urban areas (<1% of the catchment area), a slight decrease in agricultural areas (around 1.6%), and a slight increase in grasslands. An open mining pit in the downstream part of the catchment was closed down around 1996/1997 and was transformed into a lake (see Winter et al., 2021). On this basis, we expect no major trends in catchment functioning in terms of N cycling from land use changes, at least within our observation period.

Forest damage: The mortality rate of trees in the forests of Saxony Anhalt (the state where the Selke catchment is located) remained relatively stable below 1% in the years from 1991 until 2017 (State-of-the-forest Report Saxony Anhalt, 2020). Therefore we do assume forest dieback to be relatively stable across our long-term control period. Only with the beginning of the severe drought, mortality rates increased above 1% and even towards >4% in 2019, which we discuss in our manuscript.

Temperature Increase: With global warming, also temperatures in the Selke catchment increased significantly in the long term. Increasing temperatures have been found to cause an increase in microbial N cycling, including increased rates of mineralization, nitrification, and denitrification (Dai et al., 2020). During the severe summer drought, soils were too dry for denitrification and plant uptake, so the temperature increase did not play an important role anymore. However, just as mentioned by the reviewer, given sufficient soil moisture, increasing temperatures can potentially intensify biogeochemical processes within a catchment. We have added this to our discussion section.

Extension of vegetation period: With climate change, also the length of the vegetation period increases. Menzel et al. (2001) reported an increase of up to 0.2 d yr^{-1} across Germany, but with high inter-annual variability. However, over the 23 years of data used for our study, this corresponds to <5 days, which does likely not have a stronger impact than the overall hydro-meteorological variability. Furthermore, over the period from 1997-2018, no clear trend in the overall length of the vegetation period could be observed from the Germany-wide averages (Figure 1 below). We, therefore, argue that the extension of the vegetation period very likely impacts N cycling but not in a way that it has fundamentally changed catchment functioning in terms of N retention within the observation period of our study.

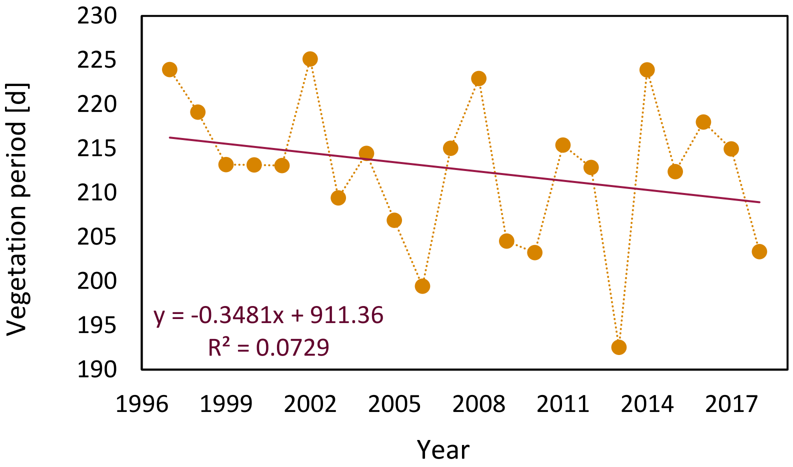


Figure 1. Annual averages of the vegetation period in Germany. Data source: Umweltbundesamt (https://www.umweltbundesamt.de/daten/klima/veraenderung-der-jahreszeitlichen#weiterfuhrende-informationen, retrieved 7/25/2022)

Nevertheless, the long-term L-Q relationship as well as the N_{ret}-Q relationship across all years previous to the drought, could be very well fitted (R^2 0.91 – 0.96; Figure 2 d-f and Figure 4). This indicates that despite changing temperatures, nitrate load export and retention have been dominantly controlled by discharge. Temperature increase and variability, along with other extreme events, such as floods, are likely to explain a part of the residuals in the L-Q and the N_{ret}-Q relationships; however, none of these years stood out to the same degree as 2018 and 2019 in the upper Selke. Consequently, we argue that the unprecedented 2018-2019 drought also caused unprecedented changes in N export and retention that are not only controlled by low discharge but also by changes in catchment functioning. This can explain the deviation between 2018 and 2019 from the long-term L-Q and N_{ret}-Q relationships.

We will add a paragraph on long-term stability with a focus on rising temperatures to our Discussion (chapter 4.3 Exported nitrate loads and catchment retention capacity).

3. Scenario of forest dieback

This point was already raised by referee #1: Omit the simplistic scenario of forest dieback. You only simulate reduce N uptakes but there are surely more effects on the N cycle here: e.g. mineralisation of dead organic material, altered soil characteristics, etc.

We fully agree and have removed the part about the forest dieback scenarios and only left a short part on forest dieback (also regarding catchment stability) in the Discussion.

4. Increased N-mineralization during droughts

You speculate that mineralization during droughts might increase, based on a single study that found increased rates of depolymerisation during droughts in a montane grassland in Austria. I would be more careful when transferring these findings to the forested Selke catchment. E.g.

depolymerisation in montane grasslands might principally be temperature dependent. In forest soils, I would rather argue that mineralization is hampered by soil moisture deficits during droughts and subsequently reduced microbial activity. Your data only shows increased mineralization during the entire dry-wet cycle which could be due to onset of strong mineralization during re-wetting. Then mineralization could be more intense, because you have an organic N-pool accumulated during the drought.

Thank you for this comment. We understand the point that evidence from a single case study in a different ecosystem is a little weak for the base of our argumentation and have therefore removed the sentence about depolymerization. Overall, our line of argumentation is very much in line with what the reviewer suggests here: "the rewetting of dry soils in autumn can cause a peak in Mineralization that transforms accumulated organic material into mobile inorganic N" (Line 376-377). We have further added the reviewer's argument that mineralization is likely more intense due to the accumulated pool of organic N under very dry summer conditions.

5. Groundwater data to illustrate longer term N-effects

You hypothesize that in the upper Selke short TTs lead to visible effects of N-retention rather than in the low Selke, where longer TTs might cause longer term effects. This is a logic outcome of your transit time model. I think this could nicely be proved by groundwater data. Do you have well or spring data that could be used to support your hypothesis here? Recent papers have looked on groundwater nitrate responses after droughts and e.g. found instantaneous or delayed reactions depending on aquifer types.

This is an interesting point. We found one observation well in Wilsleben in the lower Selke catchment, maintained by the LHW, which supports our results:

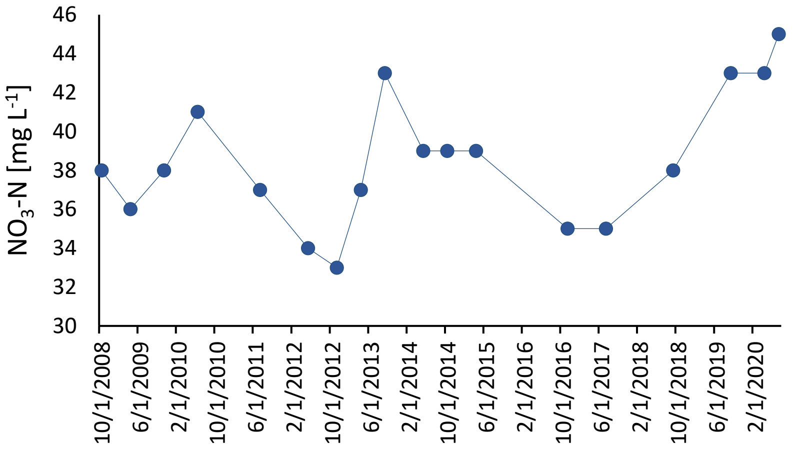


Figure 2. Groundwater observation well in Wilsleben (lower Selke catchment). Data provided by the LHW.

Nitrate-N concentrations in the groundwater from 2019 onwards were among the highest values of the time series (starting in 2008). So, while the decrease in nitrate retention (i.e., higher nitrate concentrations) in the downstream part of the catchment has not (yet) reached the stream network, there is already some indication for higher nitrate concentrations in the groundwater. We added this plot to the Supplement and added the information to our Discussion (line 471 onwards). However, there is high variability in nitrate concentrations measured at this well, and therefore we see it as an indication that supports our results rather than a real proof of concept, which would require data from more than just one well. Other studies, just as mentioned by the reviewer, found a similar increase in groundwater nitrate concentrations looking at a large number of wells (e.g., Jutglar et al., 2021), which we have therefore mentioned in our Discussion as well.

6. Shape of the Retention-Discharge-Relationship

I disagree with the shape of the function between retention capacity and log-scaled discharge, as presented in Figure 4 and in the conceptual framework in Figure 5. A linear relationship is not meaningful here, maximum retention is 1 (at zero discharge), and your figure implies higher values. The data of HD shows this shape shape quite nicely. So put in the upper limit and an asymptotical approach of the function to this value when it comes to very low Q.

Thank you for this well-thought-out advice. To address this comment, we have put a lot of thought into the ideal N_{ret}-Q relationship and are now confident that we can present a solution that certainly improves the quality of our manuscript. To ensure that the retention capacity does not exceed 1, we changed it to an asymptotical approach, just as suggested, using the following systematic derivation:

According to the general framework of the C-Q relationship, we assume that discharge (Q) is log-normally distributed and related to nitrate loads (L) as:

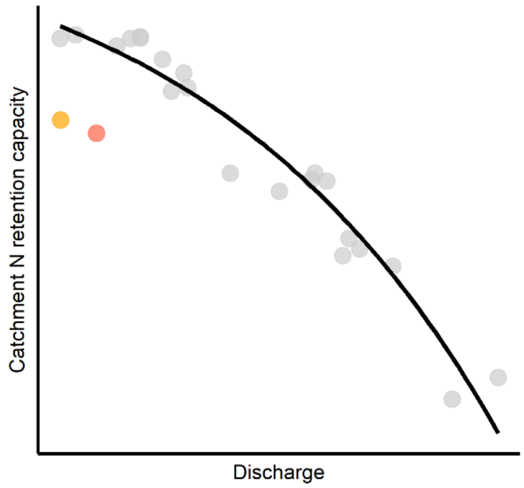
$$L = \alpha Q^{\beta+1}, \qquad (1)$$

with β+1 being the L-Q slope. With $N_{ret} = 1 - N_{out}/N_{in}$ and assuming N input (N_{in}) to be constant, N retention (N_{ret}) can be calculated as:

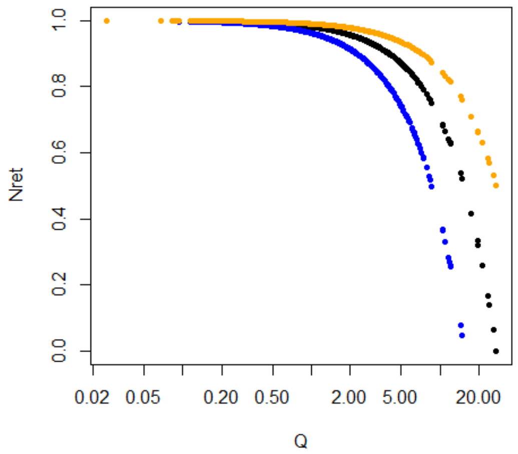
$$N_{ret} = 1 - \frac{\alpha Q^{\beta+1}}{Nin} \qquad (2)$$

Therefore, Nret ~ $Q^{\beta+1}$ can be described as a non-linear relationship, and Nret ~ Q is linear if $\beta=0$ or, in other words, if the L-Q slope equals 1. The result is an N_{ret}-Q relationship that asymptotically approaches 1, and that is zero if $L = N_{in}$.

We will use the fitted L-Q slopes for the Selke sub-catchments (1.14, 1.21, and 1.01 for SH, MD, and HD, respectively) to fit individual N_{ret}-$Q^{\text{L-Q slope}}$ relationships, as can be seen here in the example of the discussion plot:



Just as foreseen by the reviewer, the new shape matches better N retention at low discharge, especially in the lower Selke. Even more so for the soil leachates, where even under the 2018-2019 drought, values are relatively close to the new long-term N_{ret}-Q relationship. On the other hand, the sensitivity of N retention towards changes in load export strongly decreases towards low Q, as can be derived from eq. 1. Considering the low sensitivity at low discharge, the decreased N retention under drought at the upper Selke is even more remarkable. However, at the lower Selke, any drought-induced reduction in N retention is now largely covered by the low sensitivity, especially given the relatively high N input. The following graph shows a theoretical example of N retention given the same Q and the same N input but one scenario with doubled nitrate loads (blue) and one time with halved nitrate load export (orange):



In this case, drought-induced increases in nitrate leaching from the soils (relative to Q) are more telling if depicted as a simple L-Q plot, which we have therefore added to the Supplement. Hence, we decided to make all suggested corrections but to additionally move the $N_{ret\text{-}soil}$-Q panels of Figure 4 (a-c) to the Supplement (Figure S4). We have based our Discussion more on the catchment N retention (in line with the comment of reviewer #1) and on the decreased rates of denitrification and plant uptake.

Returning to the main comment of reviewer 2: We have changed the shape of the N_{ret}-Q relationship throughout our manuscript and adapted our Method and Results section accordingly.

Minor issues (line by line):

L99: Check throughout the manuscript if you introduce abbreviations, I did not find this for TTs

We added the introduction for TTs and checked the entire manuscript to see if any other introduction for an abbreviation was missing.

L103 (Fig. 1): Colourscale for discharge anomaly could be clearer

We added a grey color for values around average discharge and changed the color for low discharge from orange to red to highlight anomalies and to make their difference clearer.

L128: you speak about multi-year drought, although you only analysed the first two years of this event. I agree that the drought lasted longer, in some areas this event is somehow present up to now. But in your analysis this is a two-year event..

We changed our wording to two-year drought throughout our manuscript.

L204: these numbers refer to the subcatchments? This is not clear..

That is right, and we also agree that it was not clear to the reader. We have now added that these numbers refer to the sub-catchments SH, MD, and HD, respectively.

L292: I understand that more details of the mHM-SAS-model are given in the supporting material. But still I would like to have a couple of sentences also in the main document that explain what a/b ratios and SAS functions are and why they are indicative of water age.

We fully agree. Therefore, we added an additional explanation of SAS functions, the a/b ratio, and their relation to nitrate transport in the Method section from line 189 onwards. For more details on what exactly we have written, we refer to our response to reviewer #1.

L297: a median of a median, really?

We agree with the reviewer that this sounds quite clunky. It is not, as one might assume, a typo. Instead, our rationale is that we show median TTs for the separate drought years and compare it with the median TTs of previous years (Figure 3). In that sense, we found the median of median TTs from previous years the most meaningful number to compare with the median of drought years. The mean would not be appropriate due to the non-normal distribution of median TTs.

L303: you present a conceptual drawback of the model: it cannot handle TTs that are larger than the simulation period. Nevertheless you model fits are quite nice on the entire time series from the very begin of grab sampling. Does this mean that your model is right for the wrong reasons and that you overcome structural deficiencies by calibration? I think this should be a point for Discussion.

Thank you for this important remark. The maximum TT is restricted by the time frame of the simulation rather than the actual age of the oldest water, which is unknown (Nguyen et al., 2022). In order to have sufficiently old ages, to create an initial age distribution in storage, and to minimize the effect of initial conditions, we replicated model input data from the 1993-1996 period ten times and ran the model with these data for warming up. In other words, the model states obtained from this model run were used as initial conditions for the actual model run from 1997-2020. We added this point to the Method section from line 198 onwards.

L339: I think the post-drought nitrate pulses in soils do not only propagate through catchments and find their way to rivers, but that runoff generation processes are altered, too. This has been documented before, i.e. a change of HOF/SOF to more SSF in a forest catchment after drought (https://doi.org/10.1016/j.jhydrol.2012.07.010).

Thank you for this interesting point. We added a discussion of the results from the suggested study to our manuscript. Specifically, we added that "increased concentration variability indicates that biogeochemical and hydrological processes (e.g., runoff generation processes; Lange & Haensler, 2012) within the catchment changed during the drought, affecting the availability and transport of nitrate."

L351, L398: You claim that upper Selke is "very low in nitrate" during droughts. What is "very low"? I think in forest catchments a value of 1 mg/l Nitrate-N is not exceptional. Also here you should compare with other studies in forest streams.

We fully agree that 1 mg L^{-1} is not exceptionally low. Please note that median nitrate concentrations during the drought were 0.2 and 0.4 mg L^{-1} in SH and MD, respectively (line 236). Nevertheless, we agree that the term "very low" is not the best choice here. Therefore, we reframed the sentence to "relatively low" to point out that concentrations are lower than during previous years.

L467: N uptake by denitrification?

Thank you for this remark. We have now specified that we mean N uptake by plants and removal via denitrification.

References

Benettin, P. and Bertuzzo, E.: tran-SAS v1. 0: a numerical model to compute catchment-scale hydrologic transport using StorAge Selection functions, Geosci. Model Dev., 11, 1627–1639, 2018.

Dai, Z., Yu, M., Chen, H., Zhao, H., Huang, Y., Su, W., Xia, F., Chang, S. X., Brookes, P. C., Dahlgren, R. A., and Xu, J.: Elevated temperature shifts soil N cycling from microbial immobilization to enhanced mineralization, nitrification and denitrification across global terrestrial ecosystems, Glob. Change Biol., 26, 5267–5276, https://doi.org/10.1111/gcb.15211, 2020.

Hannappel, S., Köpp, C., and Bach, T.: Charakterisierung des Nitratabbauvermögens der Grundwasserleiter in Sachsen-Anhalt, Grundwasser, 23, 311–321, 2018.

Jutglar, K., Hellwig, J., Stoelzle, M., and Lange, J.: Post-drought increase in regional-scale groundwater nitrate in southwest Germany, Hydrol. Process., 35, e14307, https://doi.org/10.1002/hyp.14307, 2021.

Lindner, M., Maroschek, M., Netherer, S., Kremer, A., Barbati, A., Garcia-Gonzalo, J., Seidl, R., Delzon, S., Corona, P., Kolström, M., Lexer, M. J., and Marchetti, M.: Climate change impacts, adaptive capacity, and vulnerability of European forest ecosystems, For. Ecol. Manag., 259, 698–709, https://doi.org/10.1016/j.foreco.2009.09.023, 2010.

Menzel, A., Estrella, N., and Fabian, P.: Spatial and temporal variability of the phenological seasons in Germany from 1951 to 1996, Glob. Change Biol., 7, 657–666, https://doi.org/10.1111/j.1365-2486.2001.00430.x, 2001.

Nguyen, T. V., Kumar, R., Musolff, A., Lutz, S. R., Sarrazin, F., Attinger, S., and Fleckenstein, J. H.: Disparate Seasonal Nitrate Export From Nested Heterogeneous Subcatchments Revealed With StorAge Selection Functions, Water Resour. Res., 58, e2021WR030797, https://doi.org/10.1029/2021WR030797, 2022.

Rivett, M. O., Buss, S. R., Morgan, P., Smith, J. W., and Bemment, C. D.: Nitrate attenuation in groundwater: a review of biogeochemical controlling processes, Water Res., 42, 4215–4232, 2008.

Rode, M., Halbedel née Angelstein, S., Anis, M. R., Borchardt, D., and Weitere, M.: Continuous in-stream assimilatory nitrate uptake from high-frequency sensor measurements, Environ. Sci. Technol., 50, 5685–5694, https://doi.org/10.1021/acs.est.6b00943, 2016.

Samaniego, L., Kumar, R., and Attinger, S.: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale, Water Resour. Res., 46, https://doi.org/10.1029/2008WR007327, 2010.

Sapkota, A., Haghverdi, A., Avila, C. C. E., and Ying, S. C.: Irrigation and Greenhouse Gas Emissions: A Review of Field-Based Studies, Soil Syst., 4, 20, https://doi.org/10.3390/soilsystems4020020, 2020.

Trauth, N., Musolff, A., Knöller, K., Kaden, U. S., Keller, T., Werban, U., and Fleckenstein, J. H.: River water infiltration enhances denitrification efficiency in riparian groundwater, Water Res., 130, 185–199, https://doi.org/10.1016/j.watres.2017.11.058, 2018.

Winter, C., Lutz, S. R., Musolff, A., Kumar, R., Weber, M., and Fleckenstein, J. H.: Disentangling the impact of catchment heterogeneity on nitrate export dynamics from event to long-term time scales, Water Resour. Res., 57, e2020WR027992, https://doi.org/10.1029/2020WR027992, 2021.