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 the constructive feedback. In the following, we provide a point-by-point response to the reviewers' suggestions. We are confident that all points raised can be appropriately addressed and will help 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 will consider this in the Introduction and also in the Discussion of our work.

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 will introduce SAS functions in the introduction around line 73 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 195 onwards. Specifically, we will add 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 will add 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 a direct validation. Therefore, to address this point, we will calculate and include the propagation of the parametric model uncertainty in our soil N retention metric. In the revised manuscript, we will add this model uncertainty in the form of error bars to the soil N retention capacity in Figure 4 a-c. Furthermore, we will shift the focus from a direct comparison of catchment N vs. soil N retention more towards 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 how to solve it. We will follow this suggestion and take 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 will therefore add information on its potential impact in the discussion of our work (chapter 4.1).

Specifically, we will add that irrigation would increase the soil moisture content and might therefore buffer drought impacts, such as the decrease in plant uptake and denitrification. Moreover, irrigation might lead to mobilization of nitrogen as nitrate, which would otherwise be retained in the upper soil until rewetting in autumn. As such 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. We will therefore refer to “year” instead and emphasize that it starts in May throughout the figures of our manuscript.

Line 160: This is results, not methods.

We will shift that part to the result section.

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 will add more information on SAS functions, the a/b ratio and sub-catchment differences in the NSE efficiencies here.

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 physically-based model 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 interval 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. We will add a short form of this explanation in the section around line 202.

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 will make these points clearer in the revised manuscript.

Line 231: instead – word choice?

We will replace "instead" by "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 will add here information on the NSE. Specifically, simulated nitrate concentrations showed a good fit to the measured sensor data during drought with a NSE of 0.89, 0.88 and 0.79 in SH, MD and HD respectively.

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

We will add this information around line 195.

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 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 will add that information in the discussion part of our manuscript.

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 will adapt our wording by adding site names in parentheses throughout the manuscript.

Line 336-337: This is an incomplete sentence

We will delete the incomplete sentence, as we already elaborate on the seasonality of nitrate export below in line 344.

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 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 line 343 is somewhat an oversimplification and will therefore change it to "high nitrate concentrations following drought can pose a threat to water quality".

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 the 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 for denitrification in the groundwater, not in the soils nor in the stream. We see now that this was not sufficiently clear in the manuscript and will reframe 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 from 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 will add a short form of these arguments to our discussion around line 358.

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 been analyzed already 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 finding 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 artefact 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 that is even larger and therefore integrates more influencing factors. Nevertheless, also our combination of data-driven analysis and process-based modelling showed that the differences between upstream and downstream are quite consistent. We will further elaborate on these sub-catchment specific differences in section 4.2 (Sub-catchment specific contributions to the integral nitrate response to drought; line 392 onwards).

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 will follow them and will i) separate panel A and B from C, ii) place discharge in C in a separate combined panel and iii) shift 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 will correct 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, which represents 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 will do that.

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.

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

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.

Musolff, A., Schmidt, C., Selle, B., and Fleckenstein, J. H.: Catchment controls on solute export, Adv. Water Resour., 86, 133–146, https://doi.org/10.1016/j.advwatres.2015.09.026, 2015.

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.