

Reply: Anonymous Referee #1

General comments

This study definitely addresses an important question in the context of Northern Europe, and provides an interesting answer: catchment-scale change in nitrate reduction can reach 10% and climate changes have a greater effect compared to land use changes. I suggest four main improvements to the manuscript:

- Place the study in a more international context. Most references are from Denmark, where the same models and maps are used. What about other countries in the Baltic area? Although I understand that nitrate reduction maps may not be common in other regions of the world, it would be interesting to do a literature review on nitrate retention and changes in flowpaths as a result of a changing climate and land use. This would greatly improve the introduction and the discussion.

Thank you for this suggestion, we have added a section in the introduction of the linkage of nitrate, Climate change and land use:

It is therefore very relevant to investigate the potential error arising when nitrate reduction maps are assumed to be constant in time. No studies have been reported on that issue. Even as, the link between climate change, land use change and nitrate reduction has been established in previous studies (e.g. Fleck et al., 2017; Mas-Pla and Menció, 2019; Olesen et al., 2019; Ortmeyer et al., 2021; Sjøeng et al., 2009). Ortmeyer et al. (2021) used a water balance model combined with a lumped-parameter nitrate mass model for an area in Germany, finding that nitrate concentrations in the groundwater increased towards the end of the century by up to 89 % as a result of changes in temperature, evapotranspiration and precipitation. Mas-Pla and Menció (2019) found that climate change in turn affects groundwater recharge and thus the dilution of nitrate in the subsurface in a study in Catalonia. While, Paradis et al. (2016) found that new agricultural practices under changing climate conditions led to substantial nitrate increases on an Island in eastern Canada.

Fleck, S. et al., 2017. Is Biomass Accumulation in Forests an Option to Prevent Climate Change Induced Increases in Nitrate Concentrations in the North German Lowland? , 8(6): 219.

Mas-Pla, J., Menció, A., 2019. Groundwater nitrate pollution and climate change: learnings from a water balance-based analysis of several aquifers in a western Mediterranean region (Catalonia). Environ Sci Pollut Res Int, 26(3): 2184-2202. DOI:10.1007/s11356-018-1859-8

Olesen, J.E. et al., 2019. Nitrate leaching losses from two Baltic Sea catchments under scenarios of changes in land use, land management and climate. Ambio, 48(11): 1252-1263. DOI:10.1007/s13280-019-01254-2

Ortmeyer, F., Mas-Pla, J., Wohnlich, S., Banning, A., 2021. Forecasting nitrate evolution in an alluvial aquifer under distinct environmental and climate change scenarios (Lower Rhine Embayment, Germany). Science of The Total Environment, 768: 144463. DOI:https://doi.org/10.1016/j.scitotenv.2020.144463

Paradis, D. et al., 2016. Groundwater nitrate concentration evolution under climate change and agricultural adaptation scenarios: Prince Edward Island, Canada. Earth Syst. Dynam., 7(1): 183-202. DOI:10.5194/esd-7-183-2016

Sjøeng, A.M.S., Kaste, Ø., Wright, R.F., 2009. Modelling future NO₃ leaching from an upland headwater catchment in SW Norway using the MAGIC model: II. Simulation of future nitrate leaching given scenarios of climate change and nitrogen deposition. Hydrology Research, 40(2-3): 217-233. DOI:10.2166/nh.2009.068

- Improve the calibration strategy. This study uses a parameter-rich model, which is calibrated in one single catchment and with a poor evaluation of how the model is simulating interannual climate variability. This type of model is subject to equifinality, and it is important to evaluate it properly to make sure that it provides the right answer for the right reason. For example, would it be possible to evaluate the model in different sub-catchments with different land use % before testing land-use change scenarios? In the same way, evaluate the model in different climatic conditions during the calibration period (making the most of the past interannual climate variability) before testing future climate scenarios?

Thank you for this comment. The section was unfortunately somewhat unclear, and it was not very well described the steps we took during the calibration. We have added more information on and restructured the entire the calibration procedure in the section 3.1-3.3 and inserted additional on the validation in section 4.1.

The model was calibrated in 3 steps, the parameter-rich hydrological model (Mike She), was first calibrated by identifying five sensitive parameters (among these was a single soil parameter) through a sensitivity analysis of 28 free (and 43 tied) parameters. A global search engine (Shuffled Evolution Complex) was used. Thus, limiting the risk of equifinality. Secondly, the sensitive and calibrated soil parameter was thereafter transferred to Daisy. The Daisy model has previously been manually calibrated (a manual calibration is the only way to calibrate this model due to its 1D column formulation) to the catchment, and this setup was used in this study and performance was evaluated after changing the soil parameter calibrated by Mike She. First after the calibration of these two models, the redox interface location was calibrated as a final step, using the nitrate arrival percentage obtained at the downstream station.

It is a very valid and good suggestion to evaluate the performance of the model in different subcatchment with different land use %. The most important land use difference for nitrate is forest versus agricultural areas, but unfortunately this catchment is very agriculturally dominated, and therefore does not contain large enough forest areas to give a strong enough difference in the signal for different subcatchments. Furthermore, the only discharge station close to a larger forested area, is very doubtful as it is highly regulated and therefore measurements are uncertain.

Thank you for pointing out the very relevant comment on testing the model's climate variability using dry and wet periods. This is very good idea. Unfortunately, our studies area has a generally homogeneous precipitation time series during the model simulation. Any relatively low/wet precipitation years are not occurring successively, but are isolated in the time series, making it difficult to select relatively dry/wet periods.

- Show time series of river discharge and nitrate concentration + model fit

We have added a graph showing the main station hydrograph as well as the performance statistics of the calibration and validation for water balance and RMSE in the supplementary material. The way the model is setup there is no time series of nitrate concentrations given as output of the model, but a comparison of the resulting nitrate root zone leaching to other studies has been added in L338. And a comparison of the nitrate load at Kratholm to the measurements has also been added in the text (L355).

- Spatial variability and representativeness of the study area. Why only studying one catchment and not the entire country? Is it a problem of data availability, computation time? If studying more catchments is not possible, I would suggest to present the results in different subcatchment with different land use / soil types to see if they have similar responses (preferably subcatchments with a nitrate monitoring station). This would also help assess whether the <10% change in nitrate reduction is big or not compared to current spatial variability.

Due to both computational time and data availability it was not possible to run the setup for more than one catchment. Especially Daisy modelling is very time consuming as all combinations of climate/land use and soil type must be setup as an individual 1D model, which then inputs to the different grids in Mike She.

Unfortunately, it is not straightforward to evaluate results from different sub-catchments. Most of the sub-catchments are very similar in relation to land use and soil type, and there are no nitrate station data available within the catchment to support the validity of reduction percentage on a subcatchment level. We have, however, added some reflection on the importance of the 10 % change in nitrate reduction compare to other studies of mitigation effect see comment for L20.

Specific comments L9:

“Nitrate reduction maps have been used routinely in Northern Europe for calculating efficiency of remediation measures and impact on climate change on nitrate leaching and are as such valuable tools for policy analysis and mitigation targeting.” This sentence is too long.

Good point. This sentence has been split into: *Nitrate reduction maps have been used routinely in Northern Europe for calculating efficiency of remediation measures and impact of climate change on nitrate leaching. These maps are therefore valuable tools for policy analysis and mitigation targeting.*

L11&14 “Nitrate maps . . .” -> “nitrate reduction maps”

Thank you. This has been corrected.

L20 “The study, however, also showed that the reductions maps are products of a range of complex interactions and that the combination of the choices made for selected scenarios, model formulations and assumptions are critical for the resulting span in reduction capability.” Sentence too long and unclear. Suggest to discuss whether these differences <10% are important to consider for management. How big is this 10% difference in comparison to the spatial variability across different regions of Denmark?

Yes, this is not a clear sentence. We have changed it according to the suggestion and added reflections on the importance for management in L19: *The study, however, also showed that the reductions maps are products of a range of complex interactions between water fluxes, nitrate use and timing. What is also important to note, is that the choices made for future scenarios, model setup and assumptions may affect the resulting span in reduction capability. To account for this uncertainty multiple approaches, assumptions and models could be applied for the same area, however as these models are very time consuming this is not always a feasible approach in practice. An uncertainty in the order of 10% on the reduction map may have major impacts on practical water management.*

L593: *10% error on the reduction map may potentially have major impacts on practical water management. Considering for instance the baseline scenario in Table 4, where the average N-reductions vary between 55% and 67% reduction, this implies that the net impact of a 100 kg N reduction in leaching from the root*

zone will vary between 45 kg and 33 kg (i.e. 30%). Such changes are larger than the effects of sophisticated mitigation measures (Hansen et al., 2017).

Hansen AL, Refsgaard JC, Olesen JE, Børgesen CD (2017) Potential benefits of a spatially targeted regulation based on detailed N-reduction maps to decrease N-load from agriculture in a small groundwater dominated catchment. *Science of the Total Environment*, 595, 325-336.

L29 “depending on the actual hydrobiogeochemical conditions the removal may mainly occur in groundwater or in surface water systems such as lakes or wetlands” Reference needed. Please add a paragraph with references showing that the retention process studied in this paper is a dominant process in the context of the study.

Thank you for this comment, we have added a reference to sentence mentioned above:

It can be expressed as a percentage removal and depending on the actual hydrobiogeochemical conditions the removal may mainly occur in groundwater or in surface water systems such as lakes or wetlands (Huno et al., 2018; Quick et al., 2019).

Huno, S.K.M., Rene, E.R., van Hullebusch, E.D., Annachhatre, A.P., 2018. Nitrate removal from groundwater: a review of natural and engineered processes. *Journal of Water Supply: Research and Technology-Aqua*, 67(8): 885-902. DOI:10.2166/aqua.2018.194

Quick, A.M. et al., 2019. Nitrous oxide from streams and rivers: A review of primary biogeochemical pathways and environmental variables. *Earth-Science Reviews*, 191: 224-262.
DOI:<https://doi.org/10.1016/j.earscirev.2019.02.021>

And a reference in L84, about the dominant process in the study area: *The reduction in this catchment has previously been shown to be dominated by saturated zone reduction processes (Hansen et al., 2009).*

Hansen, J.R. et al., 2009. An integrated and physically based nitrogen cycle catchment model. *Hydrology Research*, 40(4): 347-363.

L75 “one of the best nutrient time series in Denmark, providing a long and near-complete data set” Please specify length, frequency.

More information on the station has been added: *The discharge station at Kratholm has one of the best nutrient time series in Denmark starting in the 1980s, with near-daily sampling from 1989 (Windolf et al., 2016). The station, therefore, provides a long and near-complete data set for nutrient modelling as well as an extensive water discharge time series (Trolle et al., 2019).*

Trolle, D. et al., 2019. Effects of changes in land use and climate on aquatic ecosystems: Coupling of models and decomposition of uncertainties. *Science of The Total Environment*, 657: 627-633.
DOI:<https://doi.org/10.1016/j.scitotenv.2018.12.055>

Windolf, J. et al., 2016. Successful reduction of diffuse nitrogen emissions at catchment scale: example from the pilot River Odense, Denmark. *Water science and technology : a journal of the International Association on Water Pollution Research*, 73(11): 2583-9. DOI:10.2166/wst.2016.067

L76 “The average discharge amounts to 4.4m³/s and the load is approximately 14 kg NO₃-N/ha/year” what period? Any trend?

More information has been added here: *In 2005-2009, the average discharge amounts to 4.6m³/s and the load is approximately 14 kg NO₃-N/ha/year. A decreasing trend in nitrate loads has been observed previously during 2000-2013 by Windolf et al. (2016), possibly due to implementation of mitigation measures in the catchment.*

Windolf, J. et al., 2016. Successful reduction of diffuse nitrogen emissions at catchment scale: example from the pilot River Odense, Denmark. *Water science and technology : a journal of the International Association on Water Pollution Research*, 73(11): 2583-9. DOI:10.2166/wst.2016.067

L 77 “The geology is mainly a result of previous glaciations like till deposits. Aquifers are generally confined and the phreatic groundwater tables are shallow.” Move next to the sentence about soil types.

The sentence has been moved.

L79 “There were 226 measurements of the redox depth in the area” Please briefly present the method to measured redox depth.

Thank you, a sentence has been added on this: *226 measurements of the redox depth are available from boreholes in the area, the redox depths were mainly interpreted based on sediment color as described by e.g. Ernstsens and Mørup (1992), and a few by measurements of reduced compounds.*

Ernstsens, V., Mørup, S., 1992. Nitrate reduction in clayey till by Fe(II) in clay minerals. *Hyperfine Interactions*, 70(1): 1001-1004. DOI:10.1007/BF02397497

L114 “Calibration was carried out against data from four discharge stations and 455 groundwater wells from the period 2004-115 2007, see Karlsson et al. (2016) for details” please present the calibration strategy, objective function, parameter exploration algorithm, etc. did you calibrate Daisy and MIKE SHE together or separately?

Thank you for this comment. The method section is indeed somewhat confusing. We have tried to structure it more clearly by moving section and paragraphs and adding some introductory sentences (L121-L137). We have also added more information in the entire section.

L139 “Each time the accumulated input of nitrate reach 0.5 kg N” is it 0.5 kg/ha, or kg/km²?

This has been clarified in the paper: *Each time the accumulated input of nitrate reach 0.5 kg N within the model cell (200mx200m), a particle is released from the water table and is allowed to follow the groundwater flow.*

L 165 “subsequently compared with the measured redox depth in boreholes” where is this comparison?

This line has been deleted. See reply for L251.

L215 “More information on the land use scenarios can be found in Olesen et al. (2014) and Karlsson et al. (2016).” Please provide a summary how these scenarios were built.

Thank you for pointing out this missing part. We have added more information in line 289-293: *The land use scenarios were created during workshops with researcher, farming industries, environmental protection agencies and government representatives. During the workshops, participants identified possible paths of developments for the land use in Denmark considering the balance of agricultural marked value on one side and priorities in the society on the other (e.g., environmental concerns or recreational use). From the workshop four scenarios that describe agricultural management in the period 2080-2099 was created,*

[I241 “For all the crops in the region the DAISY model is able to reproduce the observed harvested N” performance in evaluating the spatial variability in harvested N?](#)

Thank you for this comment. Unfortunately, there is no information available to evaluate the harvested N spatially, as information on this is only registered on a regional level. Therefore, we cannot test the Daisy performance on the spatial variability.

[I251 “To select the most appropriate combination, the cumulative distribution of the resulting redox depth of a given parameter combination is plotted against the distribution of observed redox depths. The observed redox depth is both compared to the simulated values at the actual point of observational measurement as well as to the total fractional distribution of the whole catchment \(not shown\). Based on this analysis the best combination with the correct NAP was found to be \$f=0.01\text{yr}\$ and \$\text{min.redoxdepth}=3\text{m}\$.” I found this difficult to understand, add a figure in SI? This paragraph should be moved to Materials & methods section](#)

Thank you, yes this is unclear. The paragraph above has been moved to section 3.4. We have also added some more explanation (L226-239): *As the calibration of these two parameters may result in non-uniqueness, all possible combinations (realisations) of the two parameters resulting in observed NAP, are identified. For all realisations the cumulative distribution of the redox depth is found at the location, where observations of redox depth are available from boreholes, as well as the cumulative distribution of the entire catchment. These two graphs are subsequently compared with the cumulative distribution of the actual measured redox depth in boreholes. The realization with the best representation of the fractional distribution of the observed redox depth for both on-site and especially catchment scale is chosen for the final redox depth parameters.*

The reason for comparing calculated redox depths to cumulative distributions for actual measurement locations and the entire catchment distribution is due to several issues. First, measured redox depths are very local point measurements, and large variations in space (within a few meters) are often reported (e.g. Ernsten, 1996; Hansen et al., 2008), and a measurement may not be representable for the area or model scale, where numerous measurements together are more likely to represent to the correct fractional distributions in the catchment. Furthermore, the calculated redox depth may be applicable on catchment scale, on which scale it is also calibrated, but less trustworthy on location scale.

Ernsten, V., 1996. Reduction of Nitrate By Fe²⁺ in Clay Minerals. *Clays and Clay Minerals*, 44(5): 599-608. DOI:10.1346/CCMN.1996.0440503

Hansen, J.R., Ernsten, V., Refsgaard, J.C., Hansen, S., 2008. Field scale heterogeneity of redox conditions in till-upscaling to a catchment nitrate model. *Hydrogeology Journal*, 16(7): 1251-1266. DOI:10.1007/s10040-008-0330-1

The remaining paragraph in L251 is moved as suggested.

[I255 “the depth to the redox interface is fairly shallow” avoid “fairly” in a technical paper.](#)

Thank you, this has been corrected.

L480 “also worth noting, that all combinations of land use and climate change scenarios may not be equally likely or even plausible in the future” Interesting remark. Please add a reference.

This is more a reflection on the fact that as climate changes in one way or the other there may be some political and financial strategies that are more likely to occur than others. We have no reference for this, and have therefore modified the sentence, so that it now states: *However, one could speculate that not all combinations of land use and climate change scenarios may be equally likely or plausible in the future, as decisions on land use application made by local farmers or through national regulations are made concurrently to adapt or mitigate changes in the climatic conditions.*

L511 “The indication that errors can be up to 10% is based on only a single case study with one catchment”. Wasn’t it possible to do the same study in several catchments of DK? If not, is it an issue of data availability? Computation time? If doing the same analysis in other catchments is not possible, I would suggest presenting the N retention in different subcatchments with different land use / soil type / topography

See response to main comment 4.