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

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.

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).

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.

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.*

Thank you. This has been corrected.

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. Ernstsen and Mørup (1992), and a few by measurements of reduced compounds.*

Ernstsen, 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,*

l241 "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 a 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.

l251 "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=0.01yr and min.redoxdepth=3m." 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. Ernstsen, 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.

Ernstsen, 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., Ernstsen, 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.

l255 "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.

Reply:

In the manuscript, the authors investigate the impact of changes in climate and land use on maps of nitrate reduction in groundwater in a Danish catchment. Such maps are important tools to support management strategies that deal with nitrate pollution. Therefore it is highly relevant to investigate the potential error made in current practices that use static maps and thus neglect the effect of changes in climate and land cover. This study compares maps of nitrate reduction in groundwater produced for different climate and land use scenarios within a modelling framework. I think that a number of point need to be clarified and examined to ensure the robustness and reproducibility of the results. I provide here a summary of my main concerns:

1) More information on the data and methods used is needed. In particular, the manuscript refer to numerous past studies for the data and methodology, which makes it difficult for the reader to have a clear understanding of the data and methods. The authors should provide in the manuscript a summary description of all data and method used (they can then refer to past studies for more details).
Thank you for this comment. We have added new information on the methodology and restructured the entire method section. We have also added more information on the data and the data sources in the data section (2. Study site).

2) The calibration includes a number of 'manual' adjustments to the parameter values and identify a single parameterization. I think that it would be valuable to account for the uncertainty in the model parameter values and to determine whether the changes observed in the nitrate reduction maps due to changes in the climate are appreciable given the uncertainty due to parameter values. Given the presumably large number of calibrated parameters, the issue of equifinality is likely to arise, i.e. combinations of different parameter values could lead to the same model performances, but produce different nitrate reduction maps. In particular, only groundwater parameters are adjusted to match the Nitrate Arrival Percentage (NAP), and the value of these groundwater parameters could compensate for deficiencies in the values of the soil parameters (in particular soil denitrification parameters). I think this should be at least discussed in more details in the manuscript and I refer e.g. to Wade et al. (2008).
Thank you for this comment. Yes, the section was unfortunately somewhat unclear and it was not very well described the steps we took during the calibration to ensure equifinality. We have added some extra introductory sentences to the method section (L121-L137), as well as restructured the entire section and changed the headlines. We have also added more information on the calibration procedure in the section 3.1-3.3 and on the validation in section 4.1, and an additional figure (Figure S1) in the supplementary material showing the stream hydrograph and performance.

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.

3) From the manuscript, I understand that, in the model, tile drains can be located in non-agricultural areas, which I find surprising. Some explanation on this are required, since tile drains appear to have a large impact on the model results.

Yes, this true. We did not explain this. We have added text on this issue in L145-149:

The study is heavily tile drained, and it can be assumed that drainage will always be present when it is need in the agricultural areas. However, the actual site-specific location of tile drains are unknown and therefore drains are specified across the entire catchment at a depth of -0.5 meters . Drain flow is however only activated when groundwater level rise above drain level. Apart from representing tile drainage the drainage system also represents small ditches and stream, too small to incorporate into the river system following the approach of Troldborg et al. (2010).

Troldborg, L. et al., 2010. DK-model2009 - Modelopstilling og kalibrering for Fyn. GEUS Report.

I provide below detailed comments.

ABSTRACT

L20-22 'Th study, however, [...] in reduction capability: ' this sentence needs to be revised. 'complex interactions' is vague and the analyses presented in the manuscript do not explore the effect of model formulations on the nitrate reduction maps.

We have reformulated the sentence: *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.*

SECT. 1 (INTRODUCTION)

L42: I suggest citing and discussing the study by Knoll et al. (2020), that establishes a map of groundwater redox conditions for Germany through machine learning, and the study by Tesoriero (Tesoriero et al., 2015), that investigates the redox conditions in groundwater in the Chesapeake Bay watershed in the USA.

Thank you, we have added these references. L49-54: *Heterogeneities in geology and drainage systems are responsible for substantial local spatial variations in nitrate reduction. However, the spatial variation of nitrate reduction in the groundwater system has so far only been investigated in a handful of studies (e.g. Højberg et al., 2015a; Knoll et al., 2020; Kunkel et al., 2008; Merz et al., 2009; Tesoriero et al., 2015; Wriedt and Rode, 2006). Different approaches have been used in these studies from nitrate groundwater modelling (Højberg et al., 2015b; Merz et al., 2009; Wriedt and Rode, 2006), data driven machine learning (Knoll et al., 2020) or statistical modelling (Tesoriero et al., 2015).*

L59-60: 'A severe problem [. . ..] resulting flow pathways.': This statement should be better explained and supported by some reference.

We have reformulated this sentence and added a reference:

A severe problem in this respect is, however, that the nitrate reduction maps may not be constant in time as the reduction taking place at a given location depend on resulting flow pathways (Hansen et al., 2014b).

SECT. 2 (STUDY SITE)
L 75: Please add further details on the type, characteristics (such as frequency) of the nutrient data.

Thank you, this has been added in L98-L104: *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). 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.*

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

L81: A definition of the criteria to identify the redox depth (i.e. to separate aerobic from anaerobic conditions) is missing.

This information has been added, L105-106: *226 measurements of the redox depth are available from boreholes in the area, the redox depths were mainly interpretated based on sediment colour as described by e.g. Ernstsen and Mørup (1992), and a few by measurements of reduced compounds.*

Ernstsen, 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

SECT. 3 (METHODS)
- The authors need to justify their choice of coupling the Daisy and MIKE SHE model. Why not using one model or the other? Why Daisy/MIKE SHE are particularly appropriate for this study?

Thank you for this comment. We have added some reflections on this in L 124-130: *Both Mike She and Daisy have been used extensively in the danish area, and Mike She forms the basis of the national nitrate and groundwater model (Bruun et al., 2003; Hoang et al., 2010; Højberg et al., 2015a; Højberg et al., 2010; Højberg et al., 2013; Højberg et al., 2015b; Troldborg et al., 2010). Mike She is a fully coupled integrated groundwater-surface water model and this integration is important for assessing the feedback between unsaturated and saturated zone, especially under changing climate. However, Mike She does not simulate crops development and nitrate leaching from the root zone, and therefore information from an agrological model, like Daisy, is necessary.*

Bruun, S., Christensen, B.T., Hansen, E.M., Magid, J., Jensen, L.S., 2003. Calibration and validation of the soil organic matter dynamics of the Daisy model with data from the Askov long-term experiments. Soil Biology and Biochemistry, 35(1): 67-76. DOI:http://dx.doi.org/10.1016/S0038-0717(02)00237-7

Hoang, L. et al., 2010. Comparison of the SWAT model versus DAISY-MIKE SHE model for simulating the flow and nitrogen processes. In: conference, T.I.S. (Ed.).

Højberg, A.L. et al., 2015a. En ny kvælstofmodel. Oplandsmodel til belastning og virkemidler. Metode rapport (A new nitrogen model. Catchment model for loads and measures. Methodology Report – In Danish). DOI:Available from http://www.geus.dk/DK/water-soil/water-cycle/Documents/national_kvaelstofmodel_metoderapport.pdf

Højberg, A.L. et al., 2010. DK-model2009 – Sammenfatning af opdateringen 2005-2009, Geological Survey of Denmark and Greenland.

Højberg, A.L., Troldborg, L., Stisen, S., Christensen, B.B.S., Henriksen, H.J., 2013. Stakeholder driven update and improvement of a national water resources model. Environmental Modelling & Software, 40(0): 202-213. DOI:http://dx.doi.org/10.1016/j.envsoft.2012.09.010

Højberg, A.L. et al., 2015b. National kvælstofmodel - Oplandsmodel til belastning og virkemidler. Metode rapport, revideret udgave september 2015 (National nitrogenmodel - Catchment model for load and measures. Method report, revised version September 2015). DOI:ISBN 978-87-7871-418-3

Troldborg, L. et al., 2010. DK-model2009 - Modelopstilling og kalibrering for Fyn. GEUS Report .

- It is also not clear which parameters are calibrated and for this I think that a table that summarizes the model parameters and their calibrated value should be added (in the main text or in the supplementary information).

Thank you for this comment. The calibration and parameters have already been reported in detail elsewhere and we would therefore like to avoid going into detail with the parameters, so the the paper do not get too long. We, however, do acknowledge that some information is needed here. We have therefore added a general statement of the parameter types included in the calibration at L157: *After a sensitive analysis on 28 free parameters with 43 tied parameters; a total of five parameters were chosen for calibration, of these, one soil parameter in the unsaturated zone, one drainage parameter and three saturated zone parameters.*

- What are the nitrogen inputs to the system? From L124 I understand that in cropland areas mineral fertilizers only are considered, is this correct? What about the N input for the areas with other land uses such as grass or forest? What about nitrogen biological fixation and nitrogen atmospheric deposition which can also be important inputs of nitrogen to soils?

Thank you for the comment, yes that was indeed not clear. We have expanded this explanation in L199-208: *Daisy also simulates nitrate leaching for each soil column that represents a unique combination of soil type, climate, crop rotation and groundwater depth. Crops are fertilized with mineral and organic nitrogen dependent on the farm type and soiltype. The crop recommended nitrogen rate for the years 2004-2007 was used to setup the fertilization scheme. Nitrate leaching input are simulated on daily basis based on the leaching from the permutated crop rotations simulated for the dominating soil type within a 200m x 200 m square grid (Karlsson et al., 2016). Because of the close feedback mechanism between nitrogen yields and nitrate leaching, the simulated mean nitrogen yields were recalibrated to observed annual mean nitrogen yields on Funen (Statistikbanken, 2015) for the dominating soil type for the period 2004-2007. The calibration is conducted by adjusting the crop parameters, following the methodology of Styczen et al.*

(2004). Nitrogen concentrations of yields were extracted from table values of mean nitrogen contents for different crops (Møller et al., 2005). For crop rotations including clover grass and peas nitrogen biological fixation is calculated using Høgh-Jensen et al. (2004) and nitrogen atmospheric deposition is included as input to the soil using standard Daisy settings for dry and wet deposition (Hansen et al., 2012).

Hansen, S., Abrahamsen, P., T. Petersen, C., Styczen, M., 2012. Daisy: Model Use, Calibration, and Validation. Transactions of the ASABE, 55(4): 1317. DOI:https://doi.org/10.13031/2013.42244

Høgh-Jensen, H., Loges, R., Jørgensen, F.V., Vinther, F.P., Jensen, E.S., 2004. An empirical model for quantification of symbiotic nitrogen fixation in grass-clover mixtures. Agricultural Systems, 82(2): 181-194. DOI:https://doi.org/10.1016/j.agsy.2003.12.003

Karlsson, I.B. et al., 2016. Combined effects of climate models, hydrological model structures and land use scenarios on hydrological impacts of climate change. Journal of Hydrology, 535: 301-317. DOI:http://dx.doi.org/10.1016/j.jhydrol.2016.01.069

Møller, J. et al., 2005. Fodermiddeltabel - Sammensætning og foderværdi af fodermidler til kvæg. 64.

Statistikbanken, 2015. Statistical regional registrated annual mean yields. (In Danish) https://www.statistikbanken.dk/jord3.

Styczen, M. et al., 2004. Standardopstillinger til Daisy-modellen. Vejledning og baggrund, Institut for Vand og Miljø, DHI.

- The grid resolution for Daisy/MIKE SHE needs to be clearly defined. Is it 200m x200m (L 126)?
See comment above, we have added more information on this in the cited section (L203).

- L110-111 (and also L304-305): aren't tile drains usually located in agricultural areas? My understanding is that here tile drains are distributed uniformly independently of the land use. This is an important model assumption, since it appears that tile drainage has a large impact on the study results.
See response to main comment 3)

- L115 '455 groundwater wells': Which data were derived from the groundwater wells?
Thanks, we added L153: *and 455 groundwater wells with hydraulic head measurements from the period 2004-2007*

- L119-120 'following methods proposed by Allen et al. (1998) and Styczen et al. (2014)': it is required to add more explanation here (brief description of the methods, parameters that are calibrated with the methods). I also have the same comment regarding L128-129.
Thank you. Yes, the section is very unclear. The complete section on Daisy modelling and calibration have been rewritten: *The Daisy model setup for the Odense is contains on roughly 12,000 1D Daisy columns, and the water balance module is based on a previous calibration of the catchment (Børgesen et al., 2013), where root zone leaching and groundwater abstraction is compared with river discharge (Børgesen et al., 2013; Refsgaard et al., 2011). The model is setup so that each column represent unique combinations of soil type, climate, crop rotation and groundwater depth. The Daisy model uses the same climate input and soil parameter setup as MIKE SHE and the sensitive and calibrated unsaturated soil parameter from MIKE SHE were therefore transferred to Daisy. A more detailed description of the Daisy setup can be found in Karlsson et al. (2016). The water balance performance of Daisy was evaluated in the same calibration (2004-2007) and validation periods (2000-2003/2008-2009) as MIKE SHE.*

Daisy also simulates nitrate leaching for each soil column that represents a unique combination of soil type, climate, crop rotation and groundwater depth. Crops are fertilized with mineral and organic nitrogen dependent on the farm type and soiltype. The crop recommended nitrogen rate for the years 2004-2007 was used to setup the fertilization scheme. Nitrate leaching input are simulated on daily basis based on the leaching from the permutated crop rotations simulated for the dominating soil type within a 200m x 200 m square grid (Karlsson et al., 2016). Because of the close feedback mechanism between nitrogen yields and nitrate leaching, the simulated mean nitrogen yields were recalibrated to observed annual mean nitrogen yields on Funen (Statistikbanken, 2015) for the dominating soil type for the period 2004-2007. The calibration is conducted by adjusting the crop parameters, following the methodology of Styczen et al. (2004). Nitrogen concentrations of yields were extracted from table values of mean nitrogen contents for different crops (Møller et al., 2005). For crop rotations including clover grass and peas nitrogen biological fixation is calculated using Høgh-Jensen et al. (2004) and nitrogen atmospheric deposition is included as input to the soil using standard Daisy settings for dry and wet deposition (Hansen et al., 2012).

Børgesen, C.D. et al., 2013. Udviklingen i kvælstofudvaskning of næringsstofoverskud fra dansk landbrug for perioden 2007-2011. Evaluering af implementerede virkemidler til reduktion af kvælstofudvaskning samt en fremskrivning af planlagte virkemidlers effekt frem til 2015, DCA - Nationalt Center for Fødevarer og Jordbrug, Tjele, Denmark.

Hansen, S., Abrahamsen, P., T. Petersen, C., Styczen, M., 2012. Daisy: Model Use, Calibration, and Validation. Transactions of the ASABE, 55(4): 1317. DOI:https://doi.org/10.13031/2013.42244

Høgh-Jensen, H., Loges, R., Jørgensen, F.V., Vinther, F.P., Jensen, E.S., 2004. An empirical model for quantification of symbiotic nitrogen fixation in grass-clover mixtures. Agricultural Systems, 82(2): 181-194. DOI:https://doi.org/10.1016/j.agsy.2003.12.003

Karlsson, I.B. et al., 2016. Combined effects of climate models, hydrological model structures and land use scenarios on hydrological impacts of climate change. Journal of Hydrology, 535: 301-317. DOI:http://dx.doi.org/10.1016/j.jhydrol.2016.01.069

Møller, J. et al., 2005. Fodermiddeltabel - Sammensætning og foderværdi af fodermidler til kvæg. 64.

Refsgaard, J.C. et al., 2011. Vandbalance i Danmark - Vejledning i opgørelse af vandbalance ud fra hydrologiske data for perioden 1990-2010, Copenhagen, Denmark.

Statistikbanken, 2015. Statistical regional registrated annual mean yields. (In Danish) https://www.statistikbanken.dk/jord3.

Styczen, M. et al., 2004. Standardopstillinger til Daisy-modellen. Vejledning og baggrund, Institut for Vand og Miljø, DHI.

- L121 'such that they produce similar actual evapotranspiration and stream flow for the simulation period': a precise definition of what is meant by 'similar' is needed.

See comment to L119-120

- L134 'the procedure described in [. . .]: Please summarize the procedure.

See comment to L119-120

- L165 'compared with the measured redox depth in boreholes': a description of this comparison is missing. Evt. Supmat.

Thank you, yes this is unclear. We have 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. Ernstsen, 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.

Ernstsen, 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., Ernstsen, 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

- L179-181: I do not understand this statement (in particular it is not clear to me what 'direct arrival percentage' means).
We have reformulated this sentence to make it more clear: *If this assumption is valid the calculation the reduction potential in each grid cell is the same with/without the stuck particles. Unfortunately, this assumption may not always be valid. Furthermore, the arrival percentage estimated by the two methods are not the same as not all particles are released in the complex particle arrival count, the data from which is the only way to calibrate the nitrate model. For the two methods to be comparable it is therefore necessary to exclude the particles that are stuck in the unsaturated zone.*

- Table 1: Please define reference evapotranspiration and specify how it was calculated.
The reference evapotranspiration (also sometimes denoted as potential evapotranspiration) as referred to in Table 1, is calculated based on variables from the climate models. We have added information on how it was calculated in L276-277: *The reference evapotranspiration is calculated using FAO Penman– Monteith formula based on the climate model outputs for temperature, radiation, water vapour wind speed and water pressure.*

- L215-216: I would suggest to briefly summarize how these scenarios were established and to refer to Oleson et al. (2014) and Karlsson et al. (2016) for more details as currently done.
Thank you. This has been added in L289-293: *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,

- L222 '50 model simulations': In table 3 I see only 45 and not 50 scenarios, please clarify.

Yes this is a mistake, it has been corrected to 45.

SECT. 4 (RESULTS)

- L323: 'net precipitation' should be defined.

This has been added in L413: *Hence, the overall net precipitation (precipitation-actual evapotranspiration) may change slightly across the climate models,*

- L453-455: It is not clear to me how these numbers are derived from Figure 9.

Thank you, no this is unclear. We have specified in the text: *Even though it is not possible to completely separate the signal of these three components, a cautious estimation can again be achieved by subtracting the blue bars with the red bar result from LU0 in the reference period, so that the signal from the climate model bias is tried to be removed.*

SECT. 5 (DISCUSSION AND CONCLUSIONS)

- L480: replace 'the full range' by 'a range' as a limited number of scenarios were used, which cannot comprehend all possible futures.

This has been corrected as suggested.

- L515-516 'that should be addressed along with other known sources of uncertainty such as climate model projections, land use projections and hydrological model structure uncertainty.': This discussion needs to be expanded. In particular, uncertainty in model parameter values can also affect the results.

We have expanded this section L710-719: *The indication that errors can be up to 10% is based on only a single case study with one catchment, one model and a limited number of land use and climate change scenarios. While similar results may be found when applying the same approach for catchments governed by the same dominant flow processes and land use types, like the one investigated in this study. The error must be expected to be site and context specific and therefore causes projection uncertainties that should be addressed along with other known sources of uncertainty such as climate model projections, land use projections, parameter uncertainties, geological uncertainty, and hydrological model structural uncertainty (Hansen et al., 2014; Karlsson et al., 2016). Furthermore, during calibration and setup of the model assumptions must be made, adding to the uncertainty. Parameter estimation are here done in a stepwise fashion, and the catchment scale calibration of Daisy along with a particle tracking approach limits the evaluation of performance of the nitrate component to mean catchment figures, and the dynamic of the nitrate system is thus impossible to verify. To account for this a full solute transport solution would be necessary but was unfortunately not possible in the framework of this study; but would be a relevant next step in investigating uncertainties and improve model verification.*

Hansen, A.L., Gunderman, D., He, X., Refsgaard, J.C., 2014. Uncertainty assessment of spatially distributed nitrate reduction potential in groundwater using multiple geological realizations. Journal of Hydrology, 519, Part A: 225-237. DOI:http://dx.doi.org/10.1016/j.jhydrol.2014.07.013

MINOR EDITS:
- L143 'section 0': please add the correct section number.
This has been added

- L170 and in the figure captions: add 'cell' after grid.
Corrected

- L230 'no 1': do the authors refer to scenario 1 in Table 3? Please clarify.
Yes, it refers to scenario 1. We have added more explanation: *Baseline: The term baseline refers to results from the specific model run combination (scenario number1, Table 3), where current land use scenario (LU0) is combined with the observational climate data.*

- L236: replace 'm2/s' by 'm3/s'.
This has been corrected.

- L345: remove 'impact' after 'change'.
This has been corrected.

- L348: there is something wrong here. Maybe 'different' need to be removed?
Some words are missing, we have added: *different due to*

- L364-365: Please correct by 'the change in the drain flow fraction' (two occurrences).
This is corrected

- L397 'this is also found for one of the models': remove 'also' (possibly replace by 'indeed').
This is corrected

REFERENCES:
Knoll, L., Breuer, L., & Bach, M. (2020). Nation-wide estimation of groundwater redox conditions and nitrate concentrations through machine learning. Environmental Research Letters, 15(6). https://doi.org/10.1088/1748-9326/ab7d5c
Tesoriero, A. J., Terziotti, S., & Abrams, D. B. (2015). Predicting Redox Conditions in Groundwater at a Regional Scale. Environmental Science and Technology, 49(16), 9657–9664. https://doi.org/10.1021/acs.est.5b01869
Wade, A. J., Jackson, B. M., & Butterfield, D. (2008). Over-parameterised, uncertain "mathematical marionettes" - How can we best use catchment water quality models? An example of an 80-year catchment-scale nutrient balance. Science of the Total Environment,

400, 52–74. https://doi.org/10.1016/j.scitotenv.2008.04.030

The authors investigate how nitrate reduction is affected by climate and land use changes for one Danish catchment. This is important as nitrate reduction maps are usually considered to be constant in time, which might not be appropriate for water quality management. With a modelling approach using MIKE SHE and Daisy, the authors show that climate has a stronger impact on nitrate reduction.

General comments:
- The introduction is too brief; it should inform why land use and climate changes are relevant to NO3-reduction. This is not explicitly mentioned. The key word "denitrification" also needs to be included.

Thank you for pointing out this issue. We have included the denitrification term and added a paragraph on climate change/land use influences on reduction in L73-80: *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

Thank you for this comment. Yes, the section was unfortunately somewhat unclear. We have restructured and added extra information to the method section and changed the headlines, as well as added some extra introductory sentences to the method section (L121-L137). The reason for defining the period terms at the end of the chapter is due to the fact that the future scenarios have not yet been introduced in the Study site chapter, and that calibration/validation periods are not the same as the complete observational period.

Thank you for pointing out this weakness. We have restructured and added text to the method section to improve the paper and added information and restructured the results section (4.1) on validation of the model. As suggested, we have moved figures to the supplementary material and added a conclusion (see response to L516).

Thank you, we have gone through the text to streamline definitions.

This sentence has been revised: *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.*

Corrected.

Thank you, this sentence has been revised.

This is a valid and important question, but unfortunately not one that can be answered easily. One approach could be testing multiple model setups or types on the same study area, however due to computational and time limitations this is however not always feasible. We have added a few reflections on this in the abstract: *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. It is therefore important to acknowledge if such errors are deemed acceptable in relation to the purpose and context of specific water management situations.

Introduction:
- L35: "The amount of nitrate reduction occurring in groundwater depends on the flow paths and the depth to the redox interface." is very brief considering for example the Damköhler number. This part would benefit from a bit more in depth. What about availability of electron donors?

We have added some more text on this: *In the groundwater zone, nitrate reduction (nitrate reduction) takes place when nitrate containing water migrates from aerobic to anaerobic conditions and inherent reduced compounds are available (Hansen et al., 2014; Postma et al., 1991). For quaternary sediments these reduced compounds are mainly organic carbon and pyrite and ferrous ion from clay minerals (Ernstsen and Mørup, 1992; Postma et al., 1991). This transition zone between aerobic and anaerobic conditions is denoted the redox interface.*

Ernstsen, 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
Hansen, A.L., Christensen, B.S.B., Ernstsen, V., He, X., Refsgaard, J.C., 2014. A concept for estimating depth of the redox interface for catchment-scale nitrate modelling in a till area in Denmark. Hydrogeology Journal, 22(7): 1639-1655. DOI:10.1007/s10040-014-1152-y
Postma, D., Boesen, C., Kristiansen, H., Larsen, F., 1991. Nitrate Reduction in an Unconfined Sandy Aquifer: Water Chemistry, Reduction Processes, and Geochemical Modeling. Water Resources Research, 27(8): 2027-2045. DOI:10.1029/91wr00989

- L36: "In areas with Quaternary sediments characterized by groundwater dominated flow patterns and a relatively shallow redox interface, the N-reduction in groundwater can be the dominant removal process." Please provide a reference

Thank you, a reference has been provided: *In areas with Quaternary sediments characterized by groundwater dominated flow patterns and a relatively shallow redox interface, the nitrate reduction in groundwater can be the dominant removal process (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.

- L43: I do not know why showing maps is considered as "a new approach". Merz et al. 2009, for example, also showed retention maps and NO3 half-life times.

Yes, that is true. The sentence has been revised: *An approach for utilizing and illustrating the results and the spatially varying nitrate removal fractions (percentages) are through a nitrate reduction map*

– L47: "produced N-reduction maps with a 100 m spatial resolution for a 101 km2 catchment in Denmark," is not helping the argumentation, can be removed.

The sentence has been revised: *Hansen et al. (2014) produced nitrate reduction maps for a 101 km^2 catchment in Denmark, showing that nitrate reduction may vary from 20% to 70% between neighboring agricultural fields located only a couple of hundred meters apart.*

Hansen, A.L., Christensen, B.S.B., Ernstsen, V., He, X., Refsgaard, J.C., 2014. A concept for estimating depth of the redox interface for catchment-scale nitrate modelling in a till area in Denmark. Hydrogeology Journal, 22(7): 1639-1655. DOI:10.1007/s10040-014-1152-y

- L50: This sentence has to be checked for grammar. It is also partly redundant with the next sentence.
Thank you, this sentence has been removed.

- L55: "the effect hence is relatively large" please be more specific, "effect" is too vague here.
We have specified that it is mitigation effects.

- L57: This sentence is not clear to me. You are saying that N-reduction maps can be used more easily than hydrological models, but actually those models are used to produce the maps. This contradicts.
Thank you for this comment. This was indeed not very clear. We have revised the sentence in L68-L70:
Using nitrate reduction maps based on a single model run, is clearly a much faster method than running multiple complex hydrological simulation models for large ensembles of scenarios and is therefore a practical tool for policy analysis

And a few places in the section above we have changed "nitrate maps" to singular form, to indicate more clearly that nitrate reduction maps are based on a single run.

- L60: Please revise the sentence.
Thank you, this sentence has been revised: *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.*

Methods:
- L74: "best", "long", "near-complete" Please, specify
Thank you, the details of the timeseries have been specified: *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). The average discharge amounts to 4.4m3/s and the load is approximately 14 kg NO3-N/ha/year.*

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." Reference or more details needed

Time period for the averages are added, as well as a reference on trends in the nitrate time series. See the comment above.

- L79: reformulate "There were measurements"

Sentence has been revised: *226 measurements of the redox depth are available from boreholes in the area, the redox depths were mainly interpretated based on sediment colour as described by e.g. Ernstsen and Mørup (1992), and a few by measurements of reduced compounds.*

Ernstsen, 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

-L84: revise "100 meter redox depth map", I assume you mean the resolution

Thank you, this has been corrected.

- L84: "This map" reference unclear, as citations mismatch, it is not clear to me.

Thank you, this has been revised: *A recent redox depth map of Denmark was created in 2019, where measurements and system variables were used in a machine learning environment to create a detailed redox depth map in 100 meter resolution. This newer map also indicates that the redox depth in the study area is predominantly shallow with 1-10 meter depth, and very few sites of 10-15 meters depth (Koch et al., 2019a; Koch et al., 2019b).*

Koch, J., Ernstsen, V., Højberg, A.L., 2019a. Dybden til redoxgrænsen, 100 m grid. In: (GEUS), D.N.G.U.f.D.o.G. (Ed.), Copenhagen, GEUS. DOI:https://data.geus.dk/geusmap/?mapname=denmark#baslay=baseMapDa&optlay=&extent=-139925.69284407864,5929490.944444444,1254925.6928440786,6520509.055555556&layers=redox_dybde_100m_grid

Koch, J. et al., 2019b. Modeling Depth of the Redox Interface at High Resolution at National Scale Using Random Forest and Residual Gaussian Simulation. 55(2): 1451-1469. DOI:10.1029/2018wr023939

- L110: Please give a reference or indication why drains are needed, if locations are unknown. How do you define the "drain level" without this information?

A more detailed explanation is now provided in L145-149:

The study is heavily tile drained, and it can be assumed that drainage will always be present when it is need in the agricultural areas. However, the actual site-specific location of tile drains are unknown and therefore drains are specified across the entire catchment at a depth of -0.5 meters . Drain flow is however only activated when groundwater level rise above drain level. Apart from representing tile drainage the drainage

system also represents small ditches and stream, too small to incorporate into the river system following the approach of Troldborg et al. (2010).

Troldborg, L. et al., 2010. DK-model2009 - Modelopstilling og kalibrering for Fyn. GEUS Report.

- L119 typo "percolation"
Thank you, corrected.

-L119: "following methods" specify. Do you use several?
Thank you. Yes, the section is very unclear. The complete section on Daisy modelling and calibration have been rewritten: *The Daisy model setup for the Odense is contains on roughly 12,000 1D Daisy columns, and the water balance module is based on a previous calibration of the catchment (Børgesen et al., 2013), where root zone leaching and groundwater abstraction is compared with river discharge (Børgesen et al., 2013; Refsgaard et al., 2011). The model is setup so that each column represent unique combinations of soil type, climate, crop rotation and groundwater depth. The Daisy model uses the same climate input and soil parameter setup as MIKE SHE and the sensitive and calibrated unsaturated soil parameter from MIKE SHE were therefore transferred to Daisy. A more detailed description of the Daisy setup can be found in Karlsson et al. (2016). The water balance performance of Daisy was evaluated in the same calibration (2004-2007) and validation periods (2000-2003/2008-2009) as MIKE SHE.*

Daisy also simulates nitrate leaching for each soil column that represents a unique combination of soil type, climate, crop rotation and groundwater depth. Crops are fertilized with mineral and organic nitrogen dependent on the farm type and soiltype. The crop recommended nitrogen rate for the years 2004-2007 was used to setup the fertilization scheme. Nitrate leaching input are simulated on daily basis based on the leaching from the permutated crop rotations simulated for the dominating soil type within a 200m x 200 m square grid (Karlsson et al., 2016). Because of the close feedback mechanism between nitrogen yields and nitrate leaching, the simulated mean nitrogen yields were recalibrated to observed annual mean nitrogen yields on Funen (Statistikbanken, 2015) for the dominating soil type for the period 2004-2007. The calibration is conducted by adjusting the crop parameters, following the methodology of Styczen et al. (2004). Nitrogen concentrations of yields were extracted from table values of mean nitrogen contents for different crops (Møller et al., 2005). For crop rotations including clover grass and peas nitrogen biological fixation is calculated using Høgh-Jensen et al. (2004) and nitrogen atmospheric deposition is included as input to the soil using standard Daisy settings for dry and wet deposition (Hansen et al., 2012).

Børgesen, C.D. et al., 2013. Udviklingen i kvælstofudvaskning of næringsstofoverskud fra dansk landbrug for perioden 2007-2011. Evaluering af implementerede virkemidler til reduktion af kvælstofudvaskning samt en fremskrivning af planlagte virkemidlers effekt frem til 2015, DCA - Nationalt Center for Fødevarer og Jordbrug, Tjele, Denmark.
Hansen, S., Abrahamsen, P., T. Petersen, C., Styczen, M., 2012. Daisy: Model Use, Calibration, and Validation. Transactions of the ASABE, 55(4): 1317. DOI:https://doi.org/10.13031/2013.42244
Høgh-Jensen, H., Loges, R., Jørgensen, F.V., Vinther, F.P., Jensen, E.S., 2004. An empirical model for quantification of symbiotic nitrogen fixation in grass-clover mixtures. Agricultural Systems, 82(2): 181-194. DOI:https://doi.org/10.1016/j.agsy.2003.12.003

Karlsson, I.B. et al., 2016. Combined effects of climate models, hydrological model structures and land use scenarios on hydrological impacts of climate change. Journal of Hydrology, 535: 301-317. DOI:http://dx.doi.org/10.1016/j.jhydrol.2016.01.069

Møller, J. et al., 2005. Fodermiddeltabel - Sammensætning og foderværdi af fodermidler til kvæg. 64.

Refsgaard, J.C. et al., 2011. Vandbalance i Danmark - Vejledning i opgørelse af vandbalance ud fra hydrologiske data for perioden 1990-2010, Copenhagen, Denmark.

Statistikbanken, 2015. Statistical regional registrated annual mean yields. (In Danish) https://www.statistikbanken.dk/jord3.

Styczen, M. et al., 2004. Standardopstillinger til Daisy-modellen. Vejledning og baggrund, Institut for Vand og Miljø, DHI.

- L121: when is the simulation period? Calibration was mentioned, but what about validation period?

This has been added: *The water balance performance of Daisy was evaluated in the same calibration (2004-2007) and validation periods (2000-2003/2008-2009) as MIKE SHE.*

- L129: this is not a sentence

The sentence has been corrected: *Nitrogen concentrations of yields were extracted from table values of mean nitrogen contents for different crops*

- L135: bad title

The title has been changed to: *3.3 Calibration of the nitrate model (Phase 3)*

- L139: "If the particle penetrates the redox interface, the nitrate is assumed to be removed completely and instantaneously by denitrification." Please, reflect more on this assumption.

This is a generally accepted assumption in geochemistry (Postma et al., 1991; Hansen et al., (2014). We have added reference to validate this assumption.

Hansen AL, Christensen BSB, Ernstsen V, He X, Refsgaard JC (2014) A concept for estimating depths of the redox interface for catchment scale nitrate modelling in a till area. Hydrogeology Journal, 22, 7, 1639-1655.

Postma D, Boesen C, Kristiansen H, Larsen F (1991) Nitrate reduction in an unconfined sandy aquifer: water chemistry, reduction processes, and geochemical modeling. Water Resources Research 27:2027–2045.

- L143: section 0
Corrected to section 3.4

- L175-181: Please, specify how you can state that this was or was not the case? How did you further investigate the stuck particles? How many particles get stuck? It seems quite a lot if the correction causes changes between -7 and 9%.

Yes, this issue is very frustrating. The issue has been observed before, when running Mike She with particle tracking. From the tracking file we can see that these particles do not leave their cell of origin. We have tried different approaches to eliminate the stuck particles (e.g. placement of the particles in different locations within the cell). Unfortunately, there was no fix for this issue, even as it was reported to the responsible company behind Mike She (DHI). We have estimated the number of stuck particles in the model an average of 9 % of particles.

This sentence has been revised and more information added: *Mike She is a commercial modelling tool and therefore there is no possibility to access the modelling code in order to correct this numerical error, or in any other way account for this model limitation. Therefore, it was necessary to introduce a correction scheme.*

We have added a better explanation on how this correction was done: *The actual fate of these stuck particles (reduced/non-reduced) are unknown. At an early stage the assumption was made that the captured particles, if they had moved correctly through the system, would be subject to a fate similar to the non-captured particles, i.e. that the relationship between reduced/non-reduced was the same. If this assumption is valid the calculation the reduction potential in each grid cell is the same with/without the stuck particles. Unfortunately, this assumption may not always be valid. Furthermore, the arrival percentage estimated by the two methods are not the same as not all particles are released in the complex particle arrival count, the data from which is the only way to calibrate the nitrate model. For the two methods to be comparable it is therefore necessary to exclude the particles that are stuck in the unsaturated zone. The correction factor is therefore introduced to eliminate the particles that are stuck from changing the reduction map. The correction uses a simple linear equation, where a correction factor is manually fitted so that the arrival percentage (originating from the reduction map multiplied by the nitrate input) matches the particle arrival percentage. These corrections are done individually for all reduction maps, and the correction causes a change in the reduction in the range of -7% to 9% with a mean of 2%.*

Thank you for this comment, it is entirely true that the newer projection may be less uncertain. Unfortunately, the study was conducted some years ago and is therefore using the climate projections available at the time. We, unfortunately, do not have the possibility to rerun the complete modelling suite with the updated climate change scenarios.

Thank you for your question. The bias correction is done because the regional climate model operates on much larger scale than hydrological models. Local precipitation patterns are thus not necessarily correct represented in the regional models. Bias correction to local observations is therefore a common approach for dealing with this issue e.g., Chen et al. (2011), Pasten-Zapata et al., (2019); Refsgaard et al., (2016). We have added the word downscaled to the sentence to help.

Chen, J., Brissette, F.P., Leconte, R., 2011. Uncertainty of downscaling method in quantifying the impact of climate change on hydrology. Journal of Hydrology, 401(3–4): 190-202. DOI:http://dx.doi.org/10.1016/j.jhydrol.2011.02.020

Pasten-Zapata, E., Sonnenborg, T.O., Refsgaard, J.C., 2019. Climate change: Sources of uncertainty in precipitation and temperature projections for Denmark. Geological Survey of Denmark and Greenland Bulletin, 43: e2019430102-01-e2019430102-06. DOI:https://doi.org/10.34194/GEUSB-201943-01-02

Refsgaard, J.C. et al., 2016. Climate change impacts on groundwater hydrology – where are the main uncertainties and can they be reduced? Hydrological Sciences Journal, 61(13): 2312-2324. DOI:10.1080/02626667.2015.1131899

- L200: a "combined" median model?

Thank you, this has been explained better: *The four selected realizations represent a wet, +19% in precipitation (ECHAM-HIRHAM5), a dry, -11% decrease in precipitation (ARPEGE—RM5.1), a warm, +3.4 °C temperature increase (HadCM3-HadRM3) and a model representing a median projection, +10% in precipitation and +2.1 °C in temperature (ECHAM5-RCA3).*

- L210 I do not understand "3% point reduction". What is point telling here?

We wanted to indicate that it is percentage point difference and not a difference in percentage. We have change it to p.p. which may be a more common way to write it.

- L218f.: I do not understand this combination that was done: "compared to the climate model results found for the reference period 1990-2009 using the same land use scenarios, resulting in 32 scenarios." I think formulations are overcomplicated and Table 3 should be placed here.

Thank you for this comment. We have moved table 3 to this location and revised the sentence:
All 20 combinations of future climate projections (4) and land use (5) were specified as input to the hydrological model. The model was run for both future (2088-2099) and reference period (1990-2009) resulting in 40 scenarios. Additionally, the model was run with observed climate for the period 1990-2009 (5 scenarios using observed land use and the four land use scenarios).

Results:
- L240: "observed trends in nitrate yields" where are these trends shown?

Yes, trends is the wrong word. We have changed it to: *observed values of nitrate yields*

- Table 3: I think this should be presented in the Methods? 3.6 scenarios? What does the grey shade mean? How can the climate scenarios be used for 1990-2009? Or is it necessary? I do not see them later in the presented maps.

Thank you, the table has been moved accordingly. The grey shades indicate the scenarios displayed in the coming figures 3,4,6,7,8, and the baseline scenario. We have added text in the figure caption to explain.

Thank you for you question on the climate scenarios. Climate models are commonly run for both a reference period in the present time and a future period. As climate model results are affected by biases, even though they are bias corrected/downscaled, the way of dealing with this issue is mainly solved by displaying only changes from past to future. This is done under the assumption that the climate model

biases are constant (the same for present and future conditions), thereby changes in the impact results are deemed more trustworthy than actual values from the future projections. This is the reason why figure 4-9 all focus on the changes from past to future.

We have added an extra explanatory sentence in line 312: *Future climate model runs are always compared with results from this period for the relevant climate model to ensure that climate model biases do not dominate the results.*

- Table 4: Do you have an idea why the standard deviations of all models are that similar (Table 4, 0.36-0.39)? Can you comment on that, please?

This could be related to the fact that most changes in the reduction map are happening for values close or around the mean. Areas with 0% og 100% reduction are perhaps less likely to change reduction potential, as they are either very close to surface water (0% reduction) or located at areas with a deep groundwater level (long deep flow paths and limited drain flow) (100% reduction).

- L295: "To investigate to what degree land use changes and climate change affect the reduction map, the difference between these scenarios and the reference scenario is shown in Figure 4." Does not seem to fit here if the next section title is "4.4 Impact of land use change on reduction maps".

Yes this is confusing. We have moved this sentence and incorporated it into L391:

To investigate the impact of land use change on the reduction maps, only land use is changed while climate remains constant, shown as the difference between land use changes scenarios and the baseline scenario (Figure 4, top row).

- Figure 3 and Figure 4 seem a bit redundant to me, considering that 4 is just the difference between the map shown in Fig2 and Fig3. I think one Figure could be economized here by merging or moving to the supplements

Thank you for this suggestion, we have moved Fig. 3 to supplementary material, and moved the sentence in L274-276 connected to the figure with it.

- I think it is not necessary to show Fig7 and Fig8, especially because Fig.8 is mainly a reprint of Fig.12 in Karlsson et al. 2016, while the maps in Fig.7 do not allow to recognize more details than the general observation of land use changing little, two climate scenarios becoming wetter and two drier, which is also clear from Fig.5. Again I think redundancy should be reduced and plots removed or to the supplements.

Thank you for this suggestion, we have moved Fig. 7 and 8 to supplementary material.

Discussion/Conclusions:

- L510: "such effects" reference unclear. Please explain further how 10% change in nitrate reduction over almost a century relates to the uncertainties of nitrate reduction maps. Is it really essential to consider changes in nitrate reduction for management, if the tool itself is already quite uncertain?

Thank you, we have made this sentence more clearly: *The uncertainty of using a fixed reduction map for future scenarios should of course be seen in the context of the inherent uncertainties of the nitrate reduction maps (Hansen et al., 2014b).*

We have also added some reflections on the impact for management, L593-595:

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, A.L., Gunderman, D., He, X., Refsgaard, J.C., 2014b. Uncertainty assessment of spatially distributed nitrate reduction potential in groundwater using multiple geological realizations. Journal of Hydrology, 519, Part A: 225-237. DOI:http://dx.doi.org/10.1016/j.jhydrol.2014.07.013

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.

- L512-516: "single case study" how representative is this case? What do you expect for other sites? Compare to other studies. General spatial differences between nitrate reduction could be considered.

Thank you for this valid point. Unfortunately, we are not aware of any other studies evaluating uncertainties from application of a fixed reduction map vs. a full nitrate modelling scenario estimation. We have added a few remarks:

The indication that errors can be up to 10% is based on only a single case study with one catchment, one model and a limited number of land use and climate change scenarios. While similar results may be found when applying the same approach for catchments governed by the same dominant flow processes and land use types, like the one investigated in this study. The error must be expected to be site and context specific and therefore causes …

- L513-516: You mention uncertainties of input data (climate and land use) and model structure (though very briefly) but do not discuss the uncertainties related to the model and its parameters.

Thank you, yes this is not mentioned. We have added text on this issue in L606-615:

The error must be expected to be site and context specific and therefore causes projection uncertainties that should be addressed along with other known sources of uncertainty such as climate model projections, land use projections, parameter uncertainties, geological uncertainty, and hydrological model structural uncertainty (Hansen et al., 2014b; Karlsson et al., 2016). Furthermore, during calibration and setup of the model assumptions must be made, adding to the uncertainty. Parameter estimation are here done in a stepwise fashion, and the catchment scale calibration of Daisy along with a particle tracking approach limits the evaluation of performance of the nitrate component to mean catchment figures, and the dynamic of the nitrate system is thus impossible to verify. To account for this a full solute transport solution would be necessary but was unfortunately not possible in the framework of this study; but would be a relevant next step in investigating uncertainties and improve model verification.

- L516: This is not a nice ending. I would collect the conclusions in an extra Conclusion section.

Thank you for this suggestion. We have added a conclusion section:

Nitrate reduction maps are valuable tools used for calculation of remediation and climate change effects on nitrate leaching, and are generally considered constant in time, even though the timing of nitrate leaching,

and flow paths may change. In this study we investigate the potential consequence for estimation nitrate climate and land use change impact projections when assuming a fixed reduction map. For an agricultural dominated catchment in Denmark, the Daisy model was used to provide nitrate leaching input, while the hydrological model Mike She was used to simulate the flow regime and nitrate flow path through particle tracking. Four land use scenarios and four climate change projections were evaluated. The main finding of the study was:

- *Changing climate conditions lead to reduction map changes of up to 10%; whilst effects from land use changes where minor. However, land use effects may be underestimated due to drainage formulations in non-agricultural areas.*
- *Thus, the uncertainty of the reduction maps is dependent on both model setup and assumptions, the catchment flow regime as well as affected by the span of the chosen land use and climate change scenarios.*
- *The error will therefore be specific for the study site and context and it should, consequently, be tackled along with other sources of uncertainty, like geological, parameter, and model structure uncertainties that are not evaluated in this study.*

Data/code availability: - "owned by the DMI" - what does this mean? Where and how to access it?
The data is currently not pubplily available from the Meteorological service, but will be in the future. We have added this information in L576-569:

An exception to this is the observed climate data input which is currently the property of the Danish Meteorological Institute (DMI), but will be made publicly available through https://www.dmi.dk/frie-data/ before end of 2023.

Figures:
- F1: typo at "Market driven". Odd start of the caption as the Figure is showing the study area and land use scenarios and not a "red square".
Thank you, we have corrected the typo and the formulation in the figure text.

- F2: I think the map titles should be linked to the legend color bars. Maybe increase letter size of legend, add unit at left panel. To me it is unclear, what the text on the y axis refers to "Observed climate. . .", it seems unconnected. In the caption specify what the reduction refers to e.g. from. . . to. . ..
Thank you for these suggestions. We have changed the figure according to the comments.

- F5: I suggest to remove redundant subplot titles, this should be explained in the caption.
It has been removed

- F9: The caption is unclear to me. What combinations are shown? What is meant here "or the reduction map from the scenarios"?
We have tried to explain better what we mean: *Bars denote the change in nitrate flux at the catchment outlet that arises from using either a fixed nitrate reduction potential map (baseline) or using a reduction map based on the individual scenarios.*

References: - There seems to be an error in the display of the references – is it double or some other problem?

Yes, it is double, sorry. This has been corrected.