REVIEWER 2

The paper is written very clearly and well presented overall. As my only major comment, I am under the impression that the discussion on the global results is too optimistic. I suggest that the authors should keep the discussion in Section 3.2 more in line with what is actually shown in the results. I have further minor comments below.

We sincerely appreciate the reviewer's comment. We have improved the clarity of our global evaluation and have adjusted the tone to be more balanced (Section 3.2). We would also like to emphasize that, as previously mentioned to Reviewer 1, we identified and addressed a bug in the script used to calculate the fluxes presented in Tables 2-4 and the comparison of soil NH_3 against emission inventories (Figures 6 and S8). It is important to note that these updates have not altered the main findings of the manuscript; if anything, they have contributed to a clearer comparison between global emissions and emission inventories. We provide a detailed account of these modifications in response to the reviewer's specific comments below.

We have responded to each comments sequentially with italicised text showing the reviewer's comments and plain text showing our response. New text added to the manuscript is coloured blue, and any text removed from the manuscript is struck through. The locations of changes are stated.

Line 41: The literature can be made updated. The scenario community is exploring scenarios without much relying on CDRs (e.g. Riahi et al. 2021), a shift in paradigm from the time of Fuss et al. (2014).

We thank the reviewer for pointing out to Riahi et al. (2021), we have added this work in the introduction accordingly.

Line 46: The authors may include the most recent State of CDR report. https://www.stateofcdr.org/

Thanks for bring Smith et al, (2023) to our attention. Added as suggested.

Lines 48-55: The authors could also point out the release of phosphorus as another biogeochemical consequence of ERW, as discussed in Goll et al. (2021).

Added as suggested.

Section 1 Line 51

> Basalt is an ideal abundant silicate rock for ERW because of its potential co-benefits for crop yields and capacity to reverse soil acidification (Kantola et al., 2017; Beerling et al., 2018) and supply plant-essential nutrients like phosphorus (Goll et al., 2021).

Line 135: What happened to other factors such as soil temperature and water content, which were just mentioned a few lines above?

The reviewer raises a valid point. As discussed in Section 2.2.1, it is acknowledged that factors such as soil temperature and water content can also influence soil N_2O nitrification fluxes. However, the primary focus of our work is to update the nitrogen cycling in CLM5 that specifically responds to changes in soil pH, which we consider a key property affected by basalt application based on our field trials in the US Corn Belt. Throughout the manuscript, we have taken into account Reviewer 1's feedback and made it clearer that our implementations represent an initial approximation for an ERW parameterization in a land model. We recognize the need for further research and improvements in the ERW schemes as well as the nitrogen cycling within CLM5, particularly as more experimental data becomes available. We remain hopeful that future work will allow for the introduction of a more comprehensive parameterization, including the consideration of other relevant factors, in the nitrogen cycling component of the model. We highlighted this in section 2.2.1.

Section 2.2.1 Lines 142-144

> However, further work is needed to evaluate the sensitivity of the model to this specific parameterization under other soil conditions, as well as to incorporate the influence of other environmental factors, such as water content and temperature.

Line 141: Is the soil pH kept at the nominal value throughout the simulation period?

The reviewer is correct: In the Control Run, soil pH is kept constant to the nominal value provided by the Harmonized World Database. We have made this clearer in the manuscript.

Section 2.3 Lines 244-246

> Specifically, CLM5 acquires spatially distributed delta pH values, and adjusts the initial soil pH accordingly. Thus, in the "Control" Run soil pH is kept constant to the nominal values provided by Harmonized World Soil Database, whereas in the "ERW" Run is modified following the ERW model projection.

Line 171: Usually "taken up", not "uptaken"

Changed as suggested.

Line 262: "qualitative" may be replaced with a more appropriate word. I understand what the authors try to say, but numerical comparisons are always quantitative.

Changed as suggested.

Section 2.6 Lines 276-278

> It is important to note that our CLM5 model-inventory comparison ~~should not be considered quantitative, but rather qualitative~~ should be considered as an approximation because our simulations do not match the meteorological years [..].

Line 290: "increases", not "increased"

Corrected

Lines 293-295: Please describe what kind of calibrations are needed to get a better agreement.

We have added the following text to clarify the type of calibrations we were referring to.

Section 3.1 Lines 300-304

> We note that in this project CLM5 has not been ~~calibrated~~ tuned specifically for the Energy Farm conditions or across the U.S., rather used as in the released version as the objective is to use the model at a global scale, across many crops, regions and for future climate projections. As a result, the land management practices, such as planting and harvesting times, as well as fertilizer application frequency and rate, employed in our simulations may not precisely match those implemented at the Energy Farm.

Line 297: it should be "with respect to".

Corrected.

Line 305: Why is the range from the 2019 soy simulations so large?

Thank for this comment, as it prompted to look at the soybean results with further detail. As a first step, we rerun the simulations turning both synthetic and manure fertilizers off to make the comparison to the Energy Farm soybean observations more accurate, as soybean is not fertilized at the Energy Farm. As indicated in the manuscript, we did not tune CLM5 for the Energy Farm conditions, as our aim was to use the model at a global scale, across many crops and regions, and for future climate projections. However, we acknowledge that a closer approximation should have been done for the field site comparison. The adjusted simulation did not change the result in Figure 3 b) but reduced the variability. It also improved the comparison of the soil N_2O (in magnitude) with the Energy Farm observations.

We updated Figure 3 a) Soy 2019 and Figure 3b), and added in the caption what the error bars represented. In Section 3.1, we also discussed that soybean simulations did not consider fertilizer application.

Section 3.1 Lines 304-305

> To facilitate a more direct comparison for soybean, we made an exception and turned off synthetic and manure fertilizers because the Energy Farm does not employ fertilizer application for soybean crops.

Line 344: Maybe "simpler validation"

We believe the reviewer refers here to Line 354, where the word 'validation' is used with 'briefer'. We modified it as suggested and 'briefer validation' reads now as 'simpler validation'.

Line 378: The numbers are "r", not "r2". I would rather see it as a relatively poor correlation.

We agree with the reviewer about considering *r* values in the range of 0.3-0.4 was quite optimistic. An *r*-value of 0.4 suggests that there is some degree of association between the variables, but the relationship is not very strong.

We have modified the discussion accordingly.

Section 3.2 Lines 391-393

> The global *r* values range between 0.3 and 0.4 across the inventory and models, suggesting that CLM5 does not exactly replicate the spatial patterns reported on the emission inventories.

Section 3.2 Lines 396

> Our global *r* values lie between 0.4–0.6 across all inventories, indicating a fair correlation.

Section 3.2 Lines 405

> The global *r* values are 0.5–0.6, indicating a fair correlation between CLM5 in all three emission inventories.

Line 379: I think that the statement here is also too optimistic. NMB from EDGAR is 147%.

The NMB for EDGAR has been updated and is now 112%. We understand that this is a large bias but want to remind the reviewer that there are significant differences among emission inventories too, in terms of magnitude, spatial distribution and seasonality, as discussed in the manuscript (Section 3.3 Lines 438-444). None of the emission inventories can be considered as the definitive ground truth. Our soil agriculture NO_x emissions (2.2 Tg N/yr) align reasonably with reported values (0.4–3.5 Tg N yr^{-1}), which is acceptable given the wide range reported in the literature. In response to reviewer's suggestion, we have moderated the level of optimism in our model results throughout Section 3.2, as shown through several of our responses.

Line 379: NMBs for NO are 21%, 71%, 142%, and -11%. What are the reasons for the very high bias?

The NMB values for soil NO have been corrected to address the calculation issue, resulting in updated values of -5%, 6%, 57%, and 117%. The corresponding text has been revised accordingly to accurately reflect these corrected NMB values.

Section 3.2 Lines 396-399

> Our estimate is higher than ~~three~~ two emission inventories (~~CAMS,~~ CEDS and EDGAR) with a global NMB value ~~between~~ of ~~21~~–57 and ~~147~~ 117%, but close to the CAMS (NMB=6%) and the adjusted HEMCO (NMB=-~~11~~ -5%) estimates.

We acknowledge the substantial biases between CLM5 and certain emission inventories (eg, EDGAR), as well as among the emission inventories themselves. These disparities can be attributed to several factors discussed in Section 3.2 (Lines 438-444). These factors include differences in the timing and duration of fertilization considered, the inclusion of various agricultural sources (e.g., synthetic and/or manure application, manure management), and systematic uncertainties within the global inventories (e.g., emission factors, environmental conditions, fertilizer types and rates). In this work, it is not our objective to comprehensively address every individual bias in this study, we use the model-inventory comparison to provide context to the CLM5 output.

Line 382: What is the basis for good correlation?

When evaluating model performance on a spatial scale, the closer the correlation coefficient to 1 the stronger agreement between the model and observations (or emission inventories in this case). In line with the reviewer's comment, we acknowledge that our previous reporting of model performance against emission inventories may have been optimistic. It is important to recognize that our model does not fully capture the underlying processes or mechanisms responsible for the observed spatial patterns in the various emission inventories. As indicated above by the reviewer, we have adjusted the tone of the discussion in Section 3.2 accordingly.

Line 383: Table 2 indicates 142%.

This value has now changed in Table 2 and the text cites the correct value.

Line 402: I think that the discussion should continue with a comparison between CLM5 and inventories. For example, there is a rather large difference in the estimate of NH3 emissions in China between CLM5 and inventories.

We appreciate the reviewer's feedback and acknowledge the significant disparities between CLM5 and emission inventories, particularly regarding NH_3, on a regional scale. In response to this comment, we have addressed these differences in more detail and provided an expanded discussion in Section 3.2 of our manuscript to underscore their importance. We also strengthen the disparities among emission inventories.

Section 3.2 Lines 416-419

> Emission inventories show a similar regional distribution of emissions, with a higher proportion of agriculture emissions in China and India. For example, for NH_3 emissions, CAMS, CEDS, and EDGAR indicate that India is the largest emitter, accounting for 23–30% of global emissions, followed by China with 16–17%.

Section 3.2 Lines 444-446

> It is important to ~~note~~ acknowledge that substantial differences among emission inventories also exist in terms of their magnitude, spatial distribution, and seasonality.

Lines 437-439: Here again I think that the statement is too optimistic and unsubstantiated. I would rather see it as a mixed outcome.

As suggested by the reviewer we have tone down the statement to represent a more realistic outcome.

Section 3.2 Lines 454-456

> We concluded that CLM5 ~~captures well~~ provides a reasonable representation of the magnitude and seasonality of direct agriculture nitrogen emissions within the major hotspot regions (North America, Brazil, Europe, India, and China), which are relevant to our study. We note that there may be some limitations and uncertainties associated with the model's performance as well as current emission inventories in capturing the full complexity of these emissions. Further investigations and validation efforts are warranted to enhance our understanding of regional variations in agricultural nitrogen emissions.

Line 481: Table 3 indicates that CLM5 generally gives a lower estimate of NH3 emissions than emission inventories do. Does this imply that the effect of ERW on NH3 emissions can be larger than what is indicated from CLM5?

As mentioned above, we have made the necessary updates to Table 3 to address the calculation issue. In this updated version, CLM5 shows lower NH_3 estimates for Brazil and China, while higher estimates are observed for India and North America, in comparison to other emission inventories. The reviewer's observation regarding these biases is valid, as they may potentially impact the results. However, it is important to note that our focus is on comparing the Control and ERW scenarios, and any inherent model biases are expected to offset each other. Furthermore, we have considered the Reviewer 1's suggestion and added a note to emphasize that our study represents an initial implementation to evaluate ERW in the land model N cycle, and we acknowledge the need for further research and improvements to address any potential biases.

Section 5 Lines 512-520

> We acknowledge the need for further improvement in the CLM5 nitrogen cycling representation and the ERW parameterizations. In a comprehensive evaluation of CLM5 nitrification and denitrification processes, Nevison et al. (2022b) emphasized that the nitrification:denitrification ratio (2:1) in CLM5 is likely to be unrealistically low, even when considering the missing N mineralization term in potential nitrification (Section 2.2.1). Consequently, CLM5 underestimates the fraction of gross mineralization leading to nitrification and overestimates NH_4^+ uptake by plants. Additionally, CLM5 underestimates NO_3 assimilation by immobilizing bacteria. To enhance the confidence in our land model simulations, it is thus crucial to gather more experimental data from ERW field trials as well as observational constraints on soil nitrogen fluxes and flux ratios.
>
> Our study represents a first implementation of an ERW parametrization in a land model N cycling, which has enabled us to understand the implication of large-scale deployment of ERW with croplands on direct soil nitrogen trace gas emissions.

References

Goll et al. (2021) Potential CO2 removal from enhanced weathering by ecosystem responses to powdered rock. Nature Geoscience 14 (8):545-549. doi:10.1038/s41561-021-00798-x

Riahi et al. (2021) Cost and attainability of meeting stringent climate targets without overshoot. Nature Climate Change 11 (12):1063-1069. doi:10.1038/s41558-021-01215-2