Reviewer 2

The manuscript presents lysimeter results comparing conventional tillage (CT) and no-inversion tillage (NIT) across three soil types under the same crop rotations over a 2–3 year period (2013–2016). The topic is timely, and the results clearly indicate differences in nitrogen losses between CT and NIT. That said, the manuscript would benefit from substantial revisions to improve clarity, structure, and the strength of its mechanistic interpretation.

➔ Many thanks for the thorough and constructive review! We appreciate the your feedback and valuable suggestions and have revised the manuscript extensively to improve its clarity, structure and mechanistic interpretations.

In several places, the writing reads as a sequence of observations rather than a cohesive argument that explicitly links proposed mechanisms to the reported patterns. In addition, key pieces of evidence are sometimes introduced late in paragraphs, where readers would instead expect a clear synthesis or “take-home” message for the section.

➔ The structure of the results section was revised completely to improve clarity of presented results. We also added more relevant information on initial soil properties for the three soil types. We now show results also by soil type to support the interpretation and we added an ANOVA investigating variance in annual leaching loads following recommendations from reviewer 1.

➔ To strengthen the linkage between observations and mechanisms, also the discussion section was revised and restructured extensively.

Relatedly, the manuscript includes considerable discussion of carbon’s role in mediating nitrogen loss under different tillage practices, but it is not clear that carbon content was directly measured in the study soils. The contribution could be strengthened considerably by incorporating additional measurements—such as soil carbon quantity and forms of nitrogen (and, ideally, complementary water chemistry where feasible)—to better support the claims about nitrogen retention and the drivers of observed losses.

➔ Very good point. We added information on initial SOC and N-contents for the three soil types. Inclusion of measured soil organic carbon contents in an ANOVA highlights the relevance of SOC as a mediator of nitrate leaching. The newly added ANOVA revealed that variance in annual NO_3^- leaching between 2013 and 2016 was most strongly influenced by precipitation amount, the treatment effect, SOC and their interactive terms.

At present, I think the manuscript requires major revision before it can be considered for publication. While the dataset is interesting and potentially valuable, the current presentation feels closer to a field report of patterns than a study that systematically evaluates multiple factors shaping nitrogen dynamics. Several important variables (e.g., carbon content, aggregate stability, nitrogen species, litter quantity/quality, rooting depth, and the impacts of plowing) are discussed as likely explanations but do not appear to be directly measured, aside from porosity. Where porosity is reported (notably in the appendix), it would also help to clearly indicate which values correspond to each treatment (NIT vs. CT), and to ensure those results are integrated into the main narrative as appropriate.

➔ We agree that the presentation of results and interpretations lacked clarity in the submitted manuscript. We therefore restructured and rewrote the results section completely, also improving readability of graphics. The revised results section starts with a presentation of the main finding: nitrate leaching is higher with CT than with NIT. In the following subsections we present and describe results that support our interpretation of mechanisms driving the observed treatment effect. Our data shows that the observed treatment effect on nitrate leaching is driven by differences in

nitrate concentrations in seepage water and not seepage water amounts. We argue that this difference in nitrate concentrations is attributable to a difference in flow pathways, differences in aggregate stability, differences in crop residue amounts and presumably differences in denitrification. While our data supports the hypotheses that differences in flow pathways and residue amounts have contributed to the observed difference in N leaching, it is not suffiencient to prove the relevance of either of these mechanism in the experiment. In the last sub-section of the discussion, we refer to these limitations of our study and points at possible ways to overcome them in future studies.

➔ In general, the discussion section was revised extensively to improve readability and clarity in our line of argumentation.

Finally, the manuscript would benefit from careful attention to consistency in figures and text, including standardized styling (colors and formatting) and consistent treatment labeling/abbreviations throughout.

➔ Previous figures were replaced by new ones to improve readability, also paying attention to consistency in formatting/labelling/colors.

Overall, I see real promise in this work, particularly in highlighting management differences and the importance of site context. To make the manuscript publication-ready, I encourage the authors to (1) strengthen the linkage between mechanisms and observations in the writing, (2) add or more directly incorporate key supporting data (especially soil nitrogen, and carbon if central to the interpretation), and (3) more fully integrate relevant supplemental information (for example, the temperature data) into the analysis and discussion.

➔ By restructuring, rewriting the results section with the additions of ANOVA, we strengthen the link between observed treatment differences and influencing mechanisms

➔ To support the line of argumentation some of the material from the appendices was moved to the main text:

- Climate information was moved to section 2 (methods and data)
- Soil profile information was moved to section 2 (methods and data)
- The graphic on dry matter yields by harvest year and treatment was moved to the results section.

➔ Figures showing soil moisture, water tension and soil temperature in 10 cm depth were integrated into the results section. In our interpretations of observed treatment effects and influencing mechanisms, we refer to these graphics specifically.

Lines 35-40: I think it would strengthen the authors argument if you can add a line about how widespread these tillage practices are (i.e. what proportion of farms or what percent of land undergoes method X or Y via management). This would strengthen your paper's relevance and impact.

➔ We added a paragraph to the introduction section providing this information

Graphical abstract: It currently does not suggest based on the drawing that one method goes deeper than the other. Furthermore, the abstract doesn't show contrast with regards to changes to soil structure which appear to be central in the author's argument.

➔ There was actually no difference in tillage depth. Tillage depth was 20cm both in CT and NIT. The graphic was modified to indicate differences in macropore continuity and aggregate stability

Line 42: N2O should be subscript (similar for all other chemical formulas brought into the text and used in your graphs/figures, use correct notation and subscript!). The sentence after this makes no substantial addition to the introduction nor does it connect to the next paragraph. Reconsider structure here.

➔ This part on reported co-benfits of reduced tillage practices was shortended and revised to improve clarity

Line 49: “addressed the inconsistency in the impacts of tillage on nitrate leaching currently documented.” Or rewrite another way, its very awkwardly phrased at the moment.

➔ Was revised accordingly

Line 51: This last sentence does not connect well with the next paragraph nor is it a good concluding sentence.

➔ True, this is actually not a concluding sentence. The concluding sentence was in the next paragraph, which was now moved and revised.

Line 53: Why is this sentence separate from everything else?

➔ Exactly. This sentence had to be moved upwards (see previous comments)

Line 55: This is not a complete paragraph. Plant uptake is something that could decrease the amount of leachate correct? Then this wouldn’t be a driver but instead a factor that affects the magnitude of nitrate leachate. In this case calling these “processes that influence nitrate leaching dynamics” would be more accurate.

➔ Thanks for the recommendation. This was revised accordingly.

Line 79: Any time you place the variable from your equation into your main text, keep the same text style as the one you used in the equation description (is it subscript a or normal a?).

➔ We thoroughly checked the whole manuscript to eliminate all such inconsistencies in style and formatting.

Line 95-97: this is exactly why you need to have a section in the introduction that can speak to how widely used this practice is globally. It enhances your research’s relevance to global community while emphasizing its importance and makes sure content like this is already introduced so that it doesn’t take up space later in the manuscript where you need to focus on the experimental details instead.

➔ We have moved this part to the introduction section.

Line 105: Did each treatment receive the same cover crop treatment? It would be helpful to have a schematic here of the experimental design or a table in the SI.

➔ Yes, the rotation was the same in each treatment. The treatment difference was only related to the soil management (NIT vs. CT.). We understand that the description of the crop rotation implemented in the experiment was not very clear. This was now revised and to improve clarity we added a table listing crops in the rotation and respective fertilizer amounts. The experimental plan was added to the appendix.

Line 115: The listing of the cover crop rotations is confusing with some have a (+) and others not having this. Improve the clarity of what you are trying to express here. Is there a particular order (numbering may be better). Is there a set duo? (note this in the text). At the moment this listing is somewhat disorienting and difficult to parse. Additionally, italize scientific names.

➔ This section was revised to improve its readability. A table was added listing crops by year and N fertilization amounts.

Line 120: These values should be supplied in the SI or some note on using these to normalize your results should be made (or if they weren’t that different from each other between the treatments, this should be noted).

➔ This was revised. A new table lists this information. There were no differences in fertilization levels between treatments.

Line 125: Be consistent throughout your text with regards on how you will be referring to your nitrogen species (chemical formula? Written name? Also if you choose chemical formula add respective charges and subscripts – be consistent!).

➔ We thoroughly checked the whole manuscript to eliminate all such inconsistencies in style and formatting.

Figure 1 & similar style comments for figure 2

➔ All results figures were redone to improve readability and consistency in style.

Where is the key for the first top figure? (monthly precip and mean temp)

➔ Climate information was removed from the results and is now presented in section 2 “Methods and data”.

Label these figures as Fig 1a 1b 1c etc same for figure 2

➔ Revised figures are labelled as recommended.

Figure axes labels are blurry and so is a few of the graphs included here (fix please!)

➔ All figure were revised to improve readability and are provided in vector format.

Label evapotranspiration as evapotranspiration.

➔ The graphic was redone.

Monthly precipitation data is provided in Month:Year format and the following figures that are meant to be aligned over the same time periods as this precip data has Year:Month:Day. Change these to have the same formatting or change these to all be over one axes if appropriate (which I believe it is).

➔ The revised graphics are aligned over a common time axis (i.e. bi-weekly in accordance with water sampling intervals).

I would further suggest placing the name of the crop that was collected over the top of the colored boxes.

➔ Yes, this was done.

Be consistent in all of your graph labelling – will you capitalize or not? Don’t short hand write the labels, use full words to convey what you are showing and explain in detail the units in the captions if necessary.

➔ We thoroughly checked the whole manuscript to eliminate all such inconsistencies in style.

Line 185: reformat the figure caption:
coloured background shading is in reference to cropping periods: purple (phacelia), yellow (sugar beet) … ; grey dashed vertical lines denote ploughing times; …etc

➔ Done.

Get rid of the “=” they are hard to read/decipher with everything else that’s being listed in this caption.

➔ Done.

For figure 2 make one large legend for parts b, c, d (label these figures accordingly)

➔ The figure was replaced by new figures, all of which are labeled.

Figure 3 should be colored to match earlier graphs. However, choose a color palette that is color blind safe. Also denote significant differences on your figure in some manner (asterisks or letters).

➔ A color-blind palette is now used and where appropriate, p-values or asterisks were added to the new graphics to indicate statistical signifcances.

Line 205: Your mechanistic connection here between pore structure and matrix water flow is not very clear. Are you saying that increased tillage under CT would reduce pore connectivity or increase pore presence? What about pore structure is facilitating the matrix water flow in CT vs NIT. You leave your readers without a clear connection between the two and then just citing various literature that suggest similar trends. It will really help drive your point if you make the connection clearer here.

➔ This part of the discussion was extensively revised for clarification.

Line 210: This sentence is just a list of literature, integrate it into the earlier sentence and cut down on the number of citations or still include them but not as an in-line reference just in the ().

➔ In the revised version of this discussion sub-section we make clear in what respect we refer to these references.

Line 215: “In our study, also..” the structure of this sentence is awkward, edit this to improve delivery of the point you are trying to make here.

➔ As part of the extensive revisions of the discussion, this was revised and improved.

Line 225: I think what your results may be suggesting is that it is the combination of crop and soil management technique that directly affects the rate and timing of your N leachate. You are implying here that the nutrient quality and quantity of the sugar beet reincorporated into the soil after harvest had an impact in N mobilization. Do you have data on rooting depth? If roots are going deeper than what the plow can do this may add to your story here.

➔ Unfortunately, we do not have data on rooting depth from this experiment, but with the addition of information on SOC and with the ANOVA we could point out that the treatment effect is significantly influences by precipitation and soil organic carbon content.

Line 247 to 252: This sentence is way too long

➔ This sentence was revised to improve readability.

Line 252: Include Dr. Colunga’s meta-analysis work on conservation (reduced) tillage and conventional tillage in sandy arid environments with carbon deprived soils https://doi.org/10.1016/j.still.2024.106310

➔ We added this reference in the introduction section where we refer to benefits of reduced tillage for carbon sequestration.

Section 4.2 – I think the big message here is that the benefits of CT or NIT with regards to N mobilization and leaching are extremely context dependent. You can improve the delivery and conclusion of this section by coming towards a conclusion like this rather than having a paragraph that just details a variety of different impacts from these practices in different arid and low C environments. This makes this section currently just feel like a list rather than a cohesive argument coming to a point. Enhance the delivery of that context dependence message by incorporating the literature suggested earlier but also noting that this is dependent on regional practices (crops of cultivation, rooting and biomass size, etc).

➔ Yes, thank you for the comment. We agree that the structure of the discussion was not well established and we have revised and re-written the discussion to improve readability and clarity in the line of argumentation.

Figure E1: labels are upside down and poor resolution of axes and repetitive keys make this super messy.
Figure E2: similar as E1 but it is super blurry and almost impossible to read.
Figure E3 is similar as E1

➔ These figures were redone with consistent color-coding and labelling. We include vector graphics in the revised manuscript.

Figure F1 has a completely different color scheme. There is no indication here of significant different between sugar beet harvest in the NIT or CT. The bottom axes has the labels off centered and the double ticks seem unnecessary and messy.

➔ This figure was removed and replaced by a new figure in the main text in alignment with the previous color scheme.