**Reviewer #1:**  
  
Comments to the Author  
I appreciate the effort the authors put in the preparation of this manuscript. It is very clear and the conclusions and discussion are supported by the data and analysis. Here are my suggestions:  
  
LN 87 I suggest replacing the word “managerial” with “management”

Done

LN 124 I suggest replacing “save” with “except”

Done

LN 242 “which may have better randomized spatial patterns in the soil”. I’m not sure what this means. Can this be reworded? I can guess, but the way it is written is not clear to me.

We changed the wording to be more explicit:

The Central site had several treatments in small plots and the field was divided into four quadrats to create the four blocks with each treatment being randomly assigned within the block; this arrangement may have better randomized treatments with respect to natural patterns of soil texture variability spatially within the site.

LN 250. It is not possible to have 1.8 g of OM per g of soil. Is this per 100 g soil? So a 1.8% (or 4.6) OM.

Yes this was a typo, thank you for catching it. We changed the text to read: 1.8 to 4.6% for clarity.

LN 282 place a comma after do.

This sentence was changed based on another reviewer’s comment, and the ‘do’ was eliminated as a result:

To our knowledge there are limited studies examining the potential for CCs to reduce flood damage in the Midwest, but the one we are aware of accounts for only the increased evapo-transipiration with the use of CCs (Antolini et al., 2020).  
  
Figure 1 I suggest removing this figure; it is not very informative.

The figure was moved to supplementary material.

Figure 4. Could be more informative if the average for Cover vs. No Cover was included and the individual plots are being more transparent or grayed out.

We changed the individual plots to transparent dotted lines and added the mean values for each cover crop treatment as solid thick lines.   
  
I would note that there is no reporting of yield or soil water - in this paper - so there is no evidence that the cover crops have actually impacted these cropping systems. However, I do understand that this is somewhat outside the scope of this particular contribution.

We agree many readers may wonder the same thing. Unfortunately there are no measurements of soil water available for these sites, but we added a citation to direct the reader to reports of yields.   
  
**Reviewer #2 – Comments and responses**

Dear authors,  
The manuscript had the potential to be useful to readers of this journal. However, it is not suitable for publication. The study was supposed to investigate the effects of CCs on soil water storage at various sites and to provide possible causal models. However, much time was spent explaining the effects of soil texture. There was very little discussion of the results (causes) and there were too many speculations not supported by the data presented. More detailed comments on how to improve the manuscript is provided in the attached document.  
Thank you.  
  
Thank you for the helpful suggestions. We have de-emphasized the soil texture results as they were not the main focus of this study. We added text to direct the reader more explicitly to the discussion of causes. We were unable to identify specific speculations the reviewer deemed inappropriate in the comments. We found soil texture varied by cover crop treatment (we do not imply cover cropping *caused* the soil texture to vary, and we reworded that section to make that clear), which may be what the reviewer is referring to. If there are other specific instances where the reviewer feels our claims are not supported by our data we are happy to address them.

L40 - This is common enough to not need an explanation. i.e. no need for abbreviation.

We want our work to be accessible to international audiences, and therefore do not want to make assumptions about knowledge of governmental acronyms used in the United States.

L48 – Secale cereale

Corrected.

L50 - Avoid using first person pronouns.

To our knowledge, this journal does not specifically ban the use of first person pronouns. We believe it aids in readability in this instance.

L53 - This suggests several depths were sampled. If so, what was the extent? If not, please clarify that only 10-18 cm depth was sampled (rather than an increment.

We removed the word ‘increment’.

L63-64 Please rephrase. Something along the lines of "At two trials, soil water content at field capacity under CC management was 2.5 and 2.4% higher, respectively, compared with NC".

Changed.

L89 - Since this is the first Nichols et al., 2020 in the manuscript, shouldn't this be 2020a?

Yes, this citation was changed to 2020a and the references were updated accordingly.

L120 - Soil textural analysis will be a very good addition to the data. Although the soil series were identified, it is important to remember that the translocational and depositional processes can affect the current particle size distribution.

The average soil texture information for each location was added to the table.

L128 - This should be Nichols et al., 2020b

Changed.

L128 - At what depths?

We changed the sentence to read:

*All sites had sub-surface drainage at approximately 1.2 m*

L130 - The map is not clear (the highlighted counties are very clear and it makes it difficult to understand what is going on here)

We eliminated the map from the manuscript and put it in supplementary material, and eliminated the county outlines on the map.

L139 - A brief explanation (before referring the readers to another publication for detailed explanation) will be helpful for readers.

We changed the text to the following:

*Cover crop biomass sampling occurred each spring at every trial by removing two or four aboveground biomass from an area of 0.25-0.36 m2 from each plot, depending on the trial; details about methodology are reported elsewhere (Nichols et al. 2020b).*

L145 - Please state how many days for clarity since this will be dependent on soil texture, among other things.

We added the following text:

*Two to four*

L149 - Why was this depth chosen and not any other depths?

Soil surface conditions play a huge role on water movement and retention. Why was the surface not sampled? It could provide important information.

This is explained in the introduction (L96-99). We reworded our objectives to better highlight the deeper depth as an explicit goal of this study (see below), but we do not think it is appropriate to repeat our logic in the materials and methods section.

*Given the need to quantify long-term benefits of cover cropping, the scarcity of Midwest-specific data at depths relevant to crop water-use, and lack of a framework for organizing relevant knowledge, the objectives of our study were to:*

*(1) Determine what aspects of a soil’s hydrological characteristics are affected by long-term cover cropping at a depth relevant to crop production, and whether those effects are consistent across sites.*

*(2) Use our findings to propose a causal model connecting CCs to changes in soil properties to aid in targeting future research.*

L158 – 159 - Not really important. Please delete.

We believe this information is important for understanding how the statistical models were chosen, and other reviewers requested information be added to this sentence, implying they found it pertinent. We therefore left the text.

L170-171 - Please be careful here. The method in the cited study is for the determination of bulk density in forest soils (this study was not conducted on forest soils so this citation is not accurate). Please use this statement and citation instead "Soil bulk density was analyzed using the core method (Grossman and Reinsch, 2002)".

The reference was changed.

L175 - This does not need to be a stand-alone subsection. Just mention the method used for the analysis with the citation and include it in the subsection just above this. Further, the percentages need to be presented (either in a new table or added to the current Table 1)

We combined the soil texture and organic carbon sections into one. The texture information was added to Table 1.

L181 - See comment above for soil textural analysis. Just state the method and citation. This will enhance brevity.

See above.

L187 - "Analysis" is a better term.

Changed.

L198 - Biologically meaningful? What does this mean?

This sentence was deleted.

L218 - It also depends on other soil intrinsic factors.

We agree, however the majority of studies assume -330 cmH2O which ignores ALL other factors. In areas with shallow water tables, such as Iowa, the depth to the water table is one of the main considerations determining the matric potential at field capacity. We provide citations supporting our choice. If there is a particular factor the reviewer thinks we need to consider in addition to the water table, we are happy to consider it.

L220 - Higher matric potentials

Based on previous feedback, we believe ‘less negative’ avoids the confusion that results from a lower absolute value also being a higher numeric value in the case of negative numbers.

L220 - These values are not assumed. They are measured values in most cases. Please be careful here.

We provide citations supporting our decision to use -100 cmH2O.

L222 - Please be careful here too. It is a leap to go from -33 kPa soil water pressures not being an accurate measure of field capacity to -10 kPa being a better approximation. It is not supported by data. I suggest removing this sentence. The sentence before this is enough justification

We do not understand this suggestion. We believe we provide sufficient justification.

L236 - This should have been added to Table 1 rather than having it in the form of a figure. Also, this figure does not tell the reader anything about the depths.

We added the texture data to Table 1.

Also, this is presented in such a way that makes it seem like the CCs had any effects on the soil texture. This is not true.

The reviewer is mistaken – cover cropped plots had higher sand components than the no-cover plots at two of the trials, as described in L235-245. We do not imply the cover cropping caused different soil textures. We hope removing the figure will allow the reader to concentrate on the text description, and de-emphasize this component of the analysis.

L258 - It was just not significant. This will enhance brevity.

The problem is it was significant, both with and without a sand co-variate, but in one case significantly lower and in the other significantly higher. Stating it was not significant would not be true, although we realize it is a confusing result. The wording was changed in response to another reviewer’s suggestions.

L267 - Why measure water content at different pressures and not present the moisture retention curve.

We moved the retention curve figure to the beginning of this section so it is the first figure the reader sees.

That is a better way of presenting this data rather than the current table.

We believe the reviewer is referring to the figure, not a table. We changed it to a bar graph so it is easier to interpret, and moved it after the water retention curve.

L282 - This sentence suggests that you need to cite more than one reference here)

We reworded it to make it clear we are only aware of one study:

*To our knowledge there are limited studies examining the potential for CCs to reduce flood damage in the Midwest, but the one we are aware of accounts for only the increased evapo-transipiration with the use of CCs (Antolini et al., 2020).*

L284 - Why was it higher? Any causal mechanism?

We added text to indicate causal mechanisms are discussed later in the paper

L284 - 2% is just fine (delete the 'vol')

We found people were confused if the field capacity increased by 2% of the original value (i.e. they thought it increased 2% *of* 35%). To avoid that confusion we added vol% to all mentions to make it clear it is an absolute increase, not a relative increase.

L318 - How does below-ground biomass measurement get into your model if it wasn't measured in the first place?

The causal model is designed to help identify measurements that need to be taken. We realize the word ‘model’ may be confusing, as we also used statistical models.

**Reviewer #3 – Comments and responses**

General comments: This article seeks to address the impacts of long-term cover cropping on soil physical properties. The paper is well written and organized, and is a good fit for the journal. The authors utilize a rich dataset from four trials, but the analysis used and presentation of results are weak. Additionally, I believe there is a significant error in analysis that may have large impacts on the authors’ results and their interpretation (see detailed comments below). Overall, the results and discussion section needs much improvement, and the authors should improve the overall discussion of the impact of their work.

It is not clear to the reader a) what the major finding is and b) why that finding is important to the scientific community. I recommend that the paper be considered for publication only after major revisions are made. Specific comments are included below to aid the authors in improving their paper.

Thank you for the kind words and suggestions. We have reorganized the presentation of data, changed figures, and re-written the results and discussion to better emphasize the take-home messages. The major error the reviewer is referring to was the result of rounding to the nearest percent for purposes of figure labeling – the texture values added to 99 instead of 100. The analysis was not affected, and we added a table with the values in supplementary material to demonstrate the values did indeed sum to 100. We hope these changes help the reader understand our work and its significance.

Specific comments:

Core idea 1: Change 14 cm to 10-18 cm to match methods section

Done

Line 147: How many replications at each trial? Also, the authors do not discuss the impact of soil sample

collection in the middle of the plots vs. nearer the plant.

We added text about the total number of plots:

*This process was repeated for each plot (eight plots at East-grain, eight at West-grain, 10 at Central-silage, and 10 at Central-grain; Table 1)*

We also added text to indicate why we chose to sample between planted rows:

*At all trials, one sample was taken in the middle of the plot between planted rows in a non-wheel traffic row in order to get a sample unaffected by planting activities.*

Line 201: Note that the alpha parameter is not exactly the inverse of the air entry potential but is *related*

to it. From Van Genuchten (1980) “The value of alpha was found to be… ***approximately*** the inverse of

the pressure head at which the retention curve becomes the steepest.” Other follow-up studies have

shown that alpha values are related to air-entry potential but not exactly the same as that value.

We changed the text to include this:

*a approximates the inverse of the pressure at which the retention curve is steepest (e.g., air-entry potential; van Genutchen 1980),*

Line 206: The residual water content is not the same as the water content at permanent wilting point.

The residual water content is the water content at which no additional water is lost with increase matric

potential and has nothing to do with plant water uptake. The water content at permanent wilting point

is the water content associated with the matric potential at which plants cannot extract additional soil

water and is entirely dependent upon the plant species. While these values may be similar for a given

soil-plant system, to equate them to one another is incorrect.

We eliminated the mention of permanent wilting point from the text:

Residual water contents (θr) are estimated by the model, and can be greater than -15,000 cmH2O (SSSA, 2008).

Line 209: Replace “the data” with “measured values.” The use of the term “the data” is overused and

becomes confusing.

We changed this line, as well as other instances where ‘measured values’ will help the reader.

Line 215: Replace “extracted directly from the data” with “calculated from the measured data.”

Done.

Lines 215-222: I am glad to see that more researchers are finally adopting the use of -10 kPa (-100 cm)

as a standard value for the field capacity, and that the authors clearly state that the field capacity of a

soil is dependent upon the depth to the water table. I image some reviewers may push back against the

authors on this point, but the authors are both technically correct and do a good job of defending their

use of -100 cm.

Thank you. Another reviewer did indeed push back, so we appreciate this validation.

Results and Discussion: There is very little text in this section, and the text that is here is choppy and not

very informative. The authors should thoroughly improve this section, including a more thorough

presentation of results and improved discussion of their interpretation of those results.

Line 234: After “commercial field trials” add in parenthesis “East and West trials” for clarity.

This entire section was re-written to address this and the next comment, as well as other reviewer’s comments.

Line 234: Was a statistical test done to determine significance here? There are several instances where

this is not clear. I suggest the authors identify every instance of the word “significant” and insert pvalues.

Done.

If statistical tests were not used, do not use the word “significant.”

We added p-values to each instance of reporting statistical results.

Line 236: Table 1 contains no relevant soil information.

We added soil texture information to Table 1, but again this entire section was re-written and the reference to Table 1 was eliminated.

Figure 2: Big issue here! Only three of the eight pie charts shown here contain soil textural fractions

which sum to 100%. The one trial where both charts sum to 100% (Central-grain) shows no difference at

all in texture between the CC and no-CC plots. Since most of the authors results hinge on differences in

texture between plots, I suggest the authors double-check their textural analysis results and re-run any

statistical analysis or modeling that could have been influenced by this error. The results of the study

could be strongly affected.

This was due only to the default behavior of rounding in R (by default R rounds 2.5 down to 2), which was only used to create labels for the figures. We added a table with the texture values to supplemental material that includes texture values to the hundredths place precision with a column of summed values to demonstrate the components add to 100%. The analyses were not affected by the creation of the figure, so the results are still valid.

Line 253: It is not clear to me why the authors made the decision to neglect the organic matter analyses.

Hardly any information or explanation is given.

We added a more detailed explanation of results, with p-values.

Line 273: Replace “could not be predicted based on” to “are likely unrelated to.”

Changed.

Line 289-292: Two sentences does not make a paragraph. The authors need to expand this section to

include much more discussion.

Figure 5: The boxes below the figure labeled “Micropore” and “Macropore” are not helpful, especially

since the colors shown only apply for the “No Cover” columns. I suggest removing them, as they do not

add any useful information to the figure.

Lone 305: Again, a single sentence does not a paragraph make.

Conclusions: This study presents an enormous amount of data and supplemental material, but looking

only at the conclusions drawn I would have guessed that the