Reviewer #1: Summary: Maize can be grown in a continuous maize system, in a simple rotated maize system, or in an extended maize rotation. Previous research has shown an increase in yield in the extended maize rotation compared to a simple rotated maize system. Differences in resource distribution, microbial communities, and nutrient cycling have also been observed between these two systems. The authors investigate how the extended rotation influences maize root growth and its implications on yield. Over the investigated time period, maize yield is shown to be on average higher in the complex/extended rotation system compared to the simple rotation. This increase in yield also corresponds to an increase in rooting depth, less root biomass reported in the top 15 cm, lower soil penetration resistance, and lower soil moisture in the complex system compared to the simple rotation. The authors claim that the complex rotation system results in a cheaper and steeper root system which helps increase the likelihood the plant can withstand unfavorable growing conditions. They suggest the changes in soil physical, chemical, and biological properties contribute to this resilience.

Major comments:

1. The paper claims that the difference in root growth is due to the increased complexity of the extended rotation. However, it's unclear how the difference in fertilization method (inorganic vs manure) and differences in tillage practices (moldboard plowing after alfalfa and before corn) influence soil physical, chemical, and biological properties. Past studies on the same field acknowledge the possibility that fertilization by manure and moldboard plowing can contribute to differences in soil properties. These factors are mentioned at the end of the discussion, but it would be helpful to mention this in the introduction and in the results when discussing root growth and soil physicochemical properties.

We agree, we reworked the introduction and added text to explicitly acknowledge these complexities.

2. Due to the presence of significant root biomass at maize planting, the authors use the difference in root biomass from the beginning and end of the growing season to compare the two systems. This method leads to a negative root biomass in the 0-15 cm soil depth. While I acknowledge roots from a previous growing season could lead to an overestimate of the maize root biomass, the negative value suggests that a significant amount of the alfalfa roots decomposed by maize maturity. It does not seem accurate to take the difference in this case. Also, given that alfalfa has a favorable C:N ratio and the application of composted manure, the alfalfa roots are in very favorable conditions for rapid decomposition.

This is a good point, and prompted us to rework our root biomass analyses. We instead look at extreme cases of assumed background root decomposition. In all but the most extreme case (no background root decomposition in the short rotation and 100% background decomposition in the extended rotation) the statistics show the short rotation maize produced more root biomass than the extended rotation maize in the 0-15 cm depth. In other words, the conclusion does not change, only the magnitude of the difference. This lends confidence that our results are robust.

3. In addition, the negative root mass in the top 15 cm is ~125 kg ha-1 which accounts for a majority of the difference in total root biomass (~180 kg ha-1) between the two systems. Factoring in that there must be root growth in the top 15 cm that connect the root biomass at 15-30 cm soil depth, I would guess there isn't a significant difference in total root biomass.

Indeed, using our new analytical approach we found there was no conclusive evidence that there was a difference in total root biomass between the two cropping systems, only the way it was distributed in the soil. We added text to clarify this finding in the redrafted manuscript.

4. The root biomass addition method is also confounded by the fact that microbial activity decreases with soil depth. The deeper alfalfa roots may not decompose as readily as those in the top 15 cm. So, subtracting out the root biomass at planting at all soil depths may be accurate at deeper depths but not accurate for shallower soil depths.

By using the ‘extreme assumptions’ ranges noted above, we addressed this issue. The differences at the deeper soil depths were sensitive to the assumed decomposition, so we do not claim differences between the cropping systems in root production at those depths (>15 cm).

5. The method for collecting root biomass described by the authors likely has significant error and likely contributes the large error bars in Figure 3. Root biomass is collected from 4 samples collected from 32 mm diameter (1.26 inch) soil corers. This is roughly 32 cm2 (5 square inches) area in total. There is likely an extremely high variability in root biomass especially at lower depths where roots are much sparser and would depend on the ability to capture a root in that segment. Thus, it's hard to compare the systems at depths below 15 cm.

We agree with the reviewer that there may have been true differences between cropping systems at depths below 15 cm that we were unable to detect due to our sampling scheme. We added text to the manuscript to acknowledge this possibility. We note that the coefficients of variation of the measurements showed no pattern with regard to depth, so taking more subsamples in each plot may not have reduced the variation at lower depths, and therefore may not have resulted in more statistical power – it’s difficult to say.

6. Authors suggest the complex system is more resilient to extreme weather events which increases yield. The authors provide broad characterization of the average weather conditions for the measured years (hot vs cold and wet vs. dry). It would be nice to include some measure of the extreme weather events from those years (heavy rainfall, heat waves, droughts, frost, etc.) as these events and their timing may contribute to differences in yield.

We initially tried to parse out relationships between the yield differentials between the two rotations and more nuanced growing condition events, however it very quickly became complicated because ‘extreme’ events may only be extreme relative to the crop stage, and many do not impact fields uniformly (e.g. hail, extreme wind). It was therefore difficult to quantify how ‘favorable’ or ‘unfavorable’ a given year was without coupling the study with a crop model that can express drought, heat, or excess water stress in a quantitative way. In the re-drafted manuscript, we softened our language to avoid suggesting extended crop rotations can buffer against all types of extreme events.

7. The differences in root biomass are only reported for 2019 and 2020 when there is not significant difference in yield. It's unclear whether in a year where there is a greater difference in yield results in root biomass trends still hold true.

This is a good point - we added text and additional data demonstrating these patterns in a year with a significant difference in yield to address this concern.

8. In line 162, the authors describe the complex rotation maize roots as "more functional." And, in line 245, the authors claim the complex rotation achieved a more efficient root system with less resource investment. However, roots respond to resource availability as noted by the authors in a separate part of the discussion. Previous publications from the same field experiment note differences in POC, microbial activity, and soil physical properties. It would be more appropriate to frame the difference in root system architecture to be a result of differences in soil properties rather than the maize plant investing in steeper and cheaper roots especially since the maize genotype is the same. In the Lazicki et al., they note that differences in root length density seem to correlate with POC content.

This is a good point. In the re-drafting of the manuscript we frame the changes in soil physical, chemical, and biological properties due to cropping systems as creating opportunities for dynamic responses by maize roots, which may be linked to yield enhancement under certain conditions.

Minor comments:
1. Degradation of the alfalfa root biomass will impact N and P availability and likely influences rooting. In addition, the weather conditions will influence microbial activity and litter decomposition rates.

Yes, that’s true. To accommodate differences in nutrient addition between cropping systems, N fertilization followed protocols associated with the Late Spring Nitrate Test, and P fertilization followed recommendations based on soil samples drawn from individual plots.

2. In table 3, it's unclear how the timing of maize growth advantage is determined/calculated.

We added a superscript to clarify this – early refers to time periods before the maximum growth rate occurred, and late to time periods after the maximum growth rate occurred.

Also, are the p-value based significances comparing values (ratios and timing) between years? I was confused by how p values were determined for this table.

We added references within the table caption to clarify where the values and significances came from and added text to clarify that the ratios of absolute values are presented only to aid in visual comparisons of the values/patterns.

3. In table 3, please include the values for 2020 and 2019. Even though it is not significant, the root information is mainly for those two years and would help put into context how root growth may impact yield.

We reworked the entire table to include the root data and simplify its visual message.

4. The authors note that soil from the complex system resulted in plants with finer roots when grown in the greenhouse. It would be nice to add in a small discussion about how microbiome composition can influence the plant allocation to coarse vs fine roots. For example, root order has been shown to impact microbiome composition, and AMF have been shown to induce lateral root growth in plants. This may be especially important given that a previous publication focused on these fields (King and Hofmockel) observed higher microbial biomass in the more complex rotation.

We did not quantify differences in coarse versus fine roots in our study, so we chose not to include this in our discussion.

5. Inorganic N is highly mobile in soils, and increased wetness/soil moisture may result in more leaching and less N taken up by the plant in the simple system versus the complex system which has more sustained release through the organic N source (manure).

Yes, see Figure 2 in Tomer and Liebman (2014, http://dx.doi.org/10.1016/j.agee.2014.01.025).

6. There is a difference in the inorganic N applied at V6 between the simple and complex systems. Is this due to differences in soil nitrate measurements at that time? If so, this could point to differences in soil N over the growing season that may influence yield as mentioned in the previous comment.

The difference in applied N at V6 was due to differences in soil NO3-N concentrations and recommendations for side-dressing from the Late-Spring Nitrate Test protocol that were cropping system-specific. We added this to the management table to clarify this.

7. At line 98, the authors note that the difference between system additions was not statistically significant. Even if it is not statistically significant, there is a noticeable difference as shown in Figure 3, and this trend is consistent with previous root data (root length density measurements from Lazicki et al.) from the same fields. Lazicki et al observe a more even distribution of roots throughout the soil column in the complex system versus the simple system.

We reworked the statistical analysis and in the process addressed this issue. Additionally, we referenced Lazicki et al.’s work more explicitly, adding a supplemental figure summarizing their findings.

8. The higher penetration resistance at depths below 30 cm in the complex system may suggest that the biopores created from deeper alfalfa roots contribute to the differences in rooting depth and the more even distribution of root biomass along the soil profile.

We added this to the discussion. We are unsure whether the biopores would show up in the resistance measurements, so changed the terminology to refer to the penetration resistance of the bulk soil, as we believe that is what is being measured.

Reviewer #2: This study enhances understanding of the 'rotation effect'in agricultural production systems, and provides new data on yields and roots by long term positioning experiment. However, it is unclear what the physiological and ecological mechanisms are, and further supplementary data is needed to support the Conclusions.

Reviewer #3:
This manuscript sets out to show whether changes in root systems explain the rotation effect. It reports many measurements and concludes that chemical and/or biological factors are responsible but physical factors are not. These conclusions at the end of the Discussion are hard to justify because neither chemical nor biological factors were measured.

We rewrote the results and discussion to clarify the evidence that supports our conclusions.

The strength of the manuscript is the extensive data collection, the mostly complete statistical analysis and clear expression. The weakness is that the control of the rooting depth and its connection with yield are not clearly explained. This comes about because the long-term experiment was not designed to clarify the topics discussed in this manuscript. The authors should make revisions that more closely link these data to yield. The catchy title deserves retention; it is a little deceptive because cause and effect are unclear and this should be clarified in the abstract and conclusions.

We hope the redrafting of the discussion section has led to a clearer proposed link between our data and yield. Additionally we adjusted the language in the abstract to clarify our evidence and the conclusions we draw from it.

The manuscript implies that the reported 'complex rotation system' is representative of complex rotation systems in the Midwest US. Clear evidence for such equivalence should be reported but if it is not available the manuscript should address only the particular four-course rotation system that was studied and not extrapolate the results to other, undefined, 'complex' rotations.

This is a valid concern. We believe we address this in the introduction where we define ‘general’ extended rotations for the Midwestern region of the US as those including a small grain and/or forage:

Midwestern maize-based systems fall into three main categories: (1) continuous maize systems, wherein maize is grown for two or more consecutive seasons; (2) simple rotated maize systems, wherein maize is rotated with soybean; and (3) extended maize rotations, wherein maize is grown in a rotation with two or more years between maize crops, often including a small grain such as oats (Avena sativa) and/or a forage crop such as alfalfa (Medicago sativa).

This applies strongly to references to papers where it is not clear if the complex rotation is identical to the one reported here.

All citations extend from this definition presented in the manuscript:

The maize yield advantage accrued from extending short rotations to include small grains and forage legumes has received less attention compared to the continuous maize penalty, but has likewise been well-documented (Liebman et al., 2008; Stanger and Lauer, 2008; Coulter et al., 2011).

While it is almost impossible for a rotation to be identical due to regional differences in markets, soils, etc. all of our ‘extended’ rotation references refer to maize rotations that include a small grain and/or forage.

Value-neutral names of the systems would be two-course and four-course rotations.

We changed the terminology to ‘short’ and ‘extended’.

Attributing maize yield to a system is less convincing than to a particular part of the system such as crop species or management operation. It is difficult to unscramble such effects from the data but the authors are in the best position to do so.

We rewrote the introduction and parts of the discussion to make it clear why it is difficult to attribute such a complex characteristic such as yield to a single change, when extending rotations involves a slew of changes that cannot be controlled for individually. We try to present the factors that may be contributing, and in reality is unlikely it is one factor.

The emphasis on nitrogen is unbalanced. It appears on the second line of the abstract and in many parts of the manuscript but there are no reported measurements of soil mineral N. This is a matter of concern since there is much published evidence that crops growing after alfalfa benefit from residual N arising from biological N fixation. If there are no such data available for this experiment, are there published or unpublished data from other studies to fill the gap? The N supply to maize in the four-course rotation also includes a large amount of N in manure, as reported in Table 1 but N-supply is not discussed as a possible reason for the higher yield.

We added citations and text to demonstrate these yield enhancements are observed even when nitrogen is not limiting. We also added text to Table 1 to demonstrate how the amounts of N applied to the two systems were balanced by the side-dressing which was based on a late-spring nitrate test. While the forms of nitrogen supply differed, the absolute amounts did not.

Please provide as much information as possible about the oats and alfalfa; were they cut or grazed and is it possible to report the amount of N contained in alfalfa, both above-ground and below-ground.

We added information indicating the management of the oat straw and alfalfa (oat straw and alfalfa hay were harvested, baled, and removed from the research site, no grazing occurred on these plots). We do not have data on the N content of the below-ground alfalfa biomass. Osterholz et al. 2018 does an in-depth exploration of nitrogen pools in these systems and found they could not explain the differences in maize nitrogen dynamics/yield – we added a more explicit citation of that work.

At L107 there is mention of biological and physical effects on root systems but less prominence is given to chemical factors. The nearest that the manuscript comes to reporting chemical effects is the large amount of manure-N. This should be strengthened.

We added mention of ethylene build up and its potential role in root growth. See above comment concerning N.

The shallow rooting depths (Figs 2 and S2) in some seasons is explained by excess soil water. Could the deeper maize roots in the four-course rotation be due to dewatering of the soil profile by the previous alfalfa crop?

The shallow rooting depth in 2018 was due to high water tables. Our soil moisture sensors at 45 cm depths indicate no difference in soil water contents at that depth between the two rotations, and indicates the soil was fully recharged with water following the alfalfa crop termination the previous fall(s). In general, in this region the entire soil profile goes through a resaturation event before spring's maize planting activities.

The discussion of the effect of alfalfa should include mention of hydrogen fertilisation. Presumably there was hydrogen fertilisation by the alfalfa rhizobia but any contribution by soybean rhizobia would depend on whether they were HUP-plus or HUP-minus.

We considered mentioning this in the original manuscript draft, but in the end felt it was a distraction. The soybean inoculant was HUP-minus, so any differences would be a result of the different quantity of roots releasing H2, and we have no data on the soybean nor alfalfa roots in these plots. While certainly and interesting area of research, we opted to omit it in this manuscript.

The end of the Discussion suggests that roots in 'more complex' rotations are better buffered against unfavorable growing environments. In Fig. 1 the four-year rotation seems to perform best in the wet seasons of 2016-2018. This apparent inconsistency should be discussed.

While the absolute yields of the four-year rotation were highest in 2016 and 2018 as the reviewer points out, we were interested in how the four-year rotation performed RELATIVE to the two-year system. The size of the points on the right panel represents this RELATIVE difference, and it shows that this relative performance was higher in a range of weather conditions (for example in 2013, a hot and dry year, and in 2018 (a hot and wet year).

Detail

Report the statistical significance of the higher yield, rooting depth, root biomass and penetration resistance in the Abstract and Highlights.

Word limits precluded the inclusion of statistical significances in all instances, but we added the significance in several places within the abstract. We hoped that the context would convey that any results reported as being different can be assumed to be significantly so.

The Conclusions (L258) that the result is 'novel' is inconsistent with many papers referred to in the manuscript.

We adjusted that language to better articulate the exact component of our study we feel is novel.

The renumbering of lines makes it difficult to review the manuscript. Please stick to one system in any resubmission.

Yes, we apologize, the line numbers restarted after the landscape table which required a section break. We have corrected this issue.

The references are reported in several different formats and should be standardised.

We corrected the references, switching between citation systems is always a challenge.