Diagram

Description automatically generated

Figure showing the ratio of average NASS acreages to estimated field acres from 2008-2021 by individual crop, using threshold 1. Top (A) is Illinois, middle (B) is Michigan, Wisconsin is bottom (C). Log2 fold change applied.

Takeaway: Major crops do quite well, lower acreage crops less so. It’d nice if I had a way to show relative acreage; violin plots don’t do a great job of this, but I could maybe replace the boxplots with beeswarm (proportionally sized points)

Graphical user interface, chart

Description automatically generated

Illinois line plots of summed acreages over time; black is NASS, grey is CDL

Graphical user interface, chart, histogram

Description automatically generated

Michigan line plots of summed acreages over time; black is NASS, grey is CDL

Timeline

Description automatically generated

Wisconsin line plots of summed acreages over time; black is NASS, grey is CDL

Chart, scatter chart

Description automatically generated

Figure showing the average sum crop acreages across the thresholds from 2008-2021, with the black line indicating the average NASS acreage

Takeaways:

I’ve checked and double checked the CoA in terms of making sure I’m counting and joining the categories correctly; Acres Harvested or Bearing, with a few key words to exclude # of farms. For example, irrigated and non-irrigated seem to be two different types of acreages, when previously I had only included acreages harvested (and excluded irrigated, as I was unsure if they were lumped in the acreages harvested); all evidence I can find says they are grown as seperated acreages.

I also want to triple double check why some counties like Waushara are showing a lower NASS than CDL for all years; it is not unusual for NASS to be lower than CDL for non-CoA years (in fact, it may be best to simply eliminate non-CoA years all together), but it should be greater than or about equal to CDL for 2008, 2012, and 2017. This could also be a CDL vs CoA thing for grassland/hay specifically. I want to look at the statistics for alfalfa in Waushara county.

For counties like Oceana, inspection shows a patchwork of extremely small, poorly delineated areas. It’s not surprising to me that we would miss a large chunk of fields, especially if we’re working on the assumption that fields are more consistent across 14 years. The other aspect is that the CDL has a notoriously hard time differentiating between grassland/pasture and alfalfa/hay categories for crops (see Lark et al. 2021), which is a large factor in parts of Michigan and Wisconsin (and makes the RPBB question even harder in terms of capturing accurate grassland exposures).

Solutions:

-Break up the thresholds into smaller sets (instead of 14 years, perhaps do average years threshold? Ex: 6-8 years for Michigan

-Have smaller focal window sizes to match the average field sizes in these landscapes

-lower threshold for exclusion of edge pieces, lone patches of acreage

Question is whether I should go ahead and attempt those changes for Michigan OR leave as is and just suggest these changes in manuscript. Those changes would require extra time and effort, and re-running/re-writing some chunks of the script. But it would be faster as most everything is written. So the question is whether we present the method as is specifically as a good fit for areas with high corn/soy and suggest how it can be adapted to fit other areas, or if we present the method as dependent on the average field life and show that.