Spatial Analysis of Corn Yields June 2020

Purpose

Determine if a spatial approach to corn yield analysis can be useful in screening producer yield reports to RMA

Background

CAE has provided us with sub-county yield data with locations that are specific enough to enable us to map the yields. The goal is to see how well we can predict yields at a location based on surrounding reports.

First Step – PRISM-ELM Corn Yield Study

Apply the PRISM mapping system to corn yield reports and assess strengths and weaknesses of this approach as well as data quality.

Contents

Slides 2-6

Summary of mapping methods, performance statistics, discussion of high local yield variation, and next steps

Slides 7-10

Examples of the extent and density of yield reports used in PRISM

Slides 11-13

Screenshots of the PRISM interface showing examples of high local yield variation

Slides 14-15

PRISM maps of 12-year mean corn yield and 70% prediction interval

Slides 16

PRISM mapping performance statistics by year

Slides 17-40

2007-2018 PRISM maps of yield as actual and percent of the 12-year average

Mapping Methods

Data Filtering Criteria

Corn for grain, non-irrigated practice Only used reports from CLU acreage sizes <= 4 grid cells Period of record >=3 yr, cause of loss = 3, prevented planting flag = 0 Valid Range: 0-399 bu/ac, all others omitted

Mapping Procedure

For each year, 2007-2018 (pass 1):

- Interpolate yield with PRISM using PRISM-ELM ESI grid as predictor
- Produce a 70% regression prediction interval (PI70) grid
- Average the resulting yearly yield grids to create a 12-yr average yield grid
- Average the resulting yearly PI70 grids to create a 12-yr average PI70 grid

For each year, 2007-2018 (pass 2):

- Interpolate each year's yield with PRISM using 12-yr average yield grid as predictor
 - This is similar to the "Climatologically-Aided Interpolation" (CAI) method that we employ in our climate mapping, where the local long-term average spatial pattern is a good first guess of the spatial pattern in any given year. In this instance, the 12-year average yield map is the "climatology."

Performance Statistics

- The PRISM MAE (mean absolute cross-validation error) of the interpolation averages is about 19 bu/ac, meaning that on average, a PRISM prediction at a yield report location would be within about 19 bu/ac of the actual reported yield value, when that report is omitted from the dataset.
 - This relatively low value is somewhat misleading, because each yield value is given equal weight in the average, over-emphasizing data-dense areas. As the PRISM interface examples show, there can be substantial variation well beyond 19 bu/ac within a very small area.
- The 12-year average PRISM PI70 (70% prediction interval) ranges from less than 20 to about 50 bu/ac across the corn growing region. PI70 is a measure of the scatter of the data around the linear regression line.
 - PI70 is greatest where and when yield reports are sparse (early years, outlying areas), and smallest in data-dense areas.
- The NRCS NCCPI (National Commodity Crop Productivity Index) was also tried as a predictor grid in place of ESI, but with little improvement in results. This is a topic of future work. (Statistics for those runs are not reported here.)

High Local Variation in Corn Yields

In all years, there was a large amount of local variation in yields that could not be explained by local variations in ESI or 12-year average yield grids. Given this lack of local explanatory power, the PRISM yield maps are spatially smoother than the data would indicate.

- Why this High Local Variation? Possibilities include:
- <u>Data QC issues</u>: Erroneous transcription of reports, yields modified by program adjustments (non-physical yields), irrigated (or partial irrigated) practice being reported as non-irrigated, outcomes other than grain (e.g., silage) reported as grain, and others.
- <u>Causes we cannot explain with current data</u>: Crop variety, soil amendments, management decisions, etc.
- <u>Scale-dependent variations</u>: Imprecise locations of yield reports (currently to the nearest 4 800-m grid cells), and variations in soil type, fertility, and topographic settings that occur at scales finer than the 800-m level.

Next Steps

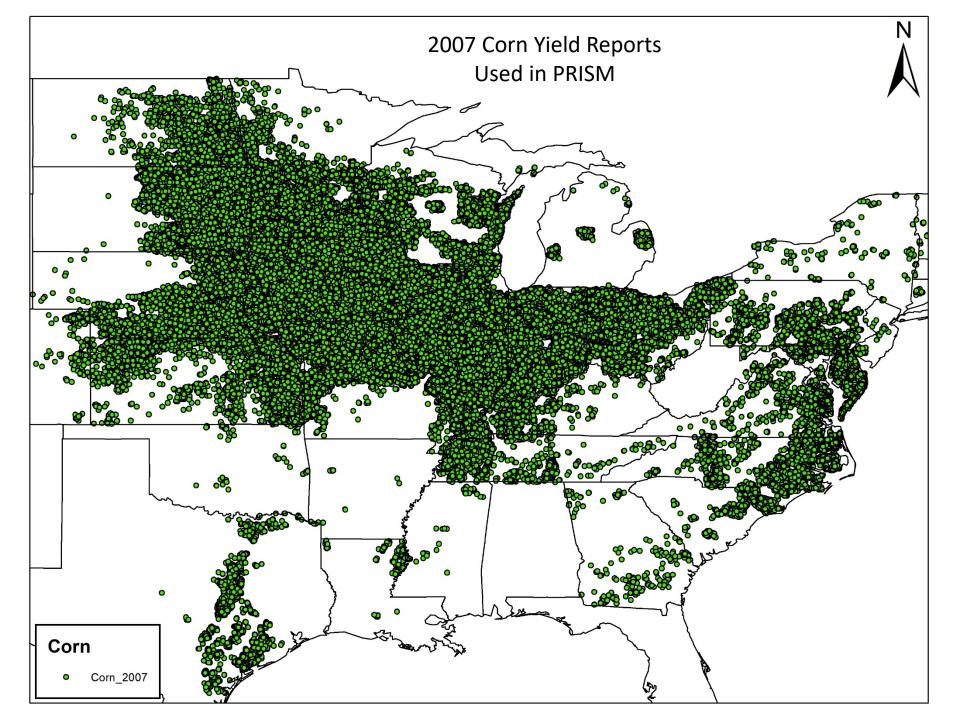
New dataset

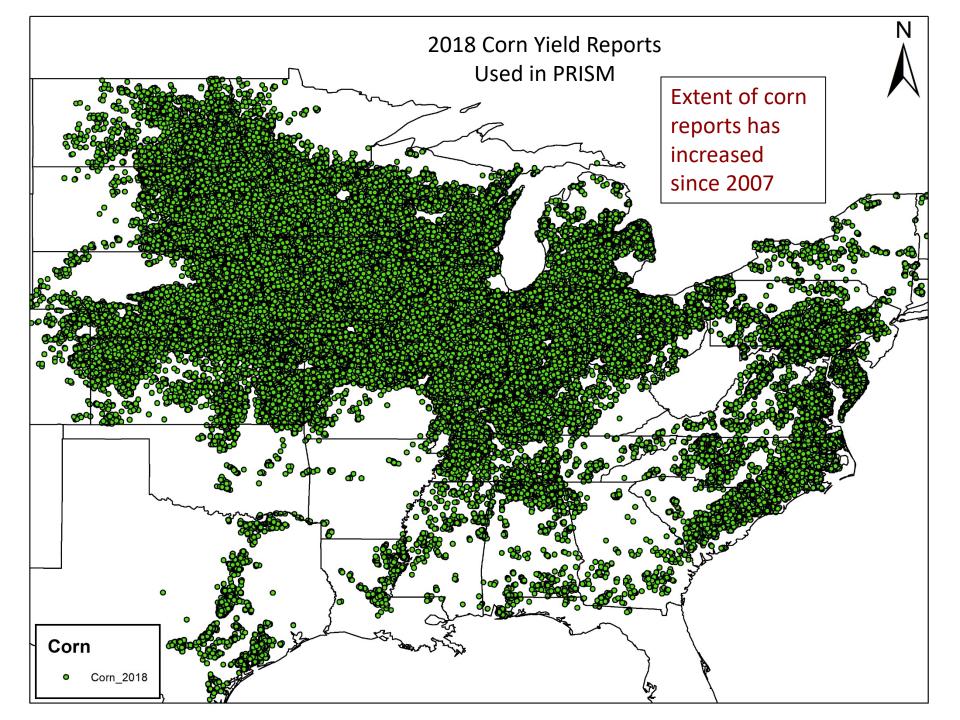
- CAE has delivered version 2 of the sub-county level data, with additional information and better QC screening. This dataset will be analyzed, and mapped with PRISM.
- Various methods of determining which data are to be mapped will be explored.

Explore fine-scale variations in local yield patterns

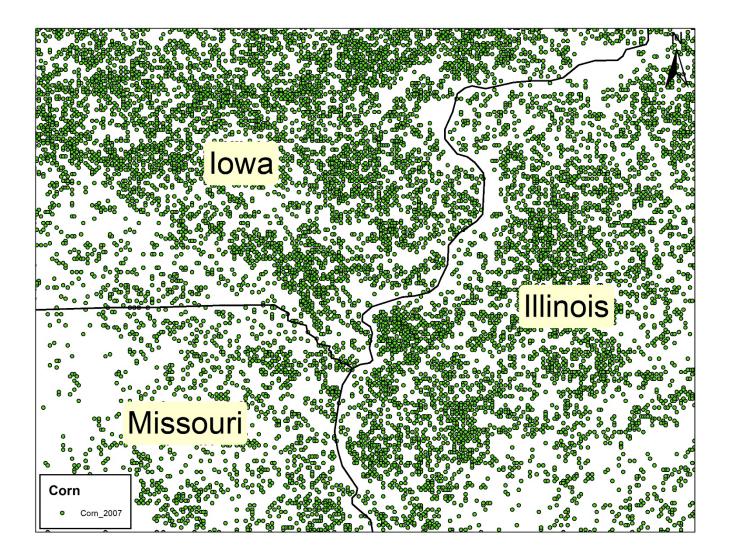
We are pretty good at predicting large-scale yield patterns, but not yet very good at smallscale ones.

- So far, we have limited the yield reports used for mapping to those from CLU acreages that are no larger than four 800-m grid cells. To tighten up the location precision, we will attempt to limit them further to CLU areas corresponding to one 800-m grid cell.
- The NRCS NCCPI native grid resolution is 30 m; we will attempt to overlay yield reports in local areas where the variation is high to determine if there is increased explanatory power when using this fine-grid information instead of an 800-m version. However, it is recognized that even yield reports located to the nearest 800-m grid cell may still not be sufficiently precise to support a 30-m assessment. Is it possible to obtain a limited amount of yield data with highly precise locations?
- If local yield variations are at all tied to local variations in growing conditions, they should be relatively persistent in time, at least in a relative sense. It should therefore be possible to create long-term patterns that best predict these variations. However, management variations from year to year will still be difficult to explain .

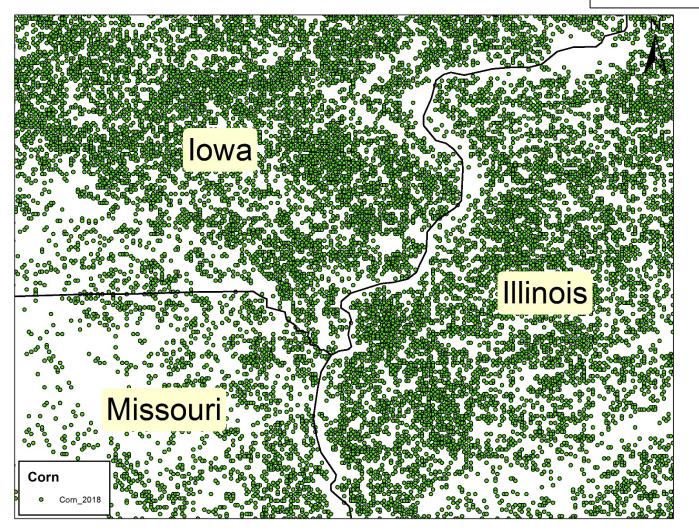




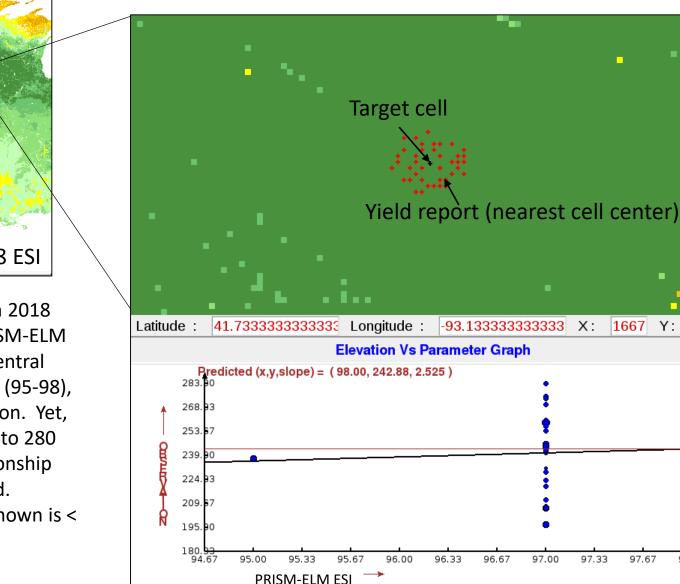
2007 Corn Yield Reports Used in PRISM – Zoom View



2018 Corn Yield Reports Used in PRISM – Zoom View Density of corn reports has increased since 2007



PRISM Interface Screen Shots Showing Wide Yield Variations over Short Distances PRISM-ELM vs. Corn Yield

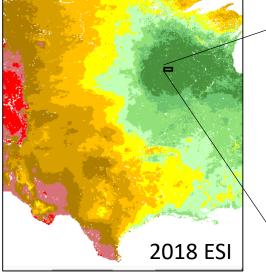


1667 Y:

97.67

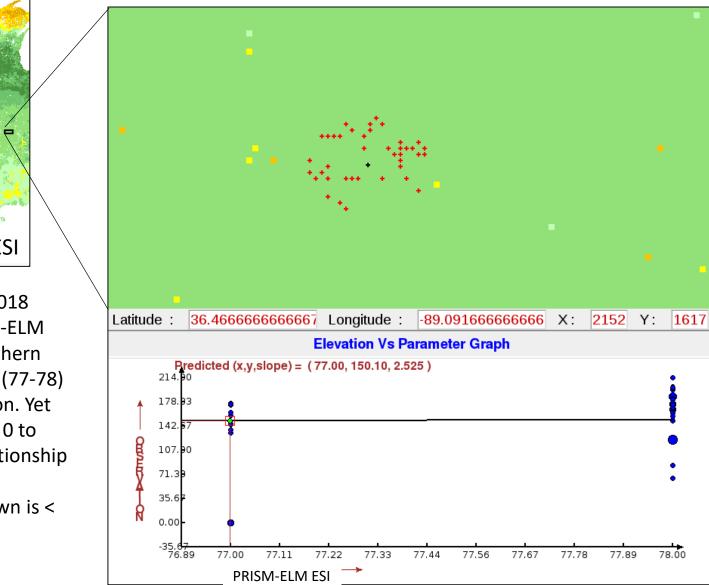
98.00

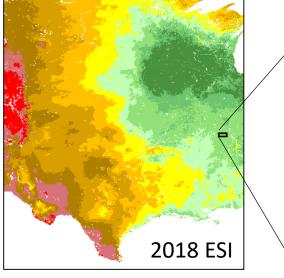
985



Relationship between 2018 yield reports and PRISM-ELM ESI for a location in central Iowa. ESI is very high (95-98), with very little variation. Yet, yield varies from 195 to 280 bu/ac. No local relationship between ESI and yield. Diameter of cluster shown is < 10 km.

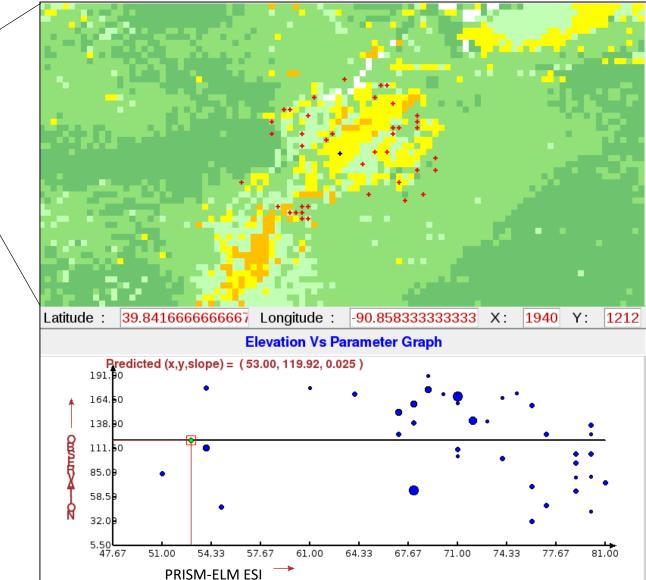
PRISM Interface Screen Shots Showing Wide Yield Variations over Short Distances PRISM-ELM vs. Corn Yield

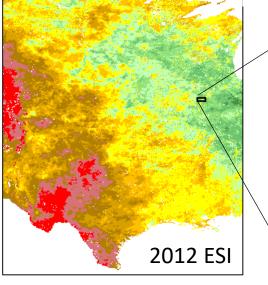




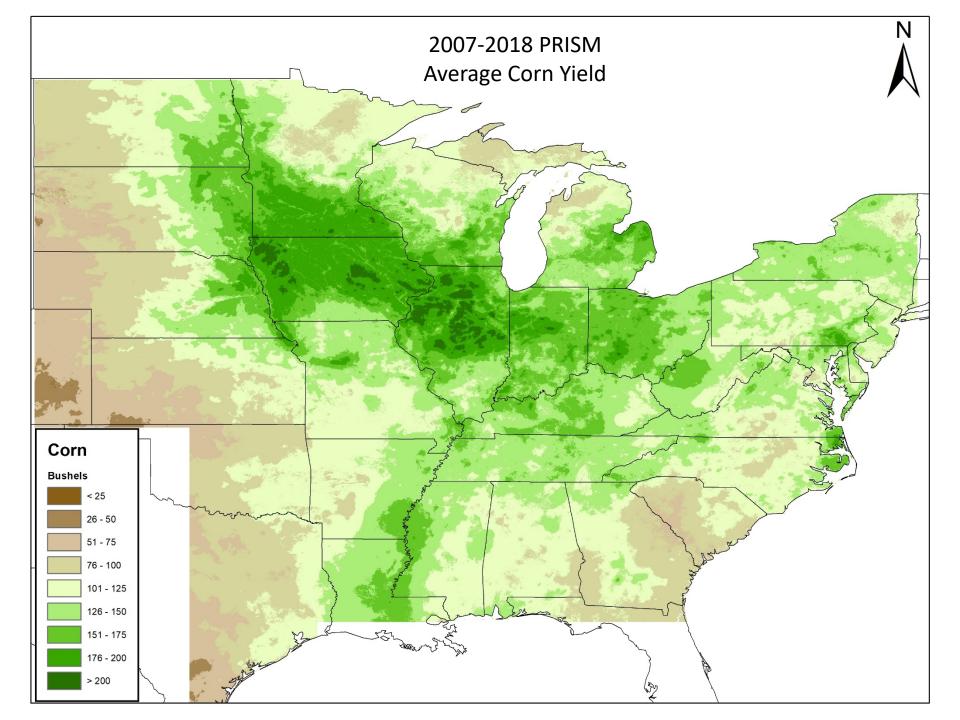
Relationship between 2018 yield reports and PRISM-ELM ESI for a location in southern Illinois. ESI is fairly high (77-78) with virtually no variation. Yet yield varies widely from 0 to 210 bu/ac. No local relationship between ESI and yield. Diameter of cluster shown is < 10 km.

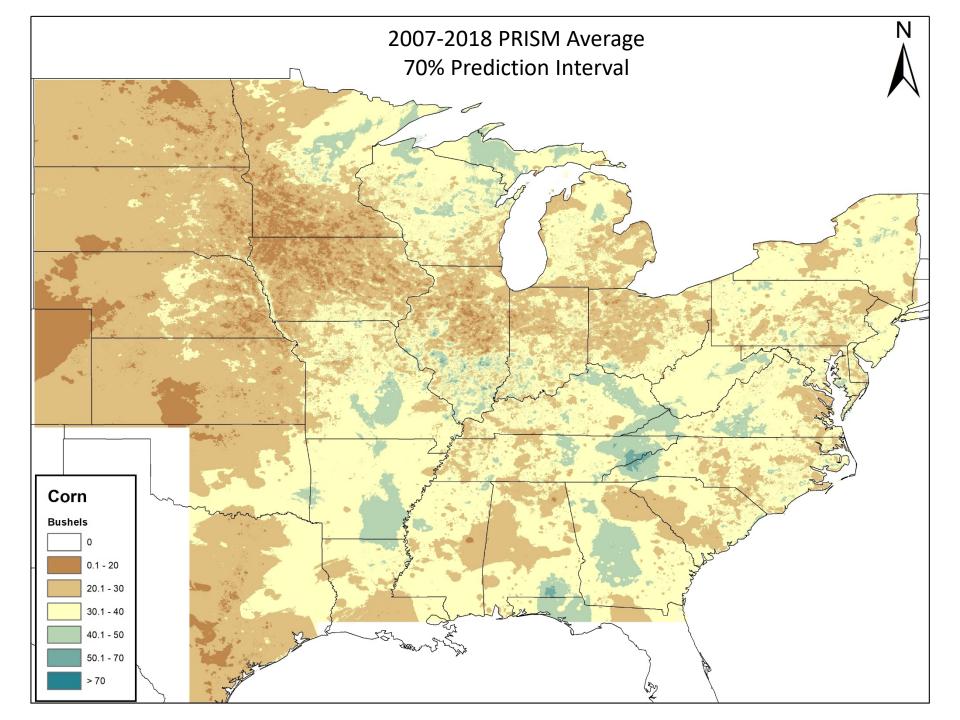
PRISM Interface Screen Shots Showing Wide Yield Variations over Short Distances PRISM-ELM vs. Corn Yield





Relationship between 2012 yield reports and PRISM-ELM ESI for a location in NW Illinois. ESI varies fairly widely (51-81). Yield also varies widely from 30 to 190 bu/ac, but shows no local relationship to ESI. Irrigation? This is a drought year. Diameter of cluster shown is ~ 20 km.





PRISM Mapping Performance Statistics

Year	No. of Reports Included*	Mean Absolute Cross- Validation Error (bu/ac)*
2007	213,207	16.64
2008	174,322	18.60
2009	248,354	19.23
2010	270,194	18.74
2011	284,778	17.62
2012	205,991	20.29
2013	237,956	20.36
2014	249,051	19.53
2015	323,836	19.22
2016	334,248	18.52
2017	302,675	22.50
2018	305,337	20.43

* Reported for the PRISM Central US region, which extends as far east as the longitude of middle of Lake Michigan.

