Crop Yield Calculations for Estimating Nutrient Application and Projecting Future Demand

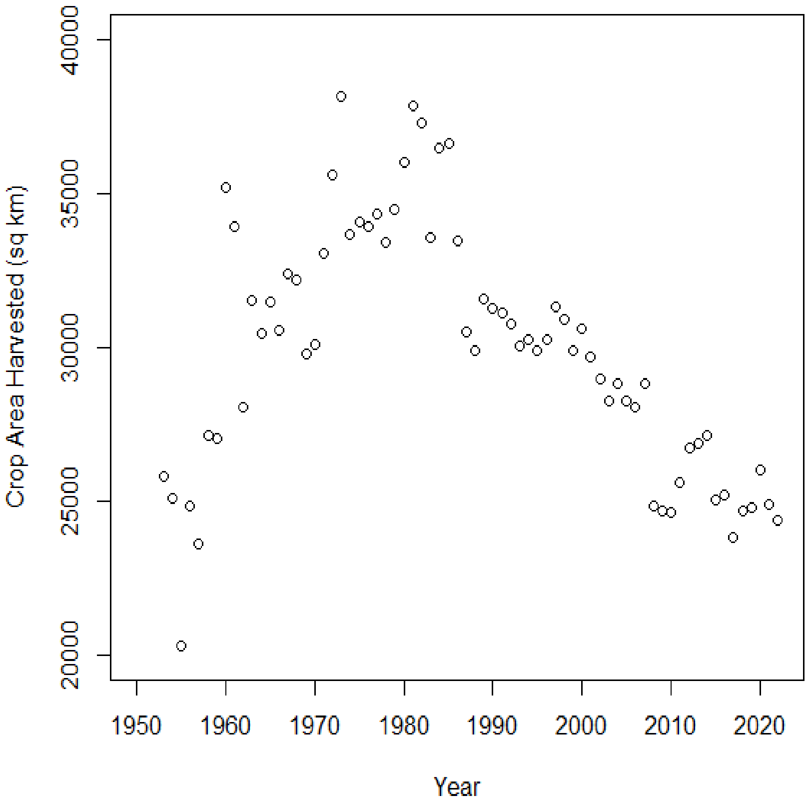
Joseph Delesantro

ORISE Fellow, CBPO Modeling Team

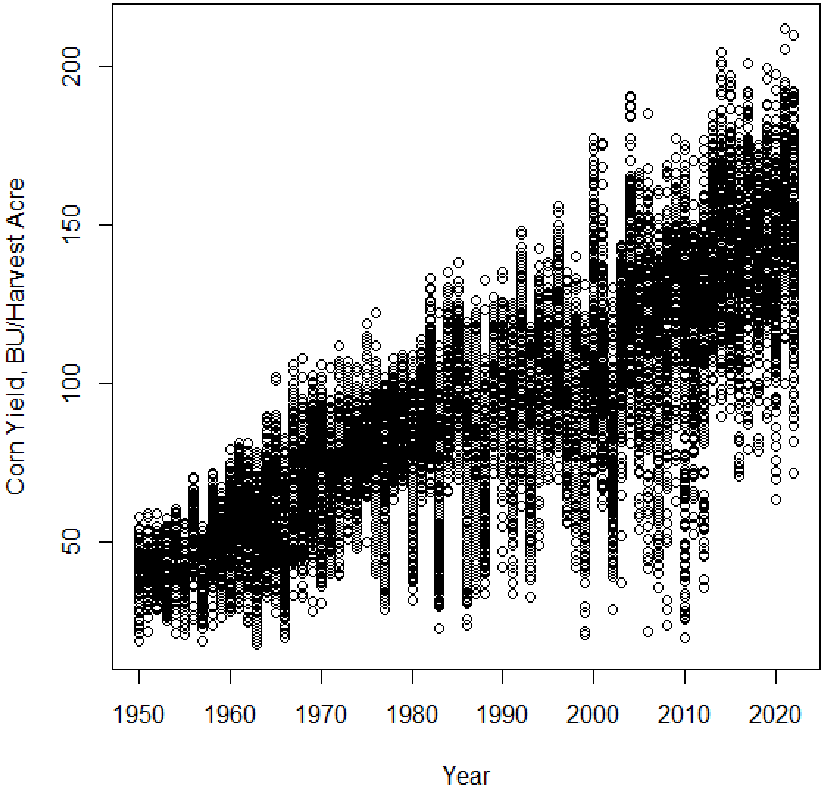
Photo: WILL PARSON/CHESAPEAKE BAY PROGRAM

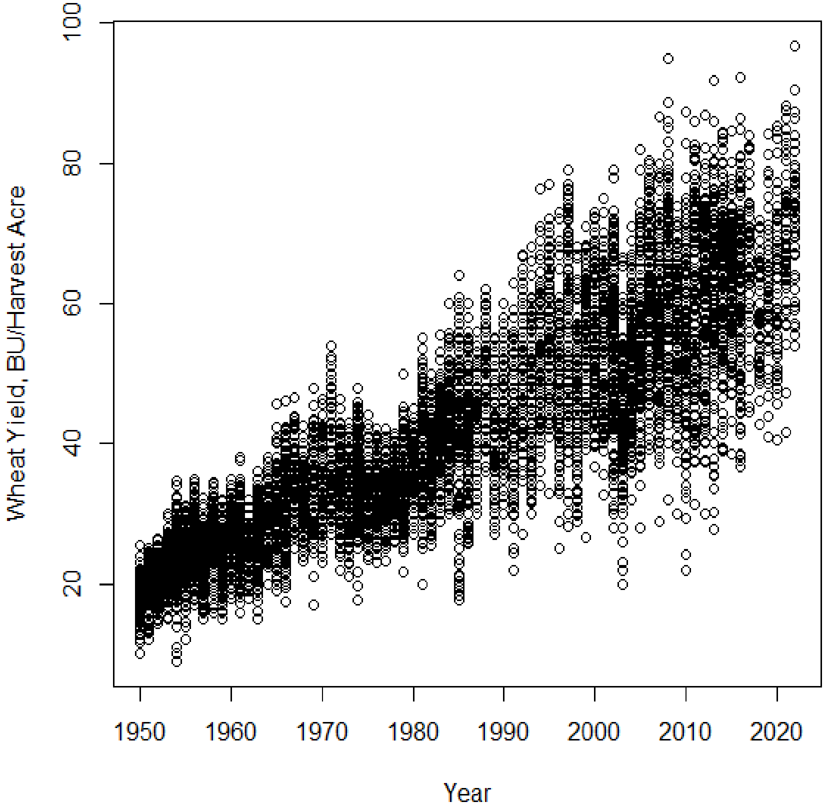
Changing cropland area… AND changing crop yields

CBW Cropland Area



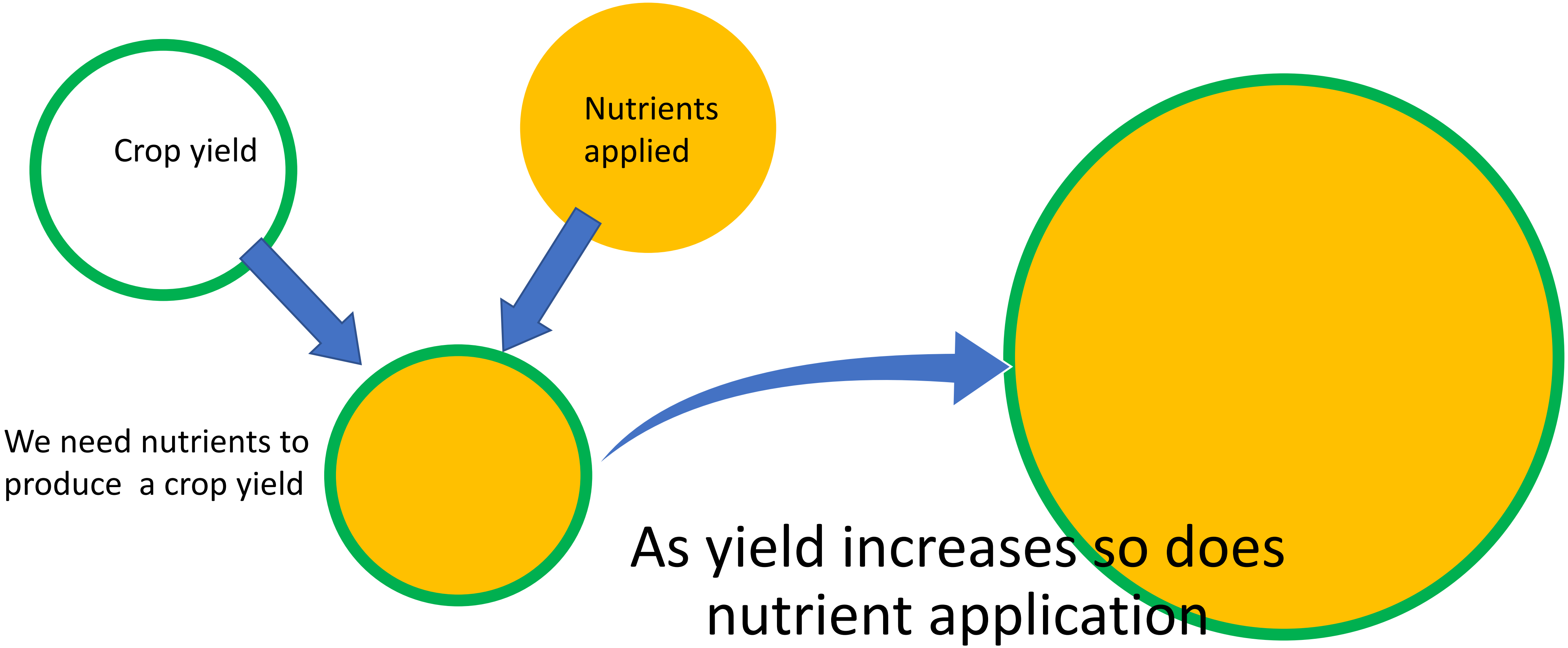
Example CBW County Crop Yields



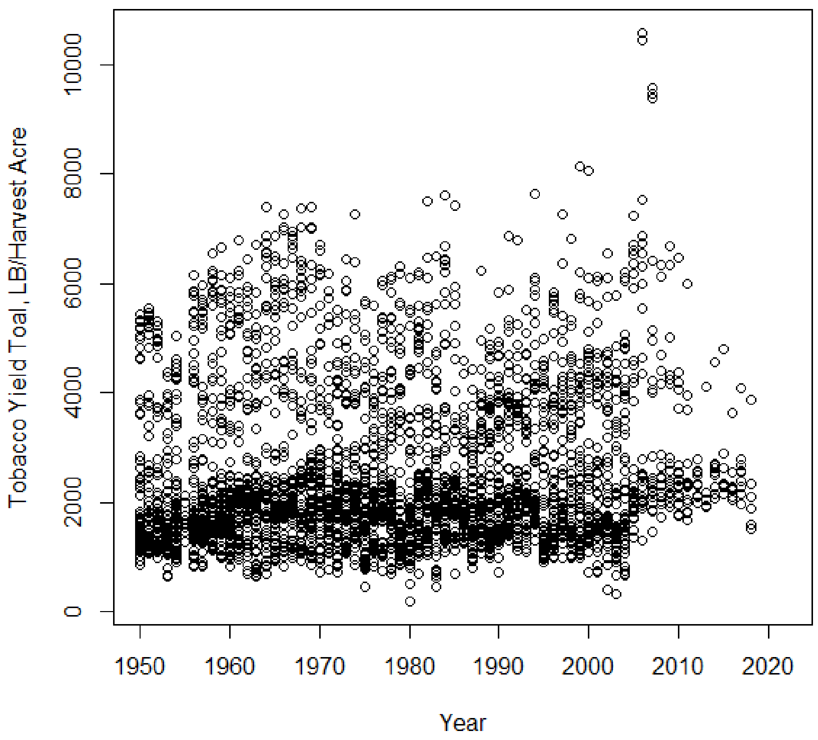


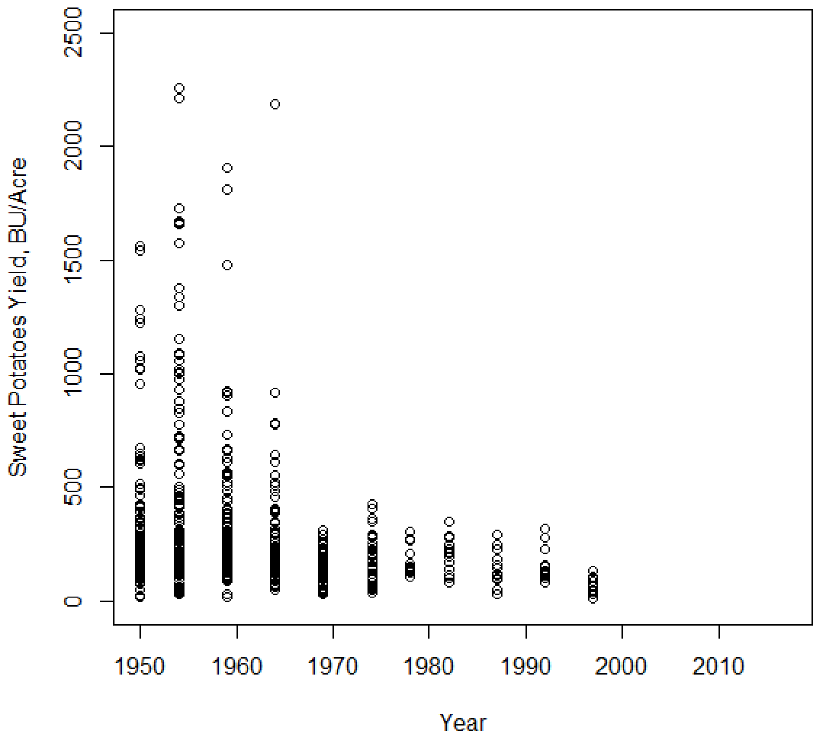
Why crop yields matter

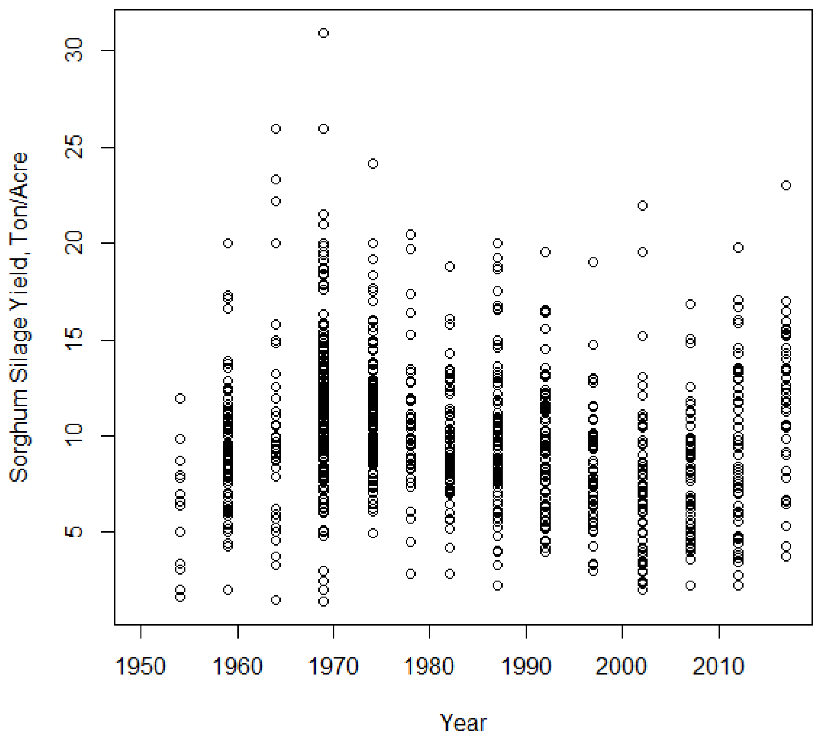
- Yields and nutrient applications are tied together



Changing cropland area… AND changing crop yields

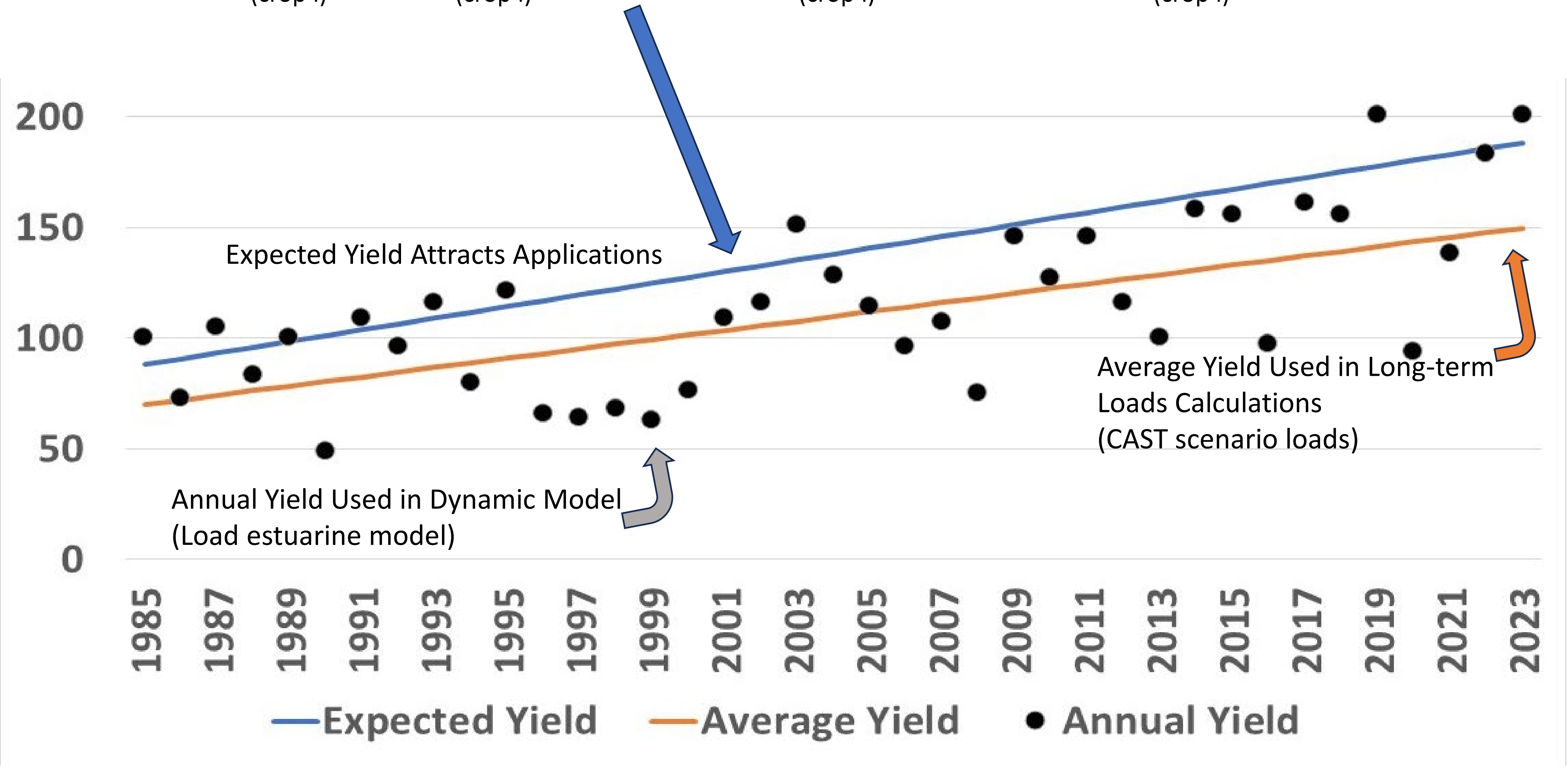






*EXAMPLE DATA ONLY

$$\text{N applied}_{(\text{crop } i)} = \text{Acres}_{(\text{crop } i)} * \text{Expected Yield}_{(\text{crop } i)} * \text{lbs N/unit yield}_{(\text{crop } i)}$$

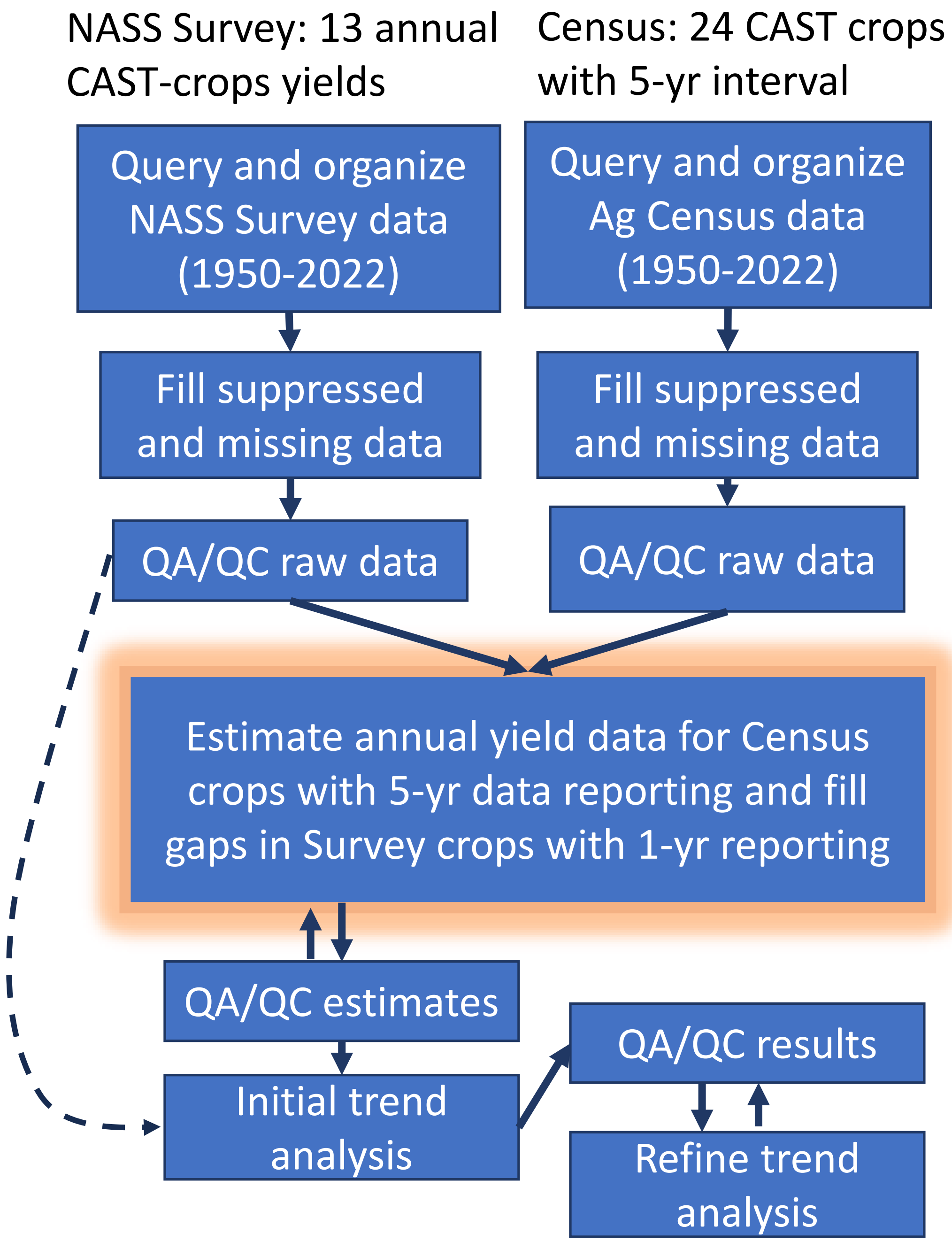


Planned path for investigation

Goals:

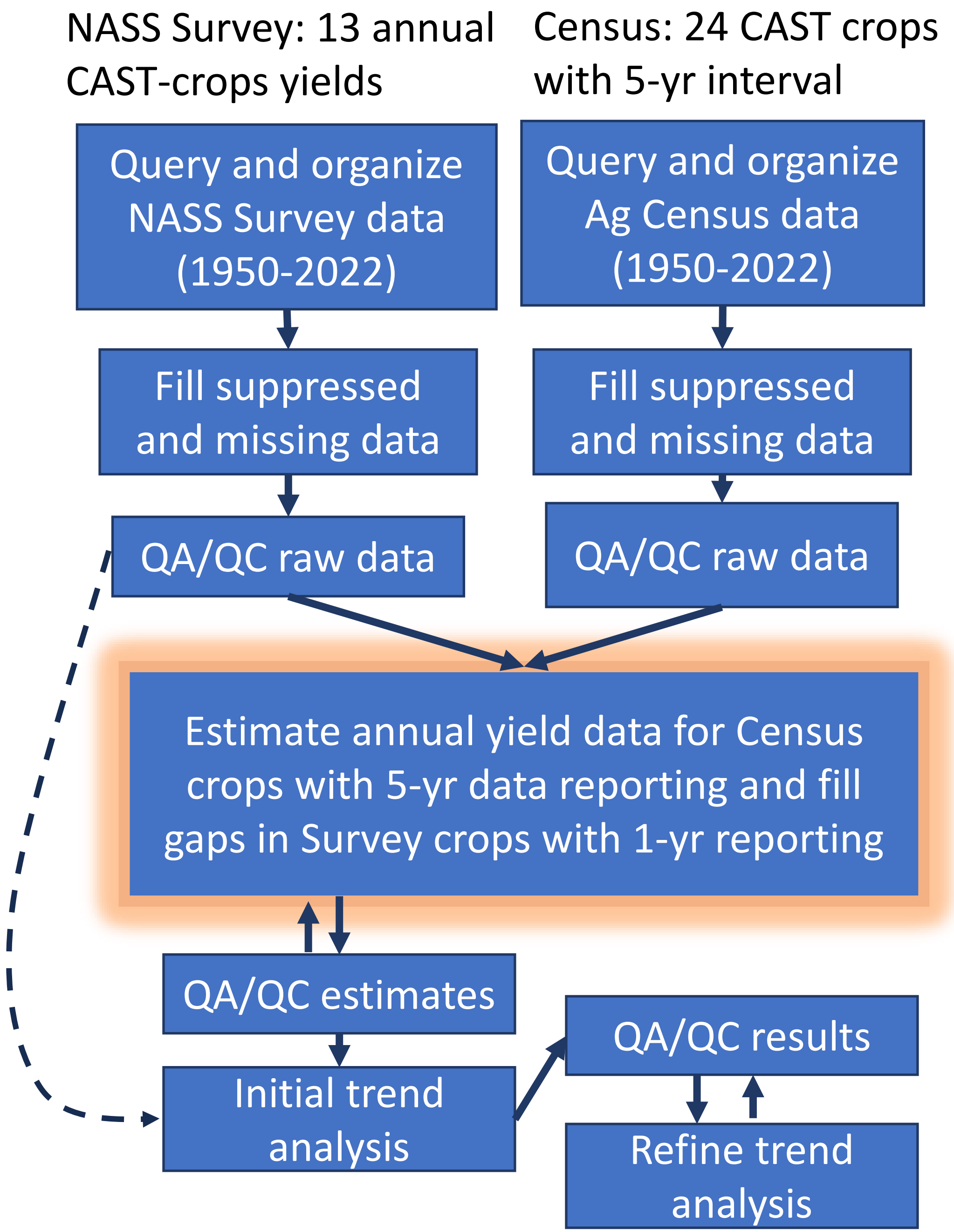
- Estimate farmer yield expectations at the county level which drive the application of nutrients.
- Estimate various yield trends to support several potential scenarios.

Approach: Use trend analysis of long-term annual crop yields.



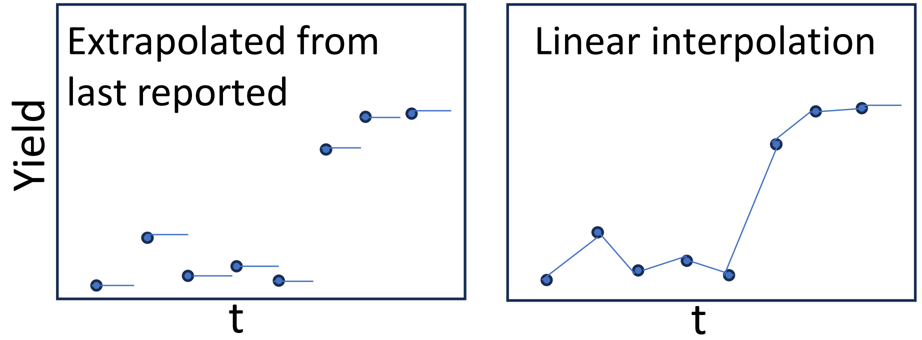
Crop data collection

- 94 CAST-crops with both a potential yield and N-application
  - Excludes pasture, fallow, unmanaged or wild covers
- “Complete” data for 24 of these CAST-crops
  - Complete = data spanning >85% of period 1950-2017
  - **91% of crop land area, 95% of N applied to crop land**
- Partial data for an additional 40 crops
  - Partial = partial spatial range, partial time range, state-level only
  - 2.2% of crop land area, 3% of N applied to crop land
- No yield data for 31 crops
  - 6% of crop land area, 2% of N applied to crop land



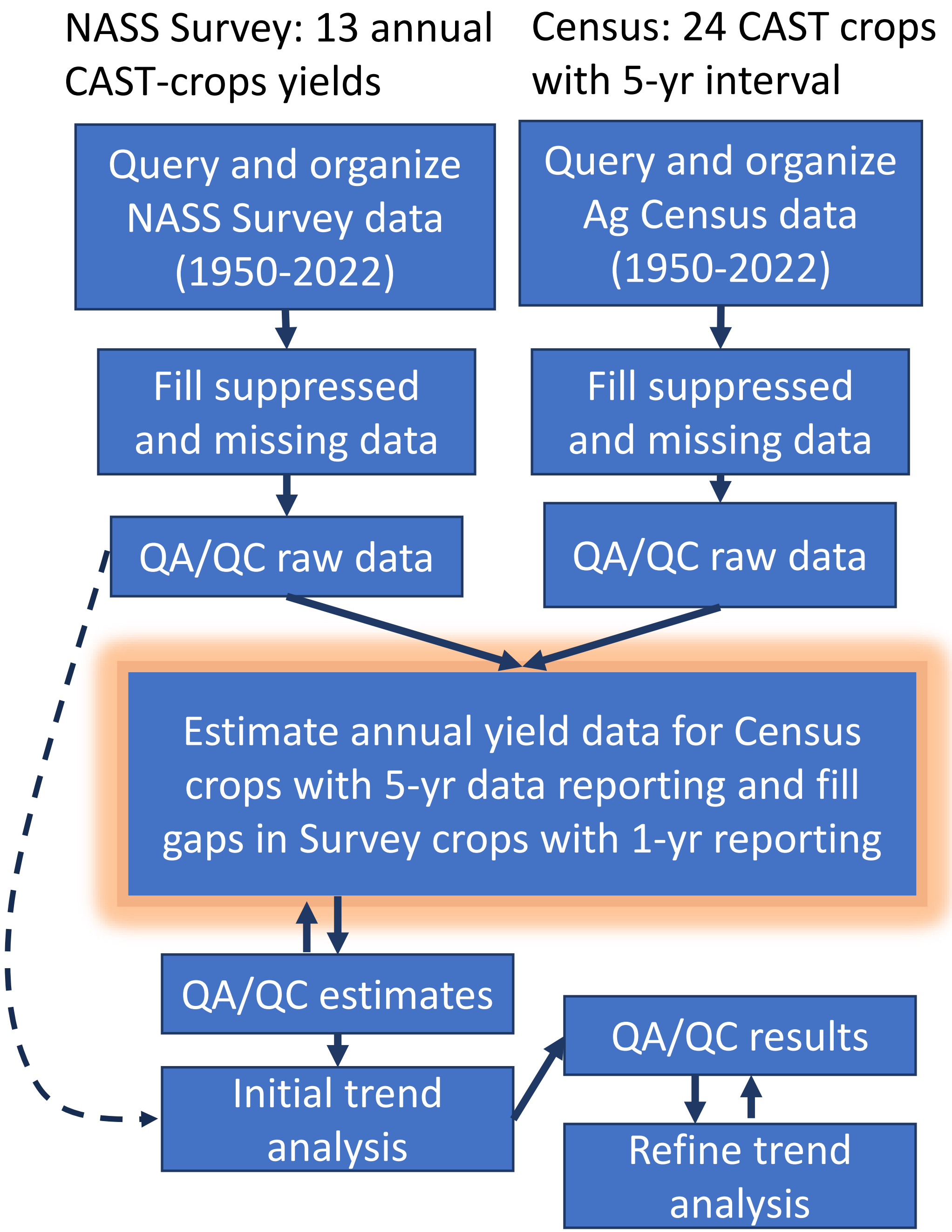
Annual estimation of yields from available data and environmental variables

- Data gaps or coarse data can introduce uncertainty and errors in annual estimates. Conceptual examples:



- Consecutive Census years coincident with bad weather/yield years can skew trend analysis.

  "USDA-NASS Census reported crop yields which reflected lower than expected yields due to negative weather influences. These reports have a dampening effect on the increase of crop yields in more recent years." –Mark Dubin



Aggregate to growth regions for more consistent yield data

The problem

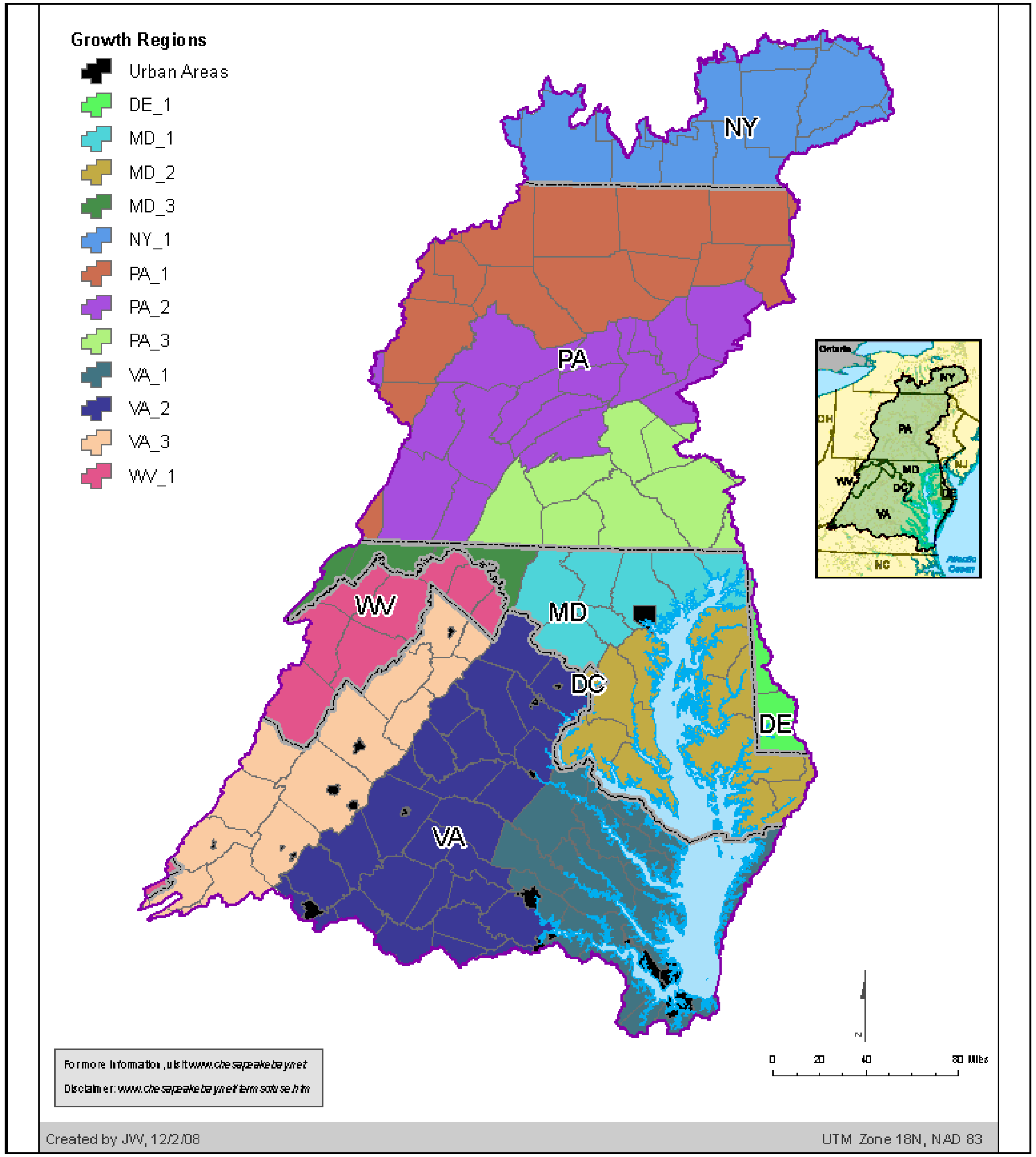
- At county level survey response data can be noisy
- There are gaps in county level data for many reasons

The solution

- Aggregating to the growth regions averages out noise and creates more complete time series
    - Aggregation=Weighted average by the county cropland area. Same method applies to yields and predictors.

Trade offs

- Compresses variation in geospatial predictors and reduces spatial statistical power (from many counties to few growth regions)
- The growth regions become the most important geospatial predictor



Potential statistical modeling methods

$\text{Yield}_{\text{crop } i} \sim f(\text{growth regions, time, weather})$

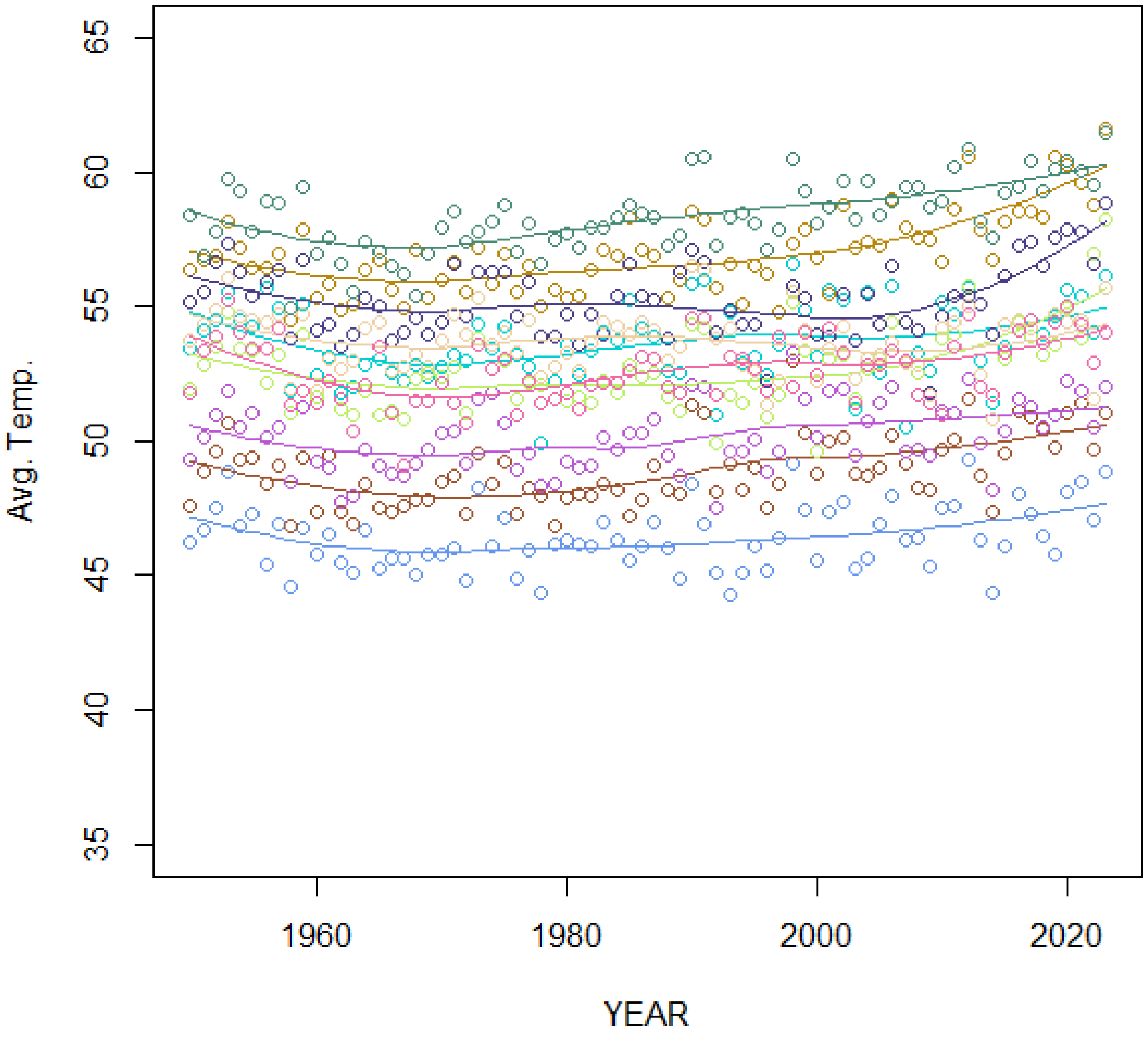
$\text{Yield}_{\text{crop } i} \sim f(\text{growth regions: weather, GR: time})$

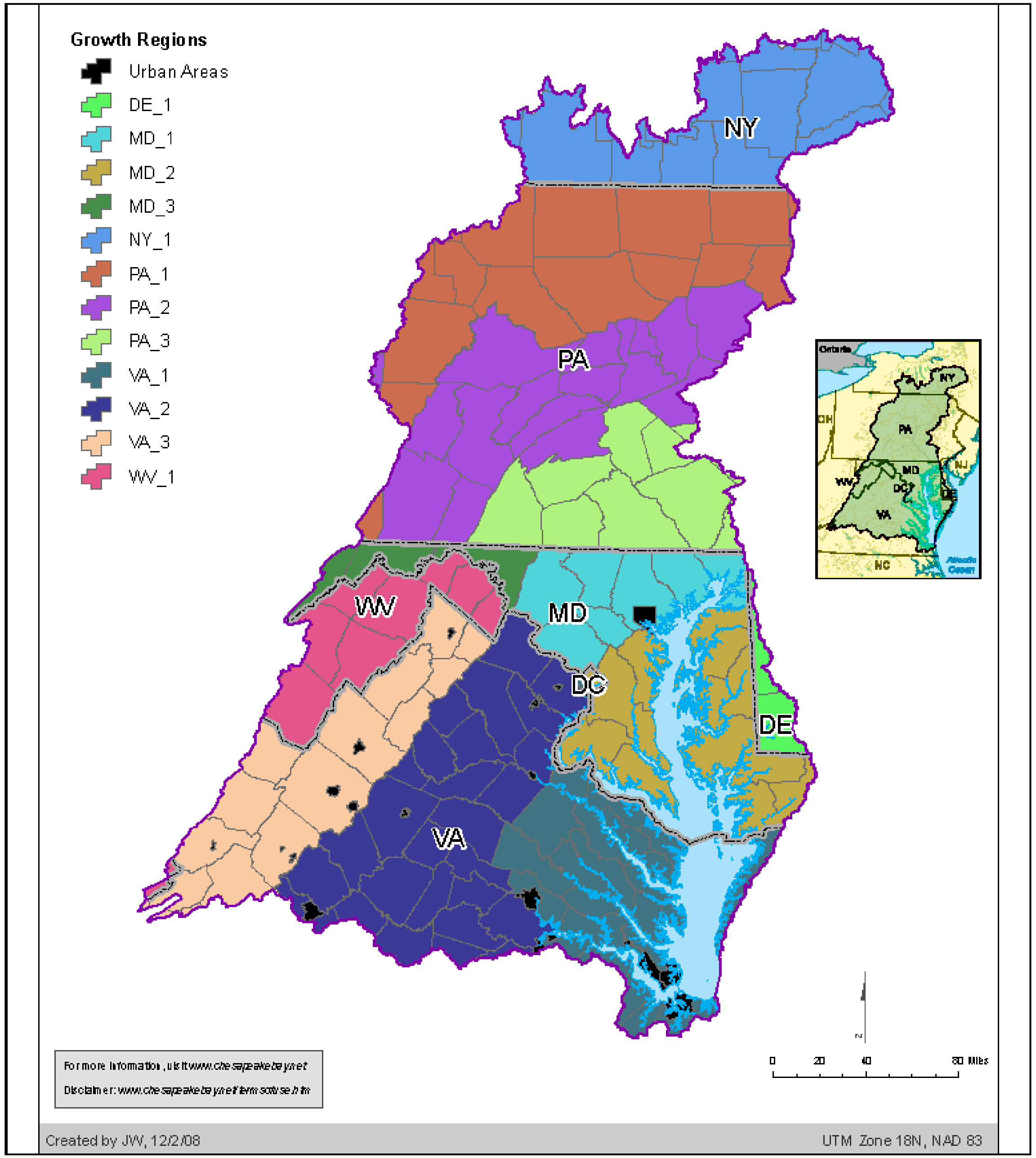
$\text{Yield}_{\text{crop } i\text{, growth region } j} \sim f(\text{time, weather})$

$\text{Yield}_{\text{crop } i\text{, growth region } j} \sim f(\text{time, weather, crop yields})$

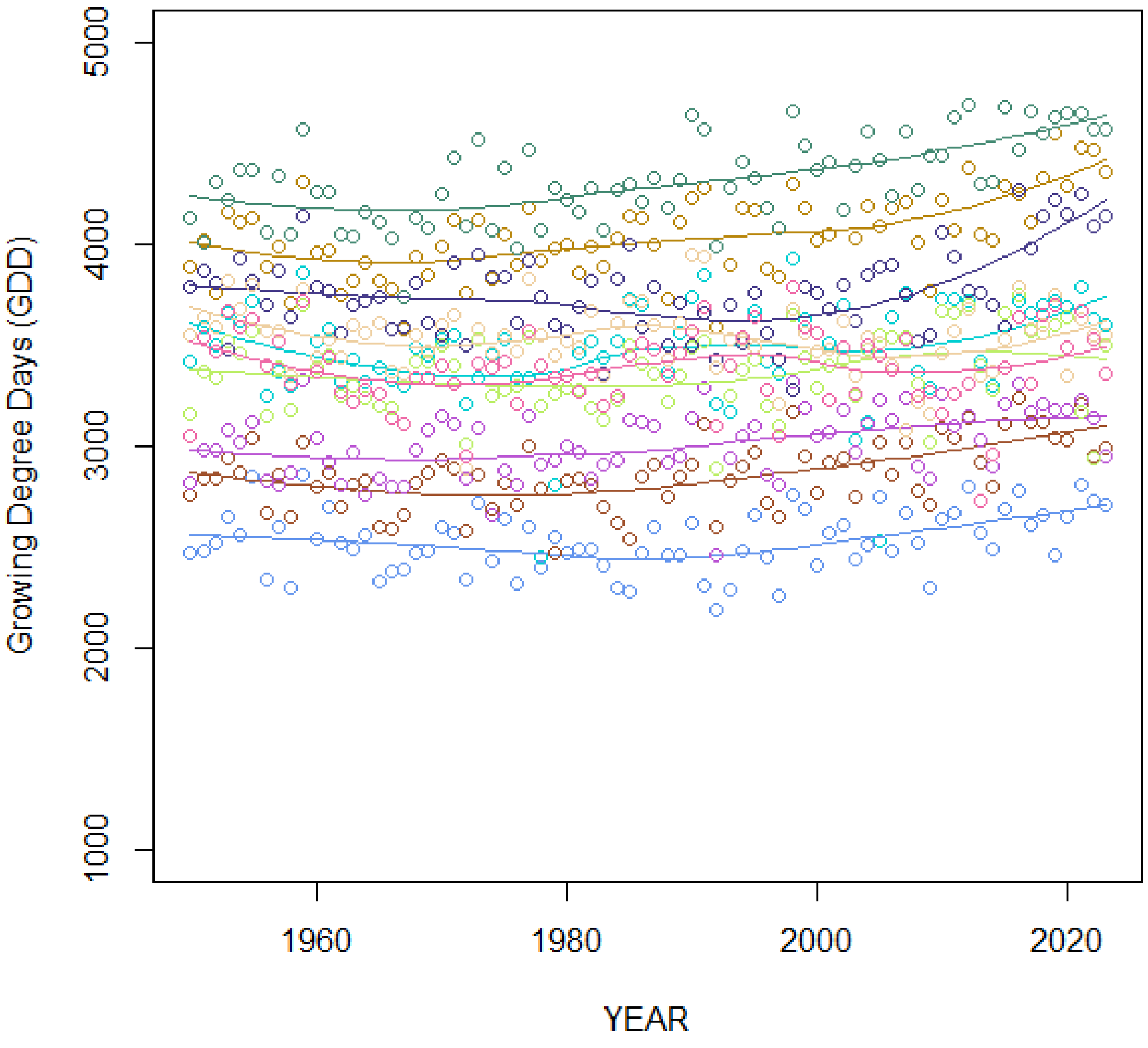
Corn-grain
Corn-silage
Barley
Alfalfa
Oats
Wheat
Soy

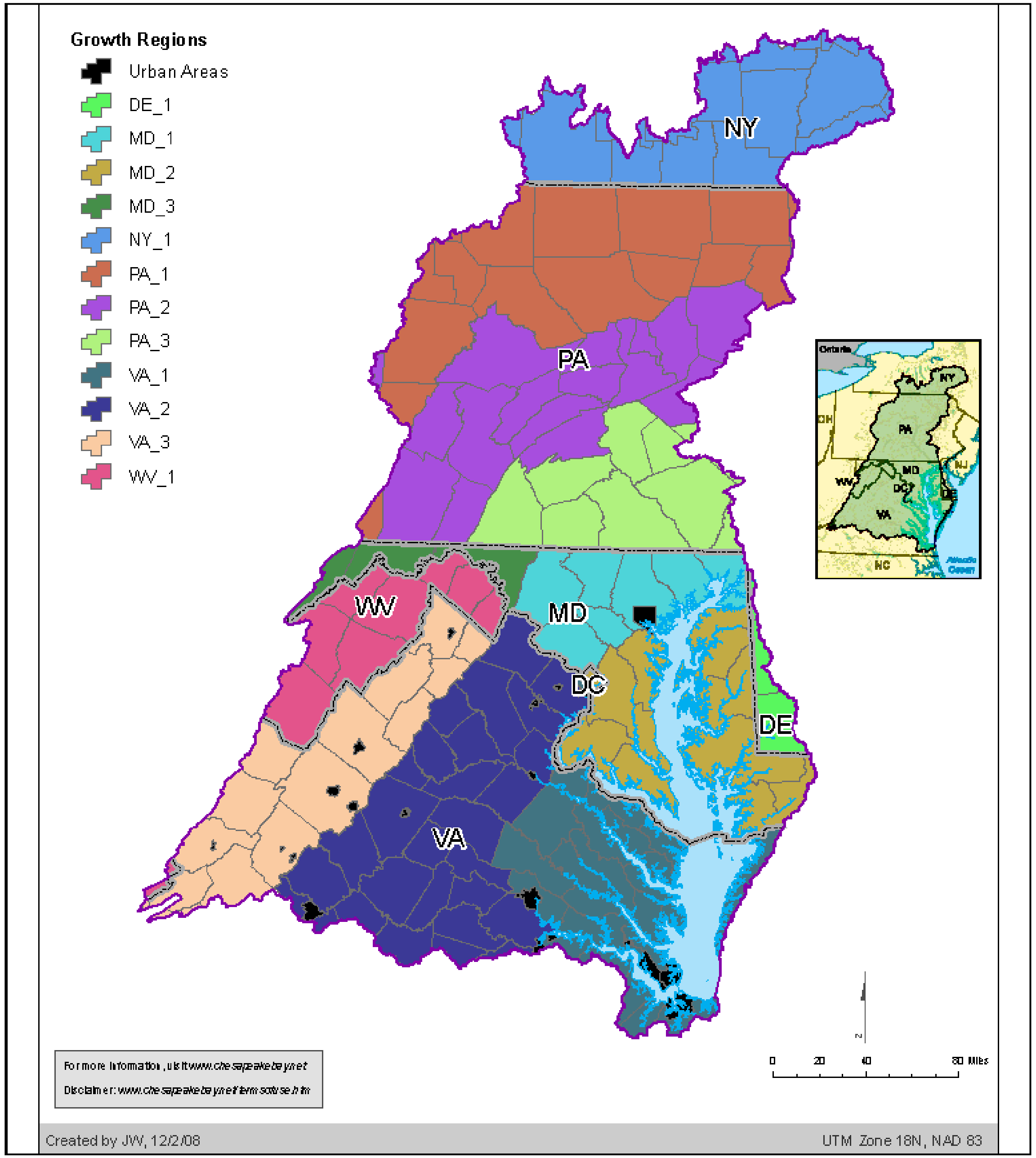
Weather data: Temperature



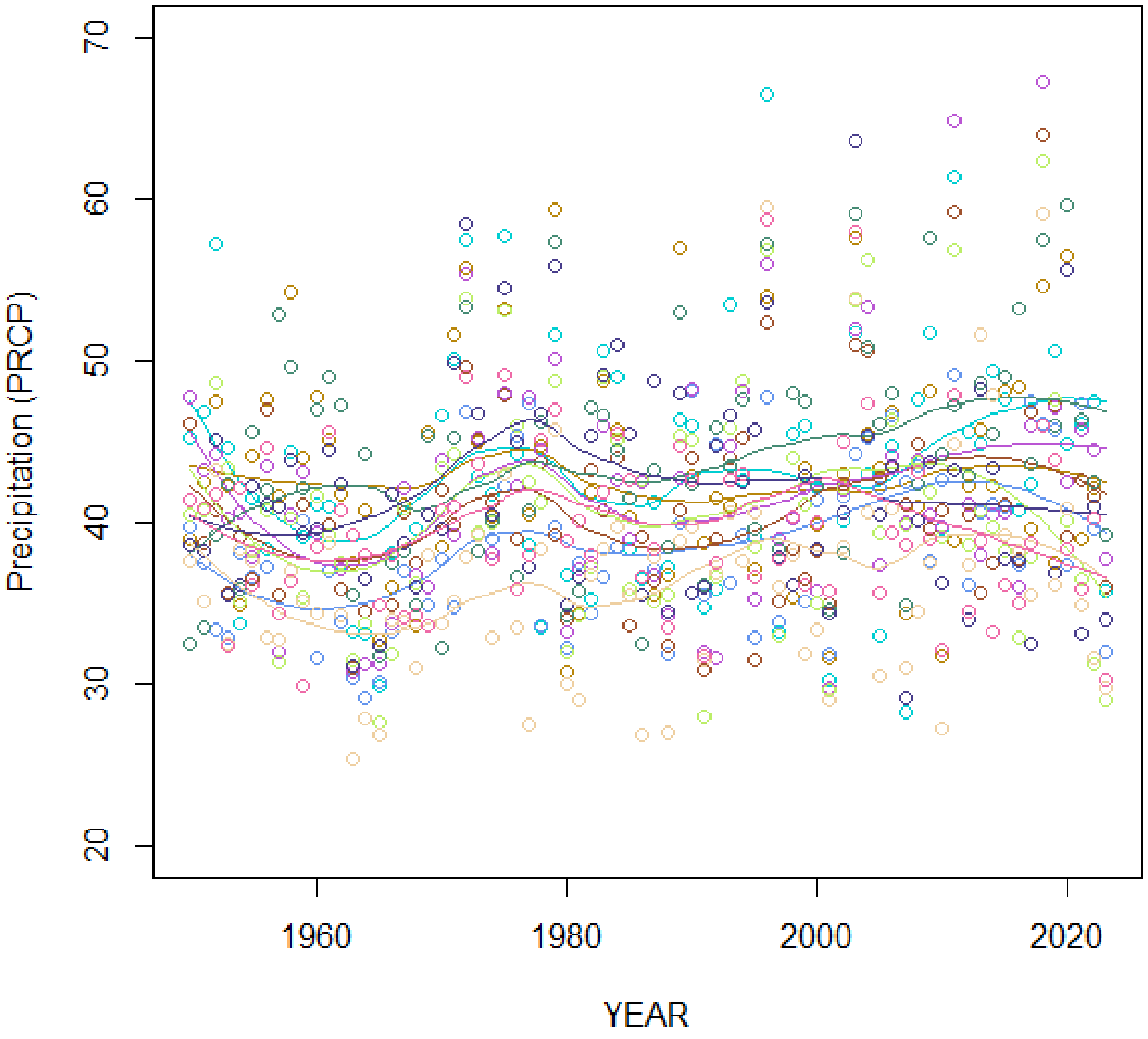


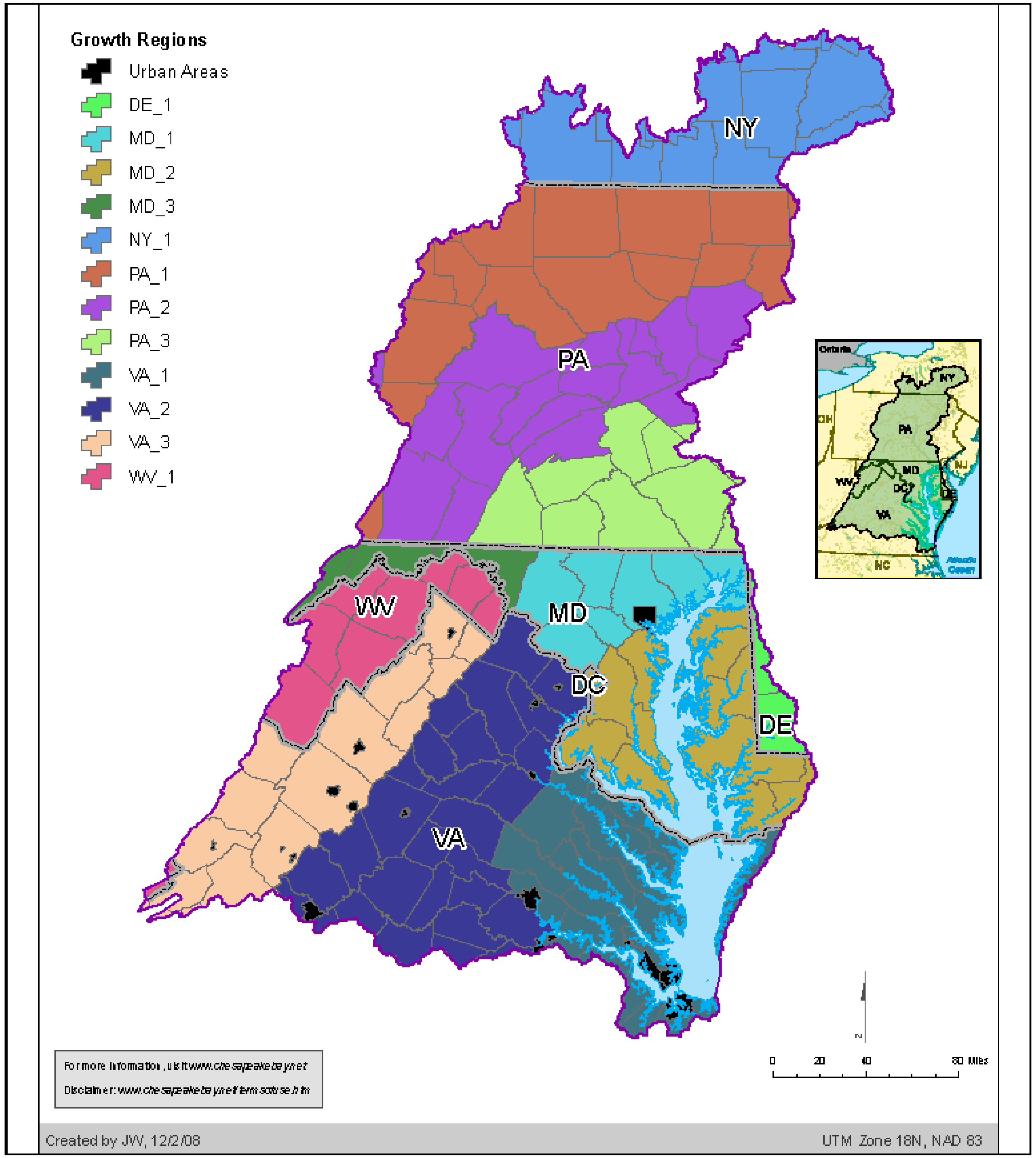
Weather data: GDD



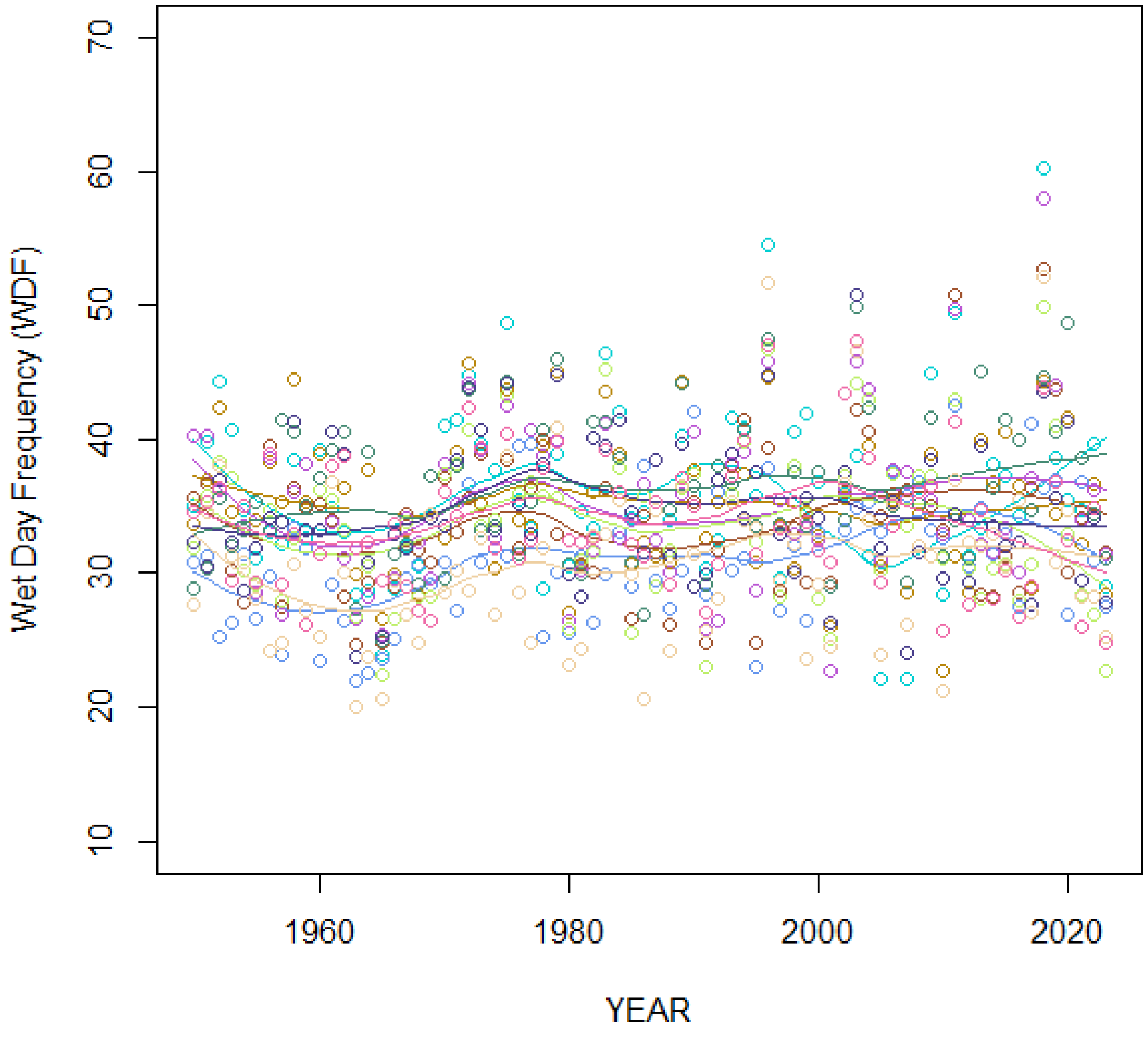


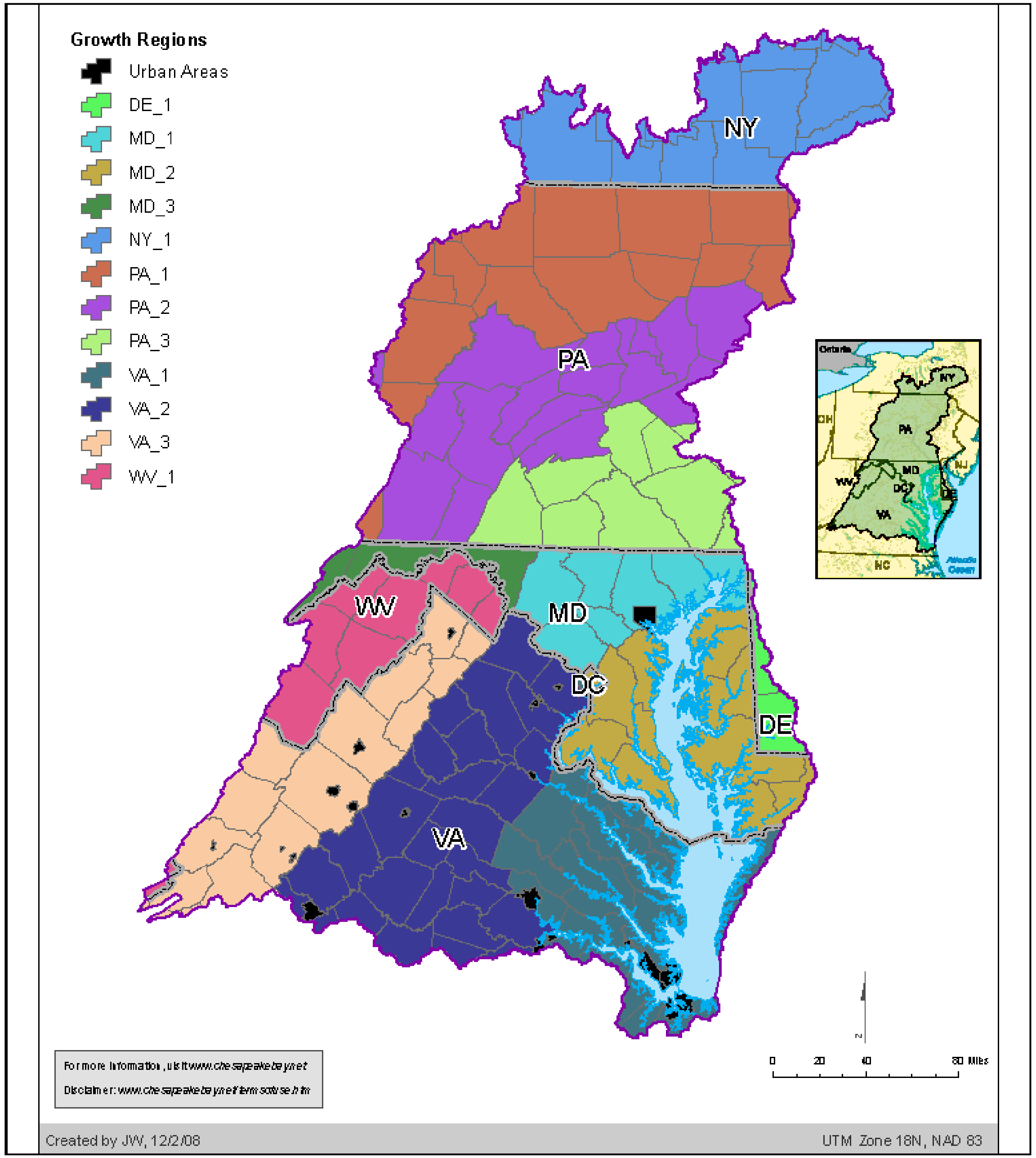
Weather data: Precip.





Weather data: Wet Day Freq.





Preliminary statistical modeling

All forms for multivariate linear models, bootstrapped (LOO).

Lots of good options, need to evaluate based on:

- consistency/stability
- complexity
- variation explained

$Yield_{crop\ i} \sim f(\text{growth regions, time, weather})$ $R^2 \sim 0.46$

- Single model per crop with greater consistency in estimate yields between growth regions

$Yield_{crop\ i} \sim f(\text{growth regions: weather, GR: time})$ $R^2 \sim 0.63$

- Mixed effects model

$Yield_{crop\ i,\ growth\ region\ j} \sim f(\text{time, weather})$ $R^2 \sim 0.71$

- Generates many discrete models (i x j), requires automated model selection
- Stepwise model selection optimizing Bayesian Information Criteria

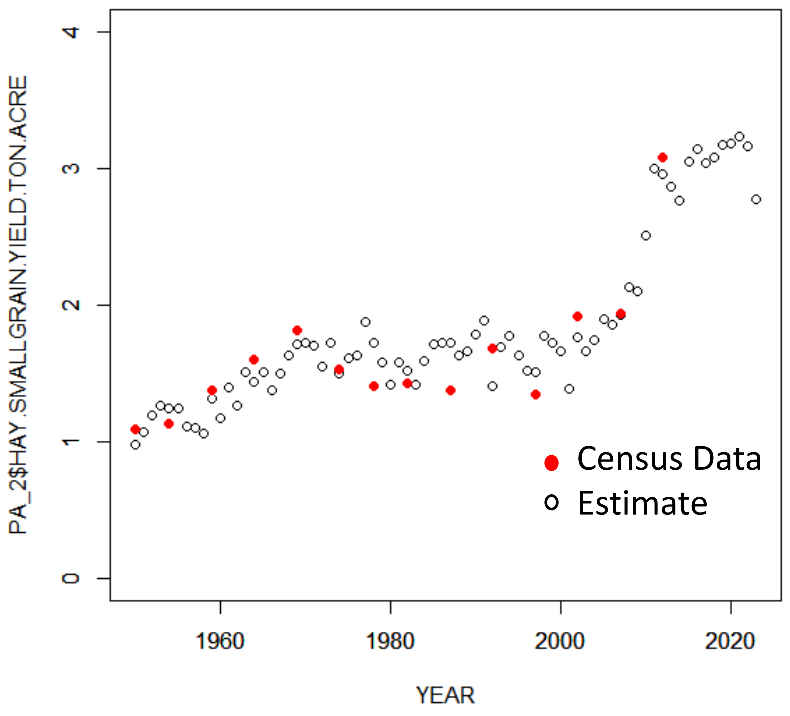
$Yield_{crop\ i,\ growth\ region\ j} \sim f(\text{time, weather, crop yields})$ $R^2 \sim 0.88$

- Greater opportunity for over parametrization
- Limited ability for projection outside crop data range

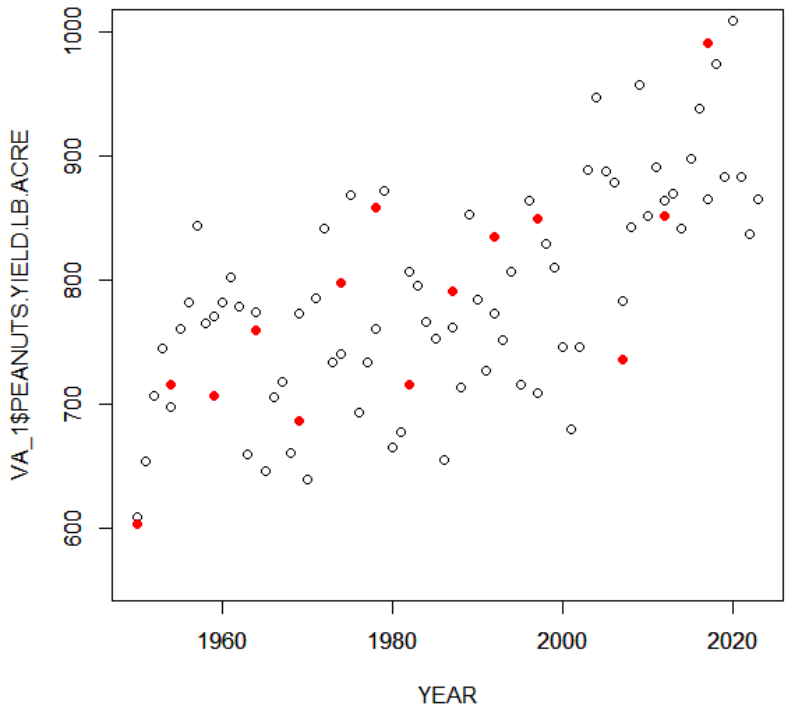
Corn-grain
Corn-silage
Barley
Alfalfa
Oats
Wheat
Soy

Estimate examples

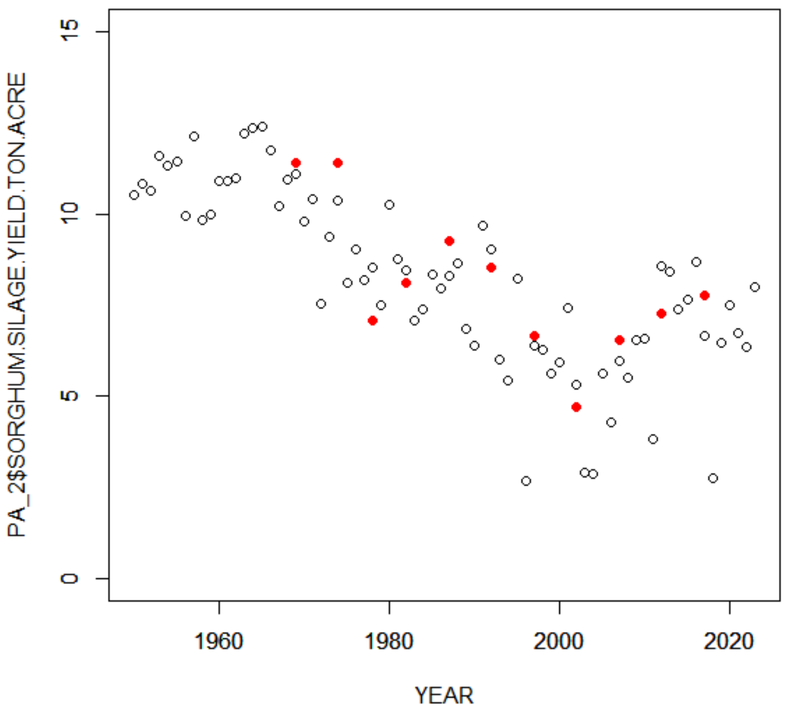
Hay.SmallGrain ~ 1.24e-02*YEAR + 8.84e-04*GDD - 2.47, R^2=0.86



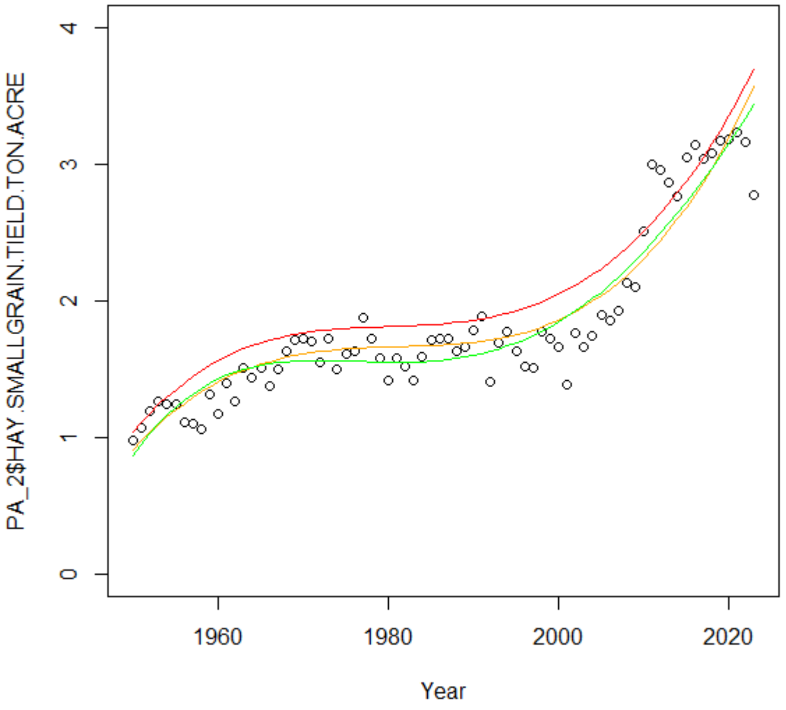
Peanuts ~ 9.23*PRCP + 14.46*AvgTMAX -530.54 , R^2=0.32

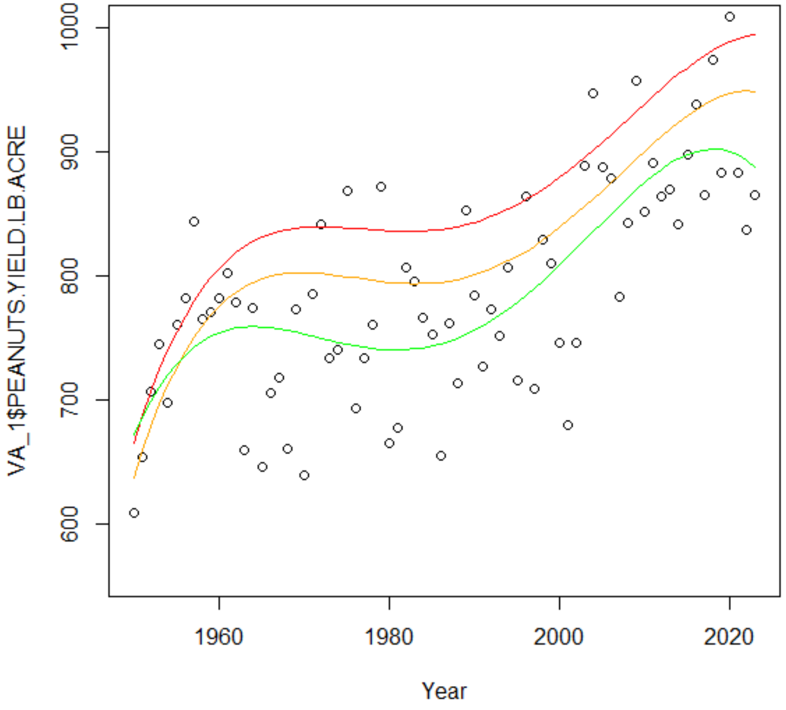


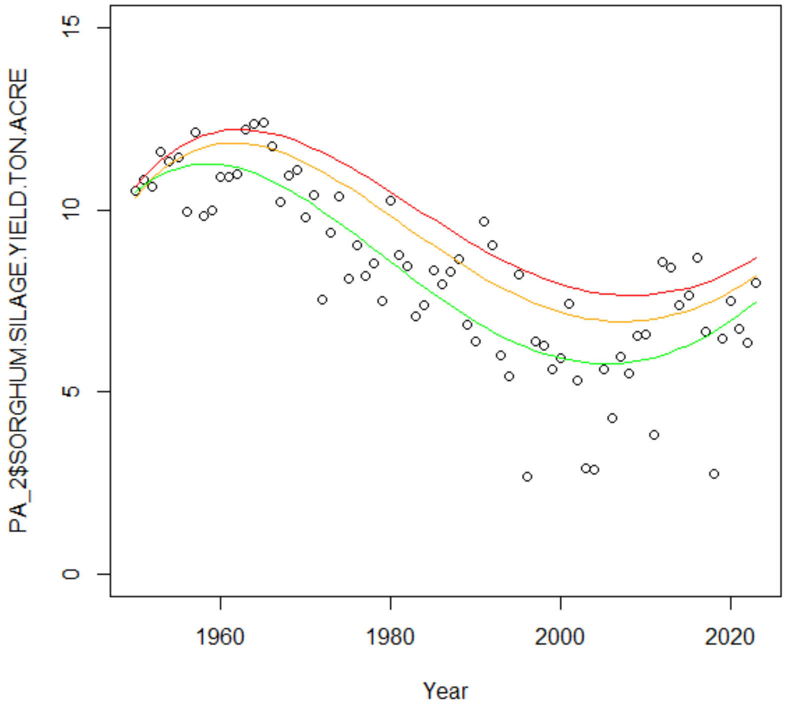
Sorghum.Silage~ -0.11*YEAR -0.20*PRCP + 0.22*AvgTMIN + 226.15, R^2=0.80



Apply trend analysis to annual estimates



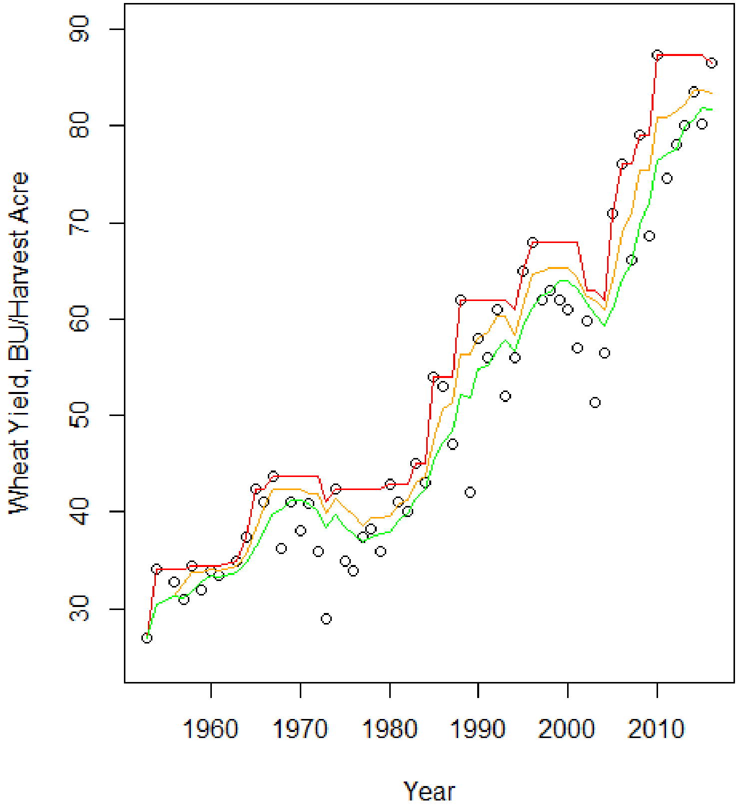




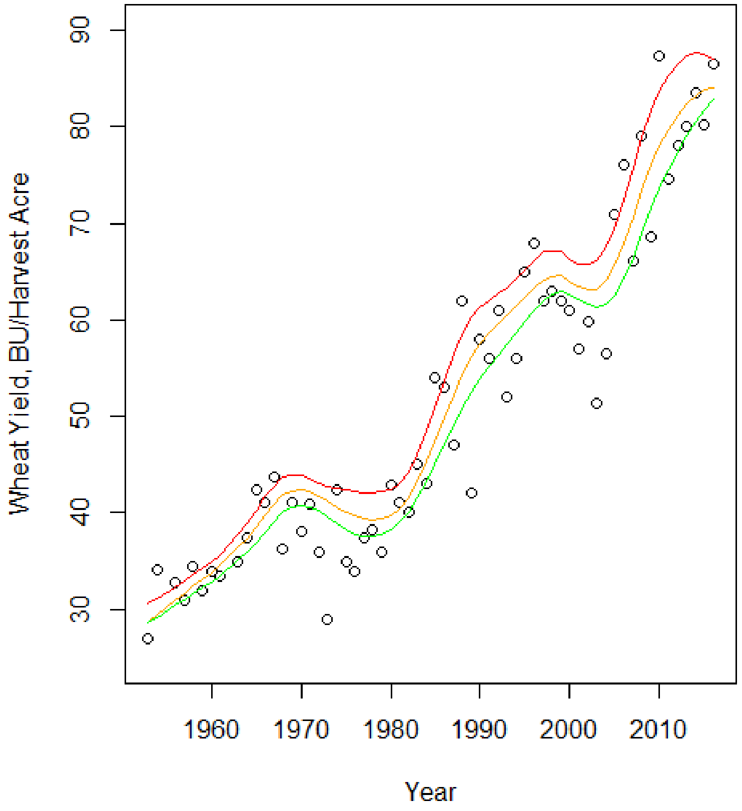
- Max yield expectation (5-year max)
- Weather independent yield expectation (Average of best 3 of past 5 years)
- 5-yr mean yield

Various trend analysis options

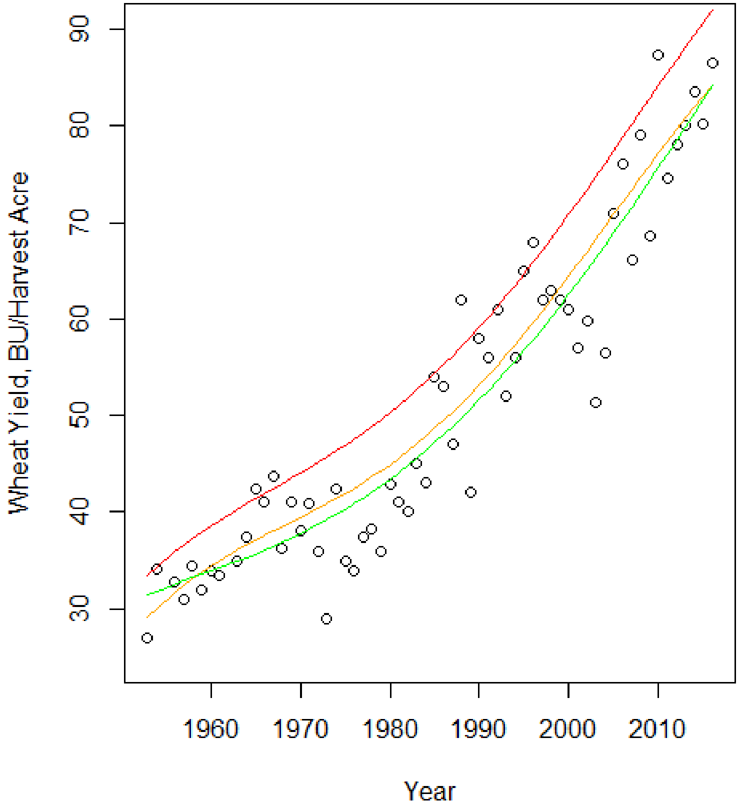
Moving window



Moving window smoothing



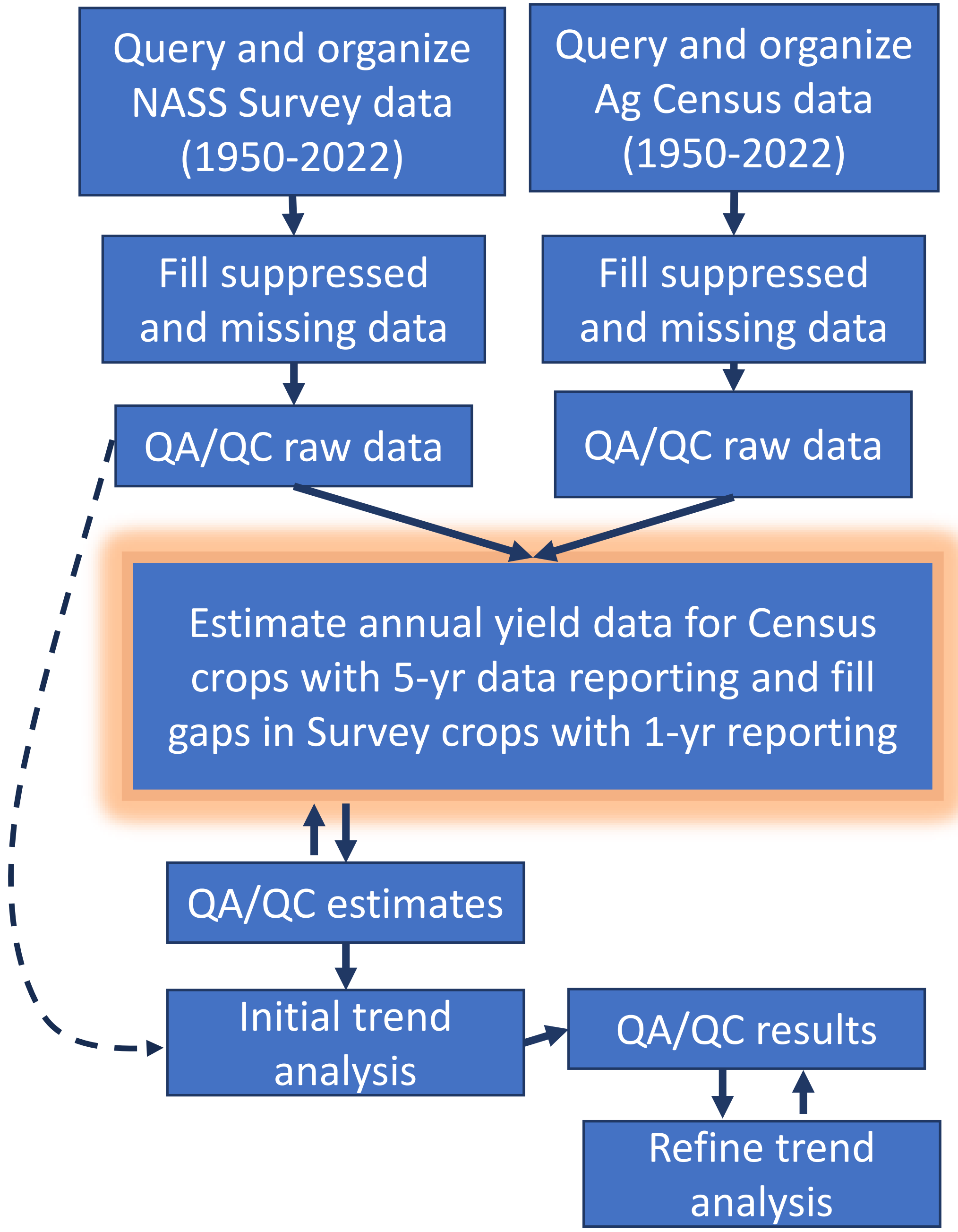
Regression, weighted residuals



- Max yield expectation (5-year max)
- Weather independent yield expectation (Average of best 3 of past 5 years)
- 5-yr mean yield

What's next

- Further evaluation of methods
  - Further refinement of methods and input data
- The last 5% of crop land N application
  - Estimation method can be applied to many crops with partial data
- QA/QC results
- Decide on final estimation method
- Decide on final trend analyses



Questions?

Example of Nutrient applications

Year	1990	
County	A	B
Crop Yield (BU)	100	200
Nutrients applied	150	300

Year	2020	
County	A	B
Crop Yield (BU)	200	250
Nutrients applied	300	375