STRATIFICATION OF AN AGRICULTURAL AREA SAMPLING FRAME USING GEOSPATIAL CULTIVATION AND CROP PLANTING FREQUENCY DATA LAYERS

Claire G. Boryan, Zhengwei Yang and Patrick Willis

USDA National Agricultural Statistics Service (NASS)

Claire.Boryan@nass.usda.gov

OUTLINE

- RATIONALE
- BACKGROUND
- OBJECTIVES
- STUDY AREA
- DATA
- METHODOLOGY
- RESULTS/DISCUSSION
- CONCLUSIONS
- FUTURE RESEARCH
- REFERENCES

RATIONALE

- The National Agricultural Statistics Service (NASS)
  - provides timely, accurate and useful statistics in service to U.S. agriculture
  - collects and disseminates data on all facets of agriculture

- NASS conducts June Agricultural Survey (JAS) and other agricultural statistics programs based on NASS Area Sampling Frame (ASF) [1]

- The ASF stratification is based on percent cultivated cropland at the primary sampling unit (PSU) level

- However, percent cultivation information does not provide crop specific predictive planting information, which is critical for sampling and crop estimation

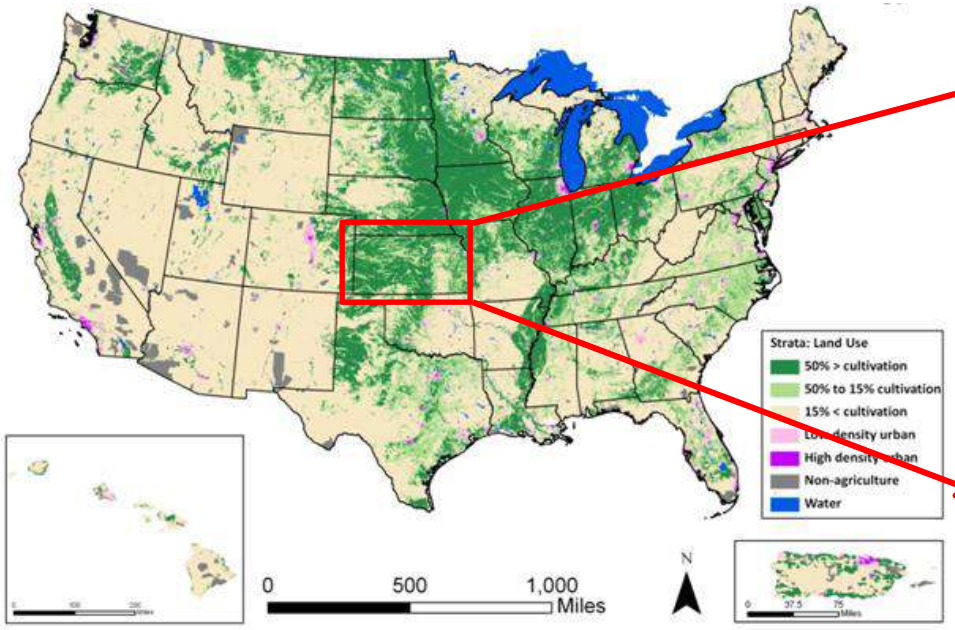
- To improve the ASF's predictability, we investigated using Crop Planting Frequency Data Layers for ASF development

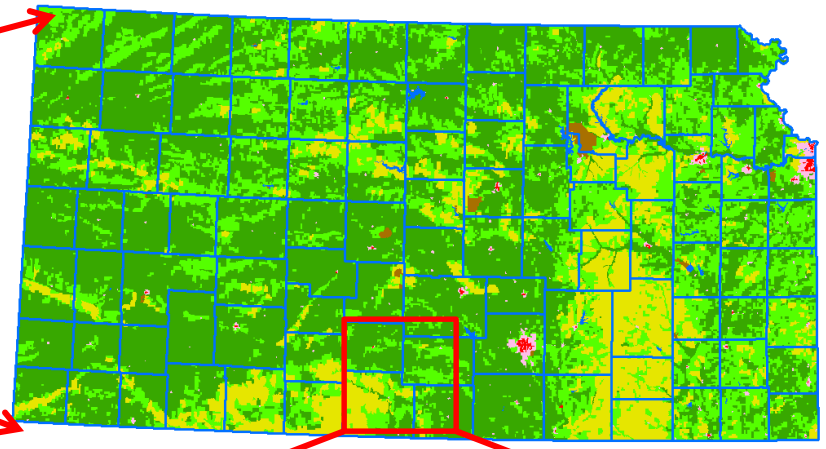
OBJECTIVES

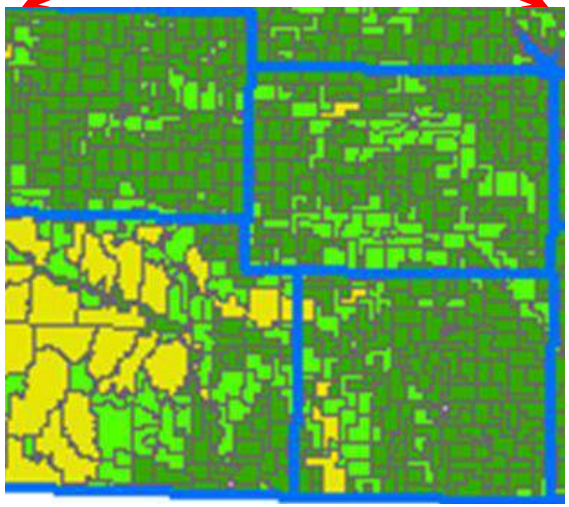
- To develop a new ASF stratification with predicative crop planting information at the ASF Primary Sampling Unit level:
  - Develop three crop specific ASF stratifications using the NASS Geospatial Crop Planting Frequency Data Layers and the NASS Cultivated Layer derived from multi-year NASS Cropland Data Layers (CDLs) and
  - Evaluate them using 2014 Farm Service Agency (FSA) Common Land Unit (CLU) Data

BACKGROUND

NASS AREA SAMPLING FRAME (ASF)







> 50% Agriculture

> 15% and < 50% Agriculture

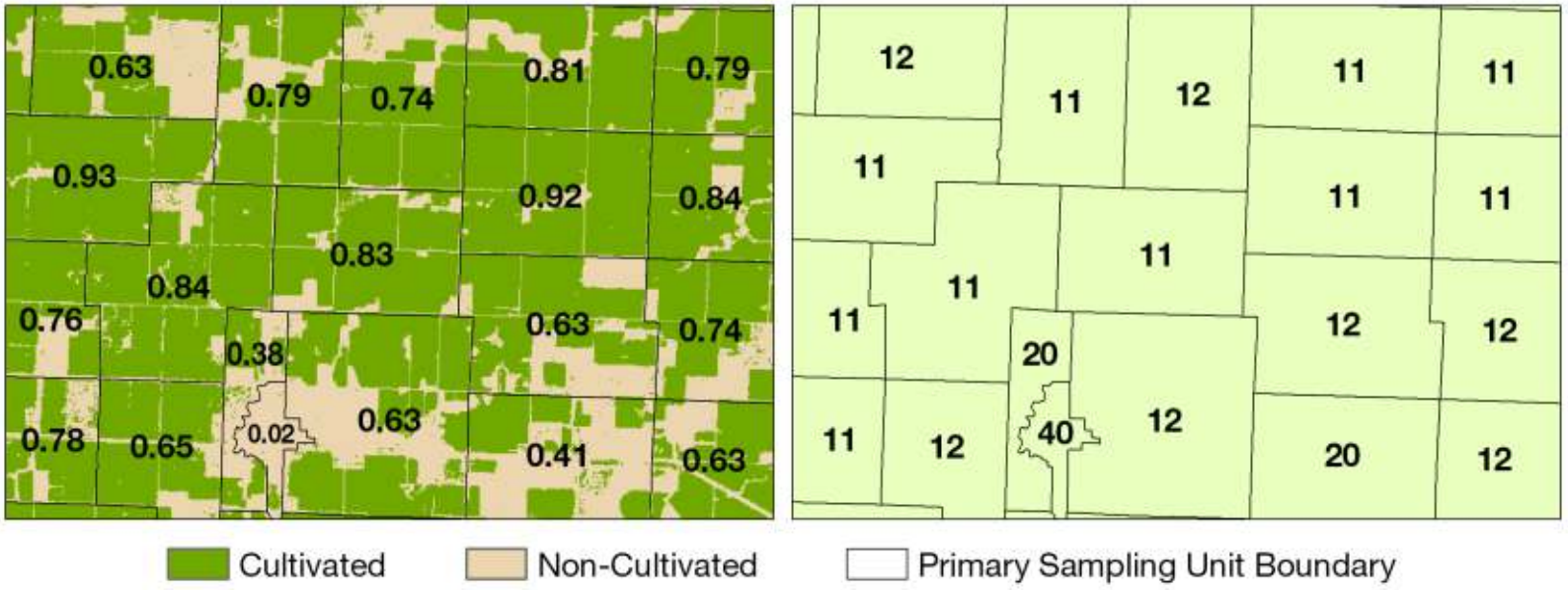
< 15% Agriculture

No Ag Potential

Agri-Urban

Commerical

BACKGROUND
ASF BASED ON NASS CULTIVATED LAYER [2-3]



ASF PSUs, with percent cultivation, overlaying the NASS cultivated layer

Each PSU labeled with a stratum category based on the state specific stratum definitions

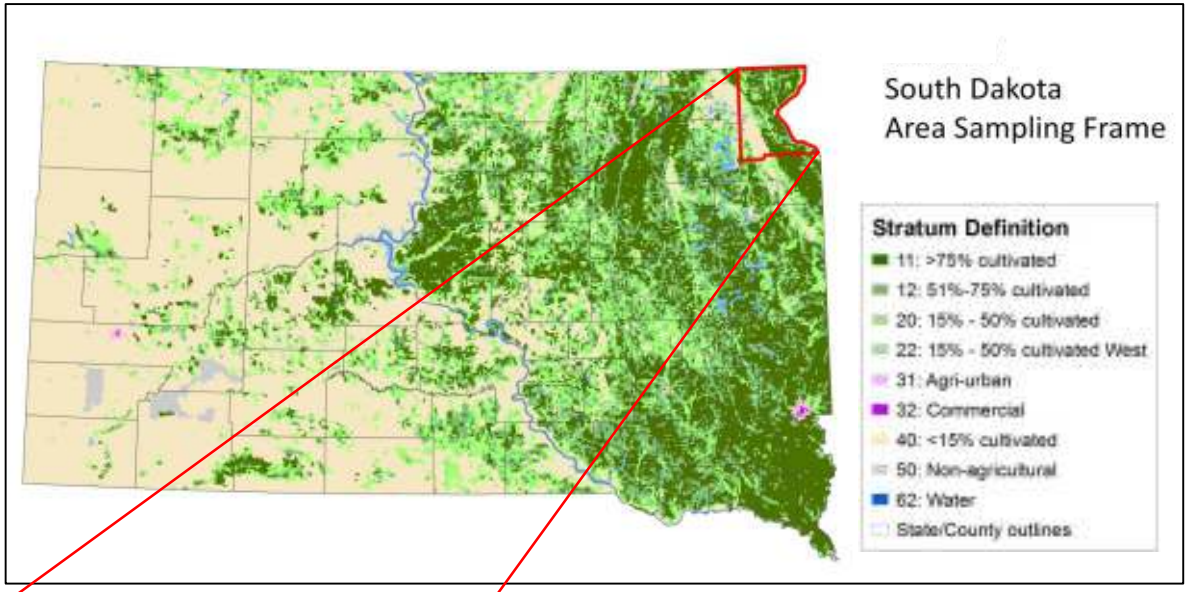
STUDY AREA

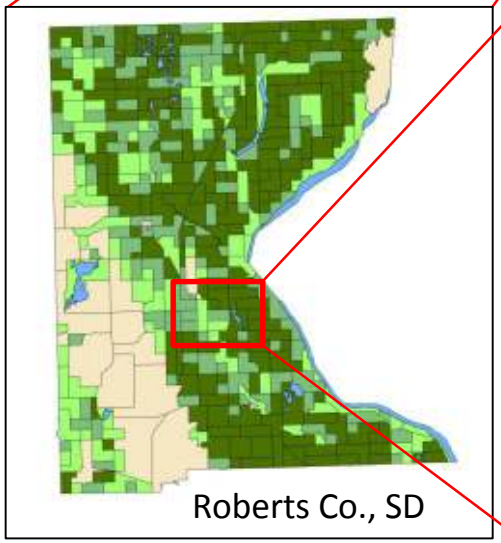
SOUTH DAKOTA (SD) - U.S.A

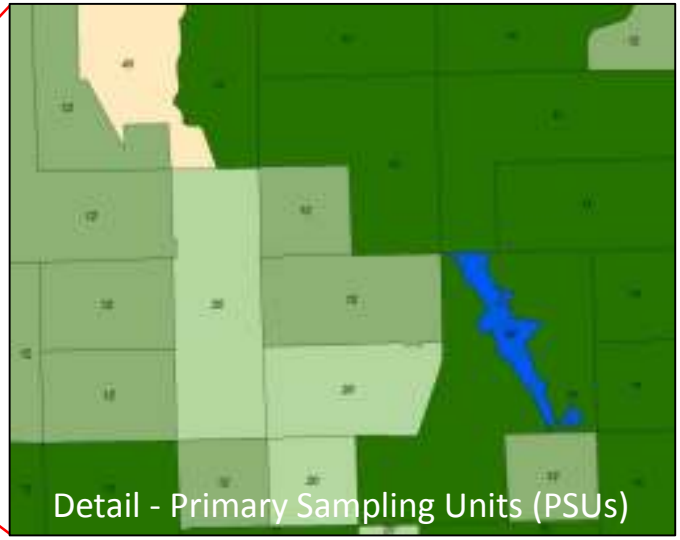
- **Large quantities of corn, soybeans, and wheat**
- **Greater than 95% coverage with Farm Service Agency Common Land Unit data for *in situ* validation and assessment of ASF performance**

DATA

NASS SOUTH DAKOTA AREA SAMPLING FRAME



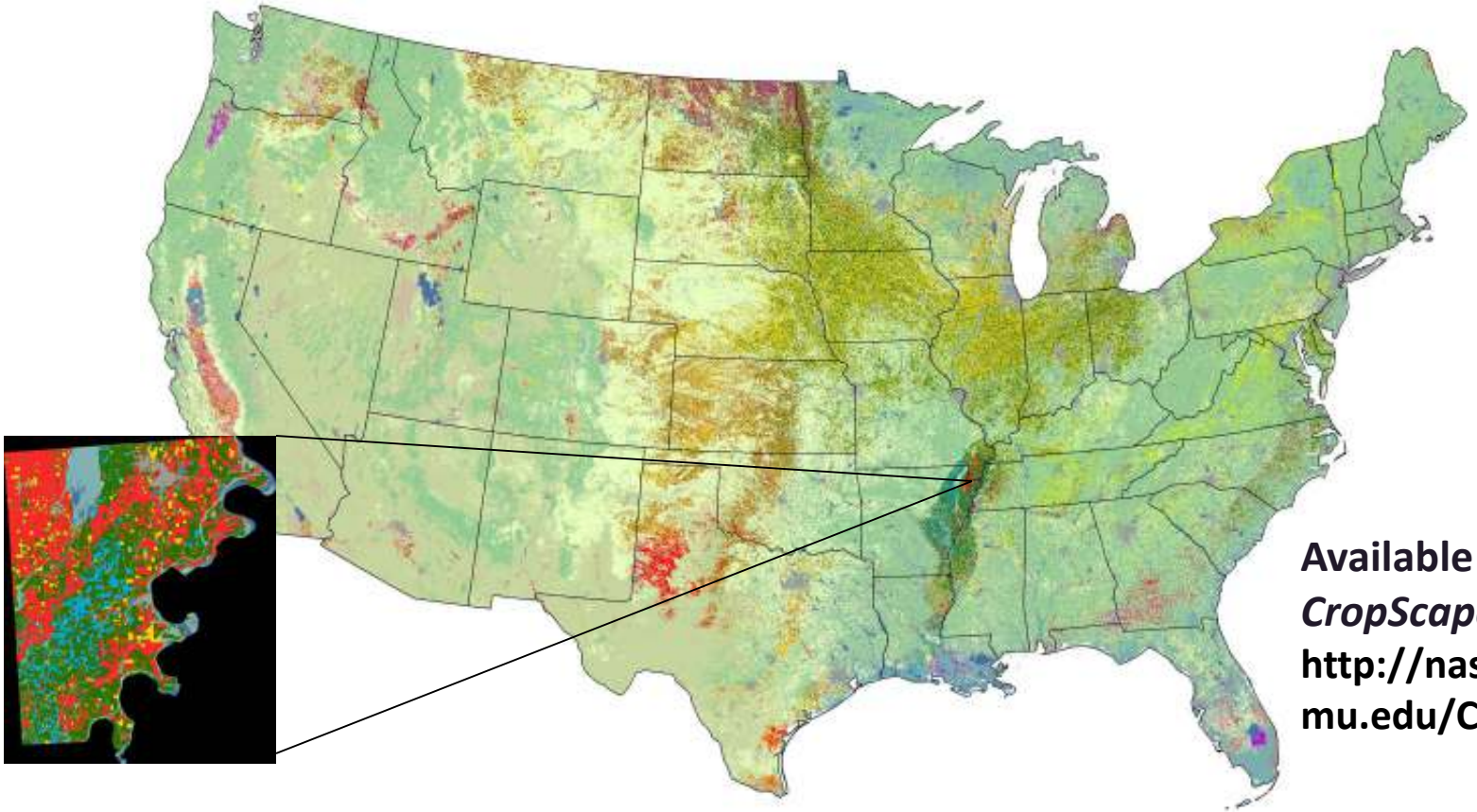




DATA

THE NASS CROPLAND DATA LAYER (CDL) [4]

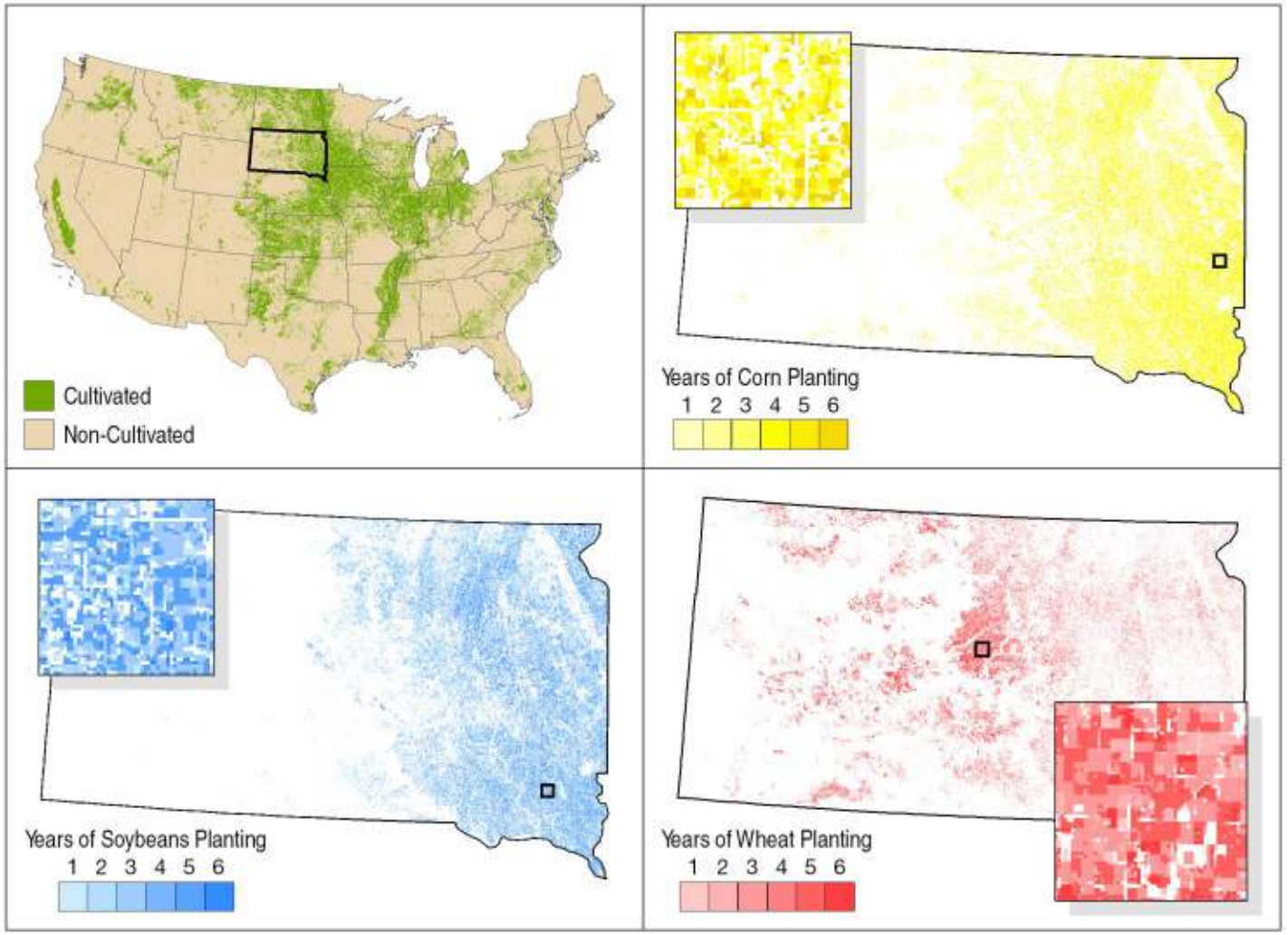
The CDL product is an annually produced, 30 m, geo-referenced, crop specific, land cover data set



Available on NASS's *CropScape*
http://nassgeodata.gmu.edu/CropScape/

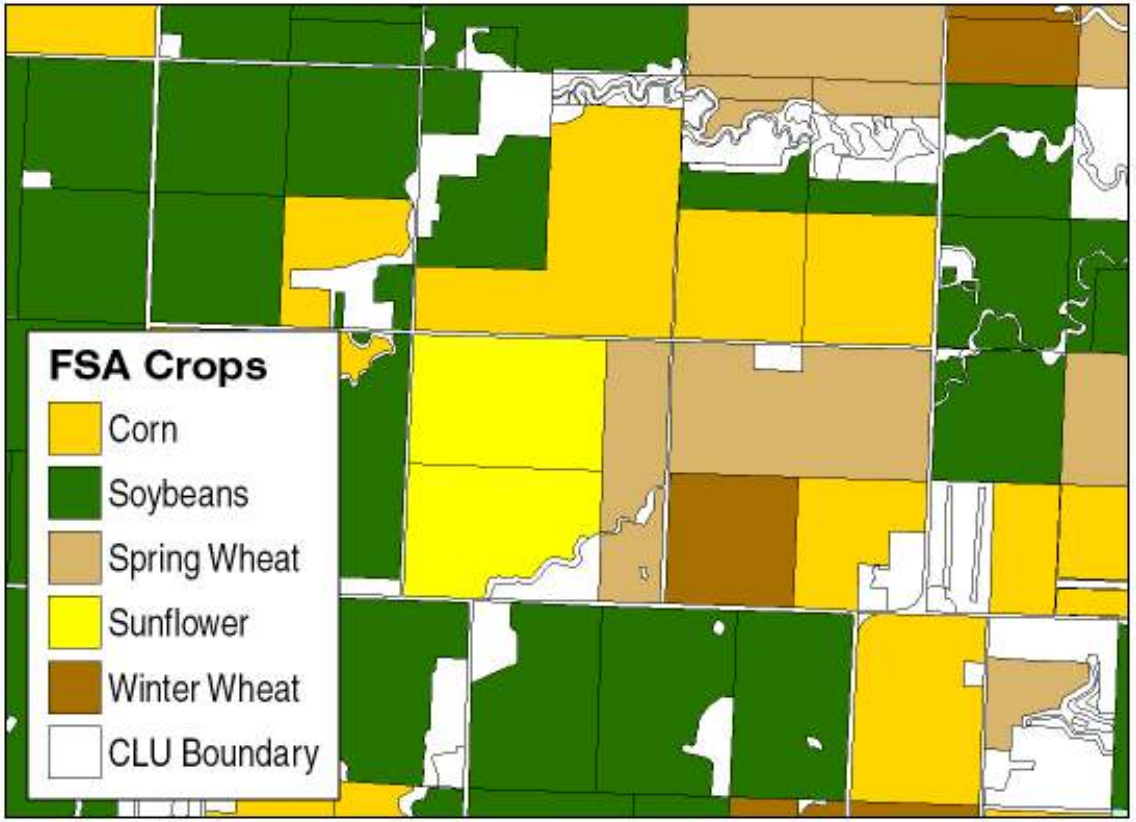
DATA

NASS CULTIVATED LAYER and CROP FREQUENCY DATA LAYERS [5] [6]



DATA

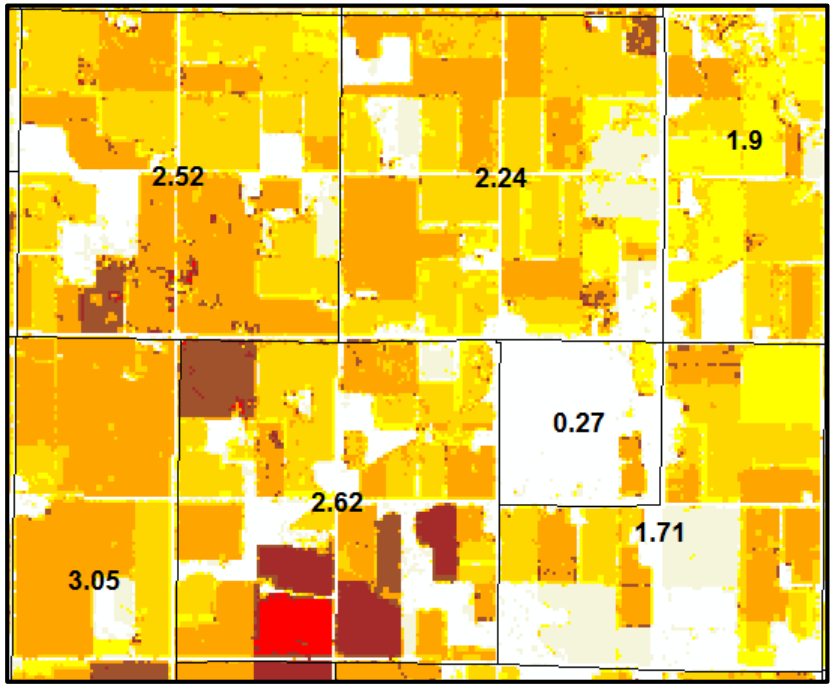
FARM SERVICE AGENCY (FSA) COMMON LAND UNIT (CLU) & 578 ADMINISTRATIVE "CROP" DATA



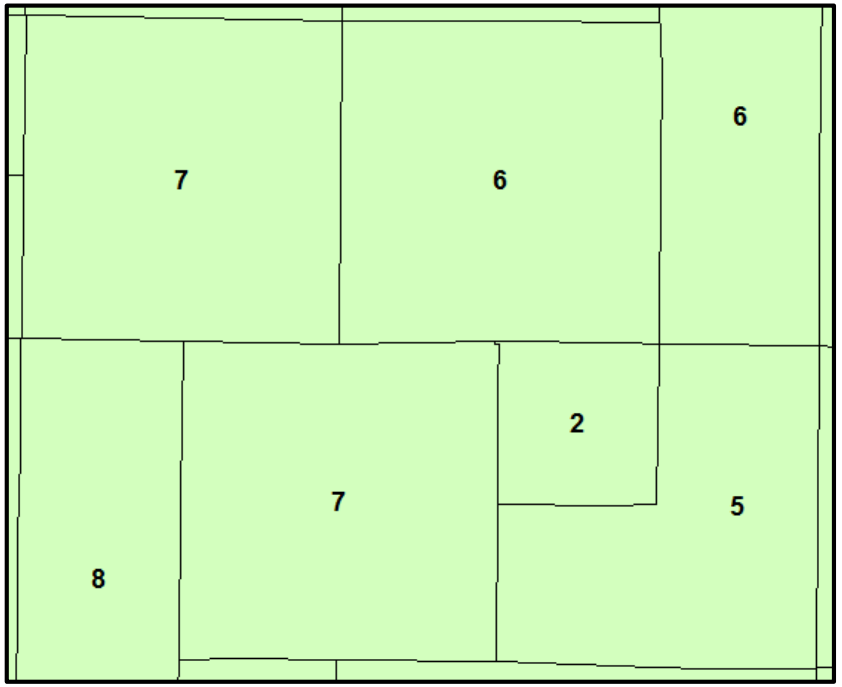
METHODOLOGY

CALCULATE MEAN CORN FREQUENCY (YEARS) AT THE PSU LEVEL

(Using the 2008-2013 Corn Frequency Layer)



PSUs with corn frequency means over the 2008-2013 SD Corn Planting Frequency Data Layer



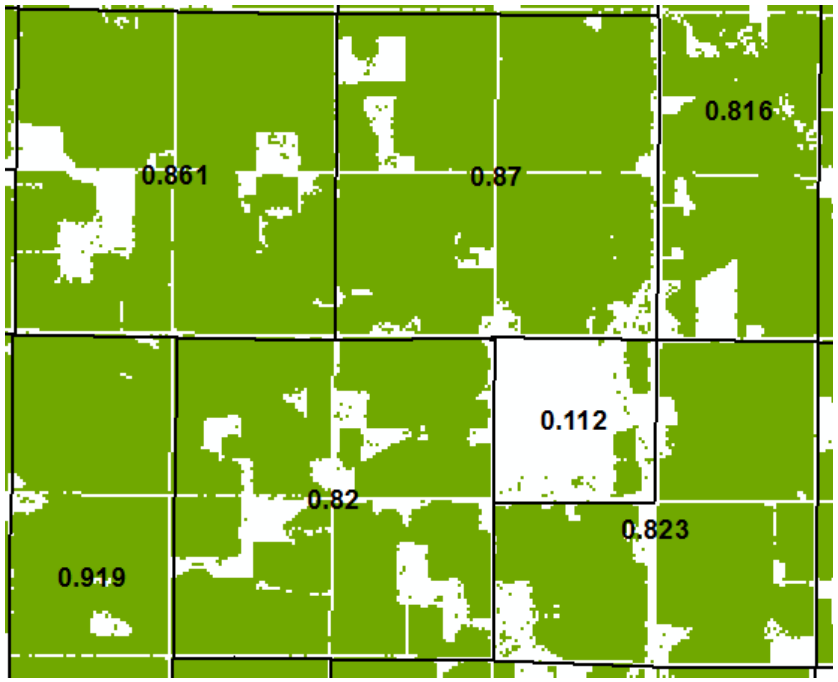
Each PSU labeled with a stratum category based on corn planting frequency

Zero Years Corn
One Year Corn
Two Years Corn
Three Years Corn
Four Years Corn
Five Years Corn
Six Years Corn

METHODOLOGY

CALCULATE PERCENT CULTIVATION AT PSU LEVEL

(Using the 2013 Cultivated Layer)



Area Frame: PSUs with percent cultivation based on the SD 2013 Cultivated Layer

Not Cultivated

Cultivated

METHODOLOGY

STRATIFICATION

K-means clustering was used on the mean corn planting frequency variable to create eight strata (same number as the current SD ASF)

STRATA	NUMBER OF PSUs	CORN PLANTING FREQUENCY RANGE (Number of years out of 6)
1	4634	0.0000 - 0.2558
2	2443	0.2560 - 0.6508
3	2497	0.6510 - 1.0550
4	2500	1.0551 - 1.4593
5	2268	1.4596 - 1.8645
6	1909	1.8651 - 2.2901
7	1544	2.2902 - 2.8491
8	356	2.8506 - 4.8648

METHODOLOGY

SUB-STRATIFICATION

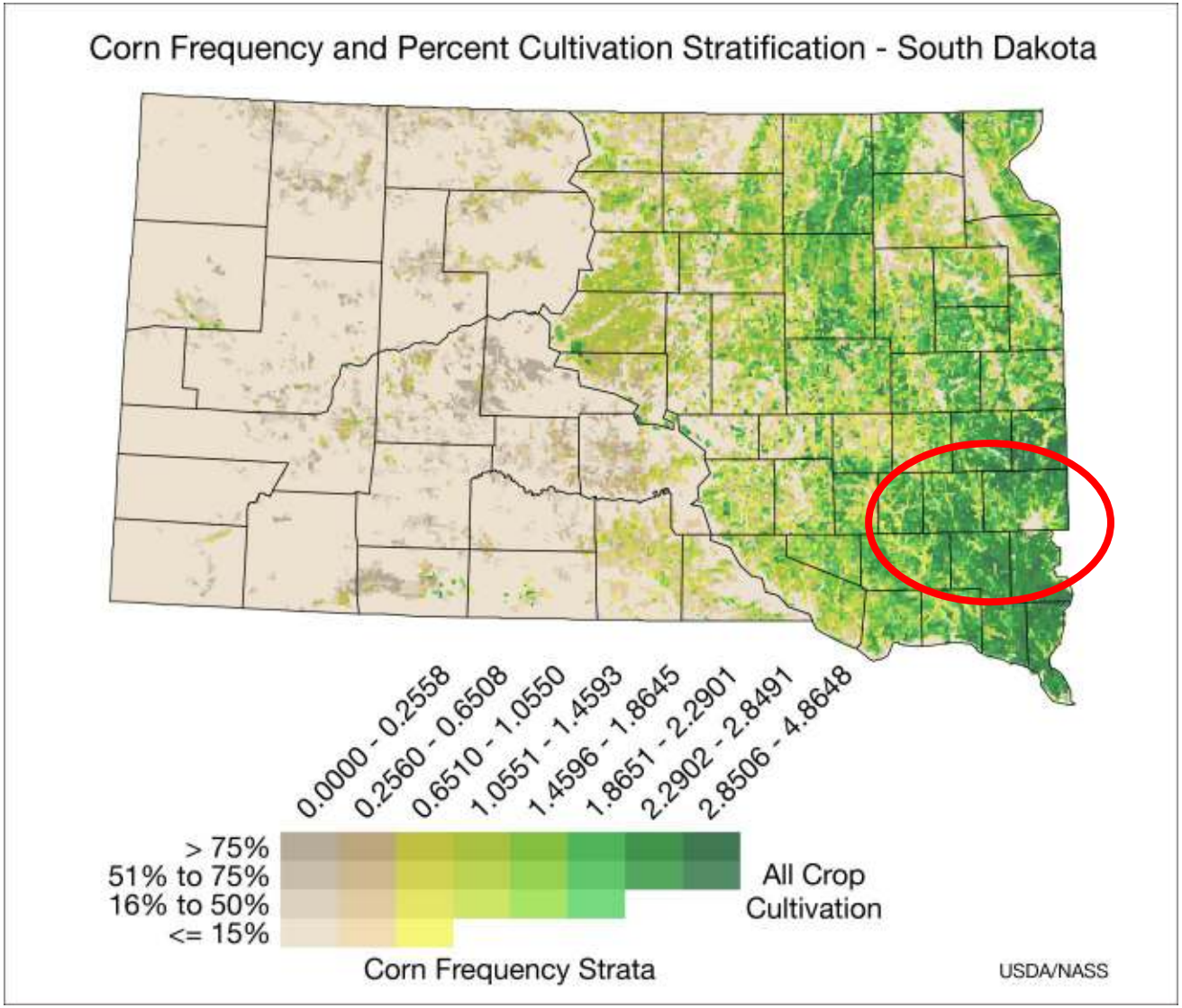
SD corn planting frequency strata are sub-stratified by percent cultivation

STRATA	CORN PLANTING FREQUENCY RANGE (Number of years out of 6)	NUMBER OF PSUs
1	0.00-0.2558	4634
2	0.256-0.6508	2443
	• • •	
8	2.8506 - 4.8648	356

STRATA	NUMBER OF PSUs	PERCENT CULTIVATION
1a	293	>75%
1b	379	>50 % - 75%
1c	1291	>15 % - 50%
1d	2671	≥ 0.0 - 15%
Total	**4634**	N/A

Strata	NUMBER OF PSUs	PERCENT CULTIVATION
2a	214	>75%
2b	376	>50 % - 75%
2c	1768	>15 % - 50%
2d	85	≥ 0.0 - 15%
Total	**2443**	N/A

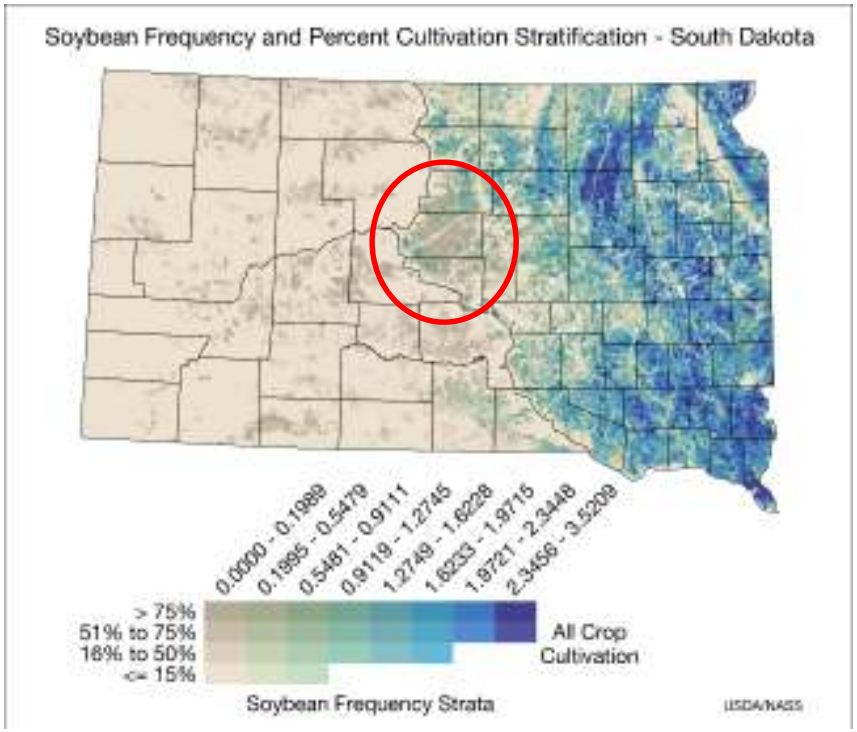
RESULTS
CORN STRATIFICATION



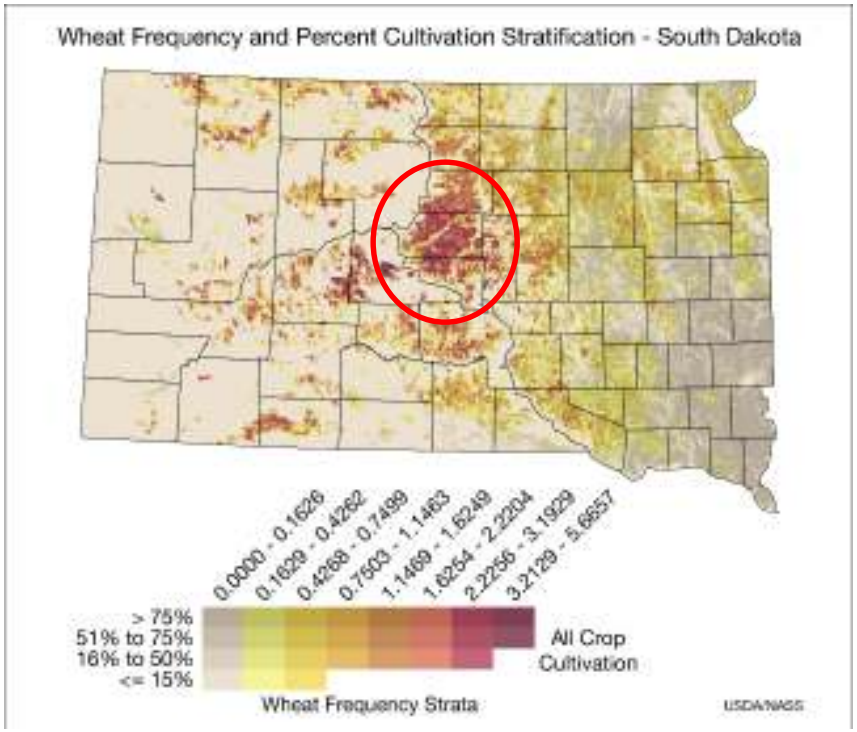
RESULTS

SOYBEAN AND WHEAT STRATIFICATIONS

SOYBEAN



WHEAT



RESULTS

2014 FSA CLU Mean Corn Acres/Percent Corn Increases from stratum 1 - 8

STRATA	NUMBER of PSUs	CORN PLANTING FREQUENCY RANGE (Number of years out of 6)	MEAN CORN ACRES	MEAN CORN PERCENT
1	4634	0.0000 - 0.2558	58.402	2.575
2	2443	0.2560 - 0.6508	203.782	12.041
3	2497	0.6510 - 1.0550	288.300	19.077
4	2500	1.0551 - 1.4593	400.515	26.975
5	2268	1.4596 - 1.8645	490.888	32.978
6	1909	1.8651 - 2.2901	638.074	39.241
7	1544	2.2902 - 2.8491	812.784	44.667
8	356	2.8506 - 4.8648	922.534	51.578
1-8	18151	0.0000 - 4.8648	N/A	N/A

RESULTS

PERCENT CULTIVATION & PSU DISTRIBUTION

- Percent cultivation substrata of the new SD ASF design has an overall accuracy of 72.365%
  - validated using the 2014 FSA CLU & 578 data
- Accuracy is consistent with the accuracy of the current SD ASF of 74% based on percent cultivation
  - validated using 2014 June Agricultural Survey data
- The new ASF's PSU distribution is consistent with the current NASS SD ASF

RESULTS

PREDICTION OF CROP SPECIFIC PLANTING PATTERNS IN 2014

- The sub-stratification of the eight primary corn, soybeans and wheat frequency strata provides additional crop specific information at the PSU level than the current NASS ASF based on percent cultivation

- The novel ASF method creates strata where high concentrations of corn, soybean and wheat are likely planted which was confirmed using the 2014 FSA CLU data as *in situ* validation data

CONCLUSIONS

- It was observed that all three ASF designs consistently predicted corn, soybean and wheat planting patterns well as verified by the 2014 FSA CLU and 578 administrative data

- The new stratification based on crop frequency and cultivation is crop type independent and applicable to all major crops

- These results indicate that the new stratification method and the three novel SD crop specific ASF designs have great potential to improve ASF accuracy, efficiency and crop estimates

FUTURE RESEARCH

- Future research will include 1) conducting estimation based on the new ASF stratifications using the 2014 FSA CLU & 578 data as proxy survey data which can be compared with the 2014 JAS estimates and 2) creating and assessing a multivariate stratification based on the NASS Crop Frequency Data Layers

REFERENCES

[1] Cotter, J., Davies, C., Nealon, J. and Roberts, R. 2010. "Area Frame Design for Agricultural Surveys," in *Agricultural Survey Methods* (eds R. Benedetti, M. Bee, G. Espa and F. Piersimoni), John Wiley & Sons, Ltd, Chichester, UK. doi: 10.1002/9780470665480.ch11

[2] Boryan, C. Yang, Z. Di, L. and Hunter, K. "A new automatic stratification method for U.S. agricultural area sampling frame construction based on the Cropland Data Layer," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, VOL. 7, NO. 11, NOVEMBER , 2014

[3] Boryan, C. and Yang Z. 2014a, "Operational implementation of a new automatic stratification method using geospatial cropland data layers in the NASS area frame section", Proceedings of IGARSS 2014, IGARSS 2014 & 35th Canadian Symposium on Remote Sensing, Quebec City, Canada, July 13-18, 2014.

[4] Boryan, C., Yang, Z., Mueller, R., and Craig, M. 2011. "Monitoring US Agriculture: The US Department of Agriculture, National Agricultural Statistics Service Cropland Data Layer Program," *Geocarto International* 26, (5): 341-358.

REFERENCES

[5] Boryan, C., Yang Z., and Di, L., 2012. "Deriving 2011 cultivated land cover data sets using USDA national agricultural statistics service historic cropland data layers," Proc. of IEEE International Geoscience and Remote Sensing Symposium, July 22-27, 2012, Munich, Germany.

[6] Boryan, C. G., Yang Z., and Willis P. 2014 -C. "US Geospatial Crop Frequency Data Layers," Proc. Of the Third International Conference on Agro-geoinformatics (Agro-geoinformatics 2014), August 11-14 2014, Beijing, China, DOI:10.1109/Agro-Geoinformatics.2014.6910657.

Thank you

Questions?

Claire G. Boryan, Zhengwei Yang and Patrick Willis
USDA/NASS/RDD
claire.boryan@nass.usda.gov