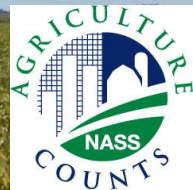# A NASS Cropland Data Layer based stratification tool
for use in
NASS Area Sampling Frame construction

Claire G. Boryan, Geographer
USDA\NASS\Research and Development Division
Fairfax, VA

NASS Overview

Provider of timely, accurate, and useful statistics in service to U.S. agriculture

What is the Cropland Data Layer (CDL)?

The Cropland Data Layer product is a raster-formatted, geo-referenced, crop specific, land cover product.
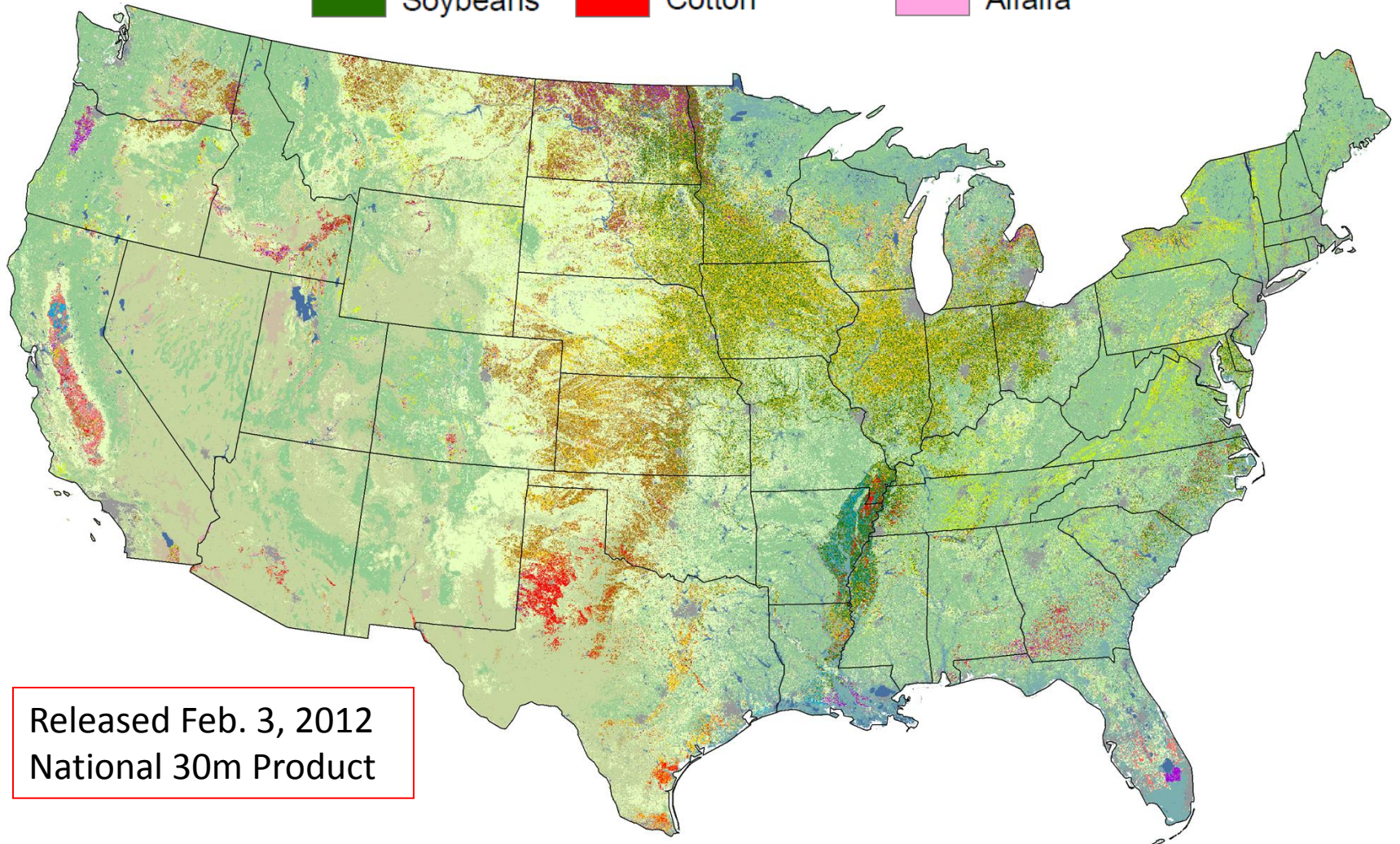
Released Feb. 3, 2012
National 30m Product

Total crop mapping accuracies for historic CDLs range from 85% to 95% for the major crops
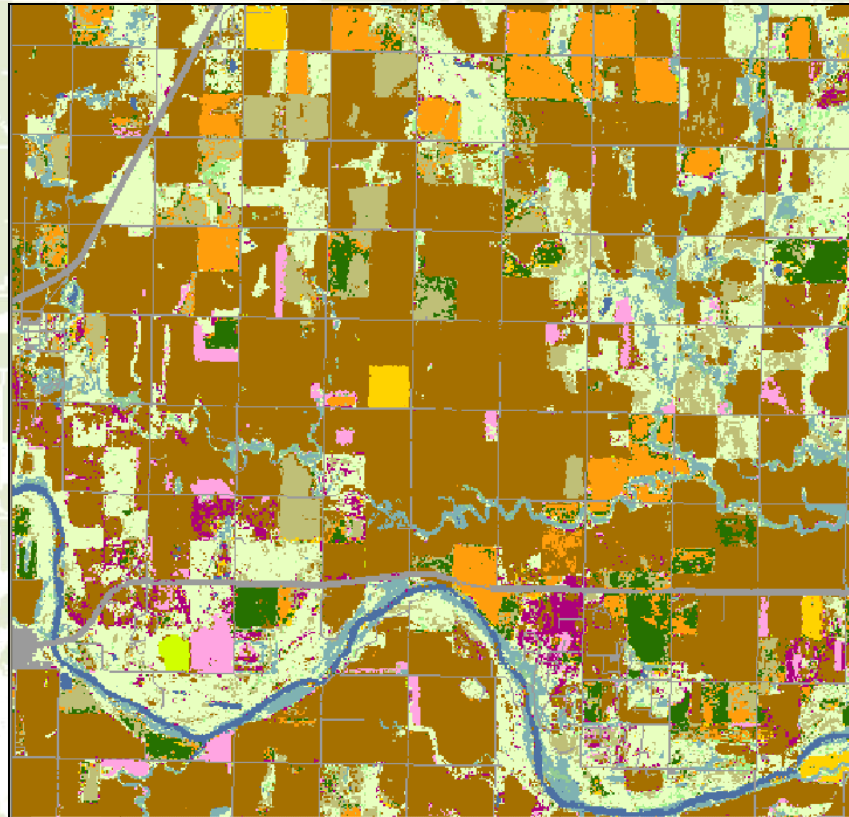
Corn		Winter Wheat		Rice	
Soybeans		Cotton		Alfalfa	

CropScape Portal

nassgeodata.gmu.edu/CropScape

NASS has used area frames for agricultural surveys since 1954

Stratification of Illinois 2006
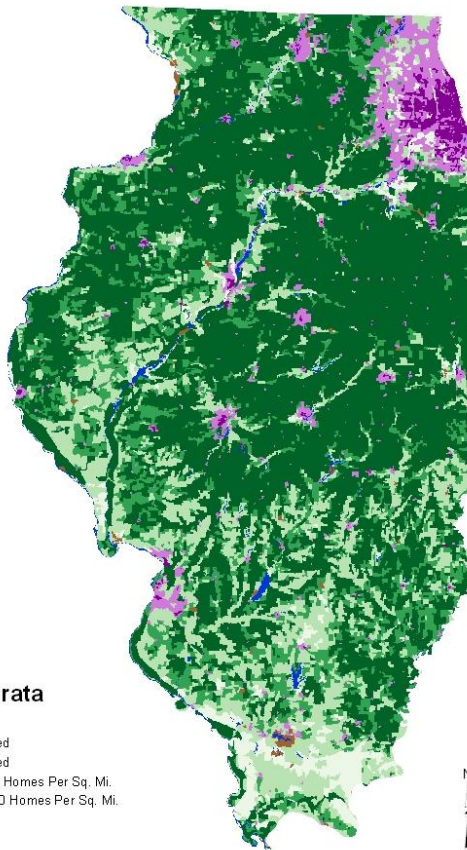
Land Use Strata
- >75% Cultivated
- 51 - 75% Cultivated
- 25 - 50% Cultivated
- Agri-Urban: > 100 Homes Per Sq. Mi.
- Commercial: > 100 Homes Per Sq. Mi.
- < 25% Cultivated
- Non-Agriculture
- Water

Projection: NAD 1983, UTM Zone 16

Prepared by Area Frame Section, 2006

USDA
NASS

Stratification of Arizona 1984
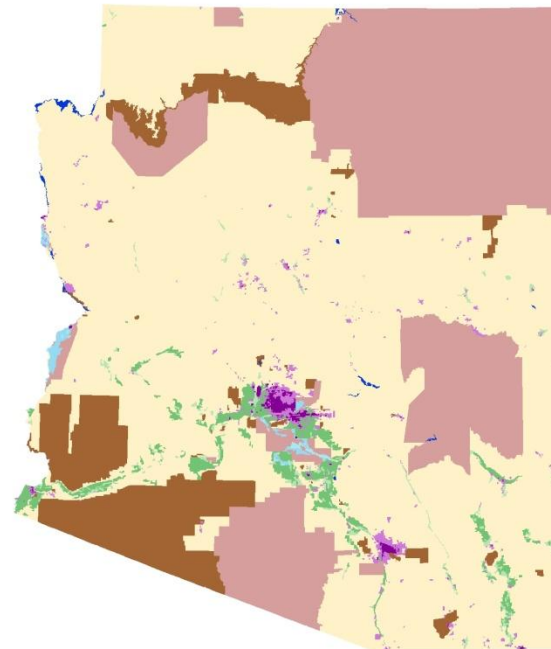
Land Use Strata
- >50% Cultivated
- >50% Cultivated--Native American
- Agri-Urban: >20 Homes/Sq Mi
- 15-50% Cultivated
- 15-50% Cultivated--Native American
- Commercial: >20 Homes/Sq Mi
- Native American
- Public/Private Lands
- Non-Agricultural
- Water

Projection: Geographic Latitude and Longitude

Prepared by Area Frame Section, 2006

USDA
NASS

NASS Area Sampling Frames

Area Sampling Frames (ASF) have been used as the primary mechanism for conducting agricultural surveys since 1954 and are considered the backbone to the agricultural statistics program of the NASS.

The NASS ASFs are based on a stratification of land cover in the U.S. by percent cultivated cropland, and are the statistical foundation for providing estimates with complete coverage of U.S. agriculture.

General land-use stratification codes and definitions represented in the NASS Area Sampling Frames

Land-Use Strata Codes and Definitions Stratum	Definition
11	General Cropland, greater than 75% cultivated.
12	General Cropland, 51-75% cultivated
20	General Cropland, 15-50% cultivated.
31	Ag-Urban, less than 15% cultivated, more than 100 dwellings per square mile, residential mixed with agriculture.
32	Residential/Commercial, no cultivation, more than 100 dwellings per square mile.
40	Less than 15% cultivated
50	Non-agricultural.
62	Water

US map illustrating the implementation years of current NASS Area Sampling Frames
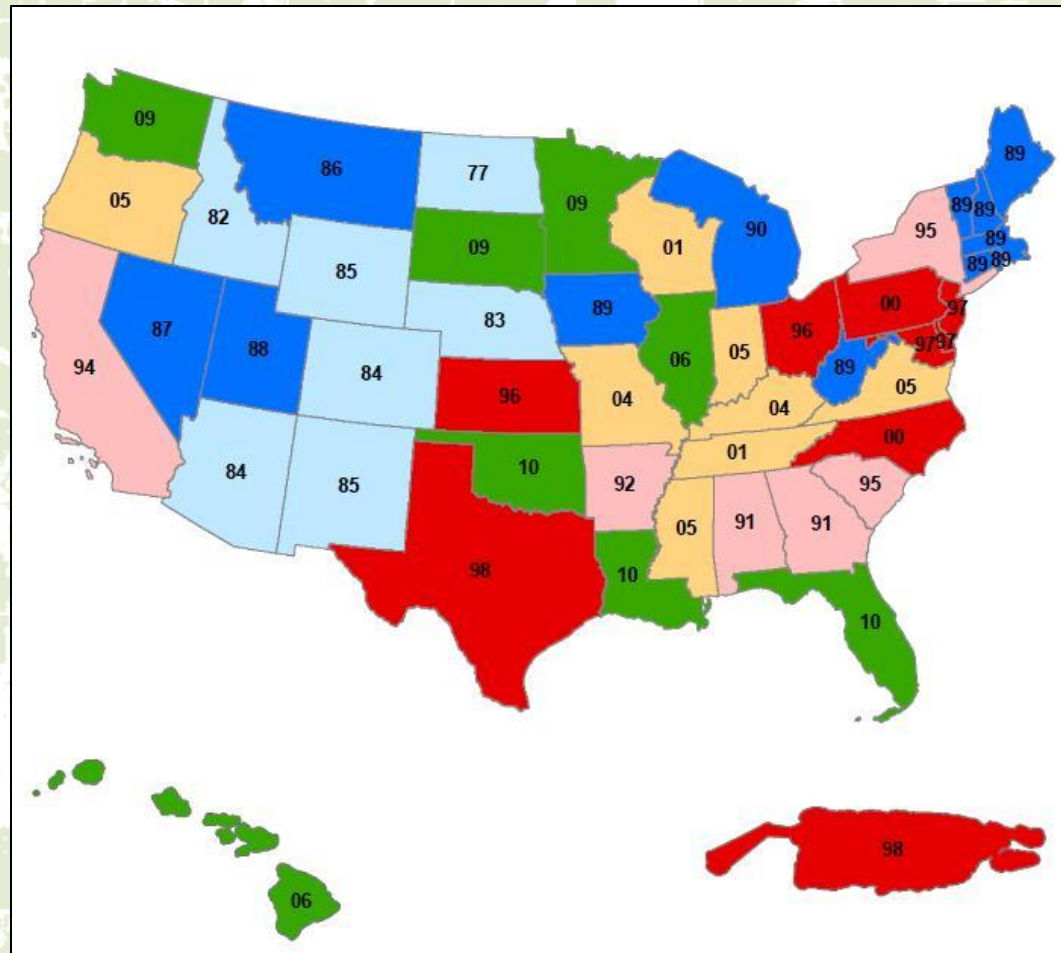
A new automated stratification method has been developed to utilize the NASS Cropland Data Layer in the construction of the NASS Area Sampling Frame
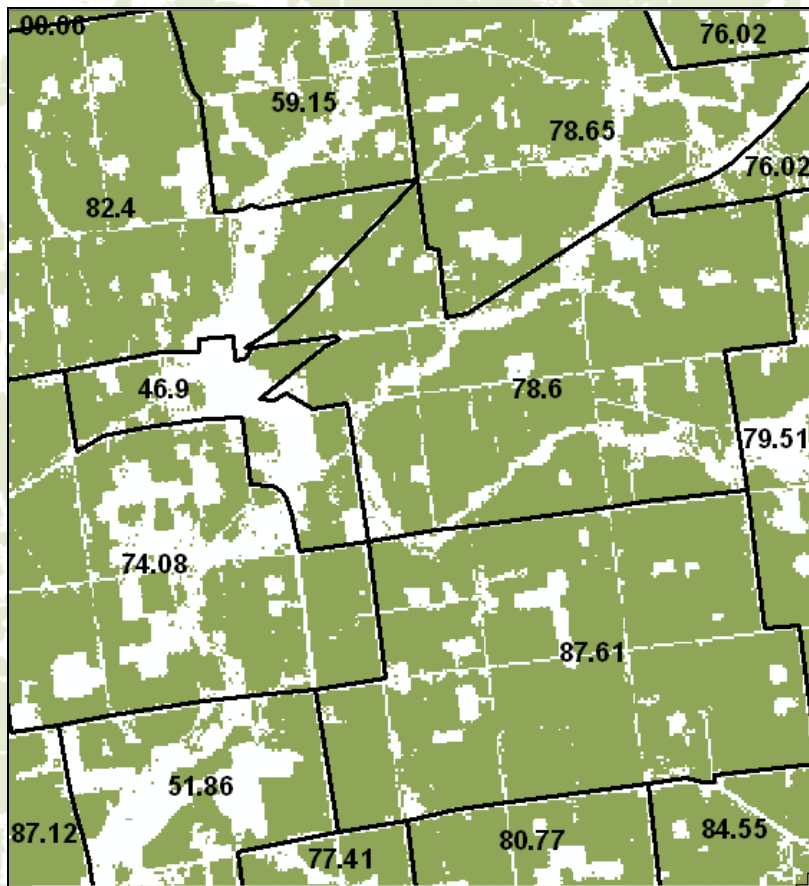
Primary Sampling Units with CDL percent cultivation
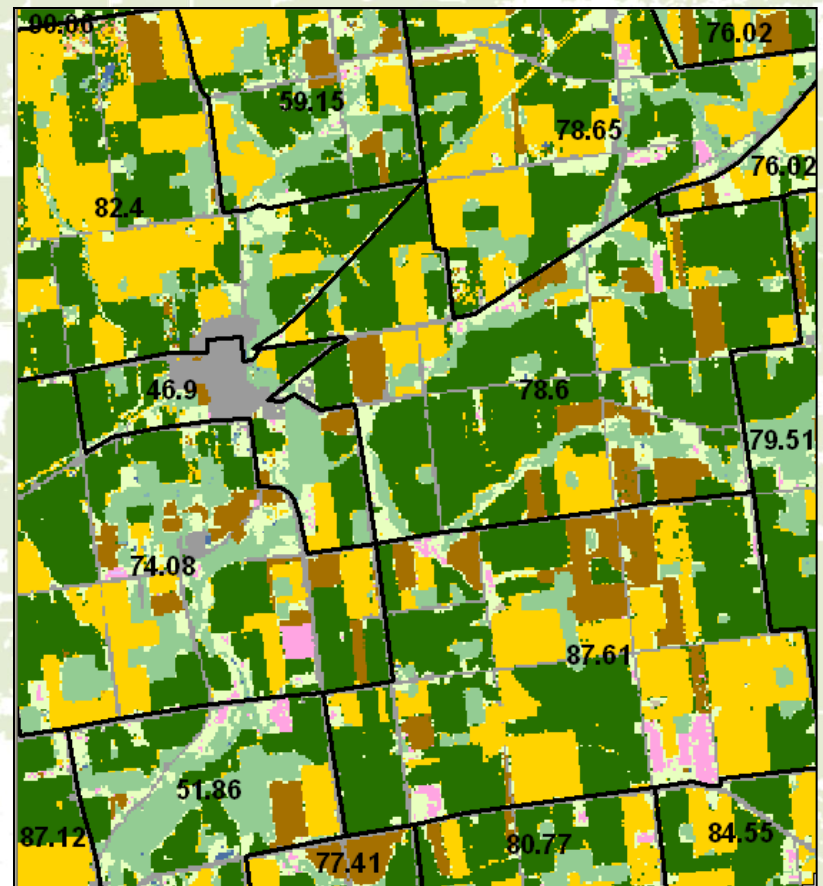
Primary Sampling Units with CDL percent cultivation, overlaying a 2010 CDL image product

Cropland Data Layer (CDL) based stratification of a NASS Area Sampling Frame (ASF)
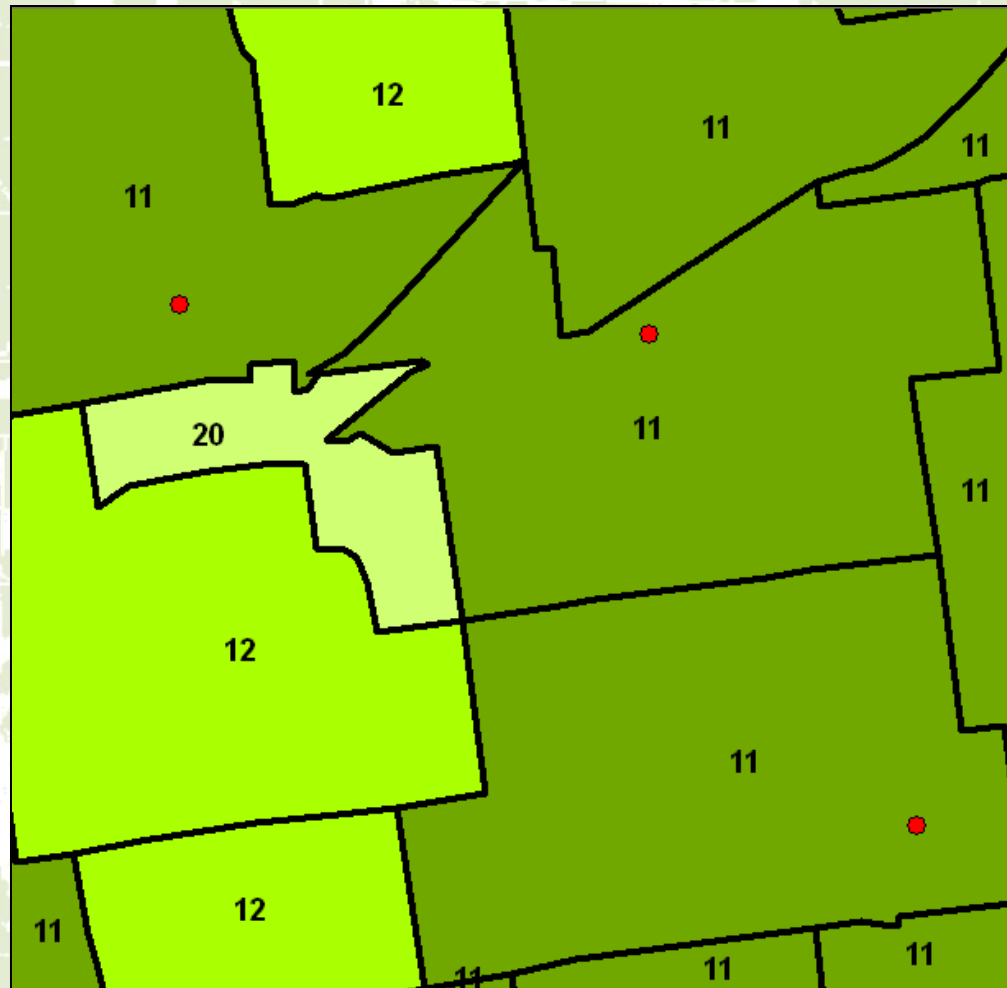
Red dots are location points of in situ validation collected during the 2010 June Area Survey

Objective

The objective of this investigation was to determine the utility of the automated Cropland Data Layer (CDL) based stratification method for use in Area Sampling Frame (ASF) construction.

CDL based stratification of NASS ASF Primary Sampling Units (PSUs) was successfully conducted for Arizona, Georgia, Ohio, Oklahoma, Virginia.

Area Frame manual stratification
matches
CDL based automated stratification
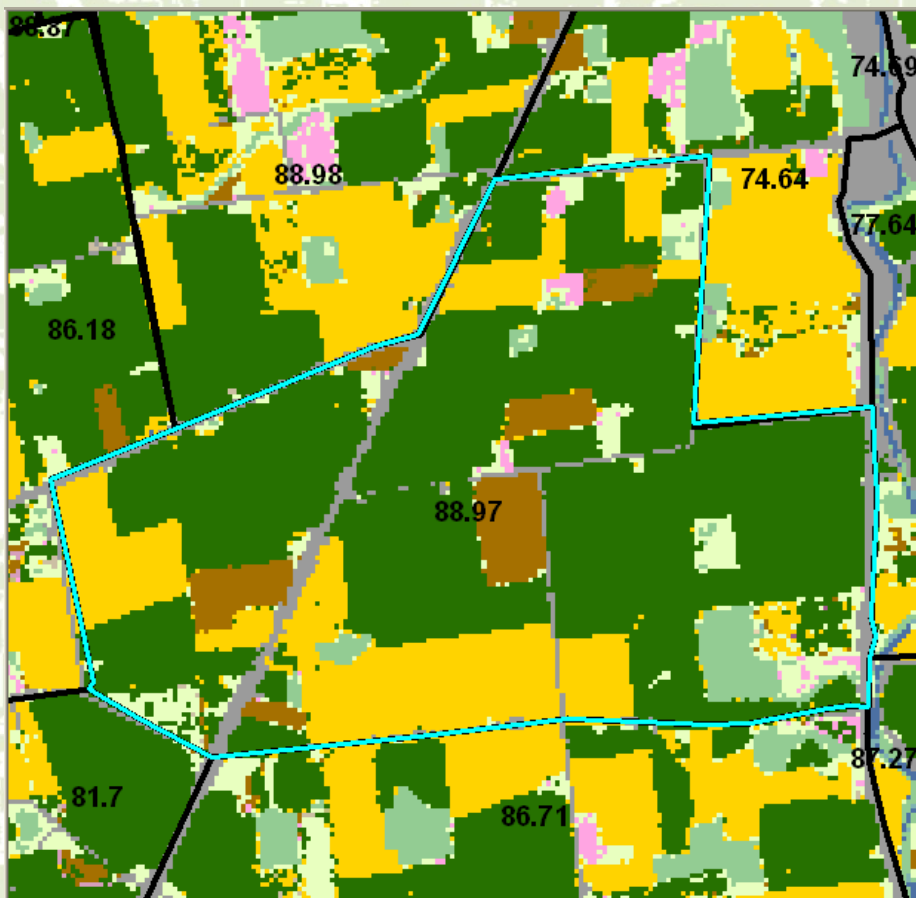
CDL percent cultivation – 88.97%
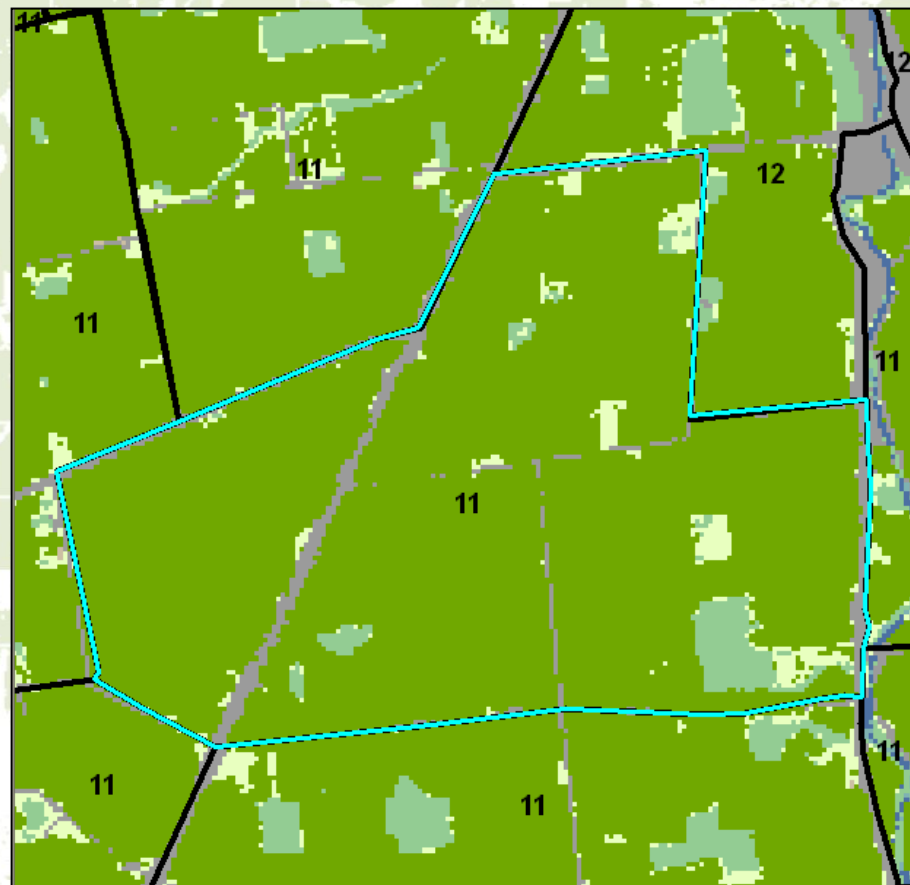
AF stratification – 11 (manual)
CDL stratification – 11 (automated)

Area Frame manual stratification
does not match
Cropland Data Layer based automated stratification
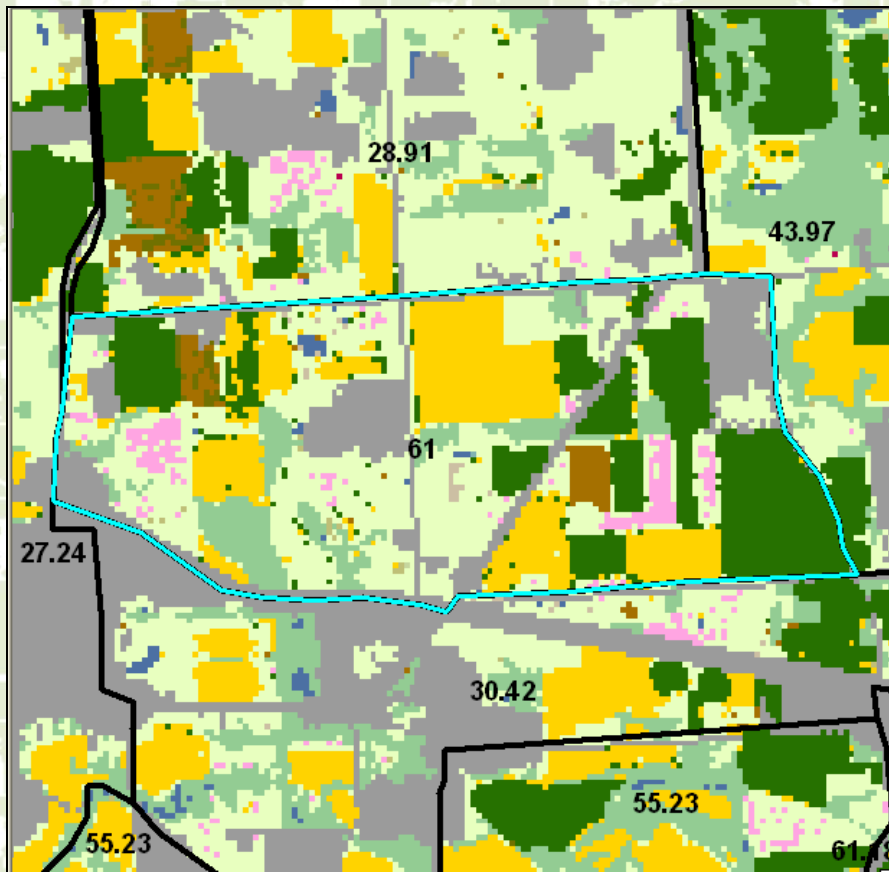
CDL percent cultivation – 61%
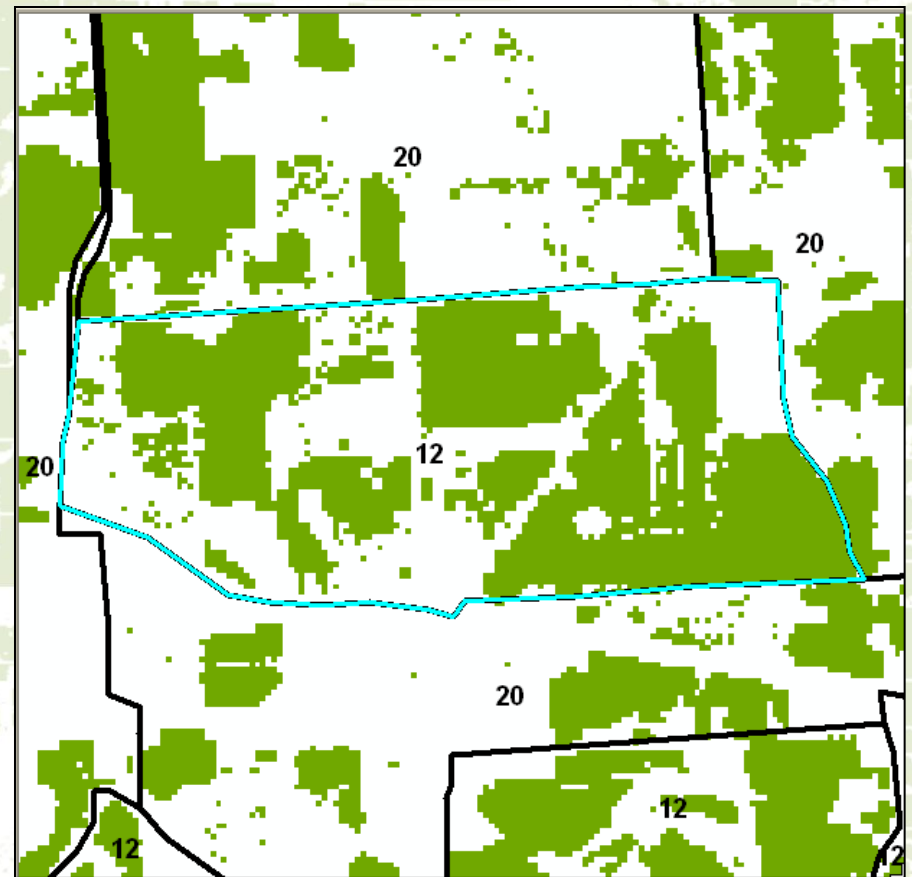
AF stratification – 11 (manual)
CDL stratification – 12 (automated)

Evaluation

Stratification accuracy was measured using in-situ data collected by enumerators during the 2010 June Area Survey (JAS) in the five states evaluated.

Accuracy measures were derived by comparing the strata definitions reported by JAS enumerators with the original ASF manual stratification and the CDL based automated stratification.

Evaluation

To determine if the percentage differences between the original Area Frame stratification method and the CDL based stratification method were statistically significant at a 95% confidence level, a two-tailed proportion test was used.

These tests were performed with Chi-Square and Fisher's Exact tests when the sample sizes were less than five

Evaluation

The hypotheses of the significance tests were H_0: $p_1=p_2$ and H_a: $p_1 \neq p_2$. The null hypothesis stated that there was no difference in the results of the two stratification methods while the alternative hypothesis stated that the results of the two stratification methods were significantly different.

The tests were performed and p values were calculated for each state and each stratum with a confidential level of 95%.

Area Frame vs. CDL Stratification Oklahoma 2010

Stratum	Survey Ratio (% Cultivated)	AREA FRAME Stratification			CDL Stratification			P-value
		Total Segments	Total Reported	Percentage (p_1)	Total Segments	Total Reported	Percentage (p_2)	Ha: $p_1 \neq p_2$
11	>=75%	140	47	34%	43	27	63%	0.001
12	51% - 75%	48	9	19%	77	30	39%	0.024
20	15% - 50%	74	26	35%	98	42	43%	0.305
40	< 15%	61	61	100%	105	96	91%	0.027
Total		323			323			

Area Frame vs. CDL Stratification Oklahoma 2010

Area Frame vs. CDL Stratification
Five State - Strata Summary, 2010

Stratum	Survey Ratio (% Cultivated)	AREA FRAME Stratification			CDL Stratification			P-value
		Total Segments	Total Reported	Percentage (p_1)	Total Segments	Total Reported	Percentage (p_2)	Ha: $p_1 \neq p_2$
11	>75%	250	131	52%	128	103	80%	0.000
12	51% - 75%	83	24	29%	119	53	45%	0.025
13	>50%	171	90	53%	91	69	76%	0.000
20	15% - 50%	371	177	48%	387	219	57%	0.000
40	< 15%	322	305	95%	472	407	86%	0.000
Total		1197	727	61%	1197	851	71%	0.000

Five State Analysis - Arizona, Georgia, Ohio, Oklahoma, Virginia.

Area Frame vs. CDL Stratification
Five State - Strata Summary 2010

■ AREA FRAME Stratification   ■ CDL Stratification

Percent Segments

Stratum & Percent Cultivation	AREA FRAME	CDL
>75% (11)	52%	80%
51% - 75% (12)	29%	45%
>50% (13)	53%	76%
15% - 50% (20)	48%	57%
< 15% (40)	95%	86%
Total	61%	71%
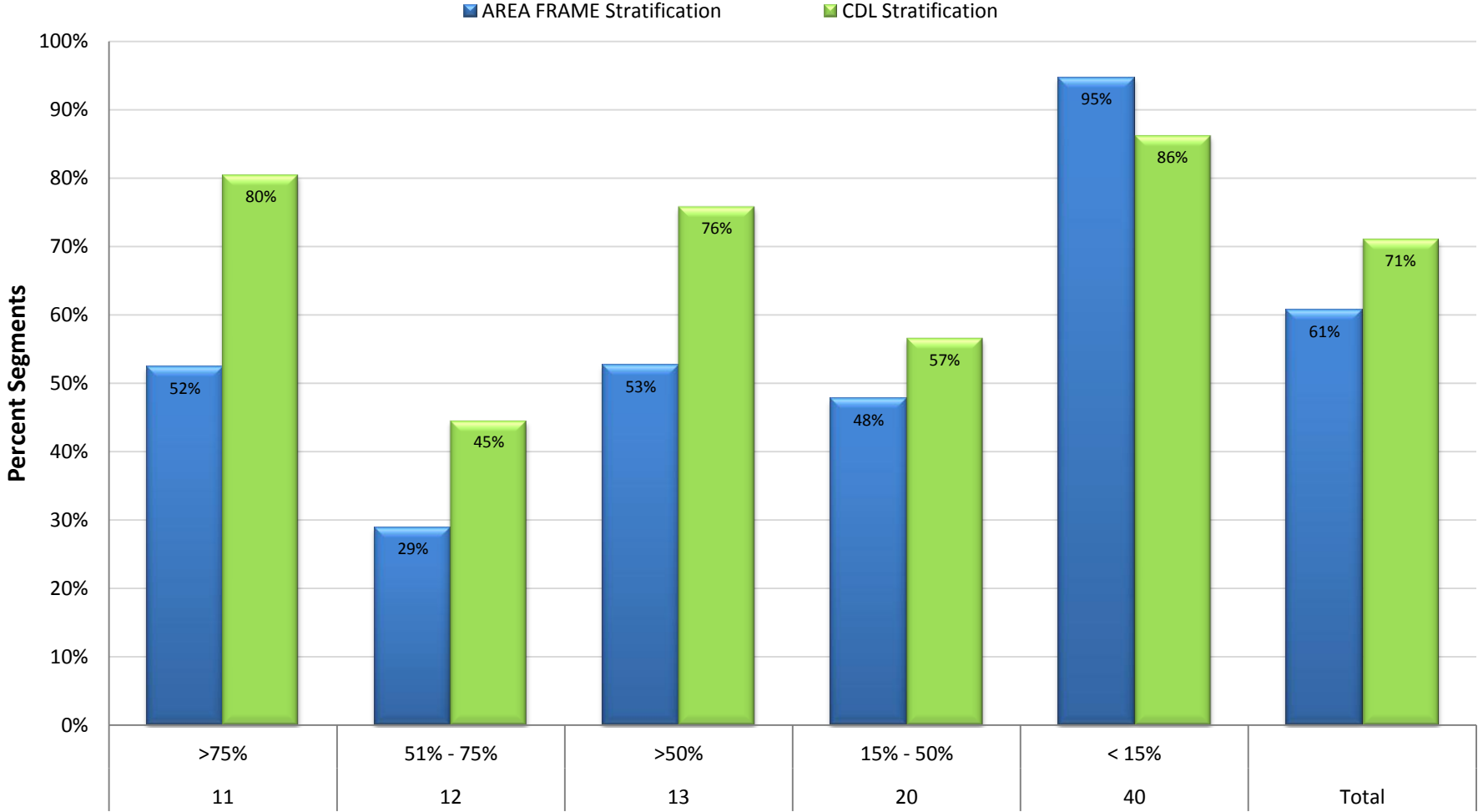
Stratum & Percent Cultivation

Conclusion

Results of the five state analyses indicated that the new automated Cropland Data Layer (CDL) stratification method performed well in determining U.S. percent cultivation **in moderate to highly intensive cropped areas** and weaker in non agricultural areas.

The strength of the CDL product and the CDL based stratification method is the **objective and consistent identification of cultivated cropland**.

Conclusion

The Cropland Data Layer based stratification method can be used for

- ***review of current Area Sampling Frames***

- ***as a change detection technique***

- ***as the primary method of stratification***

The Cropland Data Layer based automated stratification method should improve the efficiency, reduce the cost and improve the precision of the June Agricultural Survey estimates.

Thank you!
Questions?

Claire G. Boryan, Geographer
National Agricultural Statistics Service USDA
3251 Old Lee Hwy
Fairfax VA
703-877-8000 x 107
Claire_boryan@nass.usda.gov