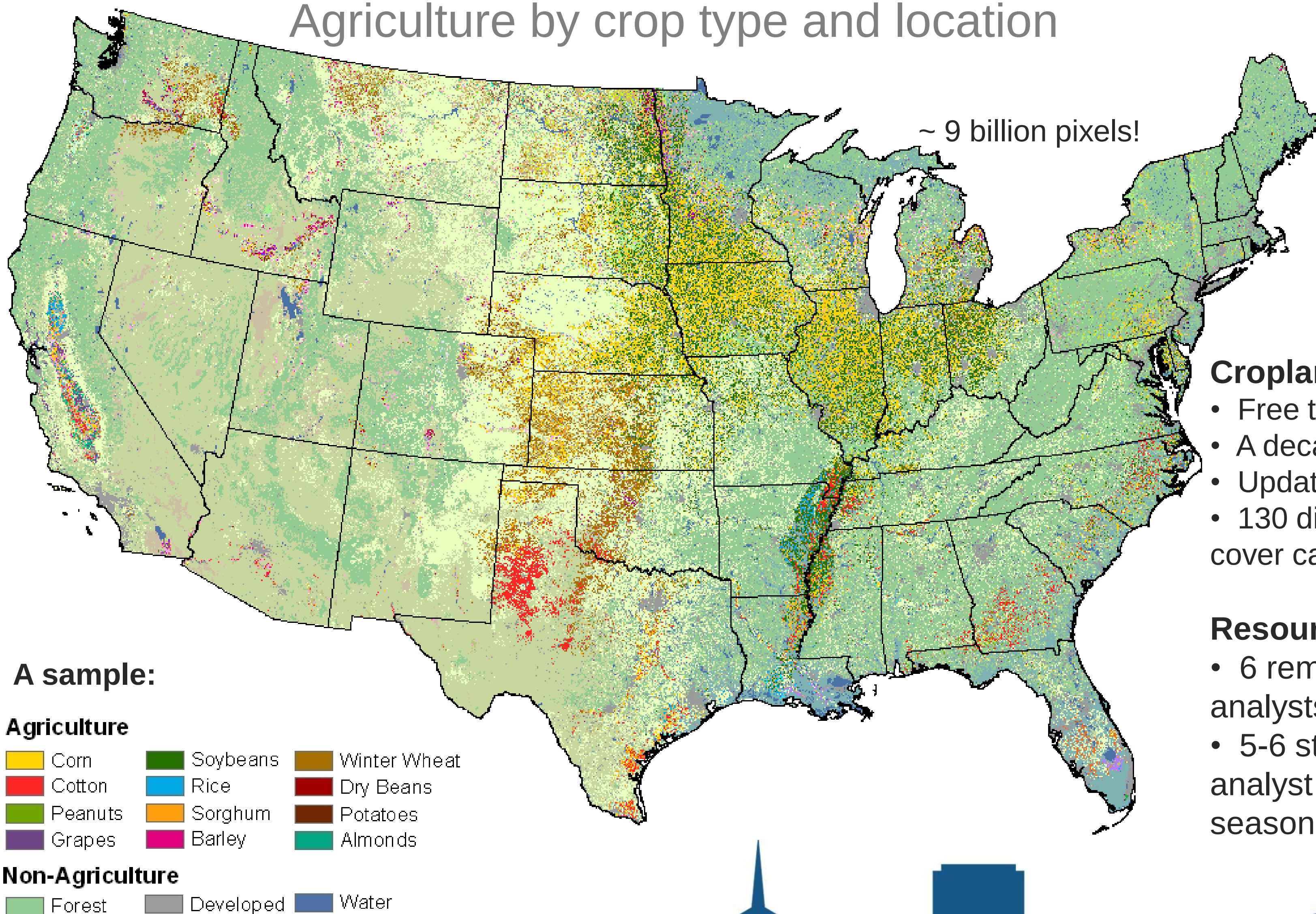
An Innovative Approach to Integrating SAS® Macros with GIS Software Products to Produce County-Level Accuracy Assessments

Audra Zakzeski
USDA NASS

Robert Seffrin
USDA NASS

Introducing the Cropland Data Layer

Agriculture by crop type and location



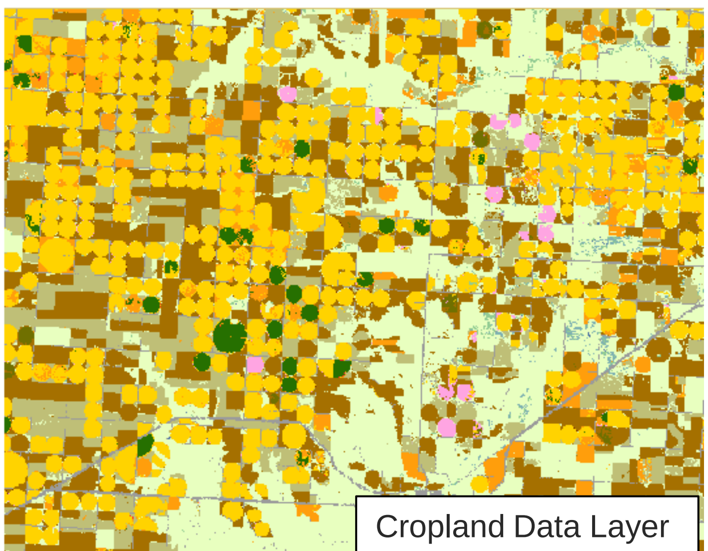
Cropland Data Layer

- Free to the public
- A decade of history
- Updated annually
- 130 different land cover categories

Resources

- 6 remote sensing analysts/statisticians
- 5-6 states per analyst during growing season

Satellite Imagery to Land Cover Map



Processing Steps

1. Gather satellite imagery from April - October
2. Gather farm/crop data from Farm Service Agency
3. Take sample of satellite pixels for each different type of land cover and develop seasonal profile of the land
4. Construct decision tree of all land cover profiles
5. Apply decision tree to all satellite image pixels and areas of the land resulting in final Cropland Data Layer
6. Derive accuracy assessments for all states & types of land cover using validation data.

SAS GLOBAL FORUM

		Correct	Total	Accuracy	Error	Kappa
Overall		7,309,535	8,689,040	84.1%	15.9%	0.831
FSA Crops		5,624,069	6,734,428	83.5%	16.5%	0.822
Principal Crops		2,738,630	3,385,003	80.9%	19.1%	0.780
Tilled Crops		3,602,592	4,508,691	79.9%	20.1%	0.779
Forage		1,223,482	1,466,009	83.5%	16.5%	0.589
Vegetables		413,725	592,861	69.8%	30.2%	0.554
Orchards		797,995	1,000,151	79.8%	20.2%	0.732
Berries		2,527	4,881	51.8%	48.2%	0.078

Accuracy Assessments

Category		Correct	Total	Producer	Omission	Kappa	Total	User	Commission	Kappa	Bias
Corn	1	279,637	312,084	89.6%	10.4%	0.892	306,419	91.3%	8.7%	0.909	-1.8%
Cotton	2	555,887	567,233	98.0%	2.0%	0.979	589,463	94.3%	5.7%	0.939	3.9%
Rice	3	827,370	834,301	99.2%	0.8%	0.991	840,327	98.5%	1.5%	0.983	0.7%
Sorghum	4	5,088	13,472	37.8%	62.2%	0.377	6,778	75.1%	24.9%	0.750	-49.7%
Sunflower	6	51,832	62,330	83.2%	16.8%	0.830	62,241	83.3%	16.7%	0.832	-0.1%
Sweet Corn	12	1,427	6,188	23.1%	76.9%	0.230	2,782	51.3%	48.7%	0.513	-55.0%
Mint	14	1,235	2,117	58.3%	41.7%	0.583	2,008	61.5%	38.5%	0.615	-5.1%
Barley	21	83,791	132,542	63.2%	36.8%	0.628	95,457	87.8%	12.2%	0.876	-28.0%
Durum Wheat	22	136,946	166,874	82.1%	17.9%	0.817	167,081	82.0%	18.0%	0.816	0.1%
Spring Wheat	23	11,984	15,053	79.6%	20.4%	0.796	15,355	78.0%	22.0%	0.780	2.0%
Winter Wheat	24	374,442	512,300	73.1%	26.9%	0.715	492,364	76.0%	24.0%	0.745	-3.9%
Other Small Grains	25	2	38	5.3%	94.7%	0.053	52	3.8%	96.2%	0.038	36.8%
Rye	27	7,136	16,224	44.0%	56.0%	0.439	8,805	81.0%	19.0%	0.810	-45.7%
Oats	28	69,894	141,268	49.5%	50.5%	0.489	99,494	70.2%	29.8%	0.698	-29.6%
Canola	31	350	586	59.7%	40.3%	0.597	355	98.6%	1.4%	0.986	-39.4%
Safflower	33	57,458	69,848	82.3%	17.7%	0.821	62,974	91.2%	8.8%	0.912	-9.8%
Alfalfa	36	1,040,280	1,099,990	94.6%	5.4%	0.938	1,125,969	92.4%	7.6%	0.913	2.4%
Non-alfalfa Hay	37	150,876	233,510	64.6%	35.4%	0.638	185,297	81.4%	18.6%	0.809	-20.6%
Sugarbeets	41	34,067	42,941	79.3%	20.7%	0.792	40,586	83.9%	16.1%	0.839	-5.5%
Dry Beans	42	13,377	26,456	50.6%	49.4%	0.505	19,228	69.6%	30.4%	0.695	-27.3%
Potatoes	43	15,995	21,514	74.3%	25.7%	0.743	20,634	77.5%	22.5%	0.775	-4.1%
Other Crops	44	9,203	13,001	70.8%	29.2%	0.708	10,510	87.6%	12.4%	0.875	-19.2%
Sweet Potatoes	46	1,991	2,867	69.4%	30.6%	0.694	2,626	75.8%	24.2%	0.758	-8.4%
Misc. Vegs. & Fruits	47	9,193	16,920	54.3%	45.7%	0.543	12,480	73.7%	26.3%	0.736	-26.2%
Watermelons	48	1,334	11,624	11.5%	88.5%	0.114	2,894	46.1%	53.9%	0.460	-75.1%
Onions	49	33,809	48,526	69.7%	30.3%	0.695	41,161	82.1%	17.9%	0.820	-15.2%
Cucumbers	50	297	638	46.6%	53.4%	0.465	667	44.5%	55.5%	0.445	4.5%
Peas	53	2,299	9,035	25.4%	74.6%	0.254	4,467	51.5%	48.5%	0.514	-50.6%
Tomatoes	54	296,161	328,490	90.2%	9.8%	0.898	331,775	89.3%	10.7%	0.888	1.0%
Herbs	57	2,724	5,654	48.2%	51.8%	0.482	3,676	74.1%	25.9%	0.741	-35.0%
Clover/Wildflowers	58	31,534	38,233	82.5%	17.5%	0.824	34,425	91.6%	8.4%	0.916	-10.0%
Sod/Grass Seed	59	3,262	6,744	48.4%	51.6%	0.483	3,396	96.1%	3.9%	0.961	-49.6%
Fallow/Idle Cropland	61	366,184	486,556	75.3%	24.7%	0.738	497,254	73.6%	26.4%	0.721	2.2%
Cherries	66	9,602	15,968	60.1%	39.9%	0.601	12,802	75.0%	25.0%	0.750	-19.8%
Peaches	67	791	3,249	24.3%	75.7%	0.243	1,243	63.6%	36.4%	0.636	-61.7%
Apples	68	1,542	1,989	77.5%	22.5%	0.775	1,953	79.0%	21.0%	0.790	-1.8%
Grapes	69	175,999	201,880	87.2%	12.8%	0.869	198,892	88.5%	11.5%	0.882	-1.5%
Other Tree Nuts	71	3,573	9,810	36.4%	63.6%	0.364	4,177	85.5%	14.5%	0.855	-57.4%
Citrus	72	27,319	35,162	77.7%	22.3%	0.776	30,121	90.7%	9.3%	0.907	-14.3%
Pecans	74	264	1,804	14.6%	85.4%	0.146	1,120	23.6%	76.4%	0.236	-37.9%
Almonds	75	393,278	438,669	89.7%	10.3%	0.891	426,927	92.1%	7.9%	0.917	-2.7%
Walnuts	76	79,044	102,035	77.5%	22.5%	0.772	94,614	83.5%	16.5%	0.833	-7.3%
Pears	77	4,020	4,856	82.8%	17.2%	0.828	4,618	87.1%	12.9%	0.870	-4.9%
Aquaculture	92	349	1,275	27.4%	72.6%	0.274	544	64.2%	35.8%	0.641	-57.3%
Open Water	111	92,622	95,363	97.1%	2.9%	0.971	97,267	95.2%	4.8%	0.952	2.0%
Perennial Ice/Snow	112	144	258	55.8%	44.2%	0.558	193	74.6%	25.4%	0.746	-25.2%
Developed, Open Spa	121	51,357	62,278	82.5%	17.5%	0.823	85,034	60.4%	39.6%	0.601	36.5%
Developed, Low Inten:	122	22,688	33,557	67.6%	32.4%	0.675	37,599	60.3%	39.7%	0.602	12.0%
Developed, Medium Ir	123	27,740	34,231	81.0%	19.0%	0.810	35,148	78.9%	21.1%	0.788	2.7%
Developed, High Inten	124	6,036	7,911	76.3%	23.7%	0.763	7,469	80.8%	19.2%	0.808	-5.6%
Barren Land	131	78,964	106,912	73.9%	26.1%	0.735	107,363	73.5%	26.5%	0.732	0.4%
Deciduous Forest	141	4,895	16,520	29.6%	70.4%	0.295	11,442	42.8%	57.2%	0.427	-30.7%
Evergreen Forest	142	369,695	401,401	92.1%	7.9%	0.917	427,489	86.5%	13.5%	0.858	6.5%
Mixed Forest	143	22,813	49,303	46.3%	53.7%	0.460	40,331	56.6%	43.4%	0.563	-18.2%
Shrub/Scrub	152	800,386	892,433	89.7%	10.3%	0.885	928,685	86.2%	13.8%	0.846	4.1%
Grassland/Herbaceou	171	200,317	257,306	77.9%	22.1%	0.768	400,938	50.0%	50.0%	0.484	55.8%
Woody Wetlands	190	2,452	4,770	51.4%	48.6%	0.514	7,014	35.0%	65.0%	0.349	47.0%
Herbaceous Wetlands	195	5,008	10,363	48.3%	51.7%	0.482	19,430	25.8%	74.2%	0.257	87.5%
Pistachios	204	67,452	88,515	76.2%	23.8%	0.760	79,731	84.6%	15.4%	0.844	-9.9%
Triticale	205	24,109	46,411	51.9%	48.1%	0.518	31,555	76.4%	23.6%	0.763	-32.0%
Carrots	206	11,712	19,846	59.0%	41.0%	0.589	17,864	65.6%	34.4%	0.655	-10.0%
Asparagus	207	1,452	4,592	31.6%	68.4%	0.316	1,654	87.8%	12.2%	0.878	-64.0%
Garlic	208	15,213	20,594	73.9%	26.1%	0.738	17,333	87.8%	12.2%	0.877	-15.8%
Cantaloupes	209	6,249	13,111	47.7%	52.3%	0.476	9,095	68.7%	31.3%	0.687	-30.6%
Olives	211	14,419	20,471	70.4%	29.6%	0.704	16,175	89.1%	10.9%	0.891	-21.0%
Oranges	212	6,986	9,961	70.1%	29.9%	0.701	8,335	83.8%	16.2%	0.838	-16.3%
Honeydew Melons	213	3,606	5,943	60.7%	39.3%	0.607	4,812	74.9%	25.1%	0.749	-19.0%
Broccoli	214	315	3,523	8.9%	91.1%	0.089	1,672	18.8%	81.2%	0.188	-52.5%
Peppers	216	1,487	5,051	29.4%	70.6%	0.294	2,204	67.5%	32.5%	0.674	-56.4%
Pomegranates	217	8,724	14,062	62.0%	38.0%	0.620	9,936	87.8%	12.2%	0.878	-29.3%
Nectarines	218	195	725	26.9%	73.1%	0.269	325	60.0%	40.0%	0.600	-55.2%
Greens	219	1,403	6,061	23.1%	76.9%	0.231	5,235	26.8%	73.2%	0.267	-13.6%
Plums	220	4,512	12,327	36.6%	63.4%	0.366	6,870	65.7%	34.3%	0.656	-44.3%
Strawberries	221	2,434	4,382	55.5%	44.5%	0.555	2,575	94.5%	5.5%	0.945	-41.2%
Squash	222	224	2,189	10.2%	89.8%	0.102	728	30.8%	69.2%	0.308	-66.7%
Apricots	223	275	572	48.1%	51.9%	0.481	548	50.2%	49.8%	0.502	-4.2%
Vetch	224	792	2,904	27.3%	72.7%	0.273	1,238	64.0%	36.0%	0.640	-57.4%
Dbl. Crop WinWht/Co	225	203,341	241,209	84.3%	15.7%	0.838	273,879	74.2%	25.8%	0.735	13.5%
Dbl. Crop Oats/Corn	226	52,278	76,538	68.3%	31.7%	0.680	83,426	62.7%	37.3%	0.623	9.0%
Lettuce	227	6,808	16,385	41.6%	58.4%	0.415	11,338	60.0%	40.0%	0.600	-30.8%
Dbl. Crop Lettuce/Dur	230	1,549	6,188	25.0%	75.0%	0.250	4,085	37.9%	62.1%	0.379	-34.0%
Dbl. Crop Lettuce/Car	231	391	757	51.7%	48.3%	0.516	1,349	29.0%	71.0%	0.290	78.2%
Dbl. Crop Lettuce/Upl.	232	312	859	36.3%	63.7%	0.363	654	47.7%	52.3%	0.477	-23.9%
Dbl. Crop Barley/Sorg	235	252	797	31.6%	68.4%	0.316	480	52.5%	47.5%	0.525	-39.8%
Dbl. Crop WinWht/Sor	236	13,962	26,290	53.1%	46.9%	0.530	21,397	65.3%	34.7%	0.651	-18.6%
Dbl. Crop Barley/Corr	237	938	2,975	31.5%	68.5%	0.315	1,380	68.0%	32.0%	0.680	-53.6%
Dbl. Crop WinWht/Cot	238	63	345	18.3%	81.7%	0.183	73	86.3%	13.7%	0.863	-78.8%
Blueberries	242	93	237	39.2%	60.8%	0.392	173	53.8%	46.2%	0.538	-27.0%
Cabbage	243	26	344	7.6%	92.4%	0.076	162	16.0%	84.0%	0.160	-52.9%

← Save your eyes – Don't try to read that – it's just a visual

Things to measure for each crop per state

1. Producer Accuracy/Omission Error
   - Probability a validation pixel is correctly mapped.
   - Occurs when a pixel from the validation is excluded from the correct category
2. User Accuracy/Commission Error
   - Probability a CDL pixel matches the validation pixel
   - Occurs when a pixel is included in the incorrect category
3. Confidence Value
   - Measure of how easy it was to classify a pixel – calculated in ERDAS Imagine
4. Acreage Estimate & Coefficient of Variation
   - Calculated in SAS using Linear Regression
5. Pixel Count Acreage
   - Count of pixels for each land cover

SAS GLOBAL FORUM

Step 1: Batch file to run ERDAS Imagine

Calculating producer accuracy, user accuracy, confidence value, and pixel counts happens in ERDAS Imagine.

Use SAS to create batch files calling ERDAS Imagine and accuracy calculations

```
%LET DirEst = File path for Estimate file;
%LET DirCnty = File path for state/county outlines;
%LET IMGWork = ERDAS Model .exe file path;
%LET Model = Preexisting ERDAS .pmdl model file path;
%LET Zone1 = &DirCnty/Specific State Outline;
%LET Zone2 = File path for CDL File #1;
%LET Class = File path for CDL File #2;
%LET MTXout = File path for preexisting matrix file (.mtx);
%LET IMGlaunch = &IMGwork &Model -s -m &Zone1 &Zone2 &Class &MTXout;
OPTIONS NOXWAIT XSYNC;
```

Define file path locations

```
DATA _NULL_;
FILE "File location to be created batch file" LRECL=600;
PUT 'Set Imagine_Batch_Run=1";
PUT 'ECHO off ';
PUT 'TITLE ' "&StYr2. &Tabulated";
PUT 'COLOR 79 ';
PUT 'ECHO Mosaic running from N:\Estimates\Acreage\Temp\Tabulate.bat";
PUT 'TIME /t';
PUT %unquote(%str(%')&IMGLanch %str(%));
RUN;
```

Code to create batch files to run ERDAS Imagine

Step 2: Reformat Matrix File

- The resulting matrix file created by the batch file is converted to column formatted data.

```
DATA _NULL_;
SET MtxColNames;
LENGTH List Rename $2000;
ARRAY Class(*) &ClassRng;
DO i = 1 to HBOUND (Class);
List = Trim(List)||’ c’||PUT(Class[i],z3.);
Rename = Trim(Rename)||’ c’||
      PUT(i-1,z4.)||’c=’||PUT(Class[i],z3.);
END;
CALL SYMPUT(‘ClassList’,List);
CALL SYMPUT(‘Rename’, Rename);
RUN;

PROC DATASETS LIBRARY=Work NODETAILS NOLIST;
MODIFY MtxMain;
RENAME &Rename;
QUIT;
```

County		1 (Corn)	2 (Cotton)	3 (Rice)	4 (Sorghum)
6	1 (Corn)	1523	54	74	12
6	2 (Cotton)	25	6841	45	36
6	3 (Rice)	0	68	842	95
6	4 (Sorghum)	1	8	25	751

County ID	Validation Category	Classified CDL Category	Pixel Count
6	1 (Corn)	1 (Corn)	1523
6	1 (Corn)	2 (Cotton)	25
6	1 (Corn)	3 (Rice)	0
6	1 (Corn)	4 (Sorghum)	1
6	2 (Cotton)	1 (Corn)	54
6	2 (Cotton)	2 (Cotton)	6841
6	2 (Cotton)	3 (Rice)	68
6	2 (Cotton)	4 (Sorghum)	8
6	3 (Rice)	1 (Corn)	74
6	3 (Rice)	2 (Cotton)	45
6	3 (Rice)	3 (Rice)	842
6	3 (Rice)	4 (Sorghum)	25

These values are used to calculate producer and user accuracy

Step 3: Creating the Template

Reference Map

Map of Pixel Counts

Bar Chart/Scatterplot Legend

Bar Chart/Scatterplot

Producer Accuracy Choropleth Map

User Accuracy Choropleth Map

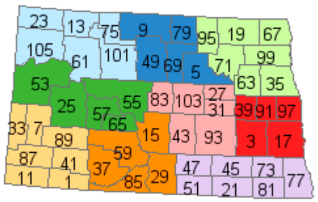
Legend

Confidence Value Choropleth Map

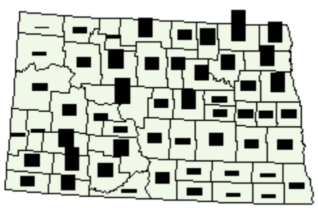
*All data to be displayed is for display purposes only.

Accuracy Assessment Dashboard

County FIPS and Districts



Pixels merged



County estimate bars

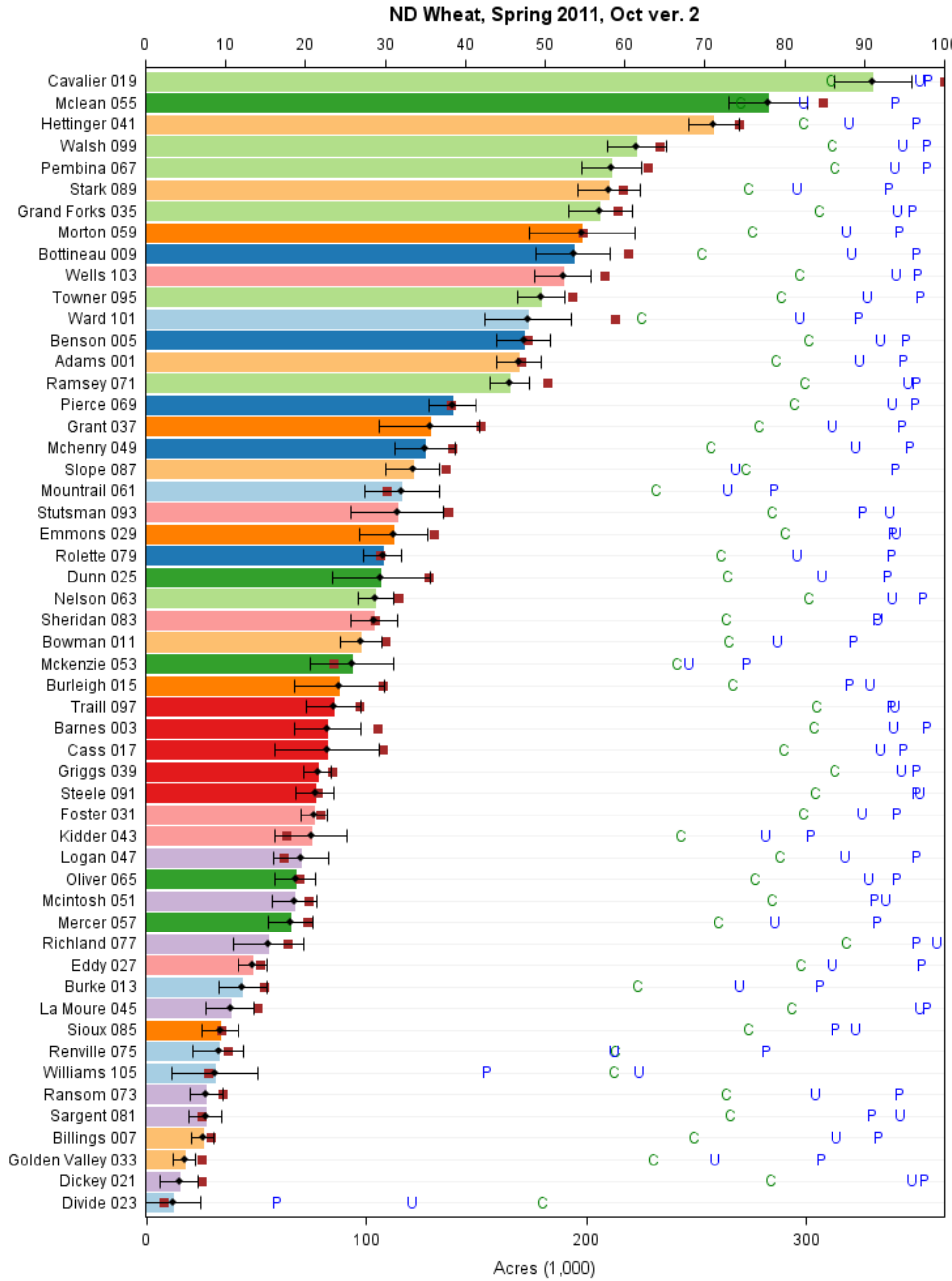
Scale, percent (top)

- P Accuracy, Producer
- U Accuracy, User
- C Confidence, mean

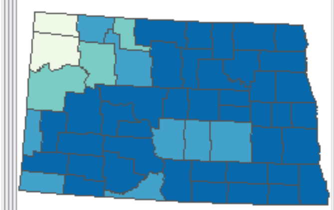
Scale, acres (bottom)

- CDL estimate
- 2 Standard errors
- Pixel acres
- FSA physical
- Stat. final

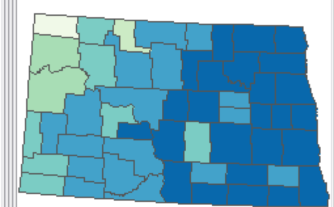
*All data displayed is for display purposes only.



Accuracy, Producer



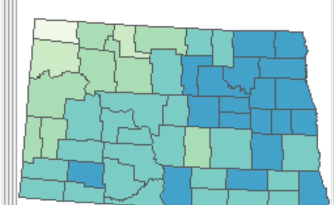
Accuracy, User



Choropleth ranges

- 0 – < 50
- 50 – < 60
- 60 – < 70
- 70 – < 80
- 80 – < 90
- 90 – 100

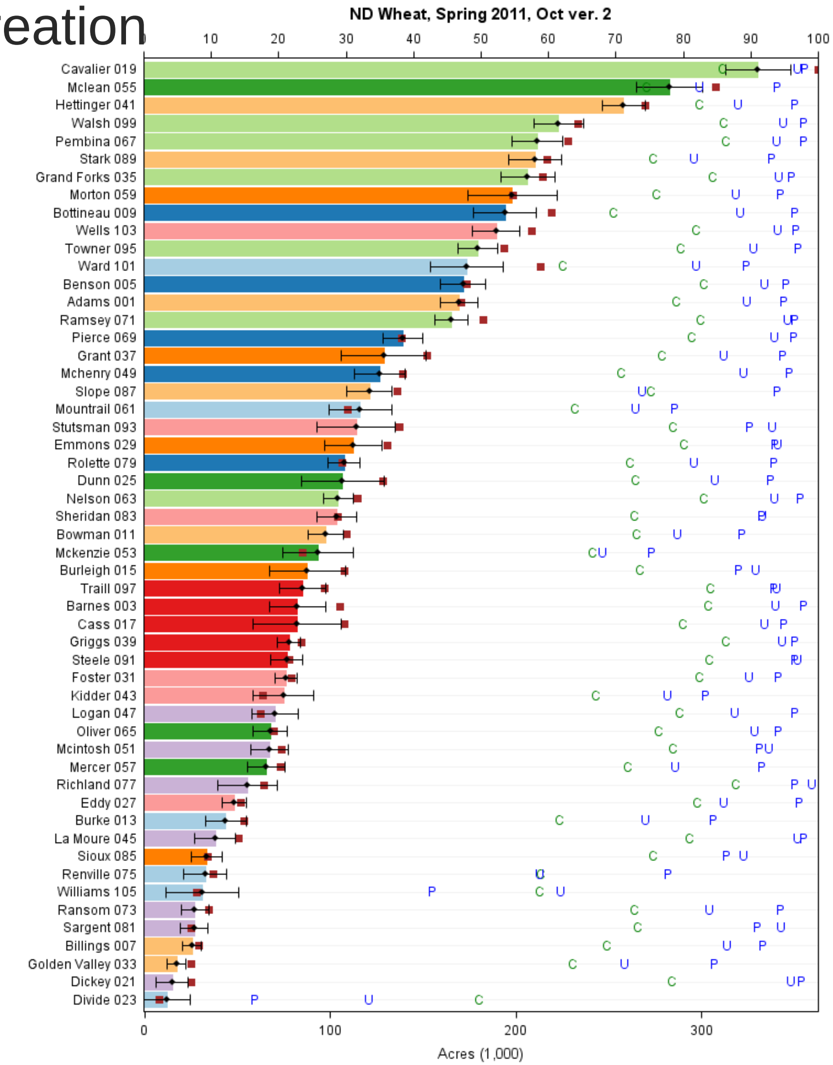
Confidence, Mean



Step 4: Constructing Pieces

- Bar Chart/Scatterplots Template Creation

```
PROC TEMPLATE;
    DEFINE STATGRAPH Bar_Chart_county;
    DYNAMIC _yAxis “Category var” _xAxis “Continuous var”;
    MVAR _Title “Title” _SF_Counts “StatFinal count”;
    begingraph;
      entrytitle _Title;
      layout overlay /
        x2axisopts= …
        xaxisopts= …
        yaxisopts= …
    REFERENCELINE Y= _yAxis / datatransparency=0.9;


barchart x= _yAxis / options…

scatterplot y=_yAxis x=Pixel_Acres / options …

scatterplot y=_yAxis x=Estimate / options …

scatterplot y=_yAxis x=Conf_Mean / options …     C

scatterplot y=_yAxis x=Conf_Mean / options …     P

scatterplot y=_yAxis x=Conf_Mean / options …     U

ENDLAYOUT;
ENDGRAPH;
END;
RUN;
```



*All data displayed is for display purposes only.

Step 5: Assembling HTML Columns

- Column 1 – 2 maps, 1 legend

```
filename htmlpath "C:\Temp";
ods listing close;
ods tagsets.htmlpanel nogtitle path=htmlPath
     file="Final.html" style=Paired
     options(panelborder='1' panelcolumns='4 3')
```

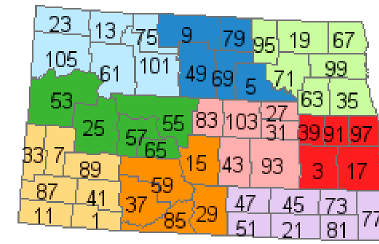
Defines html panel 4 rows, 3 columns

```
ods tagsets.htmlpanel event=row_panel(start);
ods tagsets.htmlpanel event=column_panel(start);
```

Start construction Row 1, Column 1

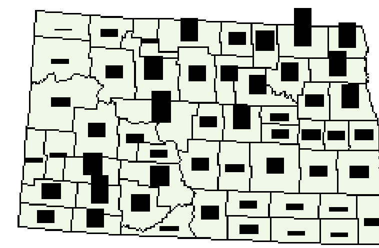
```
DATA Anno_Map_ASD …
TITLE1 "County FIPS and Districts";
PROC GANNO ANNOTATE=Anno_Map_ASD;
```

Annotated County Map



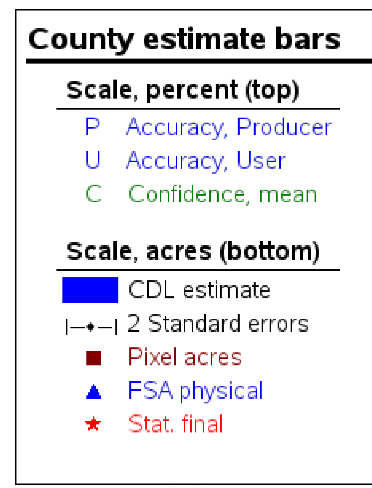

```
TITLE "Pixels merged";
PROC GMAP …
```

Pixel Counts by County



```
DATA My_anno…
goptions xpixel=240 ypixel=300;
PROC GANNO ANNOTATE=My_anno;
```

Legend – created manually, stored as picture



SAS GLOBAL FORUM

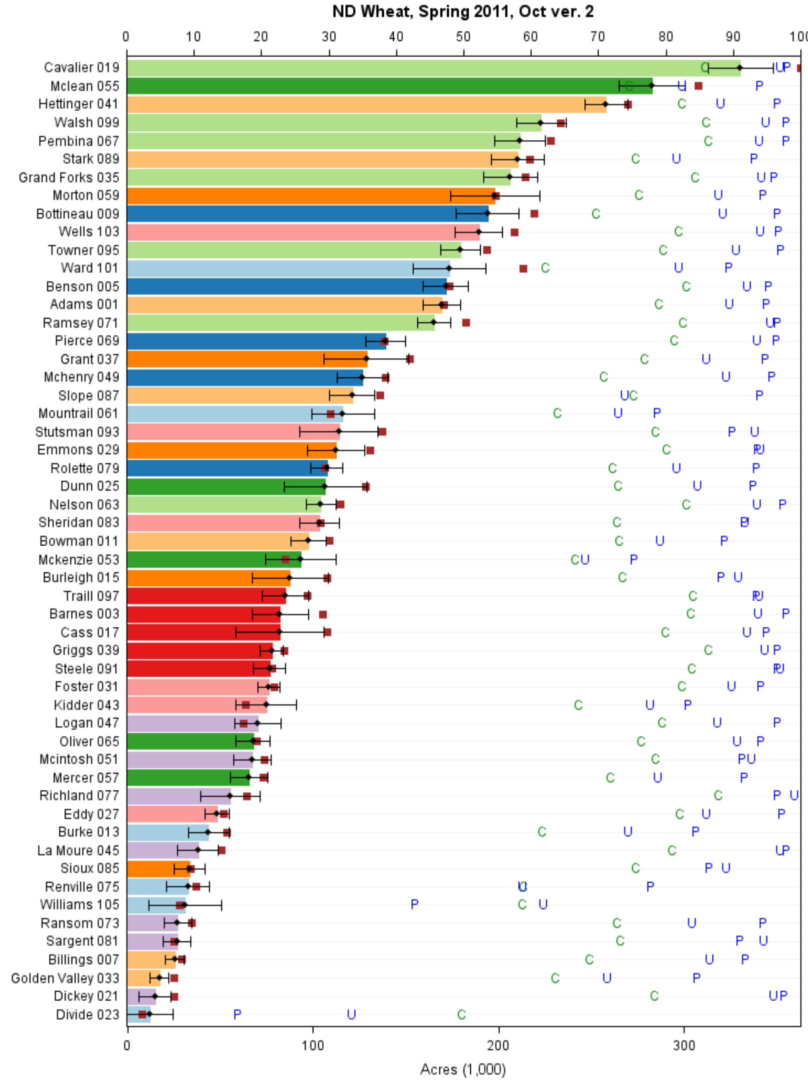
Step 5: Assembling HTML Columns

- Bar Chart / Scatterplots Placement

```
ods tagsets.htmlpanel event=column_panel(finish);
ods tagsets.htmlpanel event=column_panel(start);


ods graphics / reset noborder width=750px height=1000px
imagename=”&File” imagefmt=png noscale;


PROC SGENDER DATA=RevGraph.ND12 TEMPLATE=Bar_Chart_county;
DYNMAIC _yAxis=”StAsdCty” _xAxis=”Estimate”;
FORMAT Estimate Pixel_Acres Thous. StAsdCty $SAC_Fips.;
WHERE GeoLevel=’Cy’ AND ssYYmmmv=”&SSyyMMMv” and
CatName04=”Swht’;
RUN;
```

Calling ODS tagsets starts editing column 2.

Inserts the data into the bar chart/scatterplots template created in previous step and sizes it appropriately



SAS GLOBAL FORUM

Step 5: Assembling HTML Columns

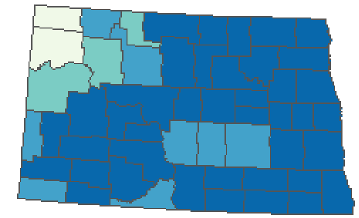
- Column 3 – 3 maps, 1 legend

```
ods tagsets.htmlpanel event=column_panel(finish);
ods tagsets.htmlpanel event=column_panel(start);
```

Starts editing Column 3

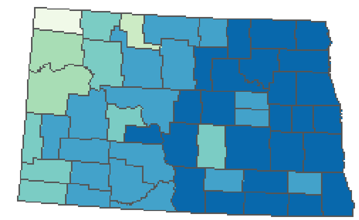
```
TITLE "Accuracy, Producer";
PROC GMAP …
```

Choropleth Map of Producer Accuacy



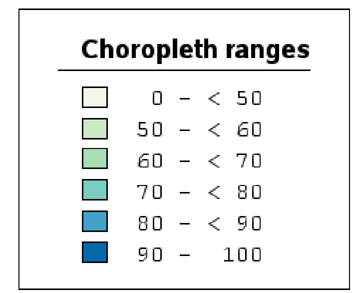
```
TITLE "Accuracy, User";
PROC GMAP …
```

Choropleth Map of User Accuacy



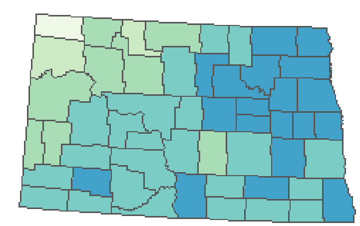
```
Data My_anno …
Goptions xpixels=240 ypixels=200;
TITLE;
PROC GANNO ANNOTATE=My_anno;
```

Legend – created manually, stored as picture




```
Title "Confidence, Mean";
PROC GMAP…
```

Choropleth Map of Confidence



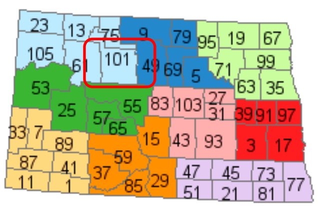
```
ods tagsets.htmlpanel event=column_panel(finish);
ods tagsets.htmlpanel event=row_panel(finish);
ods _all_ close;
ODS LISTING;
```

Closes ODS

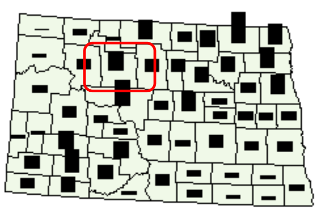
SAS GLOBAL FORUM

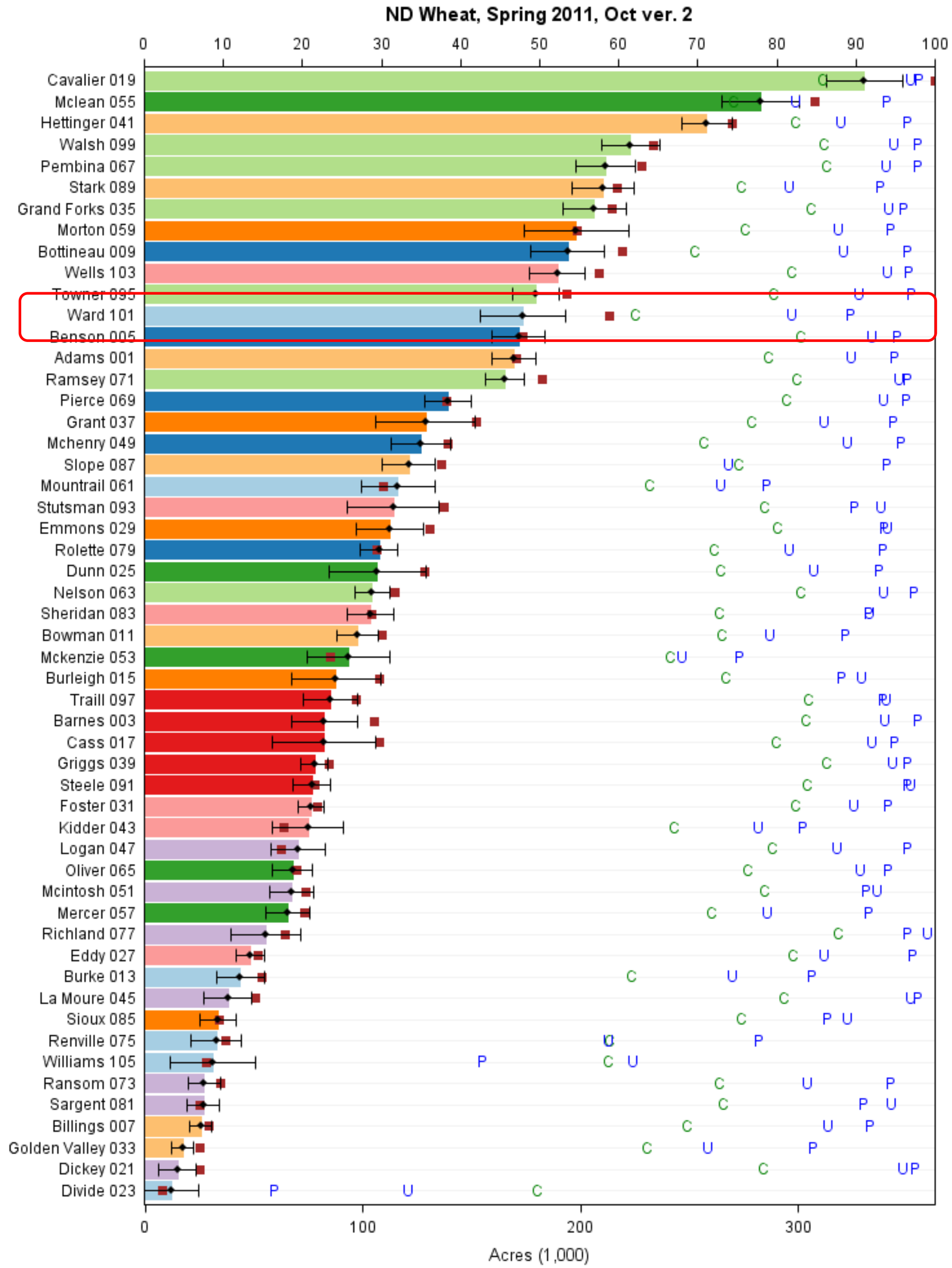
Using the Accuracy Assessment Dashboard

County FIPS and Districts



Pixels merged





County estimate bars

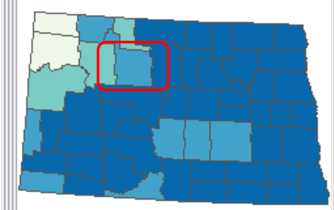
Scale, percent (top)

- P Accuracy, Producer
- U Accuracy, User
- C Confidence, mean

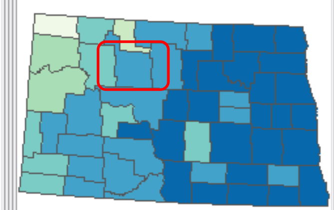
Scale, acres (bottom)

- CDL estimate
- |—•—| 2 Standard errors
- ■ Pixel acres
- ▲ FSA physical
- ★ Stat. final

Accuracy, Producer



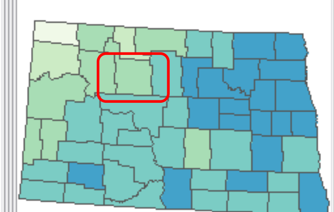
Accuracy, User



Choropleth ranges

Range
0 – < 50
50 – < 60
60 – < 70
70 – < 80
80 – < 90
90 – 100

Confidence, Mean



*All data displayed is for display purposes only.

SAS GLOBAL FORUM

San Francisco, CA

April 28–May 1, 2013

Audra Zakzeski
USDA – NASS
3251 Old Lee Hwy, Room 305
Fairfax, VA 22030
(703) 877 - 8000
audra_zakzeski@nass.usda.gov

Robert Seffrin
USDA – NASS
3251 Old Lee Hwy, Room 305
Fairfax, VA 22030
(703) 877 - 8000
robert_seffrin@nass.usda.gov