

Interactive Outlier Review and Regression Analysis in Stat Studio

Bob Seffrin
Statistician

**United States Department of Agriculture
National Agricultural Statistics Service
Research and Development Division
Spatial Analysis Research Section**



NASS Overview

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

NASS - Data and Statistics - Microsoft Internet Explorer

Address: http://www.nass.usda.gov/Data_and_Statistics/index.asp

USDA United States Department of Agriculture
National Agricultural Statistics Service

The 2002 Census of Agriculture is the most comprehensive source of statistics portraying our nation's agriculture

Home About NASS Newsroom Publications Data and Statistics Census Surveys Help Contact Us

You are here: Home / Data and Statistics

Data and Statistics

Quick Stats (Agricultural Statistics Data Base)

NASS publishes U.S., state, and county level agricultural statistics for many commodities and data series. Quick Stats offers the ability to query by commodity, state(s) and year(s), providing the most up-to-date statistics including all revisions. The query dataset can be downloaded for easy use in your database or spreadsheet.

- Query our Quick Stats Data Base

Additional Crops County Resources

Maps of crops county estimates for acreage and yield are available from NASS as both CSV data files and maps.

County data from Quick Stats data is also available in pre-extracted data sets by year and by crop.

Census of Agriculture

To query Census of Agriculture data, choose from the Census years below. To view the Census publications, click here:

- Data Queries for 2002, select below:

Select a Census Query

- Data Queries for 1997, 1992, 1987

Interactive Data

NASS provides a variety of tools for interacting with our Census datasets.

Interactive Statistical Maps Interactive Census Maps for 2002 Census Highlights

Table Lens Table Lens Application for 1997 Census Data

Last modified: 12/30/05

NASS Home | USDA.gov | FEDSTATS | Economics Statistics System (ESS) | Site Map
FOIA | Accessibility Statement | Privacy Policy | Non-Discrimination Statement | Information Quality | FirstGov | White House

2001 Wildlife Damage Survey

7.7 Percent of Crop Value Lost to Deer and Geese

Maryland farmers lost \$17.2 million of corn, soybeans and wheat to deer or geese during 2001, translates to Maryland farmers losing 7.7 percent of the crop value to deer and geese. Soybeans account for the greatest economic loss, totaling \$9.1 million, 11 percent. Corn losses were \$6.6 million, 5.8 percent and wheat \$1.5 million, 5.6 percent. Deer damage resulted in losses of \$13.6 million, 6.1 percent, while geese losses were \$3.6 million, 1.6 percent.

Production losses totaled 6.0 million bushels. Corn losses were 3.2 million bushels, soybean losses are 2.2 million bushels and wheat accounted for 0.6 million bushels. Production losses to deer were 4.7 million bushels and geese 1.3 million bushels.

In terms of yield, losses to deer were most severe in Central and Western Maryland, while geese damage greater on the Eastern Shore. Corn yield losses of 9.6 bushels per acre and 7.4 bushels per acre were reported in Central and Western Maryland, respectively. The Lower Eastern Shore reported the highest soybean loss of 6.1 bushels per acre.

Sixty-two percent of farms reported deer or geese damage to one or more crops. Damage was reported on 58 percent of farms raising corn, 58 percent of farms growing soybeans and 27 percent of farms with wheat.

Maryland 2001 Crop Loss from Deer

Region	Crop	Acres Harvested	Harvested Yield (bushels)	Average Yield Loss (bushels)	Production Loss (bu)	Economic Loss (\$)
Western Maryland	Corn	5,500	124,9	7.4	40,700	83
	Soybeans	300	36.7	9.9	3,003,250	2,473
	Wheat	200	45.2	2.3	460.2	1
Central Maryland	Corn	114,200	2,624	9.9	1,201,250	2,473
	Soybeans	92,200	34.2	3.9	360,750	1,479
	Wheat	38,200	63.3	3.3	126,250	339
Southern Maryland	Corn	25,200	122.9	4.9	146,250	299
	Soybeans	43,200	39.0	3.3	142,250	594
	Wheat	16,000	57.0	0.3	14,400	16
Upper Shore	Corn	157,200	159.2	5.1	800,750	1,241
	Soybeans	212,200	39.9	2.4	186,000	2,232
	Wheat	84,800	64.0	1.1	99,250	213

NEWS RELEASE

NATIONAL AGRICULTURAL STATISTICS SERVICE
United States Department of Agriculture - Washington, DC 20250
Ag Statistics Hotline: (800) 727-9540 • www.nass.usda.gov

Contact: Ellen Dougherty, (202) 690-8122
Jeff Geuder, (202) 720-2127

USDA FORECASTS RECORD-SETTING CORN CROP FOR 2007

Washington, Aug. 10, 2007 – U.S. history in 2007, according to the U.S. Department of Agriculture's National Agricultural Statistics Service, is that the nation produced 13.1 billion bushels, 10.6 percent more than in 2006. Based on conditions as of August 10, 2007, the nation is expected to produce 13.1 billion bushels, up 3.7 bushels from last year. The average yield per acre behind the 160.4 bushels per acre is 160.4 bushels per acre for grain, up 1.1 bushels per acre from 159.3 bushels per acre in 2006. The average yield per acre for corn is 159.3 bushels per acre, up 1.1 bushels per acre from 158.2 bushels per acre in 2006. The average yield per acre for soybeans is 34.2 bushels per acre, up 0.3 bushels per acre from 33.9 bushels per acre in 2006. The average yield per acre for wheat is 63.3 bushels per acre, up 0.3 bushels per acre from 63.0 bushels per acre in 2006. The average yield per acre for cotton is 49.9 bushels per acre, up 0.1 bushels per acre from 49.8 bushels per acre in 2006. The average yield per acre for rice is 10.0 bushels per acre, up 0.1 bushels per acre from 9.9 bushels per acre in 2006. The average yield per acre for sorghum is 10.0 bushels per acre, up 0.1 bushels per acre from 9.9 bushels per acre in 2006. The average yield per acre for grain sorghum is 10.0 bushels per acre, up 0.1 bushels per acre from 9.9 bushels per acre in 2006. The average yield per acre for dry beans is 10.0 bushels per acre, up 0.1 bushels per acre from 9.9 bushels per acre in 2006. The average yield per acre for dry peas is 10.0 bushels per acre, up 0.1 bushels per acre from 9.9 bushels per acre in 2006. The average yield per acre for chickpeas is 10.0 bushels per acre, up 0.1 bushels per acre from 9.9 bushels per acre in 2006. The average yield per acre for lentils is 10.0 bushels per acre, up 0.1 bushels per acre from 9.9 bushels per acre in 2006. The average yield per acre for mung beans is 10.0 bushels per acre, up 0.1 bushels per acre from 9.9 bushels per acre in 2006. The average yield per acre for cowpeas is 10.0 bushels per acre, up 0.1 bushels per acre from 9.9 bushels per acre in 2006. The average yield per acre for soybeans is 34.2 bushels per acre, up 0.3 bushels per acre from 33.9 bushels per acre in 2006. The average yield per acre for wheat is 63.3 bushels per acre, up 0.3 bushels per acre from 63.0 bushels per acre in 2006. The average yield per acre for corn is 159.3 bushels per acre, up 1.1 bushels per acre from 158.2 bushels per acre in 2006. The average yield per acre for rice is 10.0 bushels per acre, up 0.1 bushels per acre from 9.9 bushels per acre in 2006. The average yield per acre for sorghum is 10.0 bushels per acre, up 0.1 bushels per acre from 9.9 bushels per acre in 2006. The average yield per acre for grain sorghum is 10.0 bushels per acre, up 0.1 bushels per acre from 9.9 bushels per acre in 2006. The average yield per acre for dry beans is 10.0 bushels per acre, up 0.1 bushels per acre from 9.9 bushels per acre in 2006. The average yield per acre for dry peas is 10.0 bushels per acre, up 0.1 bushels per acre from 9.9 bushels per acre in 2006. The average yield per acre for chickpeas is 10.0 bushels per acre, up 0.1 bushels per acre from 9.9 bushels per acre in 2006. The average yield per acre for lentils is 10.0 bushels per acre, up 0.1 bushels per acre from 9.9 bushels per acre in 2006. The average yield per acre for mung beans is 10.0 bushels per acre, up 0.1 bushels per acre from 9.9 bushels per acre in 2006. The average yield per acre for cowpeas is 10.0 bushels per acre, up 0.1 bushels per acre from 9.9 bushels per acre in 2006.

WISCONSIN AGRICULTURAL STATISTICS SERVICE

P.O. Box 8034 Madison, WI 53708-8034
In cooperation with WI Department of Agriculture, Trade and Consumer Protection

2002 Dairy Producer Opinion Survey

November 2002

Wisconsin Milk Production to Recover

Milk production is expected to increase in Wisconsin during the next few years according to a survey conducted by the Wisconsin Agricultural Statistics Service. This statewide survey of producers asked for their plans with the assumption that milk prices for the next five years will be at the same level as the past five years. The survey was conducted during May and June 2002.

Based on the survey, 60 percent of producers expect to keep the same herd size, 20 percent plan to increase herd size, and 20 percent intend to discontinue milking by 2007. Actual results will depend on future milk prices, input prices, financing availability, crop yields, and other factors.

The number of herds projected for 2007 shows that the diversity of small to large herds will continue. The most prevalent herd size will remain at 50 to 99 cows.

<http://www.nass.usda.gov/0800> - 2002 Census of Agriculture - SVG Interactive Mapping - United States - Microsoft Internet Explorer

National Agricultural Statistics Service

2002 Census of Agriculture

United States | All data items are from Chapter 2 - Table 1. Area Summary Highlights: 2002 Selected crops harvested - Land in orchards (acres)

State: United States - County Level | Data Item: Selected crops harvested - Land in orchards (acres)

United States Total: 5,330,439

State Total:

County Total:

Download data as CSV | XML | PDF

Help | Print | Return to

Legend

Scale: National | Zero or Data Withheld

(Changes the data range based on National or State level)

Comparisons: 6 | 20,001 to 40,000

Color: Green | 100,001 >=

Source: USDA-NASS 2002 Census of Agriculture © USDA-NASS 2005-2006

Navigate: Mouse-over a specific state/county to view the state/county level data. Right click to zoom (option-click for MAC users). Hold the Alt key and click+drag to pan. For additional assistance with this application, [click here to view the support page.](#)

All Milk Price, Wisconsin Annual Average, 1985 - 2002 1/2

Wisconsin Dairy Herds by Herd Size

Milk cow herd size	May 2002 herds	May 2007 herds (projected) 1/2	Change 2007/2002
1 - 29	2,800	1,440	-45
30 - 49	4,700	3,440	-27
50 - 99	7,400	5,600	-24
100 - 199	1,900	2,080	+10
200 - 499	700	600	+29
500+	200	440	+120
Total	17,500	15,900	-20

1/2The May 2007 projection is based on farmers' opinions May-June 2002, with the assumption that milk prices for the next five years will be at the same level as the past five years.

Percent of Herds by Size Group 2007 Projection

Wisconsin Dairy Farmer Plans for May 2007 1/2 by Herd Size

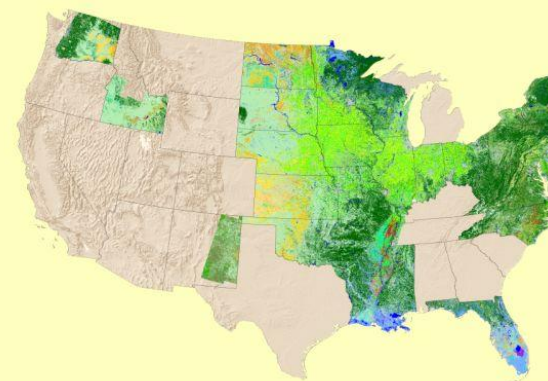
Herds	Keep same herd size	Increase herd size	Discontinue milking
2,800	47	17	36
4,700	71	9	20
7,400	65	19	16
1,900	53	37	10
700	35	59	8
200	22	78	0
17,500	62	29	20

1/2The May 2007 projection is based on farmers' opinions May-June 2002, with the assumption that milk prices for the next five years will be at the same level as the past five years.

Research and Development Division

Geospatial Information Branch Spatial Analysis Research Section

USDA
Cropland Data Layers
The First Decade: 1997 - 2006



NASS - Research and Science - Windows Internet Explorer
 http://www.nass.usda.gov/Research_and_Science/index.asp
 USDA United States Department of Agriculture
 National Agricultural Statistics Service

Search NASS
 All NASS
 Advanced Search
 Search Tips
 Browse NASS by Subject
 Crops and Plants
 Demographics
 Economics
 Environmental
 Livestock and Animals
 Charts and Maps
Research and Science
 Education and Outreach
 Statistics by State
 Select a State

You are here: Home / Research and Science
Research and Science

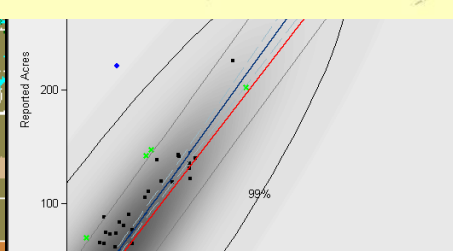
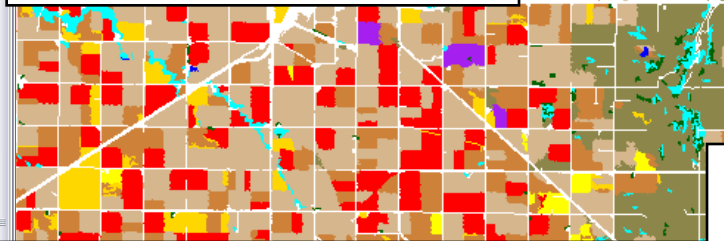
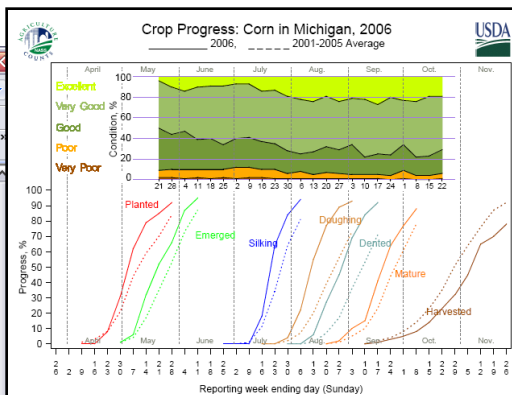
Spatial Data
 Vegetation Condition Images
 Cropland Data Layer
 Image Gallery (2003) available for these states: Arkansas, Illinois, Indiana, Iowa, N. Dakota, Mississippi, Missouri, Nebraska, Wisconsin
 Land Use Strata for Selected States

Census of Agriculture
 2002 Census Map Gallery
 2002 Maps: Gallery | Star Tree | List
 Interact with Data (1997)
 "Linked Micromap" Plots (1997):
 Corn | Cotton | Hay | Soybeans | Wheat

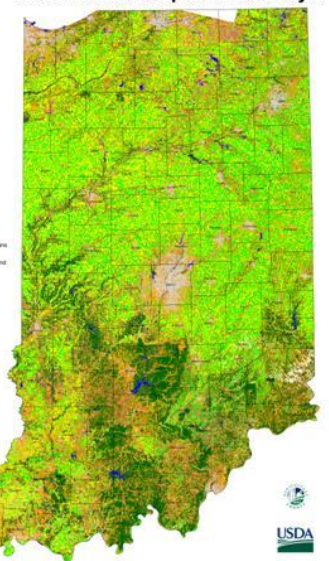
Animated Maps
 Crop Acreage
 Vegetation Condition

Also See
 Research Fellow and Associate Program
 Seasonal Summary of Crop Progress and Condition
 Remotely Sensed Data
 Crop Yield
 Future Vision

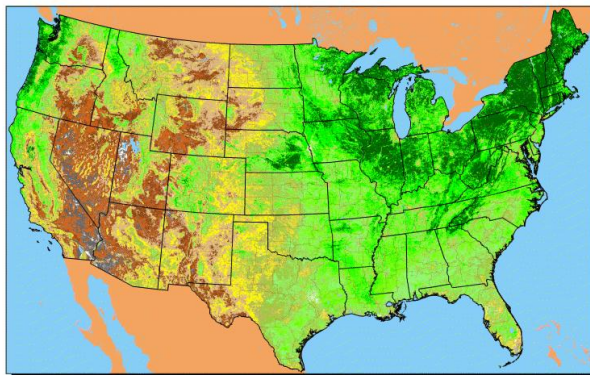
Media Help
 To view animated map files you must have Quicktime installed on your computer.



2006 Indiana Cropland Data Layer

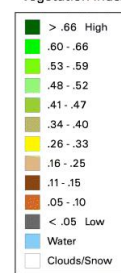


Conterminous U.S. Vegetation Condition - 2007
 Period 33 (7/31 - 8/13)



No Water Vapor Correction Applied

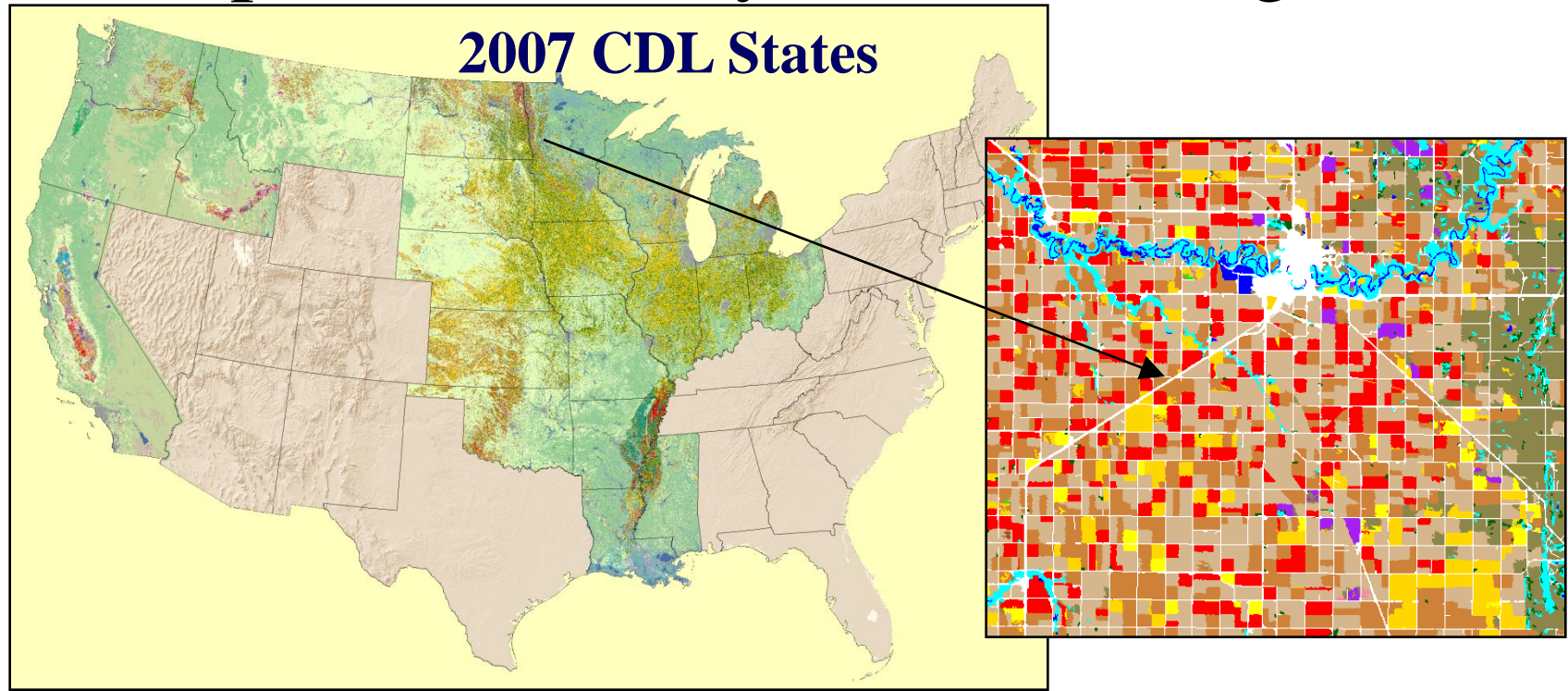
Vegetation Index



Agricultural Statistics Districts
 1:15,000,000
 Digital Imagery: NOAA-17 AVHRR
 Reanalysis: NCEP/NCAR
 Composite Imagery: USDA-ERS Data Center
 Database: NASS, ARS, USDA
 For Additional Images, Please See:
 The National Agricultural Statistics Service
 Geospatial Information Research
 http://www.nass.usda.gov/research/units/

available
 ld
 modified: 10/04/07
 Map | USA.gov | White House
 Information Quality | Guidance Documents

Cropland Data Layer (CDL) Program



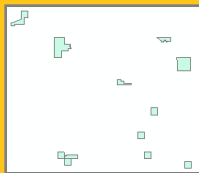
- State specific land cover classifications emphasizing row crop agriculture
 - Some regions done annually (Corn Belt, The Delta)
 - Others “one-and-done” (California, Northwest)
- Within NASS, CDL used to
 - Increase precision on survey derived acreage estimates
 - Improve county level acreage estimates

Cropland Data Layer and Acreage Estimation Processing Flow

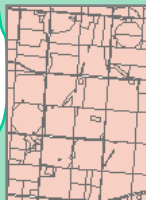
Input Vector Data

Input Raster Data

NASS JAS segments



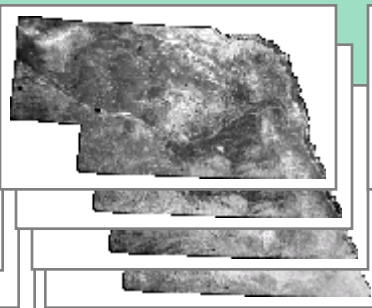
FSA CLU



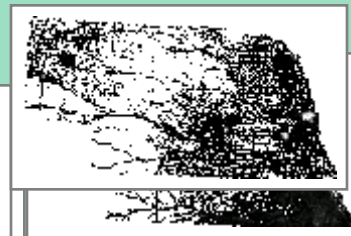
IRS Resourcesat-1 raw AWiFS summer time series



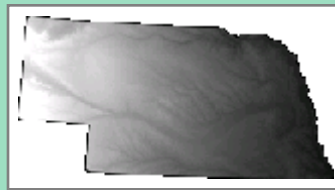
NASA Terra MODIS 16-day NDVI prior fall and summer time series



USGS NLCD circa 2001 Impervious & Canopy



USGS NED Elevation



Tabular Data

JAS eData

STATE	YEAR	STATE	YEAR	STATE	YEAR
01	2001	01	2002	01	2003
02	2001	02	2002	02	2003
03	2001	03	2002	03	2003
04	2001	04	2002	04	2003
05	2001	05	2002	05	2003
06	2001	06	2002	06	2003
07	2001	07	2002	07	2003
08	2001	08	2002	08	2003
09	2001	09	2002	09	2003
10	2001	10	2002	10	2003
11	2001	11	2002	11	2003
12	2001	12	2002	12	2003
13	2001	13	2002	13	2003
14	2001	14	2002	14	2003
15	2001	15	2002	15	2003
16	2001	16	2002	16	2003
17	2001	17	2002	17	2003
18	2001	18	2002	18	2003
19	2001	19	2002	19	2003
20	2001	20	2002	20	2003
21	2001	21	2002	21	2003
22	2001	22	2002	22	2003
23	2001	23	2002	23	2003
24	2001	24	2002	24	2003
25	2001	25	2002	25	2003
26	2001	26	2002	26	2003
27	2001	27	2002	27	2003
28	2001	28	2002	28	2003
29	2001	29	2002	29	2003
30	2001	30	2002	30	2003
31	2001	31	2002	31	2003
32	2001	32	2002	32	2003
33	2001	33	2002	33	2003
34	2001	34	2002	34	2003
35	2001	35	2002	35	2003
36	2001	36	2002	36	2003
37	2001	37	2002	37	2003
38	2001	38	2002	38	2003
39	2001	39	2002	39	2003
40	2001	40	2002	40	2003
41	2001	41	2002	41	2003
42	2001	42	2002	42	2003
43	2001	43	2002	43	2003
44	2001	44	2002	44	2003
45	2001	45	2002	45	2003
46	2001	46	2002	46	2003
47	2001	47	2002	47	2003
48	2001	48	2002	48	2003
49	2001	49	2002	49	2003
50	2001	50	2002	50	2003

FSA 578

STATE	YEAR	STATE	YEAR	STATE	YEAR
01	2001	01	2002	01	2003
02	2001	02	2002	02	2003
03	2001	03	2002	03	2003
04	2001	04	2002	04	2003
05	2001	05	2002	05	2003
06	2001	06	2002	06	2003
07	2001	07	2002	07	2003
08	2001	08	2002	08	2003
09	2001	09	2002	09	2003
10	2001	10	2002	10	2003
11	2001	11	2002	11	2003
12	2001	12	2002	12	2003
13	2001	13	2002	13	2003
14	2001	14	2002	14	2003
15	2001	15	2002	15	2003
16	2001	16	2002	16	2003
17	2001	17	2002	17	2003
18	2001	18	2002	18	2003
19	2001	19	2002	19	2003
20	2001	20	2002	20	2003
21	2001	21	2002	21	2003
22	2001	22	2002	22	2003
23	2001	23	2002	23	2003
24	2001	24	2002	24	2003
25	2001	25	2002	25	2003
26	2001	26	2002	26	2003
27	2001	27	2002	27	2003
28	2001	28	2002	28	2003
29	2001	29	2002	29	2003
30	2001	30	2002	30	2003
31	2001	31	2002	31	2003
32	2001	32	2002	32	2003
33	2001	33	2002	33	2003
34	2001	34	2002	34	2003
35	2001	35	2002	35	2003
36	2001	36	2002	36	2003
37	2001	37	2002	37	2003
38	2001	38	2002	38	2003
39	2001	39	2002	39	2003
40	2001	40	2002	40	2003
41	2001	41	2002	41	2003
42	2001	42	2002	42	2003
43	2001	43	2002	43	2003
44	2001	44	2002	44	2003
45	2001	45	2002	45	2003
46	2001	46	2002	46	2003
47	2001	47	2002	47	2003
48	2001	48	2002	48	2003
49	2001	49	2002	49	2003
50	2001	50	2002	50	2003

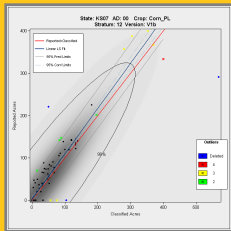
All Input layers gridded to common cell size, map projection and areal extent



Extract JAS intersecting pixels



Customized for acreage estimation



Pixel count v. reported acreage

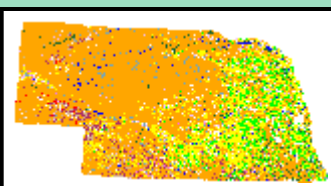
Link and assess data sets



Ground truth

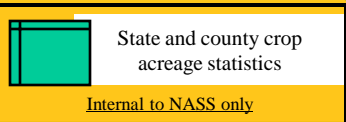


Manages and visualizes datasets



-- Cropland Data Layer --

datagateway.nrcs.usda.gov



State and county crop acreage statistics

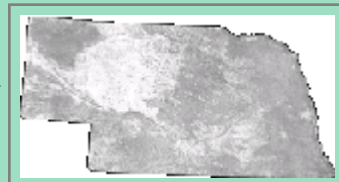
Internal to NASS only



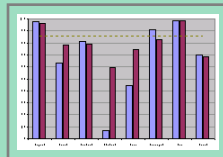
Derives decision tree-based classification rules



Generated rule set



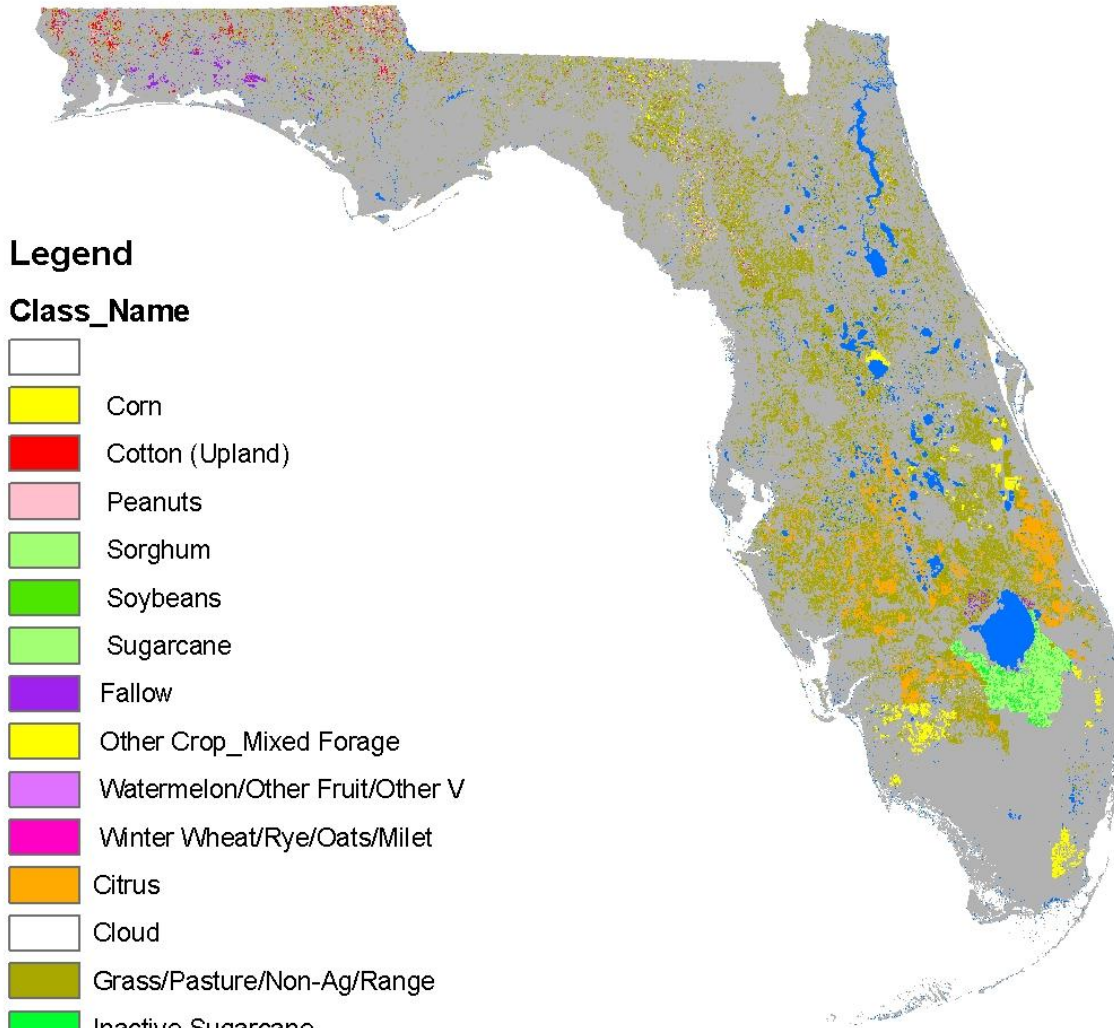
Confidence Layer



Accuracy Assessment




















Diagnostics

Cropland Data Layer, 2004

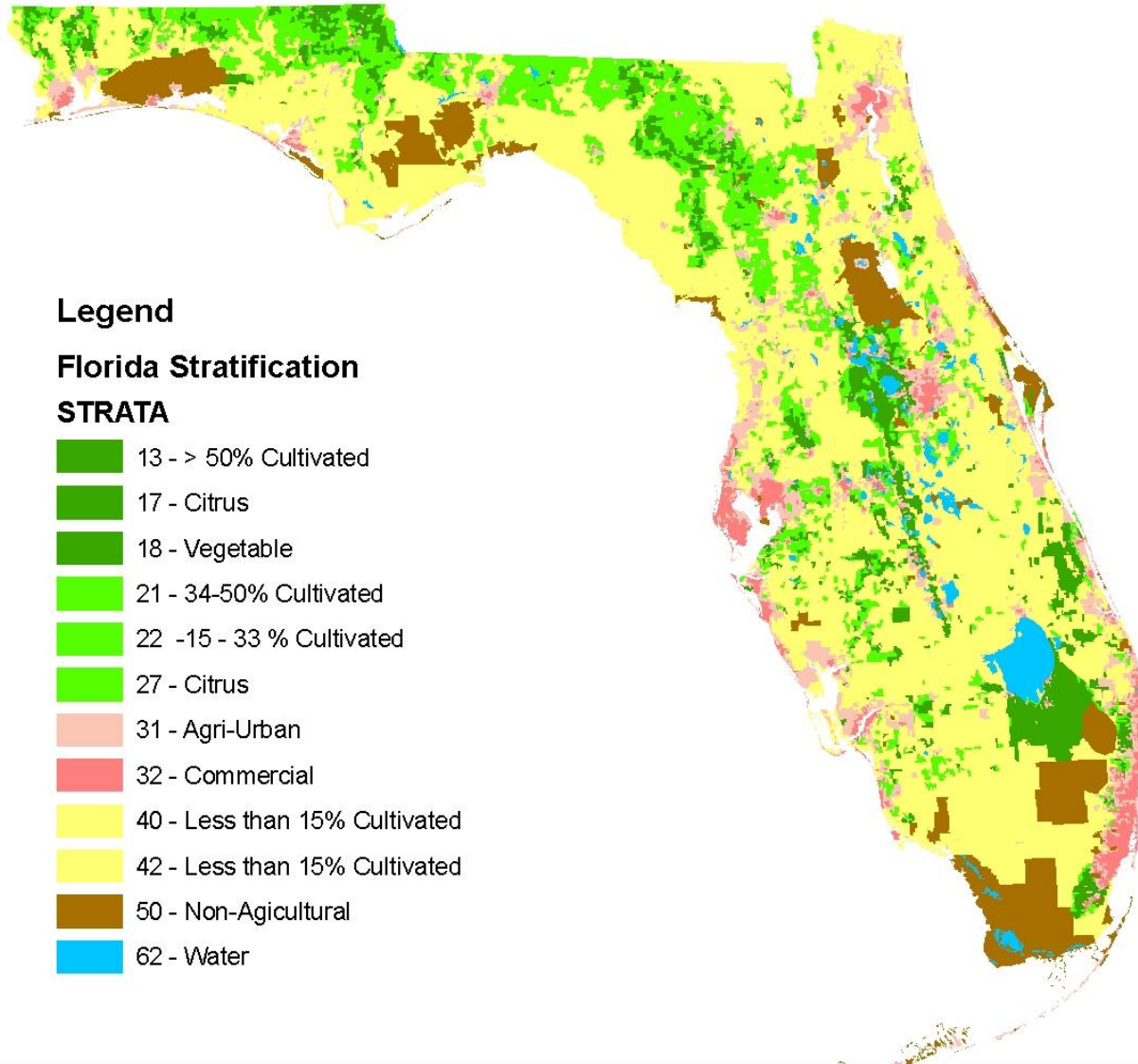


Legend

Class_Name

-  Cloud
-  Corn
-  Cotton (Upland)
-  Peanuts
-  Sorghum
-  Soybeans
-  Sugarcane
-  Fallow
-  Other Crop_Mixed Forage
-  Watermelon/Other Fruit/Other V
-  Winter Wheat/Rye/Oats/Millet
-  Citrus
-  Cloud
-  Grass/Pasture/Non-Ag/Range
-  Inactive Sugarcane
-  NLCD-Unconsolidated Shore
-  NLCD_Aquatic Beds
-  NLCD_Bare Rock/Sand Clay
-  NLCD_Barren

Area Frame Stratification for Florida



JAS Questionnaire

- Enumerators account for all land usage in segment
 - Draw off field location by direct observation
 - Directly link questionnaire to segment photo

PAGE 2

SECTION D - CROPS AND LAND USE ON TRACT

17

How many acres are inside this blue tract boundary drawn on the photo (map)?

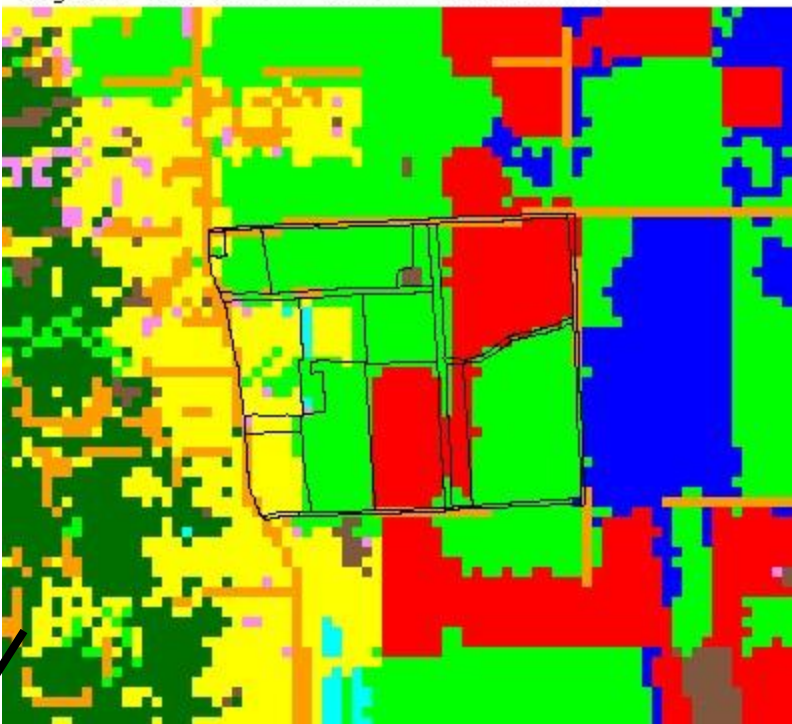
Now I would like to ask about each field inside this blue tract boundary and its use during 2000.

FIELD NUMBER		01	02	03	04	05
1.	Total acres in field	828	828	828	828	828
2.	Crop or land use. [Specify]					
3.	Occupied farmstead or dwelling	843				
4.	Waste, unoccupied dwellings, buildings and structures, roads, ditches, etc.	---	---	---	---	---
5.	Woodland	831	831	831	831	831
6.	Pasture	Permanent (not in crop rotation)	842	842	842	842
		Cropland (used only for pasture)	856	856	856	856
			857	857	857	857

PAGE 2 SECTION D - CROPS AND LAND USE ON TRACT 17

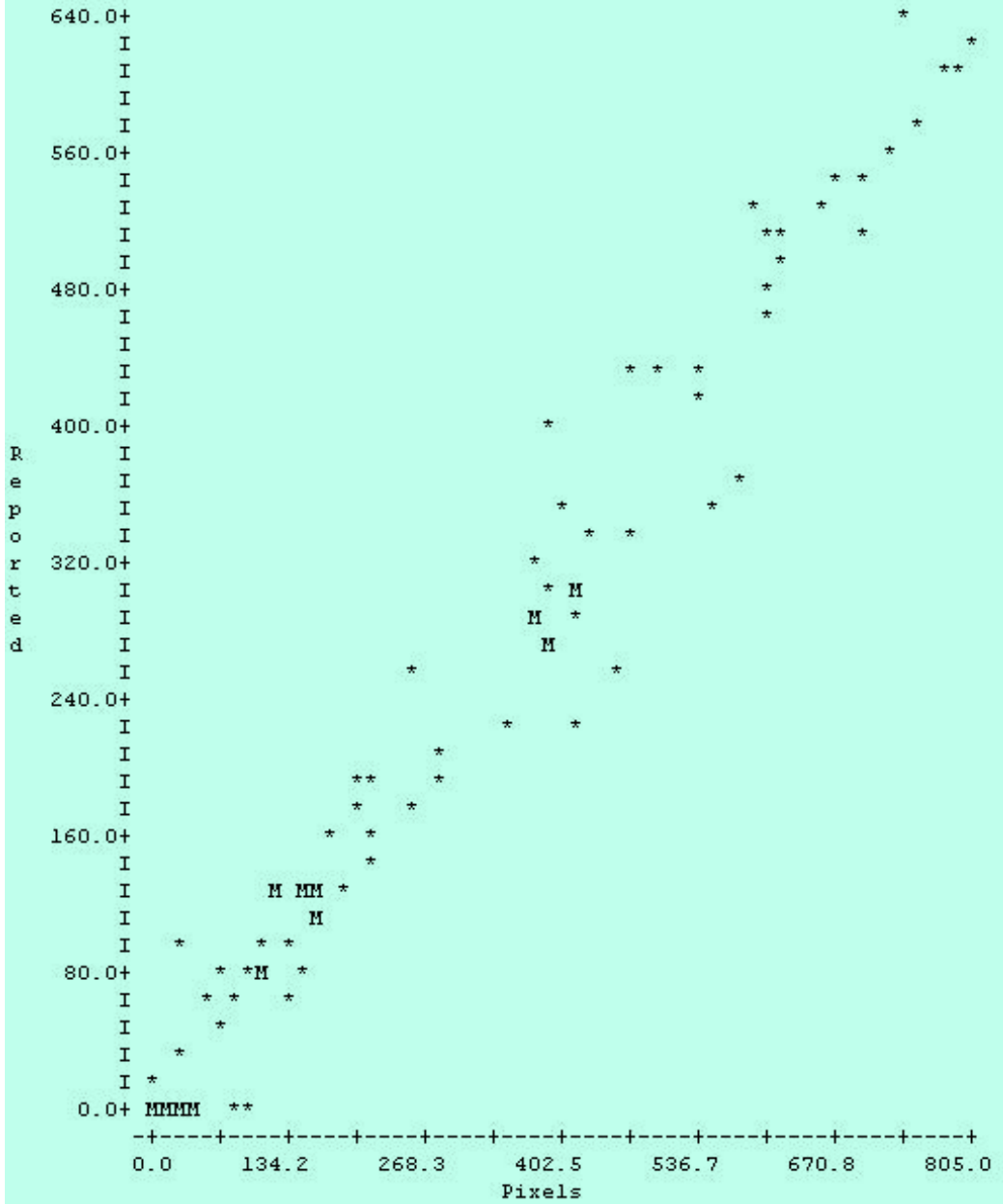
How many acres are inside this blue tract boundary drawn on the photo (map)?
 Now I would like to ask about each field inside this blue tract boundary and its use during 2000.

FIELD NUMBER	01	02	03	04	05
1. Total acres in field	828	828	828	828	828
2. Crop or land use. [Specify]					
3. Occupied farmstead or dwelling	843				
4. Waste, unoccupied dwellings, buildings and structures, roads, ditches, etc.					
5. Woodland	831	831	831	831	831
6. Pasture	842	842	842	842	842
Permanent (not in crop rotation)					
Cropland (used only for pasture)	856	856	856	856	856
8. Idle cropland - Idle all during 2000	857	857	857	857	857
9. Two crops planted in this field or two uses of the same crop.	Yes No	Yes No	Yes No	Yes No	Yes No
[Specify second crop or use]					
Acres	844	844	844	844	844
10. Acres left to be planted	610	610	610	610	610
11. Acres irrigated and to be irrigated [If double cropped, include acreage of each crop irrigated]	620	620	620	620	620
16. Winter Wheat [Include cover crop]	Planted	540	540	540	540
17. Rye [Include cover crop] [Exclude ryegrass]	For grain or seed	Planted	547	547	547
18. Rye [Include cover crop] [Exclude ryegrass]	Planted	548	548	548	548
19. Rye [Include cover crop] [Exclude ryegrass]	For grain or seed				



Crop Type	Y	X
	Enumerated JAS Acres	Classified Pixels
Rice	227.0	273
Soybean	337.0	541

Plot of COTTON Pixels(X) Reported(Y)
Strata 11



Pre-SAS Regression output

Data Listing

Table for COTTON
Strata 11

SEGMENT	PIXELS:X	REPORT:Y	ESTIMATE	RESIDUAL	HAT	RSTUD	DFFITS	COVRAT
1	391.0	265.0	303.40	-38.40	0.01	-1.76	-0.19*	0.99
2	27.0	0.0	19.92	-19.92	0.01	-0.91	-0.07	1.01
3	0.0	0.0	-1.11	1.11	0.01	0.05	0.00	1.02
4	0.0	0.0	-1.11	1.11	0.01	0.05	0.00	1.02
5	411.0	308.0	318.97	-10.97	0.01	-0.50	-0.06	1.02
6	0.0	0.0	-1.11	1.11	0.01	0.05	0.00	1.02
7	110.0	78.0	84.56	-6.56	0.00	-0.30	-0.02	1.01
8	345.0	230.0	267.57	-37.57	0.01	-1.72	-0.17	0.99
9	0.0	0.0	-1.11	1.11	0.01	0.05	0.00	1.02
10	579.0	374.9	449.81	-74.91	0.03*	-3.54*	-0.58*	0.93*
11	280.0	205.0	216.95	-11.95	0.01	-0.54	-0.05	1.01
12	97.0	0.0	74.43	-74.43	0.00	-3.47*	-0.24*	0.91*
13	0.0	0.0	-1.11	1.11	0.01	0.05	0.00	1.02
14	0.0	0.0	-1.11	1.11	0.01	0.05	0.00	1.02
15	605.0	480.0	470.06	9.94	0.03*	0.46	0.08	1.04*
16	592.0	522.0	459.94	62.06	0.03*	2.91*	0.49*	0.96*
17	0.0	0.0	-1.11	1.11	0.01	0.05	0.00	1.02
18	0.0	0.0	-1.11	1.11	0.01	0.05	0.00	1.02

Parameter Listing

STRATA	R-SQUARE	COEFFICIENTS				STANDARD DEVIATION (STRATUM TOTAL)	C.V.
		B(0)	STD ERR	B(1)	STD ERR		
11	0.963	6.23	2.5860	0.8104	0.0110	22597.44	1.48%
21	0.991	1.25	1.5491	0.8184	0.0157	3735.68	3.38%
TOTAL						22904.13	1.40%

TABLE OF RELATIVE EFFICIENCIES DIRECT EXPANSION		
STRATA	STANDARD DEVIATION	RELATIVE EFFICIENCY
11	116398.52	26.53
21	37671.11	101.69



CDL Estimation

Project Open Year: 2008 Project Save

Month: Jun

N: State: MO - Missouri Version: v4a

Classification File: \\Acreage\MO08\Final\mo08Jun_patrick_v4a.img

Master Categories:
Grass_Resrv
Wetland_Resv
Water_Imp_St
Wildlf_Habit
Forest_Mgmt
SkipRow_Area
Fallow_Idle
Corn_HV
Sorghum_HV
Barley_HV
Wht_Durhm_HV
Wht_Sprng_HV

1a file found Tabulate Segment 1b file found Tabulate Frame

2a file found Import Seg tab. 2b file found Import Frame tab.

3 file found Build Regress Data file found Regress Seg. Summary Open SAS Log

4 Regression no file Regress Param. Summary

5 file found Estimator Selection Close

6 no file Build Sample

7 file found Accumulate State no file Open Summary

Running: SARS Estimation on June 16, 2008
By: seffro on RDWS-RSEFFRIN

Opening project MO - Missouri, 2008, Jun... ..ready.

IML Workshop / Stat Studio Steps

- > Launch
- > Generate action menu
- > Run regression
- > Display regression equation
- > Highlight outliers
- > Review outliers

Launching IML Workshop / Stat Studio

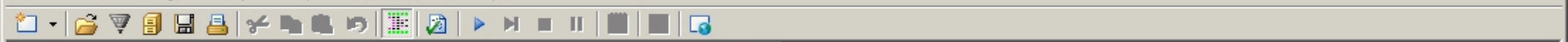
pb_Launch_Reg:

```
IF ComputerName = 'SARSBATCH2' THEN DO;  
    Program = "C:\...\statstudio.exe";  
    inProg   = "C:\...\...\ActionMenuScatterPlot_5.sx";  
END;  
ELSE DO;  
    Program = " C:\...\ IMLWorkshop.exe";  
    inProg  = "C:\...\...\ ActionMenuScatterPlot_5.iml";  
END;  
rc = OPTSETN('XWAIT', 0);  
rc = OPTSETN('XSYNC', 0);  
...
```

Launching IML Workshop / Stat Studio

```
Parameters = ' -d State=' || StatePost ||  
             ' -d Year='   || Year2   ||  
             ' -d Version=' || Version ||  
             ' -d LibDir='  || PathFinal ||  
             ' -d inFile='  || 'Regression_Build_' || Version ;  
ProgLaunch = Program || inProg || Parameters;  
rc = SYSTEM ( ProgLaunch );  
rc = OPTSETN('XWAIT', 1);  
rc = OPTSETN('XSYNC', 1);
```

RETURN;

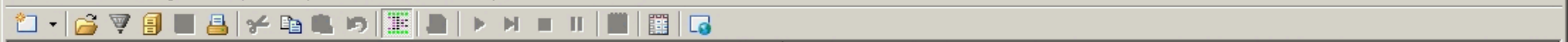


```
/*-
  Import the NBST&B.lst files across districts using
  Import_NBS-TAB.sas
*/
/*- TO DO

Add columns to DoDialogGetListItems (Strata) of seg. counts, ..
Brush scatterplot by % good ground truth
Print state, year, district, strata, cover to log/output

Program Outline:

Get command line parameters
Build variable names
Create a SAS dataset to hold deleted segments if not already ex
Sort the regression build file
```

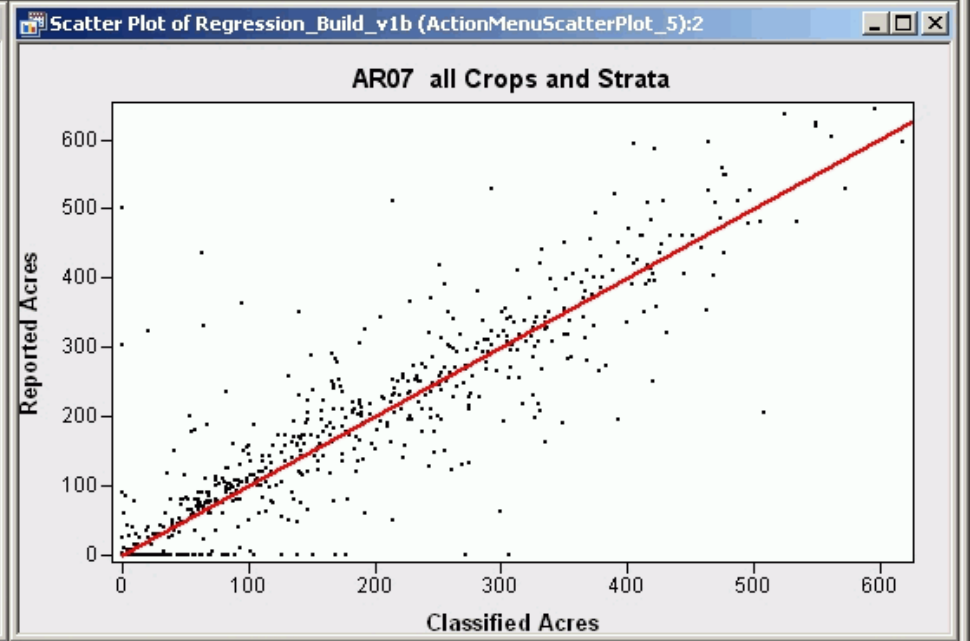
```

ActionMenuScatterPlot_5.sx
/*-
  Import the NBSTAB.lst files across districts using
  Import_NBS-TAB.sas
*/
/*- TO DO

Add columns to DoDialogGetListItems (Strata) of seg. counts, ..
Brush scatterplot by % good ground truth
Print state, year, district, strata, cover to log/output

Program Outline:

Get command line parameters
Build variable names
Create a SAS dataset to hold deleted segments if not already ex
Sort the regression build file
Create data Object, set properties
    
```



Output1

Working directory: C:\N\Estimates\Acreage\Workfiles_07_Oct\AR07

Regression_Build_v1b (ActionMenuScatterPlot_5):1

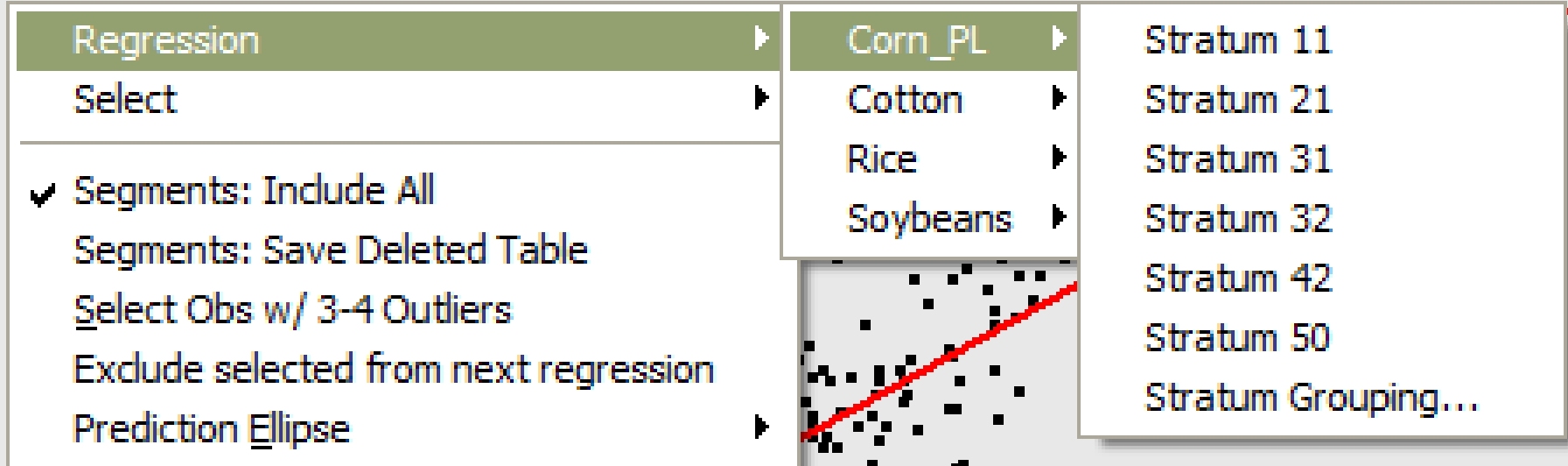
	9	CropName	Pixels	Stratum	AD	State
1368		Nom	Int	Nom	Nom	Nom
1	■ x²	Corn_PL		0 11	00	AR
2	■ x²	Corn_PL		0 11	00	AR
3	■ x²	Corn_PL		0 11	00	AR
4	■ x²	Corn_PL		0 11	00	AR
5	■ x²	Corn_PL		0 11	00	AR
6	■ x²	Corn_PL		0 11	00	AR
7	■ x²	Corn_PL		0 11	00	AR
8	■ x²	Corn_PL		0 11	00	AR
9	■ x²	Corn_PL		0 11	00	AR
10	■ x²	Corn_PL		0 11	00	AR
11	■ x²	Corn_PL		0 11	00	AR
12	■ x²	Corn_PL		0 11	00	AR
13	■ x²	Corn_PL		0 11	00	AR
14	■ x²	Corn_PL		0 11	00	AR
15	■ x²	Corn_PL		0 11	00	AR

IML Workshop / Stat Studio Steps

- > Launch
- > Generate action menu
- > Run regression
- > Display regression equation
- > Highlight outliers
- > Review outliers

Activating Action Menu

F11



The image shows a screenshot of the SAS software interface. A regression analysis menu is open, displaying the following options:

- Regression ▶
- Select ▶
- ✓ Segments: Include All
- Segments: Save Deleted Table
- Select Obs w/ 3-4 Outliers
- Exclude selected from next regression
- Prediction Ellipse ▶

The 'Corn_PL' menu is also open, showing the following options:

- Corn_PL ▶
- Cotton ▶
- Rice ▶
- Soybeans ▶

The 'Stratum' menu is also open, showing the following options:

- Stratum 11
- Stratum 21
- Stratum 31
- Stratum 32
- Stratum 42
- Stratum 50
- Stratum Grouping...

The background of the screenshot shows a scatter plot with a red regression line. The plot contains numerous data points, with a red line indicating the fitted regression model.

Generating the Action Menu

Create text like:

```
'Regression\nState AR\nDistrict 01\nCorn\nStrata 11'J
```

```
uCrop    = UNIQUE( allCrop );      { also strata, state, district }  
cntCrop  = NCOL( uCrop );         { also strata, state, district }
```

```
TotMenuItems = cntState*cntDist*cntCrop*cntStrat;  
Counts      = cntState || cntDist || cntCrop || cntStrat;  
Permies     = J( TotMenuItems, NCOL(Counts), 0 );  
Row         = 0;
```

Generating the Action Menu

Build indexes of possible combinations

```
DO i1 = 1 TO Counts[1];
  DO i2 = 1 TO Counts[2];
    DO i3 = 1 TO Counts[3];
      DO i4 = 1 TO Counts[4];
        Row = Row + 1;
        IF Counts[1] = 1 THEN Permies[Row,1]=0; ELSE Permies[Row,1]=i1;
        IF Counts[2] = 1 THEN Permies[Row,2]=0; ELSE Permies[Row,2]=i2;
        IF Counts[3] = 1 THEN Permies[Row,3]=0; ELSE Permies[Row,3]=i3;
        IF Counts[4] = 1 THEN Permies[Row,4]=0; ELSE Permies[Row,4]=i4;
      END;
    END;
  END;
END;
```

Generating the Action Menu

Define the root of menu, create arrays to hold results

```
sMenuBase_Reg = "Regression\n"J;  
sCode_Reg     = 'RUN OnLSRegression;';  
sCode_Group   = 'Group...';  
LengthFiller = "                ";  
aMenuStr      = J( TotMenuItems, 1, sMenuBase_Reg + LengthFiller );
```

Generating the Action Menu

Apply the root of menu, cascading menus to Plot

```
plot.AppendActionMenuItem( aMenuStr[1], aMenuCodeSel[1] +  
                           aMenuCodeReg[1] );
```

```
/*- Append rest of menus to roots */
```

```
DO i = 2 TO TotMenuItems;
```

```
    plot.AppendActionMenuItemToGroup( aMenuStr[1],aMenuStr[i]  
                                     ,aMenuCodeSel[i]+Code_Reg );
```

```
END;
```

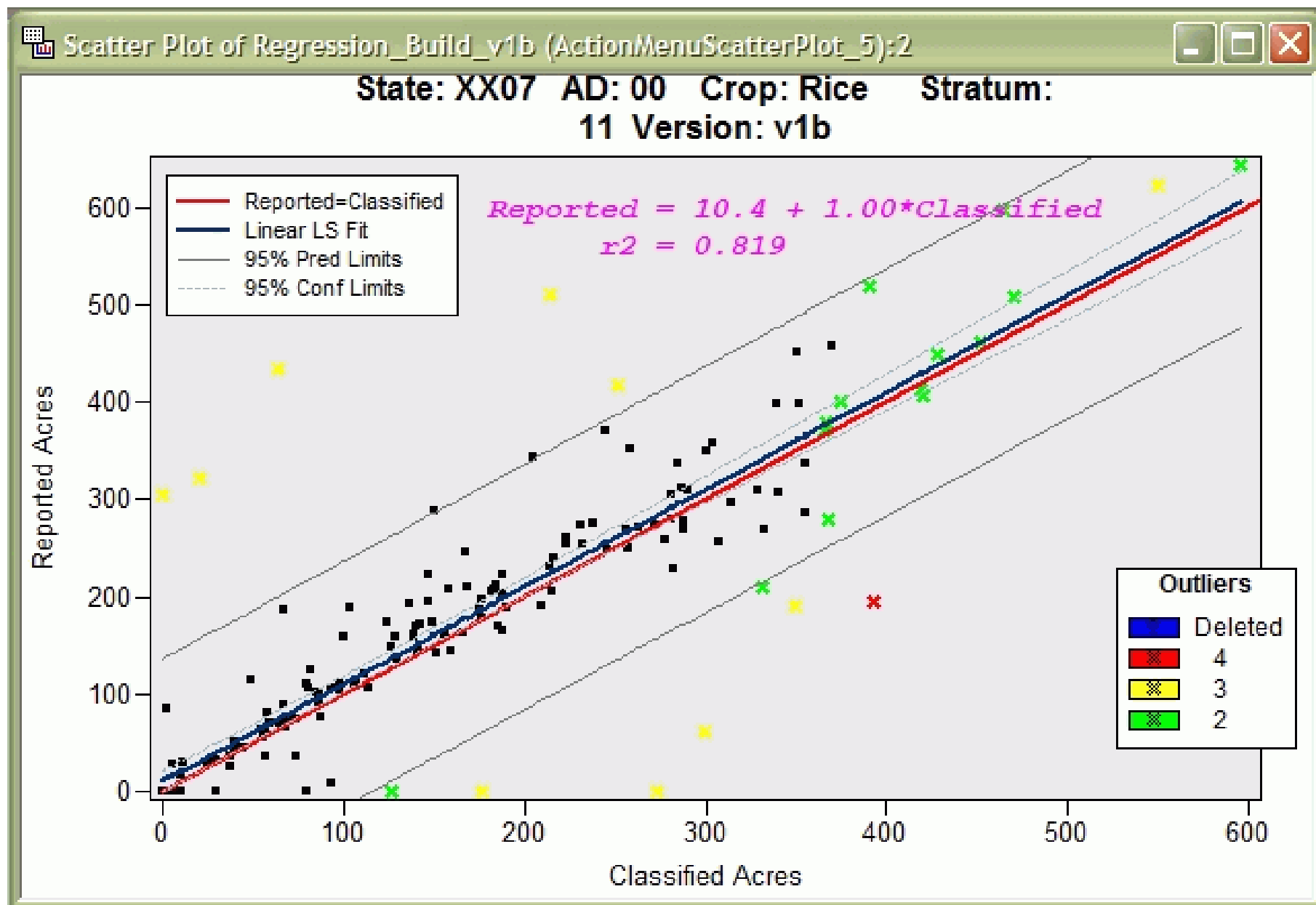
Generating the Action Menu

Create text like:

"Regression\nCorn_PL\nStrata 11"J

The image shows a screenshot of the SAS software interface. The 'Regression' menu is open, displaying several options: 'Select', 'Segments: Include All' (with a checkmark), 'Segments: Save Deleted Table', 'Select Obs w/ 3-4 Outliers', 'Exclude selected from next regression', and 'Prediction Ellipse'. The 'Corn_PL' sub-menu is also open, showing 'Cotton', 'Rice', and 'Soybeans'. The 'Strata 11' sub-menu is open, showing 'Stratum 11', 'Stratum 21', 'Stratum 31', 'Stratum 32', 'Stratum 42', 'Stratum 50', and 'Stratum Grouping...'. A red diagonal line is drawn across the plot area, which contains a scatter plot of data points.

Using the Action Menu



IML Workshop / Stat Studio Steps

- > Launch
- > Generate action menu
- > Run regression
- > Display regression equation
- > Highlight outliers
- > Review outliers

Run the Regression

SUBMIT < *...some variables to pass, SAS language will see as macro variables...> ;*

ODS OUTPUT

ANOVA = **oAnovaFits**

Fitstatistics = **oFitstatistics**

ParameterEstimates = **oParameterEstimates ;**

Run the Regression

```
proc reg data=RegIn TABLEOUT ;  
&ModelStat : model &yVarName = &RegXVarNames / cl  
  ADJRSQ AIC BIC CP EDF GMSEP JP MSE ...;  
&WhereClause  
output out = RegOut          p = &predName  
residual   = &residName      lclm = &lclmName  
uclm       = &uclmName        lcl = &lcliName  
ucl        = &ucliName        H = &Hat  
rStudent   = &rStudent      DFFITS= &DFFITS  
CovRatio   = &CovRatio;  
quit;
```

Run the Regression

PROC SQL NOPRINT;

CREATE TABLE FileStat <... *query to merge regression parameters from ODS tables...*>

ENDSUBMIT;

IML Workshop / Stat Studio Steps

- > Launch
- > Generate action menu
- > Run regression
- > Display regression equation
- > Highlight outliers
- > Review outliers

Display Regression Equation

*Reported = 10.4 + 1.00*Classified*
R2 = 0.819

Display Regression Equation

```
plot.DrawRemoveCommands("Regress Equation");
```

```
declare DataObject RegParms;
```

```
RegParms = DataObject.CreateFromServerDataSet(FileStat)
```

```
RegParms.GetVarData( "Intercept", Reg_Intercept);
```

```
RegParms.GetVarData( "Slope", Reg_Slope);
```

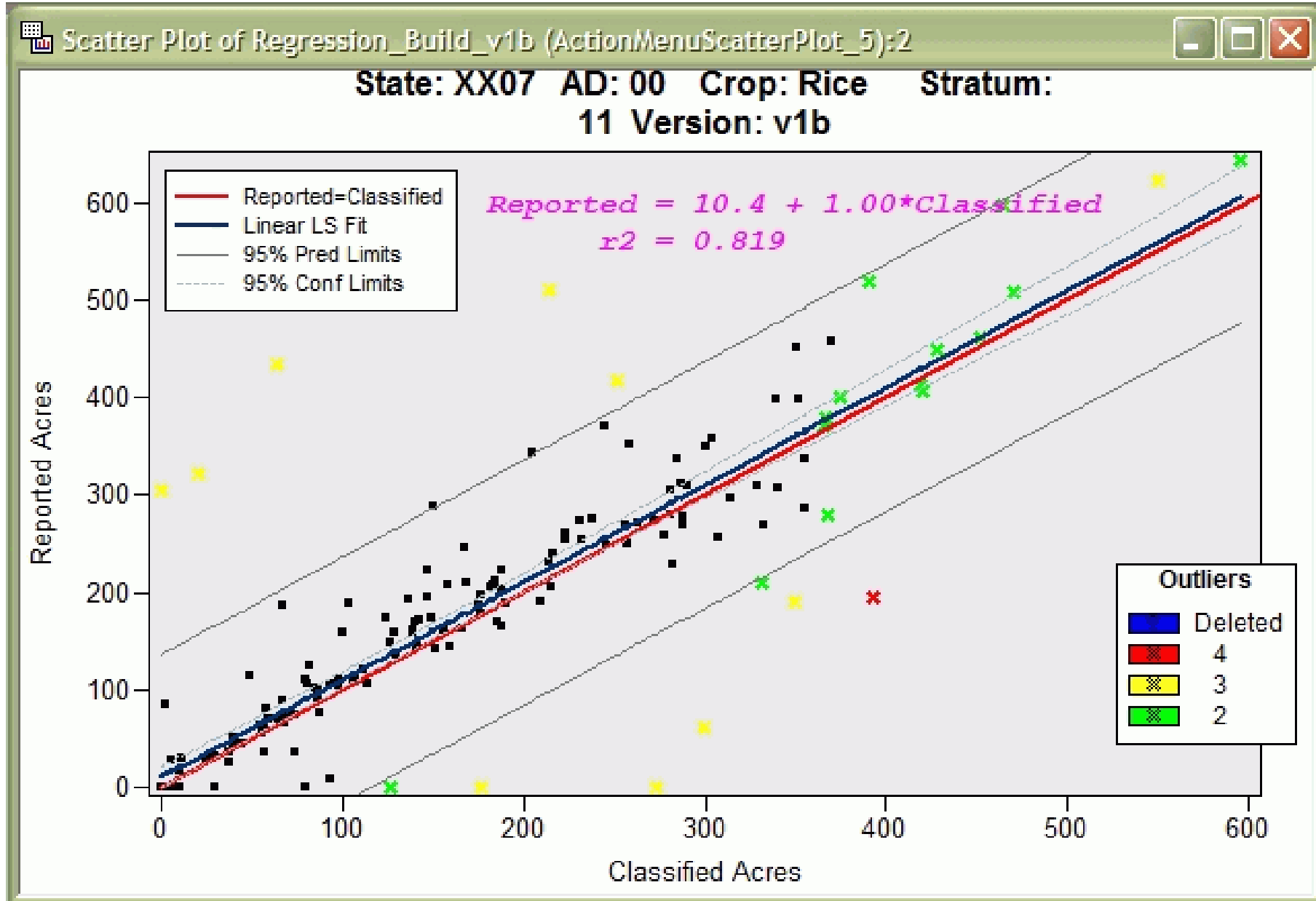
```
RegParms.GetVarData( "R_Square", Reg_r2);
```

```
Reg_Eq = CONCAT( 'Reported = '  
  , STRIP( PUTN( Reg_Intercept, '4.2' ))  
  , ' + ', STRIP( PUTN( Reg_Slope, '4.2' ))  
  , '*Classified \n\r2 = "J  
  , STRIP( PUTN( Reg_r2, '5.3' )) );
```


Display Regression Equation

```
plot.DrawBeginBlock( "Regress Equation" );  
  plot.DrawPushState();  
  plot.DrawResetState();  
    plot.DrawSetTextTypeface( "Courier New" );  
    plot.DrawSetTextStyle( STYLE_BOLDITALIC );  
    plot.DrawSetTextColor( MAGENTA );  
    plot.DrawSetTextSize( 11 );  
    plot.DrawSetTextAlignment( ALIGN_LEFT, -1 );  
    plot.DrawText( 30, 93, Reg_Eq );  
  plot.DrawPopState();  
plot.DrawEndBlock();
```

Display Regression Equation



IML Workshop / Stat Studio Steps

- > Launch
- > Generate action menu
- > Run regression
- > Display regression equation
- > Highlight outliers
- > Review outliers

Highlight Outliers

Define thresholds

```
PCR = 2;  
dobjOut.GetObsNumbersInAnalysis( AnalyCnt );  
FN      = NROW( AnalyCnt );  
HatCR   = 2*PCR/FN;  
CovCR   = 6/FN;  
DffCR   = 2*SQRT( PCR/FN );  
RstCR   = 2.0;
```

Highlight Outliers

Tally outliers

```
dobjOut.GetVarData( "Hat" , Hat );  
IF Hat[j] > HatCR THEN DO;  
    OutLier[j] = OutLier[j]+1;  
    OutIndi[j] = OutIndi[j]+1;  
END; {for all outlier indicators}  
dobj.AddVar("OutCount", "Outlier Count,1-4", OutLierTemp);  
  
Out4      = LOC(OutLierTemp=4);  
OutSel34  = LOC(OutLierTemp>2);
```

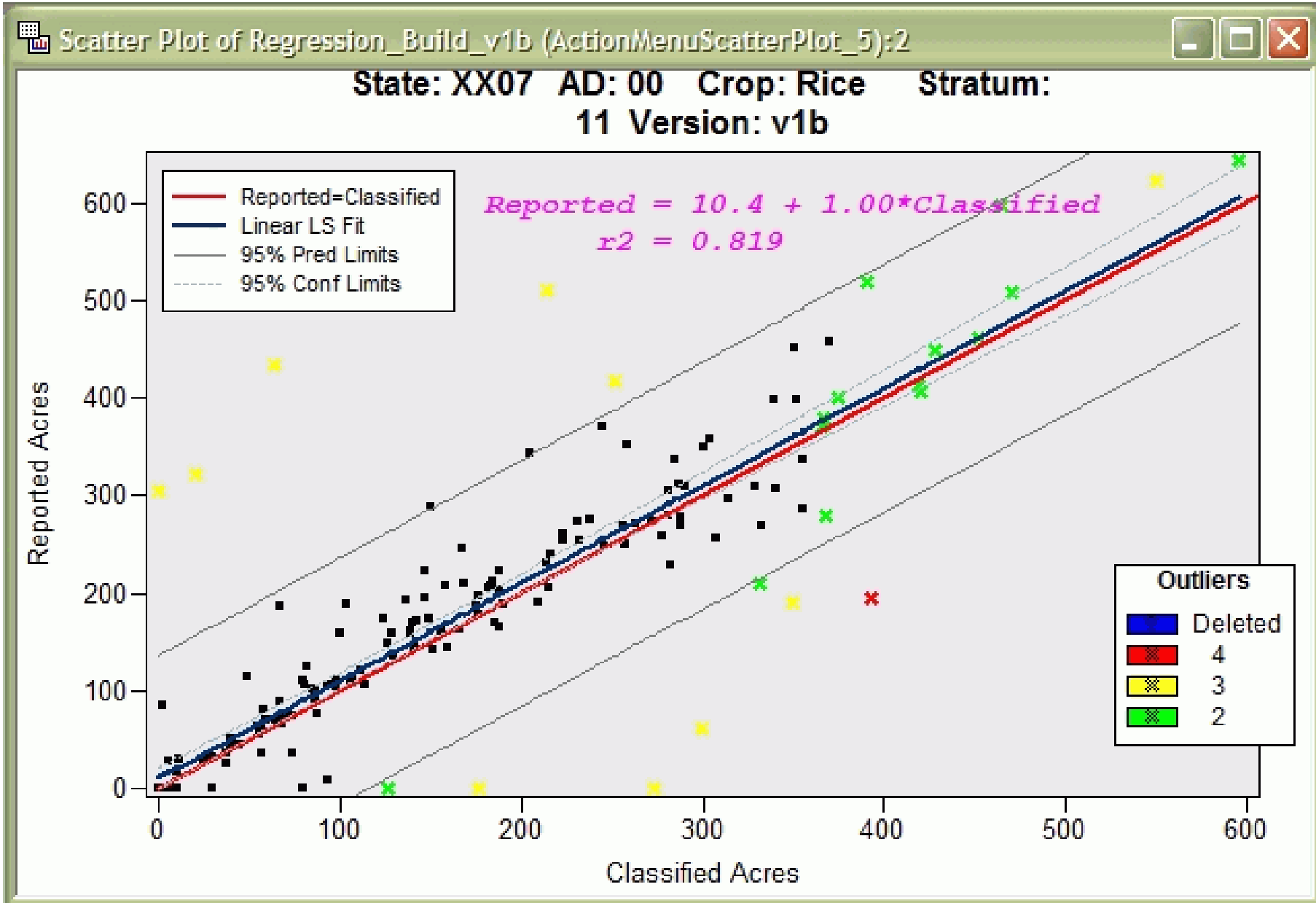
Highlight Outliers

Change marker and color

```
IF NCOL(Out234)>0 THEN  
    plot.SetMarkerShape( Out234, MARKER_X );
```

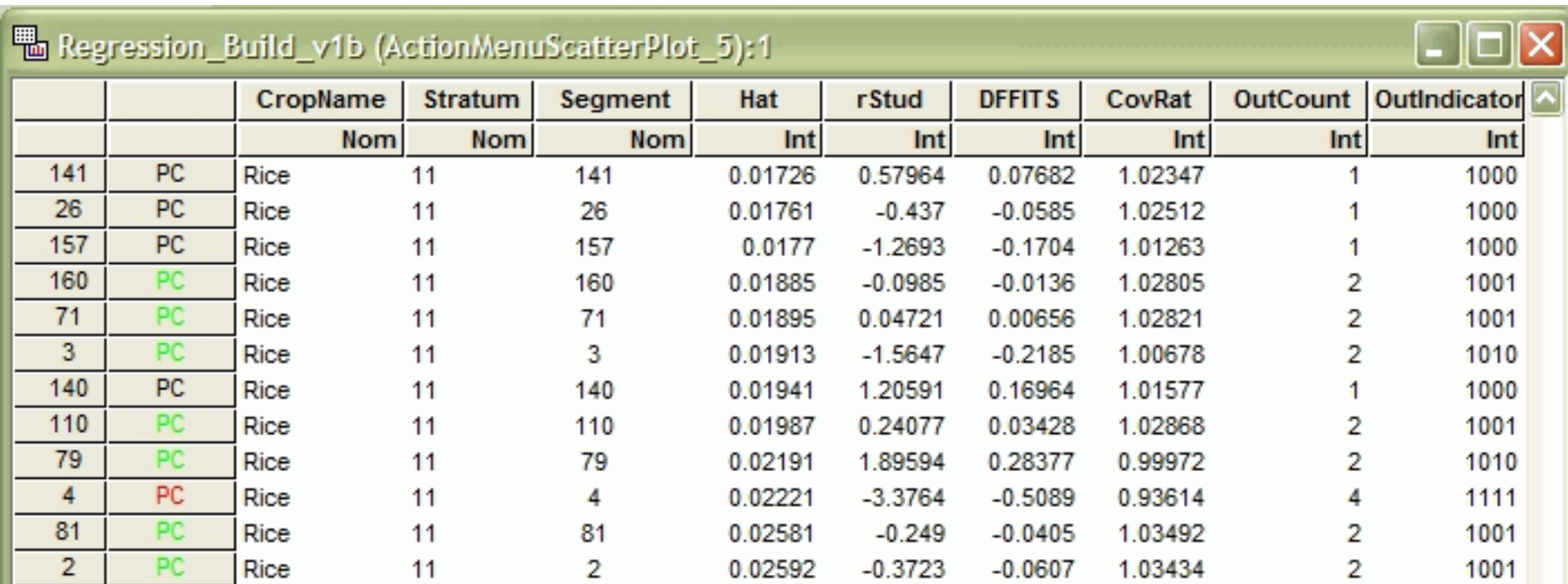
```
IF NCOL(Out4) > 0 THEN DO;  
    plot.SetMarkerColor( Out4, RED );  
END;
```

Highlight Outliers



Highlight Outliers

Table: outliers, counts, colors



Regression_Build_v1b (ActionMenuScatterPlot_5):1

		CropName	Stratum	Segment	Hat	rStud	DFFITS	CovRat	OutCount	OutIndicator
		Nom	Nom	Nom	Int	Int	Int	Int	Int	Int
141	PC	Rice	11	141	0.01726	0.57964	0.07682	1.02347	1	1000
26	PC	Rice	11	26	0.01761	-0.437	-0.0585	1.02512	1	1000
157	PC	Rice	11	157	0.0177	-1.2693	-0.1704	1.01263	1	1000
160	PC	Rice	11	160	0.01885	-0.0985	-0.0136	1.02805	2	1001
71	PC	Rice	11	71	0.01895	0.04721	0.00656	1.02821	2	1001
3	PC	Rice	11	3	0.01913	-1.5647	-0.2185	1.00678	2	1010
140	PC	Rice	11	140	0.01941	1.20591	0.16964	1.01577	1	1000
110	PC	Rice	11	110	0.01987	0.24077	0.03428	1.02868	2	1001
79	PC	Rice	11	79	0.02191	1.89594	0.28377	0.99972	2	1010
4	PC	Rice	11	4	0.02221	-3.3764	-0.5089	0.93614	4	1111
81	PC	Rice	11	81	0.02581	-0.249	-0.0405	1.03492	2	1001
2	PC	Rice	11	2	0.02592	-0.3723	-0.0607	1.03434	2	1001

IML Workshop / Stat Studio Steps

- > Launch
- > Generate action menu
- > Run regression
- > Display regression equation
- > Highlight outliers
- > Review outliers

Review Outliers

Regression

Select

✓ Segments: Include All

Segments: Save Deleted Table

Select Obs w/ 3-4 Outliers

Exclude selected from next regression

Prediction Ellipse

Corn_PL

Cotton

Rice

Soybeans

Stratum 11

Stratum 21

Stratum 31

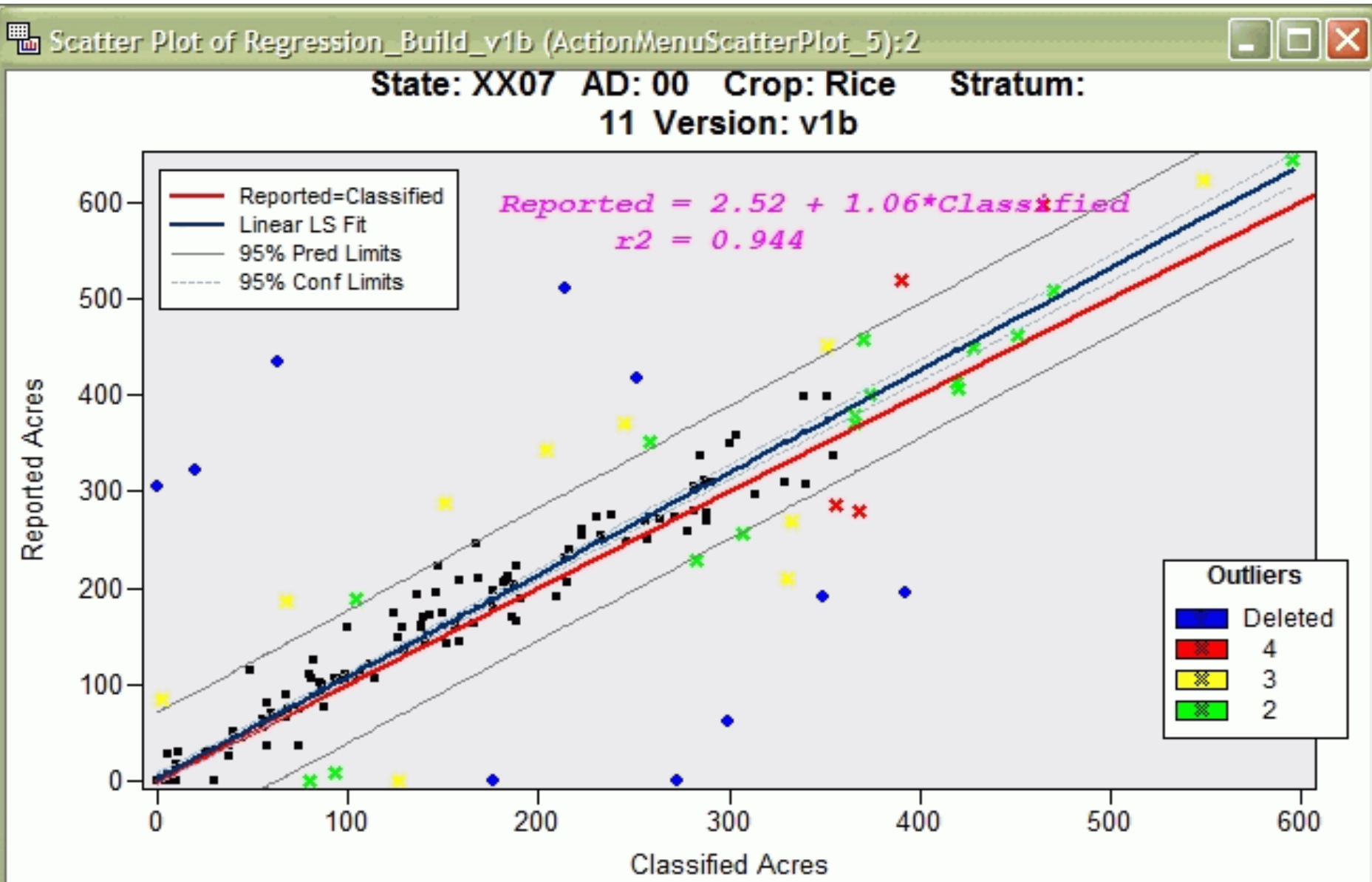
Stratum 32

Stratum 42

Stratum 50

Stratum Grouping...

Review Outliers



Stat Studio Highlights

- Interactive
- Linkage between objects
- Access to rest of SAS through SUBMIT
- Can manage graphics as named Blocks
- Power of IML
- Flexible menu

Interactive Outlier Review and Regression Analysis in Stat Studio

Bob Seffrin
Statistician

**United States Department of Agriculture
National Agricultural Statistics Service
Research and Development Division
Spatial Analysis Research Section**

