

MODIS-based Modeling of Corn and Soybean Yields in the US

David M. Johnson, Geographer
dave.johnson@nass.usdas.gov

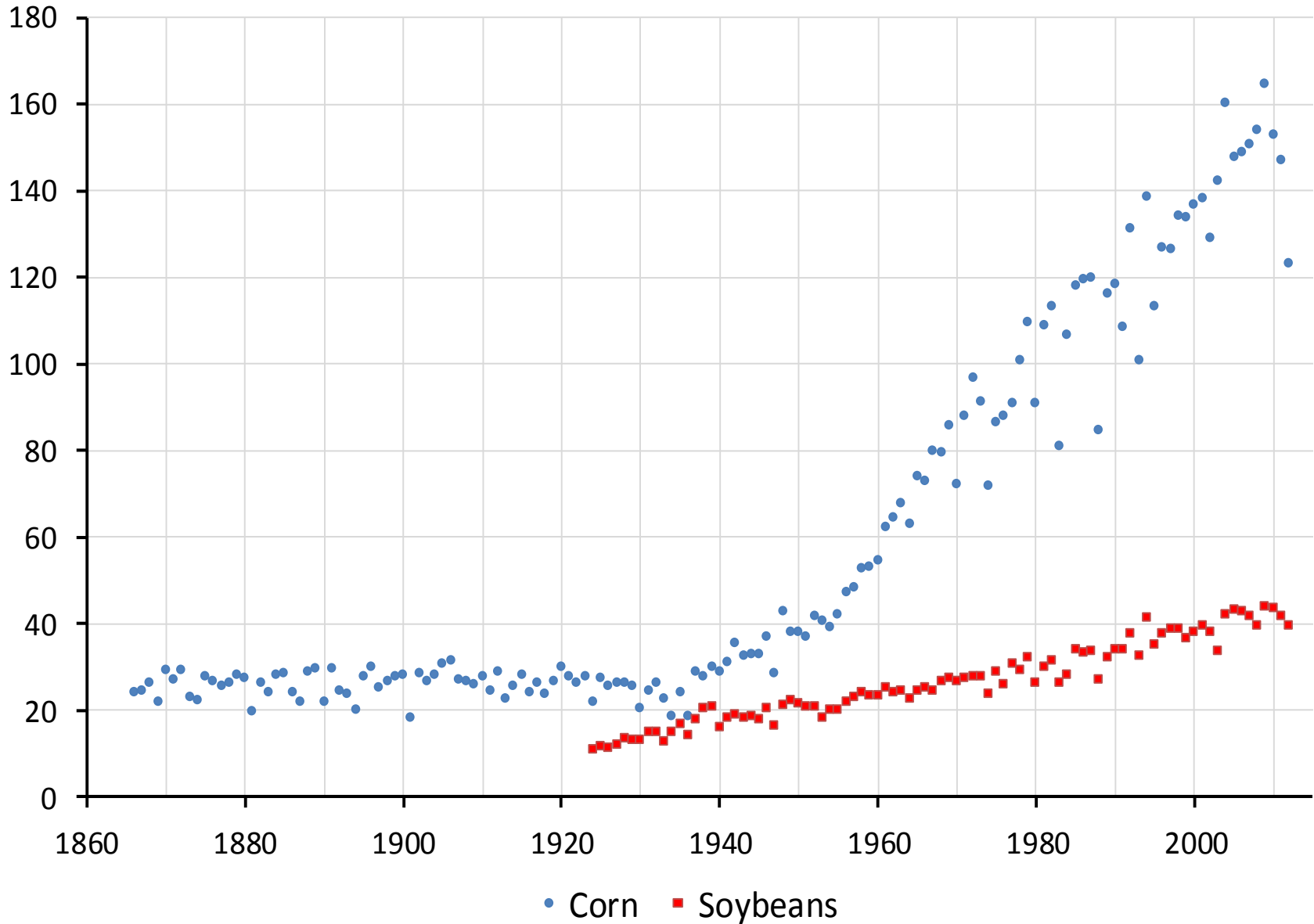
United States Department of Agriculture
National Agricultural Statistics Service
www.nass.usda.gov



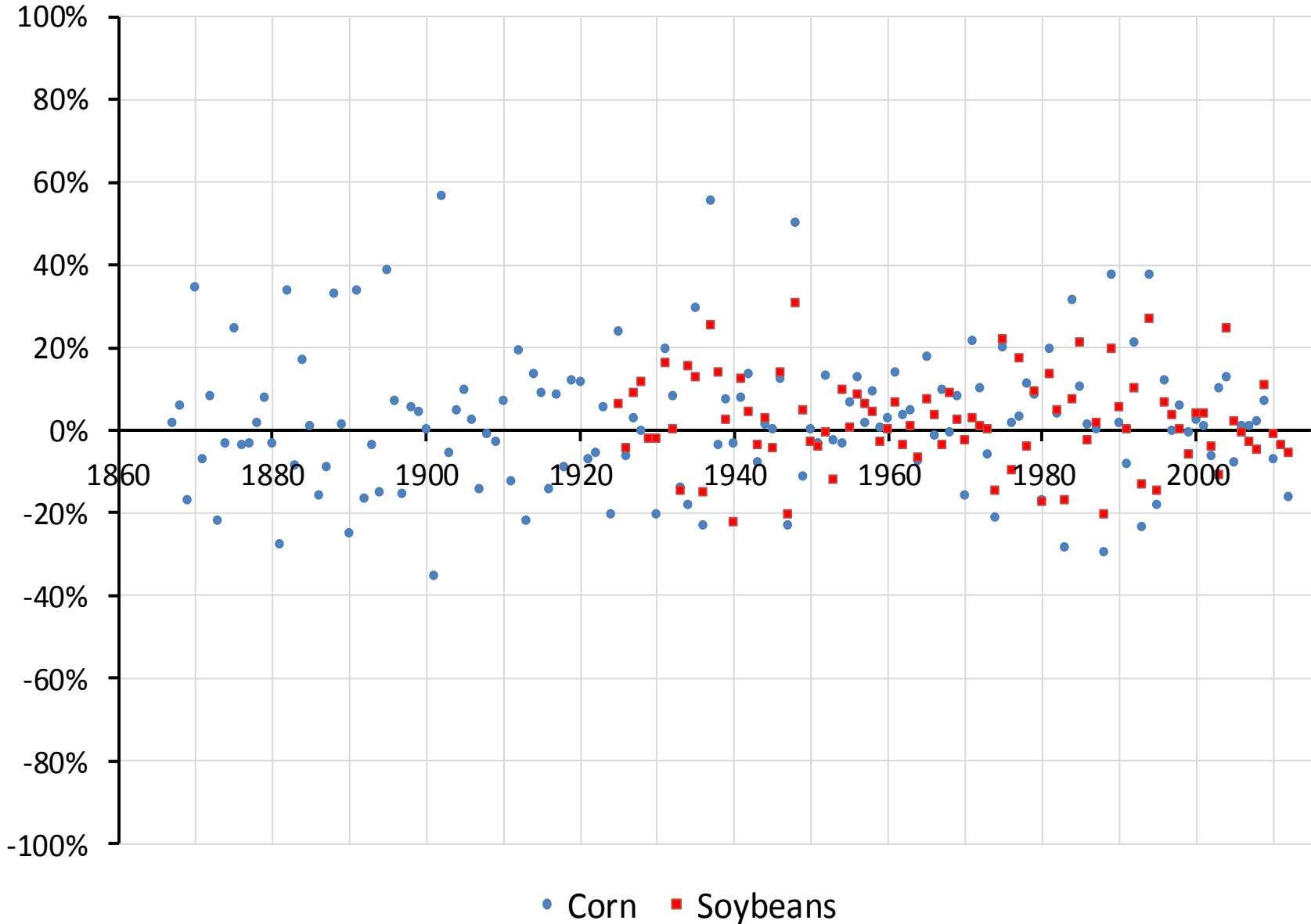
“ . . . providing timely, accurate, and useful statistics in service to U.S. agriculture.”



United States Yield (bushel/acre)



United States Yield Year to Year Change %



NASS Crop Production reports

Corn Area Planted for All Purposes and Harvested for Grain, Yield, and Production – States and United States: 2009-2011 (continued)

State	Yield per acre			Production		
	2009 (bushels)	2010 (bushels)	2011 (bushels)	2009 (1,000 bushels)	2010 (1,000 bushels)	2011 (1,000 bushels)
Alabama	108.0	116.0	114.0	27,000		
Arizona	175.0	210.0		3,500		
Arkansas	148.0	150.0	142.0	60,680		
California	180.0	195.0	185.0	28,800		
Colorado	153.0	151.0	133.0	151,470		
Connecticut	(NA)	(NA)	(NA)	(NA)		
Delaware	145.0	115.0	130.0	23,635		
Florida	100.0	105.0	100.0	3,700		
Georgia	140.0	145.0	158.0	51,800		
Idaho	180.0	180.0	185.0	14,400		
Illinois	174.0	157.0		2,053,200		
Indiana	171.0	157.0	146.0	933,660		
Iowa	182.0	165.0	172.0	2,420,600		
Kansas	155.0	125.0	107.0	598,300		
Kentucky	165.0	124.0	139.0	189,750		
Louisiana	132.0	140.0	135.0	80,520		
Maine ¹	(NA)	(NA)	(NA)	(NA)		
Maryland	145.0	106.0	109.0	61,625		
Massachusetts ¹	(NA)	(NA)	(NA)	(NA)		
Michigan	148.0	150.0	153.0	309,320		
Minnesota	174.0	177.0	156.0	1,244,100		
Mississippi	126.0	136.0	128.0	87,570		
Missouri	153.0	123.0	114.0	446,760		
Montana	152.0	135.0	130.0	3,952		
Nebraska	178.0	166.0	160.0	1,575,300		
Nevada	(NA)	(NA)	(NA)	(NA)		
New Hampshire	(NA)	(NA)	(NA)	(NA)		
New Jersey	143.0	114.0	123.0	10,010		
New Mexico	185.0	180.0	180.0	9,250		
New York	134.0	150.0	133.0	79,730		
North Carolina	117.0	91.0	84.0	93,600		
North Dakota	115.0	132.0	105.0	200,100		
Ohio	174.0	163.0	158.0	546,360		
Oklahoma	105.0	130.0	90.0	33,600		
Oregon	215.0	200.0	215.0	6,890		
Pennsylvania	143.0	128.0	111.0	131,560		
Rhode Island	(NA)	(NA)	(NA)	(NA)		
South Carolina	111.0	91.0	65.0	35,520		
South Dakota	151.0	135.0	132.0	706,690		
Tennessee	148.0	117.0	131.0	87,320		
Texas	130.0	145.0	93.0	254,800		
Utah	155.0	172.0	164.0	2,635		
Vermont ¹	(NA)	(NA)	(NA)	(NA)		
Virginia	131.0	67.0	118.0	43,230		
Washington	215.0	205.0	225.0	22,575		
West Virginia	125.0	90.0	114.0	3,790		
Wisconsin	153.0	162.0	156.0	448,290		
Wyoming	140.0	121.0	130.0	6,300		
United States	164.7	152.8	147.2	13,091,862		

(NA) Not available
¹ Area harvested for grain not estimated.

Crop Production 2011 Summary (January 2012)
USDA, National Agricultural Statistics Service



ISSN: 1936-3737

Released August 11, 2011, by the National Agricultural Statistics Service (NASS), Agricultural Statistics Board, United States Department of Agriculture (USDA).

Planted Acreage Update

Survey respondents who reported acreage as not yet planted in Minnesota, Montana, North Dakota, and South Dakota during the survey conducted in preparation for the *Acreage* report, released June 30, 2011 were re-contacted in July to determine how many of those acres were planted or still intended to be planted. Acreage estimates in this report reflect this updated information.

Corn Production Up 4 Percent from 2010
Soybean Production Down 8 Percent from 2010
Cotton Production Down 9 Percent from 2010
All Wheat Production Down 1 Percent from July Forecast

Corn production is forecast at 12.9 billion bushels, up 4 percent from 2010. If realized, this will be the third largest production total on record for the United States. Based on conditions as of August 1, yields are expected to average 153.0 bushels per acre, up 0.2 bushel from 2010, and the fourth highest yield on record. Acreage planted for all purposes is estimated at 92.3 million acres, unchanged from the June estimate. Area harvested for grain is forecast at 84.4 million acres, down less than 1 percent from June but up 4 percent from 2010.

Soybean production is forecast at 3.06 billion bushels, down 8 percent from last year. Based on August 1 conditions, yields are expected to average 41.4 bushels per acre, down 2.1 bushels from last year. Area for harvest in the United States is forecast at 73.8 million acres, down less than 1 percent from June and down 4 percent from 2010. Planted area for the Nation is estimated at 75.0 million acres, down fractionally from June.

All cotton production is forecast at 16.6 million 480-pound bales, down 9 percent from last year's 18.1 million bales. Yield is expected to average 822 pounds per harvested acre, up 10 pounds from last year. Upland cotton production is forecast at 15.8 million 480-pound bales, down 10 percent from 2010. American Pima production is forecast at 737,200 bales, up 46 percent from last year. Producers expect to harvest 9.67 million acres of all cotton, down 10 percent from 2010. This harvested total includes 9.38 million acres of Upland cotton and 287,500 acres of Pima cotton.

All wheat production, at 2.08 billion bushels, is down 1 percent from the July forecast and down 6 percent from 2010. Based on August 1 conditions, the United States yield is forecast at 45.2 bushels per acre, up 0.6 bushel from last month but down 1.2 bushels from last year.

Crop Production



United States Department of Agriculture

National Agricultural Statistics Service

ISSN: 1057-7822

Crop Production 2011 Summary

January 2012



ISSN: 1057-7822



Published no later than the 12th of each month.

Yields results primarily derived from two surveys

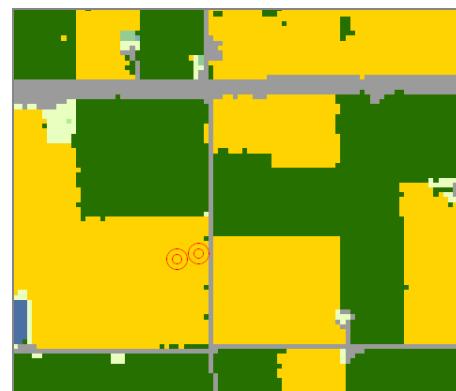
Agricultural Yield

- Farmer reported survey data of expected crop yields.
- Data obtained throughout the growing season.
- Conducted in all states except Alaska and Hawaii.
- Sample size in the 1000s per state.
- Farm operator contacts are selected from the March Crops/Stocks survey (small grains) and the June Crops/Stocks survey (late season crops and tobacco).
- Primarily telephone based.



Objective yield

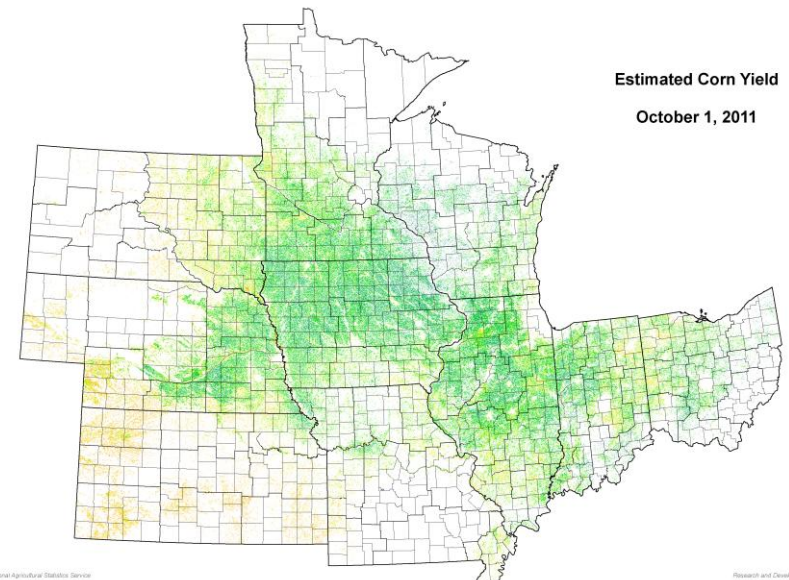
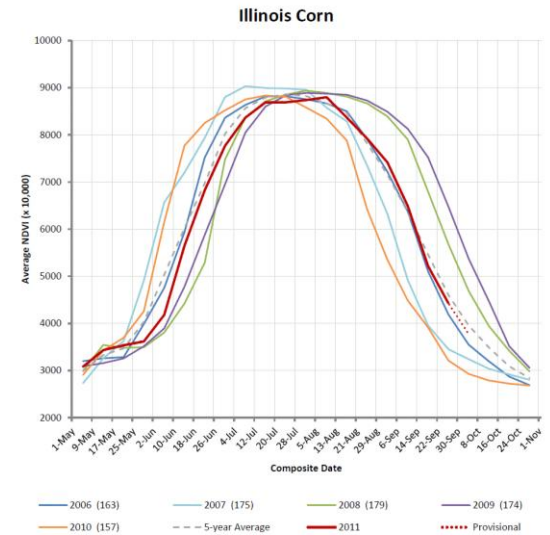
- Corn, Cotton, Soybeans, Wheat, Potatoes.
- Only done in states where the commodities are primarily found.
- Samples selected from areas found in June Area Survey (“Acreage”).
- Performed at 100s of sample sites per state.
- Biophysical plant/seed measurements obtained.
- Each plot revisited a few times per season.



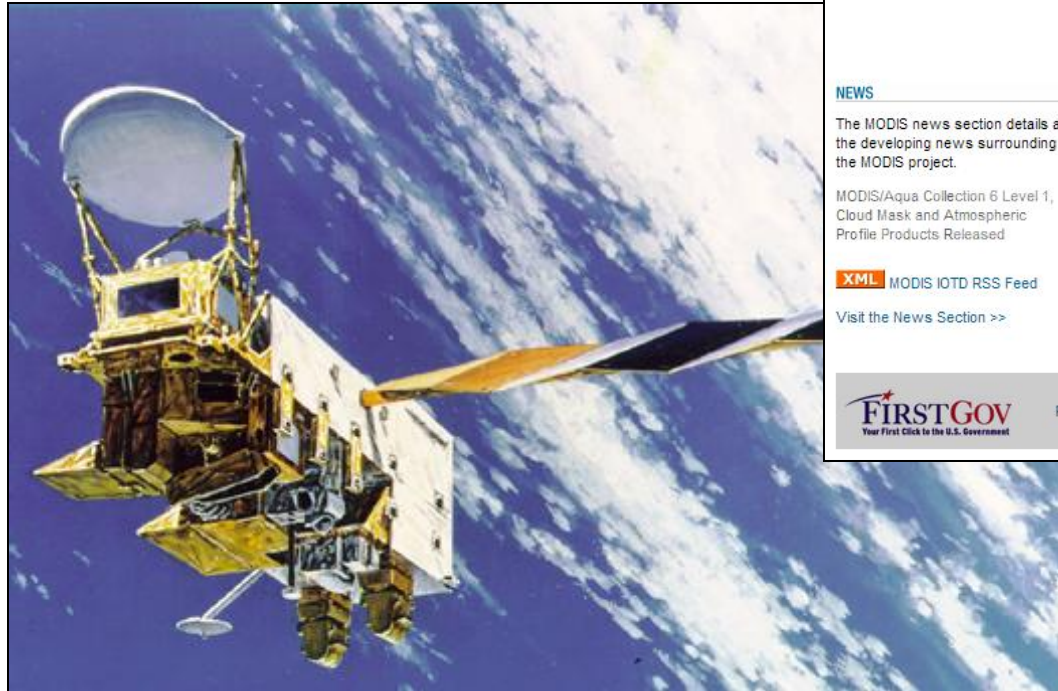
Estimating Yield from Remote Sensing


Third method for yield estimates

- There is a relationship between crop
 - Biomass, vigor, “greenness”, NDVI
 - and
 - Daytime land surface temperature
 - And the resulting corn or soybean yield
- Utilize MODIS data to obtain biomass and temperature variables
- Do it for
 - National, State, “ASD”, and County levels
 - And map at pixel level



Moderate Resolution Imaging Spectroradiometer (MODIS)



 NATIONAL AERONAUTICS AND SPACE ADMINISTRATION + NASA Homepage

SEARCH GO

MODIS Web

[+ ABOUT MODIS](#) [+ NEWS](#) [+ DATA](#) [+ IMAGES](#) [+ SCIENCE TEAM](#) [+ RELATED SITES](#) [+ SEARCH](#) [+ MODARCH](#)

DATA

The MODIS Data section contains everything from ATBDs to Product Descriptions to tutorials on ordering MODIS data from the various DAACs. Peruse the Data section today.

NEWS

The MODIS news section details all the developing news surrounding the MODIS project.


MODIS/Aqua Collection 6 Level 1, Cloud Mask and Atmospheric Profile Products Released

XML MODIS IOTD RSS Feed





[Visit the News Section >>](#)


IMAGES

[Super Typhoon Jelawat \(18W\) in the Philippine Sea](#)




DISCIPLINES

-  Atmosphere
-  Land
-  Ocean
-  Calibration

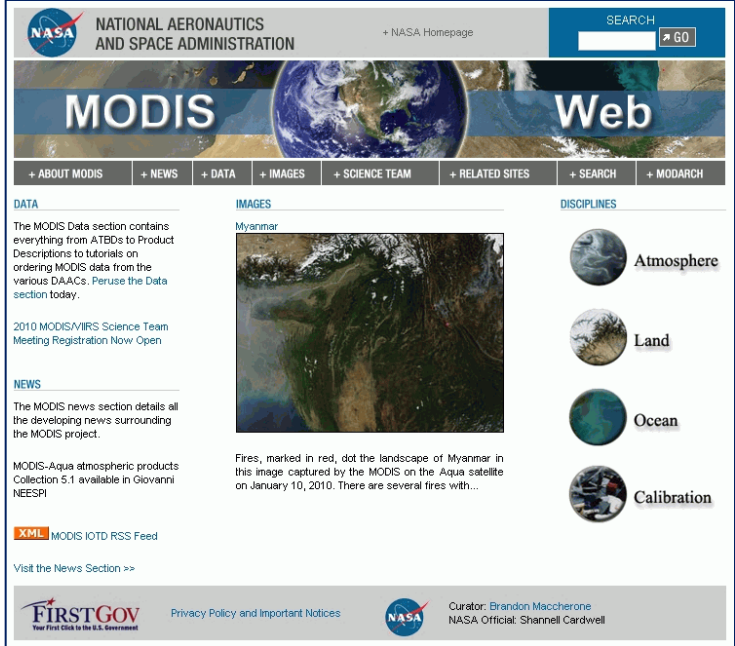
 **FIRSTGOV**
Your First Click to the U.S. Government

[Privacy Policy and Important Notices](#)

 Curator: Brandon Maccherone
NASA Official: Shannell Frazier

Why MODIS?

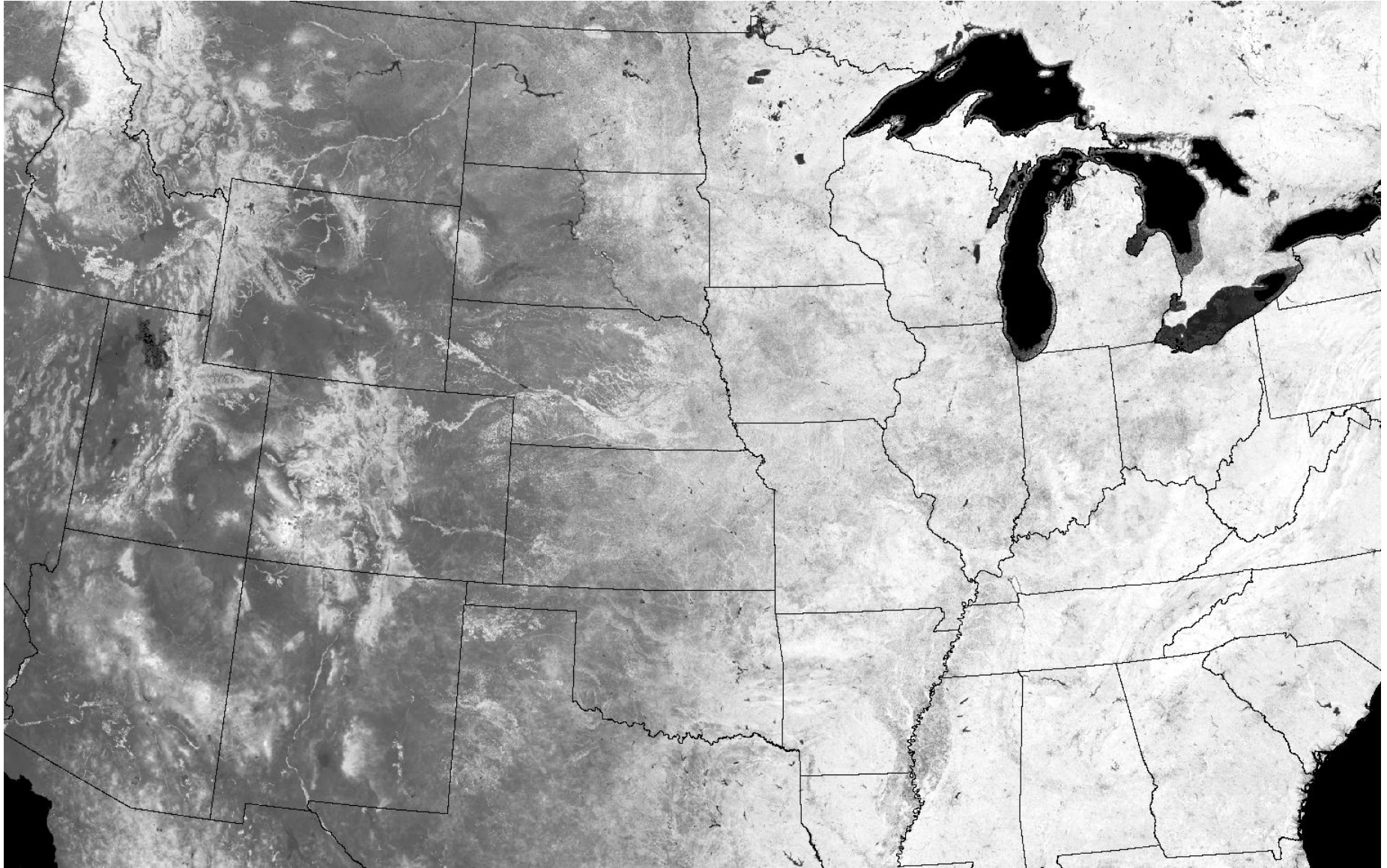
- Global coverage
- Daily revisit rate
- 15 acre ground sample resolution
 - from red and near-infrared bands
- “Best of” image mosaics automatically generated
 - 8 and 16-day temporal windows
- Timely
 - data usually available within a couple of days
- Free distribution
 - downloaded via ftp
- Robust user group
 - nearly 20,000 citations so far
- Launched in 1999 and 2002
 - Two of them
 - 10-year plus history
- 6 year design life but still functioning fine
- Similar follow-on mission
 - VIIRS



The screenshot shows the MODIS Web portal homepage. At the top, it features the NASA logo and the text "NATIONAL AERONAUTICS AND SPACE ADMINISTRATION" with a link to the NASA Homepage. A search bar is located in the top right corner. Below the header is a large banner with the text "MODIS Web" and a satellite image of Earth. A navigation menu includes links for "ABOUT MODIS", "NEWS", "DATA", "IMAGES", "SCIENCE TEAM", "RELATED SITES", "SEARCH", and "MODARCH". The main content area is divided into three columns: "DATA", "IMAGES", and "DISCIPLINES". The "DATA" section contains text about ATBDs and a link to a meeting registration. The "IMAGES" section features a satellite image of Myanmar with a caption about fires. The "DISCIPLINES" section has icons for "Atmosphere", "Land", "Ocean", and "Calibration". At the bottom, there is a "FIRST GOV" logo, a link to the "Privacy Policy and Important Notices", and the names of the curators: Brandon Maccherone and Shannell Cardwell.

modis.gsfc.nasa.gov

MODIS NDVI data example



Calculation and use of NDVI

The screenshot shows the Wikipedia article for 'Normalized Difference Vegetation Index'. The article title is 'Normalized Difference Vegetation Index' and it is from Wikipedia, the free encyclopedia. The article text states: 'The **Normalized Difference Vegetation Index (NDVI)** is a simple graphical indicator that can be used to analyze remote sensing measurements, typically but not necessarily from a space platform, and assess whether the target being observed contains live green vegetation or not.' The article also includes a 'Brief history' section, a 'Contents' table of contents, and a color-coded satellite image of the British Isles in June 2003. The image shows a color scale from -1.0 (dark blue) to 1.0 (dark red), with a legend below it. The legend is labeled 'average NDVI of June 2003' and shows a color gradient from dark blue (-1.0) to dark red (1.0). The image shows the British Isles with a color scale from -1.0 to 1.0. The caption below the image reads 'NDVI in June over the British Isles (NOAA AVHRR)'. The article also includes a 'Contents' table of contents with links to 'Brief history', 'Rationale', 'Performance and limitations', 'See also', 'References', and 'External links'. The article also includes a 'Brief history' section with the text: 'The exploration of outer space started in earnest with the launch of Sputnik 1 by the Soviet Union on 4 October 1957. This was the first man-made satellite orbiting the Earth. Subsequent successful launches, both in the Soviet Union (e.g., the Sputnik and Cosmos programs), and in the U.S. (e.g., the Explorer program), quickly led to the design and operation of dedicated meteorological satellites. These are orbiting platforms embarking instruments specially designed to observe the Earth's atmosphere and surface with a view to improve weather forecasting. Starting in 1960, the TIROS series of satellites embarked television cameras and radiometers. This was later (from 1964 onwards) followed by the Nimbus satellites and the family of Advanced Very High Resolution Radiometer instruments on-board the National Oceanic and Atmospheric Administration (NOAA) platforms. The latter measures the reflectance of the planet in red and near-infrared bands, as well as in the thermal infrared. In parallel, NASA developed the Earth Resources Technology Satellite (ERTS), which became the precursor to the Landsat program. These early sensors had minimal spectral resolution, but tended to include bands'.

$$NDVI = \frac{(NIR - VIS)}{(NIR + VIS)}$$

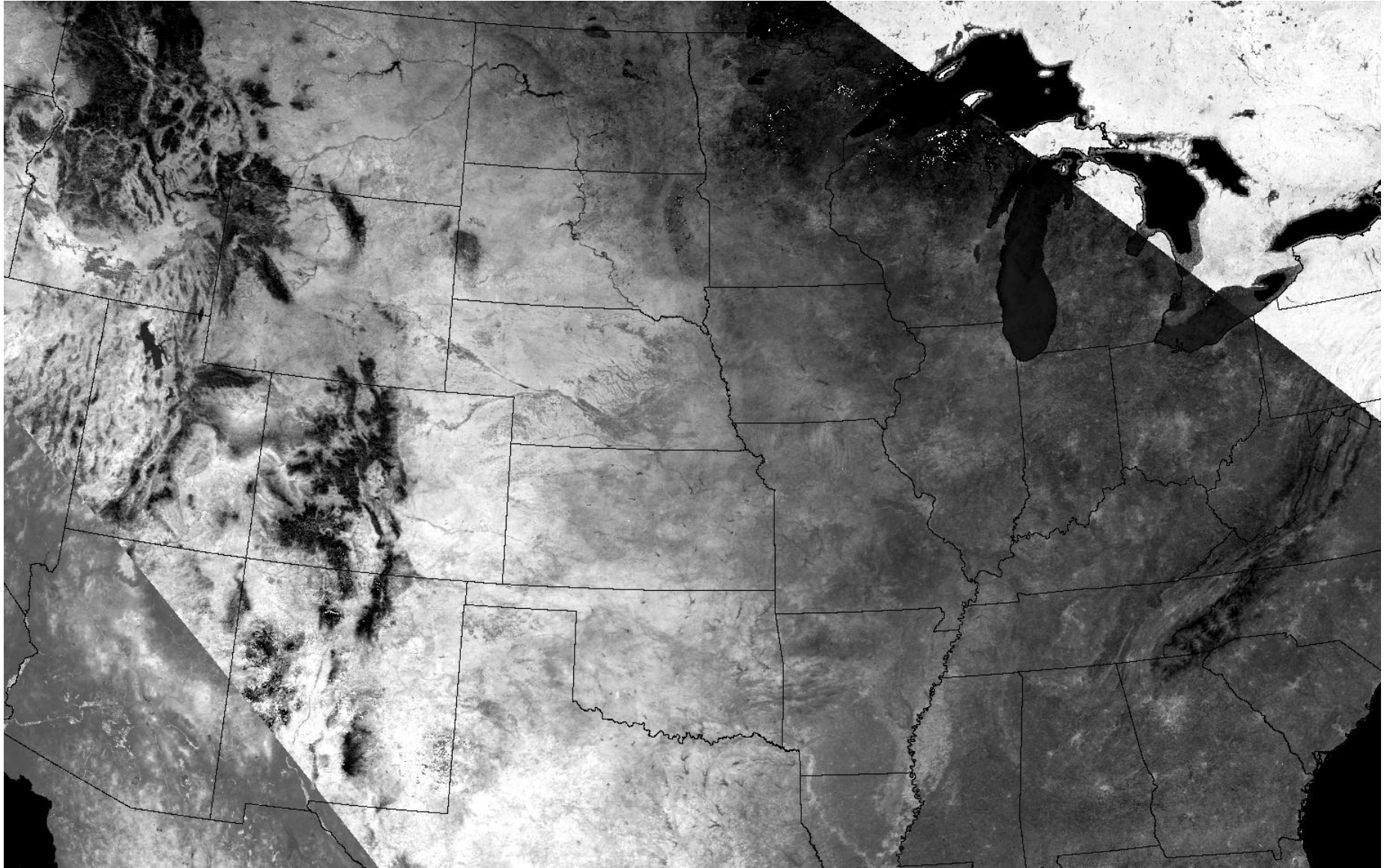
NIR = near-infrared
VIS = visible

Ranges from -1.0 to 1.0

NDVI is a related to

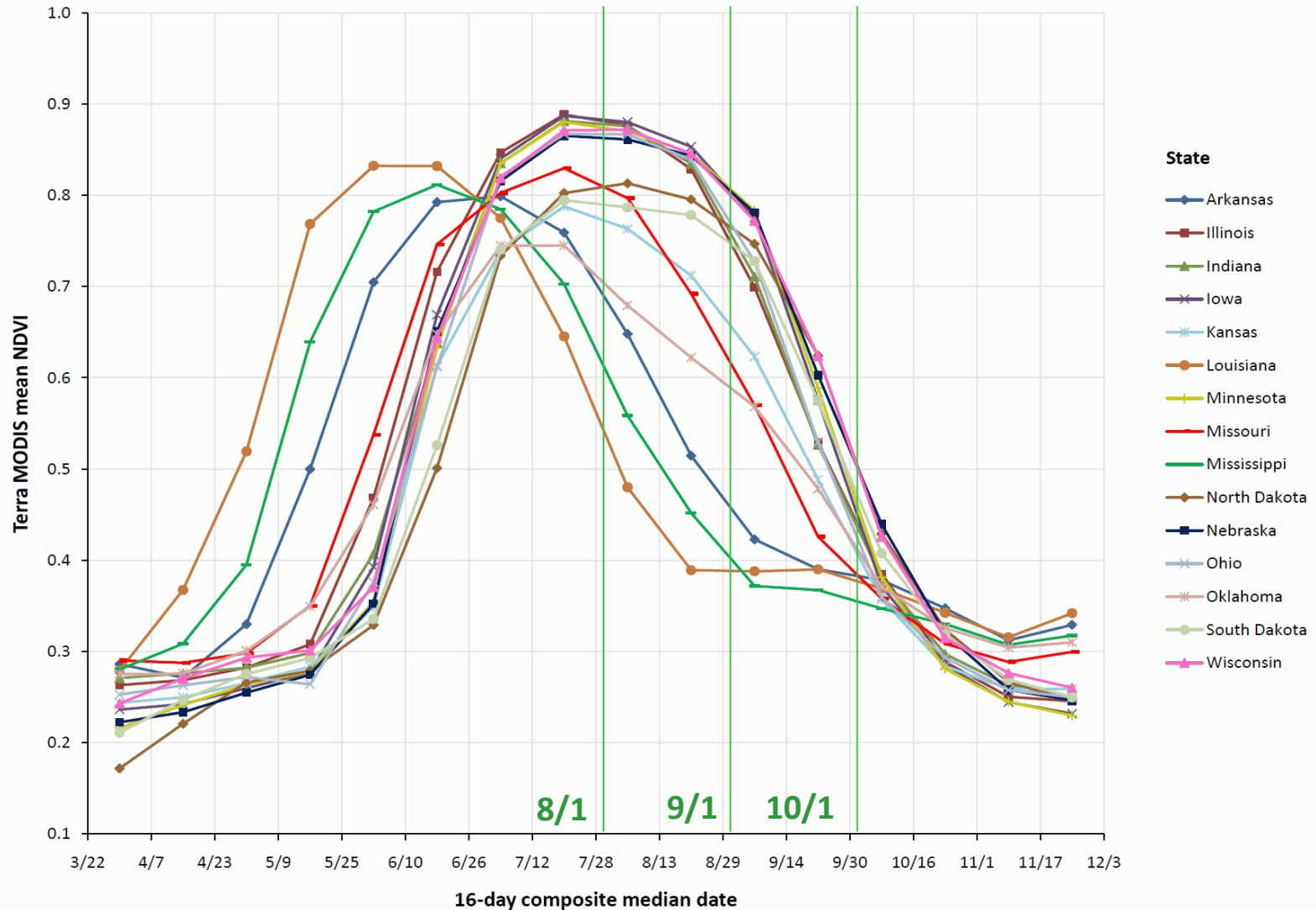
- Plant health
- Chlorophyll content
- “Greenness”
- Biomass
- Vegetation vigor

MODIS LST data example

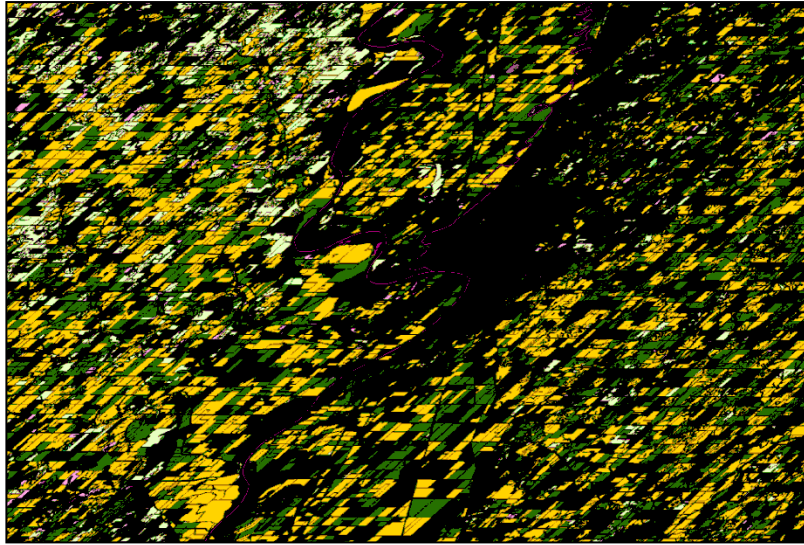


Corn phenology fundamentals

Corn 5-year average 2006-2010

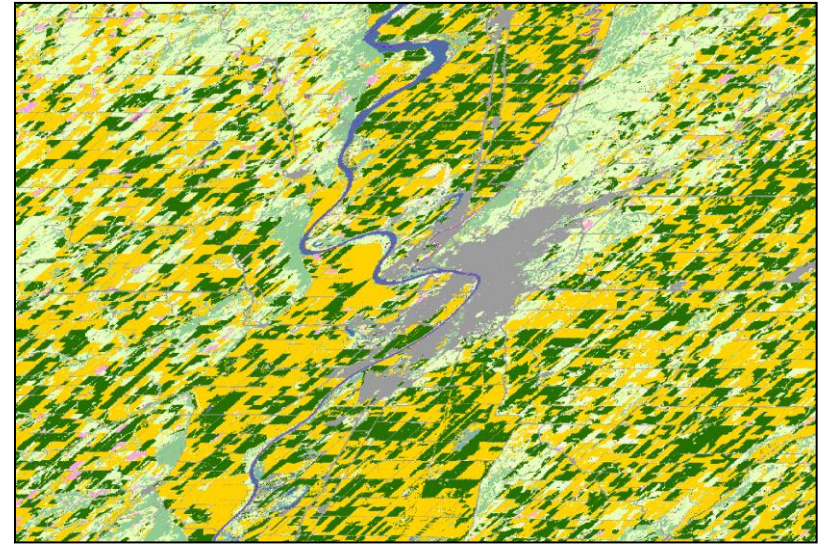


Establishing the pixels that are only corn

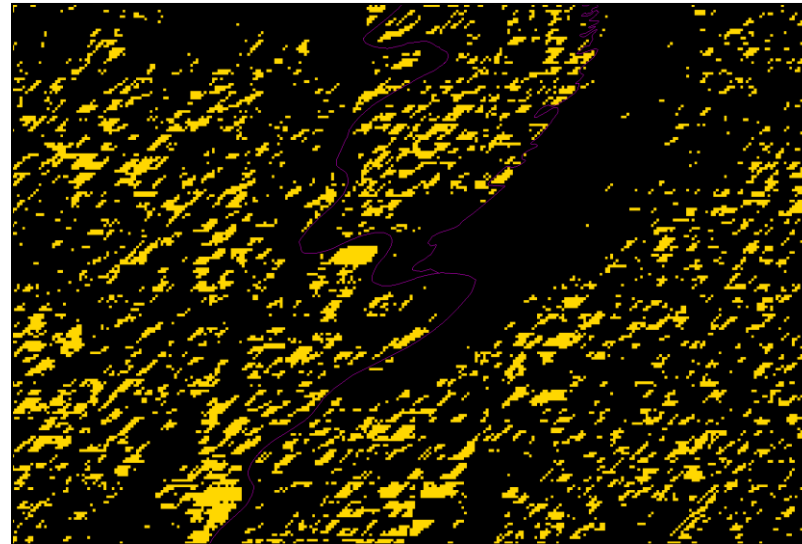


FSA CLU/578
(early season)

- or -

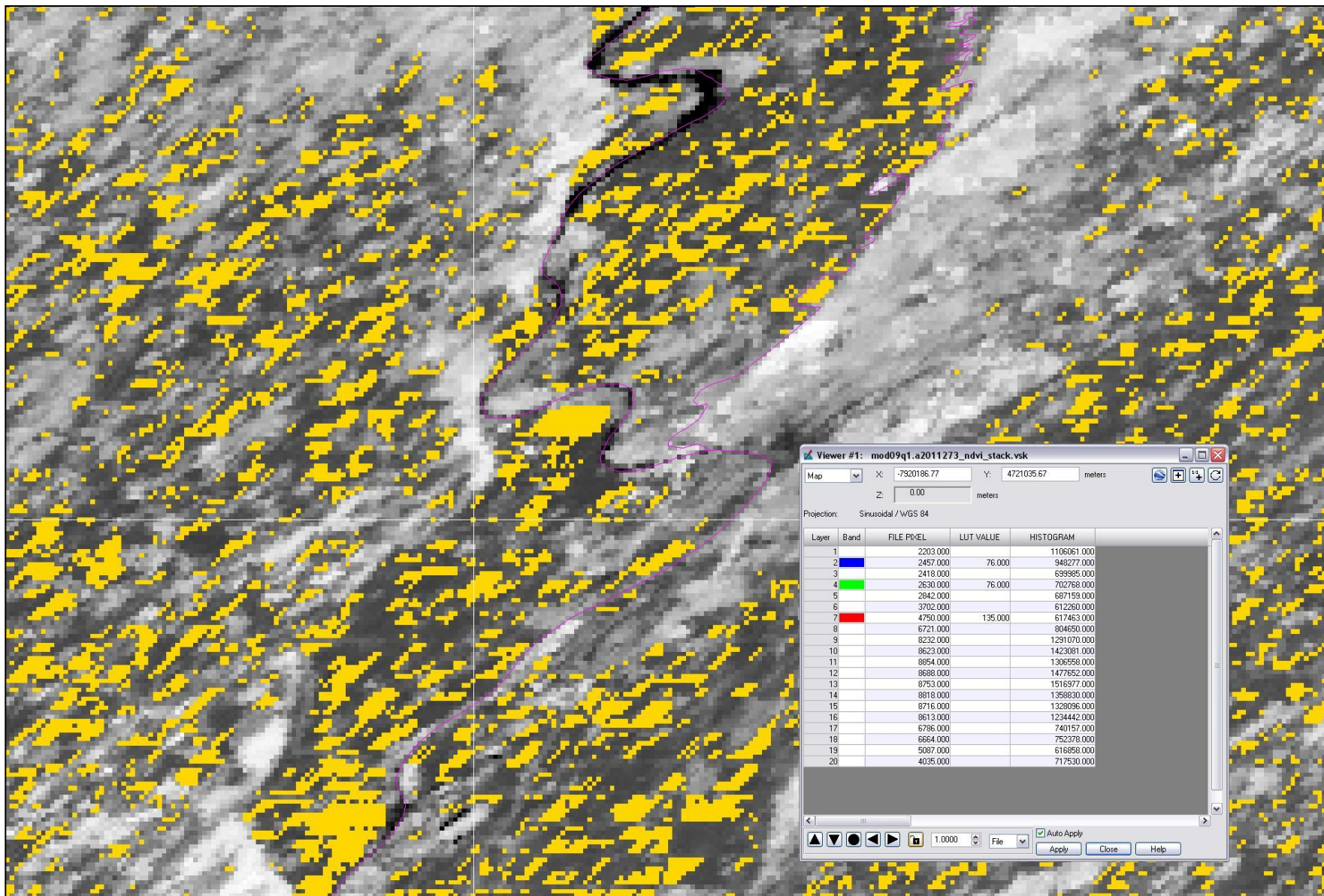


NASS CDL
(late season)



MODIS-scaled
High probability sample
of corn areas

Intersecting corn "mask" with MODIS data



County-level database developed

- Potential predictor variables (independent)
 - State (All major production Corn Belt states)
 - County (for each that had a published estimate, ~1000 of them)
 - Year (2006 – 2011)
 - 32 for each ranging every 8 days from February 18 – October 30
 - NDVI (derived from Terra “MOD09Q1” 250m)
 - Day LST (1:30 PM – from Aqua “MYD11A2” 1000m)
 - Thus 68 in total
- Forecast variable (dependent)
 - NASS published county level yield (available from NASS “Quickstats” webpage)
- Resulting database to evaluate ~5000 records

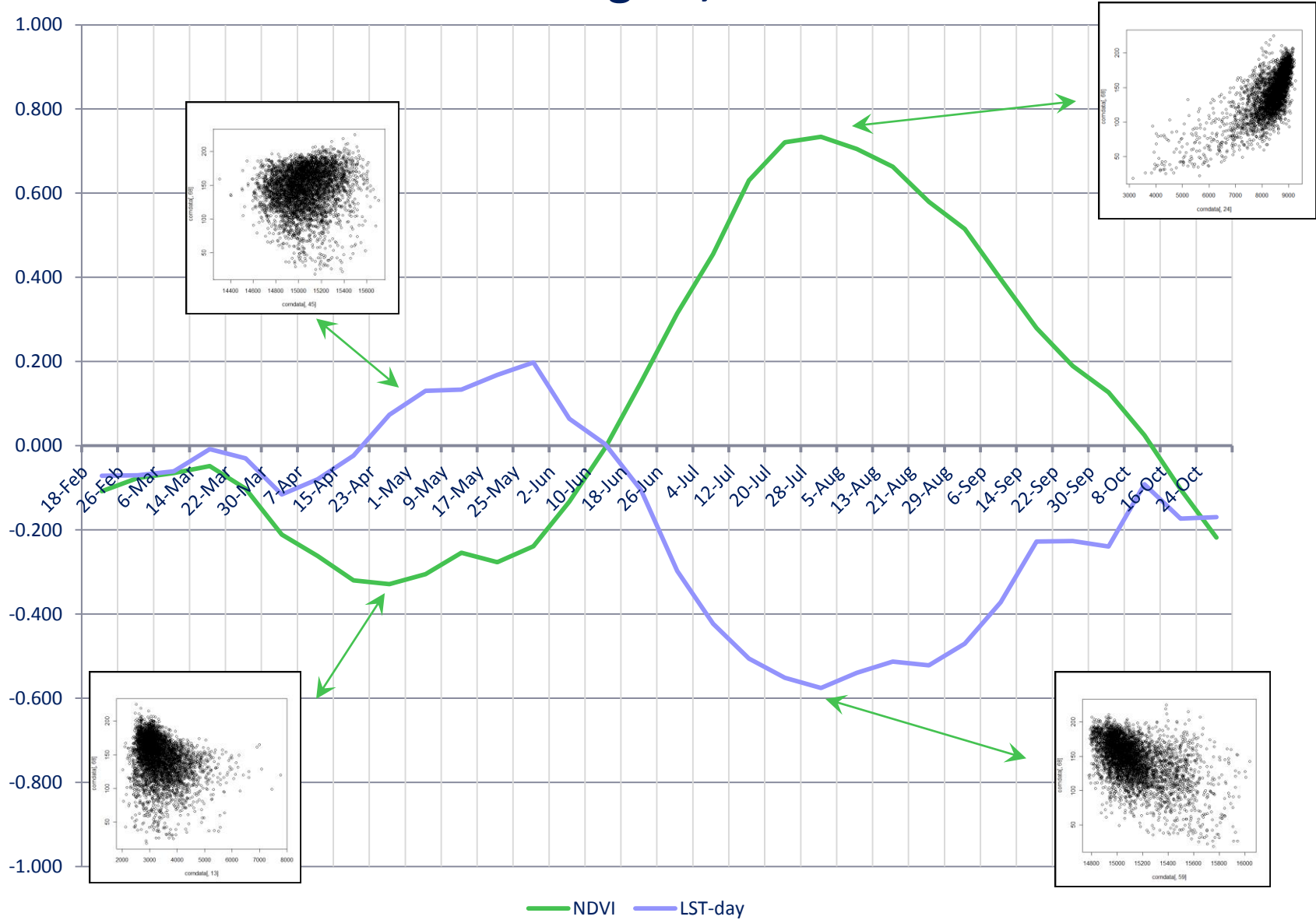
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
1	20	119	2009	2263.17	2276.58	2303.59	2348.65	2434.37	2359.19	2266.35	2288.19	2367.71	2512.1	2629.59	28
2	20	69	2009	2288.49	2311.38	2342.2	2339.95	2407.06	2383.9	2340.03	2488.27	2703.84	2684.03	2654.48	28
3	20	119	2007	2276.42	2251.23	2264.43	2352.61	2469.31	2547.4	2619.07	2684.33	2815.85	2966.01	3141.06	33
4	20	81	2009	2218.23	2247.03	2304.9	2340.96	2421.18	2349.36	2270.35	2346.69	2446.1	2476.15	2472.75	24
5	31	137	2009	2167.21	2154.48	2139.61	2156.15	2212.33	2278.63	2384.85	2470.95	2573.56	2613.23	2606.06	27
6	20	175	2007	2232.36	2255.85	2303.21	2438.8	2705.49	2732.7	2735.31	2733.36	2854.66	3053.56	3107.51	32
7	17	187	2008	1766.18	2556.21	2600.69	2643.94	2644.2	2590.61	2594.14	2706.57	2852.78	3125.75	3459.15	34
8	19	93	2009	1981.28	2104.24	2312.87	2173.23	2118.37	2141.06	2191.56	2315.99	2460.3	2716.56	2887.91	30
9	31	185	2009	2116.76	2196.12	2298.1	2295.07	2297.18	2361.89	2379.85	2411.81	2493.01	2683.23	2960.81	31
10	17	203	2008	2163.13	2259.22	2355.42	2618.68	2614.77	2551.36	2627.17	2764.98	2829.49	2942.49	3344.19	33
11	17	109	2008	2527.91	2730.11	2685.04	2709.46	2770.32	2661.12	2665.43	2793.58	2972.49	3209.61	3650.63	35
12	31	81	2009	2142.37	2188.07	2278.81	2282.22	2292.83	2342.79	2365.32	2406.9	2496	2665.16	2881.11	3
13	17	203	2007	96.494	2274.4	2601.68	2709.84	2777.16	2584.11	256.82	2783.05	3024.46	3052.39	3121.62	34
14	19	165	2009	2077.56	2064.12	2053.83	2155.59	2357.15	2276.43	2252.18	2341.4	2441.82	2643.78	3136.29	34
15	20	81	2007	2350.12	2276.67	2247.89	2383.91	2661.31	2660.21	2643.89	2631.8	2883.8	3317.63	3184.67	34
16	17	175	2008	1352.35	1906.66	2479.59	2545.58	2607.91	2503.92	2507.45	2621.58	2781.29	2988.94	3283.28	3
17	31	99	2009	2141.86	2118.68	2121.94	2139.96	2196.7	2261.97	2360.63	2444.19	2524.94	2535.12	2579.6	26
18	19	35	2009	2131.72	2203.73	2379.83	2237.11	2157.33	2182.48	2230.26	2375.26	2546.31	2704.29	2947.32	31

■ ■ ■

	BM	BH	BO	BP
98.5	15049.9	15100.7	14962.4	225
911.1	14966.6	15021.6	14898	219
29.4	15280.7	15183.5	15095.1	215
60.4	15013.3	15065.8	14938.8	212
66.6	14578	14519.3	14533.3	211
455	15331.1	15236.9	15109.8	207
40.3	14845.2	14543	14629.6	207
16.4	14537.5	14458.7	14470.1	207
678	14508.9	14480.1	14463.9	207
76.6	14873.8	14508.4	14620.5	206
68.4	14885.8	14568.7	14658.5	206
76.1	14501.9	14447.6	14462.2	205
14.6	15111.4	14740.3	14567.9	204
33.3	14608.8	14509.7	14527.9	204
56.2	15275	15216.8	15107.6	203
63.8	14841.5	14532.1	14606	203
55.4	14547.6	14450.1	14520.7	203
62.7	14564.6	14480.9	14480.3	203

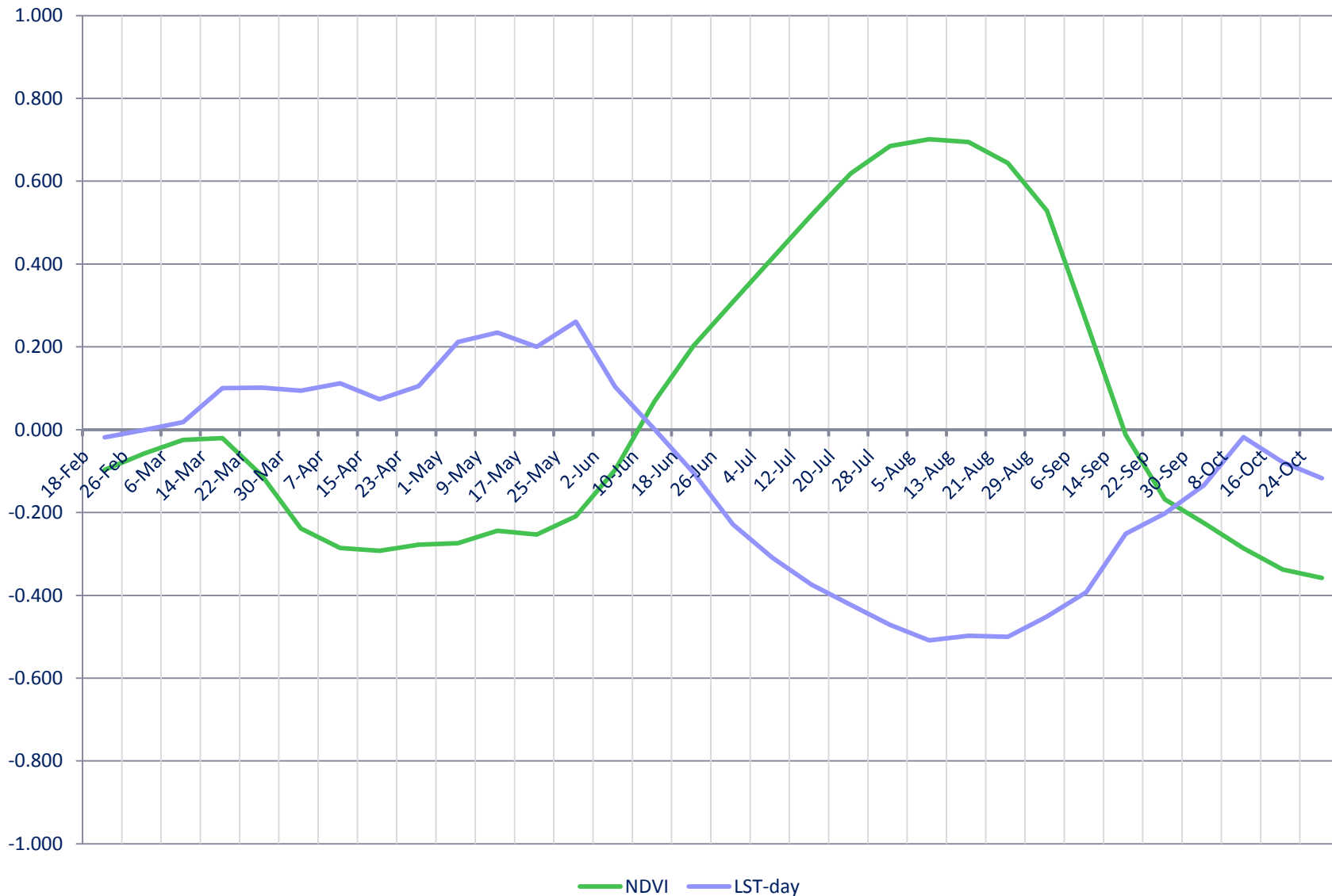
Corn yield dependence at county level

Corn Belt region, 2006-2011



Soybean yield dependence at county level


Corn Belt region, 2006-2011



Rulequest Cubist

Cubist [network version]

File Edit Help



CS_corn_5year

attribute definitions [CS_corn_5year.names]
 training cases to be analyzed (CS_corn_5year.data)
 test cases [CS_corn_5year.test]
 rule-based model [CS_corn_5year.model]
 output file [CS_corn_5year.out]

Model Construction Options

Form of Model

Rules alone
 Instances and rules
 Let Cubist decide

Use nearest instances
 Committee of members

Cross-validate folds
 Use sample of % cases
 Lock sample

Maximum rules
 Extrapolation allowed %
 Unbiased rules

OK Defaults Cancel

Results for CS_corn_5year

File Edit

Model 1:

Rule 1/1: [47 cases, mean 64.59, range 18.3 to 135, est err 17.24]

```

if
  NDVI_14 <= 5292.634
  NDVI_22 <= 6741.589
  LST_1330_02 <= 14435.63
  LST_1330_04 <= 14906.01
then
  yield = -88.46 + 0.0695 LST_1330_02 - 0.0856 LST_1330_04
    + 0.0499 LST_1330_05 + 0.0267 NDVI_31 - 0.037 LST_1330_07
    - 0.0148 NDVI_09 + 0.0031 NDVI_25 + 0.015 LST_1330_23
    + 0.0053 LST_1330_03 + 0.0024 NDVI_28 - 0.004 NDVI_12
    - 0.009 LST_1330_22 + 0.0027 NDVI_20 + 0.0025 NDVI_11
    - 0.007 LST_1330_21 - 0.0013 NDVI_24 + 0.0014 NDVI_23
    - 0.0013 NDVI_29 + 0.004 LST_1330_30 + 0.006 LST_1330_32
    - 0.005 LST_1330_24 + 0.005 LST_1330_10 + 0.001 NDVI_22
    - 0.0007 NDVI_03 - 0.0006 NDVI_01 - 0.003 LST_1330_19
    + 0.0009 NDVI_04 + 0.0005 NDVI_02 + 0.0008 NDVI_19
    + 0.003 LST_1330_26 + 0.003 LST_1330_16 + 0.001 NDVI_32
    - 0.002 LST_1330_13 - 0.002 LST_1330_31 - 0.002 LST_1330_11
    - 0.0007 NDVI_08 - 0.002 LST_1330_09 - 0.0006 NDVI_10
    + 0.0004 NDVI_21 + 0.001 LST_1330_12
  
```

Rule 1/2: [44 cases, mean 100.00, range 49 to 152, est err 9.61]

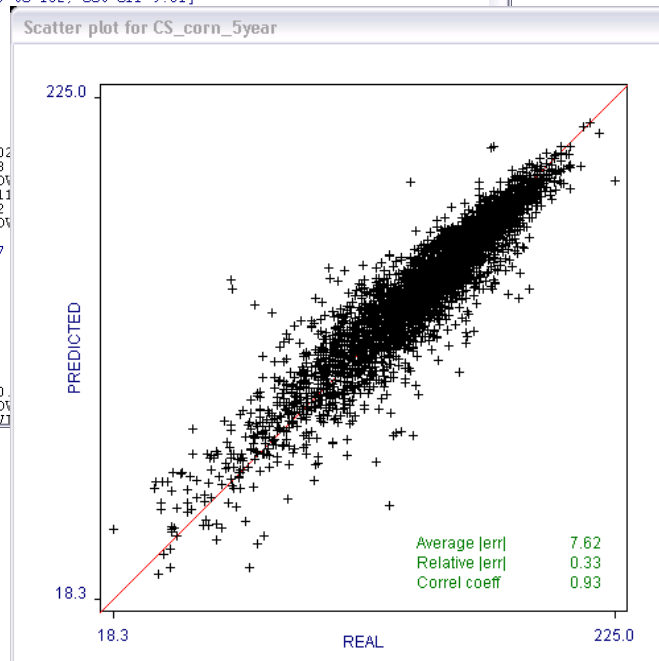
```

if
  NDVI_05 > 2600.681
  NDVI_14 <= 5292.634
  NDVI_21 <= 8340.541
  NDVI_31 > 3287.235
  LST_1330_03 <= 14082.5
  LST_1330_12 <= 15346.29
  LST_1330_32 > 14185.42
then
  yield = -1570.12 + 0.0973 LST_1330_02
    - 0.0578 LST_1330_01 - 0.038
    + 0.0373 NDVI_20 + 0.0149 NDV
    + 0.0263 LST_1330_04 + 0.0111
    + 0.0201 LST_1330_05 + 0.022
    - 0.0032 NDVI_27 - 0.0063 NDV
  
```

Rule 1/3: [37 cases, mean 104.21, range 37

```

if
  NDVI_04 > 2356.286
  NDVI_14 <= 5292.634
  NDVI_16 <= 6590.642
  NDVI_21 <= 8340.541
  LST_1330_02 > 14435.63
  LST_1330_04 <= 14906.01
then
  yield = -545.78 + 0.0109 NDVI_25 - 0
    + 0.0121 NDVI_04 - 0.0109 NDV
    + 0.0044 NDVI_16 + 0.004 NDV
  
```



Example county-level prediction output

Case No	Given Value	Predicted Value
1	1701.0	92.49 +- 24.78
2	1703.0	138.24 +- 24.78
3	1705.0	64.58 +- 24.78
4	1707.0	129.20 +- 24.78
5	1709.0	106.46 +- 24.78
6	17011.0	132.43 +- 24.78
7	17013.0	104.15 +- 24.78
8	17015.0	150.99 +- 24.78
9	17017.0	127.98 +- 24.78
10	17019.0	123.06 +- 24.78
11	17021.0	122.79 +- 24.78
12	17023.0	79.02 +- 24.78
13	17025.0	45.30 +- 24.78
14	17027.0	59.63 +- 24.78
15	17029.0	94.61 +- 24.78
16	17031.0	140.96 +- 24.78
17	17033.0	65.67 +- 24.78
18	17035.0	83.40 +- 24.78
19	17037.0	143.90 +- 24.78
20	17039.0	124.48 +- 24.78
21	17041.0	125.98 +- 24.78
22	17043.0	137.51 +- 24.78
23	17045.0	104.61 +- 24.78
24	17047.0	76.02 +- 24.78
25	17049.0	66.14 +- 24.78
26	17051.0	62.86 +- 24.78
27	17053.0	101.90 +- 24.78
28	17055.0	54.85 +- 24.78
29	17057.0	109.45 +- 24.78
30	17059.0	113.84 +- 24.78
31	17061.0	104.00 +- 24.78
32	17063.0	100.02 +- 24.78
33	17065.0	64.23 +- 24.78
34	17067.0	110.53 +- 24.78
35	17069.0	69.21 +- 24.78
36	17071.0	127.78 +- 24.78
37	17073.0	121.29 +- 24.78
38	17075.0	128.79 +- 24.78
39	17077.0	63.20 +- 24.78
40	17079.0	70.46 +- 24.78
41	17081.0	46.99 +- 24.78
42	17083.0	94.22 +- 24.78
43	17085.0	123.96 +- 24.78
44	17087.0	73.05 +- 24.78
45	17089.0	131.46 +- 24.78
46	17091.0	130.11 +- 24.78
47	17093.0	120.26 +- 24.78
48	17095.0	128.95 +- 24.78
49	17097.0	120.66 +- 24.78
50	17099.0	128.53 +- 24.78

Corn

Case No	Given Value	Predicted Value
1	501.0	45.30 +- 7.60
2	503.0	32.20 +- 7.60
3	505.0	4.31 +- 9.50
4	507.0	4.31 +- 9.50
5	509.0	4.31 +- 9.50
6	5011.0	4.31 +- 9.50
7	5013.0	4.31 +- 9.50
8	5015.0	4.31 +- 9.50
9	5017.0	36.80 +- 7.60
10	5019.0	10.88 +- 7.60
11	5021.0	39.92 +- 7.60
12	5023.0	4.31 +- 9.50
13	5025.0	4.31 +- 9.50
14	5027.0	4.31 +- 9.50
15	5029.0	19.68 +- 7.60
16	5031.0	34.14 +- 7.60
17	5033.0	28.26 +- 7.60
18	5035.0	43.82 +- 7.60
19	5037.0	39.48 +- 7.60
20	5039.0	4.31 +- 9.50
21	5041.0	39.78 +- 7.60
22	5043.0	38.52 +- 7.60
23	5045.0	11.48 +- 7.60
24	5047.0	4.31 +- 9.50
25	5049.0	4.31 +- 9.50
26	5051.0	4.31 +- 9.50
27	5053.0	4.31 +- 9.50
28	5055.0	33.53 +- 7.60
29	5057.0	28.21 +- 7.60
30	5059.0	12.47 +- 7.60
31	5061.0	4.31 +- 9.50
32	5063.0	31.07 +- 7.60
33	5065.0	4.31 +- 9.50
34	5067.0	34.07 +- 7.60
35	5069.0	41.76 +- 7.60
36	5071.0	4.31 +- 9.50
37	5073.0	36.72 +- 7.60
38	5075.0	34.29 +- 7.60
39	5077.0	41.19 +- 7.60
40	5079.0	42.98 +- 7.60
41	5081.0	26.32 +- 7.60
42	5083.0	10.81 +- 7.60
43	5085.0	34.80 +- 7.60
44	5087.0	4.31 +- 9.50
45	5089.0	4.31 +- 9.50
46	5091.0	22.51 +- 7.60
47	5093.0	43.11 +- 7.60
48	5095.0	36.57 +- 7.60
49	5097.0	4.31 +- 9.50
50	5099.0	4.31 +- 9.50

Soybeans

Weight by a 3-year average of harvested acres to derive ASD, state, and region estimates

“Voodoo Modeling”

Utilizing Rulequest Cubist software

- Learning tool to predict continuous (vs discrete) outcomes
- Allow for “composite” predictions using both
 - Instance-based
 - “Nearest neighbor”
 - Predicts the target value of a new case by finding the **n** most similar cases in the training data, and averaging their target values.
 - Model-based, via decision trees and piecewise linear regression
 - Divide and conquer strategy
 - Recursive splitting of training data to minimize intra-subset variation
 - Thus, for composite of instances and models:
 - Cubist finds the **n** training cases that are “nearest” (most similar) to the case in question. Then, rather than averaging their target values directly, Cubist first adjusts these values using the rule-based model.
- Also, does “Committee” models
 - made up of several rule-based models. Each member of the committee predicts the target value for a case and the members' predictions are averaged to give a final prediction

Model Construction Options

Form of Model

Rules alone

Instances and rules

Let Cubist decide

Use nearest instances

Committee of members

Cross-validate folds

Use sample of % cases

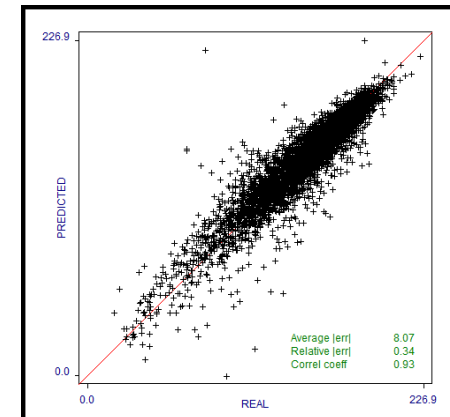
Lock sample

Maximum rules

Extrapolation allowed %

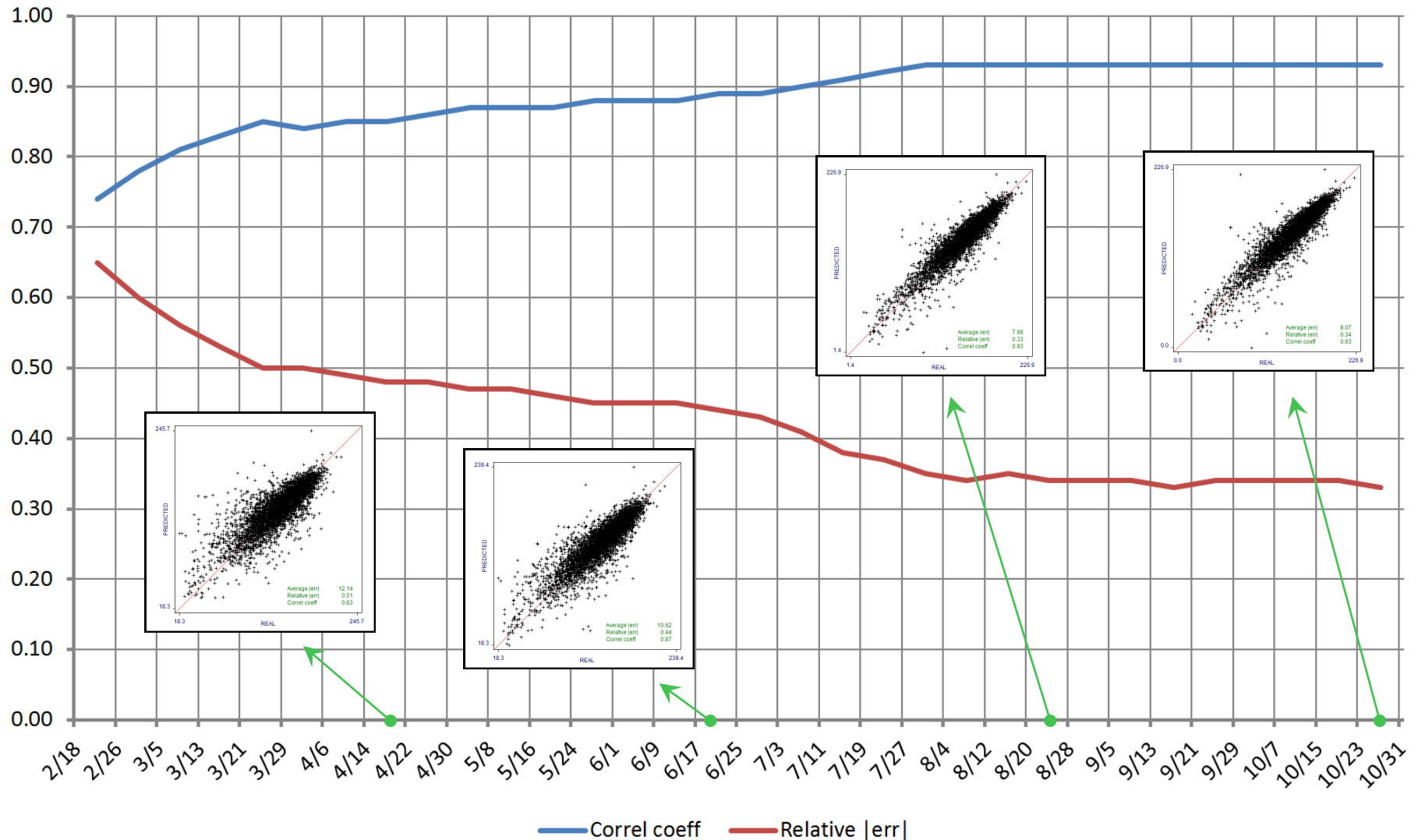
Unbiased rules

OK Defaults Cancel



Corn yield regression-tree model performance v. data timing

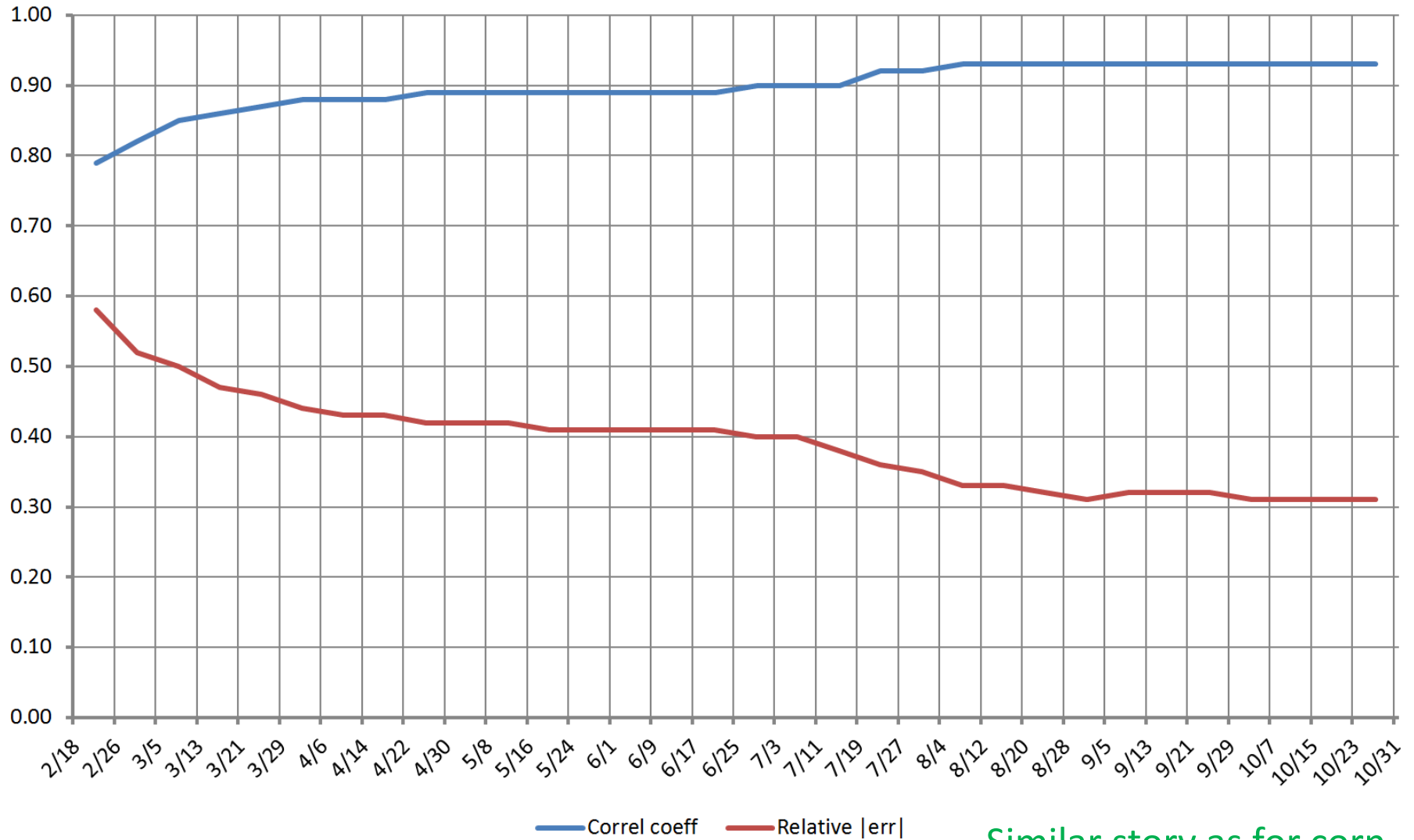
county level, speculative region, 2006-2011



The **relative error magnitude** is the ratio of the average error magnitude to the error magnitude that would result from always predicting the mean value; for useful models, this should be less than 1!

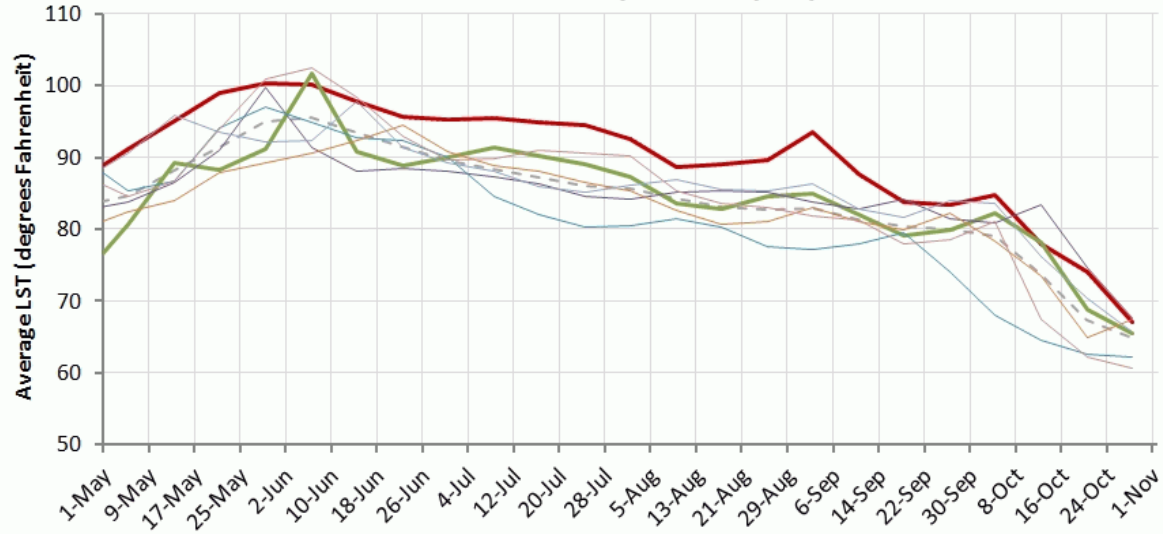
The **correlation coefficient** measures the agreement between the cases' actual values of the target attribute and those values predicted by the model.

Soybean yield regression-tree model performance v. data timing county level, speculative region, 2006-2011

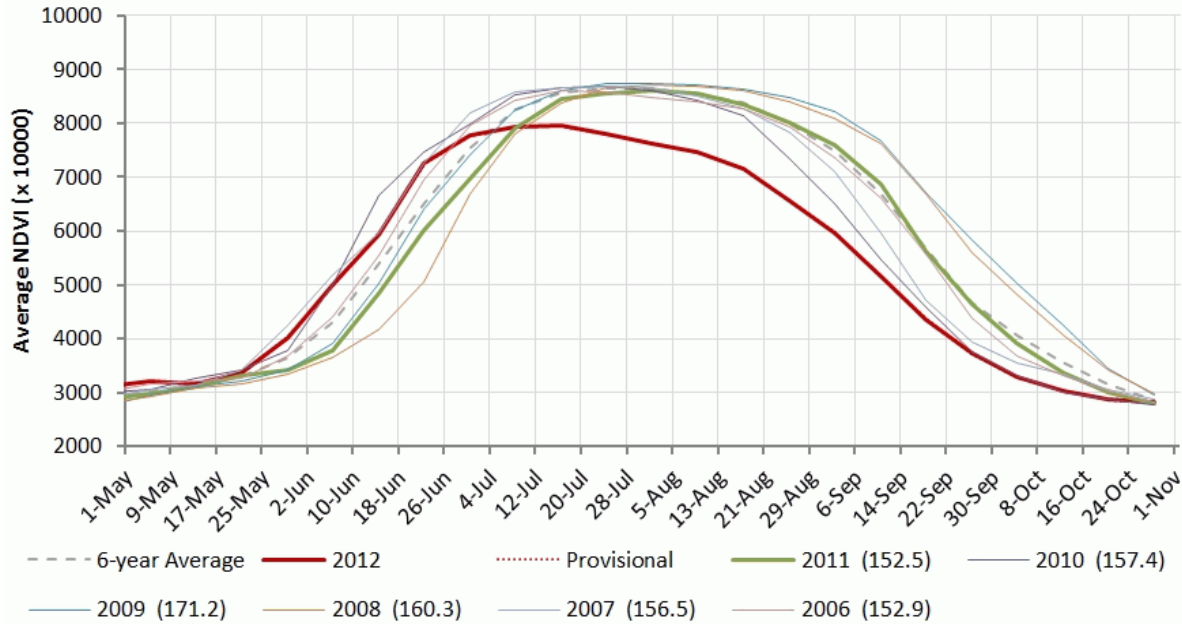


Similar story as for corn

Speculative Region - Corn Land Surface Temperature (LST)

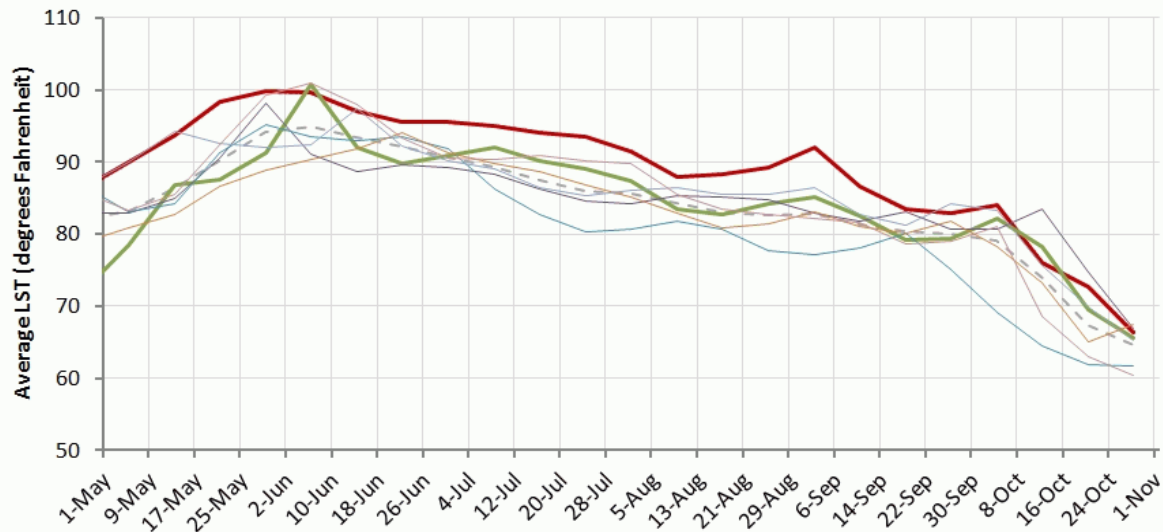


Normalized Difference Vegetation Index (NDVI)

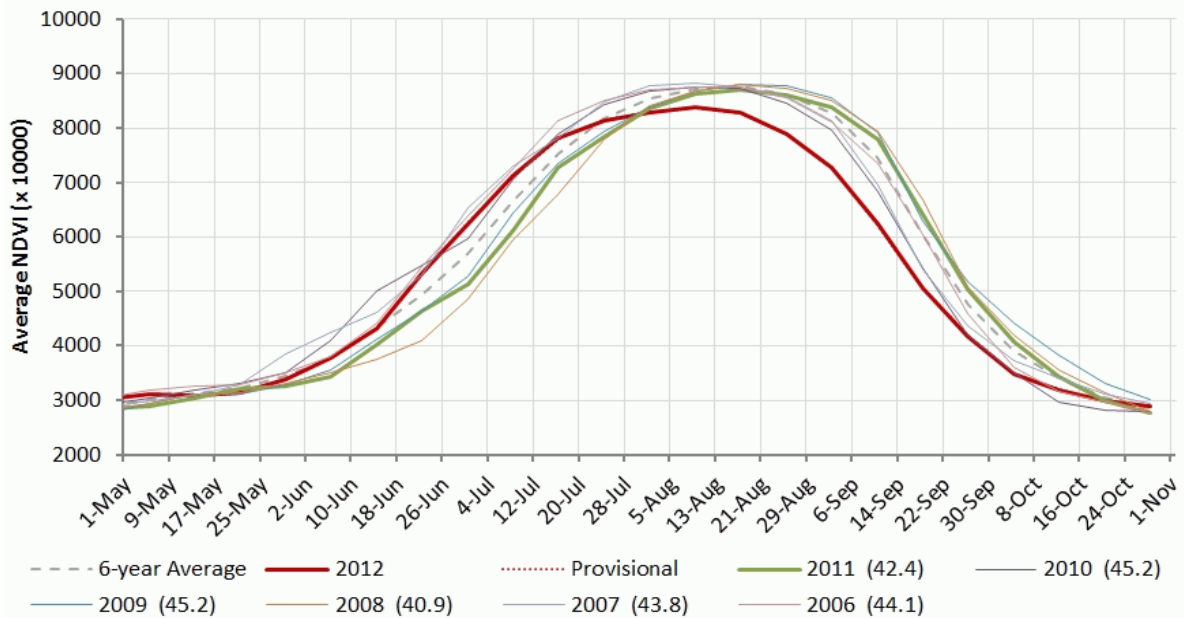


Speculative Region - Soybeans

Land Surface Temperature (LST)

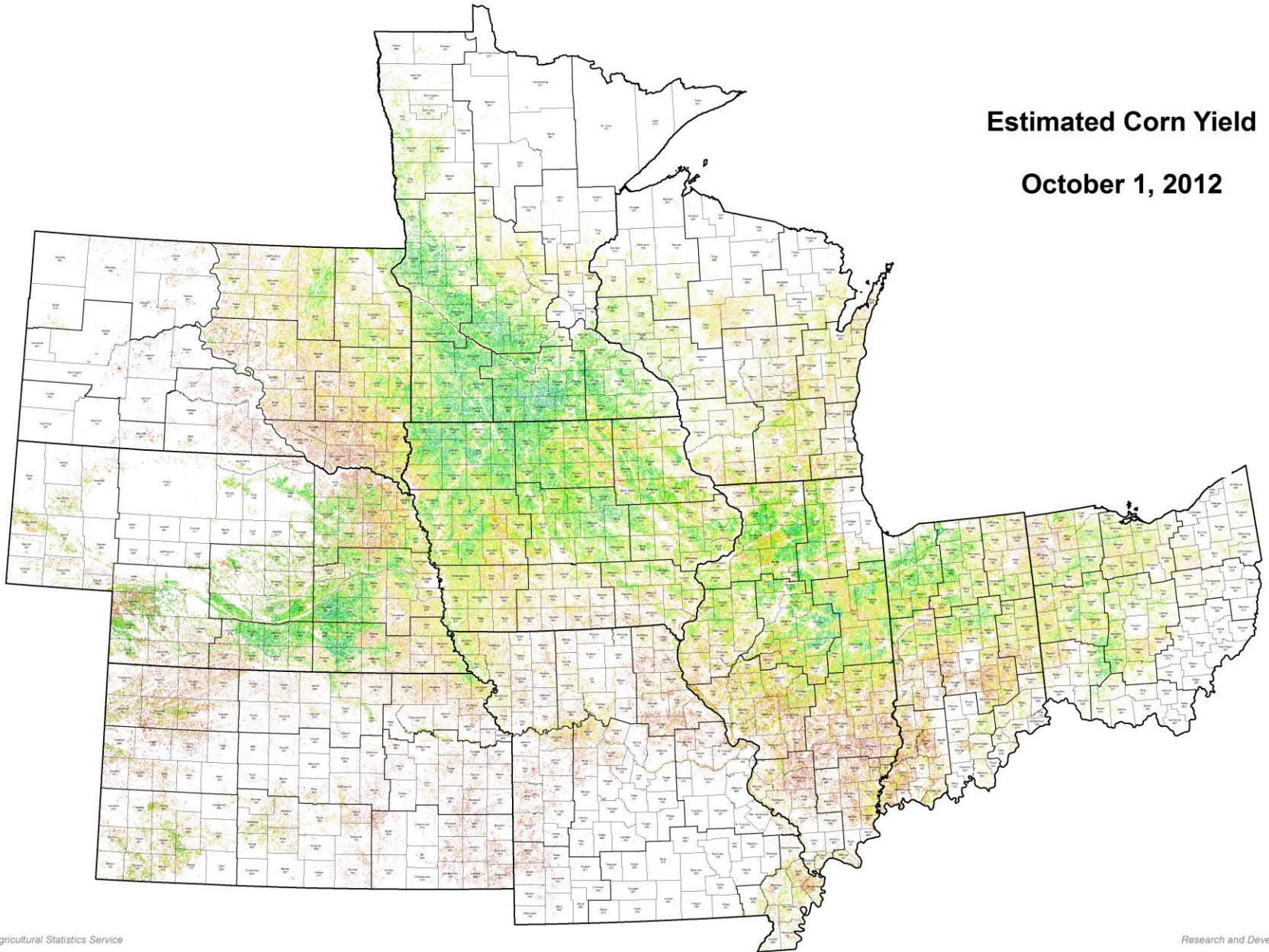


Normalized Difference Vegetation Index (NDVI)



Estimated Corn Yield

October 1, 2012

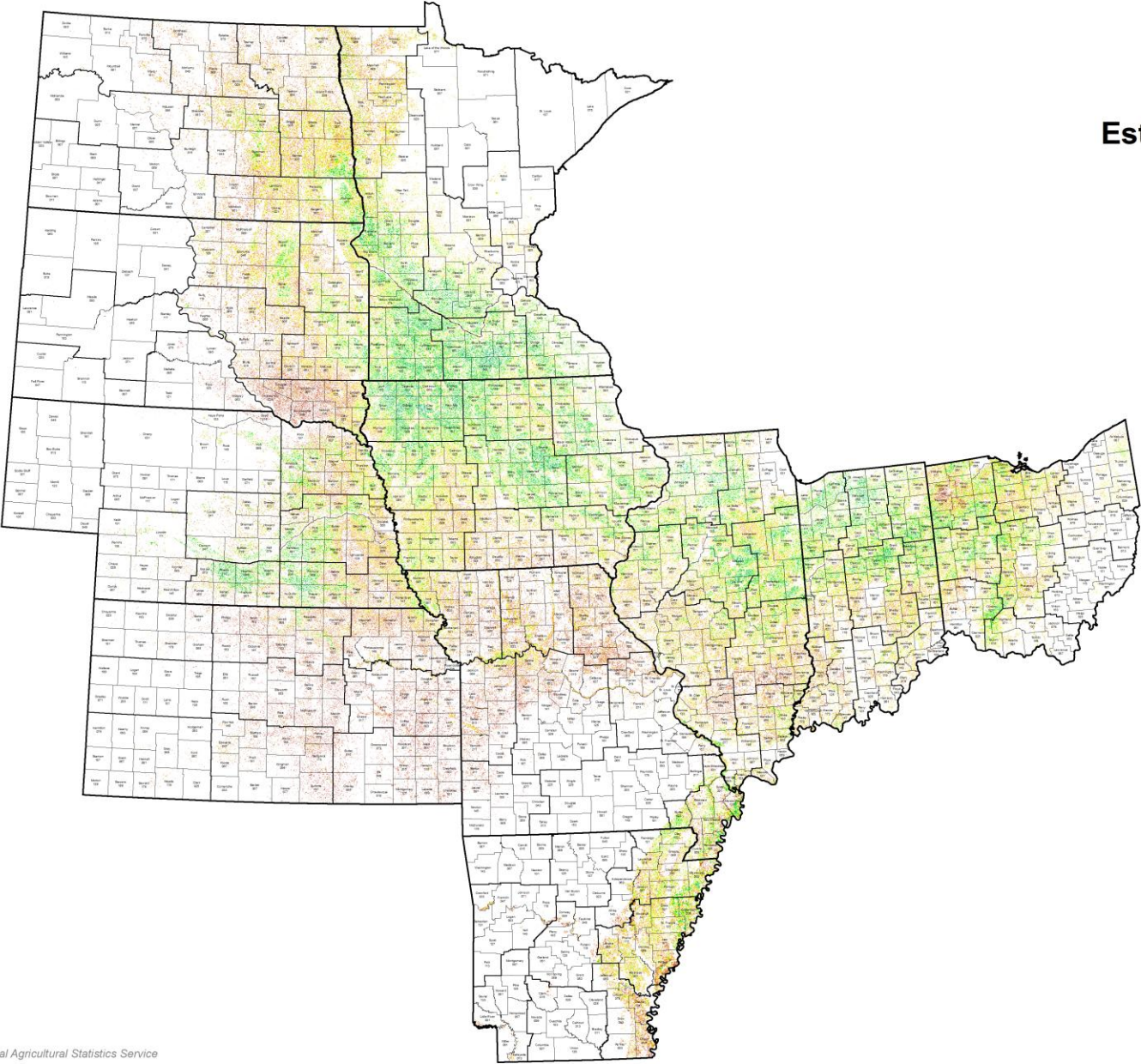


USDA/National Agricultural Statistics Service

Research and Development Division

Estimated Soybean Yield

October 1, 2012

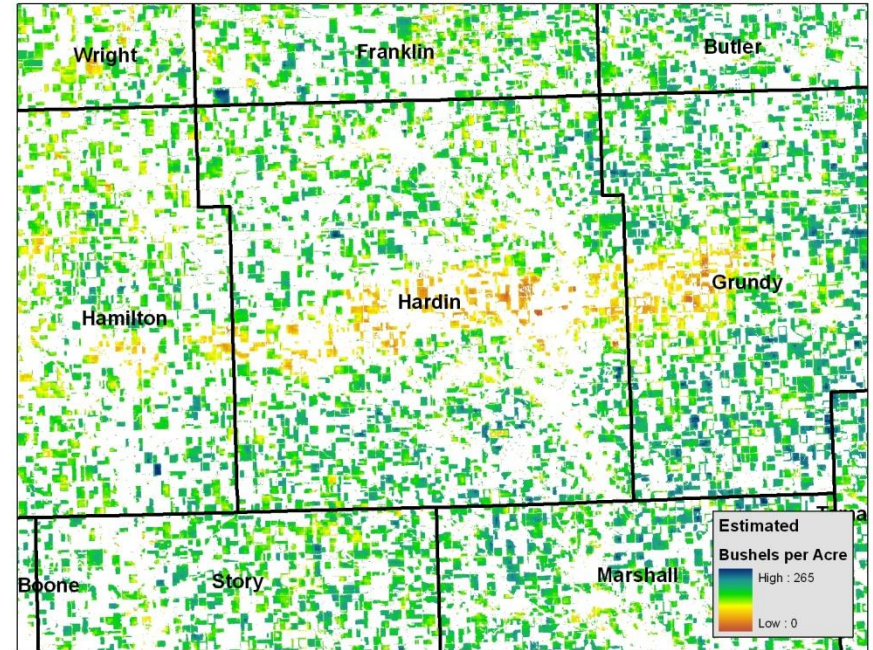
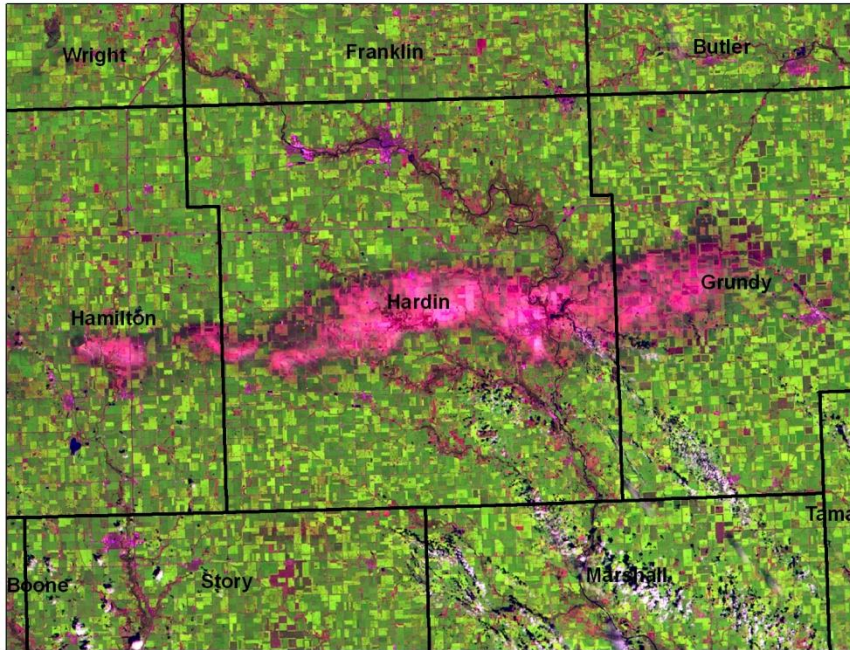


USDA/National Agricultural Statistics Service

Research and Development Division

Reality check

Scene of a large hailstorm



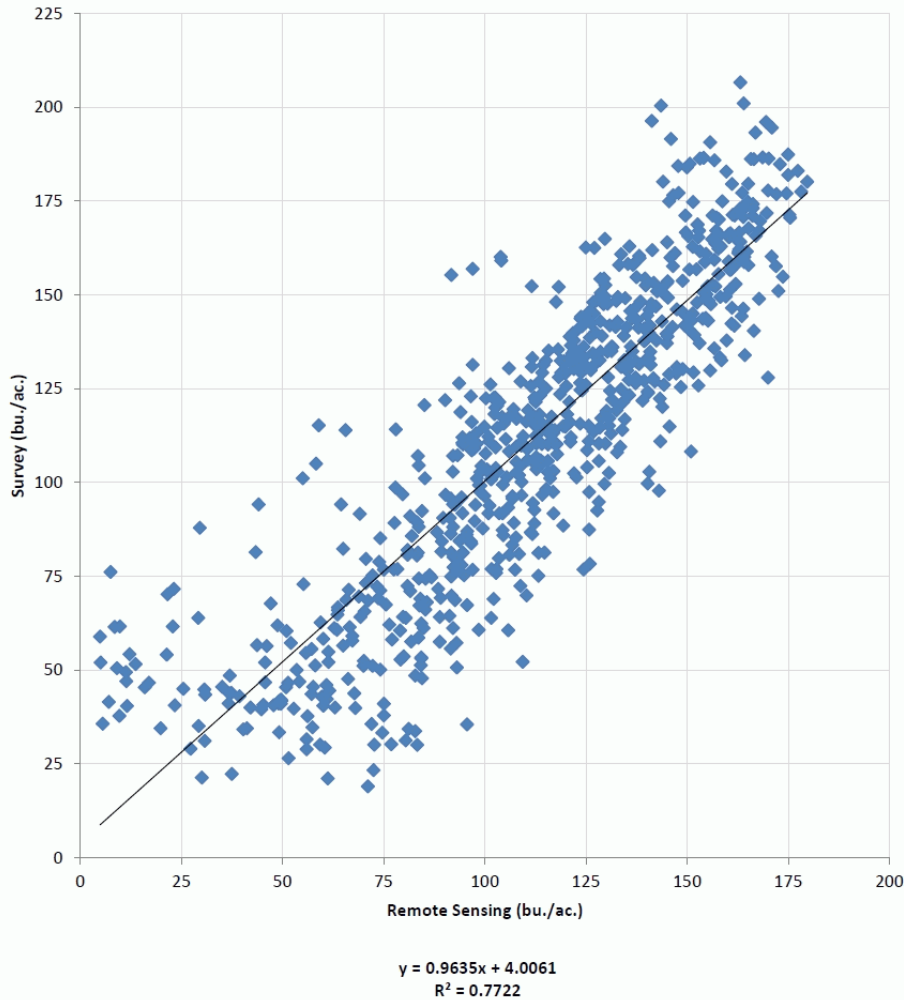
2012 Results: Remote sensing vs NASS yield

State level average error

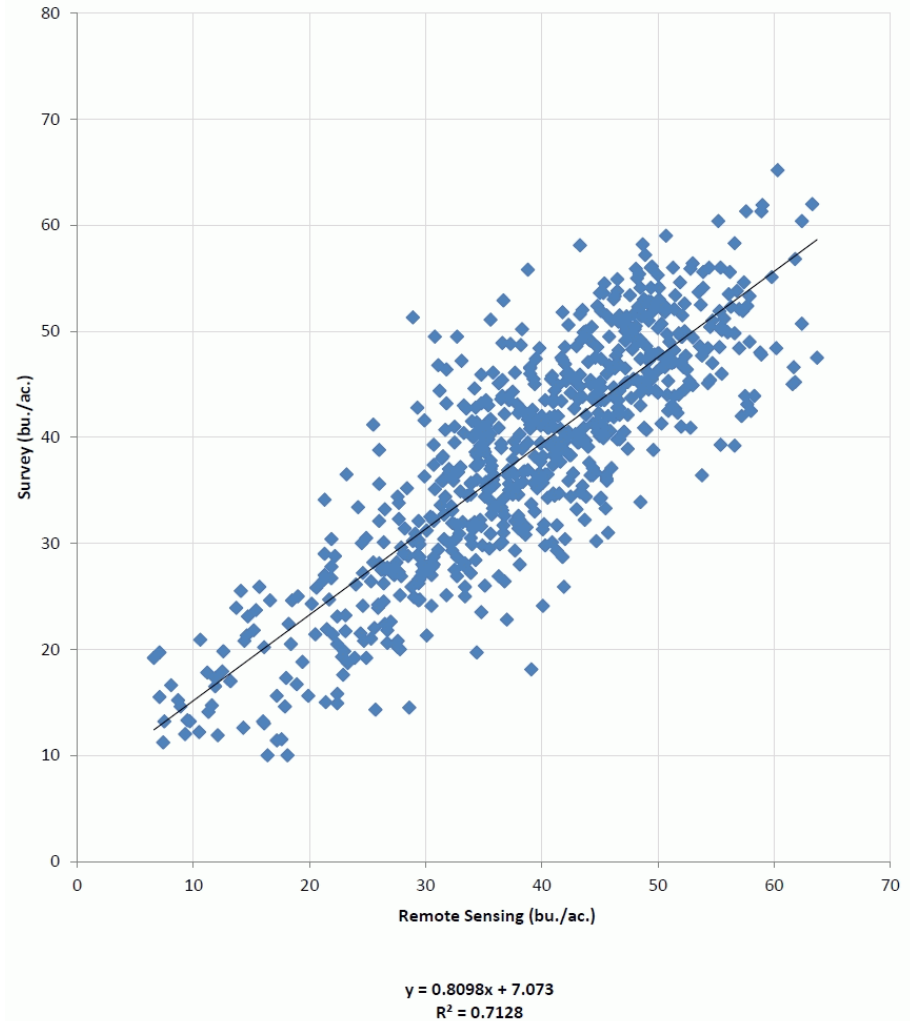
corn = 5.8 bu./ac.

soybeans = 3.1 bu./ac.

2012 County Level Yield Comparison - Corn



2012 County Level Yield Comparison - Soybeans



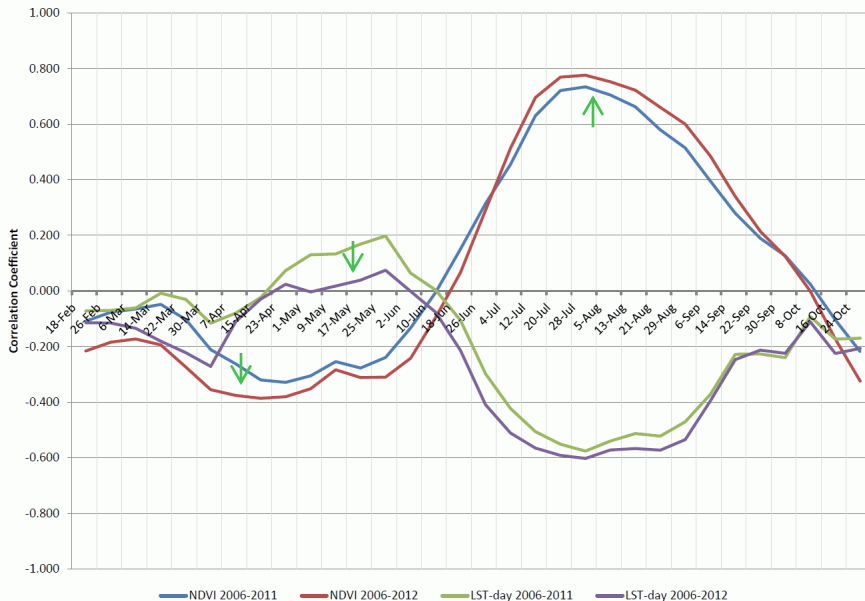
Models improvements for 2013

- Corn

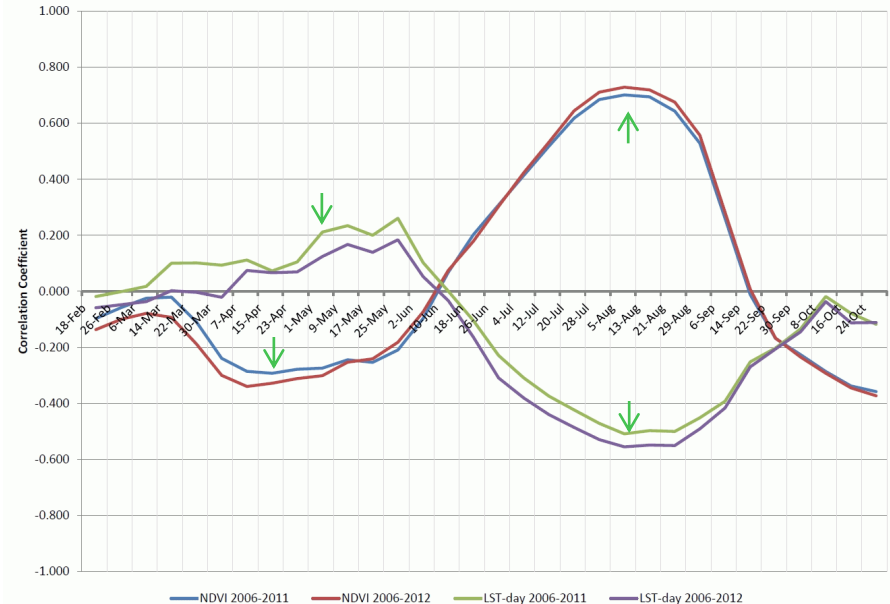
– relative err	2012	2013
– correl coeff	0.33	0.30
– correl coeff	0.93	0.95
- Soybeans

– relative err	2012	2013
– correl coeff	0.31	0.30
– correl coeff	0.93	0.94

Speculative Region Corn



Speculative Regions Soybeans



- Absolute error unchanged
 - ~8.0 bu/ac for corn, ~2.5 for soybeans

The end

