



# Enhancing Crop Insurance Program Integrity with Remote Sensing and Data Mining

Dr. Jim Hipple

Remote Sensing & GIS Advisor

USDA Risk Management Agency

Office of Strategic Data Acquisition & Analysis

Co Authors: Dr. Bert Little & Dr. Michael Schucking

Center for Agribusiness Excellence, Tarleton State University





# About the Risk Management Agency



- role is to help producers manage their business risks through effective, market-based risk management solutions
- promote, support, and regulate sound risk management solutions to preserve and strengthen the economic stability of America's agricultural producers
- operates and manages the Federal Crop Insurance Corporation (FCIC)
- provides crop insurance to American producers through 16 private-sector insurance companies sell and service the policies.

## FY 2007 Program Size

Number of Policies ..... 1.13 million  
 Premium Volume ..... \$6.55 billion  
 Crop Value Insured ..... \$67.2 billion\*  
 Acres Insured ..... 271 million  
 Data accurate as of September 25, 2007

## FY 2005 Program Size

Number of Policies ..... 1.19 million  
 Premium Volume ..... \$3.95 billion  
 Crop Value Insured ..... \$44.29 billion\*  
 Acres Insured ..... 246 million  
 Data accurate as of January 16, 2006

- RMA develops and/or approves the premium rate, administers premium and expense subsidy, approves and supports products, and reinsures the 16 companies
- sponsors educational and outreach programs and seminars on the general topic of risk management





## RMA's Goal



- Expand the use of geographical information, satellite imaging, and other technology as a means of effectively monitoring weather and other conditions that influence crop insurance payments.

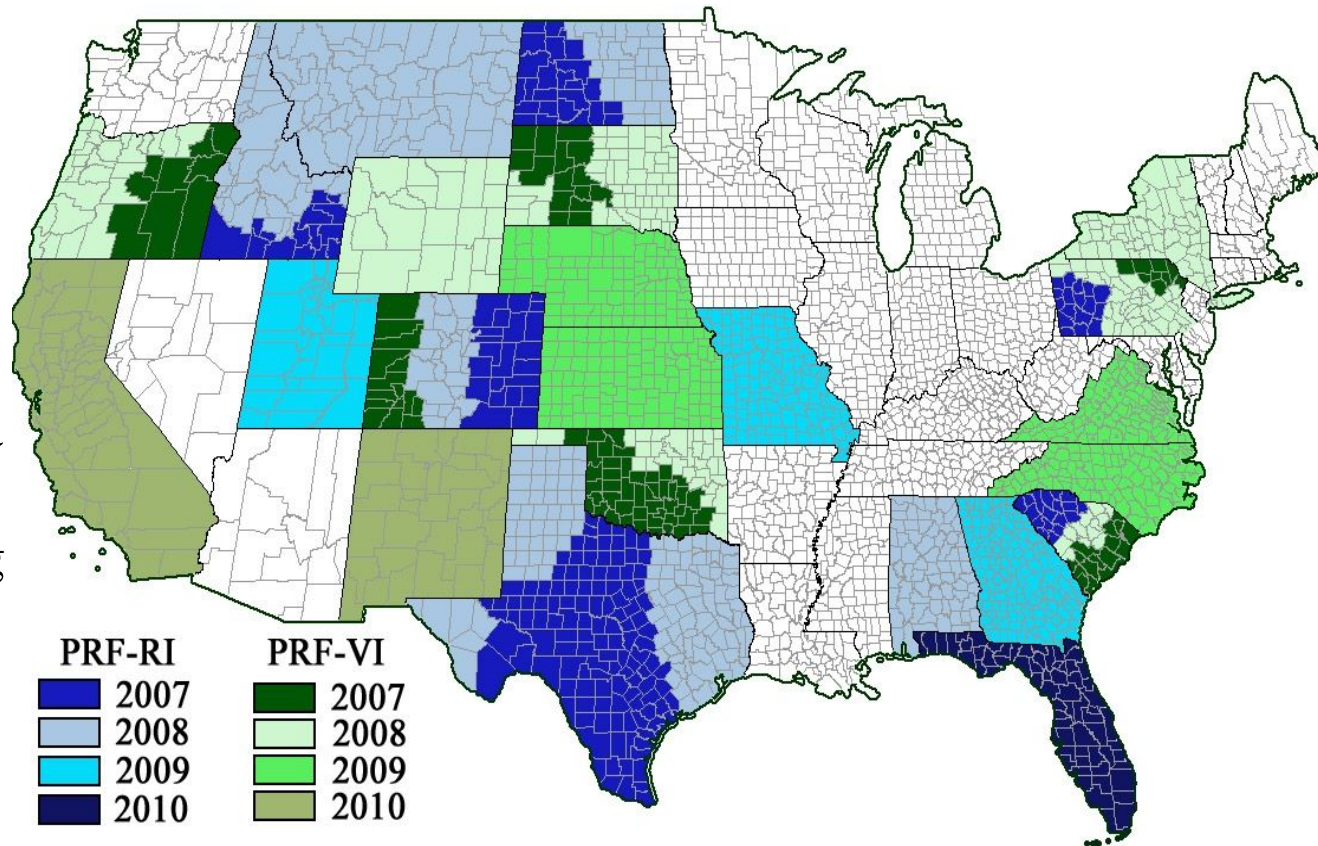




# Monitoring of RMA Pilot Programs



- Potential Pasture / Rangeland / Forage Pilot Program Expansion through 2010
- Ancillary data sources needed on a 12-month cycle for program monitoring

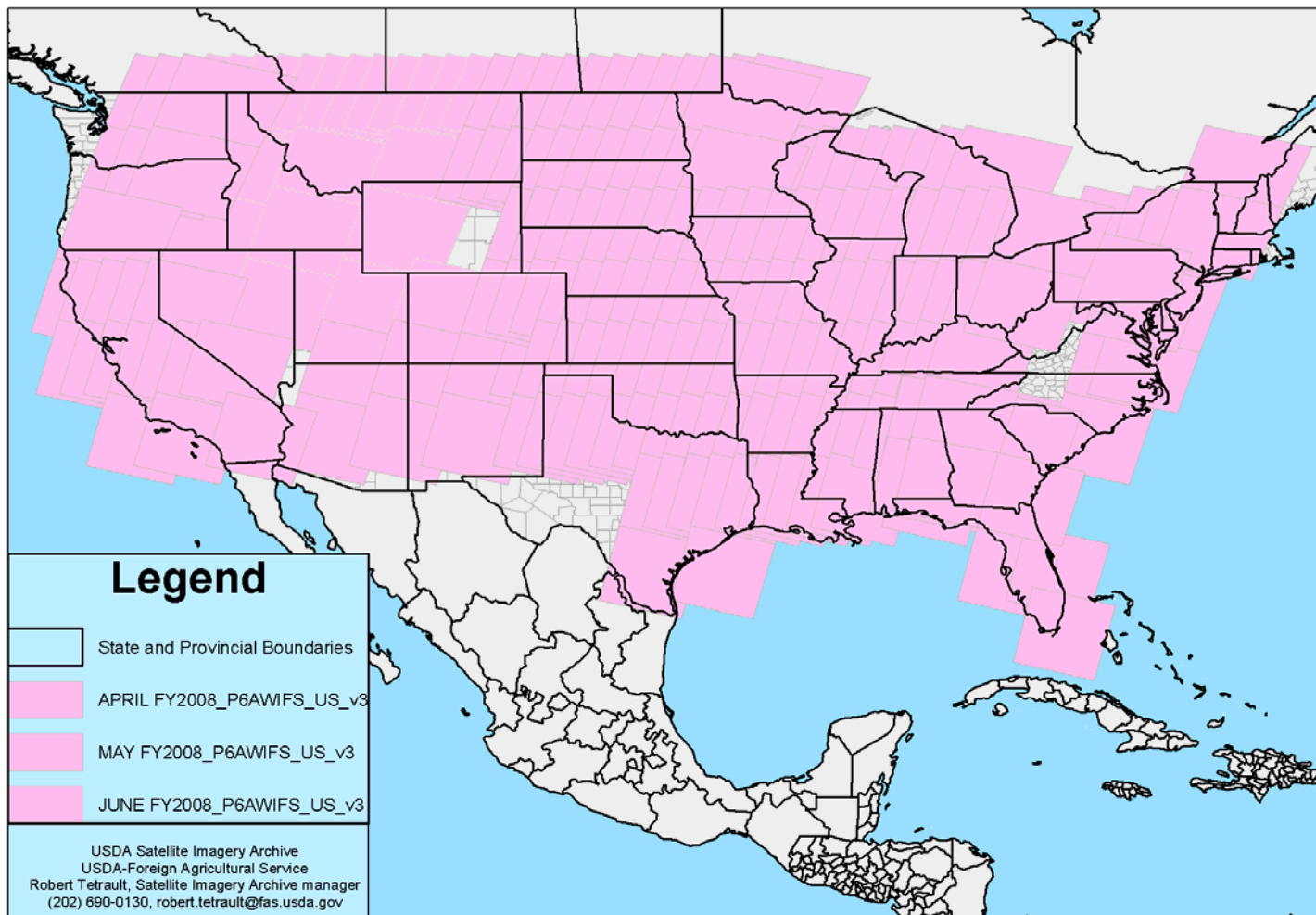




# Expansion of AWiFS Collection to meet RMAs Program Integrity Goals



## USDA's FY2008 Standing Order for P6-AWiFS CONUS





# **“Off Season” Collection Parameters (effective 10/01/2008)**





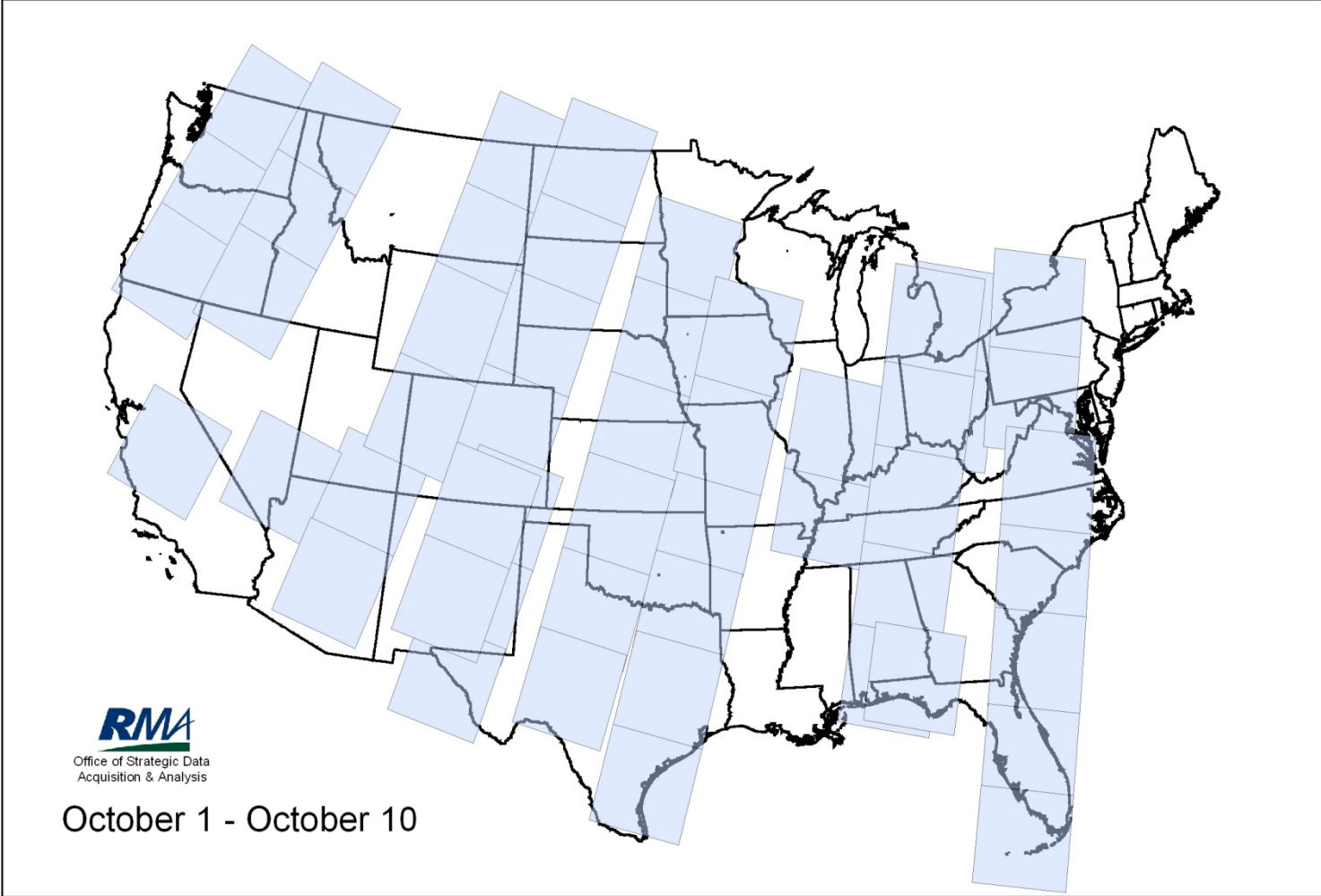
# Estimated AWiFS/LISS-3 Acquisitions



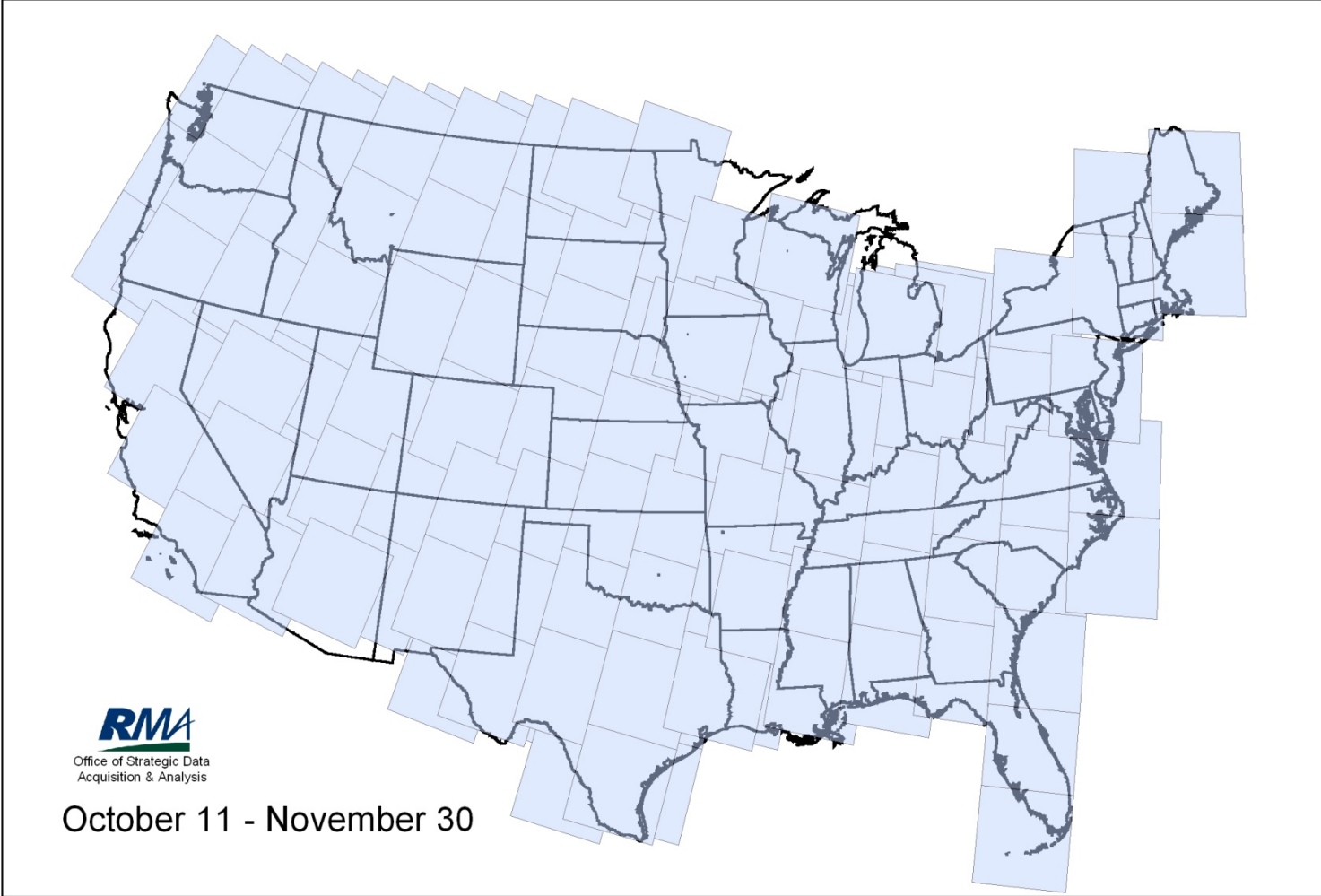
Coverage Area:	Sensor:	Processing Level:	Probable Purchases*
CONUS	P6-AWiFS	Ortho (56-m MS)	270
PRF Expansion Areas	P6-AWiFS	Ortho (56-m MS)	176
Hawaii	P6-LISS3	Ortho (23-m MS)	72
Southern Florida	P6-LISS3	Ortho (23-m MS)	8
Puerto Rico	P6-LISS3	Ortho (23-m MS)	8

Notes: Probable purchases assumes that ~50% of scenes will be not purchased because they are too cloudy.

# RMA “Off Season” AWiFS Collection



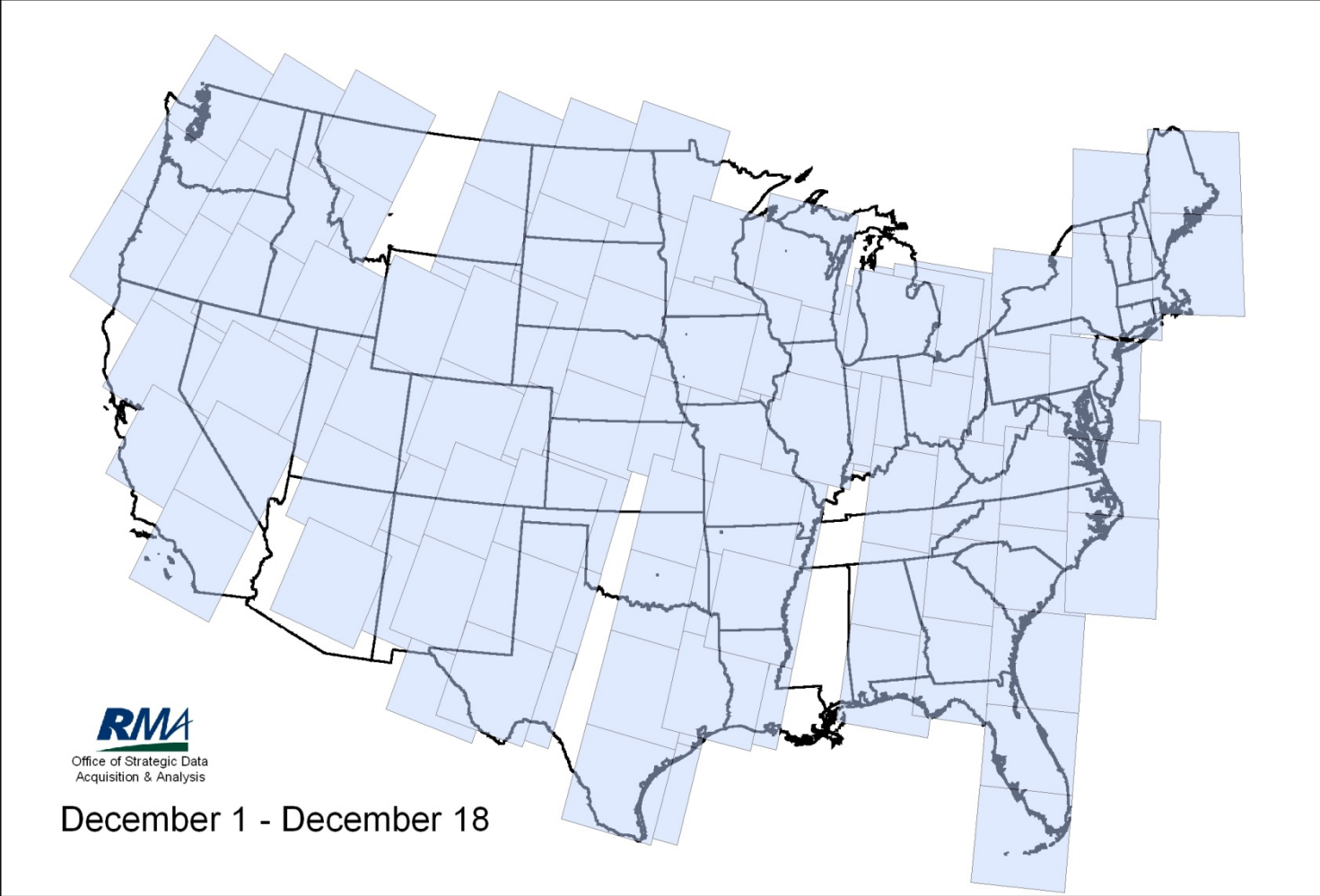
# RMA “Off Season” AWiFS Collection



Office of Strategic Data  
Acquisition & Analysis

October 11 - November 30

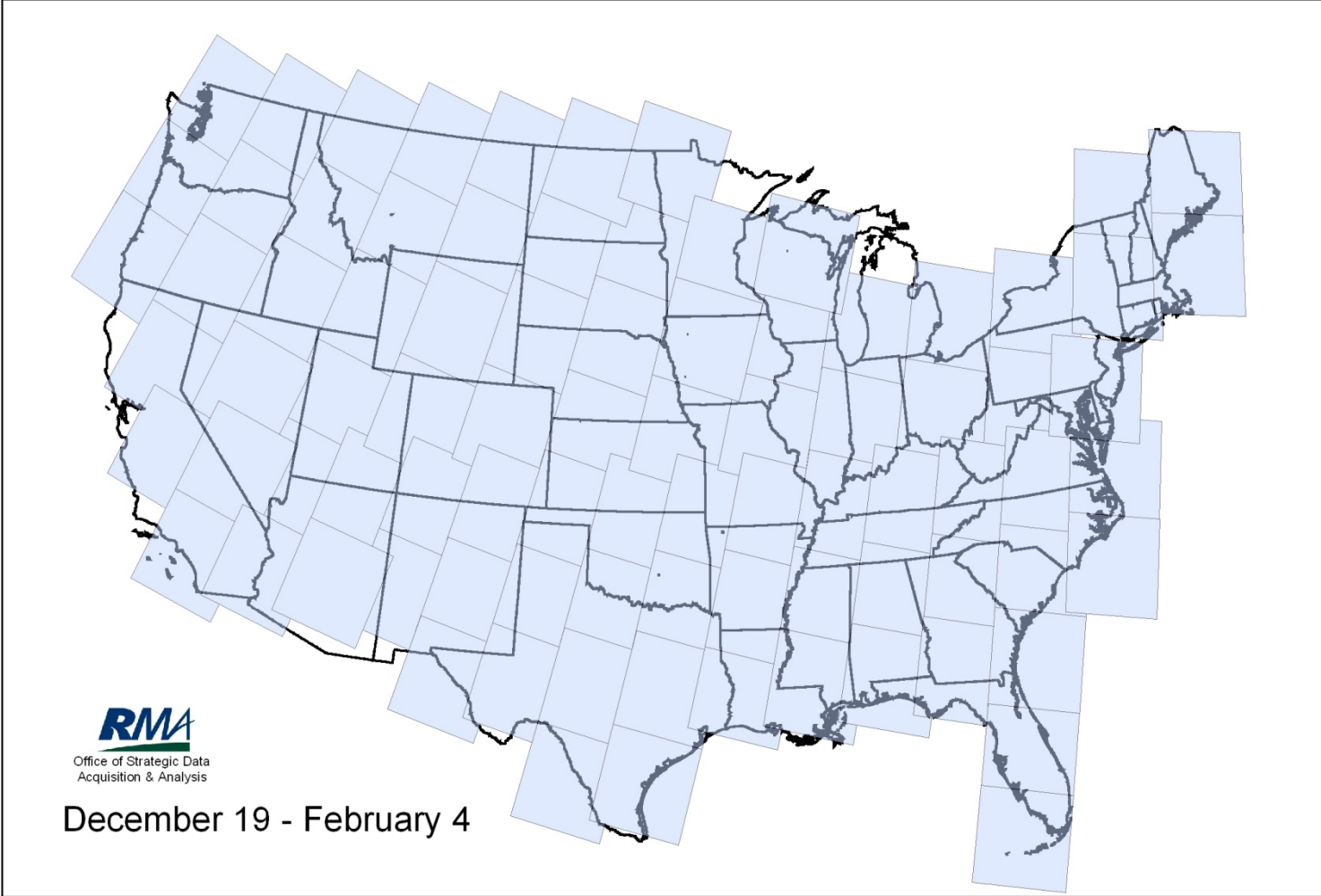
# RMA "Off Season" AWiFS Collection



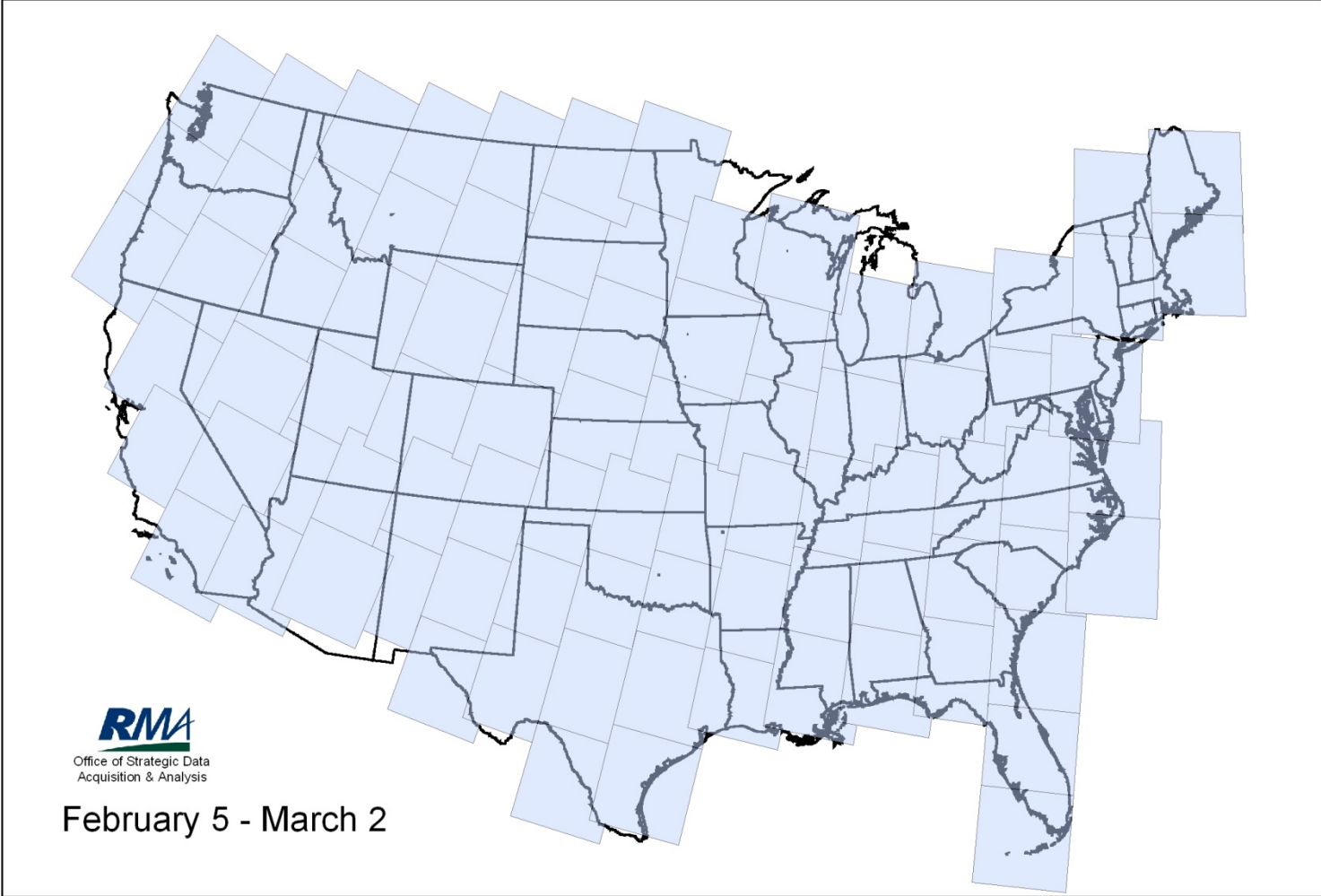
Office of Strategic Data  
Acquisition & Analysis

December 1 - December 18

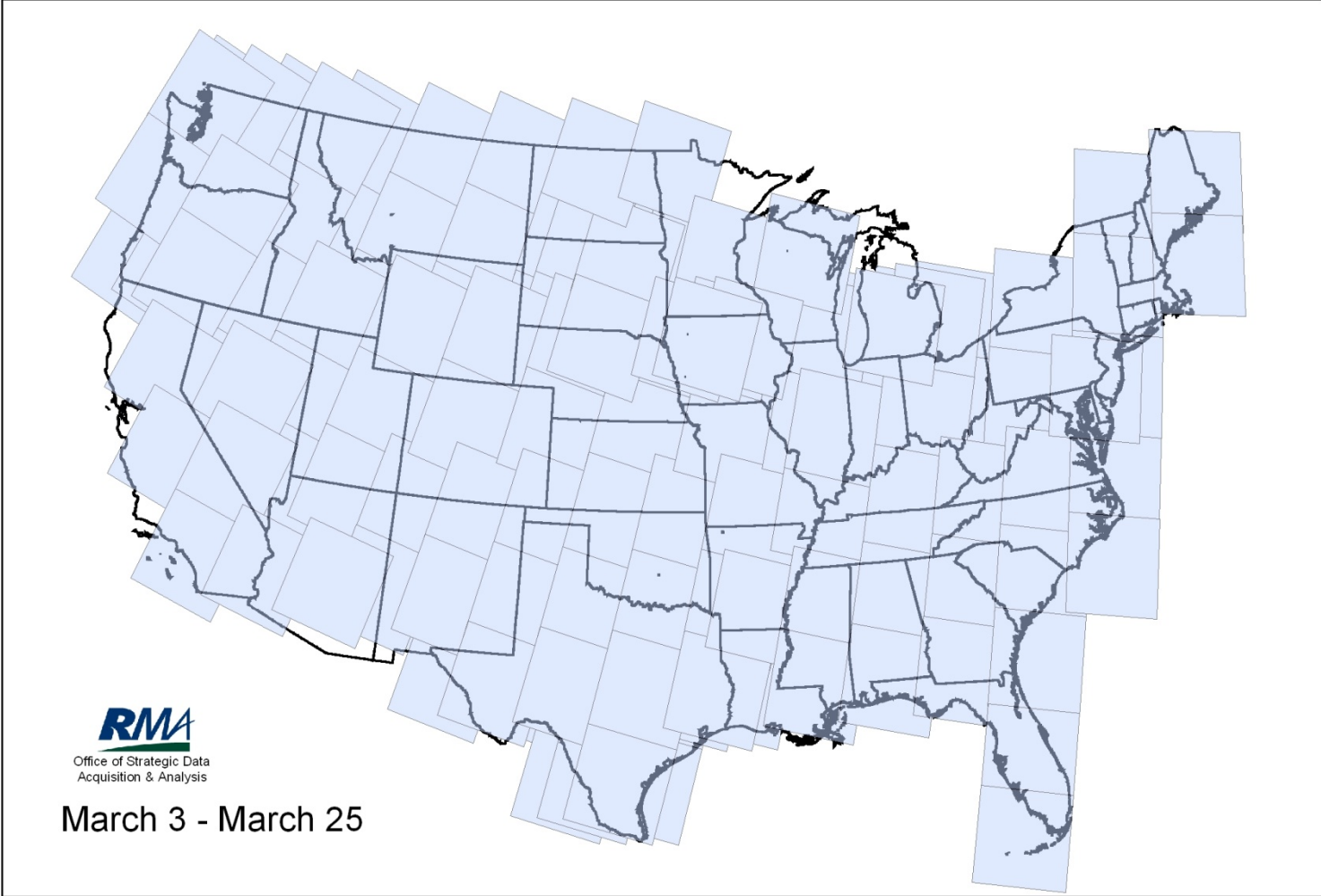
# RMA “Off Season” AWiFS Collection



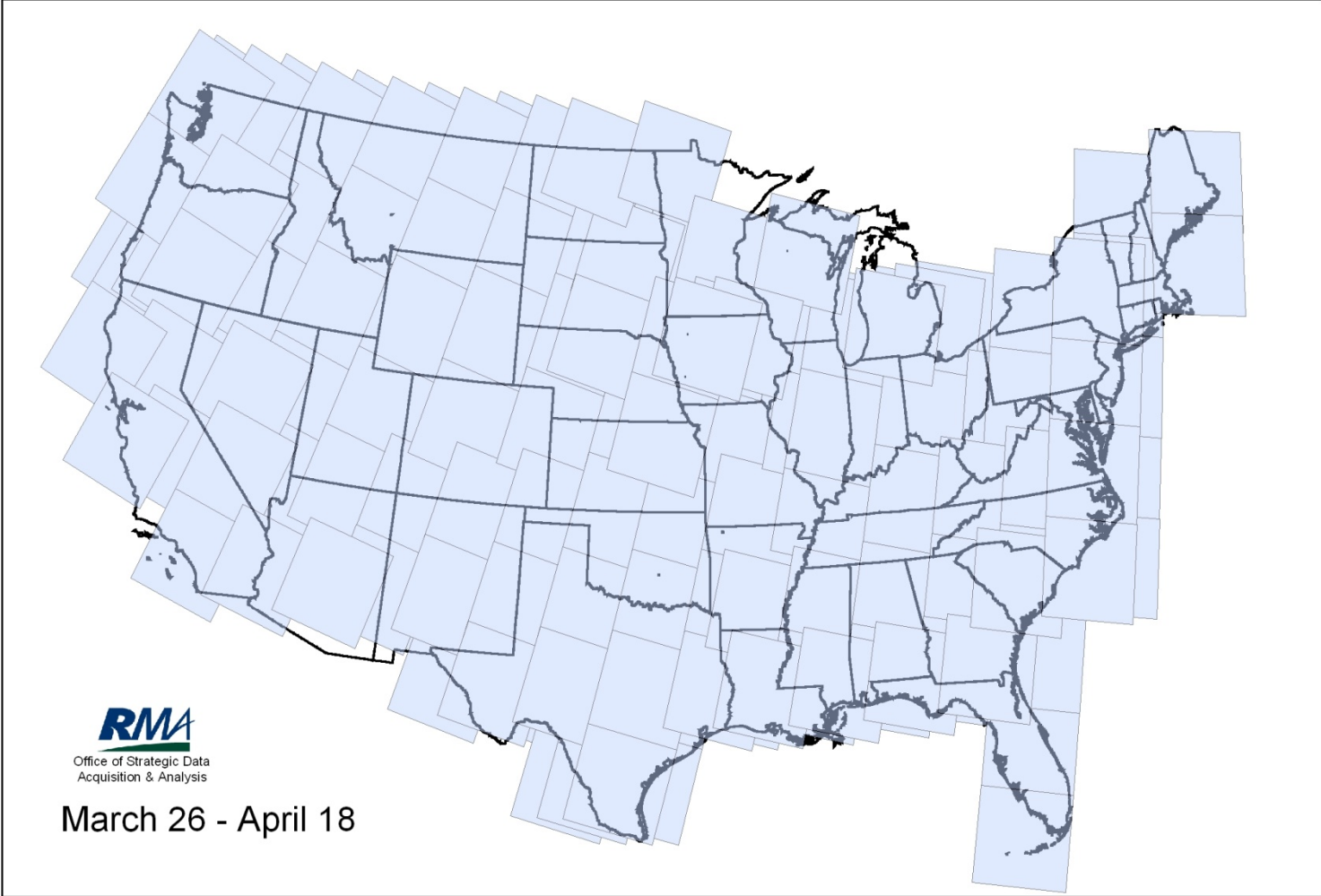
# RMA “Off Season” AWiFS Collection



# RMA “Off Season” AWiFS Collection

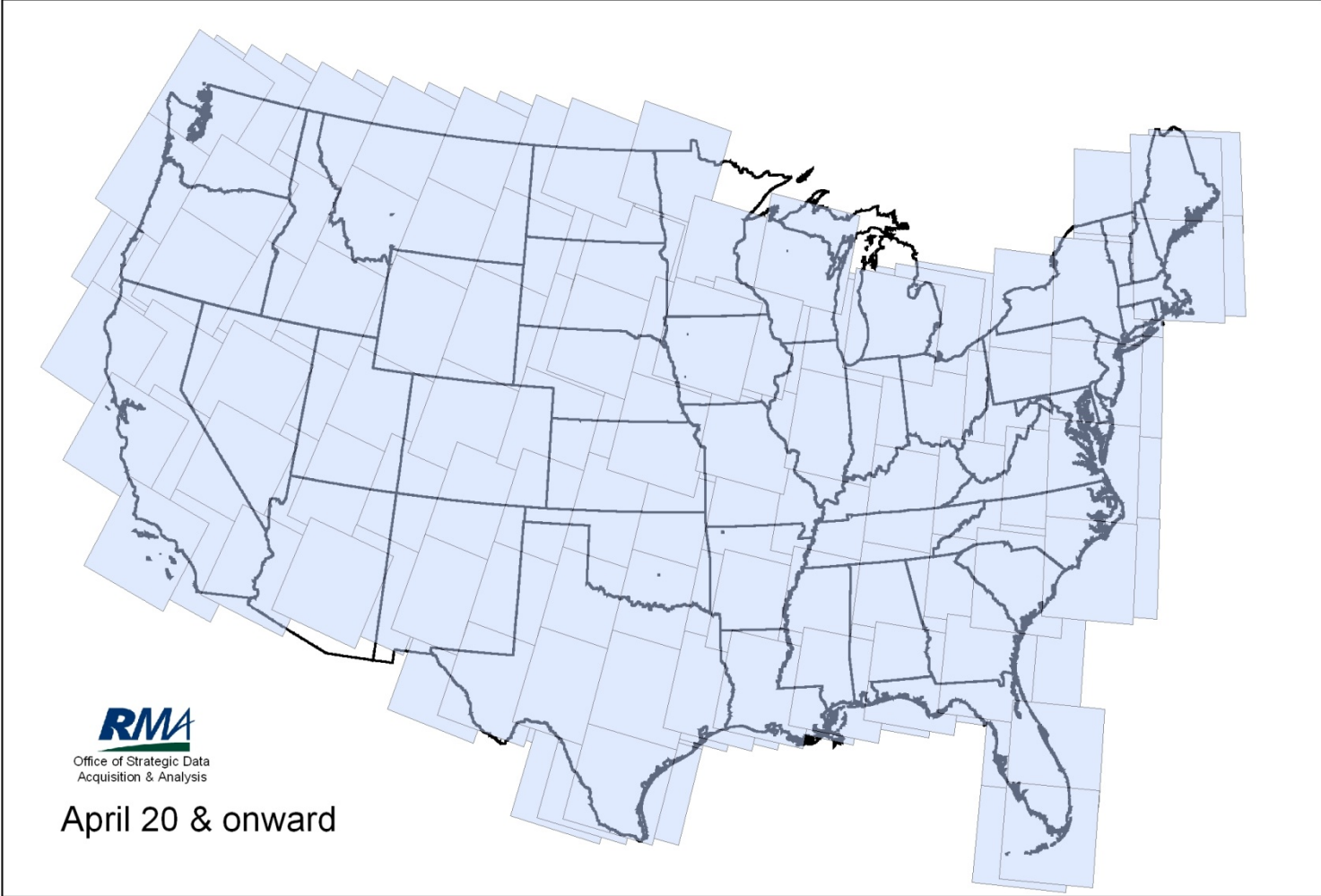


# RMA “Off Season” AWiFS Collection





# RMA "Off Season" AWiFS Collection





# Continued Processing of AWiFS





## RMA Processing Goals



- RMA/SDAA has an extensive KDD operation used to analyze patterns in crop insurance policies for increasing program integrity
- the purpose is to develop automated / semi-automated procedures to incorporate moderate resolution satellite imagery into the KDD process
- the goal is to be able to provide field-level metrics throughout the growing season on crop health





## Process

- develop automated / semi-automated procedures to preprocess IRS AWiFS (and other satellite data)
  - preprocessed to Top-of-Atmosphere-Reflectance (TOA) or *% reflectance*
    - no correction for atmospheric scattering or absorption, atmospheric gases (water vapor and ozone) and aerosols
  - TOA selected because it is a quick, low/no cost implementation with little other inputs needed & can work within our environment



## Process, cont.

- after AWiFS is preprocessed, extract data for each unique field
  - field information: USDA FSA Common Land Unit (CLU)
  - constrains: size (given each AWiFS pixel is approximately 0.70 acres), shape of field
  - data table by day of year for NDVI, NDWI, LSWI with mean & variance measure captured for each field
- data in 8-bit format (2005-2007), 10-bit (2008+)
- orthorectified data usually available to RMA from USDA Satellite Image Archive within 1 day (at most, 2 days) after acquisition



# Preprocessing Implementation



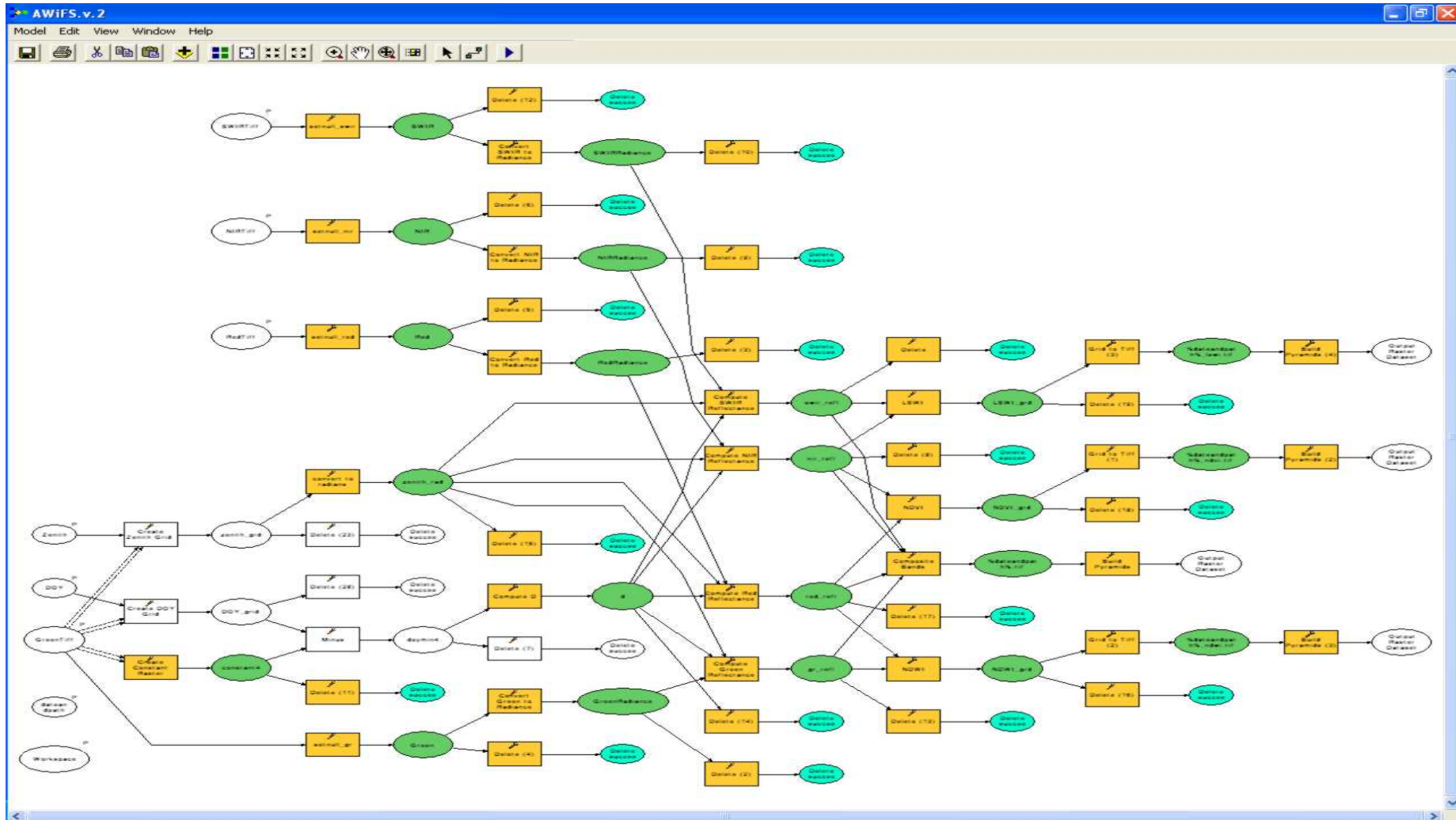


## Preprocessing Implementation

- developed in ESRI ArcCatalog ModelBuilder
- straightforward processing
- model could be used across USDA
- distributed as a ToolBox
  - **developed for AWiFS geotiff**, but can be adapted for Landsat 7 ETM+ geotiff, Landsat 5 TM geotiff , IRS ResourceSat LISS-3 geotiff



# Model Builder Preprocessing ToolBox





# Process and Results



The screenshot displays the ArcMap interface with several key components:

- ArcCatalog (Left):** Shows the file explorer with a path to 'Toolboxes\My Toolboxes\RMA Image Processing Tools'. The 'Contents' pane lists files: 'AWIFS.v.2', 'Landsat7\_geotiff.v.1', and 'LISS3.v.1'.
- Layers (Middle-Left):** A list of loaded layers including 'States', '20070402\_2510450.tif' (RGB), and three NDWI layers ('\_ndwi.tif', '\_ndwi.tif', '\_jswi.tif') with their respective value ranges.
- AWIFS.v.2 Tool Dialog (Bottom-Left):** A configuration window for the AWIFS.v.2 tool. It includes input fields for 'GreenTiff (optional)', 'RedTiff (optional)', 'NIRtiff (optional)', and 'SWIRtiff (optional)'. It also has checkboxes for 'Workspace', 'Zenith', 'dateandpath', and 'DOY'. The 'Help' tab is active, displaying the tool's description: 'Compute radiance, reflectance, ndvi and ndwi. Combine reflectance grids into a single image.'
- Main Map (Right):** A satellite image of California with state boundaries and names (Kern, San Bernardino, Los Angeles, Riverside, San Diego, Imperial, Kern, Mohave, La Paz, Yuma, Pima) overlaid. A red-shaded area is visible in the southern part of the state, corresponding to the '20070402\_2510450.tif' layer.

# Automation of Processing



The screenshot displays the RMA Automator software interface. A red bracket on the left side of the main window is labeled "RMA Automator Extension". A red arrow points from this bracket to a "Set Parameters Form" dialog box. The dialog box contains the following fields and buttons:

Input Directory	<input type="text"/>	Browse
Output Root	<input type="text"/>	Browse
Archive Directory	<input type="text"/>	Browse
gdalinfo location	<input type="text" value="C:\gdalwin32-1.5\bin\gdalinfo.exe"/>	Browse

At the bottom of the dialog box are "Save" and "Cancel" buttons. A red arrow points from the "gdalinfo location" field to the text below.

Uses GDALInfo to extract parameters from the geotiff header and then batch the AWiFS for input into the Model

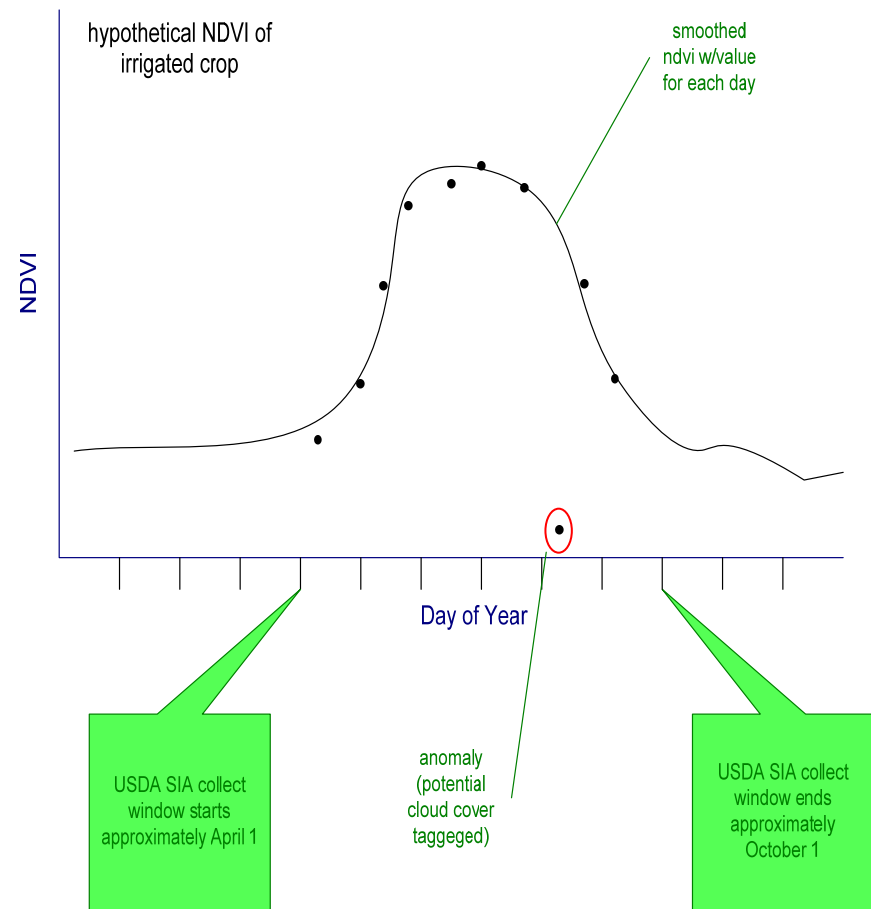
# Naming Conventions

- pull from CDINFO (or CDINFO.txt) (structure of data of the downloaded AWiFS)
- process names the files in this manner:
  - yyymmdd\_ppprrrqxxxx.tif
    - yyyy = year
    - mm = month
    - dd = day
    - ppp = path
    - rrr = row
    - q = quad (A, B, C, D)
    - xxxx = index type (ndvi, ndwi, lswi)
- example: 2007518\_263040b.tif; 2007518\_263040b(ndvi).tif; 2007518\_263040(ndwi).tif

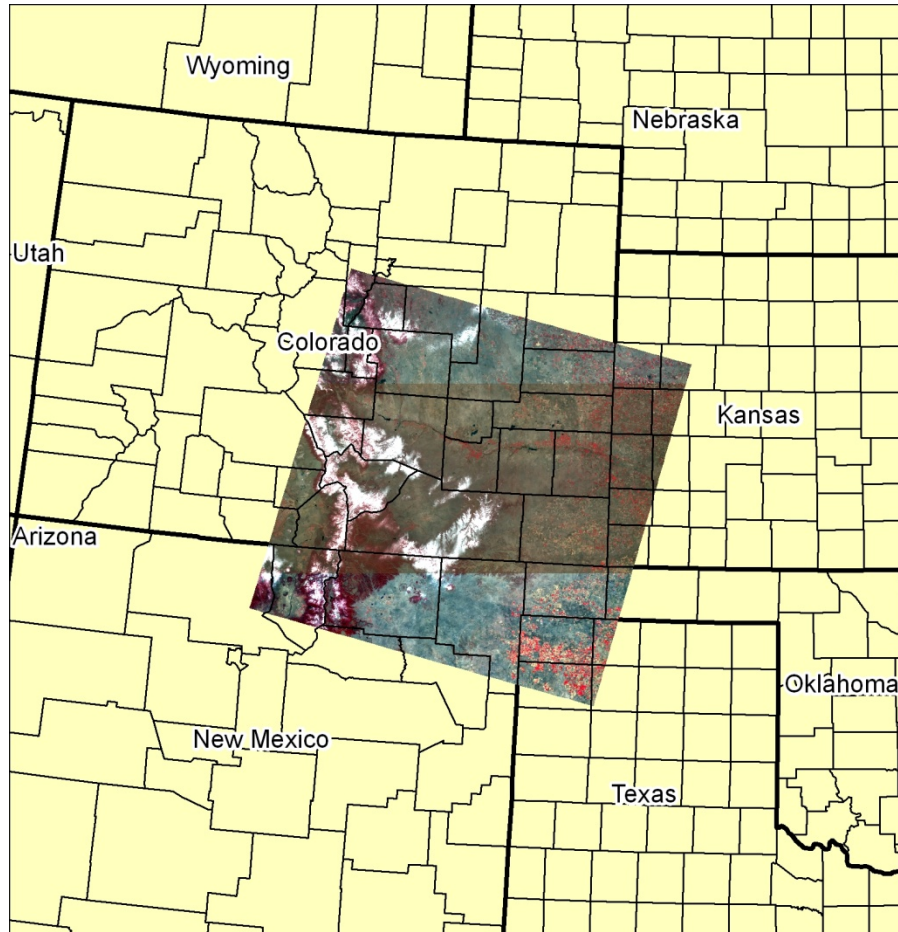
# Indices Generated



- vegetation index
  - NDVI (Normalized Difference Vegetative Index)
    - $NDVI = (nir - red) / (nir + red)$
- water index
  - NDWI (Normalized Difference Water Index)
    - $NDWI = (red - green) / (red + green)$
- land surface water index (irrigated / non-irrigated differentiator)
  - LSWI (Land Surface Water Index)
    - $LSWI = (nir - swir) / (nir + swir)$



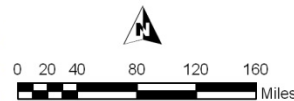
# Data Processing Examples



## AWiFS Overview

Acquisition Date: 04/18/2008  
Path: 265 Row: 045 Quad: A  
Bands: 3 (ir) / 2 (red) / 1 (green)

% Reflectance Image



# % Reflectance Image



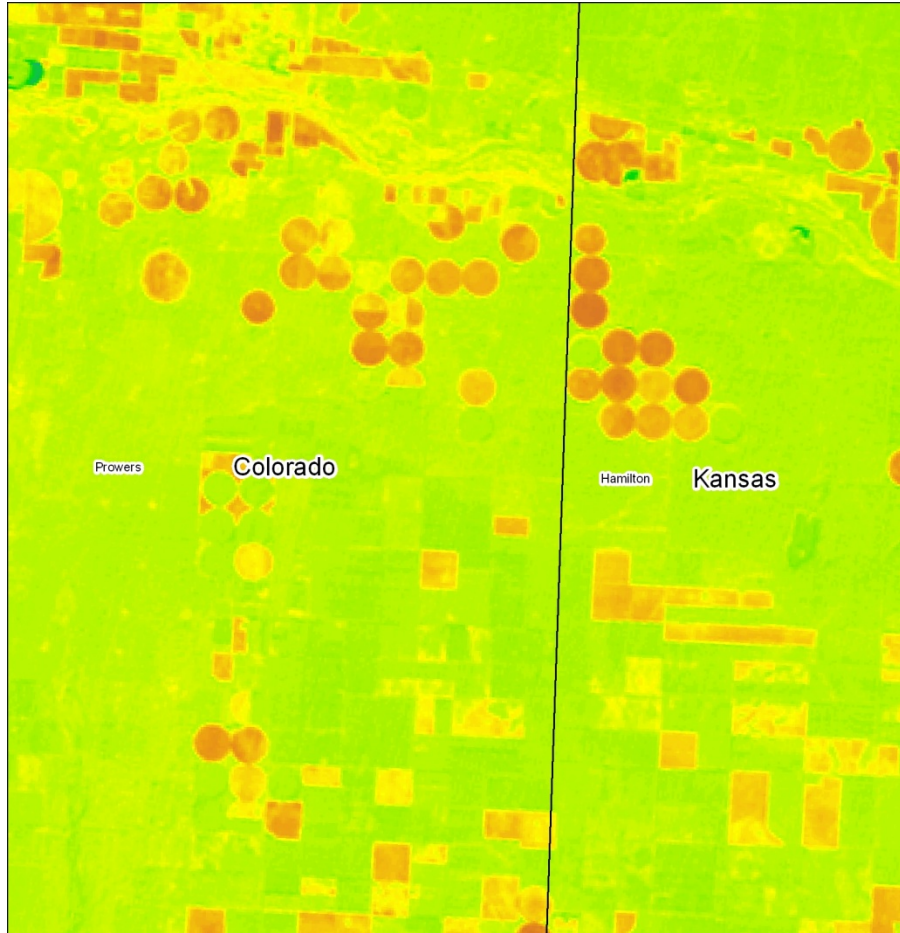
## AWiFS Overview

Acquisition Date: 04/18/2008  
Path: 265 Row: 045 Quad: A  
Bands: 3 (ir) / 2 (red) / 1 (green)

% Reflectance Image



# Normalized Difference Vegetation Index

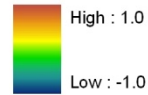


## AWiFS Overview

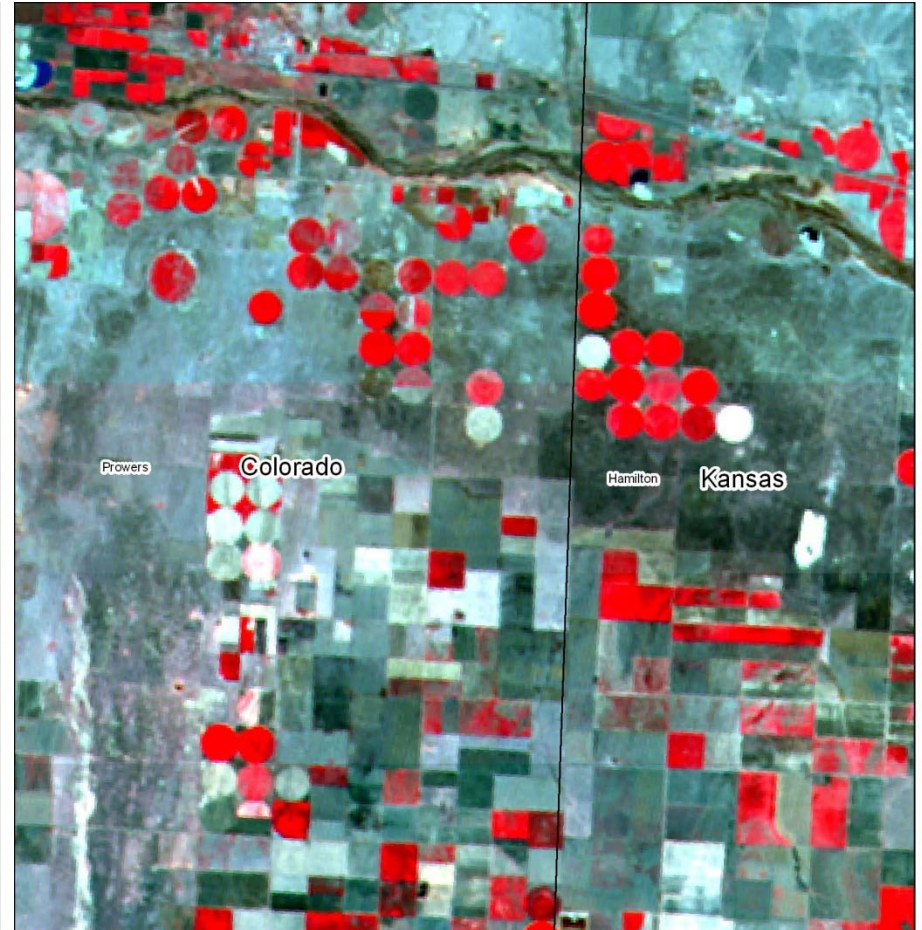
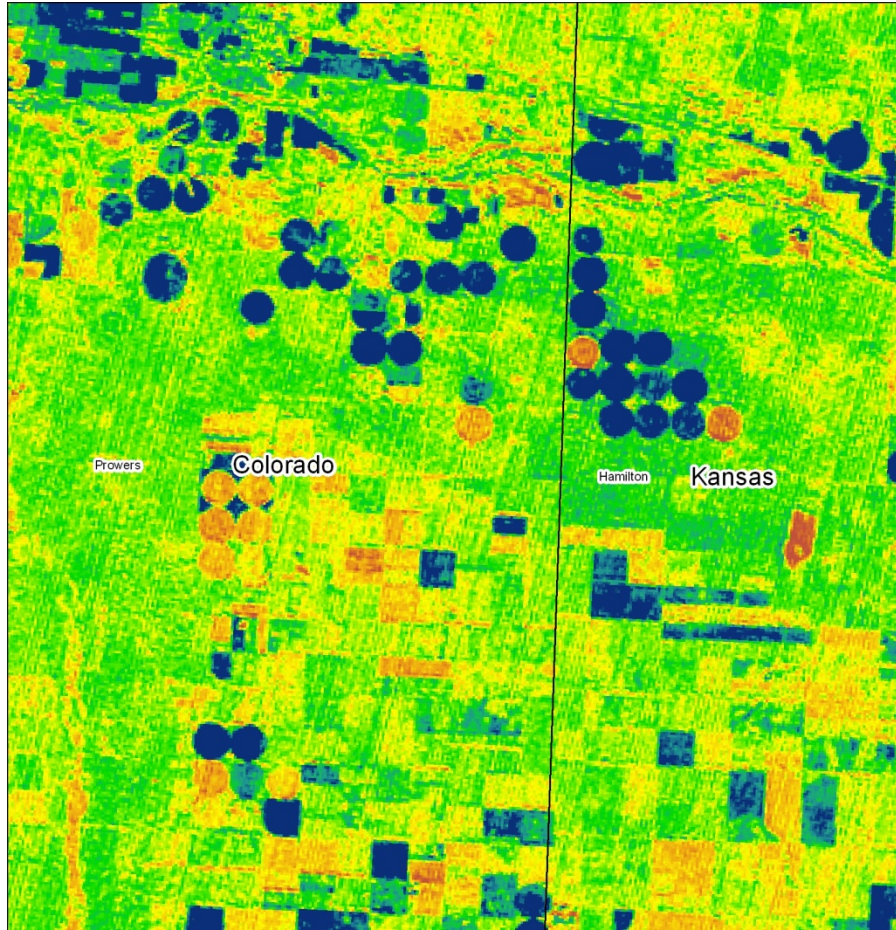
Acquisition Date: 04/18/2008  
Path: 265 Row: 045 Quad: A

NDVI Image

NDVI  
Value



# Normalized Difference Water Index







## Status

- ModelBuilder complete for AWiFS & LISS
- 50% of 2008 & 100% of 2005 US scenes AWiFS scenes processed by RMA
- 100% of 2006 - 2007 US scenes AWiFS scenes processed
  - by West Virginia University National Geospatial Development Center / NRCS under CREDA)
- **NEGATIVE:** single AWiFS scene takes 30-45 minutes to process
  - ArcGIS ModelBuilder – not that efficient!
    - Lack of support for multi-core, multi-processor under ESRI desktop products
- **POSTITIVE:** ModelBuilder models do not have the strict security review requirements in USDA of other applications that might be written (can be quickly deployed)



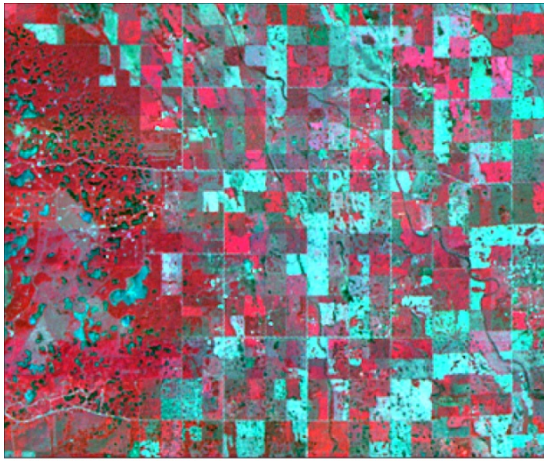


# **Extraction of Field Level Metrics / Integration into Data Mining (development ongoing)**

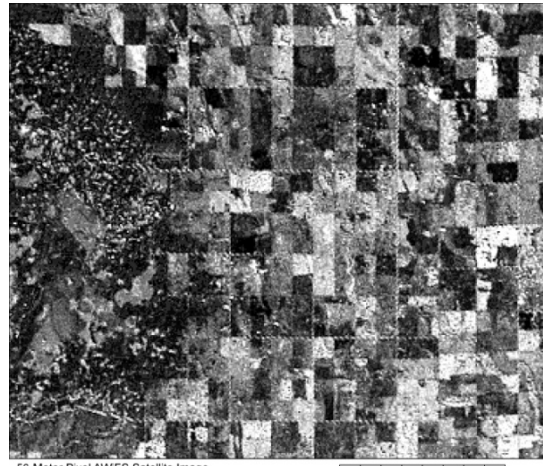




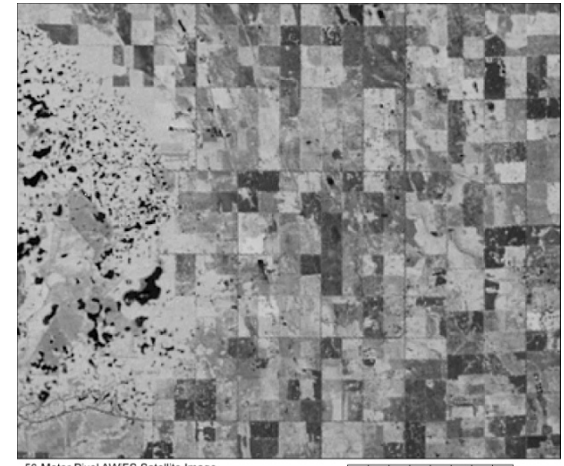
# Integrate Derived Products into Data Mining



56-Meter Pixel AWFS Satellite Image  
Displayed: Band 3-IR, 2-red, 1-green  
4-band layer-stacked geotiff in  
% reflectance with pyramids built



56-Meter Pixel AWFS Satellite Image  
Displayed: NDWI  
Normalized Difference Water Index  
 $NDWI = (red - green) / (red + green)$



56-Meter Pixel AWFS Satellite Image  
Displayed: NDVI  
Normalized Difference Vegetation Index  
 $NDVI = (nir - red) / (nir + red)$

Land Surface Water Index (LSWI)  
 $LSWI = (nir - swir) / (nir + swir)$



# Data Mining

## Starting the integration of RS data

### Current work:

- Use MODIS data to predict cotton yields in two highly homogeneous counties in west Texas
- analyze remotely sensed data variance in vegetative health in two counties (one mainly irrigated, one mainly non-irrigated) under moderate environmental stress
- analyze the ability of NDVI to predict county level yield across time, 2000 to 2006
- assess the ability of NDVI to predict yield on a day by day basis in 2006 at the farm sub-unit level

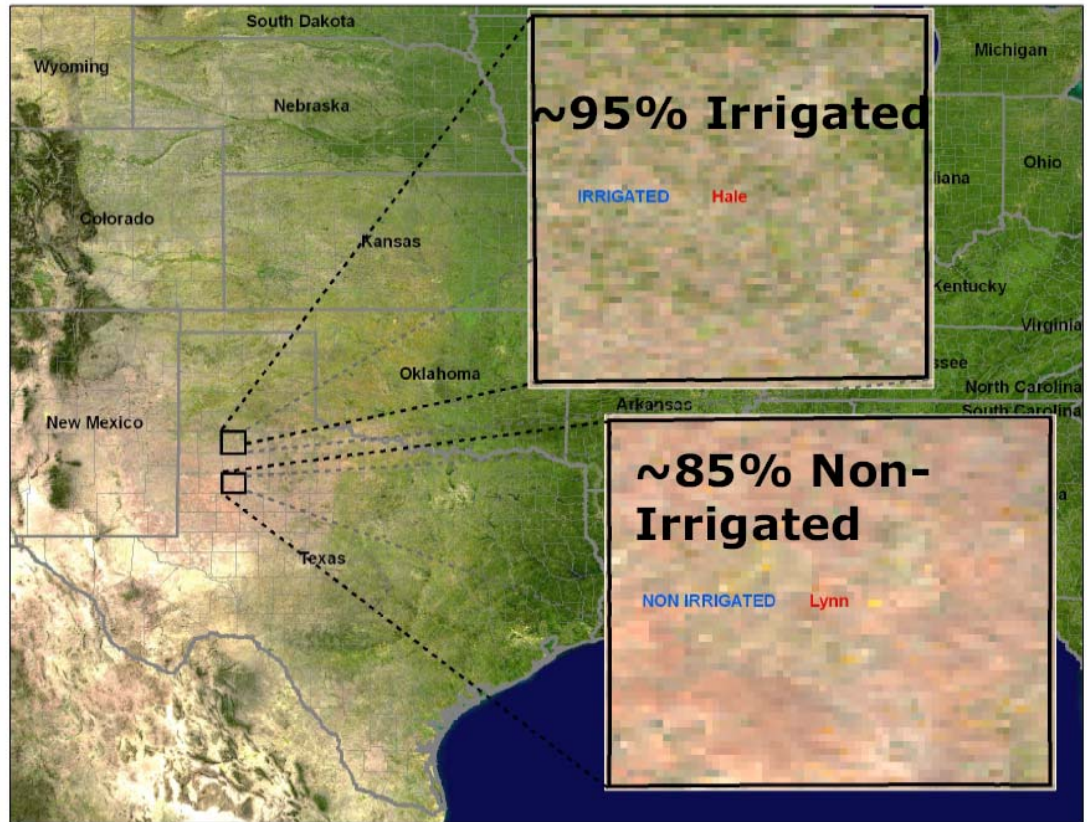
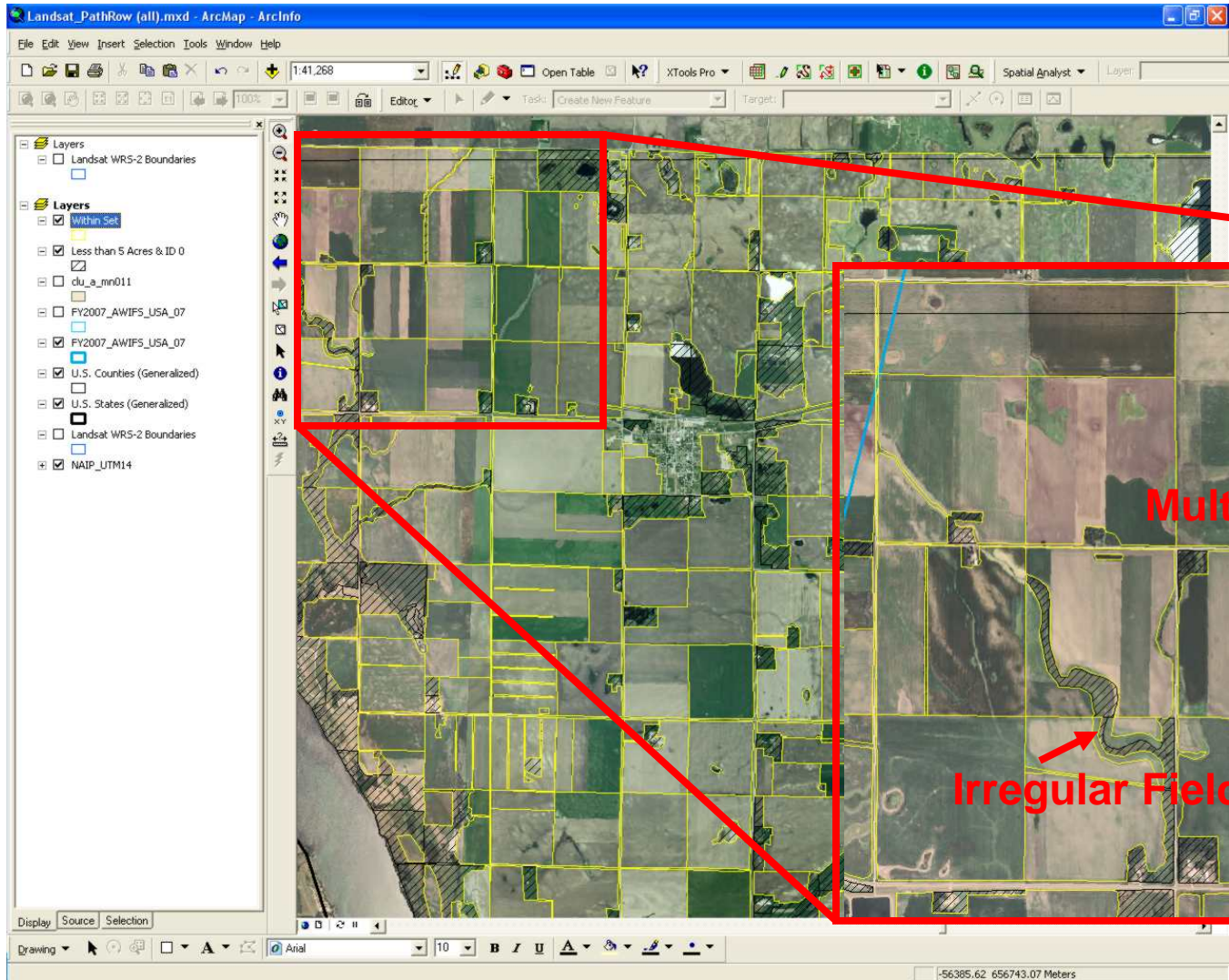


Figure 1. Counties in Texas Contrast: ~58.3% of Hale County land area (412,023 acres) is irrigated cotton agriculture (top), elevation 3200-3600 feet, 19.4" annual rainfall; and ~51.1% of Lynn County land area (570,835 acres) is dry land cotton agriculture (bottom), elevation 2881-3300 feet, 17.5" annual rainfall.

From: B Little, M Schucking, B Gartrell, B Chen, K Ross, and R McKellip (2008). "High Granularity Remote Sensing and Crop Production over Space and Time: NDVI over the Growing Season and Prediction of Cotton Yields at the Farm Field Level in Texas," *SSTDM 2008* (in press)

# CLU and Field Selection



CLU Problem Areas  
(due to spatial  
resolution of AWiFS)

Multi-Crop

Irregular Fields



# Metric Extraction & Future Direction



- working on the metric extraction procedure
  - select CLU that meet criteria of minimum size, shape
  - select CLU set that is within new image AWiFS footprint
  - calculate mean & variance values for indices & spectral bands for pixels within field boundary
  - develop ‘running’ smoothing procedure to fill in gaps
  - try to do this real-time or near real time
- look at near real time classification of crop-type cover on a per field basis
  - validate 2006 & 2007 with NASS Cropland Data Layer





## Thanks To ...

- West Virginia University National Geospatial Development Center / NRCS (Jim Thompson, Henry Ferguson & crew) for assistance on the AWiFS processing
- Bob Tetrault & Brad Doorn of FAS AND Elizabeth (Elsa) Woldemichael & Melvin Tucker of ASRC for assisting in streamlining data delivery through the USDA SIA
- the staff at ASRC, Global Marketing Insights, Inc. for putting together this forum

## Questions ...

- Dr. Jim Hipple, USDA Risk Management Agency  
james.hipple@rma.usda.gov

