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.
Determining Disaster Extent and Monitoring for Fraud

Tracking Reported Crops & Monitoring for Discrepancies

Integrating Remote Sensing, Geographic Information Systems (GIS), and Global Positioning Systems (GPS)

Forensic Reconstruction and Analysis of Crop Histories
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
“Off Season” Collection Parameters (effective 10/01/2008)
# Estimated AWiFS/LISS-3 Acquisitions

<table>
<thead>
<tr>
<th>Coverage Area</th>
<th>Sensor:</th>
<th>Processing Level:</th>
<th>Probable Purchases*</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONUS</td>
<td>P6-AWiFS</td>
<td>Ortho (56-m MS)</td>
<td>270</td>
</tr>
<tr>
<td>PRF Expansion Areas</td>
<td>P6-AWiFS</td>
<td>Ortho (56-m MS)</td>
<td>176</td>
</tr>
<tr>
<td>Hawaii</td>
<td>P6-LISS3</td>
<td>Ortho (23-m MS)</td>
<td>72</td>
</tr>
<tr>
<td>Southern Florida</td>
<td>P6-LISS3</td>
<td>Ortho (23-m MS)</td>
<td>8</td>
</tr>
<tr>
<td>Puerto Rico</td>
<td>P6-LISS3</td>
<td>Ortho (23-m MS)</td>
<td>8</td>
</tr>
</tbody>
</table>

Notes: Probable purchases assumes that ~50% of scenes will be not purchased because they are too cloudy.
RMA “Off Season” AWiFS Collection

October 1 - October 10
RMA “Off Season” AWiFS Collection

October 11 - November 30
RMA “Off Season” AWiFS Collection

December 1 - December 18
RMA “Off Season” AWiFS Collection

December 19 - February 4
RMA “Off Season” AWiFS Collection

February 5 - March 2
RMA “Off Season” AWiFS Collection

March 3 - March 25
RMA “Off Season” AWiFS Collection

March 26 - April 18
RMA “Off Season” AWiFS Collection

April 20 & onward
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
• 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
Automation of Processing

“RMA Automator Extension”

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:
  - yyyymmdd_ppprrqxxxx.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)
    - \(\text{NDVI} = \frac{\text{nir} - \text{red}}{\text{nir} + \text{red}}\)

- water index
  - NDWI (Normalized Difference Water Index)
    - \(\text{NDWI} = \frac{\text{red} - \text{green}}{\text{red} + \text{green}}\)

- land surface water index (irrigated / non-irrigated differentiator)
  - LSWI (Land Surface Water Index)
    - \(\text{LSWI} = \frac{\text{nir} - \text{swir}}{\text{nir} + \text{swir}}\)
% Reflectance Image

AWiFS Overview
Acquisition Date: 04/16/2008
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
- High: 1.0
- Low: -1.0

Proxers
Colorado
Kansas

0 0.5 1 2 3 4 Miles
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

4-band layer-stacked geotiff in % reflectance with pyramids built

Normalized Difference Water Index
$NDWI = \frac{\text{red} - \text{green}}{\text{red} + \text{green}}$

Normalized Difference Vegetation Index
$NDVI = \frac{\text{nir} - \text{red}}{\text{nir} + \text{red}}$

Land Surface Water Index (LSWI)
$LSWI = \frac{\text{nir} - \text{swir}}{\text{nir} + \text{swir}}$
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

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