How Can Remote Sensing Add to Crop Estimation?

1) Define a new, but completely known, total population
2) Compare the area sample to the known population
3) Create a relation between the sampled and known
   a) Simple Linear Regression
4) Apply the relation at any desired levels for estimation
   a) At the State level:
      - reduces variation of estimate(s)
   b) County, Sub-County, Watershed, ASD, …
      - With measurable variance estimates!
Sampling Approach Based on Area Sampling Frames
## Arkansas 2006

<table>
<thead>
<tr>
<th>Stratum</th>
<th># Segs Popn</th>
<th># Segs Sampled</th>
<th>Expans. Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>11669</td>
<td>232</td>
<td>50</td>
</tr>
<tr>
<td>21</td>
<td>2718</td>
<td>32</td>
<td>85</td>
</tr>
<tr>
<td>31</td>
<td>1308</td>
<td>4</td>
<td>327</td>
</tr>
<tr>
<td>32</td>
<td>418</td>
<td>2</td>
<td>209</td>
</tr>
<tr>
<td>42</td>
<td>18571</td>
<td>56</td>
<td>332</td>
</tr>
<tr>
<td>50</td>
<td>35</td>
<td>2</td>
<td>18</td>
</tr>
<tr>
<td>State</td>
<td>34719</td>
<td>328</td>
<td></td>
</tr>
</tbody>
</table>
JAS Questionnaire

- Enumerators account for all land usage in segment
  - Draw off field location by direct observation
  - Directly link questionnaire to segment photo
  - Able to ask questions not related to acreage
Analysis Concepts

**Direct Expansion (D.E.) = June Area Closed Estimator**

Value representing the sampling unit's contribution to the population total.

\[
D.E. = \text{Reported Data} \times \text{Expansion Factor}
\]

\[
DE = \sum \left( \frac{N}{n}x_i \right)
\]
A Basic Satellite Remote Sensing Approach

1) Locate known ground areas in the digital satellite imagery
   - ‘ground truth’ or ‘gtr’
2) Train a computer to recognize the ‘cover’ types
   - ‘signatures’ or ‘decision trees’
3) Classify the entire area of interest (State, region, …)
   - ‘known population’
4) Compare sampled areas to classification of same areas
   - ‘linear regression’
5) Create ‘revised’ estimates based on this relationship
Ground Truth Sources

• The June Area Survey itself
  – Digitize internal segment boundaries (fields)
• Analyst defined ‘extra’ categories

• Farm Service Agency administrative data
  – Match Common Land Units with 578 data
  – Select polygons with only one cover type

• National Land Cover Database
  – Non-crop categories
JAS Segment

R = Rice
S = Soyb
W = Waste/FS
CLU with Reported Crops and Landsat 30 Meter data

Renville County, MN  Landsat 5, 8/02/2000
Regression Estimator

- Regression used to relate categorized pixel counts to the ground reference data
  - Independent variable - satellite data - pixels
  - Dependent variable - JAS acreage estimate
- Satellite data - lower variance than with only JAS
- Outlier segment detection - correction or removal from regression analysis
<table>
<thead>
<tr>
<th>Crop Type</th>
<th>Y</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rice</td>
<td>227.0</td>
<td>273</td>
</tr>
<tr>
<td>Soybean</td>
<td>337.0</td>
<td>541</td>
</tr>
</tbody>
</table>

R = Rice  
S = Soyb  
W = Waste/FS
R2 = 0.971
a = intercept = 7.11
b = slope = 0.802
-------------------------------
Linear Regression
y = a + bx
-------------------------------
Seg 136 (x=273, y=227)
y = 226.11
R2 = 0.901
a = intercept = -49.93
b = slope = 0.8008

Linear Regression
y = a + bx

Seg 136 (x=541, y=337)
y = 383.30
Regression Estimator

\[ \hat{Y}_{ca(\text{reg})} = \sum_{h=1}^{H_a} N_{ah} [\bar{y}_{cah} + \hat{b}_{cah}(\bar{X}_{cah} - \bar{x}_{cah})] \]

- \( N_{ah} = \) Number of frame units in stratum \( h \) in analysis region \( a \)
- \( \bar{y}_{cah} = \) mean acres per segment from the June Area Frame Survey
- \( \bar{x}_{cah} = \) mean categorized pixel count (segments)
- \( \hat{b}_{cah} = \) coefficient from regression of acres on pixel counts
- \( \bar{X}_{cah} = \) mean categorized pixel count (scenes)

\( c = \) crop
\( a = \) analysis region
\( h = \) stratum
Acreage Estimation Methods

- Outlier segment detection - correction or removal from regression analysis
- Using pixel estimation in areas with satellite coverage but too few segments
- Using direct expansion/proration to fill in holes in satellite coverage (clouds, bad dates)
Ratio Estimators in General

**List and Area**
A ratio is an estimate of a mean, a proportion or a percent change in level. It usually takes the form of a ratio of two means or two direct expansions. The estimated ratio can then be multiplied by a "known base" to estimate a total. The general form is:

\[ Y = \frac{\bar{y}}{\bar{x}}X \]
Pixel Estimator

\[ \hat{Y}_{ca\text{ (pix)}} = \sum_{h=1}^{H_a} \lambda \left( \frac{m_{cal}}{m_{cak}} \right) X_{cah} \]

\( \lambda \) = Conversion factor (areal units per pixel)
\( m_{cal} \) = Marginal total pixels labeled to desired crop in segments
\( m_{cak} \) = Marginal total pixels categorized to desired crop in segments
\( X_{cah} \) = Number of pixels categorized to the desired crop

c = crop
a = analysis region
h = stratum
Direct Expansion

\[ \hat{Y}_{ca(dir)} = \sum_{h=H_a+1}^{H} \frac{N_{ah}}{n_{ah}} \sum_{j=1}^{n_{ah}} y_{cahj} \]

c = crop

a = analysis region

h = stratum

j = segment
Regression Analysis from Sample Estimation

Landsat TM Corn

\[ R^2(11) = .927 \]
\[ R^2(12) = .934 \]
\[ \text{Slope}(11) = .251 \]
\[ \text{Slope}(12) = .244 \]

AWiFS Corn

\[ R^2(11) = .834 \]
\[ R^2(12) = .854 \]
\[ \text{Slope}(11) = .709 \]
\[ \text{Slope}(12) = .745 \]

Slope of Acres/Pixels

\[ = 0.2224 \]

\[ = 0.7749 \]
Program Summary

Raw Satellite Images

Area Sampling Frame With PSU's

Segment Boundaries or CLU Polygons

JAS Questionnaire or 578 attributes

Categorized Images

Mosaicked CDL

Estimates