STRATIFICATION OF AN AGRICULTURAL AREA SAMPLING FRAME USING GEOSPATIAL CULTIVATION AND CROP PLANTING FREQUENCY DATA LAYERS

Claire G. Boryan, Zhengwei Yang and Patrick Willis
USDA National Agricultural Statistics Service (NASS)
Claire.Boryan@nass.usda.gov
OUTLINE

• RATIONALE
• BACKGROUND
• OBJECTIVES
• STUDY AREA
• DATA
• METHODOLOGY
• RESULTS/DISCUSSION
• CONCLUSIONS
• FUTURE RESEARCH
• REFERENCES
RATIONALE

• The National Agricultural Statistics Service (NASS)
  – provides timely, accurate and useful statistics in service to U.S. agriculture
  – collects and disseminates data on all facets of agriculture

• NASS conducts June Agricultural Survey (JAS) and other agricultural statistics programs based on NASS Area Sampling Frame (ASF) [1]

• The ASF stratification is based on percent cultivated cropland at the primary sampling unit (PSU) level

• However, percent cultivation information does not provide crop specific predictive planting information, which is critical for sampling and crop estimation

• To improve the ASF’s predictability, we investigated using Crop Planting Frequency Data Layers for ASF development
OBJECTIVES

• To develop a new ASF stratification with predicative crop planting information at the ASF Primary Sampling Unit level:
  
  – Develop three crop specific ASF stratifications using the NASS Geospatial Crop Planting Frequency Data Layers and the NASS Cultivated Layer derived from multi-year NASS Cropland Data Layers (CDLs) and
  
  – Evaluate them using 2014 Farm Service Agency (FSA) Common Land Unit (CLU) Data
BACKGROUND
NASS AREA SAMPLING FRAME (ASF)
BACKGROUND
ASF BASED ON NASS CULTIVATED LAYER [2-3]

ASF PSUs, with percent cultivation, overlaying the NASS cultivated layer

Each PSU labeled with a stratum category based on the state specific stratum definitions
STUDY AREA
SOUTH DAKOTA (SD) - U.S.A

- Large quantities of corn, soybeans, and wheat
- Greater than 95% coverage with Farm Service Agency Common Land Unit data for in situ validation and assessment of ASF performance
DATA

NASS SOUTH DAKOTA AREA SAMPLING FRAME

Roberts Co., SD

Detail - Primary Sampling Units (PSUs)
DATA

THE NASS CROPLAND DATA LAYER (CDL) [4]

The CDL product is an annually produced, 30 m, geo-referenced, crop specific, land cover data set.

Available on NASS’s CropScape
http://nassgeodata.gmu.edu/CropScape/
DATA

NASS CULTIVATED LAYER and CROP FREQUENCY DATA LAYERS [5] [6]
DATA
FARM SERVICE AGENCY (FSA) COMMON LAND UNIT (CLU) & 578 ADMINISTRATIVE “CROP” DATA
METHODOLOGY

CALCULATE MEAN CORN FREQUENCY (YEARS)

AT THE PSU LEVEL

(Using the 2008-2013 Corn Frequency Layer)

PSUs with corn frequency means over the 2008-2013 SD Corn Planting Frequency Data Layer

Each PSU labeled with a stratum category based on corn planting frequency
METHODOLOGY

CALCULATE PERCENT CULTIVATION AT PSU LEVEL
(Using the 2013 Cultivated Layer)

Area Frame: PSUs with percent cultivation based on the SD 2013 Cultivated Layer
K-means clustering was used on the mean corn planting frequency variable to create eight strata (same number as the current SD ASF)

<table>
<thead>
<tr>
<th>STRATA</th>
<th>NUMBER OF PSUs</th>
<th>CORN PLANTING FREQUENCY RANGE (Number of years out of 6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4634</td>
<td>0.0000 - 0.2558</td>
</tr>
<tr>
<td>2</td>
<td>2443</td>
<td>0.2560 - 0.6508</td>
</tr>
<tr>
<td>3</td>
<td>2497</td>
<td>0.6510 - 1.0550</td>
</tr>
<tr>
<td>4</td>
<td>2500</td>
<td>1.0551 - 1.4593</td>
</tr>
<tr>
<td>5</td>
<td>2268</td>
<td>1.4596 - 1.8645</td>
</tr>
<tr>
<td>6</td>
<td>1909</td>
<td>1.8651 - 2.2901</td>
</tr>
<tr>
<td>7</td>
<td>1544</td>
<td>2.2902 - 2.8491</td>
</tr>
<tr>
<td>8</td>
<td>356</td>
<td>2.8506 - 4.8648</td>
</tr>
</tbody>
</table>
SD corn planting frequency strata are sub-stratified by percent cultivation.

<table>
<thead>
<tr>
<th>STRATA</th>
<th>CORN PLANTING FREQUENCY RANGE (Number of years out of 6)</th>
<th>NUMBER OF PSUs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00-0.2558</td>
<td>4634</td>
</tr>
<tr>
<td>2</td>
<td>0.256-0.6508</td>
<td>2443</td>
</tr>
<tr>
<td>8</td>
<td>2.8506 - 4.8648</td>
<td>356</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>STRATA</th>
<th>NUMBER OF PSUs</th>
<th>PERCENT CULTIVATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>293</td>
<td>&gt;75%</td>
</tr>
<tr>
<td>1b</td>
<td>379</td>
<td>&gt;50 % - 75%</td>
</tr>
<tr>
<td>1c</td>
<td>1291</td>
<td>&gt;15 % - 50%</td>
</tr>
<tr>
<td>1d</td>
<td>2671</td>
<td>≥ 0.0 - 15%</td>
</tr>
<tr>
<td>Total</td>
<td>4634</td>
<td>N/A</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strata</th>
<th>NUMBER OF PSUs</th>
<th>PERCENT CULTIVATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>2a</td>
<td>214</td>
<td>&gt;75%</td>
</tr>
<tr>
<td>2b</td>
<td>376</td>
<td>&gt;50 % - 75%</td>
</tr>
<tr>
<td>2c</td>
<td>1768</td>
<td>&gt;15 % - 50%</td>
</tr>
<tr>
<td>2d</td>
<td>85</td>
<td>≥ 0.0 - 15%</td>
</tr>
<tr>
<td>Total</td>
<td>2443</td>
<td>N/A</td>
</tr>
</tbody>
</table>
RESULTS
CORN STRATIFICATION
RESULTS
SOYBEAN AND WHEAT STRATIFICATIONS
### RESULTS

2014 FSA CLU Mean Corn Acres/Percent Corn Increases from stratum 1 - 8

<table>
<thead>
<tr>
<th>STRATA</th>
<th>NUMBER of PSUs</th>
<th>CORN PLANTING FREQUENCY RANGE (Number of years out of 6)</th>
<th>MEAN CORN ACRES</th>
<th>MEAN CORN PERCENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4634</td>
<td>0.0000 - 0.2558</td>
<td><strong>58.402</strong></td>
<td>2.575</td>
</tr>
<tr>
<td>2</td>
<td>2443</td>
<td>0.2560 - 0.6508</td>
<td>203.782</td>
<td>12.041</td>
</tr>
<tr>
<td>3</td>
<td>2497</td>
<td>0.6510 - 1.0550</td>
<td>288.300</td>
<td>19.077</td>
</tr>
<tr>
<td>4</td>
<td>2500</td>
<td>1.0551 - 1.4593</td>
<td>400.515</td>
<td>26.975</td>
</tr>
<tr>
<td>5</td>
<td>2268</td>
<td>1.4596 - 1.8645</td>
<td>490.888</td>
<td>32.978</td>
</tr>
<tr>
<td>6</td>
<td>1909</td>
<td>1.8651 - 2.2901</td>
<td>638.074</td>
<td>39.241</td>
</tr>
<tr>
<td>7</td>
<td>1544</td>
<td>2.2902 - 2.8491</td>
<td>812.784</td>
<td>44.667</td>
</tr>
<tr>
<td>8</td>
<td>356</td>
<td>2.8506 - 4.8648</td>
<td><strong>922.534</strong></td>
<td>51.578</td>
</tr>
<tr>
<td>1-8</td>
<td>18151</td>
<td>0.0000 - 4.8648</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Federal Committee on Statistical Methodology Research Conference  
December 1 - 3, 2015 Washington, D.C.
RESULTS

PERCENT CULTIVATION & PSU DISTRIBUTION

• Percent cultivation substrata of the new SD ASF design has an overall accuracy of 72.365%
  – validated using the 2014 FSA CLU & 578 data

• Accuracy is consistent with the accuracy of the current SD ASF of 74% based on percent cultivation
  – validated using 2014 June Agricultural Survey data

• The new ASF’s PSU distribution is consistent with the current NASS SD ASF
RESULTS

PREDICTION OF CROP SPECIFIC PLANTING PATTERNS IN 2014

• The sub-stratification of the eight primary corn, soybeans and wheat frequency strata provides additional crop specific information at the PSU level than the current NASS ASF based on percent cultivation

• The novel ASF method creates strata where high concentrations of corn, soybean and wheat are likely planted which was confirmed using the 2014 FSA CLU data as *in situ* validation data
CONCLUSIONS

• It was observed that all three ASF designs consistently predicted corn, soybean and wheat planting patterns well as verified by the 2014 FSA CLU and 578 administrative data.

• The new stratification based on crop frequency and cultivation is crop type independent and applicable to all major crops.

• These results indicate that the new stratification method and the three novel SD crop specific ASF designs have great potential to improve ASF accuracy, efficiency and crop estimates.
FUTURE RESEARCH

• Future research will include 1) conducting estimation based on the new ASF stratifications using the 2014 FSA CLU & 578 data as proxy survey data which can be compared with the 2014 JAS estimates and 2) creating and assessing a multivariate stratification based on the NASS Crop Frequency Data Layers
REFERENCES


REFERENCES


Thank you

Questions?

Claire G. Boryan, Zhengwei Yang and Patrick Willis
USDA/NASS/RDD
claire.boryan@nass.usda.gov