Review of best practices for crop area estimation with remote sensing

Javier.gallego@jrc.ec.europa.eu
Remote sensing and Crop area estimation:

• An old love story
  (1972- ?????)

• Or better several possible love stories

• Sometimes a love-hate story
Remote sensing and crop area estimation:

- One possible story:
  - I will stand at your side every day of my life and will provide everything you need. Do not worry. I am here.
  - I will provide accurate estimates of crop area and yield and you will not need to go to the field to collect data (or very little).

- But such intense love often finishes in a violent divorce.
- Example in the European Commission (MARS ActivityB: rapid crop area change estimations):
  - Good estimates in 1990-1997
  - Later analysis reveals that there is a ± 20% margin of subjectivity for major crops
Remote sensing and crop area estimation

• Another possible story:
  • Let’s be friends. Bring your know-how, I will bring mine.

• = *Ground observations give more reliable data on a sample; remote sensing give a general view on a larger area.*

• Less romantic, but more practical
  – Example: USDA
    • Segment survey + classified images
  – Long-lasting, happy relationship
Most important research topics to improve agricultural statistics

Ranking from a survey to statistical institutes (FAO, 2010)

1. Improvement of estimation of crop area, yield and production (list or area frame surveys)
2. Use of GPS in the production of agricultural statistics
3. Methodology for the compilation of food security statistics
4. Linking area frames with list frames
5. Estimation of food stocks
6. Development of master sampling frames
7. Estimation of farm gate prices
8. Reconciliation of census data with survey data
9. Development of an integrated survey programme
10. Determination of user’s information needs for decision making
11. Methods for estimating yield of root crops, edible forest products, etc.
12. Methods for estimating crop area, yield and production of mixed and/or repeated cropping
13. Estimation of informal cross border trade data
14. Use of remote sensing
15. Use of small area estimation methods for improving agricultural statistics
16. Use of administrative data for improving agricultural statistics
Perception of African countries on research topics

• Is this low priority given to remote sensing due to an insufficient promotion?
  • A few years ago remote sensing would have been among the first priorities
  • Substantial investments in the last years to promote remote sensing for agricultural statistics in Africa

• What is happening?
  • Optimistic feasibility assessments that have not been confirmed in practice?
Possible applications of EO

Main focus of this presentation

Official statistics
- Crop Area
  - Methods are solid
  - Cost-efficiency problems
  - Few attempts

Forecasting
- A lot of research
- Yield forecasting
  - Many operational activities

Mapping
- Many
- Early warning
  - Yield indicators
Which data?

- Ground data?
- Only images?
- Ground data + images?

- It depends on the circumstances
How should we combine ground observations and remote sensing for agricultural estimates?
Target: Drafting an easy-to-read recommendations document for users.

Workshop held in Ispra June 2008.

• How often does it need to be updated?
  – When the typical classification accuracy has strong changes.
  – Example: in the EU: accuracy ~ 70-80% for main crops with medium-high resolution images.
  – When it changes to 90-95 %, the recommendations will need to be updated.
Some approaches are labeled as “Research status”
no operational applications at short term

- Crop area forecasting (estimation 3-5 months before harvest)
- Applications of SAR (radar)
- Sub-pixel analysis: the size of the pixel is of the same order or larger than the dominant field size.
  - Exception: 2-3 land cover types with strong radiometric contrast (eg: vegetation – non vegetation)
Situation 1: No or few ground data

Example: North Korea
Only the pure remote sensing approach is possible
• Margin for subjectivity: order of magnitude of the commission-omission errors.

<table>
<thead>
<tr>
<th>Landscape</th>
<th>Easy</th>
<th>Complex</th>
<th>Required Accuracy</th>
<th>Timeliness</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td></td>
<td></td>
<td>Research</td>
<td>Early</td>
</tr>
<tr>
<td>Moderate</td>
<td>(1)</td>
<td></td>
<td>(1)</td>
<td>After harvest</td>
</tr>
<tr>
<td>High</td>
<td></td>
<td></td>
<td>Research</td>
<td>Early</td>
</tr>
<tr>
<td>Moderate</td>
<td></td>
<td></td>
<td>(2)</td>
<td>After harvest</td>
</tr>
<tr>
<td>Moderate</td>
<td></td>
<td></td>
<td>(2)</td>
<td></td>
</tr>
</tbody>
</table>

(1): feasible when the priority is given to a dominant crop that has little confusion with other types of vegetation
(2): same limitation applies for the targeted groups of crops
Situation 2: A ground survey is possible

The accuracy level depends on
- Size of ground survey
- Relative efficiency of remote sensing
  - The value added by remote sensing is proportional to the size of the ground survey.

<table>
<thead>
<tr>
<th>Landscape</th>
<th>Required Accuracy</th>
<th>Single crops</th>
<th>Groups of crops</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complex</td>
<td>High</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Easy</td>
<td>High</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td>(3)</td>
<td>(4)</td>
</tr>
</tbody>
</table>

(3): Ground survey has to be carried out quickly and early and there is a short time for data cleaning.

(4): Standard situation: Regression, calibration or similar procedures recommended.
Combining ground survey and satellite images

- Main approaches: calibration and regression estimators.
- Common features:
  - combine accurate information on a sample (ground survey) with a general view from images
  - Unbiasedness is provided by the ground survey.
  - The bigger the effort for the ground survey, the higher the added value of RS.
- Difference:
  - Calibration is better adapted to ground information per point
  - Regression is better adapted to ground information per area segment
- Other approaches available: Small area estimators, ….
The USDA-NASS approach

Main data: ground observations on a sample of segments

Co-variable: classified satellite images:

- Cropland layer (Intermediate product)
- Mainly AWiIFS (56 m resolution)
- MODIS (time series) give a small contribution
- Administrative declarations of farmers: training data for classification.
- Usually 90-95% classification accuracy
  - Insufficient for a “pure remote sensing approach”
Satellite images are used for **auditing** agricultural statistics

Identifying strongly manipulated figures

- “Agricultural Attachés” of the embassies send figures and make field trips.
- Image analysts decide if the figures given by the country seem acceptable.
- Each analyst is quite free to use his personal approach.
Adapting to the EU the method used by USDA-NASS.

Images were used for
- Stratification
- Regression estimator with classified images as ancillary variable

Conclusions:
- Relative efficiency was lower than in the US, due to more complex landscape.
- Cost-efficiency with Landsat TM slightly below threshold in the 90’s

Ground data + images

Estimates
MARS Project: Rapid estimates of crop area change
(Action 4 – Activity B)

Pure remote sensing approach:
- Sample of 60 sites of 40x40 km
- 3-4 images per site every year (mainly SPOT)
- Some ground data of the previous years (for training image classification)

Good results for dominant crops:
- Example: 1-1.5 % error for the total area of cereals.
- But the margin for subjectivity was around ± 20%

Much weaker results when the changes were difficult to forecast.
LUCAS (Land Use/Cover Area frame Statistical Survey)

Main tool for land cover area estimation in the EU. (Eurostat)
Ground survey of a sample of points

Role of Remote sensing.
- Stratification
- Graphics for ground survey
- Points that cannot be reached
In most cases fully covering a country is not feasible

- Sampling necessary.
- Which is the value $K$ of a VHR image in terms of sampling error?
  - A sample of $n$ images is equivalent to a sample of $Kn$ unclustered points
- $K \approx 1/$intracluster correlation.
- The intracluster correlation is a weighted average of the correlogram.
Correlograms of several land cover types in the EU
Weights to go from Correlograms to intracluster correlation
Equivalent number of points of a 10 x 10 km site in the EU (simple random sampling)

- Equivalent number of points of a 10 x 10 km site in the EU (simple random sampling)
  - For arable land $K \approx 3$
  - For wheat $K \approx 6$
  - For sunflower $K \approx 15$
  - For land cover change (to artificial) $K \approx 30$
- The cost per point in the EU survey LUCAS is $\approx 14$ Euro
- The “sampling error” value of a VHR image 10 x 10 km roughly ranges between 40 and 500 Euro (lower than the price)
- Full coverage of satellite images is more likely to be cost-efficient