Crop Area Estimation with Remote Sensing

Some considerations and experiences for the application to **general** agricultural statistics

Javier.gallego@jrc.it

Some history: MARS Regional crop inventories

Purpose: Adapting to the EU the method used by USDA-NASS.

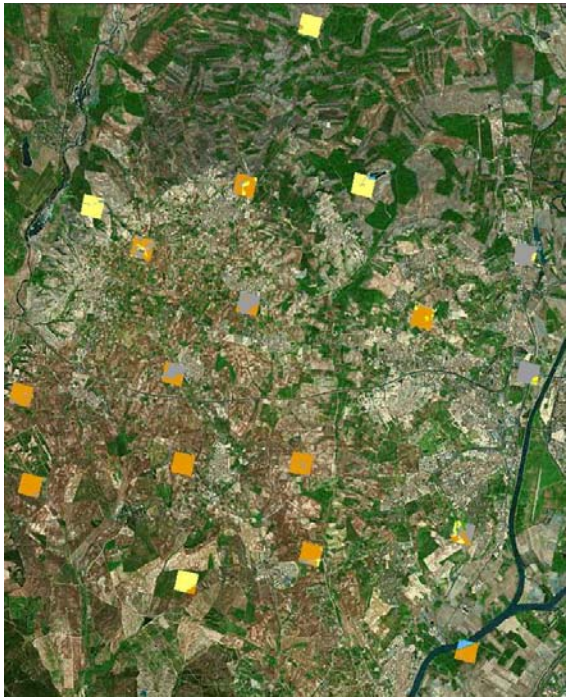
- Not running it operationally
  - This corresponds to Member States

Images were used for

- Stratification
- Supporting the ground survey
- Improving ground survey results with remote sensing
  - Regression estimator

Conclusions:

- The method could be used in the EU, but the relative efficiency was lower, due to more complex landscape.
- In 1993, the regression estimator was close to the cost-efficiency threshold with Landsat TM



Ground data

+

images

Estimates

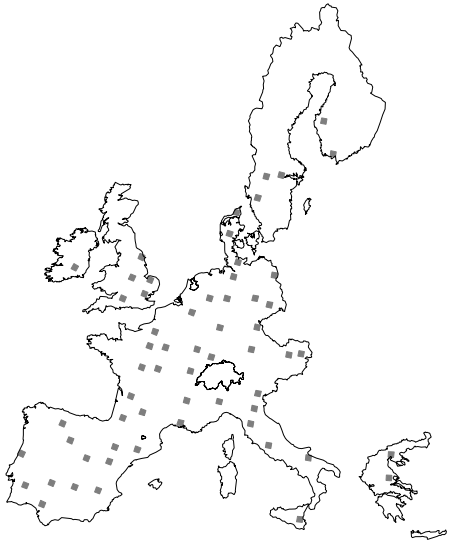
The rapid estimates of crop area change in the MARS Project (Action 4 – Activity B)

Pure remote sensing approach

Sample of 60 sites

3-4 images per site every year (mainly SPOT)

Some ground data of the previous years (for training image classification)



For 8 years it was our "star" activity

We were confirming on an objective basis to DG AGRI the figures they were expecting.

1997: Changes in agricultural policy for oil seeds

Difficult to predict area changes for rapeseed

For the first time a real challenge!!!!!!

MARS Activity B gave completely wrong figures

An expert is somebody who has made all the possible mistakes in a specific field

Niels Bohr

The MARS team became much more expert with the Action 4 / Activity B "Rapid Crop area change estimates with remote sensing"

The big mistake: believing that objective and accurate crop area (change) estimates could be obtained from satellite images without an intensive ground survey.

- It took some time to realise that the "objective" estimates were essentially subjective
  - The remote sensing team was giving the figures that the customer (DG AGRI) wanted to hear
  - Subjectivity margin: 10-20% for major crops
- Second mistake: believing that the agreement of area (change) estimates in the region could be considered as a validation of the method

A major consequence: loss of credibility of Remote Sensing

Pixel counting and similar approaches (photo-interpretation, pixel unmixing models, etc.) for area estimation:

- The margin for subjectivity is of the order of magnitude of the commission/omission errors in the classification.
- In general this is acceptable only if ground surveys are not possible or the classification accuracy is extremely good.
- The number of pixels classified in each category can be tuned by the operator $\Rightarrow$
  - area estimates by pixel counting are strongly subjective
  - You can give a good estimate if you know a priori the figure you are looking for.

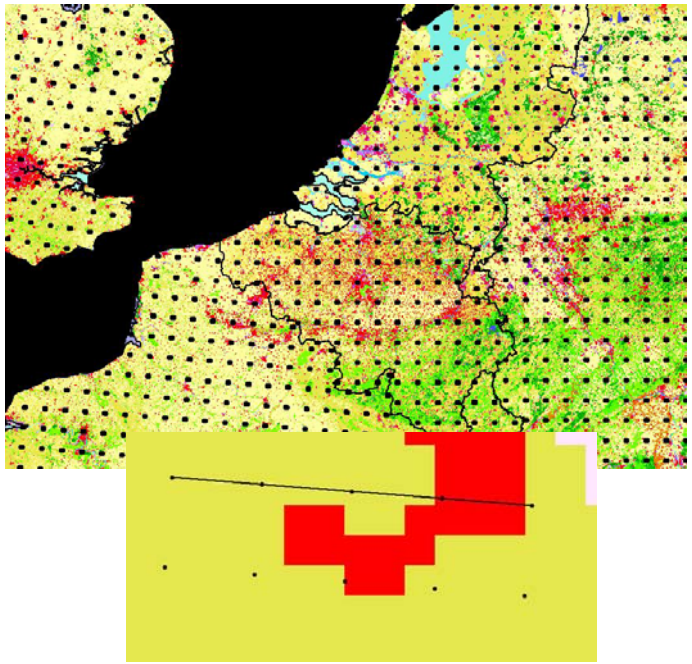
MARS "Rapid Estimates" (Action 4/Activity B): Average RMS errors of the area changes

	April	May	June	July	Aug.	Sept.	Oct.
Common wheat	1	1.2	1.9	1.6	1.5	1.4	1.4
Durum wheat	2.1	3	2.8	2.6	2.7	2.7	2.6
Barley	4	4	3.2	2.5	2.7	2.4	2.4
Rice	7.7	9.9	9.6	6.1	5.7	5	5
Maize	4	2.5	2.4	2.8	4	4.3	4
Total cereals	1.4	1.3	1	0.9	0.8	0.7	0.7
Sugar beet	6.7	4.6	4.4	2.8	4.3	2.9	3
Sunflower	16.6	12	6.5	7.4	6.3	6.7	7.3
Rapeseed	6.3	9.6	9.8	11.5	11	10.4	10.3

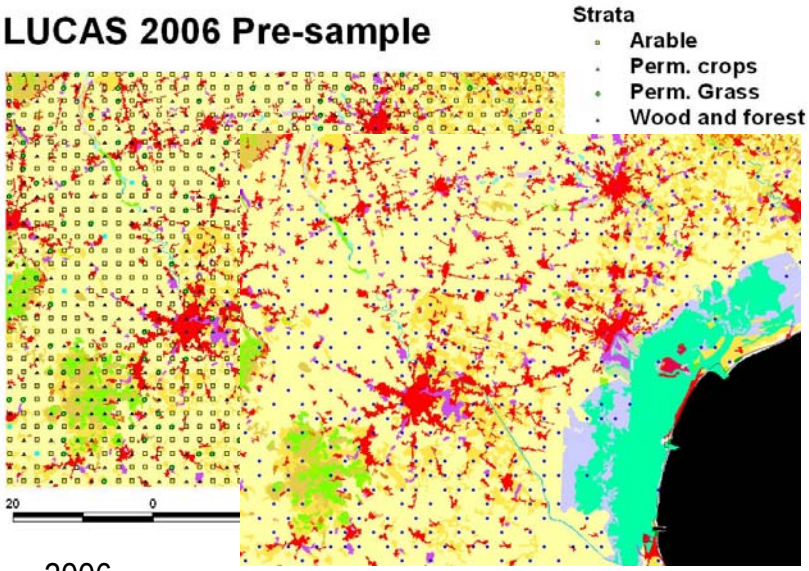
No evidence of improvement when information from satellite images were added along the year

MARS Monitoring Agriculture with Remote Sensing

LUCAS (Land Use/Cover Area frame Statistical Survey)



- 2001-2003



- 2006

Relative efficiency

	Systematic -random	Postratif	Unequal prob	Total efficiency
CEREALS	1.11	1.40	1.26	1.95
Common wheat	1.11	1.16	1.42	1.83
Durum wheat	1.43	1.29	1.41	2.60
Barley	1.15	1.17	1.40	1.88
Maize	1.21	1.19	1.43	2.06
Potatoes	1.09	1.06	1.36	1.57
Sugar beet	1.05	1.01	1.59	1.69
Sunflower	1.09	1.07	1.88	2.19

Role of Remote sensing.

- Stratification
- Graphics for ground survey

Remote Sensing and Area Frame Sampling for Agricultural Statistics:

Using classified images as co-variable is statistically sound

- Combining images with a ground survey
  - Regression estimator
  - Calibration estimator
  - Small area estimators
- Cost-efficiency depends on the landscape and type of images
  - Landsat TM had the best chance to be cost-efficient (but now it is hardly operational)
  - The more intense the ground survey, the higher the value added by satellite images
    - This does not apply when we cannot think of intense ground surveys.

The USDA-FAS approach

Satellite images are used for auditing agricultural statistics

- Agricultural Attachés of the embassies send figures and make field trips.
- Region analysts look at images and decide if the figures given by the country seem acceptable.
  - They are considering stopping looking at western/central Europe
- No specific methodology. Each analyst is quite free to use his personal approach.
- Main type of images: AWiFS (56 m resolution)
  - USDA has a framework contract for AWiFS images. Around 10 agencies in USDA use them
- Also MODIS and samples of high - very high resolution images

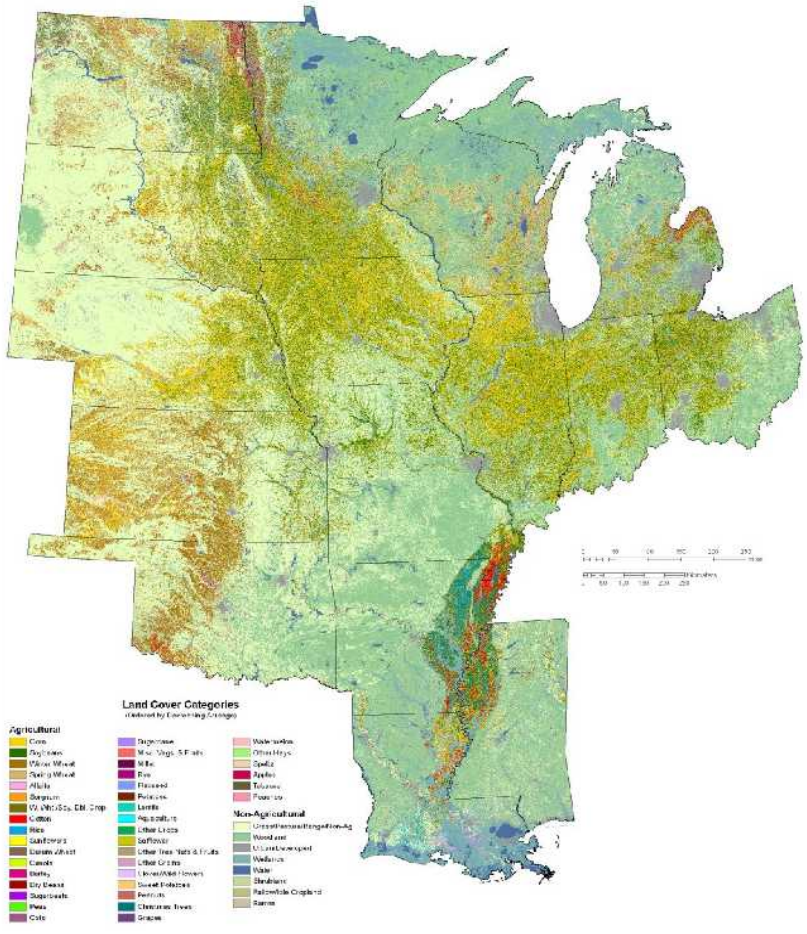
The USDA-NASS approach

Main data: ground observations on a sample of segments (Area Frame Sampling)

Co-variable: classified satellite images:

- Mainly AWiIFS (56 m resolution)
- MODIS (time series) give a small contribution
- Administrative declarations of farmers: training data for classificatic

Additional product: cropland layer (mapping, not statistics)



Subjectivity

Most statistical systems have some degree of subjectivity

- often disregarded

- Subjectivity can be approximately independent in each sampling unit
  - Decreases when the sample size grows
  - But some systematic component (bias) may remain, e.g.: unusual crops wrongly attributed to a more usual crop.
- Subjectivity in the analysis stage
  - May (partially) reduce the bias
    - Example: when observations are reviewed if they are too far from the expected value.
  - But puts a question mark on the interest of the results.

Subjectivity in pixel counting

- Intervention of the analyst
  - Mainly in tuning classification parameters
- Impact on the estimates depends on
  - Complexity of the agricultural landscape
  - Complexity of the nomenclature
  - Type of classification algorithm
    - "black box": no way of tuning (non sampling error ⇒ bias)
    - Flexible algorithm (bias becomes margin for subjectivity)
- Potential bias / margin for subjectivity
  - ~ of the order of magnitude of commission/omission errors
- Benchmark: which is the uncertainty of area by crop?
  - change of crop area from year to year if reliable statistics are available for previous years
  - If the subjectivity is smaller than the uncertainty, pixel counting can be acceptable
- ⇒ Getting estimates close to expected results (official statistics…) is not an acceptable validation if the margin for subjectivity is large

Using coarse resolution

Many current attempts to use MODIS/MERIS for area estimation (250-300 m).

- If fields are very large (most pixels are pure), previous considerations remain valid
- If most pixels are mixed, the concept of confusion matrix is easy to adapt, but
  - The contamination by co-location inaccuracy is higher
  - The error of the calibration estimator is more difficult to compute
    - $\Rightarrow$ I do not know how to do it.
  - Better using regression estimator in this case?

Accuracy in an easy landscape

	Ground				
MODIS classif		cereals + fallow	grass+ abandon	total	Producer accuracy
	crop	1470+152	57	1679	96.6%
	grass+ abandon	39+68	353	460	76.7%
	total	1729	410	2139	
	User accuracy	93.8%	86.1%		

- Pilot study Kazakhstan
- The total area of crops (Cereals+fallow) can be estimated by pixel counting with a subjectivity margin ~ ± 5%

Classification accuracy in difficult landscape

Tab. 10: Confusion Matrix – EoC ETH 2007 for all validation sites

EoC ETH	Classification					
Validation	0 - unclassified	1 – cultivated land	2 – non-cultivated land	Reference Total	Producers Accuracy	Omission Error
1 - cultivated land	3	108	46	154 (157)	70	30
2 - non cultivated land	0	40	34	74	46	54
Classification Total	**3**	**148**	**80**	228 (231)	K = 0.16	
Users Accuracy	-	73	43	OA = 62 %		
Commission Error	-	27	57			

- MODIS

ETH 10m					
Validation	1 – cultivated Land	4 – non-cultivated Land	Reference Total	Producers Accuracy	Omission Error
1 - cultivated land	116	23	139	83	17
2 - non cultivated land	71	39	110	35	65
Classification Total	**187**	**62**	249	K = 0.20	
Users Accuracy	62	63	OA = 62 %		
Commission Error	38	37			

- HR SAR

- Source: GMFS validation report

The GEOSS Best practices report

GEOSS: Global Earth Observation System of Systems

- Workshop held in Ispra June 2008.
- Recommendations document approved unanimously by the ad-hoc breakout group.
- Currently in circulation for comments (sent in August. No comments received yet).

GEOSS Best practices report

Research status (no operational applications can be foreseen 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)

GEOSS Best practices report: situations

Few ground data can be acquired

- Limitation of the accuracy (margin for subjectivity): order of magnitude of the commission-omission errors on the finest resolution.
  - Estimation possible (only indicative if ground data are not coming from a proper sampling scheme)

(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

				Nomenclature			
				Single crops		**Groups of crops**	
				Timeliness			
				Early	After harvest	Early	After harvest
Landscape	Easy	Required Accuracy	High	*Research*	*Research*	*Research*	*Research*
			Moderate	*(1)*	*(1)*	*(2)*	*(2)*
	Complex		High	*Research*	*Research*	*Research*	*Research*
			Moderate	*Research*	*Research*	*(2)*	*(2)*

GEOSS Best practices report: situations

A proper ground survey is possible.

- 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.

(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.

				Nomenclature			
				Single crops		**Groups of crops**	
				Timeliness			
				Early	After harvest	Early	After harvest
Landscape	Easy	Required Accuracy	High	(3)	(4)	(3)	(4)
			Moderate	(3)	(4)	(3)	(4)
	Complex		High	(3)	(4)	(3)	(4)
			Moderate	(3)	(4)	(3)	(4)

Stratification aspects

- Minimising variance (classical target)
  - Sample allocation
- Taking into account the identification approach/accuracy
- Yield
- Easy-difficult access (cost function)
  - Calibrate image analysis in non-accessible areas with confusion matrices in similar areas

An option to evaluate: stripe sampling for aerial photographs

Sample versus complete cover

- Total error^2 ≈ sampling error^2 + non-sampling error^2
- Classical statitiscs provide tools to compute sampling errors
  - Not always easy
  - Not always possible to get unbiased estimators
- Computing/estimating non-sampling errors (bias) is often impossible.
  - Getting an order of magnitude may already a good result
  - In remote sensing, indications come from confusion matrices (commission/omission errors)
- Wall-to-wall cover + sample of higher resolution images + sample of ground visits where possible?

Sample versus complete cover

- Sample allocation
- Taking into account the identification approach/accuracy
- Yield
- Easy-difficult access (cost function)
  - Calibrate image analysis in non-accessible areas with confusion matrices in similar areas

An option to evaluate: stripe sampling for aerial photographs