

## LPIS QA changes to the ETS guidance

Baveno, October 15, 2013

[www.jrc.ec.europa.eu](http://www.jrc.ec.europa.eu)



Serving society  
Stimulating innovation  
Supporting legislation

Joint  
Research  
Centre

## DISCLAIMER

These slides contain the proposals presented by the JRC during the workshop on 13 and 14 October 2013.

Several comments and suggestions have been and can still be made on these proposals.

These published slides do not represent any final and formal guidelines. These need to be endorsed by DGAgri and will be published in WikiCAP. Please check WikiCAP for formal documents.

The slides serve an informative role only, addressed to the participants in the scope of the LPIS workshop. They should not be used beyond the workshop scope.

15 October 2013

Joint  
Research  
Centre

2

## Outline

1. **Effective 2013**
  - QE1 reporting
  - Skipping
2. **Optional 2013**
  - Crop measurement
3. **Effective 2014**
  - Crop measurement becomes default
  - Sampling linkage
4. **Future**
  - Dedicated LPIS QA zoning
  - Peer review

## QE1: correctness of total area

2010: QE1 quantified total “error” (in absolute terms)

2011-.....: ~~absolute difference~~ quantified bias in eligibility profile

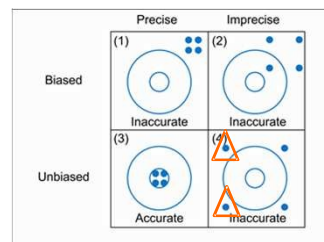
Standing problems

1.  $\text{Correct} \approx \text{Accurate} \equiv \text{unbiased} + \text{precise}$   
Without precision, compensation of absolute values occurred (e.g. included forests offset excluded grasslands)

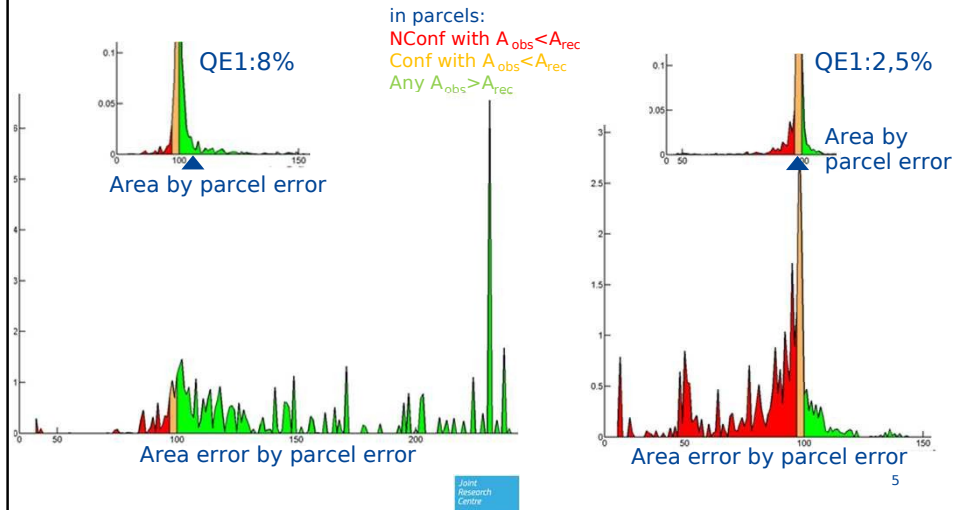
2. How to quantify the potential risk?

In their QE1 assessment, MS often conclude “unbiased” means “no financial **risk** to the fund” expressing a need for such parameter.

CONCLUSION: QE1 needs to look at both components of correctness



## Accuracy, precision and potential risk



## Quantifying the potential risk

### Assumptions:

- Not all observed area differences are errors (need to accommodate **measurement uncertainty**)
- Only **overestimate errors** represent risk for undue payment (if pre-established form and crosscheck are OK)
- Can be expressed as overestimate error rate (OER)

### Meaning of the overestimate error

The sum of overestimate errors represent what is under the left tail of the distribution, so an indicator of (lack of) precision

If LPIS QA and OTSC methodologies are both correctly applied, "areas not found" rate should match the OER

## QE1 additional indicator

Expectation: LPIS needs to be good enough so that it doesn't pose any particular risk to the fund (hence art 31a)

LPIS QA must **exclude** the possibility that 2% of the LPIS area could be "overestimate".

$\neq$  1.9% of the observed area can be "overestimate error"

= lower boundary of one-sided confidence interval of OER (LIB) > -2%

- the burden of proof is on the LPIS system
- it makes good sense to reduce the uncertainty interval of that error by measuring more reference parcels
- copy/pasting invalidates the confidence interval

Less overestimate error  $\approx$  shorter left tail of area by parcel error

- LPIS must not only be unbiased but also precise

## Formulas

For the risk assessment, we need two values:

$$1) \text{ OER} = \frac{\text{Total Overestimate Errors}}{\text{Total References}}$$

$$2) \text{ LIB} = \text{OER} - z * \sigma_e$$

where

- $\sigma_e$  is the standard deviation of the errors (see later in the slides)
- $z$  is equal to 1.6449

Expectation set at 2%:

1.  $-2\% < \text{LIB}$  : low risk
2.  $\text{LIB} < -2\%$  : 2% risk cannot be excluded

## Step-by-step procedure for QE1

1. Compute differences
2. Compute bias
3. Compare bias with  $\pm 2\%$  (QE1a)
4. Compute relative differences
5. Select non-conforming parcels
6. Compute overestimate errors
7. Compute overestimate error rate
8. Evaluate overestimate error rate's variability
9. Compute LIB
10. Compare error rate and LIB with  $-2\%$  (QE1b)

## Start: raw Measurements

Measurements [ha]	References [ha]				
3.38	3.37				
0.60	0.61				
3.01	2.89				
2.14	2.22				
4.67	4.67				
0.74	0.73				
3.24	3.48				
17.78	17.97				

Remark: We assume that the LPIS system has 1M parcels and covers 1M ha. This is important for Steps 8-10.

## Step1: Compute differences

Measurements [ha]	References [ha]	Differences [ha]			observation
3.38	3.37	0.01			
0.60	0.61	-0.01			
3.01	2.89	0.14			Grass missing
2.14	2.22	-0.08			Road inside parcel
4.67	4.67	0.00			
0.74	0.73	0.01			
3.24	3.48	-0.24			Forest inside parcel
17.78	17.97	-0.19			

## Step2: Compute bias (QE1a)

From the table, the bias is the ratio

$$\begin{aligned}
 &(\text{Total Measurements} - \text{Total References}) / \text{Total References} = -0.19 / 17.97 \\
 &= -1.06\%
 \end{aligned}$$

### Step3: Compare bias with $\pm \pm 2\%$ (QE1a)

The bias was estimated as -1.06%

As it is between -2% and 2%, this system is accurate !

### Step4: Compute relative differences

Measurements [ha]	References [ha]	Differences [ha]	Relative Differences [%]		
3.38	3.37	0.01	0.30		
0.60	0.61	-0.01	-1.64		
3.01	2.89	0.14	4.15		
2.14	2.22	-0.08	-3.60		
4.67	4.67	0.00	0		
0.74	0.73	0.01	1.37		
3.24	3.48	-0.24	-6.90		
17.78	17.97	-0.19	-1.06%		

## Step5: Select non-conforming parcels

Measurements [ha]	References [ha]	Differences [ha]	Relative Differences [%]		
3.38	3.37	0.01	0.30		
0.60	0.61	-0.01	-1.64		
3.01	2.89	0.14	4.15		
2.14	2.22	-0.08	-3.60		
4.67	4.67	0.00	0		
0.74	0.73	0.01	1.37		
3.24	3.48	-0.24	-6.90		
17.78	17.97	-0.19	-1.06%		

15 October 2013

15

## Step6: Compute overestimate errors (i.e. NC and <0)

Measurements [ha]	References [ha]	Differences [ha]	Relative Differences [%]	Errors [ha]	
3.38	3.37	0.01	0.30	0	
0.60	0.61	-0.01	-1.64	0	
3.01	2.89	0.14	4.15	0	
2.14	2.22	-0.08	-3.60	-0.08	
4.67	4.67	0.00	0	0	
0.74	0.73	0.01	1.37	0	
3.24	3.48	-0.24	-6.90	-0.24	
17.78	17.97	-0.19	-1.06%	-0.32	

16

## Step7: Compute OER

From the table, the total error was found -0.32 ha

The overestimate error rate (OER) is thus:

Total overestimate errors/Total references =  $-0.32/17.97 = -1.78\%$

## Step8: Evaluate OER's variability

(1/2)

Measurements [ha]	References [ha]	Differences [ha]	Relative Differences [%]	Errors [ha]	Errors - (Est.Errors) [ha] -0.0178 calculated from step 7
3.38	3.37	0.01	0.30	0	$0 - (-0.0178 * 3.37) = 0.060$
0.60	0.61	-0.01	-1.64	0	$0 - (-0.0178 * 0.61) = 0.011$
3.01	2.89	0.14	4.15	0	$0 - (-0.0178 * 2.89) = 0.051$
2.14	2.22	-0.08	-3.60	-0.08	$-0.08 - (-0.0178 * 2.22) = -0.040$
4.67	4.67	0.00	0	0	$0 - (-0.0178 * 4.67) = 0.083$
0.74	0.73	0.01	1.37	0	$0 - (-0.0178 * 0.73) = 0.013$
3.24	3.48	-0.24	-6.90	-0.24	$-0.24 - (-0.0178 * 3.48) = -0.178$
17.78	17.97	-0.19	-1.06%	-0.32	(chksum: must be zero !!!) 0

## Step8: Evaluate OER's variability

(2/2)

Calculate the standard deviation of the last column  $\rightarrow \sigma_e = 0.088$  ha  
(e.g. using STDEV.S in Excel  $\sqrt{\frac{\sum (x - \bar{x})^2}{(n-1)}}$  )

The standard deviation of the overestimate error rate is

$$\sigma_r = \frac{N\sigma_e}{R\sqrt{n}}$$

where

- n is the size of the LPIS QA sample (ex: 7 parcels in reality 500/ 800/1250)
- N is the total number of parcels in the LPIS system (ex: 1M parcels)
- R is the total recorded area in the LPIS system (ex: 1M ha)

(fyi:  $N/R = 1/\text{mean parcel area}$ )

Ex:  $\sigma_r = 0.088/2.65 = 3.33\%$

## Step9: Compute LIB (QE1b)

The lower interval boundary is then computed as

$$\text{LIB} = \text{OER} - z \cdot \sigma_r$$

where

- $z = 1.6449$ , i.e. the 95% quantile of the normal distribution

Ex:  $\text{LIB} = -1.78\% - 1.6449 \cdot 3.33\% = -7.26\%$

## Step10: Compare error rate and LIB with -2% (QE1b)

We found that the LIB was -7.26%.

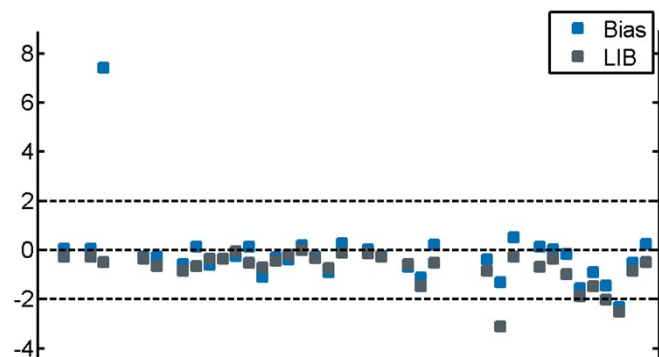
Thus, we have

$$\text{LIB} \leq -2\% \quad (\text{note: } -2\% \leq \text{OER is irrelevant})$$

This system is risky !

## Outcome

On 2012 data:



## Overestimate error and financial risk

OTSC measurements provide an independent observation of this risk

Figure: for 2012 systems

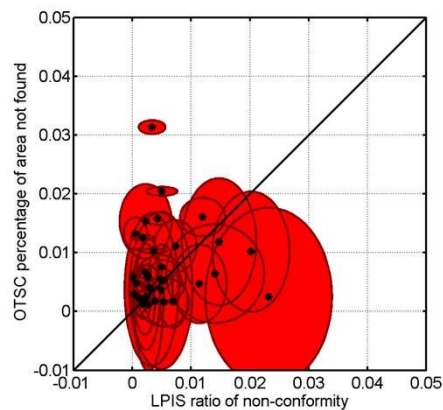
- Quid OTSC uncertainty?
- Quid LPIS QA measurements?

We cannot

- reject our assumption
- model the relation

Regarding QE7:

- $\chi^2$  will pick up any interaction
- Alternative quantitative parameter is not obvious



23

## Skipping due to erroneous scoping

Every year, there are several MS that discover population errors during the assessment processes, at a time when also the upkeep processes have been started. redo sample pre-selection is complicated

IF:

- The population submitted is complete but “contaminated”
- + that contamination is the result of an erroneous query
- + the query condition can be easily verified (screened)
- + the skipping will be applied to all contamination parcels

THEN

A dedicated skipping code can be used to enable continued inspection

24

## Consequence for 2013

### QE1:

Only processing affected: no change in observations

- Upgraded guidance on the measure (Annex I)
- Upgraded scoreboard schema
- Upgraded scoreboard word template

### Skipping

If applied: ensure the required QE4 and QE2 sample sizes

- Upgraded LPIS sample pre-selection status schema  
`<xs:element type="cap:OccurrenceType" name="parcelIsNotInScope"/>`
- New schema for describing the preparation error
- Identification of the affected parcels in the population of ETS-package

## Outline

### 1. Effective 2013

- QE1 reporting
- Skipping

### 2. Optional 2013

- Crop measurement

### 3. Effective 2014

- Crop measurement becomes default
- Sampling linkage

### 4. Future

- Dedicated LPIS QA zoning
- Peer review

## Feasibility for measurement

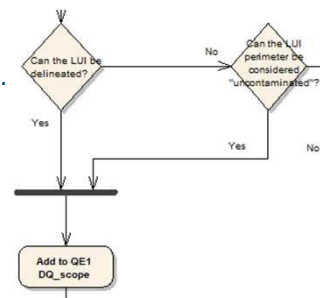
### Currently :

a No+No reference parcel is ignored but never considered non-conforming or “bad”.

### Issues :

- several MS, mostly using the cadastral reference have up to 90% No+No, not counted to QE1, perceived as bad (cfr ES intervention on DPMM 26/6/2013)
- A No+YES parcel is effectively copy/pasted and does not represent an independent physical measurement, complicating the statistical analysis.

A solution is needed where more items can be measured and none are copy-pasted

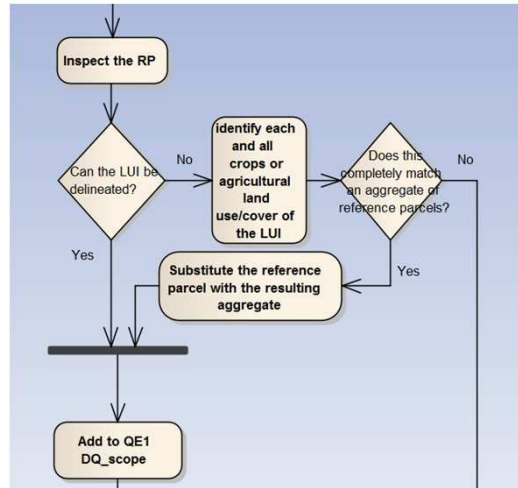


## Proposal

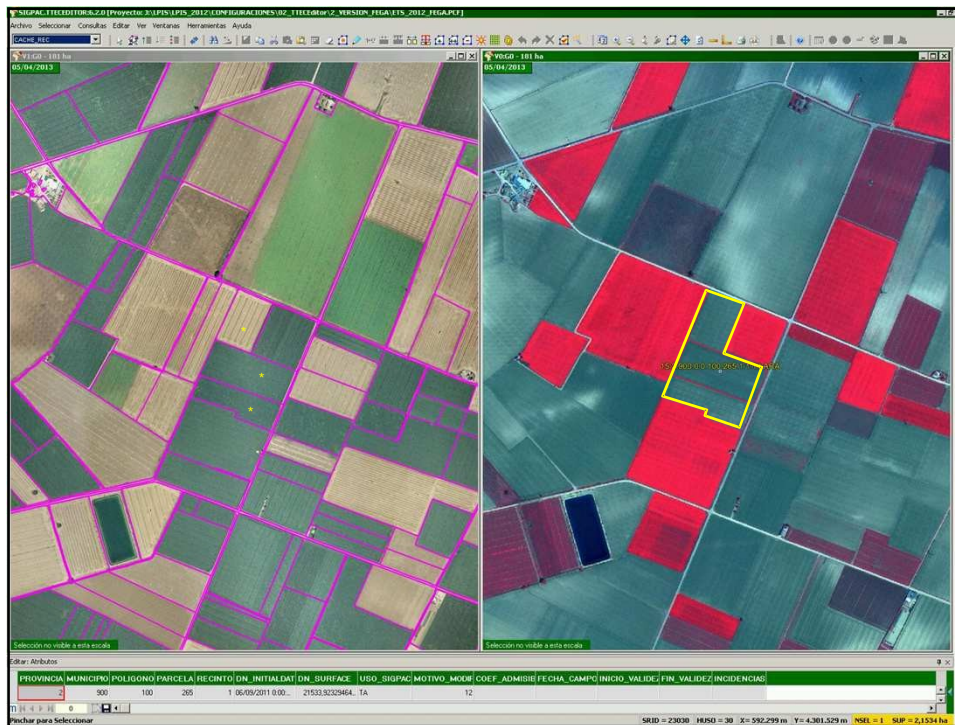
We replace the current procedure (= 5m buffer rule) and its two types of measurement by the following procedure leading to delineation

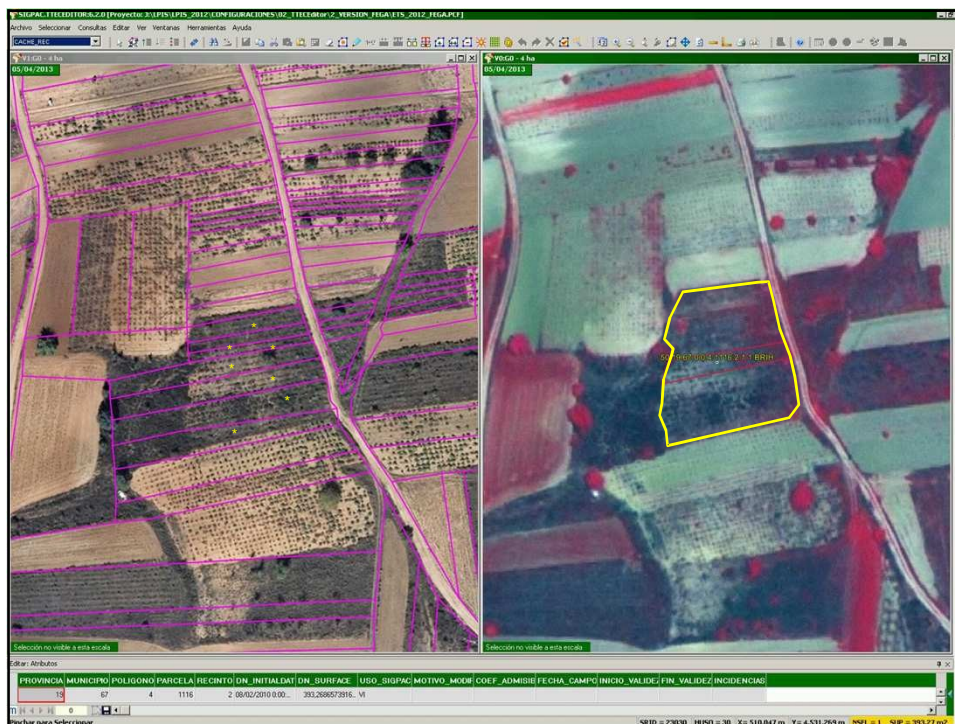
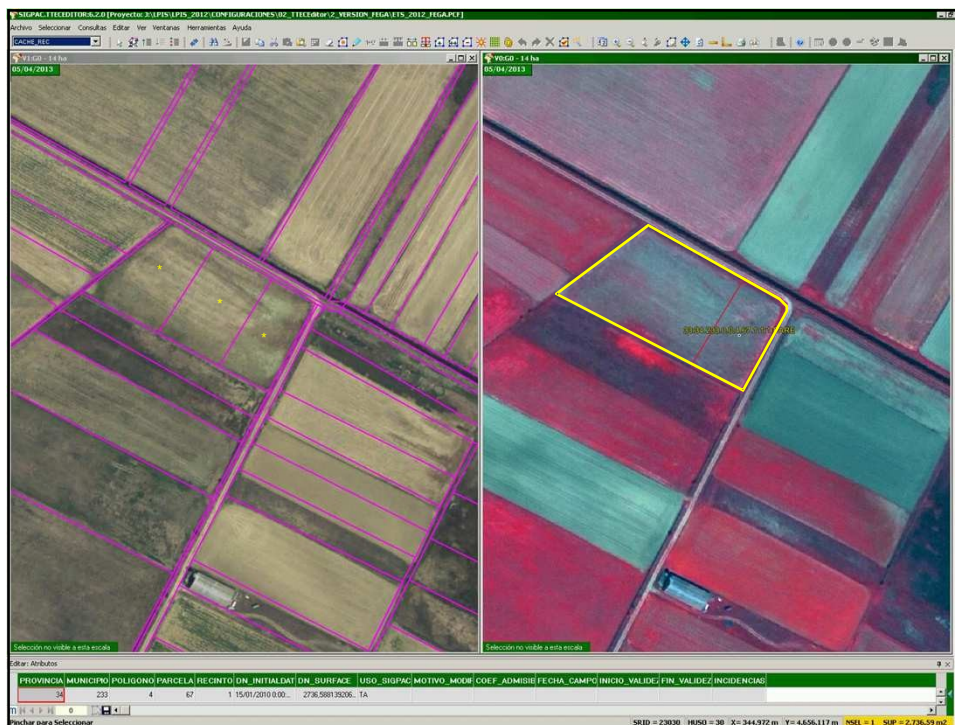
1. If the agricultural land within the LUI is not measurable, because the visible cropping pattern intersects with the LUI perimeter, Then expand the LUI to completely cover any and all visible crops, agricultural land cover type or land use units, whichever is smaller, occurring partially or completely inside the original LUI.
2. If any continuous aggregation of reference parcels matches the smallest expansion, Then, substitute the cluster for the original reference parcel in the pre-selection sample. Delineate the crop and compare  $Area_{obs}$  to the aggregate's  $\sum Area_{rec}$ .
3. Else, the reference parcel remains unmeasurable and no “area observed” will be attributed

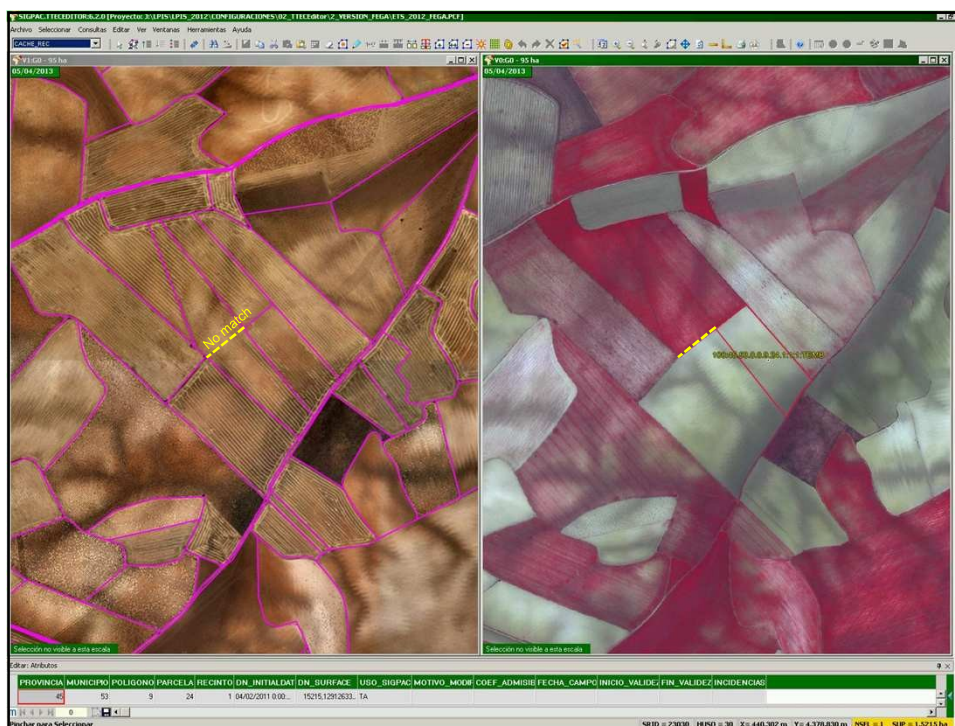
## Schema




29









  
 European Commission

## Consequences for ETS v5.4


- Aggregation/substitution action in activity diagram
  - Expanded LUI
  - Aggregation of all relevant reference parcel attributes
- New, mandatory aggregation table
 

RPID <sub>preselection</sub>   RPID(s) <sub>measured</sub>
PolygonZeroState.gml will hold both
• perimeters intersecting 100m buffer around RPID <sub>preselection</sub> ( = ETSv5.3)
AND
• perimeters of aggregate RPID(s) <sub>measured</sub> (New)

Crop/land unit delineation only

- No more 5 meter buffer
- No more copy/pasting

36

  
 Joint Research Centre

## Outline

1. Effective 2013
  - QE1 reporting
  - Skipping
2. Optional 2013
  - Crop measurement
3. **Effective 2014**
  - Crop measurement becomes default
  - Sampling linkage
4. Future – for discussion
  - Dedicated LPIS QA zoning
  - Peer review

## Sampling linkage

There need to be 127 (100-150) common parcels for  $\chi^2$  of QE7

JRC proposal (announced in Paphos) to select 127 OTSC applications, random or risk, through the random LPIS QA sample-preselection rejected by EC legal service

### Alternative proposal for 2014 :

Stratify the LPIS sample to ensure that enough reference parcels can be expected to belong to an OTSC application

Consequence:

- Flag all parcels that were declared the previous year by farmer who will be OTSC'd the current year.
- Ensure that the OTSC of the parcel likelihood is sufficient

## Considerations

Flagging applies to the previously declared parcels of the current OTSC sample

- completely anonymous
- no interference with any ongoing OTSC procedures of the MS

OTSC likelihood must be high enough

- National update imagery only usable when overlapping CwRS zones

Open issue : quid MS with partial or without CwRS?

- Use CwRS rules to “densify” OTSC into LPIS QA zones?

## Consequences for 2014 (ETS v5.4)

### Measurement guidance

- As above

### Sample linkage

- Addition of boolean flag in LPISPointZeroState.xsd
- No impact on the ETS reporting package

## Outline

1. Effective 2013
  - QE1 reporting
  - skipping
2. Optional 2013
  - Crop measurement
3. Effective 2014
  - Crop measurement becomes default
  - Sampling linkage
4. **Future**
  - Dedicated LPIS QA zoning
  - Peer review

## Dedicated LPIS QA zoning

Original discussion: use of CwRS zones causes minimum impact

- Specifications are roughly similar (area measurement)
  - No extra financial impact for existing CwRS programs
  - No extra planning provisions need for CwRS programs
- CTS guarantees sufficient OTSC density (QE7)
- Perfect integration from IACS processes possible
- OTSC risk is (mostly) unrelated to LPIS quality so can be ignored

## In practice

- CwRS program maximizes area (5% OTSC) not image quality
  - Several MS use proprietary aerial
  - Use of IKONOS (phasing out)
  - Viewing angles
- Some OTSC zoning does not respect the CTS OTSC density
  - Low QE7 sample
  - Too spread out for field observations
- Misunderstanding OTSC risk zone and LPIS QA impact
  - Unrepresentative sample
- Some MS use square scenes, other administrative boundaries
- Acquisition windows do not necessarily match
  - crop identification vs land cover delineation
- Some OTSC and LPIS administrations do not collaborate  
UNACCEPTABLE in IACS

## Proposal - for discussion

Independent and random LPIS QA zoning,

Two image options

a/ image provision through JRC-CID

b/ image provision through national aerial contract

Both options:

15kmx15km random scenes provided as part of JRC sampling process

a/ on-the-fly during the campaign

b/ at the beginning of the year

Hold 100-150 applications linked to the random OTSC sample

## Zone determination

1. Numbers controlled by representativeness of 15kmx15km scene
  - 1 zone up to 150kmx150km of territory
  - minimum distance of 250km between zone
  - Max. 3-5 zones f(parcel density, population), followed by homogeneity test
2. Scene boundary tuned by parcel distribution(ensuring efficiency)
- 3.a CID zones (tbd with image providers)
  - 100% cloud and haze free
  - Nadir view only so location on orbit
  - Weighted by weather probability
  - Decision made daily triggered by image provider supply
  - Dynamic sample pre-selection
- 3.b aerial zones
  - Scene locations delivered to MS at once
  - MS responsible for image acquisition and quality
  - Static sample pre-selection

## Q&D feasibility

Rule for zone numbers per LPIS  
 $\leq (nP * \mu) / (0.05 * A * d) \{+ e\}$

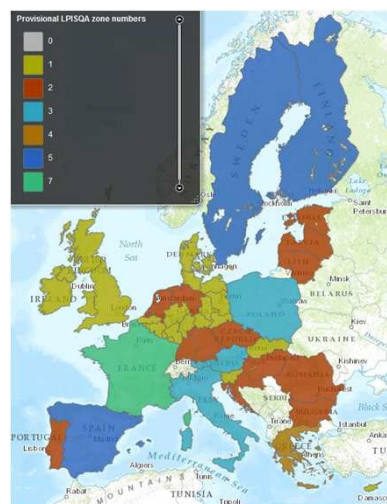
Where

- nP: LPIS QA sample size
- $\mu$ : mean parcel area (in ha)
- A: scene area (= 22500ha)
- d: density of agricultural area
- $\{+ e\}$ : optional term to ensure  
 22500km<sup>2</sup>/territory > 25%

In practice:

Total EU: ca 83 LPIS QA scenes  
 CID allocation est. 0.5 M€)

Provisional number of zones  
 See map



## Peer Review

### Status

2010+2011: individual “manual” screening by JRC

2012: advanced automatic processing

These addressed

- screening of the methodological correctness
- format checks
- internal consistency and business rules

but offered no

- detailed analysis e.g. QE3 categories
- validity of the remedial action plan

### EC challenge

- services are too limited to process 44 implementations in depth
- priority on art 31a applications

## Proposal – for discussion

Current LPIS QA set up allows for peer-reviewing: a LPIS verifies one ETS-package and report and provides feedback

### Conditions

- on voluntary basis
- random peering
- report of findings either
  - unknown to EC services (i.e. bilateral contacts)
  - anonymity of the reviewer (i.e. via LPIS QA portal)

### Practical :

JRC provides

- access to allocated ETS package (download or WFS)
- WMS access to CID imagery
- contact details



Thank you!



Wim Devos,  
statistics: Dominique Fasbender

