



# Testing the LPIS Core Conceptual Model in Practice

---



2008 MARS Annual conference "Geomatics in support of the CAP"

Ljubljana, Slovenia, 3 – 5 December 2008

- ④ Brief overview of the LPIS Core Conceptual Model
  - ⑥ Purpose And Scope
  - ⑥ The Model
  - ⑥ Status
  
- ④ LCM by Example – Slovenian LPIS
  - ⑥ Matching the model – Feature Catalogue
  - ⑥ Testing model conformance – Abstract Test Suite
  - ⑥ Testing data conformance – Executable Test Suite
  
- ④ Summary, Comments and Discussion
  - ⑥ Audit procedures
  - ⑥ LCM improvements
  - ⑥ Test suite developments and improvements

EU-wide general framework of rules and techniques for describing LPIS systems in the member states

## Purpose

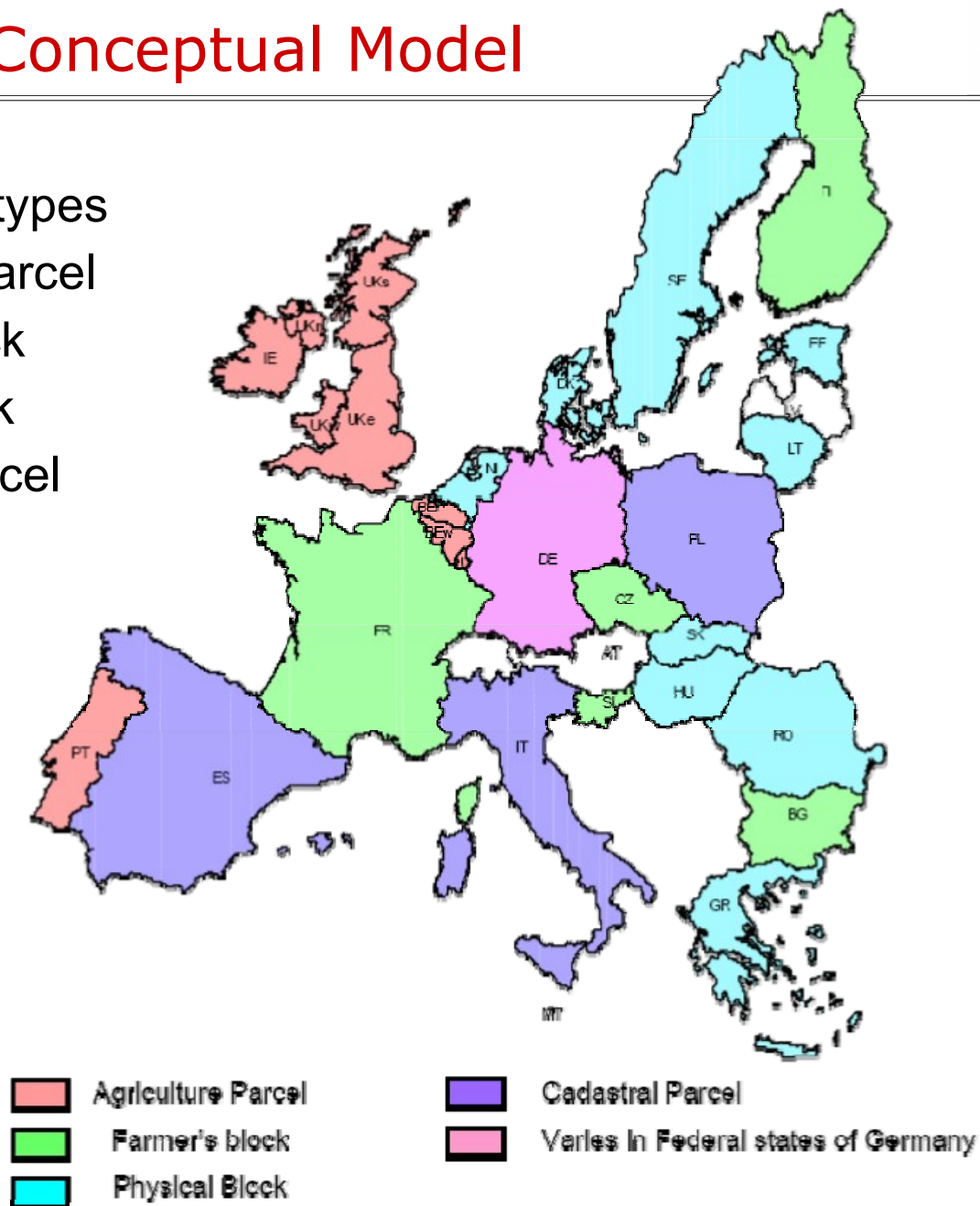
- agreement on common definitions of concept and spatial features
- proposal for standardised core data model
- open and transparent rules for quality assessment and conformity test
- promote of tools and international standards for modelling within spatial data infrastructure (SDI)

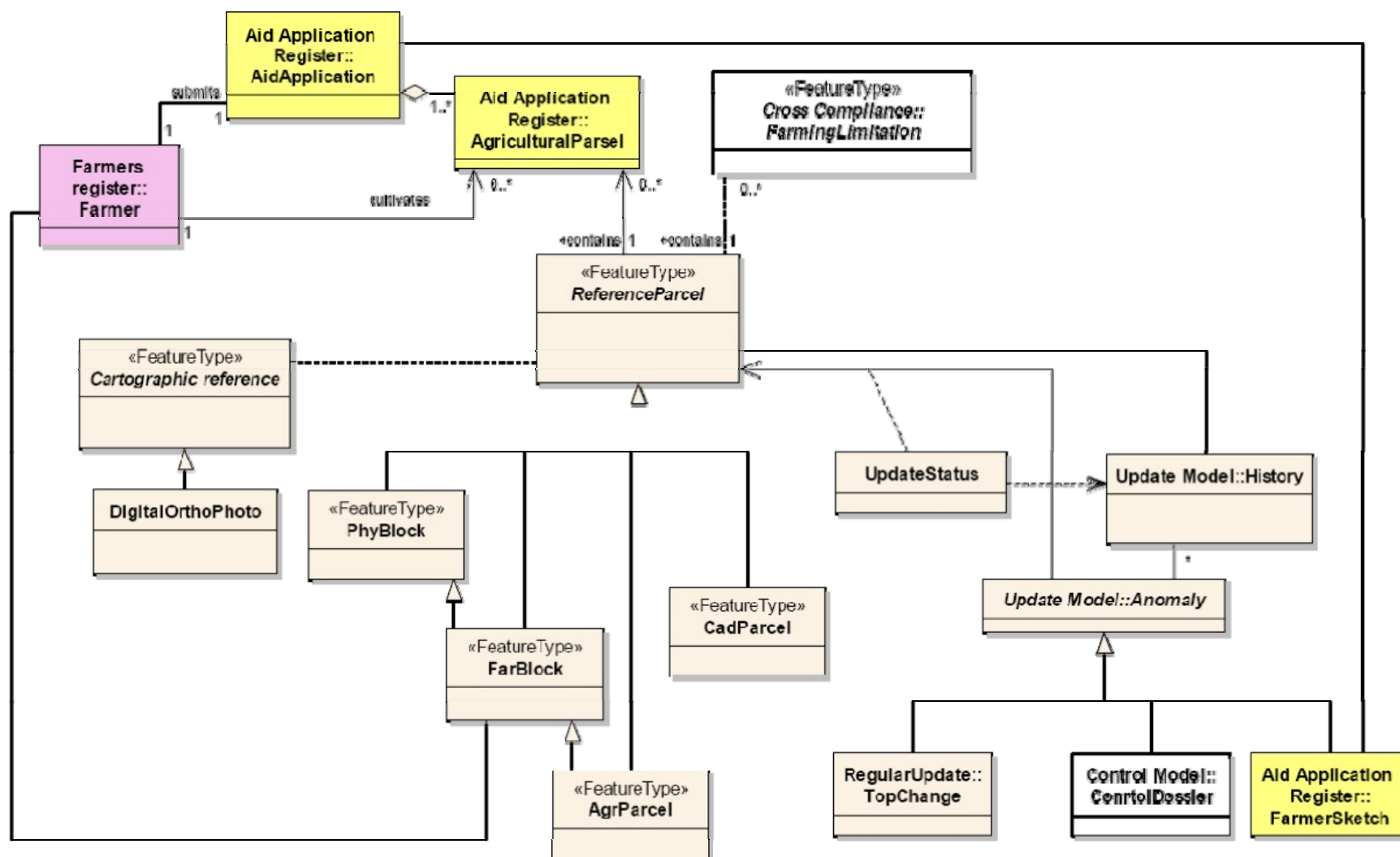
## Current scope

- analyse the European regulatory requirements for first-cut Core Conceptual Model
- analyse National LPIS implementations for 'As-is analysis
- Produce first-cut LPIS model and testing towards national implementations



- ④ four different RP types
  - ⑥ Agricultural Parcel
  - ⑥ Farmer's Block
  - ⑥ Physical Block
  - ⑥ Cadastral Parcel






- ④ announced at LPIS workshop in October 2007
- ④ draft for comments to SDIC experts available in January, 2008 via wikiCAP
- ④ first version published in discussion paper on GeoCAP web page in June, 2008
- ④ paper presented in FIG conference in Stockholm, June 2008
- ④ co-operation with Land Administration Domain Model for annex to ISO standard 19152
- ④ New version of the LCM expected January, 2009

- ④ Creation of Feature Catalogue
- ④ Abstract test suite
- ④ Executable test suite
- ④ Additional tests

FC is one of the two standard ways to describe a system's data organization (ISO 19110). The same could be achieved by creating an Application Schema. We chose FC because it is less technical

FEATURE TYPE	
Name:	<b>GERK</b>
Definition:	GERK is a graphical land use unit of farm, which is continuous piece of agricultural land, with the same land use type; in use of one single farmer. /GERK je grafična enota rabe zemljišča kmetijskega gospodarstva; je strnjena površina kmetijskega zemljišča z enako dejansko rabo, ki je v uporabi enega kmetijskega gospodarstva./
Code:	GERK_SDO
Feature Operation(s):	
Feature Attribute(s):	GERK-PID (M), KMG-MID (M), RABA_ID (M), GEOMETRY (M), ST_ENOTE_RABE (O), DOMACE_IME (M), VZROK_SPREMEMBE (M), DRZAVA (M), OPM_KMET (O), OPM_VNOS (O), OPM_MKGP (O), STANJE (M), KMG_MID_PREVZEM (O), AREA (M), PERIM (M), ST_VTX_OUTLINE (M), Y_MIN (M), X_MIN (M), Y_MAX (M), X_MAX (M), Y_C (M), X_C (M), IND_OBLIKE (M), Z_AVG (O), EXP_AVG (O), FI_AVG (O), NAGIB_AVG (O), GEOM_VALID (O), GERK_PID_H1 (O), GERK_PID_H2 (O), MEETING_ID (M), D_OD (M), STATUS (M), USER_ID (M), D_DO (O), PO_HGO (M), PO_PO (M), PO_DO (M)
Feature Association(s):	FarmedBy;
Subtype of:	ReferenceParcel; Coordinate_geometry::GM_Polygon.
LPIS_discussion	May contain one or more agricultural parcels or a part of one agricultural parcel and may be cultivated by one farmer or producers association. Does not necessarily cover a territory nationwide, but overlaps are not allowed.
LPIS_reference	2004R0796 Art 2. (26) - identification of agricultural parcels and determination of areas
LPIS_example	
LCM_Comment	RP:FarBlock

Feature attribute	
Name:	<b>AREA</b>
Definition:	calculated area based on the digitized parcel boundary.
Code:	A2100
Value Data Type:	measure
Value Measurement Unit:	m2, four decimal points
Value Domain Type:	0= 'not enumerated'
Value Domain:	
LPIS_discussion	This value is derived from geometry of GIS feature or measured in the field.
LPIS_reference	2004R0796 Art 6.(1)
LPIS_example	42.6715 m2
LCM_Comment	digitizedArea



- ④ Database contains about 40 feature tables – we selected a few:
  - ⑥ GERK (ReferenceParcel) – the most important layer
  - ⑥ RKG-KMG (FarmRegister) – directly referenced by RP
  - ⑥ RABA (LandUse) – referenced indirectly – used for checking
  
- ④ Database contains lots of attributes – it helps if we group them:
  - ⑥ core data
    - ③ identifier
    - ③ geometry
    - ③ reference area
    - ③ land cover
    - ③ farm id
    - ③ validity date
    - ③ local name
  - ⑥ versioning data
    - ③ journaling (expiry date, history status)
    - ③ transaction data (meeting id in case of Slovenia, )
  - ⑥ derived data (cached)
    - ③ digitizedArea, perimeter, average slope
  - ⑥ auxiliary data
    - ③ application-specific, not part of LCM
  
- ④ LCM is a logical model, not (necessarily) a technical one

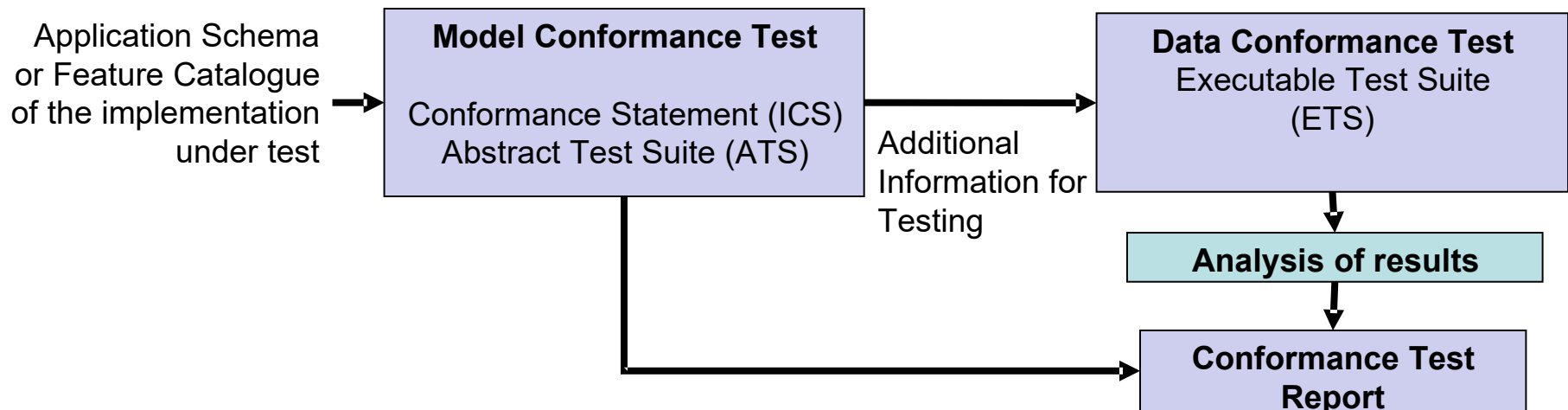
- ④ trivial task, but takes quite a lot of time
- ④ creation helps compare one's own system to the “average” implementation
  - ④ **LCM\_Comment** property provides the mapping to LCM
- ④ helps understand the relation between EC regulations and the data model (this is especially true for technical audiences)
  - ④ **LPIS\_Reference** provides a link to relevant regulations for each attribute
- ④ good opportunity for analyzing the system (and detecting its weak points)

- ④ Excel is not a very friendly way of doing this
- ④ Field names and some of the constraints could be read automatically from the database (or other data sources, e.g. via WFS describeFeatureType operation)
- ④ In the future, a more structured FC (or Application Schema) could be used to generate some executable tests



The screenshot shows a web-based schema editor interface with three tabs: Dataset, Vectors, and Attributes. The 'Attributes' tab is active. At the top, there are two buttons: 'Check Attributes' (with a green checkmark icon) and 'Perform Data Mining' (with a bar chart icon). Below these buttons, a list of attributes is displayed. The first attribute is 'REF\_CODE', which is a Character type with a length of 10. It has 'Mandatory' and 'Unique' constraints checked. There is an empty text box for 'Alias' and a larger empty text box for 'Description'. Below these, the 'Domain Codes List' is shown as 'ramсар\_REF\_CODE\_AUTO-GENERATED' with three icons: a red 'X', a pencil, and a green checkmark. Below 'REF\_CODE', there are three more attributes listed: 'NAME', 'MEASURE', and 'LABEL'. The 'LABEL' attribute is also a Character type with a length of 140, and it has 'Mandatory' and 'Unique' constraints checked. A vertical scrollbar is visible on the right side of the attribute list.

- ④ ISO 19105 Conformance and testing
- ④ **abstract test** - generalized test for a particular requirement, only system description in form of FC or application schema is examined; a formal basis for deriving executable test cases
- ④ **executable test** - specific test for a data set to meet particular requirements



Test that the implemented data model conforms to the LCM specification

④ The test suite is still under development

Example (ATS A1.1.1.1):

<b>Test purpose:</b>	verify definition of reference parcel: reference parcel demarcated by the farmer, who cultivates it (manage/execute his tenure rights: ownership, rent etc.) on multi-annual basis
<b>Test method:</b>	Inspect the documentation of the application schema or feature catalogue, verify consistency with ICS ATS NOTE: Conformant with Farmer's block definition
<b>Test reference:</b>	LCM specification
<b>Test type:</b>	Basic test

- ④ What type of reference parcel is GERK?
  - ⑥ GERK = Farmer block with more detailed land cover classification
  - ⑥ Different (overlapping) APs for different aid schemes – 1 to many cardinality?
- ④ LFA not a simple field – intersection with 3 LFA layers, calculated each year (stored in 3 fields)
- ④ Cartographic reference not directly included in the database, but can be obtained for each parcel if needed
- ④ There are three areas in the LCM: reference (M), digitised (O) and claimed (O); we have only one
  - ⑥ We included an explanation in the ATS:

A. 1.3.1.5 Area claimed inside parcel	Conformant	AREA (same as digitisedArea and eligibleArea) - For SPS, farmers can claim only whole GERKs (0% or 100%) – value can be derived from the applications database
---------------------------------------	------------	--

# Creation of FC - Land Cover

Is Slovenian reference parcel (GERK) the same as farmer's block?

- ④ The only difference is the code list of land use, which is not prescribed by the LCM

Art 44(2) Council Reg. (EC) No 1782/2003

- ④ "Eligible hectare" shall mean any agricultural area of the holding taken up by arable land and permanent pasture except areas under permanent crops, forests or used for non agricultural activities

→ The land cover concept

«FeatureType» PhyBlock
<ul style="list-style-type: none"> <li>+ landUseCover: eligibleCode</li> <li>+ isLFA: boolean</li> <li>..ReferenceParcel</li> <li>+ uniqueID: CharacterString</li> <li>+ referenceArea: float</li> <li>+ effectiveDate: Date</li> <li>+/- digitisedArea: float</li> <li>+/- claimedArea: float</li> </ul>

«CodeList» EligibleCode
<ul style="list-style-type: none"> <li>+ arableLand:</li> <li>- permanentPastures:</li> <li>+ grassland:</li> <li>+ oliveTrees:</li> <li>+ nonAgricultural:</li> <li>+ permanentCrops:</li> <li>+ forest:</li> </ul>

arable land	1100	arable land
	1160	hop fields
	1180	other permanent crops on arable land
	1190	green houses
permanent crops	1211	vineyards
	1212	nursery
	1221	intens. Orchards
	1222	ekst. Orchards
	1230	olive groves
	1240	other permanent crops
permanent pasture	1300	meadows and pastures
	1321	swampy meadows
non-utilised agriculture areas	1410	overgrown areas
	1420	forest plantations
	1500	riparian overgrowth and forest hedges
	1600	uncultivated agriculture land

Slovenian model is conformant,  
but barely...

④ ATS is a good way to decide  
which RP type is implemented

⑥ has to be improved with  
input from different MS

④ Can help find potential  
misinterpretations of certain feature types properties  
(e.g. landCover, farmerArea)

④ Are “pass” and “fail” tests OK - Is it possible to be partly conformant in a  
test?

④ Is the LCM general enough? Will it be accepted and used if most MS  
implementations fail for some reason.

Test module test	Conformity element	NOTE
A.1.1 – Reference parcel definition	Conformant	Module A.1.1 is 'Conformant' value if one of the tests A.1.1.1.1, A.1.1.1.2, A.1.1.1.3 or A.1.1.2.1
A.1.1.1	Conformant	
A.1.1.1.1	Not applicable	
A.1.1.1.2	Conformant	Farm Block with detailed land-use.
A.1.1.1.3	Not applicable	
A.1.1.2	Not applicable	
A.1.1.2.1	Not applicable	
A.1.1.3	Not applicable	
A.1.2 Eligible Land Type (land cover)	Conformant	Module A.1.2 is 'Conformant' if all tests are conformant
A.1.2.1 Eligible Land type classification	Conformant	1 to many. Arable land: 1100, 1160, 1180, 1190 Perm. crops: 1211, 1212, 1221, 1230, 1240 Perm. pasture: 1300, 1321 Non-utilized agriculture areas: 1410, 1500, 1600 (historical reasons)
A.1.2.2 Eligible Land Type (historical)	Not applicable	
A.1.2.2.1 Eligible Land type classification (historical)	Not applicable	Not evaluated, for SPS historical land use is not needed



- ④ ETS checks that all data in the system are conformant to LCM and national feature catalogue
- ④ Should be implemented for each MS separately  
– this is very easy with proper guidelines

## Example:

### Test:

farmID field references a valid farm record in the farm registry

VS

### Statement:

```
select COUNT(*) FROM GERK_SDO rp, rkg_kmg_v farm
WHERE rp.kmg_mid>0 AND rp.kmg_mid = farm.kmg_mid;
```

## Slide 17

---

**VS13**

but they cover 99% of RPs, not only 5 files of current audit procedure

Valentina Sagris, 02/12/2008

## ④ Slovenia

- ⑥ With appropriate tuning of parameters, we find some anomalies
- ⑥ Total area of anomalous data is very small (less than 0.00001%)
- ⑥ We had executable tests in place before the LCM – we knew about most of the anomalies before the tests

Test	Statement	Expected	Actual Value	Area of non-conformant RPs [m <sup>2</sup> ]	Conformance	Ratio	Area Ratio [per million]
All RP <b>uniqueIDs</b> are <b>valid</b>	select COUNT(*) FROM GERK_SDO WHERE KMG_MID>0 AND GERK_PID<0;	0	0		Conformant		
All RP <b>uniqueIDs</b> are <b>distinct</b>	select COUNT(DISTINCT GERK_PID) FROM GERK_SDO WHERE KMG_MID>0;	775679	775679		Conformant		
All RPs have <b>referenceArea</b>	select COUNT(*) FROM GERK_SDO WHERE KMG_MID>0 AND AREA IS NULL;	0	0		Conformant		
<b>referenceArea</b> of RPs should not be smaller than 25 m <sup>2</sup>	select COUNT(*) FROM GERK_SDO WHERE KMG_MID>0 AND AREA < 25;	0	71	1339	Non-Conformant	0.009%	0.27
All RPs have <b>effectiveDate</b>	select COUNT(*) FROM GERK_SDO WHERE D_OD IS NULL;	0	0		Conformant		
All <b>effectiveDates</b> of RPs are <b>valid</b>	select COUNT(*) FROM GERK_SDO WHERE KMG_MID>0 AND (D_OD < TO_DATE('2005-01-01','YYYY-MM-DD') OR D_OD > sysdate);	0	0		Conformant		
All RPs have populated land use field	select COUNT(*) FROM GERK_SDO WHERE KMG_MID>0 AND RABA_ID IS NULL;	0	0		Conformant		
RP's land use should represent agricultural land use	select COUNT(*) FROM GERK_SDO WHERE KMG_MID>0 AND (RABA_ID<=0 OR RABA_ID>=2000 AND RABA_ID<>9999)	0	9	3287	Non-Conformant	0.001%	0.67
farmID references a valid farm record	select COUNT(*) FROM GERK_SDO rp, rkg_kmg_v farm WHERE rp.kmg_mid>0 AND rp.kmg_mid = farm.kmg_mid;	775679	775679		Conformant		
RP update information references a valid user	select COUNT(*) FROM GERK_SDO g, U_USERS u WHERE KMG_MID>0 AND g.USER_ID=u.USER_ID;	775679	775679		Conformant		
RP name is not empty	select COUNT(*) FROM GERK_SDO WHERE KMG_MID>0 AND length(trim(DOMACE_IME)) <=0;	0	0		Conformant		
RP land cover is inconsistent with the type of the farm	select count(*) FROM GERK_SDO g, RKG_KMG k WHERE g.KMG_MID=k.kmg_mid AND k.VRSTA <> 'KMG' AND RABA_ID NOT IN (1130,1300,1321,1330,1410,1500,1600,1800);	0	4	38119	Non-Conformant	0.001%	7.75

VS

## Slide 18

---

**VS1**

but they cover 99% of RPs, not only 5 files of current audit procedure

Valentina Sagris, 02/12/2008

- ④ need to find a set of useful tests for all the MS
  - ⑥ experience from real-life systems are important
- ④ optimal quality acceptance levels should be established (i.e. What percentage of non-conformity is still acceptable)
  - ⑥ discussions with all MS to tune the test results
- ④ need to separate anomalies which have already been processed and are in the process of being resolved (update can take a while)

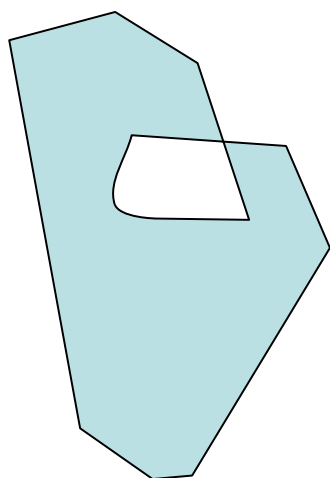
So far only the most trivial tests were identified and implemented

- ④ Testing other aspects of data quality?
- ⑥ Overlays (double declaration), topology / digitalisation errors
- ⑥ Automated cross-compliance checks – e.g. graphical intersections with spatial layers (land cover, LFA, protected areas...)
- ⑥ Testing of historic lookup functionalities
- ⑥ Testing update procedure
- ⑥ Relations to external registers
  
- ④ Find potential anomalies  
(not strictly non-conformant, but “strange” data)
  
- ④ Executable procedures to produce statistics (e.g. distribution of eligibility percentage – more useful than simple 75 / 90 rule)

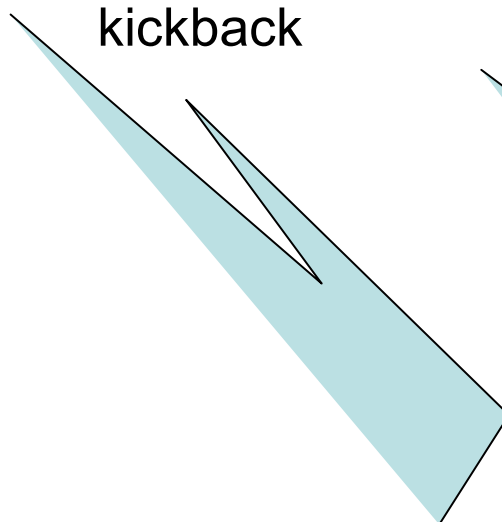
# Example - automatic topology checking

Error Code	Description	Count
101 LOOPBACK	Loop backs - self intersections (Termed 'Butterfly' polygons).	0
104 DUPLICATED_POINTS	A point that duplicates exactly the same X,Y coordinates as another point.	0
105 KICKBACKS	Digitising error leading to an inconsistency in the line.	2236
106 SPIKES	Digitising error leading to a spike inconsistency in the line. Similar to kick backs.	27232
108 SLIVER	Very small gaps between the boundaries of adjacent polygon features.	16415
109 OVERLAP	An overlap of one polygon or line feature onto another.	0
111 SHORT_SEGMENTS	A very short distance between two nodes or vertices. This distance is specified and would be expected to be the same as the cluster tolerance on the dataset.	1936
115 NEAR_POINT	A very short distance between two points.	10
119 SHORT_LINE	Length of a linear feature is smaller than a specified length.	0

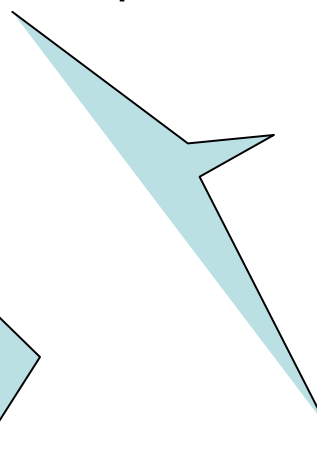
loopback



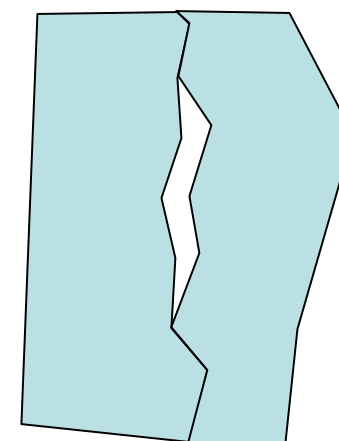
kickback

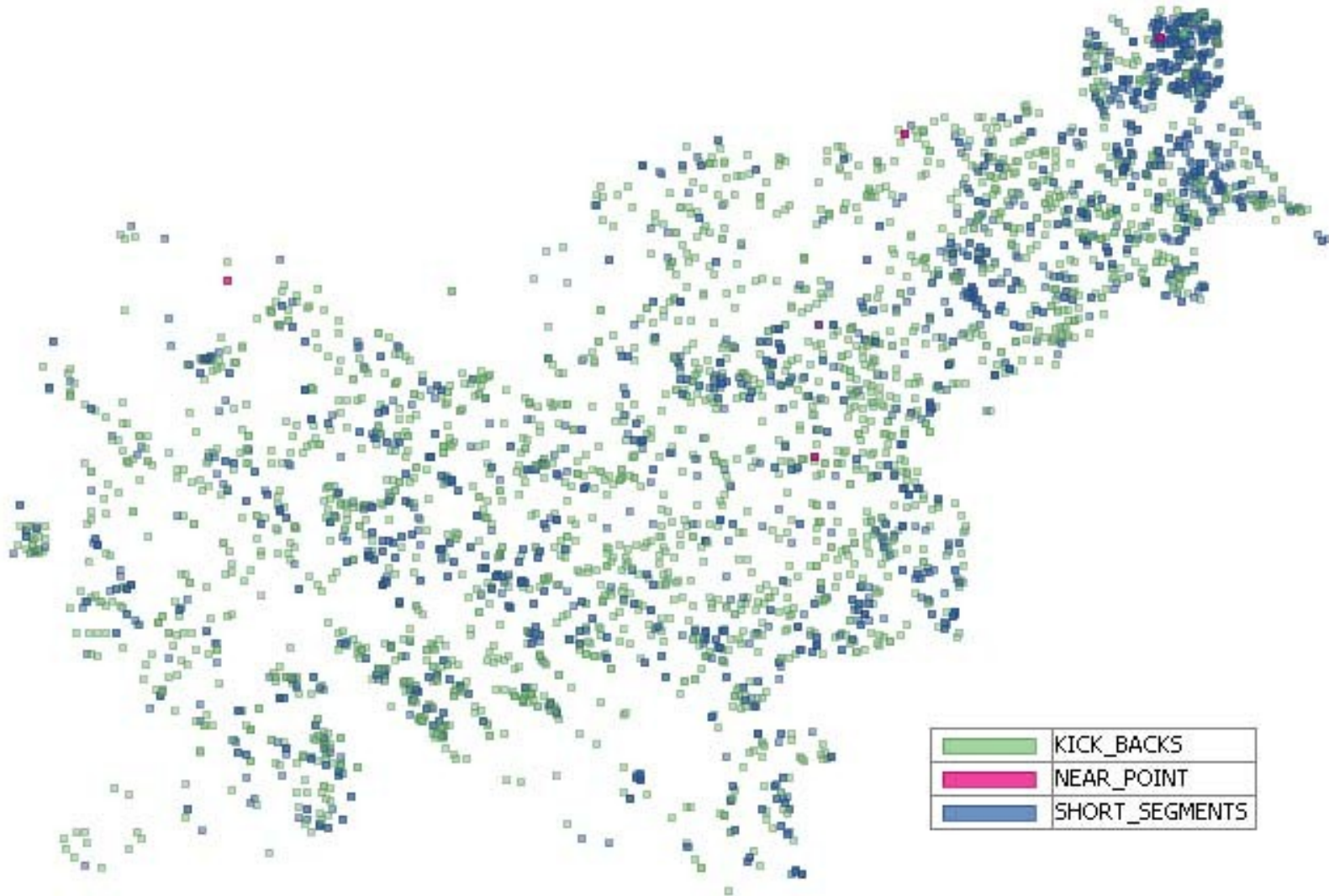


spike



sliver









- ④ Automatic or semi-automatic test scripts generated from a Feature Catalog or Application Schema
- ④ Exchange of testing experiences among member states
- ④ Analysis and examples of legislative, administrative and procedural consequences of an approach – e.g. what does it mean to use Farmer's Block:
  - ⑥ Which aspects of the updating and control should be given special attention
  - ⑥ What are the consequences of the model for the processing of Aid Applications



- ④ Information on the LPIS is currently obtained from:
  - ⑥ The annual IACS statistics/reports delivered by the MS
  - ⑥ The LPIS questionnaire sent prior to audit (influence of this step is decreasing)
  - ⑥ Presentations and clarifications requested at the start of the audit
  - ⑥ Visual inspection of a **selection of control files** to select parcels that need closer inspection
  - ⑥ The RP condition of the selection is assessed during a repetition of the OTSC or CwRS, covering a.o.
    - ③ role in the farmer declaration
    - ③ role during the administrative crosschecks
    - ③ interaction of OTSC/CwRS inspection with the update process
- ④ As a result, in the majority of cases, the control file acts as the entry point; in only a few cases a selection of all reference parcels are inspected.



## ④ Advantages

- ⑥ All RP aspects can be potentially be addressed (format, structure, functioning, performance)
- ⑥ The overall process can potentially be inspected

## ④ Drawbacks:

- ⑥ small numbers of RP are involved, representativeness is not ensured
- ⑥ there's not always enough resources to assess the complete system in full

- ④ Is current LCM general enough to cover all MS?
  - ⑥ How to treat non-conformances because of alternative solutions
- ④ How to use the results of the ATS and ETS as a “pre-audit test”
- ④ How many MS are conformant with the model
  - ⑥ now is the time to influence the LCM
  - ⑥ feedback is very welcome



- ④ Costs – (cca. 40 hours for Cosylab)
  - ⑥ FC generation (only basic layers) - 5 h
    - + overview of LCM (6 h)
    - + coordination with JRC (8 h)
    - + coordination with the MAFF Administration (6 h)
    - + problems with Excel formatting (2 h)
  - ⑥ Abstract test suite – 1 h
    - + coordination with JRC (4 h)
  - ⑥ Executable test suite – 2 h
    - + designing tests to conform with LCM – 2 h
  - ⑥ Geometry tests – 4 h
- ④ Outcome
  - ⑥ MS administration sees that our system is in principle compliant with the LCM
  - ⑥ Software developers have a better understanding of relations to other implementations and to the EC Regulations
  - ⑥ JRC is happy (collaboration helps improve LCM and tests)
  - ⑥ Some ISO and INSPIRE requirements solved as a side effect
- ④ Balance
  - ⑥ We want more feedback from tests
  - ⑥ It looks promising, but a lot still has to be done

Thank you for your attention

