Agricultural holdings identified through the Italian 6° Agriculture General Census: methodologies adopted for their localization

Giampaola Bellini
Istat - Italy

The production & use of spatial statistics for sustainability
European Forum for Geostatistics Conference (EFGS)
Statistics Portugal, Lisbon, 12-14 October 2011

6° Censimento Generale dell'Agricoltura | Istat
RACCOGLIAMO RISPOSTE, SEMINIAMO FUTURO.

PRESENTATION OVERVIEW

- **EU Reg. 1166/2008 requirements for census agricultural holding location**
- **Approach adopted in Istat**
- **Description of quality through metadata**

EU Reg. 1166/2008 requirements for census agricultural holding (AH) location

- It is necessary to provide the AH location coordinates (referred to ETRS89) to the nearest 5 minutes (approximately 3,000-7,000 hectares, depending on the location in Europe)

- Disclosure issue

2 AHs per geographical coordinate

availability of agricultural census 2010 data at a very detailed territorial level (for previous censuses the minimum level was the municipality – administrative units)

Mixed approach adopted in Istat /1

Choice was based on availability of:

- census map made by Istat
- digital cadastral maps
- commercial geo-referenced address registry

Istat has experienced spacially referenced census data production since 1991, based on census enumeration areas

Mixed approach adopted in Istat /2

Location information required by questionnaire:

- **address** *and*
- **cadastral identification code**

of the holding headquarter

Both information will be used to geocode the AHs to the 2011 Census enumeration areas

EA centroid coordinates released to Eurostat

Mixed approach adopted in Istat /3

Meeting EU precision requirement

preliminary checking of
- 2001 EAs areas (only around 1.400 units beyond the thresholds, over around 384.000)

- cadastral sheet maps & parcels characteristics

Municipalities by cadastre type

Cadastral typologies	N	%
Ex-Austro-Ungaric cadastre (TN & BZ)	333	4,1
With open map sheet perimeter	690	8,5
Ordinary	7023	86,8
Tavolare	48	0,6
Total	**8094**	**100,0**

Parcel

Map sheet

Mixed approach adopted in Istat /4

Meeting EU precision requirement

Cadastre maps *follows*

completeness in terms of municipalities mapped

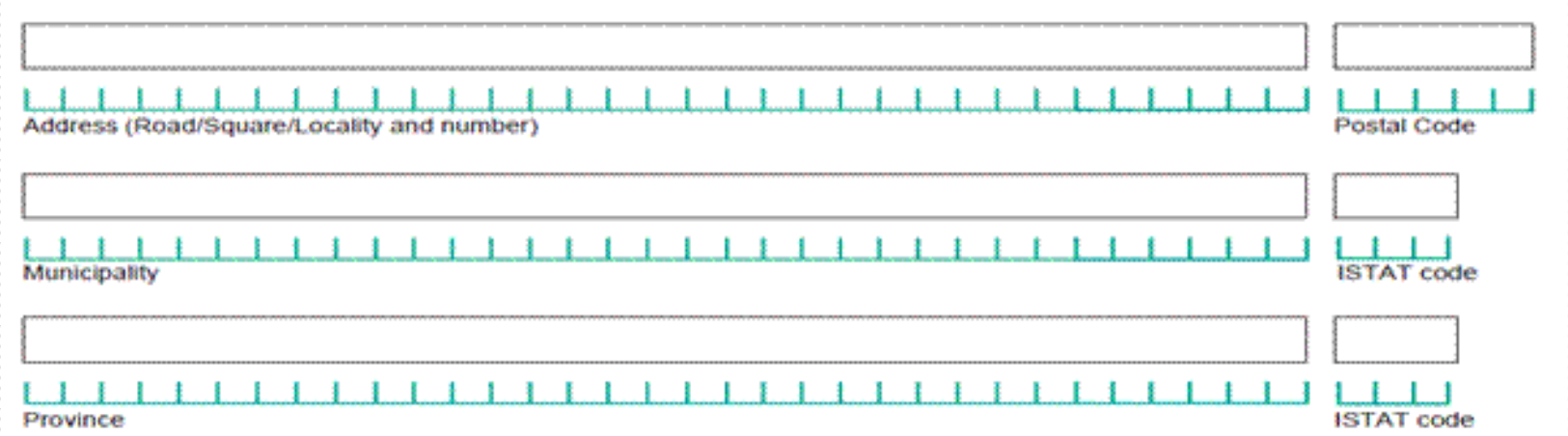
logical accuracy (1st level): an overlay operation made possible to check centroid falling out of national territory and between enumeration area polygons (0.6% out of total involved units)

logical accuracy (2nd level): check the list of the cadastral communes spatially joined with the administrative ones

Mixed approach adopted in Istat /5

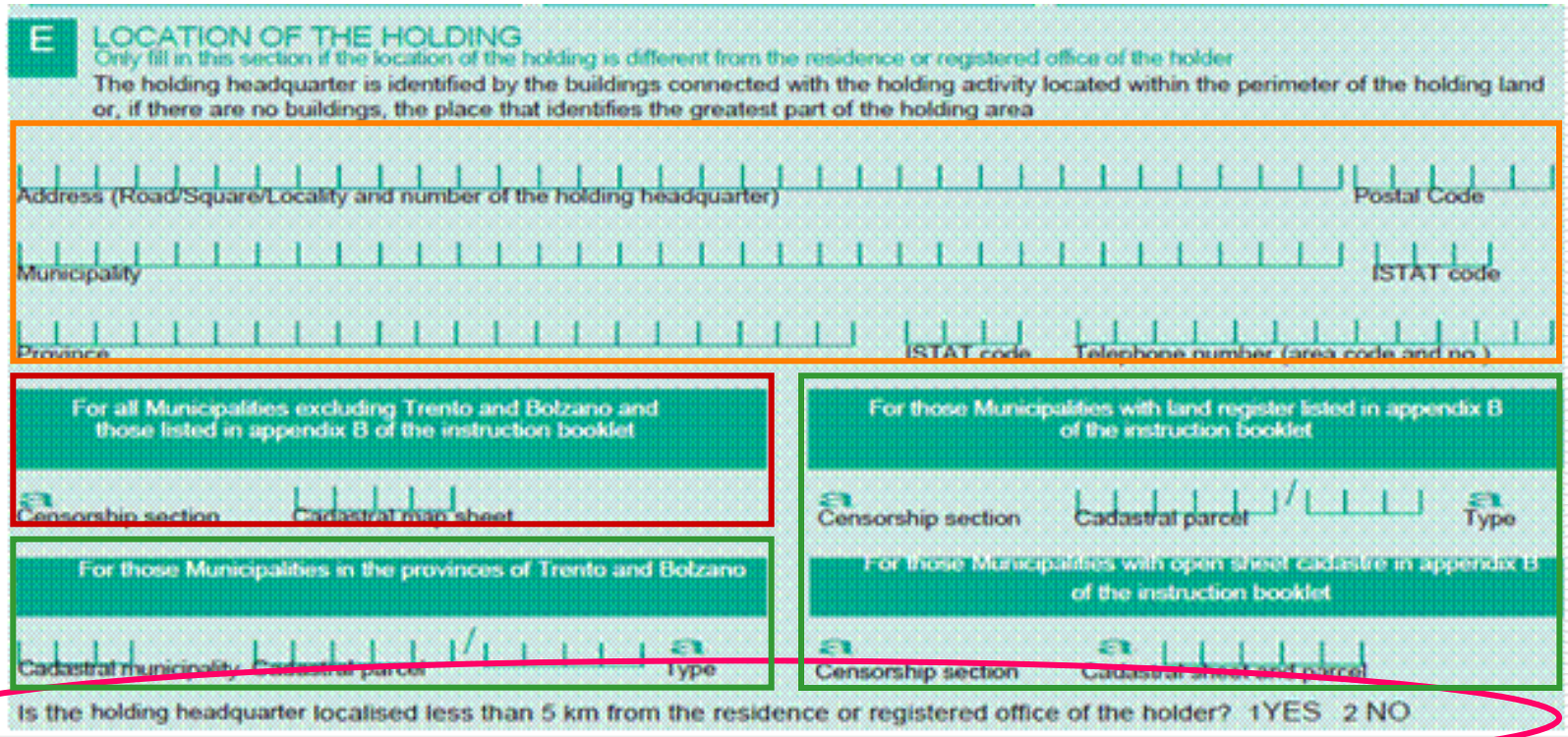
Questionnaire SectionA:

Holder residence address



Mixed approach adopted in Istat /6

Questionnaire SectionE (to be filled in case the HH location is different from the holder residence):



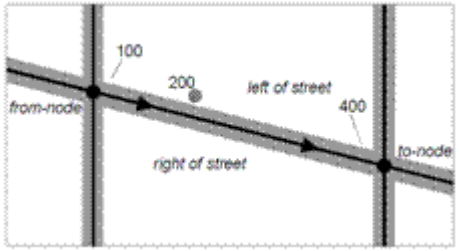
Address matching (Egon software)

Normalisation – geo-referencing process

Pre-normalised address	Normalised address
FRAZ. ROCHEFORT, 18	LOCALITA' ROCHEFORT,18
FRAZ. LIGNOD	FRAZIONE LIGNOD
FRAZ. LIGNOD-RUE TRACIASA, 45/3°	RUE TRACIASA,45/3
VIA S. GOTTARDO 4	VIA SAN GOTTARDO,4
VIA CASTELLO 1, FRAZ. POIA	VIA CASTELLO,1
Dorf 4, 39010 Kuens	DORF 4,39010 KUENS
P.ZZA G. B. LAMPI 2	PIAZZA GIOVANNI BATTISTA LAMPI,2
MEDEAZZA 6	LOCALITA' MEDEAZZA,6
Frazione Gramizzola	LOCALITA' GRAMIZZOLA
Frazione Cabosa	FRAZIONE CABOSA
VIA FONTE AVELLANA 140	VIA FONTE AVELLANA,140
Piazza Flora n. 1	PIAZZA GIUSEPPE FLORA,1

Discarded

geo-coding process



2001 EAs
(transposition table to 2010 EA)

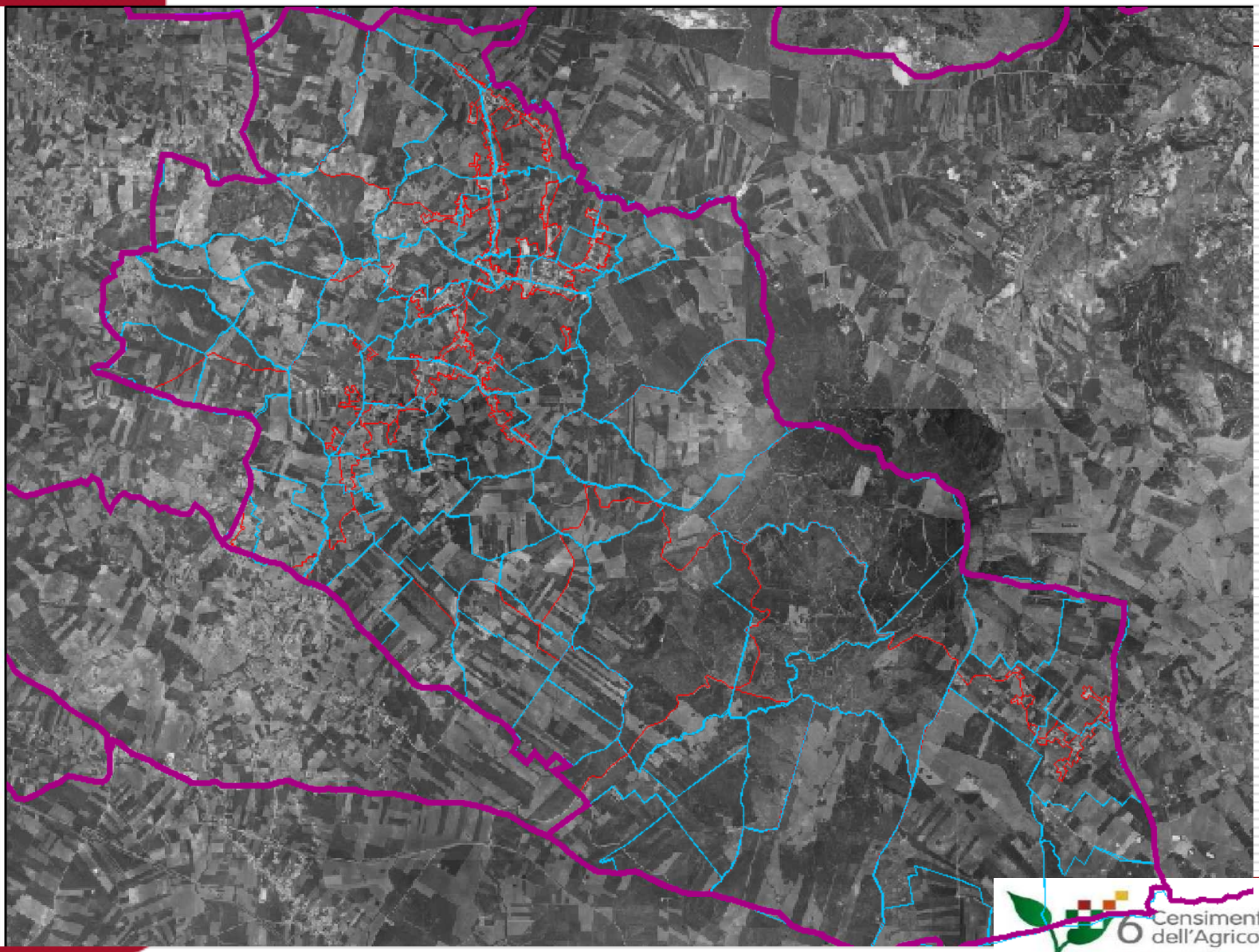
Address matching (Egon software) / 2

Another Egon function is the generations of codes :

for matching records to identify the quality of geocoding
'0' & '1' best and least quality for geocoded units
'2' & '3' for not geocoded units

for discarded records to characterize the exclusion criteria applied

Cadastral maps and census enumeration areas /1



Municipality border

Cadastral map sheets

Enumeration areas

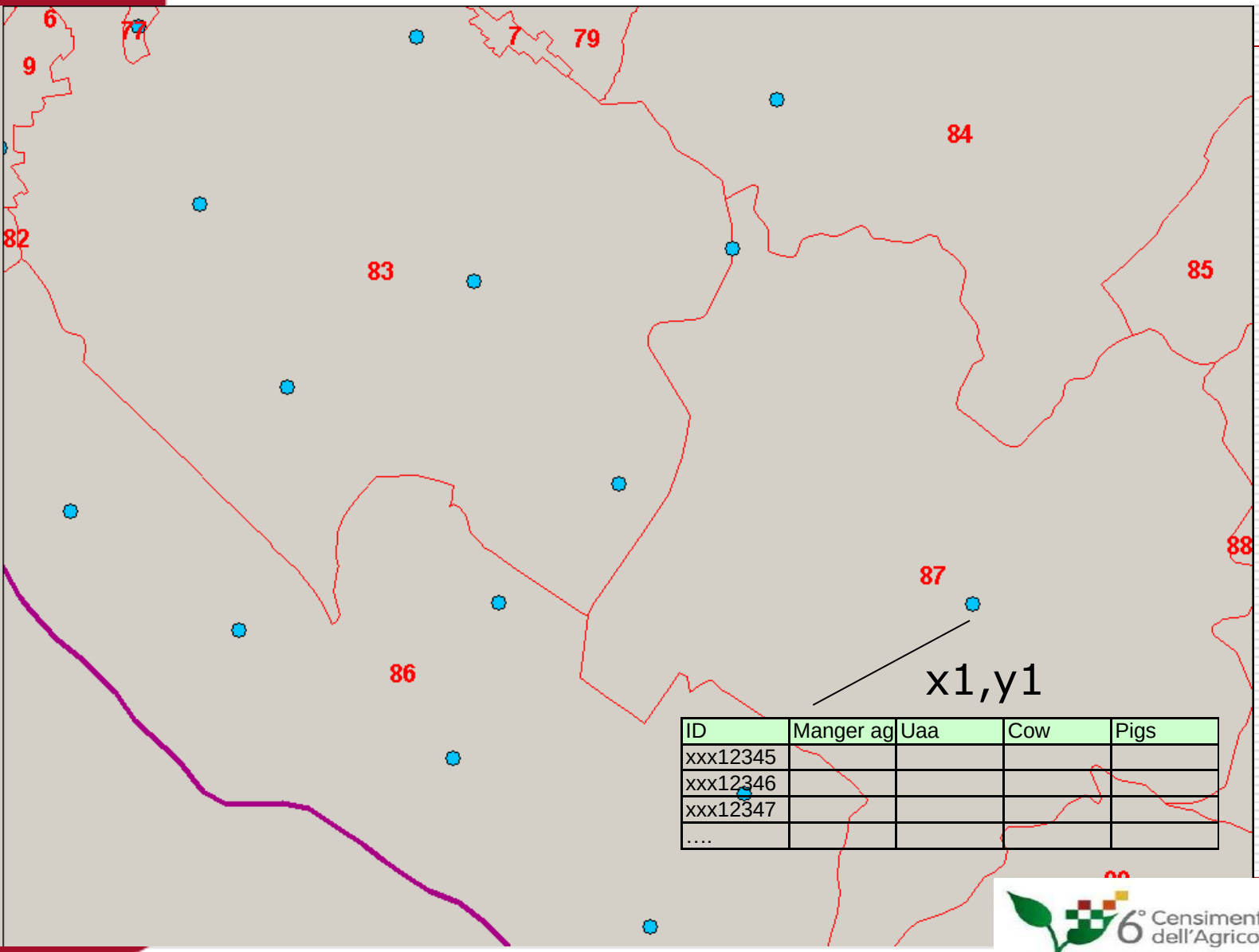
By Fabio Lipizzi

6 Censimento Generale dell'Agricoltura | Istat

RACCOGLIAMO RISPOSTE, SEMINIAMO FUTURO.

Geo-coding through cadastral map : "centroid" method

/2



ID	Manger ag	Uaa	Cow	Pigs
xxx12345				
xxx12346				
xxx12347				
….				

Cadastral map sheets

Cadastral map centroid

Enumeration area

By Fabio Lipizzi

6° Censimento Generale dell'Agricoltura | Istat

RACCOGLIAMO RISPOSTE, SEMINIAMO FUTURO.

Geo-coding through cadastral map /3

1 st step: extracted cadastral maps centroid coordinates are spatially joined to census enumeration areas (and to their centroid coordinates)

2 nd step: a unique key code is created (administrative and cadastre code)

3 rd step: the code created at step two is matched to the information collected through the questionnaire

Information collected on HH by cadastral typologies
6° General agriculture census (provisional data for 16 regions)

Cadastral typologies	N	%
Ordinary	414,299	37.2
'Tavolare'	152	0.0
Ex-Austro-Ungaric cadastre (TZ e BZ)	12,034	1.1
With open map sheet perimeter	1,937	0.2
Without cadastre insert or discarded	684,574	61.5
Total	**1,112,996**	**100.0**

Assigning geographic coordinates to residue cases

Use of the information referring to the

'distance of less than 5 km‘

In other words, if there’s the information on the 'distance of less than 5 km', to all these units will be assigned the coordinates achieved through residence address matching, under the hypothesis that the holder’s residence can be considered as the reference place of the AH

attention

For all cases treated so far we'll generate 2 pair of coordinates (ETRS89-UTM):

- detailed coordinates
- generalized coordinates (released to Eurostat)

Assigning EA centroid coordinates to residue cases

through an allocation method
(*deterministic or stochastic*)

before identifying the value to be attributed, it is better to bring back data to homogeneous subsets, cells, or classes of imputation, obtained through a series of appropriate layers

Final data quality and metadata production /1

first parameter

(*detailed coordinates*)

1 by address matching

2 by cadastral identification code matching

3 by address matching, for holder residence that is within 5 km from HH

second parameter

(*coordinates released to Eurostat*)

1 for EA coordinates generated through spatial join of the two previous cases

2 for EA coordinates generated from the third previous case and from data imputation

Final data quality and metadata production /2

two parameters on '***spatial fragmentation***' of the AH in terms of

Land units

Municipalities

will be implemented referring to the number of units involved (both for land units and municipalities)

both information have been collected through

Agriculture Census questionnaire

Final data quality and metadata production /3

AHs by number of municipalities in which are located

municipality number	AH	
	a.v.	%
1	890.816	80,0
2	167.412	15,0
3	39.342	3,5
5	10.200	0,9
6-10	5.072	0,5
11-20	135	0,0
21-30	9	0,0
31-40	6	0,0
more than 40	4	0,0
Total	1.112.996	100,0

6° General agriculture census (provisional data for 16 regions)

Conclusion

Geocoding to EAs

- meets EU requirements

- makes agriculture census results potentially integrated in the existing census map

further analysis is necessary in order to exploit the potentality offered by this spatially referenced data set

Thank you for paying attention!

Mixed approach adopted in Istat

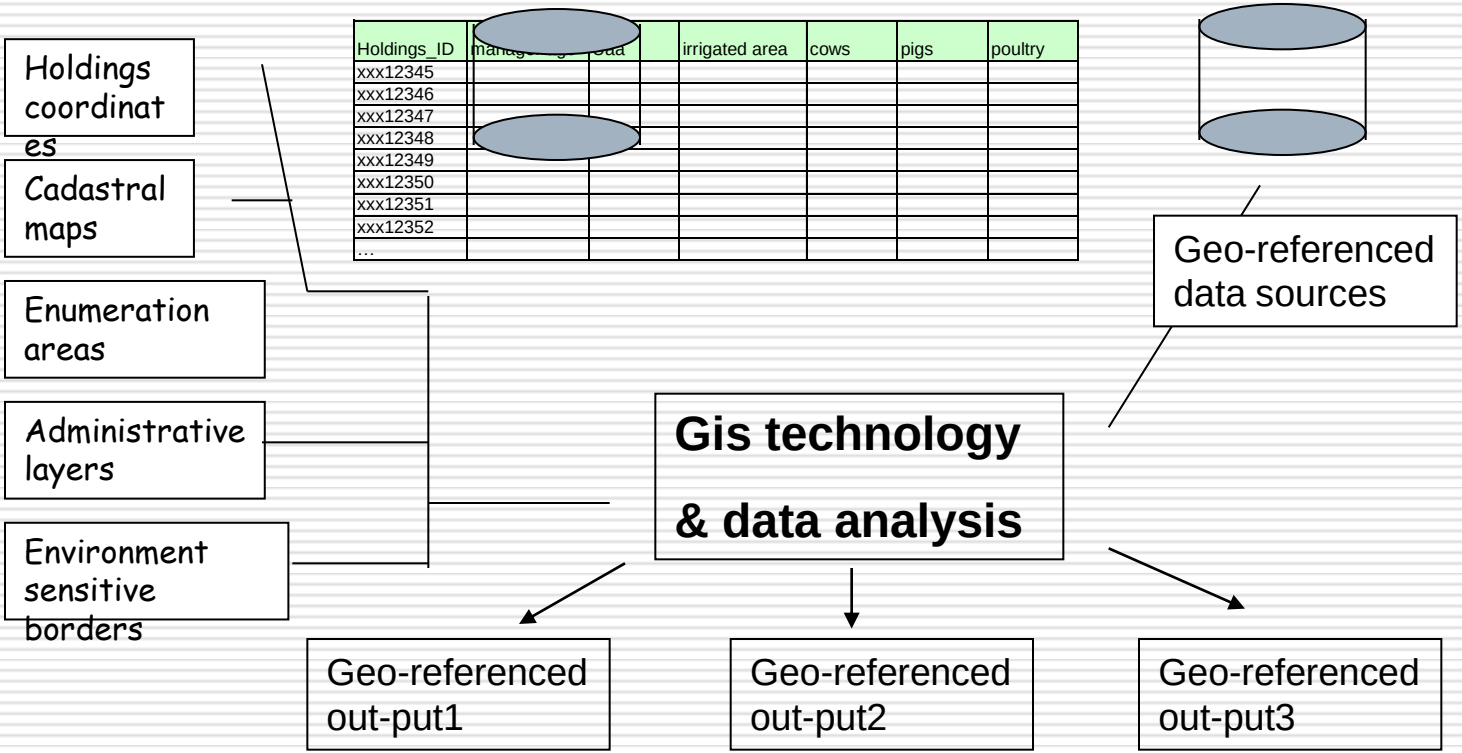
Meeting EU precision requirement

EAs and average area by locality (population census 2001)

TYPE OF LOCALITY	Enumeration area	Average area (ha)
Centri (urban areas)	258.646	6,6
Nuclei abitati (urban areas)	39.393	4,1
Case sparse (rural areas)	82.902	342,5
Productive localities	2.603	29,3
Total	383.544	79,1

What is the need for agriculture holding location

Holding location in Agriculture census 2000 required by Regulation : holdings headquarter by municipality



REGULATION REQUIREMENTS /2

Concerning the location of the holding two principles are applied:

(i) precise coordinates are not required: the longitude and latitude coordinates will not be required in terms of seconds or decimal fractions of minutes. It will be necessary to provide the location only to the nearest 5 minutes, which represent a land area of approximately 3,000-7,000 ha, depending on the location in Europe.

ii) a location with just one holding will be recoded: It is possible that in areas with very large holdings, the location specified to the nearest 5 minutes longitude and latitude may contain only one holding which would therefore be directly identifiable. To ensure that direct identification cannot take place, the locality with only one holding should be allocated to the nearest neighbouring point (chosen at random) with at least one another holding. If any of the 8 neighbouring locations do not have at least one holding, the neighbouring locations have to be extended until others are located with at least one other holding. However, whenever it is possible the agricultural holding should be allocated to the NUTS 3 region where it is situated.

APPROACHES UNDER ANALYSIS IN ISTAT

❖ Reference to administrative records such as cadastre maps and related databases managed by geographical information system techniques. Coordinates derived per holding will be the centroid of the cadastral map sheet or parcel.

• Conversion of the address to the latitude and longitude coordinates by appropriate software.

The data on location of the agricultural holding shall be provided in the European Terrestrial Reference System 1989, usually referred to as ETRS89. ETRS89 is the EU-recommended frame of reference for geodata for Europe.

A different reference system can be used providing the code indicating which one has been used.

GEO-REFERENCING CADASTRAL REFERENCE AND ADDRESSES

- ✓Geocoding throuhg cadastral maps
  - Characteristics
  - Record linkage
  - Precision issue (rough estimates: for cadastral map sheet 600 linear meters; for parcel around 100 meters)
- ✓Georeferencing addresses through Egon
  - Addresses processing
  - Precision issue

Cadastral map characteristics and related activity

Projection – Gauss Boaga Roma40 for all provinces except for Bozen and Trento (ETRS89)

Scale 1:4000

File format *cxf*

340.000 map sheets organised by province

Layers available: map sheets; parcels; buildings

Map processing

Cxf = > shape file

Coordinates transformation = > UTMED50 (census mapping units)

PRELIMINARY CONCLUSIONS AND FUTURE DEVELOPMENTS

HH location Questionnaire section content:

- Address
- Cadastral reference at map sheet level
- Distance between HH and HR (< 5KM=Y/N)

Pre-compilation of the HH location section:

Identification criterion of the HH

- HHbsUaa: building (where present); maximum(uaa) cadastral map sheet
- The procedure should run on the list of pre-census holdings ID, integrated with building information (source cadastral layer)

Pre-compilation with cadastral reference at map sheet level

- *?for all holdings*
- *?only HH matching with AC2000 and H_ID with a unique map sheet*

Georeferencing activity

- Activity run according to best location precision achieved
- Generalisation of the coordinates obtained

DATA RELEASE PLAN FOR AC2010 – FIRST PROPOSAL

Ordinary publications by administrative level

Use of ***GIS*** for the investigation of the spatial dimension of agriculture sector and agri-environmental phenomena (SAPM data available at census level)

Territorial level of interest

=> enumeration areas: integration with other census data

=> river basin areas (**EU Water Framework Directive**) and related Groundwater Directive (2006/118/EC)

=> nitrate vulnarable zones Nitrates Directive 91/676/EEC

=> protected area

= > any other area of interest

Thematic publications: integration with other available spatial data (i.e. water use in agriculture)

Coordinates generated by Egon per type and quality level

Code	Quality	Type of coordinates generated
A10	High	Street graphs Teleatlas
A20	High	Street graphs Navteq
A30	High	Provided by the client
M10	Medium	Centroid of the enumeration area 2001
M20	Medium	Centroid of the enumeration area 1991
M30	Medium	Centroid of the postal code area (zonato) and for the municipality multi-postal area Teleatlas
B10	Low	Centroide of the fraction/locality
B20	Low	Centroid of the municipality
B30	Low	Centroid of postal code area (for municipality mono-postal area and for postal area multi-municipality)