Usefulness of GIS in Strategic Planning

Summary of a lesson within an international planning workshop at the IUAV

Markus M. Hedorfer

Venice University of Architecture — Department of Urban Planning
Santa Croce 1957, I–30135 Venezia VE
Telephone: ++39–041–2572237
e–mail: hedorfer@cidoc.iuav.unive.it
FTP–site: ftp://cidoc.iuav.unive.it/pub/hedorfer

6th July 1999

Contents

List of Figures

List of Tables

Introduction

At the present day, many planners deal with geospatial information systems (GIS)[1] even if it is not always clear what exactly such a system 'does' or — in other words — which operations can be performed with this technology that is still relatively new to so–called end–users. In order to intoduce usefulness of geospatial information systems in urban and regional planning, the following three statements can be considered.

1. Each planner handles geospatial information.
2. Some planners use geospatial information systems.
3. Few planners know how to manage correctly digital geospatial information.

As misunderstandings in the field of GIS produce often serious data consistency errors, any consideration on usefulness of GIS in urban and regional planning should first be introduced by examining what exactly geospatial information handling involves. So the above–mentioned statements should be translated to the following questions to be deepened during this contribution.

1. What is geospatial information?
2. What are geospatial information systems?
3. How to manage correctly digital geospatial information?

Each section of the present contribution will treat one of these three questions. Obviously it is not possible to perform a scientific dissertation here, but only some generic indications that should help planners without any experience in GIS to know approximately what kind of support such an information system can provide.

1 What Is Geospatial Information?

The question about what can be considered geospatial information and what not, has no simple answer considering the problem from a pure scientific point of view. However, it can be said — and nobody would contradict this — that geospatial information is in any case

information that can be located on the earth's surface.

That involves that geospatial information is composed at least by the following two complementary kinds of data (see also Table 1 and Figure 1).

Locational Data Those data that encode the geometrical position and extent of a spatial object. For example, a rectangular residential building may be described by providing the x– and y–coordinates of its four vertexes, the amount of square meters of its area, and the total length of its perimeter.

[1] Usually the acronym GIS means "*geographical* information system". However the Geographical Data Committee (FGDC) at the United States Geological Survey (USGS) proposes the term "geospatial" that defined in a slightly more correct way the nature of this kind of information systems (for this purpose see e.g. the United States Content Standards for Digital Geospatial Metadata [1]).

Table 1: Example of locational, attribute, and topological data

Locational Data	Attribute Data	Topological Data
Boundings:	Class: "Residential Building"	Building 3–1 borders with:
Vertex 1 at (30; 20)	Height: 14.50 m	building 2–1
Vertex 2 at (20; 40)	Volume: 10,600.00 m	paved area 2–2
Vertex 3 at (60; 60)	Ground–Floor Height: 3.00 m	paved area 3–2
Vertex 4 at (70; 40)	Number of Floors 3	paved area 4–2
Area: 800.00 m^2	Number of Basement Floors: 1	building 4–1
Perimeter: 120.00 m	Attic Height: 2.50 m	building 5–1
	Number of Apartments: 12	garden area 5–3
	Inhabitants: 38	garden area 3–3
	Residential Volume: 27,000.00 m^3	
	Commercial Area: 750.00 m^2	

Attribute Data Those data that provide for non–metrical information about a spatial object. The imaginary residential building could hence be described by providing the discrete object class "residential building", medium height, total volume, etc.

Differing to usual CAD–software, geospatial information management software should also — even if this is not always the case — provide for a third kind of information.

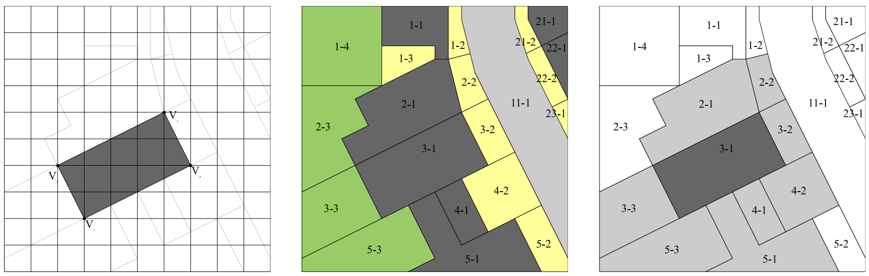
Topological Data In a very similar way to the meaning of "topology" in geometry, topological data within a GIS encodes information related to a spatial object's environment. That is, considering again the residential building, it could be said that it borders with several other areas whose locational, attribute or topological data can be used to perform specific spatial queries (questions).

That means that geospatial information lies in a close relationship with any query containing both, a *where* and a *what* component. Some examples may be the following.

- *Q: Where is the railway station?*
  A: It's about 500 meters from here moving to south west.
- *Q: Where is the City of Venice?*
  A: It's approximately at 12.33° East and 45.41° North.
- *Q: What area has the Veneto Region?*
  A: It has an area of about 18,364 km^2.

These three questions represent the most simple kind of spatial query where only absolute or relative *metrical quantities* are returned by the answers (500 meters, 12.33° East, 45.41° North, and 18,364 km^2). Other spatial queries may be the following.

Figure 1: Graphical representations of locational (on the left), attribute (in the middle), and topological (on the right) information



— *Q: By which Provinces is composed the Veneto Region?*
  A: It's composed by the Provinces of Venice, Padua, Vicenza, Verona, Rovigo, Treviso, and Belluno.
— *Q: Which Regions does the River Po cross?*
  A: It crosses the Regions of Piedmont, Lombardy, Emilia–Romagna, and Veneto.
— *Q: With which Regions does the Veneto Region share its border line?*
  A: It shares its border line with the Regions of Emilia–Romagna, Lombardy, Trentino–South Tyrol, Tyrol (Austria), Carinthia (Austria), and Friuli–Julian Veneto.

In this case the answers involve *spatial relationships* between different spatial objects: if you know where are located single provinces and if you know where are located the administrative regions, thus you also will know which provinces fall inside which regions. The same reasoning is done for interesctions between regions and rivers and between regions and their common border lines. Geospatial information systems, however, are able to return some more complex queries as the following.

— *Q: Which Townships are located completely or partially inside the 30–minutes isochronal of the Venice railway station using the public transportation system?*
  A: The Townships of Venice, Mira, Dolo, Pianiga, Vigonza, Padua, Spinea, Mirano, Martellago, Salzano, Noale, Mogliano Veneto, Preganziol, Treviso, Marcon, Quarto d'Altino, and Roncade.
— *Q: Is there a place in the Venice Township where can be located a factory of at least 100,000* m^2*, that produces some noise, and that needs to be connected by industrial railways?*
  A: Yes, there is such a place. It can be found in the southern part of the Porto Marghera industrial area.

Obviously, these spatial queries require that somewhat like 'artificial intelligence' should be implemented in the system: in the first case, a procedure must be provided that 'knows' how to compute a 30–minutes isochronal from a specific point on a network and that then performs a *polygon overlay* process to retrieve township polygons. In the second case, a particular polygon selection sequence must be prepared to isolate those polygons that fulfill given characteristics encoded as polygon attributes; it is probable that also urban planning prescriptions should be encoded to successfully perform this kind of query.

2 What Are Geospatial Information Systems?

Also for this problem, there is no clear and simple answer. Geospatial information theorists do not agree to a unique definition of GIS, neither there is a commonly accepted set of functions that each GIS should be able to perform. Nevertheless, it can be declared that

geospatial information systems are those computer-aided data management systems that are able to perform the previously mentioned spatial queries in a completely automated way.

In addition to this statement almost every GIS expert will agree that any geospatial information system should be composed of the following three components.

1. GIS Software. Software that makes you *able* to perform geospatial analysis.
2. Geospatial Data. Data to be analysed and to be used to produce *other* data.
3. Geospatial Data Handling Procedures. Procedural encoding to *perform* geospatial analysis.

This means that a GIS can never be identified only by its software part even if somebody — especially producers of low–performance 'user–friendly' GIS software — sometimes try to make believe just in this. On the contrary, the development of a GIS is a very complex activity that needs to be well planned either from the data base organisation point of view, and considering cost–benefit questions.

At this point it would be necessary to understand better what geospatial information systems exactly are — and to demonstrate the assumptions made — by developing a detailed reasoning using scientific criteria. But as this is not possible within this contribution, the few statements made until now must suffice to go on with the excursus on usefulness of GIS in planning, even if it is highly recommended to deepen the question elsewhere.

3 How To Manage Correctly Digital Geospatial Information?

While GIS theorists do very hardly agree on the previously introduced questions, the problem how digital geospatial information should be handled correctly is even more complex. Different approaches can produce equally good results, so it becomes extremely difficult to maintain that method A is better than method B.

This from the GIS expert's point of view. However, for who is approaching the world of geospatial information for the first time, things may appear in a slightly different manner. First of all, two contrasting management techniques or models can be distinguished:

1. Layer–Based Model. Geospatial information is splitted into several graphics–oriented geometrically independent object assemblage.
2. Attribute–Based Model. Geospatial information is stored as a two–dimensional continuum of geometrically dependent feature classes.

It can also be said that the layer–based model represents a more simple way that the attribute–based one to store digital geospatial data without caring too much about data consistency, maintenance, and organisation. The layer–based model can also be defined as the management technique preferred by operators who previously worked with computer–assisted design (CAD) software that uses information layers to distinguish objects that need to be treated (graphically rendered) in different ways.

On the other hand, the attribute–based data management model is a clearly more analytical way store digital geospatial data. Usually people who did not have significant experiences with CAD but consider spatial objects more from a quantitative geographical aspect are more inclined to use feature attributes to build data structures.

3.1 Layer–Based vs. Attribute–Based

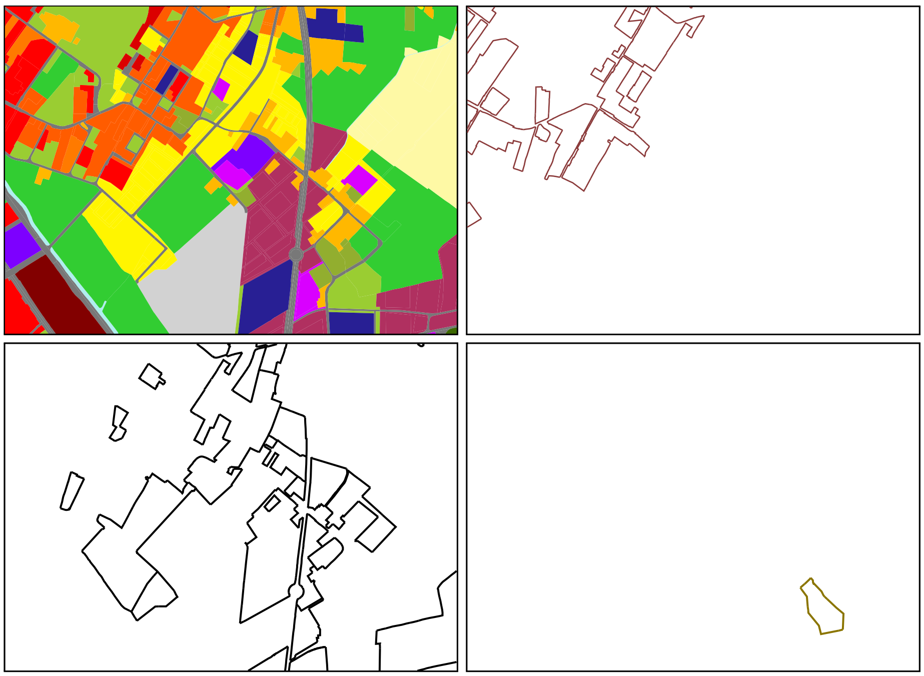
To illustrate the difference between these two data implementation techniques, it may be useful to consider a specific case and to show how the two manners deal with the related geospatial information. The example that will be considered refers to the geospatial information system developed for the Venice Township Administration in order to manage the General Programme of Urban Planning[2]. A detailed technical description is provided by the author of this contribution and Massimo Mazzanti [2] while a simplified graphical rendering of the main zoning map is given in Figure 2.

This graphical representation shows essentially four different informations: (1) *zoning classes* rendered using different colours, (2) *urban renewal areas* distinguished by brown borders, (3) *action plan delimitations* given by black borders, and (4) *woodland residues* indicated by olive green borders. Considering for the moment only graphical aspects both, the layer–based and the attribute–based data model are equally able to produce the same result. In the first case, the data administrator would prepare four different drawings (see Figure 3) containing respectively zoning polygons, urban renewal borders, action

[2]General Programm of Urban Planning and its acronym GPUP are translations made by the author for the italian term *Piano Regolatore Generale* and its acronym *PRG*.

Figure 2: A simplified graphical rendering of the Venice–Mestre General Programme of Urban Planning

Figure 3: Layer–based data encoding of the Venice–Mestre General Programme of Urban Planning



plan borders, and woodland residual borders and then display one above the other with a particular order to avoid that filled polygons would cover border lines or that thick border lines would cover thin border lines.

In the second case instead (attribute–based model) graphical rendering procedures will be more complex due to a data structure that is not graphics–oriented but tries to imitate real–world phenomena. Considering again the final graphical result to obtain (Figure 2), it can be observed that urban renewal areas, action plan perimeters, and woodland residuals are not *geometrically independent* from the main zoning; that is, spatial objects of any of these four 'layers' may share common border line segments. Translating this circumstance to real–world human perception capabilities, an urban planner observing 'his' land from the ground would be able to notice how a single land parcel has several characteristics, like its land–use destination (the zoning class), its housing quality (is it *part* of an urban renewal area or is it not?), its operational destination (is it *part* of an action plan or is it not?), or its ecologic quality (is it *part* of a woodland residual or is it not?). Never he or she would think about the parcel by performing some graphical overlay and saying "it's very strange, but by some chance the outer limit of this urban renewal area is the same as the outer limit of my parcel. . . "

So the goal of the attribute–oriented data model is to reproduce the real–world characteristics by encoding them as attributes in a so–called *feature attribute table*. For the case of land parcels, the feature class is "polygon" and thus the respective table is called *polygon attribute table*. As shown in the upper part of Figure 4, the polygon attribute table for the shaded polygons of Figure 5 is made up of the following nine different fields (items).

`RECORD` Progressive record (row) numbering,

`ZTO-ID` Progressive zoning numbering ("ZTO" is an acronym for the Italian expression *Zona Territoriale Omogenea*, Homogeneous Territory Zone),

`ZTO` Zoning class,

`RPU-ID` Progressive urban renewal area numbering ("RPU" means *Riprogettazione del Paesaggio Urbano*, Urban Landscape Re–Planning),

`RPU` Urban renewal class,

`SAF-ID` Progressive action plan numbering ("SAF" means *Strumento Attuativo di Futura Istituzione*, Future Action Planning Instrument),

`SAF` Action plan class,

`RSB-ID` Progressive woodland residual numbering ("RSB" means *Residuo Boschivo*, Woodland Residual), and

`RSB` Woodland residual class.

Graphical rendering of this kind of polygon data organisation is done by selecting, for example, all adjacent polygons that have the same value in item `RSB-ID` and displaying a common border line.

Considering again Figure 4 — this time its lower part — it can be observed that the attribute–based model needs also to be *normalised*: in order to avoid redundant data (e.g. the couple of values `ZTO-ID = 2420` and `ZTO = 'vpr'` is present three times), a new table will be created where each combination of values between object numbering (e.g. `ZTO-ID`) and object class (e.g. `ZTO`) is present only once, the object class item (`ZTO`) is deleted from the polygon attribute table and retrieved by relating the latter to the new table using as relate items the object numbering item (`ZTO-ID`) present in both tables. In this way it is ensured that data inconsistency cannot be accidentally produced and attribute data quality is guaranteed.

3.2 Topology–Oriented Attribute–Based Data Encoding

Another example of attribute–based data encoding — using topological data structures — is given by the main morphology (see Figure 6) carried out within the Venice Lagoon Experimental Geospatial Information System at the IUAV (SISALV) [3].

In this case, only three polygon attributes (see Table 2, upper part) have been used to encode morphological polygon classes. Subsequently topological information has been used to hard–encode inside the line attribute table (see Table 2, lower part) each of these three polygon attributes for the left polygon and the right polygon of each line element. So nearly any possible selection — used to perform generic spatial analysis or to generate graphical representations

Figure 4: Attribute set system of Venice–Mestre General Programme of Urban Planning

RECORD	ZTO-ID	ZTO	RPU-ID	RPU	SAF-ID	SAF	RSB-ID	RSB
5465	2420	vpr	0	no	0	no	0	no
5476	2421	vpr	0	no	0	no	0	no
3243	6516	vpr	0	no	0	no	0	no
3248	6507	vpr	0	no	0	no	0	no
2839	2731	vpr	0	no	0	no	0	no
2774	2586	vpr	0	no	0	no	0	no
2293	6566	vpr	0	no	0	no	0	no
5466	2420	vpr	0	no	0	no	0	no
5467	2420	vpr	0	no	0	no	2	yes

$\Downarrow$

RECORD	ZTO-ID	RPU-ID	SAF-ID	RSB-ID
5465	2420	0	0	0
5476	2421	0	0	0
3243	6516	0	0	0
3248	6507	0	0	0
2839	2731	0	0	0
2774	2586	0	0	0
2293	6566	0	0	0
5466	2420	0	0	0
5467	2420	0	0	2

ZTO-ID	ZTO
2420	vpr
2421	vpr
6516	vpr
6507	vpr
2731	vpr
2586	vpr
6566	vpr

SAF-ID	SAF
0	no

RPU-ID	RPU
0	no

RSB-ID	RSB
0	no
2	yes

Figure 5: Attribute–based data encoding of the Venice–Mestre General Programme of Urban Planning

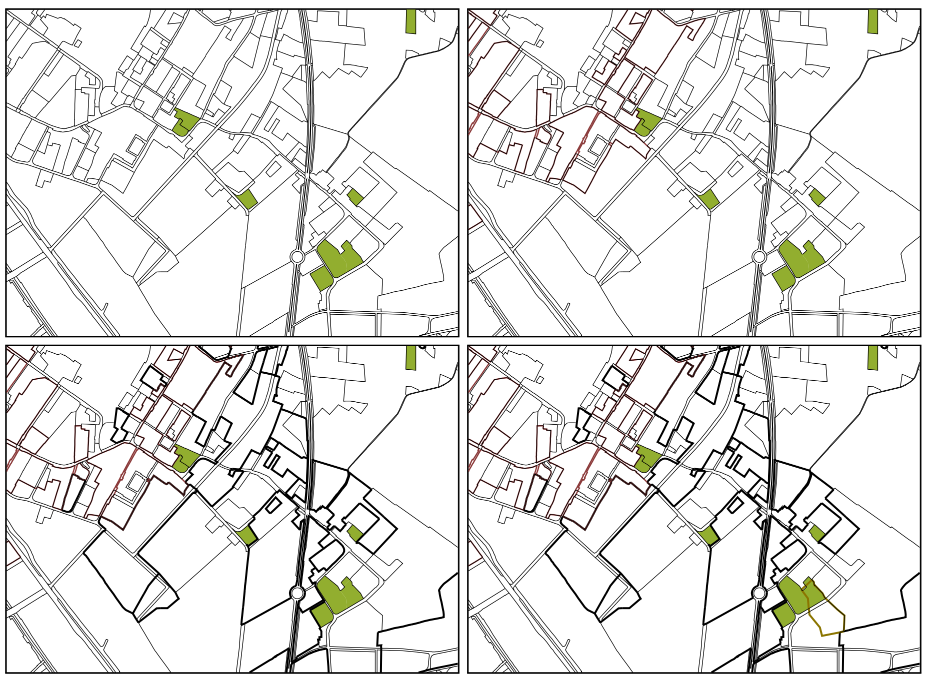


Table 2: The attribute set adopted by the SISALV for its "Main Morphology" (simplified)

Items of the polygon attribute table	
TIPO	Surface descriptor (numeric)
VALLE	Fishery number
LAGUNA	Inside/outside lagoon delimitation
Items of the line attribute table	
LP_TIPO	Value of item TIPO of the left–polygon
RP_TIPO	Value of item TIPO of the right–polygon
LP_VALLE	Value of item VALLE of the left–polygon
RP_VALLE	Value of item VALLE of the right–polygon
LP_LAGUNA	Value of item LAGUNA of the left–polygon
RP_LAGUNA	Value of item LAGUNA of the right–polygon

— will be possible, as will be illustrated now (see also Figure 7). As dry land polygons are encoded using values between 10 and 19, a selection of all dry land may be performed by executing query

```
tipo >= 10 and tipo <= 19
```

Subsequent selections of low water surfaces ($60 \div 79$), lagoon canals ($80 \div 89$), and open sea ($90 \div 99$) work in a very similar way

```
tipo >= 60 and tipo <= 79
tipo >= 80 and tipo <= 89
tipo >= 90 and tipo <= 99
```

The task becomes a few more complex if those line objects should be select that represent the outer border line of a set of polygon 'atoms' that have the same value in a single item. This is the case of the dry land border lines. The first operation to perform is to retrieve the values of item LAGUNA for the two polygons of the left and on the right side of each line object. This must be done by establishing two relate environments between the conventional LPOLY# and RPOLY#–items of the line attribute table (these two items let a line 'know' in a topological structure which are its left and right right polygon) and the record number of the polygon attribute table. Then polygon attributes are retrieved and stored permanently ('hard–encoded' as said above) in two new line attribute items for each polygon item: item TIPO will be stored as LP_TIPO and RP_TIPO, LAGUNA as LP_LAGUNA and RP_LAGUNA, and VALLE as LP_VALLE and RP_VALLE.

At this point the selection can be perfomed. In the Arc/Info selection language (a language similar to the SQL[3]) this will look like this

```
(lp_tipo >= 10 and lp_tipo <= 19)
```

[3]Structured Query Language.

Figure 6: Polygon structure of SISALV's "Main Morphology"

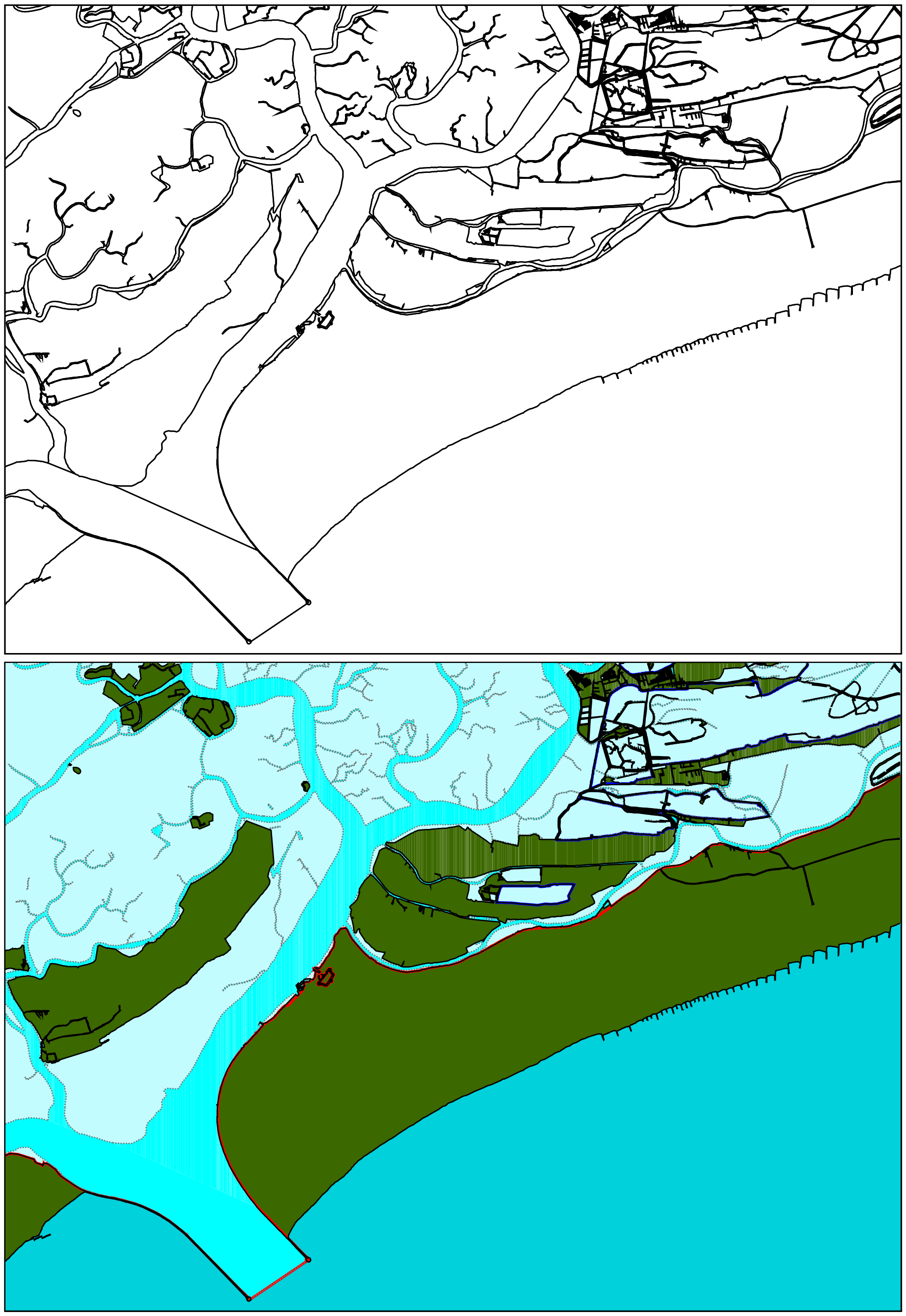
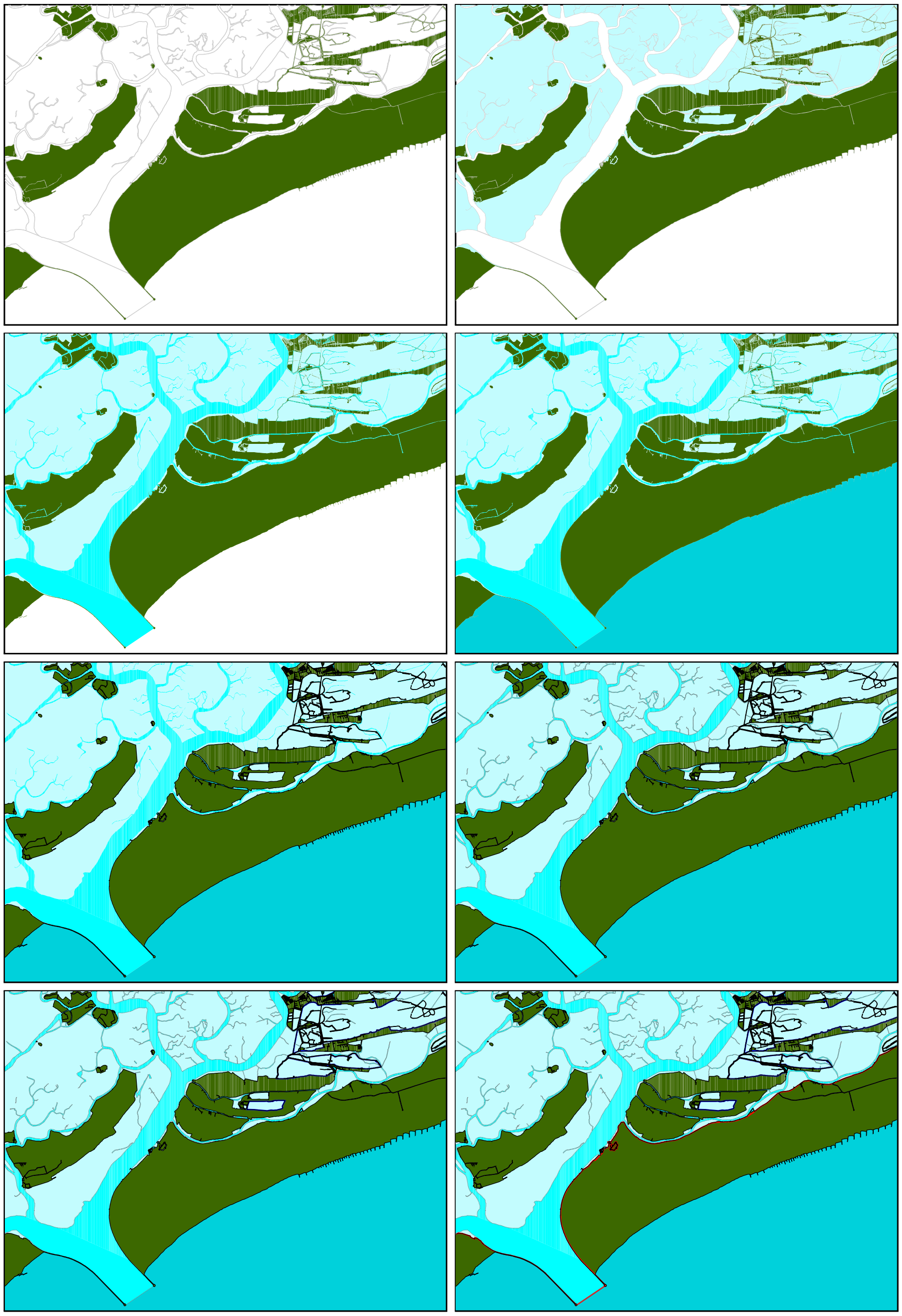


Figure 7: Example of subsequent spatial queries of polygon and attribute data



```
xor
(rp_tipo >= 10 and rp_tipo <= 19)
```

and means

"Please, select all those lines whose left polygons have values between 10 and 19 (dry land) or whose right polygons have values between 10 and 19, but not both at the same time."

A similar but even more complex selection is needed to select the lagoon canal borders (the dotted lines in the graphical representation) because some of canal border line segments are also dry land outlines:

```
(
   (lp_tipo >= 80 and lp_tipo <= 89)
   xor
   (rp_tipo >= 80 xor rp_tipo <= 89)
)
and not
(
   (lp_tipo >= 10 and lp_tipo <= 19)
   xor
   (rp_tipo >= 10 xor rp_tipo <= 19)
)
```

The 'human' meaning of this selection statement is the following.

"Please, select all those lines whose left polygons have values between 80 and 89 (lagoon canals) or whose right polygons have values between 80 and 89, but not both at the same time. To this selection please do not *add those lines whose left polygons have values between 10 and 19 (dry land) or whose right polygons have values between 10 and 19, but not both at the same time."*

Finally to select fishery outlines and the 1990 lagoon delimitation line, the principle is the same as the dry land outline selection, but in the Arc/Info selection statement is quite more simple due to the presence of the two extra attribute items `VALLE` and `LAGUNA`. The fishery outline selection

```
lp_valle <> rp_valle
```

means

"Please, select all those lines whose left polygon has a different fishery number from its right polygon"

and, last but not least, the lagoon delimitation selection

```
lp_laguna <> rp_laguna
```

means

"Please, select all those lines whose left polygon is inside the 1990 lagoon delimitation while its right polygon not or vice versa".

3.3 Attribute–Based Geospatial Analysis

A third example concerns spatial analysis to be performed in order to take some strategic decisions within a General Programme of Urban Planning. Here, the study case is the town of Castelfranco Veneto in the Province of Treviso, Italy [4], and the considered question is explained in the following problem recognition lists.

Purpose A New Residential Settlement Policy.

1. Residential settlements are extremely distributed in sub–urban and sub–rural areas.
2. Residential settlements should be more centralised.
3. Sub–urban distributed population should be transferred to clearly urban and sub–urban areas.
4. Need to distinguish proto–villages to be transformed into rural villages by transferring distributed settlements of the proto–villages' surroundings.
5. Where are the proto-villages?

Problem How To Recognise Proto–Villages?

1. Proto–villages are located in rural and sub–rural areas.
2. Proto–villages are slightly more centralised settlements than other less significant rural and sub–rural settlements.
3. There may be a critical mass of inhabitants necessary to declare a settlement as proto–village.

Data Lack There Is No Detailed Geospatial Data About Population Distribution!

Available data are only:

1. Residential Building Polygons (without volumes and heights)
2. 1991 Census Tract Polygons (with census attribute data)

Performing the Analysis The rules illustrated until now need to be transformed to specific spatial queries in order to produce new data containing the initially required information to perform the planners' decision on residential settlement redistribution.

1. Population Distribution Model (see Figure 8, upper part)
   1.1. To each residential building is attached a reference to the census tract to which it belongs.
   1.2. For each census tract the total built–up area is calculated.
   1.3. Population is distributed between (attached to) residential buildings in order to each building's 'weight' (area percentage of total built–up area).

2. Contiguous Urbanised Surfaces Simulation (see Figure 8, lower part)
   - 2.1. An appropriated buffer around each building is generated in order to simulate contiguous urbanised areas.
   - 2.2. For each contiguous urbanised area the total population is calculated.
   - 2.3. Geospatial information is displayed adopting appropriated population value ranges.

Decision Making Considering the final graphical output (Figure 8, lower part), planners can now take strategic decision on how to redistribute rural, sub–rural, and sub–urban population. A decision–supporting analysis performed within a GIS, obviously cannot substitute human intelligence during a complex decision–making process that involves not only metric, topologic and some attribute data, but also opportunity assessment and discussion, political direction definition and much more...

References

[1] Federal Geographic Data Committee. *Content Standards for Digital Geospatial Metadata*. Federal Geographic Data Committee (FGDC) at the United States Geological Survey (USGS), Reston, Virginia, USA, 1994.

[2] Markus M. Hedorfer and Massimo Mazzanti. *Comune di Venezia — Variante Generale al PRG per la Terraferma (1998) — Sistema Informativo Geografico — Documentazione Tecnica*. Unpublished, {hedorfer,massimo}@cidoc.iuav.unive.it, July 1999.

[3] Istituto Universitario di Architettura di Venezia — Dipartimento di Analisi Economica e Sociale del Territorio. *SISALV — Venice Lagoon Experimental Geographical Information System at the IUAV*, 1996–1998. Research initiative, alberta@brezza.iuav.unive.it.

[4] Franco Posocco, Paolo Ceccon, Marco Ferretto, Vittorio Guardalben, and Markus M. Hedorfer. *Comune di Castelfranco Veneto TV — Piano Regolatore Generale — Variante Generale*, July 1999. Unpublished, posocco@sit.iuav.unive.it.

Figure 8: Population Distribution Model and Contiguous Urbanised Surfaces Simulation