Technical Workshop on Monitoring Seawater Intrusion in Coastal Groundwater

INTRODUCTION TO G.I.S.

Dar es Salaam, 15 November 2011

Matteo ROSSI

SAPIENZA University of Rome (DICEA)

matteo.rossi@uniroma1.it
What a GIS is?

"Information System implemented to store, manage, analyze and present data in a topological and spatial context"

Collection, pretreatment and processing of multi-source spatial data

Maintenance and retrieval of spatial information, with the possibility of editing and updating

Manipulation and analysis, aggregation and disaggregation of data, parameter estimation, modeling

Production of reports and summary data
GIS components

Hardware

Data

Software

Users

Procedures
What GIS can be used for

**AGRICULTURE**

- Farm management
- Pest/Disease tracking
- Crop monitoring
- Yield prediction
- Soil analysis
What GIS can be used for

NATURAL RESOURCES MANAGEMENT

- Forestry
- Ecology
- Mining
- Petroleum
- Water Resources
What GIS can be used for

PLANNING AND ECONOMIC DEVELOPMENT

- Land Use/Zoning
- Emergency Preparedness
- Population Forecast
- Market Analysis
- Property Tax Assessment
- Transportation
Sources of data

Maps

Coordinates

Digital data

GPS

<table>
<thead>
<tr>
<th>Coordinates</th>
<th>Digital Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>5092218</td>
<td>1585921</td>
</tr>
<tr>
<td>5051660</td>
<td>1550442</td>
</tr>
<tr>
<td>5051873</td>
<td>1551146</td>
</tr>
<tr>
<td>5053749</td>
<td>1548144</td>
</tr>
</tbody>
</table>
Synthetic diagram of a GIS

Users → Software + Database → Results

Model

Real world
Principles of GIS

DATA MODEL

Compared to the geometric representation of real objects, a GIS has to maintain and manage all information concerning the mutual spatial relationships between different elements, such as connection, adjacency or inclusion: that is to structure the data defining their topology. In addition to these two aspects the data model, to be effective, must consider the inclusion of the descriptive data of the individual real objects, defined as attributes.

GEOREFERENCING

A fundamental characteristic of a GIS is its ability to georeference data, or to assign to each element its real space coordinates. In other words, the coordinates of an object are stored with respect to any arbitrary reference system (for example 12 cm and 5 from the bottom edge of a map from the left) or relative to the coordinate system of the device used, as the tablet digitizer or video, but they are stored according to the coordinates of the reference system in which the object is actually located (such as 121° 27' lat. E and 41° 53' long. N using the geographic system).
The importance of the reference systems congruence
How does it work?

- Link map features to Relational Databases
- Locate items from the map or by attributes
- Manages sets of elements and attributes as layers

“One to one” relationship between the graphic items and the alphanumeric database records
How does it work?

A GIS model is based on the “layer” concept;

Layers...

... can be superimposed

... are connected by the mapping (geocoding)

... contain consistent information (ie, lithology, hydrography, road network ...)

“Is the geographic world a jigsaw puzzle of polygons, or a club-sandwich of data layers?”

Couclelis 1992
Types of data managed by a GIS

**Spaceless**: do not depend in any way by a spatial context
**Discrete**: finite entities that have value only in a given spatial location
**Continuous**: taking different values in the space
Spatial data typologies

VECTOR

(x, y)

RASTER

GRID
Vector structure

Identify the main territorial elements representing them through points, lines and / or polygons

- All points of geographical entities are geo-referenced.
- Through the topological structure the attributes and information are associated to the geographical entities.
- The topological structure also allows spatial analysis in the study area.
Vector structure

**POINT elements:** entities, activities or events distributed in space identified by a single pair of coordinates
Vector structure

**LINEAR elements**: represented by a series of points of known coordinates, united to form a line

---

**Reality**

---

**Into GIS**

---

Lago Nero

Strada Statale 146

Bosco di Valle Fiorita

P1 \((x,y)\)

P2 \((x,y)\)

P3 \((x,y)\)

P4 \((x,y)\)
**Vector structure**

**SURFACE elements**: represented by a series of lines (combining points) joined in turn to form a closed polygon.
Topology

In a GIS environment, is defined as the set of spatial relations that bind together neighboring or nearby objects

The topological rules GIS features that are required to comply, are a prerequisite for proper data analysis and for their integrity

The analysis and processing of NOT topologically correct geographic data may lead to false results or undermine the procedure implemented

The topology is particularly useful in GIS because many spatial analysis procedures do not require the spatial location but only topological information about objects. Eg, to find the shortest path between two points, it is sufficient to know the list of arcs joining the two points.

"The Seven Bridges of Königsberg“, Eulero 1736
Topology

Poligonal topology

Each layer is divided into a series of polygons, which are stored as independent elements and represented by the sequence of points that define the contours.

Disadvantages

- No relations
- Junction lines are stored twice
Arc-node topology

In this kind of topology all geometries (points, lines and polygons) are recognized. Each element is stored through the polygonal arcs (lines) that make it up which are stored through the points of passage.
Topology

Arc-node topology

It’s a more complex topological organization, but softer in terms of memory required for data storage.

The lines common to adjacent polygons are not duplicated.

Through this organization, you can create relationships between adjacent polygons.

Relational topology
Topology

Some examples of common topological errors in GIS fixable through proper functions of "cleaning" or "building".

- **False intersection of two strings**
- **“Dangling” node**
- **False polygon origin**
- **False intersection between polygons**
**Raster structure**

It's probably the most simple and intuitive organization of geometric data.

The data are represented through matrixes of rectangular elements, whose dimensions are dependent on the resolution of the layer.

The position of the cell in space is the number of row and column with respect to an origin fixed (usually the cell down-left).

It's a representation modality requiring a significative use of memory, so that even the “empty” cells are still considered as belonging to the matrix and therefore included.
Raster structure

Eg a digital photo is nothing more than a matrix of pixels. Each pixel store an information consisting in the position taken by the three bands in the spectrum of the visible.
Another example may be a Digital Terrain Model: each pixel stores an information consisting in the average altitude of the land into the cell.
Data structure comparison

<table>
<thead>
<tr>
<th>Memory usage</th>
<th>VECTOR</th>
<th>RASTER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topology</td>
<td>🎉</td>
<td>🎉</td>
</tr>
<tr>
<td>Graphic precision</td>
<td>🎉</td>
<td>🎉</td>
</tr>
<tr>
<td>Cost of calculation</td>
<td>🎉</td>
<td>🎉</td>
</tr>
<tr>
<td>Image analysis and remote sensing</td>
<td>🎉</td>
<td>🎉</td>
</tr>
<tr>
<td>Update</td>
<td>🎉</td>
<td>🎉</td>
</tr>
<tr>
<td>Visualization</td>
<td>🎉</td>
<td>🎉</td>
</tr>
</tbody>
</table>
Spaceless data (attributes)

“One to one” relation between spatial geographical data and spaceless informations (attributes)

It’s then possible to access to the informations from geographic data or, at the contrary, visualize the object that satisfies certain characteristics based on specific requests (queries) made about the attributes of the data.

It’s virtually possible to store an unlimited amount of information relating to geographic data

Eg. for a Borehole:
- Coordinates
- Depth
- Yield
- Diameter
- Altitude
...
Spaceless data (attributes)

Using relational database to store attributes

- Make the storage easier and lighter
- Increase the efficiency of data queries (querying)
- Allowing multiple users
- Allow remote management
- Increase the security of data access

A query example:
What are the wells belonging to the company X that pump yield more than Y mc / h?
Spatial analysis

Using relational database to store attributes

Proximity
Which countries are 3 km and 10 km from the well?

Layers overlay
I want all the information on a given property: land use, owner, type of soil, etc.

Networking
What is the best way to reach the station from the hotel?
References

- Longley, Goodchild, Maguire, Rhind *Geographic Information Systems and Science* 2nd Ed. Wiley, 2005


- Worboys, Michael *GIS: A Computing Perspective* Taylor & Francis, 2nd Ed 2004 (Computational focus)