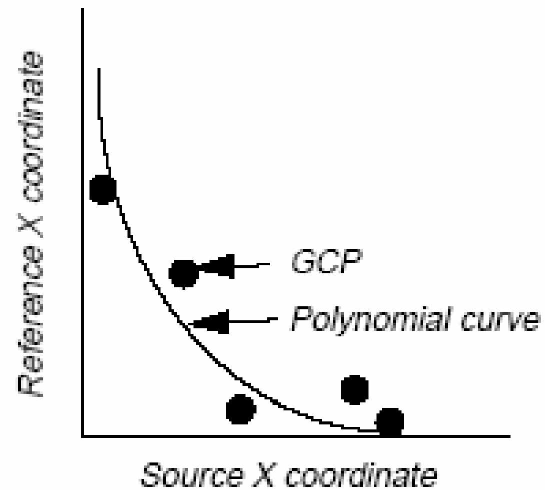
UNCERTAINTY AND ERRORS IN GIS

Christos G. Karydas, Dr.
xkarydas@agro.auth.gr
http://users.auth.gr/xkarydas

Lab of Remote Sensing and GIS
Director: Prof. N. Silleos

School of Agriculture
Aristotle University of Thessaloniki, GR



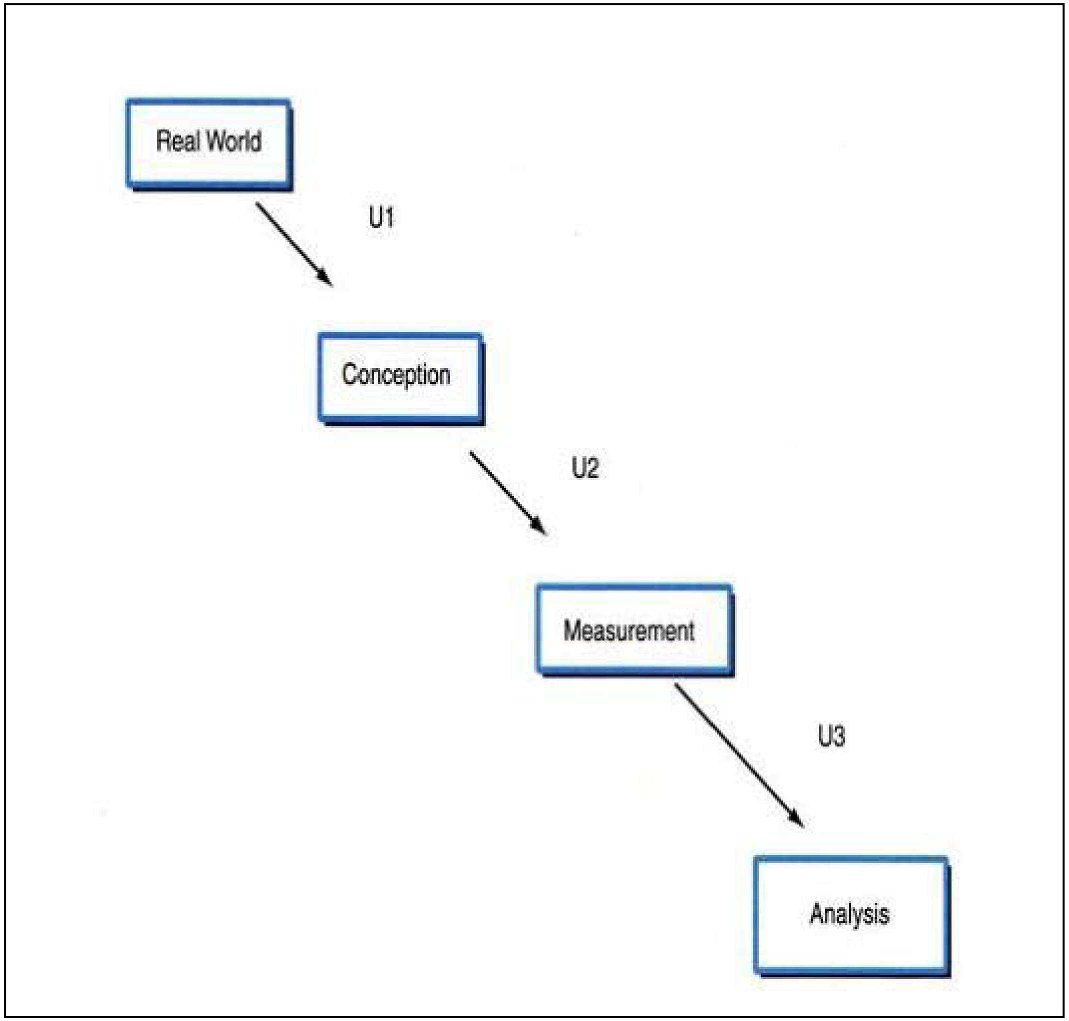
UNCERTAINTY

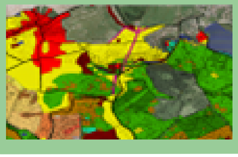
General: The phenomena are real, but we are unable to describe them exactly.

In GIS:

- a situation where digital representations are incomplete
- as a measure of the general quality of the representation

STAGES WHERE 'U' ARISES





U1: CONCEPTION

Is the defining boundary of a zone crisp and well-defined?

Is our assignment of a particular label to a given zone robust?

Examples:

- Coastal management: Where is the exact coast line in an area with tide phenomena?
- Land use: If an ‘Olive grove’ is defined as an area with more than 75% olive trees in a mapping unit, what is the result if the coverage is 74%? The area is no longer an ‘olive grove’? What is it then?

Christos G. Karydas, Lab of Remote Sensing and GIS, Aristotle University of Thessaloniki, GR

U1: CONCEPTION

The questions have:

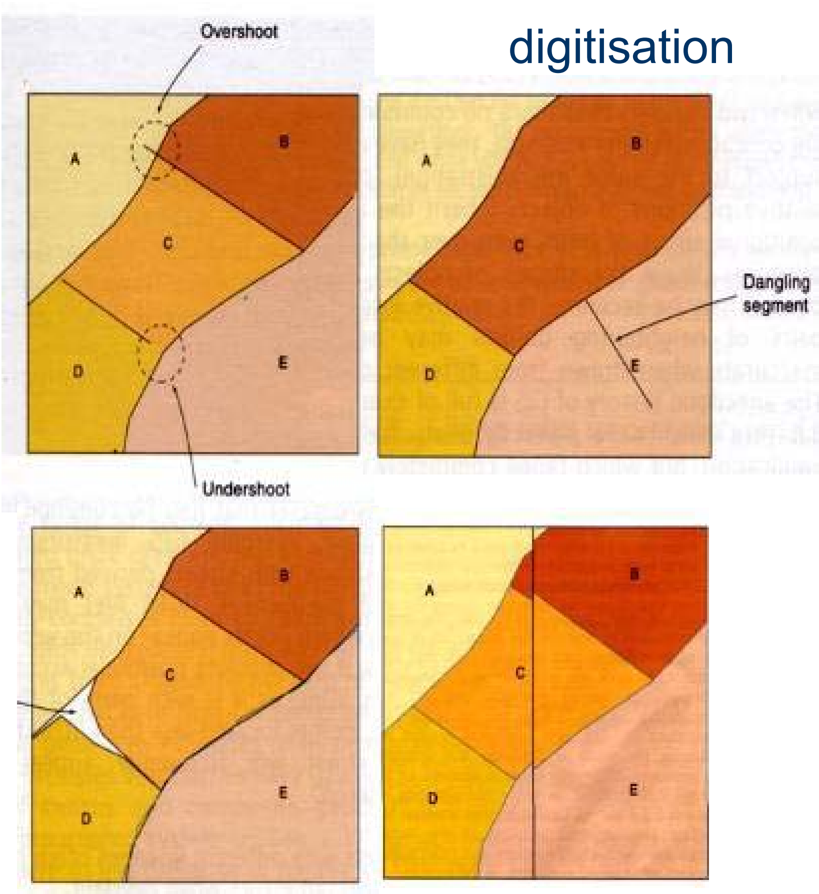
- statistical implications (can we put numbers on the confidence associated with boundaries or labels?),
- cartographic implications (how to convey the meaning of vague boundaries and labels through appropriate symbols on maps and GIS displays?)
- cognitive implications (do people subconsciously attempt to force things into categories and boundaries to satisfy a deep need to simplify the world?) -uniformity

U2: MEASUREMENTS

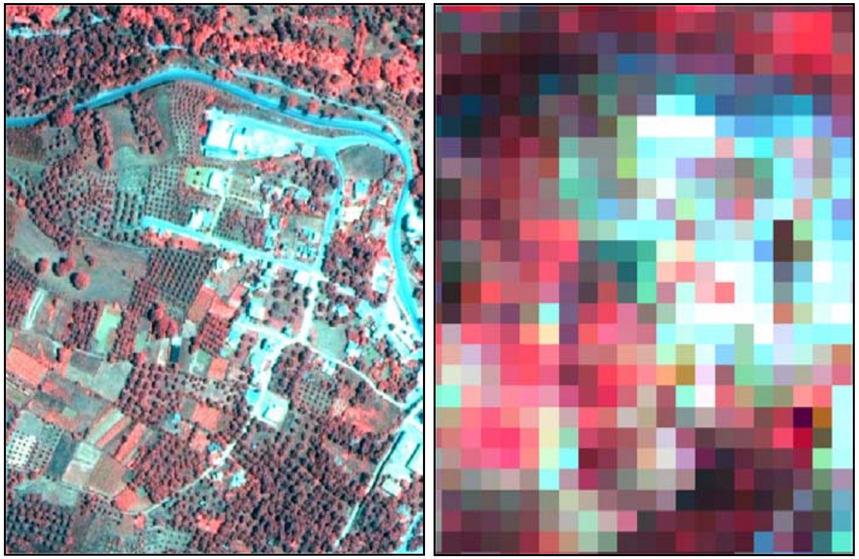
- GPS (systematic and random errors)
- Digitisation (vector or raster: human mistakes, precision)
- Pixel (sampled area, i.e. a pixel is by default the mean of included information)
- Photo-interpretation skills (do we recognise correct?)
- Transformations (continuous, e.g. NDVI, PCA; thematic, e.g. image classification)
- Projection systems (the spheroid of the Earth ‘becomes’ plane: it contains errors by default)
- Sampling (examination of part of the population; confidence level)
- …

U2: MEASUREMENTS

digitisation



pixel size effect (spatial resolution)



U2: MEASUREMENTS

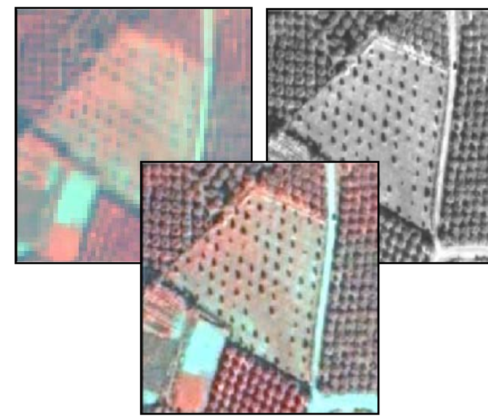
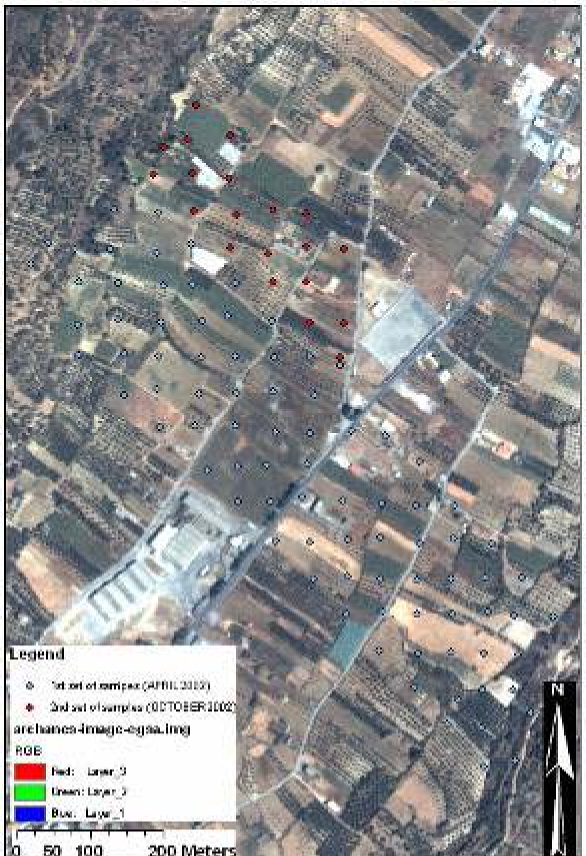


image transformations (left: classification; right: fusion)

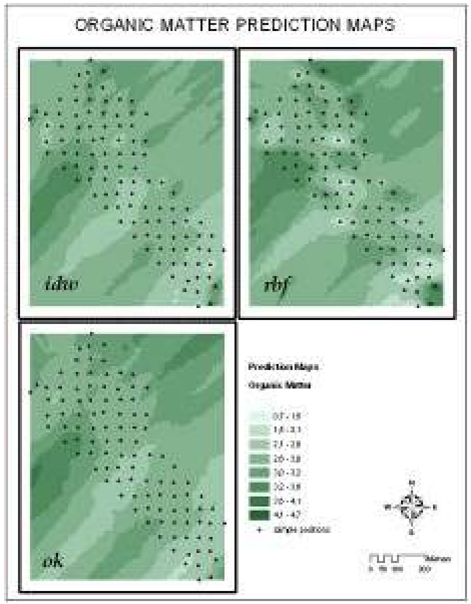
Christos G. Karydas, Lab of Remote Sensing and GIS, Aristotle University of Thessaloniki, GR

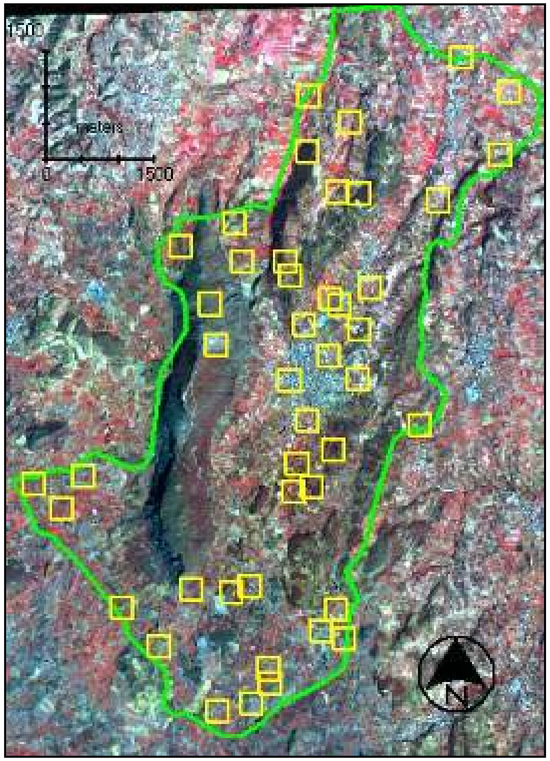
U2: MEASUREMENTS



sampling schemes;

prediction with interpolation





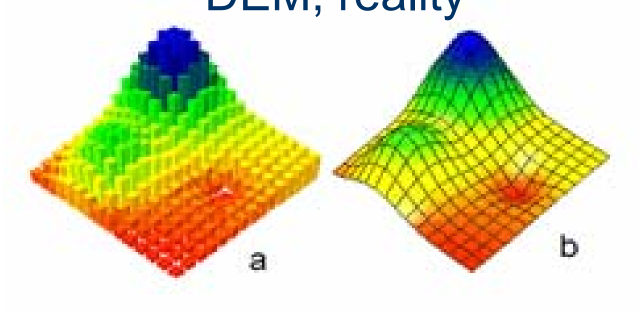
Christos G. Karydas, Lab of Remote Sensing and GIS, Aristotle University of Thessaloniki, GR

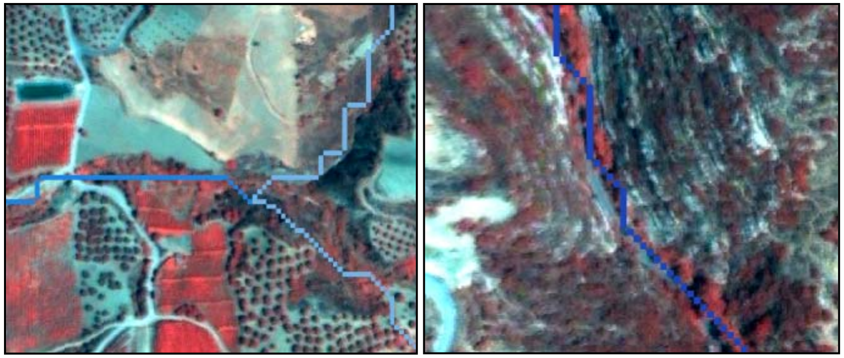
U2: MEASUREMENTS

Data models

- Vector (very crisp)
- Raster (very fuzzy)
- TIN
- Contours

DEM, reality





Spatial variance of uncertainty: drainage network from the same DEM in two sites

Christos G. Karydas, Lab of Remote Sensing and GIS, Aristotle University of Thessaloniki, GR

U3: ANALYSIS

Scale (space)

- Geographic (area of interest)
- Cartographic (linear ratio of map and real distances)
- Size (pixel or line width)
- Operational (feature or ecosystems function)

Time scale: every, when duration,…

Select the appropriate scale after thorough examination

According to CORINE land cover system (based on LANDSAT or SPOT images), the code 121 represents urban fabric (cartogr.scale 1:100000). Some errors are obvious when polygons overlay a finer resolution dataset (IKONOS, 1m)

Christos G. Karydas, Lab of Remote Sensing and GIS, Aristotle University of Thessaloniki, GR

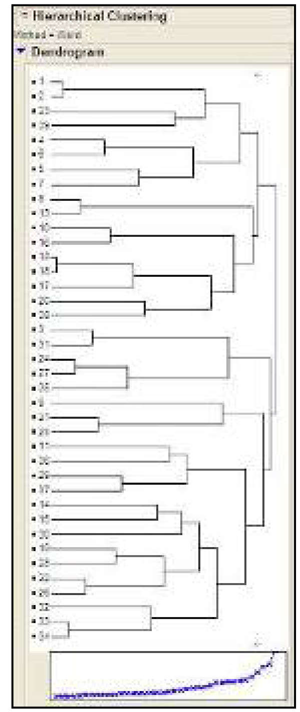
U3: ANALYSIS

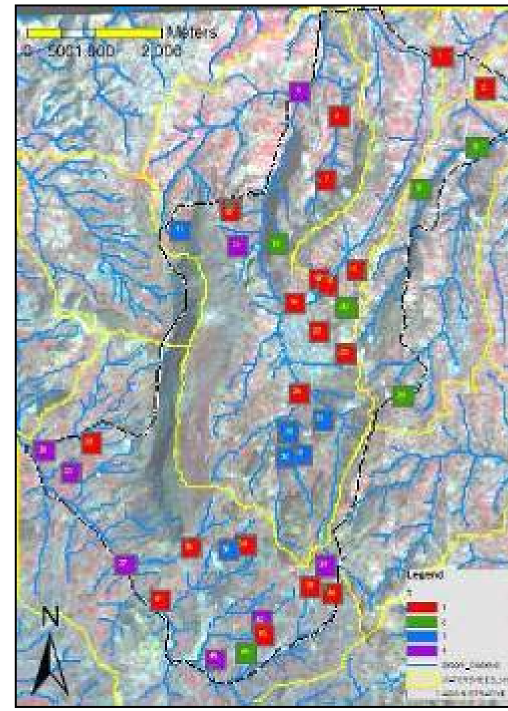
When analysing you dataset, keep in mind the following analysis fields where uncertainty may arise:

- Modified Areal Unit Problem (MAUP): when boundaries of units change
- Aggregation (smaller unit sum up to bigger units (scaling up)
- Ecological fallacy (dominant characteristics of areas assigned to individuals or point locations in those areas; scaling down)
- Atomistic fallacy (dominant characteristics of individuals assigned to entire areas; scaling up)

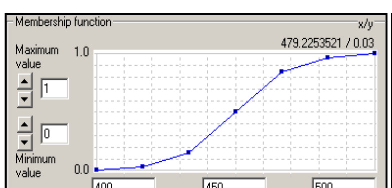
U3: ANALYSIS

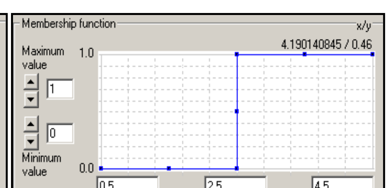
- Hierarchy theory: appropriate approach for landscape modelling, where feature relations (such as similarity) are described with dendrogramms





- Fuzzy logic: functions describing truth or false using intermediate values (not only 0 and 1)

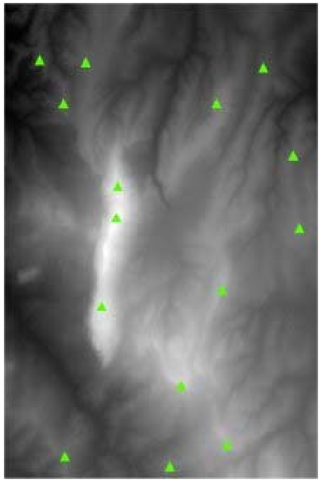


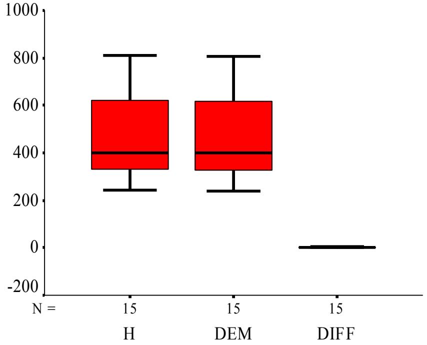


Christos G. Karydas, Lab of Remote Sensing and GIS, Aristotle University of Thessaloniki, GR

ERRORS

- error: discrepancies between recorded measurements and the truth
- precision: how exactly a location is recorded
  - repeatability of a measurement
  - the level of detail used in measurement's recording





Christos G. Karydas, Lab of Remote Sensing and GIS, Aristotle University of Thessaloniki, GR

ACCURACY

- Location accuracy
  - How close the apparent location of a map feature is to its true ground location
  - Example: map scale=reliability and detail of a map feature, rule of thumb)
- Topological accuracy
  - How well spatial relationships are maintained
  - Depends on data entry, error detection and removal

Christos G. Karydas, Lab of Remote Sensing and GIS, Aristotle University of Thessaloniki, GR

ERRORS

- Prediction errors (e.g. interpolation)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left[\left|z(x_i) - \hat{z}(x_i)\right|\right]$$

  - Mean Absolute Error is a measure of the sum of the residuals (e.g., predicted minus observed)
  - Mean Square Error: measure of goodness of control points (individually or totally)

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left[\left|z(x_i) - \hat{z}(x_i)\right|\right]^2$$

  - Goodness-of-prediction (G) estimate

$$G = \left(1 - \frac{\sum_{i=1}^{n}\left[z(x_i) - \hat{z}(x_i)\right]^2}{\sum_{i=1}^{n}\left[z(x_i) - \overline{z})\right]^2}\right) \times 100$$

- Thematic mapping errors:
  - Overall accuracy
  - Producer's accuracy
  - User's accuracy

Christos G. Karydas, Lab of Remote Sensing and GIS, Aristotle University of Thessaloniki, GR

Suggested literature

Bian, L. (1997). Multiscale Nature of Spatial Data in Scaling Up Environmental Models. Scale in Remote Sensing and GIS. D. A. Quatrochi και M. F. Goodchild. Boca Raton, Florida, Lewis Publishers, CRC Press LLC: 13-26

Csillag, F. (1997). Quadtrees: Hiearachical Multiresolution Data Structures for Analysis of Digital Images. Scale in Remote Sensing and GIS. D. A. Quattrochi και M. G. Goodchild. Boca Raton, Florida, Lewis Publishers: 247-271.