1. Introduction in GIS.

Geographical Information Systems (GIS), where developed to facilitate computer-assisted mapping and map analysis. Today it is a multidisciplinary field, merging technologies from cartography, computer-aided design (CAD) and computer graphics, surveying and photogrammetry, spatial analysis and remote sensing. The main components of a GIS are hardware, software and integration. The science and technology of GIS are expanding fast in many fields and are expected to expand further in the future.

References

- Peter A. Burrough and Rachael A. McDonnell, Principles of Geographical Information Systems (Spatial Information Systems), Oxford University Press, USA, 1998. (chapter 1).

Practical exercises.

- Introduction in the GIS software that will be used for the practical exercises. Adding layers, changing symbols, setting scale and projection, map and paper views, performing operations, exporting data.

2. Data input.

One of the most tedious processes in GIS has been to input the data in the system. Today this is changing fast as the availability of the data is increasing. In addition data input processes are becoming more advanced. Raster to vector conversion has emerged. This conversion greatly simplifies data input but it is not suitable for all applications. In addition, despite its automatic or semi-automatic nature, there are several parameters that need to be set. The quality of the output also heavily depends on data pre-processing.

References

- Peter A. Burrough and Rachael A. McDonnell, Principles of Geographical Information Systems (Spatial Information Systems), Oxford University Press, USA, 1998. (chapter 1).

Practical exercises.

- The exercise includes simple digitization using scanned maps as background. It also includes a tutorial for raster to vector conversion.

3. Data structures

The main debate in GIS has been whether to choose a raster or a vector data structure. According to the vector data structure all features are composed of a series of points, lines and areas. Vector data structure is complex and the development of algorithms is very difficult. On the other hand it is very efficient for measuring and topological operations. Raster data structure is simple and so are the operations for raster data. The question which data structure is better is problem dependent. In general applications for surveying, utilities and cadastre are more suited for vector. Raster data structure is better for applications that have less well defined geometry such as forestry and environment. Raster should also be considered when modeling is required. The current generation of GIS software makes the integration of raster and vector seamless. Conversion from one to the other are provided.

References

- Peter A. Burrough and Rachael A. McDonnell, Principles of Geographical Information Systems (Spatial Information Systems), Oxford University Press, USA, 1998. (chapter 2).

Practical exercises.

- Conversion between raster and vector using point, line and polygon data sets.

- Buffer for points, lines and polygons.

- proximity analysis (near). What is the average distance between settlements and rivers in Magnisia. Between settlements and the road network?

- Find the settlements that are within 150 meters from a river.

- Find the settlements that are within 2000 kilometers from the coastline

4. Databases

Geographical Information Systems (GISs) depend on databases. One of the main differences between a GIS system and a Database system is that in the GIS case all information has to have a spatial reference. The Relational Database model is dominant today in the GIS systems. The relational model is based on the principle that flat tables are interconnected via common fields. The main operations done in GIS systems typically include adding new fields as well as modifying or deleting existing ones. Joining two or more tables based on a common field is also a very important operation. The standard way for queering the database is the SQL language (Structured Query Language). The basic SQL syntax is

Select * from TABLE where CONDITION;

Practical exercises.

- Create a locator map showing Magnisia ([nomoi_population], definition query)

- Create a new layer with cities in Magnisia that have population more than 3000 people.

- Add the population information for each prefecture in Greece (join)

- Calculate the population density per municipality per sq. km. (you need to calculate area first.)

- Select the prefectures that have more than 100 people per sq.km.

- Select the prefectures that have more than 100 people per sq.km and at the same time more than 1 million inhabitants in total.

- Create a new layer with all the prefectures that have less than 20 inhabitants per sq.km and those prefectures that have more than 100 inhabitants per sq.km. (i.e. difference of OR and AND operators)

- Select the cities (points) where the population is more than 100 people and altitude is equal to 500 meters.

- Assuming that the population information refers to 2001 create forecast maps for 2011 and 2021 if we accept a population increase of 3% per decade.

- What is the total area of land cover "Olive Groves" in Magnisia?

- What is the total area of land cover "Continuous urban fabric", "Discontinuous urban fabric" and "Industrial or commercial units" in magnesia?

- Plan a GIS database for the plants that exist in garden of UTH. Which fields should be included? (e.g. height, spices). The collection of information will be done a Global Positioning System (GPS). You need to have a form ready to take it to the field.

5. Vector Overlay

The real power of GIS is to be able to combine geographic information in a way that it is not possible with other systems. In the GIS, the different layers of information can be combined to explore relationships between features. A typical operation of data analysis that is done on vector data structure is to find a zone around a feature (known as the *buffer* operation). Another operation is to intersect two or more layers, by super-imposing them, so that a new layer is created having the geometrical features and attributes from the two or more layers.

References.

- Peter A. Burrough and Rachael A. McDonnell, Principles of Geographical Information Systems (Spatial Information Systems), Oxford University Press, USA, 1998. (chapter 5).

Practical exercises.

- Clip

    a. Extract only the built up areas for Magnisia using the national built-up layer.

    b. Extract the parts of the roads that cross inside built up areas

- Buffer

    a. Extract only the built up areas that are within a distance of 1 km from a river.

    b. Extract only the built-up areas that are within a distance of 2000 meters from the main road network.

    c. Are there any built-up areas that satisfy both the above criteria?

    d. Buffer by population

- Intersect

    a. Find the rivers that are inside built-up areas with over 500 inhabitants.

- Union

    a. Make a new layer that consists of both the buffer zone around the main roads in Magnesia and also the built-up area.

- Suitability analysis. Propose the site for a new resort in Magnisia. Specify 5 spatial criteria and implement them.

- Propose a site for a new waste disposal site. Specify 5 spatial criteria and implement them.

6. Map Algebra

Much as we are able to formulate functions with mathematical operations between numbers in conventional algebra, we are able to do mathematical operations between maps. Map algebra is very efficient in raster data structures. This is because the mathematical operations can be easily computed in a pixel per pixel basis. The functions range from very simple operations such as the sum of two raster, were a new raster is created holding in each pixel the sum of the two input pixels. Complex functions can also easily be accommodated. This is very useful for modeling. Overall, the modeling capacity derived by map algebra is a key factor for selecting raster data structures when the objective is to do modeling.

References.

- Peter A. Burrough and Rachael A. McDonnell, Principles of Geographical Information Systems (Spatial Information Systems), Oxford University Press, USA, 1998. (chapter 5).

Practical exercises.

- Map algebra using population rasters

7. DEM

A Digital Elevation Model (DEM) is a digital representation of the continuous variation of relief over space. DEMs were originally developed for modeling relief but they are commonly used to model the continuous variation of any other attribute Z over a two dimensional surface. One example of such variation, i.e. other than elevation, can be the population or the price of land. The DEM is a raster. Each pixel of the raster holds the average elevation within its area. The standard method for representing relief has been to use contour lines, which are lines of constant elevation. A DEM can be used to provide a better and easier to perceive impression of elevation. The technique is to compute the shaded relief. During the past few years a major difference in the availability of DEM data has been made by the release of free data, with near global coverage, by the USA. This data is known as the SRTM program of NASA.

References.

- Peter A. Burrough and Rachael A. McDonnell, Principles of Geographical Information Systems (Spatial Information Systems), Oxford University Press, USA, 1998. (chapter 3).

Practical exercises.

- Starting from SRTM compute

  a. Slope

  b. Aspect

  c. Hillshade

  d. Contours

- Create a DEM from contours

- Create a population DEM from the population of cities.

- Interpolate contours of 100 m to 20 m.

- Arcsceen Walk-through

- Arcglobe walk through

- Discuss the population movie (Christman)

8. Spatial Interpolation

Interpolation is the procedure of estimating the value of properties at unsampled sites within the area covered by existing point observations. *Extrapolation* is the process of estimating the value of a property at sites outside the area covered by existing observations. Spatial interpolation and extrapolation are based in the observation that points that are closer together are more likely to have similar values compared to points that are further apart. Interpolation methods are ways to model spatial variation and make better estimates for values at unvisited sites.

Some of the interpolation methods are based on the notion that the value at an unvisited site should be a function of the weighed distance between that observation and visited sites. One question that needs to be answered is how many neighboring points should be taken into account. Another question has to do with what is the maximum distance after which observations are not interrelated. All moving average methods are by definition smoothing techniques. Maxima and minima in the interpolated surface can only occur at data points. Interpolated values are always between the two extremes.

The formula of Inverse Distance Weighting (IDW) is:

$$\hat{Z}(x_j) = \frac{\sum_{i=1}^{n} Z(x_i)d_{ij}^{-2}}{\sum_{i=1}^{n} d_{ij}^{-2}}$$

Where the x_j are the points at which the surface is to be interpolated, the x_i are the data points, d_{ij} is the distance of data point x_i to the interpolated point x_j and n is the total number of data points taken into account for the calculation.

References.

- Peter A. Burrough and Rachael A. McDonnell, Principles of Geographical Information Systems (Spatial Information Systems), Oxford University Press, USA, 1998. (chapter 8).

Practical exercises.

- Interpolate by hand using the IDW method the value at p0 given the following neighbors:

Data point	Distance to p0	value
P1	10	10
P2	10	20
P3	20	40
P4	20	20

- Interpolate the elevation of magnesia based on the spot heights. Compare with the actual 100 meter interval contours. Compare the results for power of 0.5, 2, and 3. Create a comparison map and discuss how the result is influenced.

- Interpolate the 100 meter interval contours to 20 meter interval contours. Discuss if there is gain of information.

- Interpolate the population of Magnisia that is measured at points (cities). Create contour lines and a statistical surface.

Final Examination

The final examination for the GIS class of PODEPRO 2008 consists of the following exercises. Be sure to save all the files that you create in E:\public\podepro\final_*yoursurname.* The exercises have to be completed within 90 minutes. You are allowed to use your notes, books, calculators and any computer software if you need to.

Exercise 1.

The objective is to find the sites (areas) in Magnisia that satisfy the following criteria:

1. More than 500 meters away from builtup areas [magnisia_builtup]
2. Within a distance D from the cities [magnisia_oikismoi_pop] where
   D = ((population / 10000) + 1)*1000 (in meters). The population field is [real_pop].
3. Land Cover [corine_magnisia] broad leaved forest, coniferous forest, mixed forest and olive trees (i.e. [COR] codes 311, 312, 313 and 223)
4. Not within 500 meters of a river [rivers_magnisia]
5. Within 1 km of the roads of Magnisia[roads_magnisia]
6. Name the output layer [**sites**].

Exercise 2.

1. Find the global population in 2008 given that we know the population in 2005 [pop2005] and we have an estimate of the population for 2010 [pop2010]. The interpolation method is the Inverse Distance Weighting (IDW) where the distance is perceived as the distance in years. Name the output raster [**pop2008**].

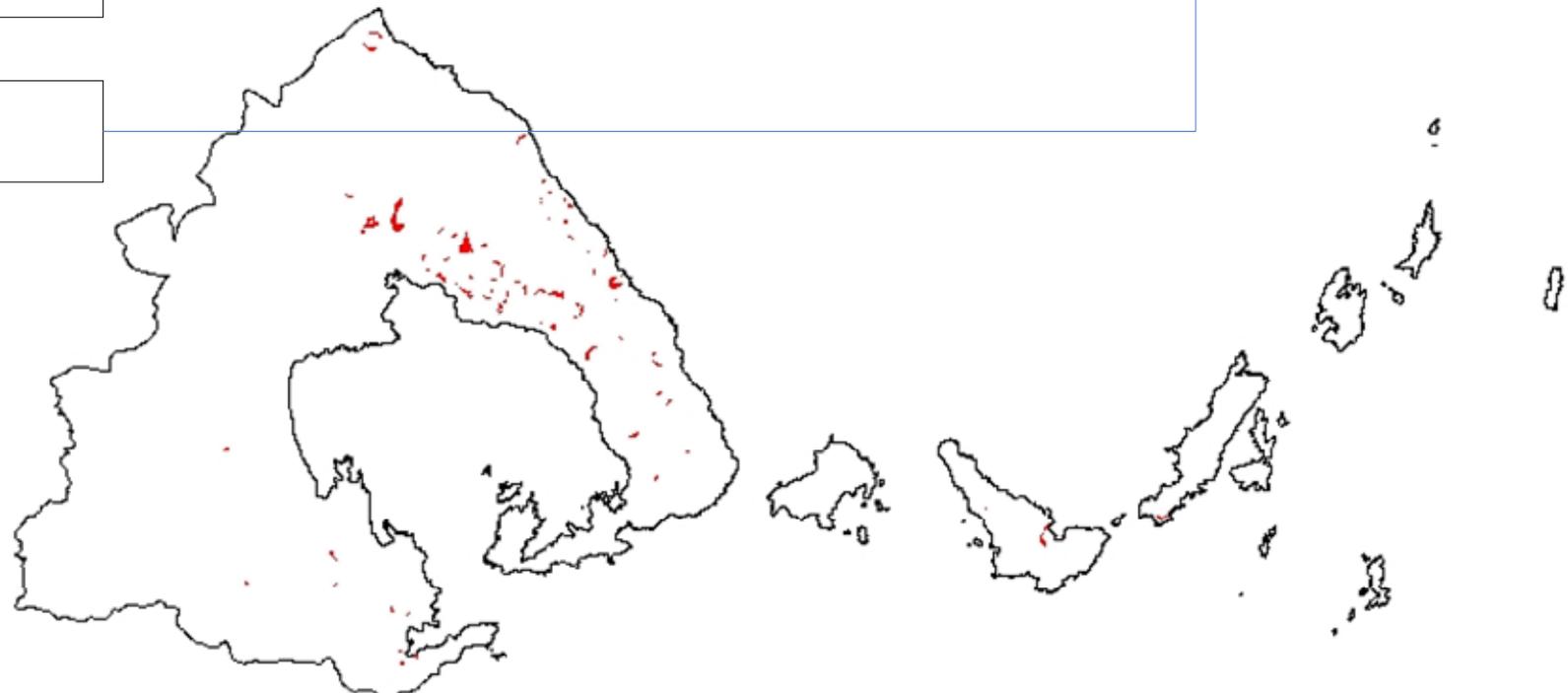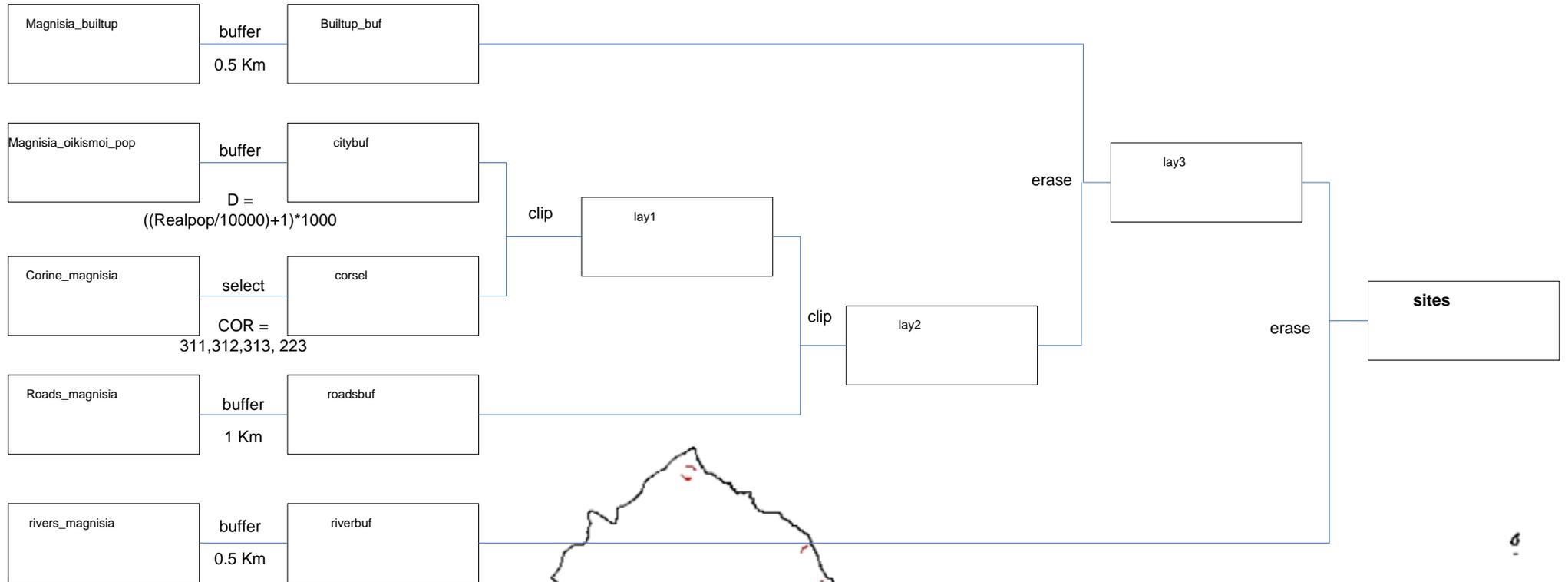   The formula of Inverse Distance Weighting (IDW) is:

   $$\hat{Z}(x_j) = \frac{\sum_{i=1}^{n} Z(x_i)d_{ij}^{-2}}{\sum_{i=1}^{n} d_{ij}^{-2}}$$

   Where the x_j are the points at which the surface is to be interpolated, the x_i are the data points, d_{ij} is the distance of data point x_i to the interpolated point x_j and n is the total number of data points taken into account for the calculation.

2. Highlight the areas where the difference in population between the years 2005 and 2008 is more than 100000 people. The difference between the two years has to be calculated in absolute terms. Create a new raster that has value of 1 if this difference is more than 100000 people and the value of 0 if not. Name the output raster [**dif100k**]

Good luck.

Exercise 1 - solution

Magnisia_builtup	buffer	Builtup_buf
	0.5 Km	

Magnisia_oikismoi_pop	buffer	citybuf
	$D = ((Realpop/10000)+1)*1000$	

Corine_magnisia	select	corsel
	$COR = 311,312,313, 223$	

Roads_magnisia	buffer	roadsbuf
	1 Km	

rivers_magnisia	buffer	riverbuf
	0.5 Km	

clip

lay1

clip

lay2

erase

lay3

erase

sites

pop2008 = ((pop2005 / 9) + (pop2010 / 4)) / 0.36111111

dif100k = con(abs(pop2008 - pop2005) > 100000, 1, 0)