Geographic Information Systems 2026-2027

Exercise 8 - Spatial interpolation

Introduction

> **GOALS OF THE EXERCISE**
>
> - Introduction to the raster data structure
> - Estimate a continuous surface based on a sample of values (points).

Source data

- The files for this exercise are in the course web page (FENIX). Download to your working area the file `Ex08_Interpolation.zip`

Layers:

The geopackage `DataIn.gpkg` contains 3 vector data sets and 3 tables:

Vector data

- `Boundary` - represents the boundary of a phreatic aquifer.
- `Wells` - represents well locations on the aquifer and its concentration of nitrate (unit is mg/l).
- `Soil` - describes the soil type. The code attribute uses the codes of soil type listed in table SoilTypeDescr.

Tables

- `Freguesias` - freguesias in the study area
- `SoilTypeDescr` - soil type description
- `Soiluse` - soil use description

Raster

- `Freguesias` - represents the civil parishes of the study area (raster data set). The code description for every freguesia is in table Freguesias.
- `Use1990`, `Use2002` and `Use2003` - describes the soil use in 1990, 2002 and 2003 (raster data sets). The soil use codes are listed in table SoilUse.

1. Get familiar with your data and your problem

1. Make a preliminary analyses of the characteristics of raster data sets

   – using **Layer properties > Source tab**, analyse the raster file characteristics: number of rows and columns, cell size, pixel type, raster extent and CRS. You will need these values later.

2. Always start by creating a diagram of operations before implementing a solution for the next problems.

2. Problem 1: legends

Create legends for Soil, Freguesias, Use1990, Use2002 and Use2003.

- Use the appropriate symbology to represent the unique categories
- In ArcGIS, after the creation of unique categories, it is possible to create a join to a non-spatial table, using a numeric field. In this case, you can join to the table `Freguesias`, to have their names represented in the legend.

3. Problem 2: conversion (vector to raster), boolean reclassification and overlay operations

Identify the zones of civil parish Estela (code = 4) with soil use Horticulture in 2003 (code = 9) and soil type arenosol (code = 1).

1. Convert the vector layer `Soil` to a raster layer `Soil_raster`
    - in **QGIS**, use the menu *Raster -> Conversion -> Rasterize (Vector to Raster)*. Use the following parameters for the conversion:
        - field to be burn: `SoilTypeCode`
        - Output raster size units: `Georeferenced units`
        - Width and Height resolution: the same of raster `Freguesias`
        - Output extent: the same of raster `Freguesias`
        - Output data type: `Int16`
        - Save file as type: `Tif files`
    - in **ArcGIS**, search for the tool in *Geoprocessing -> Spatial Analysis Tools -> Conversion tools -> Feature to Raster*. Use the following parameters for the conversion:
        - field: `SoilTypeCode`
        - Output cell size: select the raster `Freguesias.tif`
        - In the tab *Environment*, select the raster `Freguesias.tif` as source for *Output Coordinates System*, *Processing Extent* and *Raster Analysis | Cell Size*

    Note: you need to have active the license Spatial Analysis tool. If this is not the case, go to *Project -> Licensing -> Configure your licensing options*

2. Reclassify raster layers

    The raster layers `Freguesia`, `Use2003` and `Soil_raster` have to be reclassified, in order to create new rasters where the code of the areas of interest should be classified as `1`, and everything else as `0`:

    - in **QGIS**, search in the Processing Toolbox for the tool *Raster Analysis -> Reclassify by table*. Set the following parameters:
        - select in *Raster layer* the raster to be reclassified
        - In *Reclassification table*, add rows to the table with the intervals of values to be classified as `1`, and the ones to be classified as `0`
        - set *Range boundaries* to be consistent with the intervals you defined in the table

        - set Output data type as `int16`

    - in **ArcGIS**, search for the tool in *Geoprocessing -> Spatial Analysis Tools -> Reclass -> Reclassify*. Use the following parameters for the conversion:

        - select the `Reclass` field with the values to be reclassified
        - create a reclassification table, either using *Classify* or *Unique*

3. Calculate the areas that fullfil all requirements

    The previous step created raster layers are classified as binary, where pixels of value `1` represent the area of interest. By overlaying layers through a map algebra operation, it is possible to produce a result layer with the areas of interest containing being the value `1`. Which operation should be performed?

    - in **QGIS**, search in the Processing Toolbox for the tool *Raster Calculator*. Set the following parameters:

        - define the expression as the multiplication between all of reclassified layers. Double click the layers to add them to the expression panel
        - set cell size as the same of raster `Freguesias`
        - set output extent as the same of raster `Freguesias`
        - set Output CRS as the same of raster `Freguesias`

    - in **ArcGIS**, search for the tool in *Geoprocessing -> Spatial Analysis Tools -> Raster Calculator*.

        - perform a multiplication between all reclassified raster layers

4. Compare your diagrammatic solution with other possible solutions for the exercise, available at https://fenix.isa.ulisboa.pt/downloadFile/281547991170119/Ex5Diagram(Probl2and3twoSol).pdf

4. Problem 3: conversion (raster to vector)

Create a vector gds representing the regions the civil parish Estela (code = 4) with soil use Horticulture in 2003 (code = 9) and soil type arenosol (code = 1).

You can use the result from the previous step, to convert it to vector layer.

Tools needed:

- in **QGIS**, use the tool in the menu *Raster -> Conversion -> Polygonize (Raster to Vector)*
- In **ArcGIS**, use the tool *Raster to Polygon (Conversion)*

Check for the need to apply a *Dissolve* operation to aggregate polygons with the same value.

5. Problem 4: create a continuous surface from a sample of points

Based on the NO_3 concentration values observed in the wells (layer `Wells`), estimate the NO_3 concentration in the groundwater using the IDW method.

The IDW method generates a raster gds. The estimated value (the pixel value v) is a convex linear combination of the sample values:

$$v = \sum ci´vi$$

with $ci´ = ci / \sum ci$ (so $\sum ci´ = 1$), $ci = (1/di)^p$ and d_i are the distances to the sample points. The power p allows to adjust the distance weight.

1. Explore your data:

- In the layer `Wells`, calculate the basic statistics for the attribute NO_3 conc: min, max, mean, median, std. dev., count
    - In **QGIS**, use the tool Vector --> Analysis tools --> Basic statistics for Fields...
    - In **ArcGIS**, on the attribute table, use the context menu on the field name
- Create a symbology with **proportional symbols** to represent the concentration of NO_3
    - in **QGIS**, select *Graduated* symbol, and *Method: Size*
    - in **ArcGIS**, select *Proportional symbols* in Symbology

2. Create a continuous surface estimated by the IDW method

   Tools needed:

    - in **ArcGIS**, search for IDW as part of Geostatistical Wizard
    - in **QGIS**, search for IDW as part of the GDAL package

   Note: Either in ArcGIS or QGIS, there are other implementations of the IDW algorithm, although with fewer parameters to be set.

   Parameters needed (depending on the implementation):

    - Max neighbors - max number of samples to be used
    - Min neighbors - min number of samples to be used
    - Search - defines how will the method search for the closest samples
    - Angle - if an anisotropy is present, this is the orientation of the major axis

   The values of these parameters need to be decided based on the preliminary statistical analysis. Additionally, the Geostatistical Wizard provided the cross-validation method to assess your model results.

6. Problem 5: create a continuous surface from a sample of points

Do another interpolation, but changing the parameters of the model. Compare with the previous result to identify the influence of the parameters in the surface created.

7. Problem 6: reclassification

Reclassify the raster created on problem 4 in order to create a new raster gds with pixel values 1, 2, …, 5 representing the 5 classes of NO_3 concentration:]0 – 25];]25 – 50];]50 – 100];]100 – 150]; > 150.

8. Additional problem:

Build a raster gds representing the variation on forest use between 1990 and 2002.