OpenSpat Pattern Recognition Exercices

Explore and find structures

Yves Brostaux

June 2017

Exercice 1.1 - Soils of europe

Origin of the data

The Land Use and Cover Area frame Statistical survey (LUCAS) aimed at the collecting harmonised data about the state of land use/cover over the extent of European Union (EU).Among these 2 · 105 land use/cover observations selected for validation, a topsoil survey was conducted at about 10% of these sites. Topsoil sampling locations were selected as to be representative of European landscape using a Latin hypercube stratified random sampling, taking into account CORINE land cover 2000, the Shuttle Radar Topography Mission (SRTM) DEMand its derived slope, aspect and curvature.

The LUCAS topsoil database was used to map soil properties at continental scale over the geographical extent of Europe. Several soil properties were predicted using hybrid approaches like regression kriging. For those datasets, we predicted topsoil texture and related derived physical properties. Regression models were fitted using, along other variables, remotely sensed data coming from the MODIS sensor. The high temporal resolution of MODIS allowed detecting changes in the vegetative response due to soil properties, which can then be used to map soil features distribution.We will also discuss the prediction of intrinsically collinear variables like soil texture which required the use of models capable of dealing with multivariate constrained dependent variables like Multivariate Adaptive Regression Splines (MARS). Cross validation of the fitted models proved that the LUCAS dataset constitutes a good sample for mapping purposes leading to cross-validation R2 between 0.47 and 0.50 for soil texture and normalized errors between 4 and 10%.

WebSite

Data description

Data are available for the following physical properties:

- Clay content (%) in topsoil (0-20cm)
- Silt content (%) in topsoil
- Sand content (%) in topsoil
- Coarse fragements (%) content in topsoil
- Bulk density derived from soil texture datasets (obtained from the packing density and themapped clay content following the equation of Jones et al. 2003)
- USDA soil textural classes derived from clay, silt and sand maps
- Available Water Capacity (AWC) for the topsoil fine earth fraction

Note that these data are based on the LUCAS topsoil data for ca 20,000 samples across EU.

Resolution: 500m

Questions

1. Explore the relashionships between the physical properties of the soil (if the number of points is too high, try a grid sampling of the data to reduce its size)

- How are the different quantitative physical properties related to each other ? Is there some main trends ?
- See how the USDA soil textural classes (qualitative) are related to the physical properties

2. Propose an alternative typology of soils based on their quantitative physical properties

Exercice 2.1 - Aerial image interpretation

Origin of the data

The main data (00710089.tif) comes from an aerial coverage of walloon region with visible and infrared pictures. It covers a 2 x 2 km^2 area near Gembloux, with bands corresponding to Red, Green, Blue and Near-Infrared.

You have two other available informations, the digital surface model (MNS.tif) and the digital terrain model (MNT.tif), with the same extent, obtained by LiDAR technology.

The digital surface model represents the earth's surface and includes all objects on it. In contrast to a DSM, the digital terrain model (DTM) represents the bare ground surface without any objects like plants and buildings.

Questions

1. Explore the relashionships between the different bandwith refletances
   - Calculate a new reflectance information, the normalized difference vegetation index (NDVI), which is know the discriminate the live green vegetation.

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

   - How are the different bandwiths related to each other ? Is there some main trends ? Can you identify some clusters ?
2. Propose a classification of the main objects types present on the picture
   - using only reflectance informations
   - adding information about elevation of the pixels (elevation = surface - terrain)