OpenSpat, spread the spatial wor(I)d

Yves Brostaux et Bénédicte Fontez

June 2017

Perspectives



- Building a learning module on spatial data analysis, based on open and freely available tools
- Focus on pratice and peer-learning

Perspectives

The project



- Three partners
 - Liege University
 - Montpellier SupAgro
 - University of Lisboa

Funded by

- Erasmus+ Key Action 2
- Strategic Partnerships for adult education

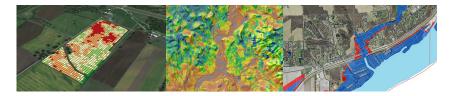
Perspectives

Why ?

Increasing availability of data



Growing set of applications



Perspectives

Need for skills

Need for skills and tools to

- access
- extract
- explore
- analyse such data.

Choice of open-source and freely available dedicated tools

- R
- QGIS

 \Rightarrow increase the potential impact and ease the spreading

Perspectives

Learning modules

Goal : analyze and interpret spatial data

- 1. Online prequel module
- 2. Face to face sessions
 - Access and manipulate spatial data
 - Spatial autocorrelation variogram variance estimation
 - Introduction to linear model
 - Regression over spatially autocorrelated variables
 - Variogramme Kriging
 - Pattern recognition with spatial constraints: clustering and classification

Perspectives

Prequel Online Course



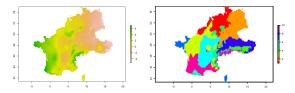
- Introduction for the OpenSpat project
- Overview of the contents of the face to face sessions
- Link with QGIS community (exchanges between QGIS and R)

Face to face sessions

- Standard day schedule
 - AM : background theory and technical introduction to the associated tools with small examples



• PM : real-size case studies in peer learning





- R markdown documents
 - illustrated with examples, *R* commands and *R* outputs Calibration of the Model Variogram

To choose the best model between the list of possible models, a visual inspection is often enough but some statistical criteria like AIC or the weighted Sum of Squares (WSS) are also used.

to go further, WSS mathematical definition:

$$WSS = \sum_{k=1}^K w(h_k) \left[\hat{\gamma}(h_k) - \gamma(h_k)
ight]^2$$

where $2\hat{\gamma}(h_k)$ and $2\gamma(h_k)$ are respectively the experimental and the model variogram values for sites separated by a lag/distance h_k . The weight, $w(h_k)$, is usually proportional to the number of site pairs at lag h_k .

```
n.fit <- fit.variogram(v, vgm("Sph"))
n.fit</pre>
```

```
## model psill range
## 1 Nug 0.00 0.00
## 2 Sph 15292.38 82946.36
```

plot(v,n.fit)

 add interactivity in the provided support documents for future reference and reproductibility

Tools

Implementation of the course with free software scenari



Plickers: create interactivity during the lesson



Perspectives

Test drive and evaluation

On this first session, 14 participants were chosen from the three different partners (young researchers and PhD students)

- different countries (France, Portugal, Belgium)
- different backgrounds (geomatics, agronomy, forestry,...)

They were assigned surveys about

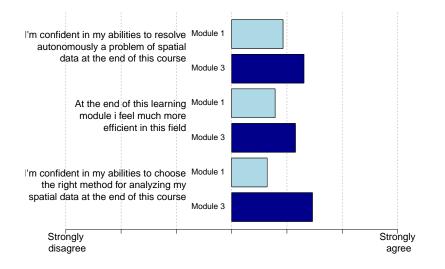
- contents and organisation of the courses
- motivation and self-efficacy

after the prequel online module and the 3 face to face modules.

Program

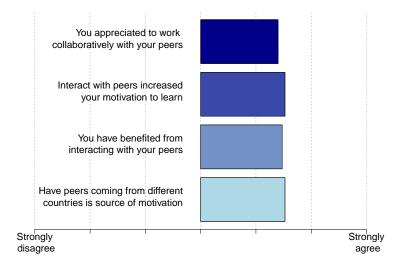
Persp

Student self-efficacy



Perspectives

Student heterogeneity as an asset



Perspectives

Open source tools

Open source and/or freely avalaible tools like

- facilitate the delimitation of technical prerequisitives (as everybody can install and learn the basics before the course)
- ease the exchange between participants
- tons of dedicated packages (sometimes confusing)
- lead to a better take home follow up (as everybody will be able to practice the tools once the course is over)

Perspectives

What's next

- Improvement of the content and organisation based on 1st session surveys
 - Clearer outline of the prequel module
 - Reorganisation of the content sequence
 - Update of the written supports
- Next year session
 - held in one place (Gembloux Agro-Bio Tech, Belgium)
 - two weeks session
 - blended learning (online support and free group work assignements)

More information on the dedicated website