GEOSTAT 2009 Split

Date:

Body:

Backgrounds:

GEOSTAT (The full name of GEOSTAT is “Spatio-temporal data analysis with R+SAGA+Google Earth”) is a 5-day training course and a workshop, initiated by the author of this report, which is run 1-2 times a year at different locations in Europe, and in collaboration with a group of colleagues (Gerard B.M. Heuvelink, Edzer Pebesma, Victor Olaya Ferrero, Roger Bivand, Olaf Conrad, Alexander Brenning). GEOSTAT aims at PhD students and young professionals interested in using open source packages for combined statistical computing and geographical analysis. The core topic of the course is applied spatio-temporal data analysis; specific sub-fields vary from day to day. The lecturers often put focus on methods that help scientists bridge the gaps between statistical and GIS analysis. Or as nicely put by one of our colleagues: “GEOSTAT is where geostatistical computing meets geographic analysis”.

Unlike many post-graduate courses, GEOSTAT is an intensive and user-oriented course. The lectures/exercises run the whole day; the programme of the course is heavily shaped by the course participants? the topics are first selected during the registration, then modified on the spot following the interests of the group. GEOSTAT is also a rather applied course? the participants typically receive practical training with the software and learn how to solve real data analysis problems. About 70% of training focuses on the R language (http://www.r-project.org), which has become a kind of Esperanto for statistical computing among researchers; the rest of syntax used comes from the open source GIS packages: SAGA, ILWIS, GRASS GIS and similar.

Another important thing about GEOSTAT is that it is a non-commercial course, which implies that none of the lecturers is contracted or receives any financial award. The main motives for GEOSTAT are indeed noble: young academic staff get hands-on training with open source software, they see many demonstrations of data processing, and learn how to generate scripts themselves. The senior researchers are there to guide them, to help them with taking the first steps, grasping the important principles and doing the right interpretation of the results.

The 5-day GEOSTAT courses have been held already five times: in 2004 in Zagreb (Croatia), in 2006 in Naples (Italy), at JRC Ispra (Italy), in 2008 in Amsterdam (the Netherlands), and in Belgrade (Serbia). The GEOSTAT 2009 has been held in the period 3-10 May 2009, at the Mediterranean Institute for Life Sciences (MEDILS) in Split. There were 19 participants from 9 countries i.e. 16 institutions: CASA Institute for Ornithology (Croatia); Université Catholique de Louvain, (Belgium); GEOdata d.o.o. (Croatia); Department of Geography and Geology,
University of Salzburg, (Austria); Ghent University (Belgium); University of Bonn, Nees Institute for Biodiversity of Plants, (Germany); GMV sistemas S.A (Spain); UC Berkeley (USA); Faculty of Agriculture, University of Zagreb, (Croatia); Institute for physical and regional planning of Dubrovnik-Neretva County (Croatia); University of Hamburg / scholarship (Germany); Meteorological and hydrological service of Croatia (Croatia); Instituto Español de Oceanografía, Centro Oceanográfico de Baleares (Spain); Iowa State University (USA); INA-Industry of Oil, E&P of Oil and Gas, Development Dept. (Croatia); Centre for Ecology and Hydrology (UK); Institute of Oceanography and Fisheries (Croatia); Fortify Group (USA).

Photo 1: Participants of the GEOSTAT 2009 Summer School, 3-10. May, MEDILS, Split.

The MEDILS institute, an international centre of excellence in the heart of the Mediterranean, has shown itself to be an excellent choice for GEOSTAT. MEDILS offers ideal working conditions and a comfortable set-up: with accommodation and catering in house, several leisure and sport facilities and access to the sea shore. The course was run in one of the MEDILS' seminar rooms (see photo); a computer cluster was also at our disposal. The majority of participants followed the complete course and exercises on their laptop computers.

The 2009 GEOSTAT Summer School was also special because it managed to bring two important developers in the world of geographical data analysis: prof. dr. Roger Bivand (Norwegian School of Economics and Business Administration, Bergen) ? the main creator of the sp, spgrass6 and maptools packages, and the main moderator of the R-sig-geo mailing list; and dr. Olaf Conrad (University of Hamburg) ? the main creator of the SAGA open source GIS. Both, in fact, met for the first time in Split (as most of the participants did) and got to learn about each other's software developments and share many new ideas.

Course programme:

GEOSTAT comprised a balanced combination of theoretical and hands-on-software training (R and open source GIS). The lectures went from 9:00 until 18:00 (with lunch and two coffee breaks), with an additional (demo/discussion) block in the evening hours (20:00-22:00). On the fifth day, a one-day workshop was held where participants could present their case studies and receive important feedback for their project/research work. On the sixth day, the group took a fieldtrip to the island of Bra?. The complete structure and objectives of the working days is given below.

1. DAY 1 ? The course started with a welcome note by B. ?argovi? (representing MEDILS)
and T. Hengl (the course moderator) who introduced the hosting institute and gave an overview of the course. R. Bivand introduced the basic concepts of spatio-temporal data and the ways to represent them in R. The lectures then continued with demos, using the Gambia data set from the geoR package, to illustrate important principles of how and why were the spatial classes designed in R, and how does this effects spatial analysis. O. Conrad further introduced SAGA (System for Automated Geosciencetific Analyses) GIS, which was the main open source GIS that was used for spatial data analysis and visualization during the course. Participants completely new (or feeling insecure about their basic knowledge) to R had a chance to take same "baby steps" in R, in the evening hours.

2. DAY 2 ? The second day focused on geostatistical mapping and use of geostatistical tools in environmental monitoring. Participants got a taste of gstat package and geostatistics in practice. The emphasis was put on regression?kriging ? the key generic spatial prediction technique used in contemporary geostatistics. In the afternoon, participants followed a mapping exercise using field-sampled concentrations of heavy metals and a set of auxiliary maps (the Meuse case study). The final maps produced were exported to KML (Google Earth), to allow interpretation of the distribution of target variables in the area of interest.

3. DAY 3 ? The third day was run by R. Bivand, who first spoke about representation of data in time and space and how the spatio-temporal classes were iteratively developed to fit both our conceptual ideas and applications. We have look at various case studies from areal panels, monitoring interpolation, surveillance, tracks, point patterns etc., and discovered that spatio-temporal data is represented in many different ways, within and across sciences. The participants were then split in five groups and were asked to list spatial classes in use, and to try to extend/improve the analysis. In the evening block, a spatio-temporal interpolation exercises using temperatures in Croatia (123 meteorological stations over 365 days) was demonstrated.

4. DAY 4 ? This day focused on using error propagation techniques for improving DEM analysis. T. Hengl first introduced the theory of error propagation, main principles and benefits of using it in GIS. The exercise focused on using point-measured elevations (the Baranja Hill dataset) to generate multiple realisations of DEM, which were then used to derive drainage network in SAGA GIS and eventually generate a probability map of drainage network. In parallel, O. Conrad gave training in SAGA GIS by focusing on the same exercises.

5. WORKSHOP ? The programme of the workshop was basically created by the participants who submitted topics/questions of interest via the course website. The course moderators then gave some live feedback and suggested solutions or links to literature/www. The presented topics ranged from: Change detection using MODIS EVI images, Interpolation of porosity values by kriging, Automatic Interpolation of LiDAR canopy values using automap, Time series analysis of irregular data, Interpolating distribution and abundance data Iowa agricultural diversity: Moving from a spatial mode to a spatiotemporal model, Modeling Patterns of Plant Diversity (Richness) in West Africa, Statistical analysis of global avian influenza spread and Vegetation patterns in Southern Kyrgyzstan's walnut forests and summer pastures.

6. FIELDTRIP ? On Saturday 9th, the group went for an excursion to the island of Bra?. The group first took a ferry to the city Supetar, then visited the monastery ?Pustinja Blaca?, which was founded in 1588 by the catholic Croatian monks.

Over 15 years ago prof. dr. Miroslav Radman [5], the Croatian maverick of molecular biology (working at the moment in Paris, France), had a dream of establishing an international
research centre of excellence in Split. Few years ago, MEDILS became a reality. The institute now increasingly hosts numerous international summer schools and promotes new directions in science. In that sense, GEOSTAT and MEDILS were a perfect match.

Preview:
> GEO

3-10.05.20 [SAM]
"Spatio-temporal SUMMER"

Registration

The GEOSTAT 2009

The objective of the