

The summer school GEOSTAT 2009

Spatio-temporal data analysis with R + SAGA + Google Earth

Mediterranean Institute for Life Sciences (MEDILS), Split, Croatia, May 3-10, 2009

In the nice location of MEDILS, which is located on the southern hills of Marjan Mt. in Split, the summer school *Spatio-temporal data analysis with R+SAGA+Google Earth* (shorter title GEOSTAT 2009) was held in the period 3-10 May 2009. The course moderators were Roger S. Bivand (Norwegian School of Economics and Business Administration), Tomislav Hengl (University of Amsterdam) and Olaf Conrad (University of Hamburg).

Generally, the GEOSTAT is training course initiated by T. Hengl, which is run 1-2 times a year at different locations in Europe in collaboration with a group of colleagues (Gerard B. M. Heuvelink, Edzer Pebesma, Victor Olaya Ferrero, Roger Bivand, Olaf Conrad and Alexander Brenning). GEOSTAT aims at PhD students and young professionals interested in using open source packages for combined statistical computing and geographical analysis. It is relative intensive (the lectures from 9:00 until 18:00, and additional evening blocks 20:00-22:00)

and user-oriented course, i.e. workshop created by participants. About 70 % of training was done in R language (<http://r-project.org>), which has become some kind of standard for statistical computing in researching community. The rest of training had been performed by open source SAGA GIS package.

The school was organized in one of the seminar room MEDILS buildings, offering all comfort for good lecturing programme. Moreover, the objective of the summer school was to promote use of open source tools for the analysis of spatio-temporal data. The course goals included a balanced combination of theoretical and hands-on-software training (R and open source GIS), one-day workshop where participants could presented own case studies from the field of geostatistics and fieldtrip at the island of Brač. Generally, the lecturing programme was organized in four days (from Monday to Thursday).

Day 1

The course started with a welcome note by B. Žargović (representing MEDILS) and by T. Hengl (representing course moderators). It was followed by R. Bivand introduction in basic concepts and representing spatial data in R. The afternoon session included selected demos that illustrated using the Gambia data set from the geoR package plus supplementary data (R. Bivand). Principles of how and why were spatial classes are designed in R, and how does this reflect on spatial analysis are given. O. Conrad introduced SAGA (System for Automated Geoscientific Analyses) program, i.e. shown spatial analysis made by SAGA as the open source GIS fully controlled from R. T. Hengl gave short review of Google Earth as the most popular browser for geographical data. The evening session ("Baby steps in R") was optional for all participants without experience in R.

Day 2

The second day had been focused on geostatistical mapping and use of geostatistical tools in environmental monitoring. It started with lecture *Geostatistical mapping theory* (T. Hengl) and emphasise on regression kriging as interpolation technique, followed by interpolation of point samples and export of maps to Google Earth (T. Hengl). Also *gstat* package had been introduced to participants as the one of the most famous geostatistical tool. The afternoon session was organized as interpolation exercise of the Meuse case study (located in the Netherlands), and led by T. Hengl and R. Bivand. The final maps were exported to KML (Google Earth file), to allow the interpretation of geographical locations of analysed variable.

Day 3

The program was run by R. Bivand, who first gave lecture *Representation in time and space*, describing the fact that observation in time present different representational issue from those in space

186



Fig. 1. Participants of the GEOSTAT 2009 Summer School, May 3-10, MEDILS, Split (Photo by Jaime Ricardo García Márquez, University of Bonn, Nees Institute for Biodiversity of Plants, Germany)

Slika 1. Sudionici ljetne škole GEOSTAT 2009, 3-10. svibnja, MEDILS, Split (snimio Jaime Ricardo García Márquez, Sveučilište u Bonu, Institut za bioraznolikost biljaka 'Nees', Njemačka)

Ljetna škola GEOSTAT 2009

Prostorno-vremenska analiza podataka upotrebom programa R, SAGA i Google Earth

Mediterranski institut za istraživanje života (MEDILS), Split, 3-10. svibnja 2009.

Na lijepoj lokaciji MEDILS-a, koji se nalazi na južnim obroncima Marjana u Splitu, održana je ljetna škola *Prostorno-vremenska analiza podataka upotrebom programa R, SAGA i Google Earth* (skraćeno GEOSTAT 2009), u razdoblju od 3. do 10. svibnja 2009. Moderator škole bili su Roger S. Bivand (Norveška škola za ekonomiju i poslovnu administraciju), Tomislav Hengl (Sveučilište u Amsterdamu) te Olaf Conrad (Sveučilište u Hamburgu).

Općenito, GEOSTAT je radionica koju je pokrenuo T. Hengl, a održava se 1-2 puta godišnje na različitim lokacijama u Europi. Rezultat je to suradnje skupine kolega (Gerard B.M. Heuvelink, Edzer Pebesma, Victor Olaya Ferrero, Roger Bivand, Olaf Conrad i Alexander Brenning). Programom ta škola kao ciljnu skupinu izdvaja doktorande i mlade znanstvenike koji su zainteresirani u svome radu upotrebljavati slobodne programe namijenjene statističkim i geografskim analizama. Radi se o relativno intenzivnoj školi (predavanja se održavaju od 9 do 18 sati, a večernje vježbe od 20 do 22 sata), čiji je program prilagođen sudionicima koji sami kreiraju radionicu. Oko 70 % vježbi radi se u programskom jeziku R (<http://r-project.org>), koji je postao svojevrsan standard za statističke analize u istraživačkoj zajednici. Ostatak vježbi izvodi se u, također slobodnom, GIS paketu SAGA.

Škola se održavala u seminarskoj sobi unutar zgrade MEDILS-a. Prostor je pružio sve pogodnosti za održavanje predavanja zahvaljujući rasporedu, opremi i klimi. Cilj škole bio je promoviranje slobodnih programskih alata u kojima je moguće načiniti prostorno-vremensku analizu podataka. To je uključilo pažljivo odabranu kombinaciju teorijskih dijelova i praktičnih vježbi, jednostavnu radionicu na kojoj su sudionici mogli prikazati svoje vlastite radove u području geostatistike te terenski izlet na otok Brač. Predavanja i vježbe moderatora škole bila su organizirana tijekom četiri dana (od ponedjeljka do četvrtka).

1. dan

Školu su otvorili pozdravnim govorima B. Žagrović u ime MEDILS-a i T. Hengl u

ime moderatora. Nakon toga je R. Bivand u uvodnom predavanju prikazao temeljne postavke rada i prikaza prostornih podataka u R-u. Poslijepodnevni program uključio je odabrane pokazne skupove podataka koji su obrađeni paketom geoR. Prikazane su osnovne postavke koje su opisale kako i zašto su prostorne klase definirane u R-u, te kako se to odražava na prostornu analizu. O. Conrad je prikazao paket SAGA (Sustav za Automatiziranu Geoznanstvenu Analizu), zatim prostorne analize koje su načinjene s tim slobodnim programom (otvoreni kod) za GIS koji je moguće potpuno kontrolirati iz R-a. T. Hengl je dao kratak pregled programa Google Earth, koji je danas najpopularniji pretraživač za geografske podatke. Večernji termin je obuhvatio vježbanje početnih koraka u R-u, a priređen je za sve sudionike koji nisu imali iskustva rada u tom programu.

2. dan

Program je bio usmjeren na upotrebu geostatističkih alata u praćenju varijabli

izmjerenih u okolišu. Prvo je T. Hengl održao predavanje *Teorija geostatističkog kartiranja* stavivši naglasak na interpolacijsku tehniku regresijskog krigriranja. Nakon toga načinjena je interpolacija nekoliko točkastih podataka, a dobivene karte su bile transformirane u oblik dostupan preko programa Google Earth (T. Hengl). Sudionicima je prikazan i paket *gstaf*, koji je danas jedan od najpoznatijih geostatističkih alata. Poslijepodnevni dio je obuhvatio vježbanje interpolacije na podatcima prikupljenim oko rijeke Meuse u Nizozemskoj, a vodili su ga zajednički T. Hengl i R. Bivand. Završne karte su transformirane u oblik zapisa KML, koji je moguće pročitati programom Google Earth, što je omogućilo uvid u geografski smještaj lokacija analizirane varijable.

3. dan

Program je vodio R. Bivand, a započeo je predavanjem *Predstavljanje (podataka) u vremenu i prostoru*. Naglašena je činjenica da je prikaz u vremenu drugačiji od prikaza u prostoru (tzv. prostorno-vre-



Fig. 2. The working atmosphere at GEOSTAT school (photo by Jaime Ricardo García Márquez, University of Bonn, Nees Institute for Biodiversity of Plants, Germany)

Slika 2. Radna atmosfera unutar učionice škole GEOSTAT (snimio: Jaime Ricardo García Márquez, Sveučilište u Bonu, Institut za bioraznolikost biljaka 'Nees', Njemačka)



Fig. 3. Photo by M. E. Gil Biraud (<http://www.flickr.com/photos/mgilbir/>)

Slika 3. Snimio M. E. Gil Biraud (<http://www.flickr.com/photos/mgilbir/>)

(so called spatio-temporal classes). It gave look at various case studies from areal panels, monitoring interpolation, surveillance, tracks, point patterns etc. In the afternoon session the participants were split in five groups, where they trained to use spatial classes on different dataset and to try to extend/improve the analysis (applying R package). In the evening block, a spatio-temporal interpolation exercises using temperatures in Croatia (123 meteorological stations over 365 days) was demonstrated.

Day 4

This last day of the summer school was focused using error propagation technique for improving DEM analysis. Theoretical review was given by T. Hengl explaining the theory of error propagation, main principles and benefits of using it in GIS. It is presented how to generate multiple DEMs using error propagation, then to derive drainage network in SAGA GIS, and eventually generate a probability map of drainage network. The exercise was organized on the Baranja Hill dataset that O. Conrad processed in SAGA. Together with T. Hengl, he also explained the combining of GIS and statistical analysis (with examples from R+SAGA/ILWIS/GRASS).

Workshop

The programme of the workshop was created by the participants. Everybody had a chance to present some specific research problem and then got some live feedback from others. The workshop was moderated by Roger Bivand and Tomislav Hengl, which also gave some live feedback and suggested solutions. The presented topics from participants had the following titles: *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*, *Modelling 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*.

Fieldtrip

The final part of school program, organized on Saturday 9th, was the excursion on the island of Brač. The tour was organized by bus, mostly crossed the

island in the north-south direction and including the visit of the monastery *Pustinja Blaca*, founded in 1588 from the catholic Croatian monks that came on the Brač retreating from the Turks from Poljica County, eastern from Split. The last priest was Nikola Miličević who organized European well known astronomical observatory inside the monastery and published his discoveries in the journal published in Vienna *Astronomische Nachrichten* (in German) in the first part of the 20th century. He even discovered in 1941 new star, naming this object as *Croatia*, but it never was officially recognized. The excursion was continued by visiting Vidova Gora, the highest peak on the Croatian islands, and driving to small town Bol with famous shape-changing sand beach.

Conclusions

This summer school offered many benefits to their participants, especially interactive team work, and can be recommended for any geoscientist interested in spatio-temporal interpolation and estimation in future. The moderators gave the best and excellent led entire program as well as group of 24 (9 from Croatia) participants. For everybody interested in school's atmosphere, the large number of photos is available at <http://www.flickr.com/photos/mgilbir/> (photos taken by Miguel Eduardo Gil Biraud).

It is necessary to point out that the GEOSTAT program included large number of open source programs that could be, very successfully, applied in spatial data analysis and GIS. Two of them are presented in details through many exercises (R and SAGA). The quality of such programs is often similar or equal to many commercial packages for similar purpose, but with the price of several hundreds or thousands euro. It is why it could be very useful to consider application of these open sources in Croatian universities or public institutes where similar analysis are regularly performed, and in such way to make possible for large number of users independently performing advanced statistical, cartographic or other (GIS) analyses of own datasets.

Tomislav Malvić

menski razredi). Prikazane su različite studije prostornih prikaza, interpolacije, praćenja svojstava podataka u prostoru, njihova rasporeda itd. Poslijepodne su polaznici bili podijeljeni u pet skupina u kojima su vježbali upotrebu prostornih klasa na različitim skupovima podataka. Upotrebom programa R pokušali su proširiti i poboljšati analizu. U večernjem terminu održana je vježba s prostorno-vremenskom interpolacijom podataka o mjerenju temperatura u Hrvatskoj sa 123 meteorološke postaje tijekom 365 dana).

4. dan

Posljednji dan bio je usmjeren na upotrebu tehnike *rasprostiranja pogriješke* u cilju poboljšavanja analiza digitalnih modela reljefa (DMR). Teorijsku podlogu, glavne postavke te koristi uporabe u GIS-u prikazao je T. Hengl. Polaznici su na testnom skupu upoznati s načinom stvaranja višestrukih DMR-ova, opisivanjem drenažnog sustava u SAGI te stvaranjem vjerojatnosnih karata za takvu mrežu. Podatke o visinama na području Baranjskoga brda u SAGI obradio je O. Conrad. Na kraju, T. Hengl i O. Conrad objasnili su kombiniranje statističkih analiza i GIS-a uz primjere upotrebe programa R, SAGA, ILWIS i GRASS.

Radionica

Program radionice oblikovali su sami polaznici. Svatko je imao priliku prikazati svoj vlastiti istraživački problem i nakon toga dobiti povratnu informaciju od nazočnih kolega. Radionicu su vodili Roger Bivand i Tomislav Hengl, koji su također dali komentare i određene savjete za poboljšavanje prikazanih analiza. Sudionici su prikazali brojne teme svojih istraživanja, kao što su primjerice (naslovi su izvorno dani na engleskom jeziku): *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*

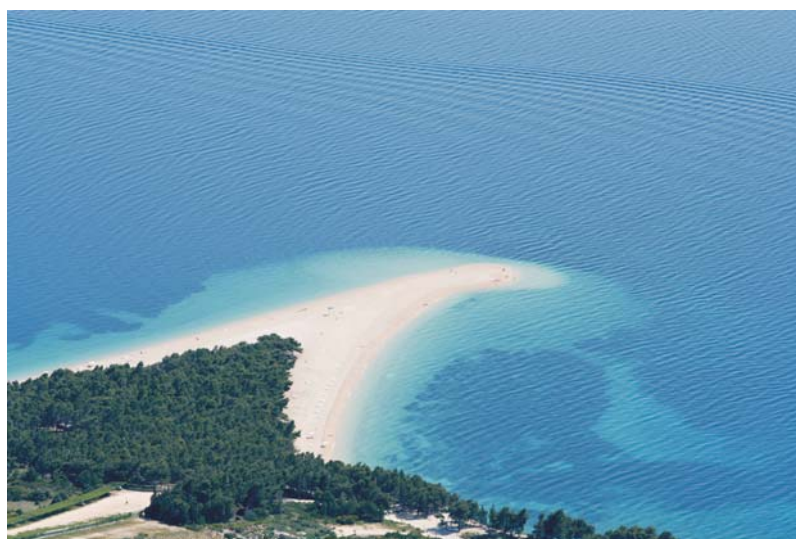


Fig. 4. Photo by M. E. Gil Biraud (<http://www.flickr.com/photos/mgilbir/>)

Slika 4. Snimio M. E. Gil Biraud (<http://www.flickr.com/photos/mgilbir/>)

West Africa, Statistical analysis of global avian influenza spread and Vegetation patterns in Southern Kyrgyzstan's walnut forests and summer pastures.

Terenski izlet

Posljednjega dana škole, u subotu 9. svibnja, organizirana je ekskurzija na otok Brač. Put je uključio vožnju autobusom po otoku, kojom se prošlo cestom sa sjeverne na južnu stranu. Posjećen je poznati samostan *Pustinja Blaca*, što su ga 1588. utemeljili katolički svećenici koji su na Brač izbjegli pred turskom najezdom u područje Poljica (Poljička republika), smješteno istočno od Splita. Posljednji svećenik bio je Nikola Miličević koji je u samostanu vodio europski poznat astronomski opservatorij, a svoja je otkrića objavljivao u bečkom glasilu *Astronomische Nachrichten* na njemačkom jeziku u prvoj polovici 20. st. Godine 1941. čak je otkrio novu zvijezdu kojoj je dao ime *Croatia*, no to otkriće nikada nije službeno zabilježeno. Izlet je nastavljen putom na Vidovu Goru, najviši vrh na hrvatskim otocima, te nakon toga, spuštanjem u gradić Bol na južnoj obali Brača, poznat po svojoj pješčanoj plaži koja mijenja svoj oblik, odnosno položaj.

Zaključak

Ljetna škola je polaznicima pružila brojne pogodnosti, među kojima je možda najveća mogućnost rada u višestrukim skupinama. Njezino pohađanje svakako se može preporučiti svim geoznanstvenicima zainteresiranima za prostorno-vremensku interpolaciju i procjenu. Moderatoru su uložili veliki trud i izvrsno vodili program za skupinu od 24 polaznika (od toga je 9 bilo iz Hrvatske). Zainteresirani za fotografski prikaz rada i atmosfere u školi može pronaći veliki broj snimaka na adresi <http://www.flickr.com/photos/mgilbir/> (fotografije je načinio M. E. Gil Biraud). Posebno treba istaknuti veliki broj slo-bodnih programa (otvorenog koda) koji se mogu vrlo uspješno primijeniti u prostornoj analizi podataka i GIS-u. Dva od njih su i detaljno prikazana na brojnim vježbama (R i SAGA). Svojom kvalitetom takvi programi često ili uopće ne zaostaju za brojnim komercijalnim paketima slične namjene, čija je cijena nekoliko stotina ili tisuća eura. Zato bi vjerojatno bilo dobro razmisliti o upotrebi takvih paketa u visokoškolskim ili javnim ustanovama u Hrvatskoj koje se bave sličnim analizama, te na taj način omogućiti velikom broju korisnika da samostalno izvode napredne statističke, kartografske ili druge (GIS) analize svojih podataka.

Tomislav Malvić