UseR!, ASDAR, r-sig-geo, gstat

Edzer Pebesma

edzer.pebesma@uni-muenster.de

ifgi
Institute for Geoinformatics
University of Münster

Tom's UvA summer school, 18 Aug 2008

The R user's conference, August 12-14, Technische Universität Dortmund, Germany

- ▶ 400+ participants, 170 papers, 13 pre-conference tutorials
- ▶ More scientific than I had thought, kaleidoscopic character
- ▶ Attitude: open and constructive, rather than closed and competitive
- ▶ 1 Spatial Statistics session, 6 sessions with papers addressing spatial issues (environmetrics, econometrics, cellular automata)
- ▶ Jul 10-12, 2009 in Rennes, Fr
- ▶ DSC every two years, more developer-oriented
- ▶ big memory in R, HPC, parallel computing, SOA

Applied Spatial Data Analysis with R

Open source: how can you participate?

Or: *how can I do something in return?*

1. Ask questions on a mailing list
   - ► Work before you post questions; search archives
   - ► Formulate well, be brief, concstructive, don't demand
   - ► Provide reproducable example
2. Answer questions on lists
3. Provide patches, code
4. Become a developer
5. Become a teacher

`r-help`: read and follow instructions *carefully* before you post.

`r-sig-geo` mailing list

... A mailing list for discussing the development and use of R functions and packages for handling and analysis of spatial, and particularly geographical, data. The list also covers mapping and cartographic issues, and interfaces between R and geographical information systems.

History:

- ▶ Thu Jul 10, 2003: Started with an email of Hisaji Ono
- ▶ Fri Oct 31, 2003: RFC by me, on design of what led to sp
- ▶ Mon Oct 18, 2008: over 1000 subscribers
- ▶ the list is not restricted to discussions on packages, spatial statistical are addressed as well
- ▶ goal: to archive Questions and Answers
- ▶ r-spatial-devel, see http://r-spatial.sf.net
- ▶ try not to waste someone else's time

`r-sig-geo` mailing list

... A mailing list for discussing the development and use of R functions and packages for handling and analysis of spatial, and particularly geographical, data. The list also covers mapping and cartographic issues, and interfaces between R and geographical information systems.

History:

- ▶ Thu Jul 10, 2003: Started with an email of Hisaji Ono
- ▶ Fri Oct 31, 2003: RFC by me, on design of what led to sp
- ▶ Mon Oct 18, 2008: over 1000 subscribers
- ▶ the list is not restricted to discussions on packages, spatial statistical are addressed as well
- ▶ goal: to archive Questions and Answers
- ▶ r-spatial-devel, see http://r-spatial.sf.net
- ▶ try not to waste someone else's time

r-sig-geo mailing list

... A mailing list for discussing the development and use of R functions and packages for handling and analysis of spatial, and particularly geographical, data. The list also covers mapping and cartographic issues, and interfaces between R and geographical information systems.

History:

- ▶ Thu Jul 10, 2003: Started with an email of Hisaji Ono
- ▶ Fri Oct 31, 2003: RFC by me, on design of what led to sp
- ▶ Mon Oct 18, 2008: over 1000 subscribers
- ▶ the list is not restricted to discussions on packages, spatial statistical are addressed as well
- ▶ goal: to archive Questions and Answers
- ▶ `r-spatial-devel`, see `http://r-spatial.sf.net`
- ▶ try not to waste someone else's time

Package gstat

- ▶ Started as a project for self-education in 1991.
- ▶ gstat stand-alone program, 1993, GPL 1997; Pebesma and Wesseling, C&G 1998
- ▶ R package submitted to CRAN in 2002, DSC 2003; Pebesma C&G 2004
- ▶ at DSC 2003: Roger Bivand organized a spatial workshop, and raised the idea for a common set of classes and methods
- ▶ about 10 people have contributed code (some critical) and patches
- ▶ hundreds of people have contributed with questions, ideas, bug reports etc.
- ▶ strong points: spatial multivariable linear model, block kriging, scalability.

Gilberto Câmara

Gilberto Câmara and **Harlan Onsrud**, 2004: *Open source GIS software: myths and realities* evaluate 70 OS GIS projects.
Here, gstat is classified as

- ▶ an *Individual-size* project (53%), as opposed to *collaborative network* or *corporation-based*.
- ▶ *Innovations-led* (19%), as opposed to *standards-led* or *post-mature*.
- ▶ with, on a 1-5 scale:
  Innovation: 4, Support: 3, Functionality: 4.

status/future gstat?

OS project status:

- ▶ transition from single-developer to multi-developer project?
- ▶ who should decide on innovations–who wants to be in the gstat core developer group?
- ▶ is this possible, or rewrite from scratch? In which language?

Functionality/code/performance:

- ▶ support multi-threading/multi-core? Also for quad/oct-tree search index using shared memory?
- ▶ replace meschach with atlas/blas?
- ▶ (further) integrate with geoR/geoRglm?
- ▶ extend variogram models? Explicit spatio-temporal capabilities?
- ▶ is it useful on the route to automatic geostatistics?