Integration of Statistics and Geographic Information Systems: the R/TerraLib case

Pedro Ribeiro de Andrade Neto
Paulo Justiniano Ribeiro Junior
Karla Donato Fook

GeoInfo 2005
Introduction

- Constantly evolving software technology
- Spatial statistics in statistical softwares
  - Point pattern analysis
  - Geostatistics
  - Areal data analysis

- Statistical softwares lack:
  - Neighborhood
  - Multiple geometry
  - etc.
Introduction

Geographic Information Systems (GIS)

• Manipulate several georeferenced data
• Geoprocessing algorithms
• GIS lack complex statistical analysis
Integration

- Technique that combines software components in order to generate more complex systems
- Saves time and resources
- Can increase individual effectiveness of both GIS and statistical softwares
Integration: loose coupling

- Two different softwares
- Data exported using files

- SPRING + SpaceStat (Anselin 92)
- ArcInfo + R (Gomes-Rubio 2005)
Integration: close coupling

- Calls direct from the GIS, and vice versa (Bivand and Neteler 2000)
- ArcInfo + S (Bao et al 2000)
- GRASS + R (Bivand and Neteler 2000)
(Fischer et al. 96) propose to use geographic databases for information exchange:

- share a common database
- preserves object identity and metadata

None of the works fits in this description.
Goal

• Provide access to a geoprocessing library in a close coupling way using a spatial database

• From an statistical program: R

• To a geoprocessing library: TerraLib

• aRT: R-TerraLib API
R Project for Statistical Computing

- Free software
- Point pattern analysis: spatstat, splancs
- Geostatistics: gstat, geoR, geoRglm
- Areal data analysis: DCluster, spdep

- 32 packages currently available in the Spatial Task View on the CRAN repository
TerraLib

- Free software
- Database interfaces
- Spatial-temporal structures
- Geoprocessing algorithms
aRT: R-TerraLib API
aRT conceptual model

- TerraLib concepts within R objects
Other functionalities

- Database management functions
- Spatial relations
- Set operations
- Metrics
- Queries
Conclusions

- TerraLib inside R
- Close coupling using databases
- aRT encapsulates TerraLib components
Current status of the project

- Version 0.4-11(2005-11-17)
- Already adheres to the sp classes
- Documentation – vignettes
- Source code and windows binary version
- Available at http://www.est.ufpr.br/aRT
Future work

- Additional interfaces to Terralib algorithms
- Option on compiling for geoprocessing functions without database dependency (via configure tool)
- Expanded functionality for temporal tables and queries
- Further drivers
- etc...
Integration of Statistics and Geographic Information Systems: the R/TerraLib case

Pedro Ribeiro de Andrade Neto
Paulo Justiniano Ribeiro Junior
Karla Donato Fook