From ArcView/XGobi to R/GGobi: Recent Developments in Exploratory Spatial Data Analysis

Jürgen Symanzik

Utah State University, Logan, UT

***e-mail: symanzik@sunfs.math.usu.edu**

WWW: http://www.math.usu.edu/~symanzik

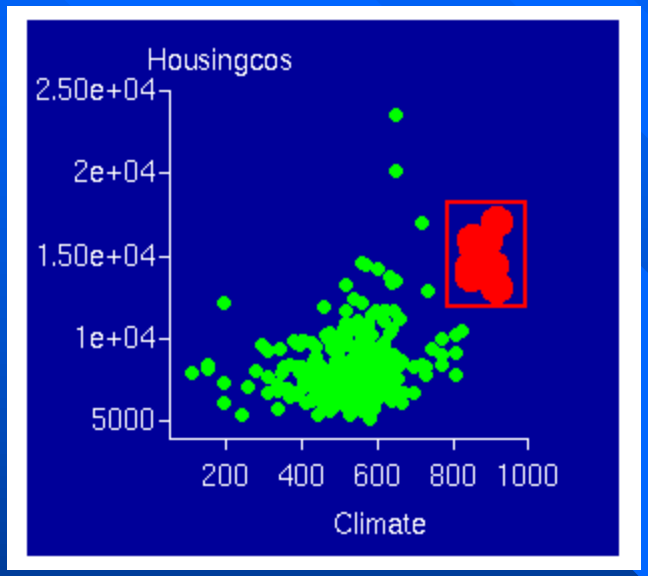
Collaborators

- Deborah F. Swayne: AT&T Labs - Research
- Duncan Temple Lang: Bell Labs, Lucent
- Dianne Cook: Iowa State University
- and many others

ArcView/XGobi/XploRe: Main Idea

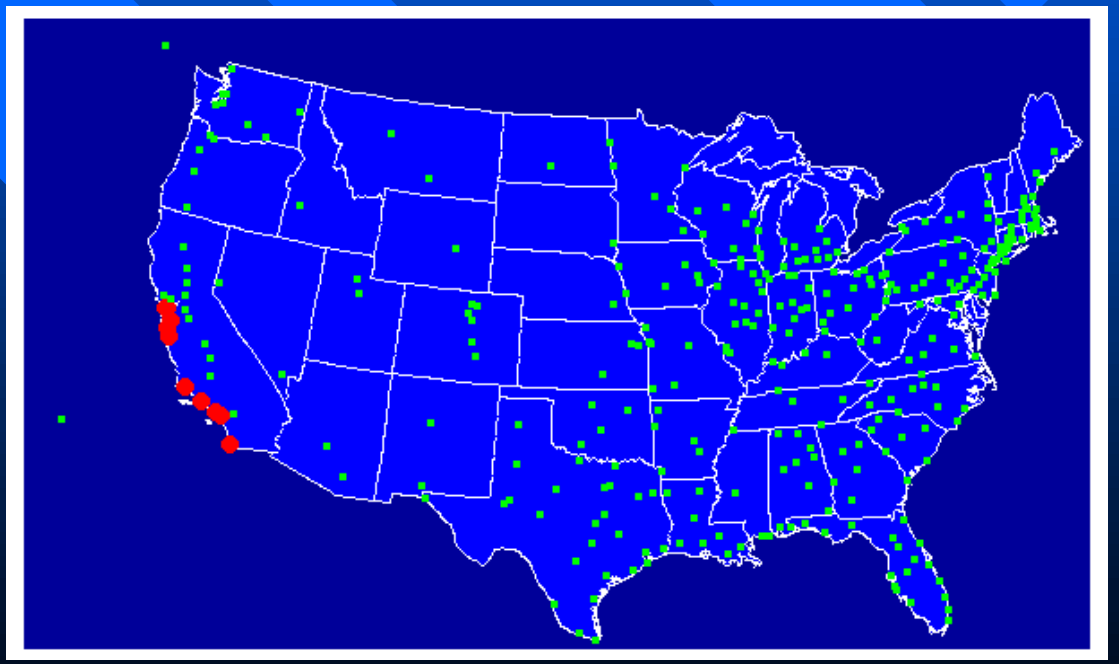
- Link three kinds of software packages:
  - ArcView: Geographic Information System (GIS)
  - XGobi: dynamic statistical graphics program
  - XploRe: statistical computing environment
- Interactive (dynamic) environment
- Close coupling through RPC's

Multivariate Link



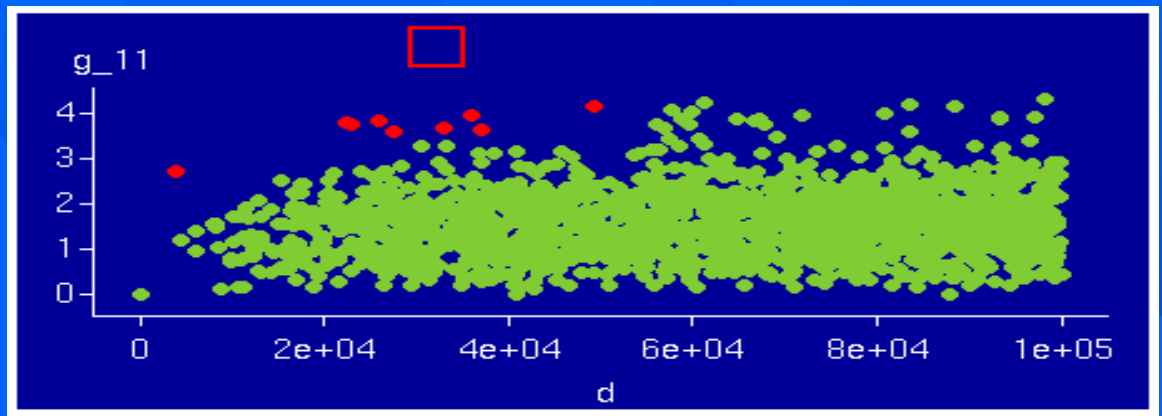
XGobi

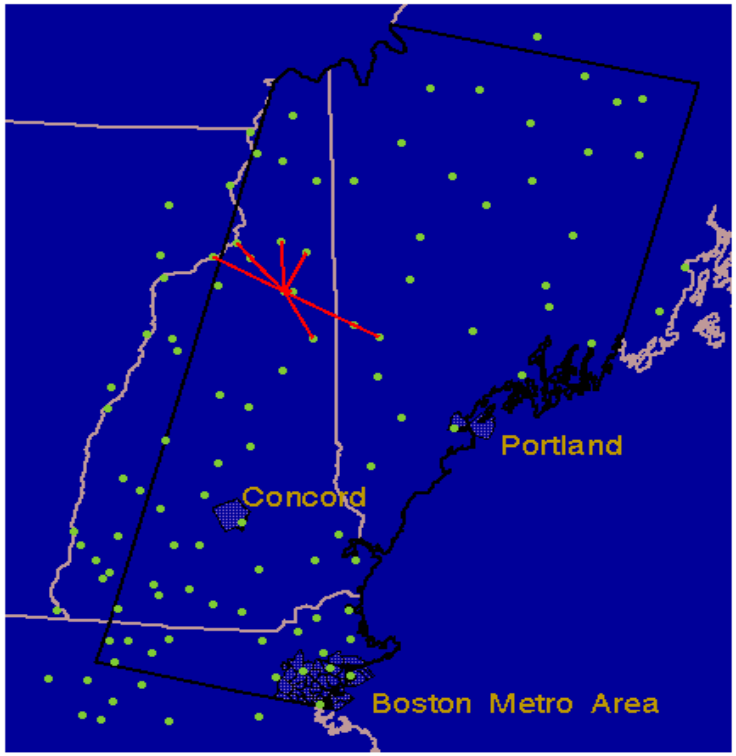
ArcView



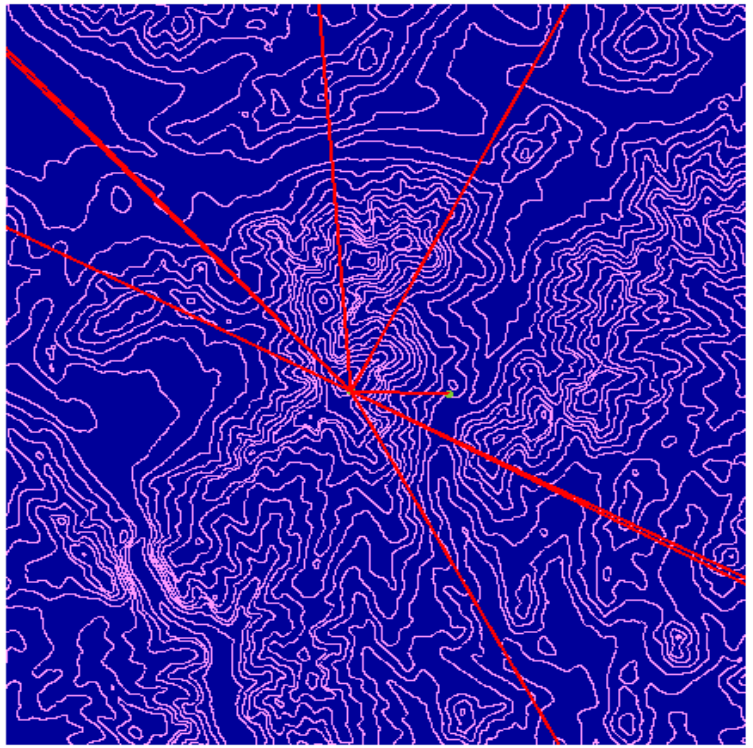
Variogram-Cloud Link

XGobi

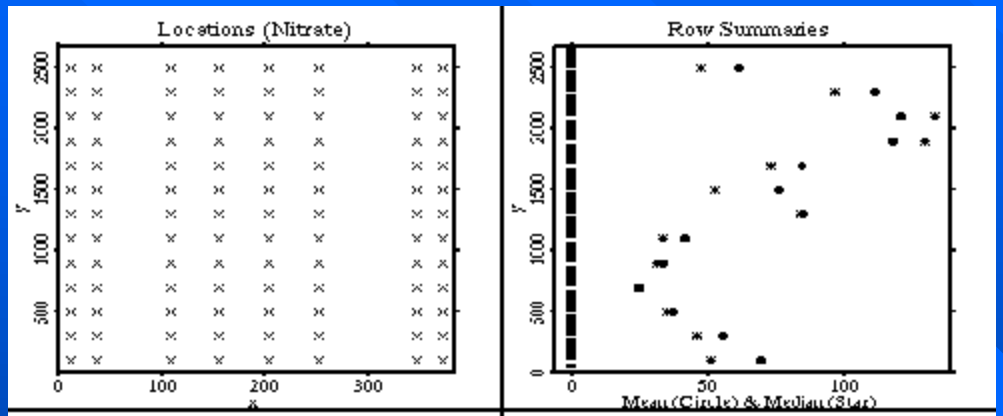




ArcView



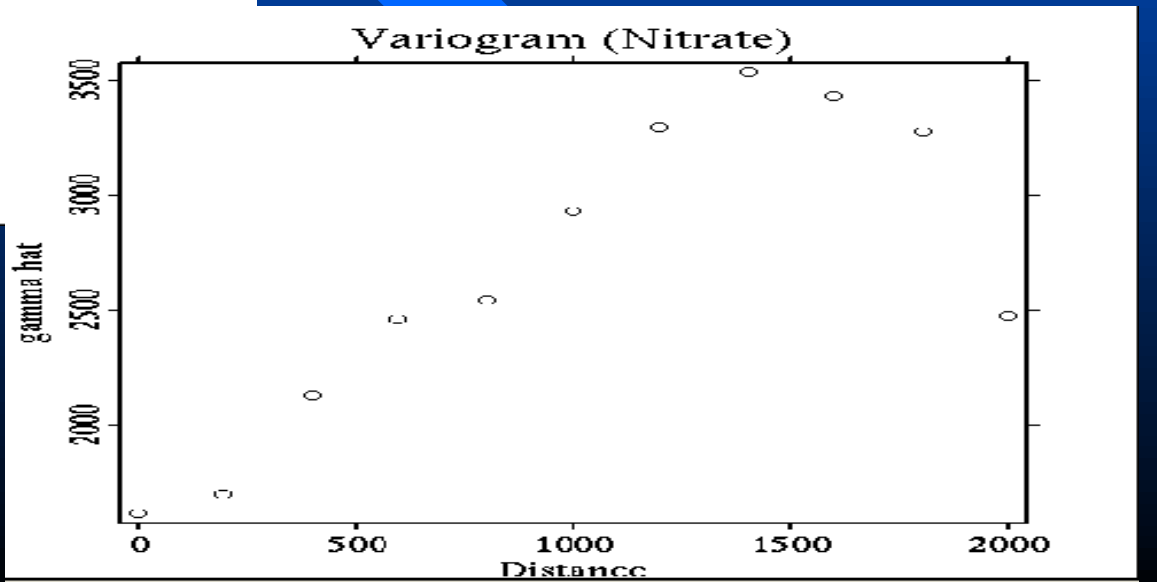
Spatial Data Analysis in XploRe



XploRe

Mean and Median Summaries

Variogram



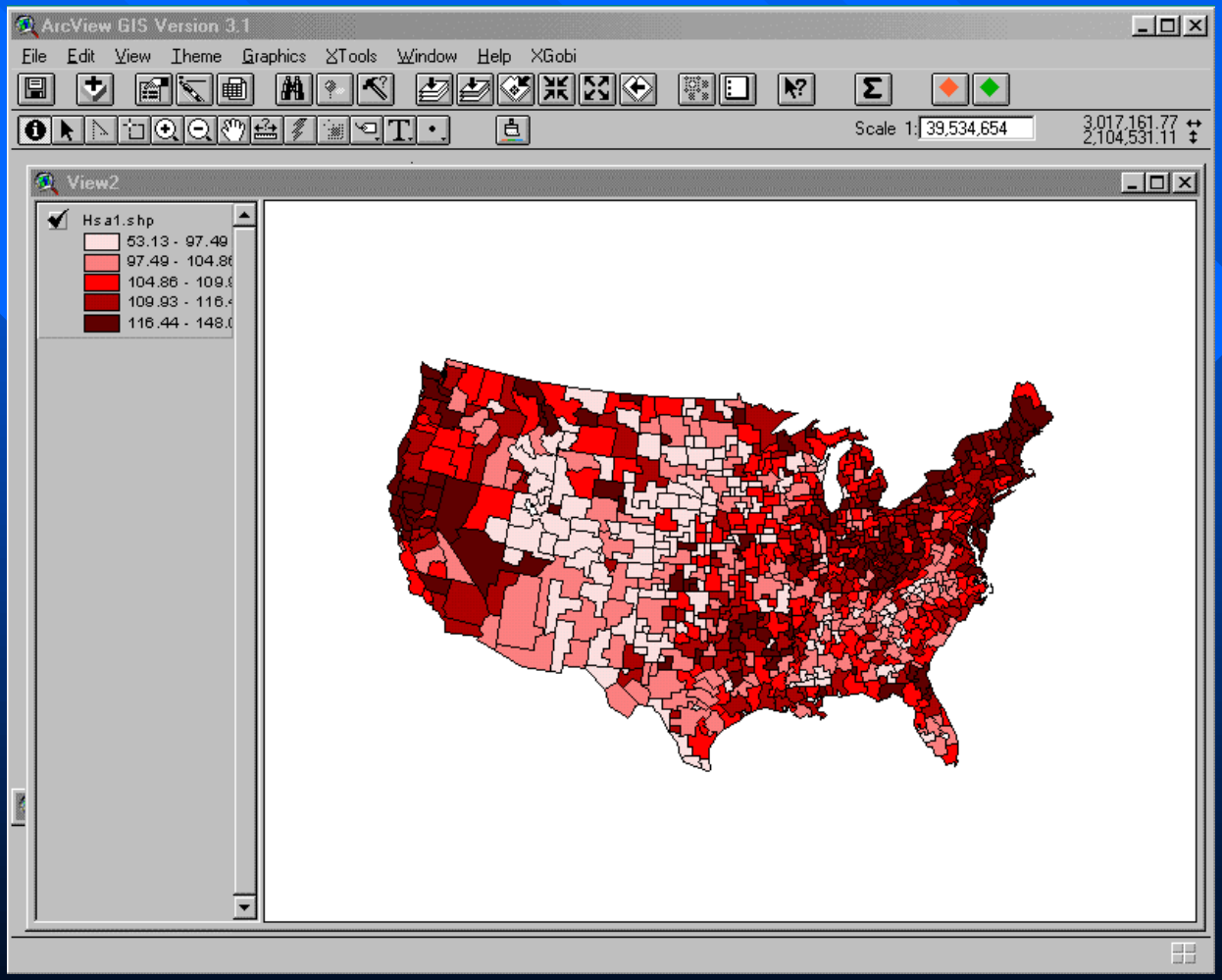
Example: NCHS Mortality Data

- Data for 798 health service areas (HSA's) in the contiguous US

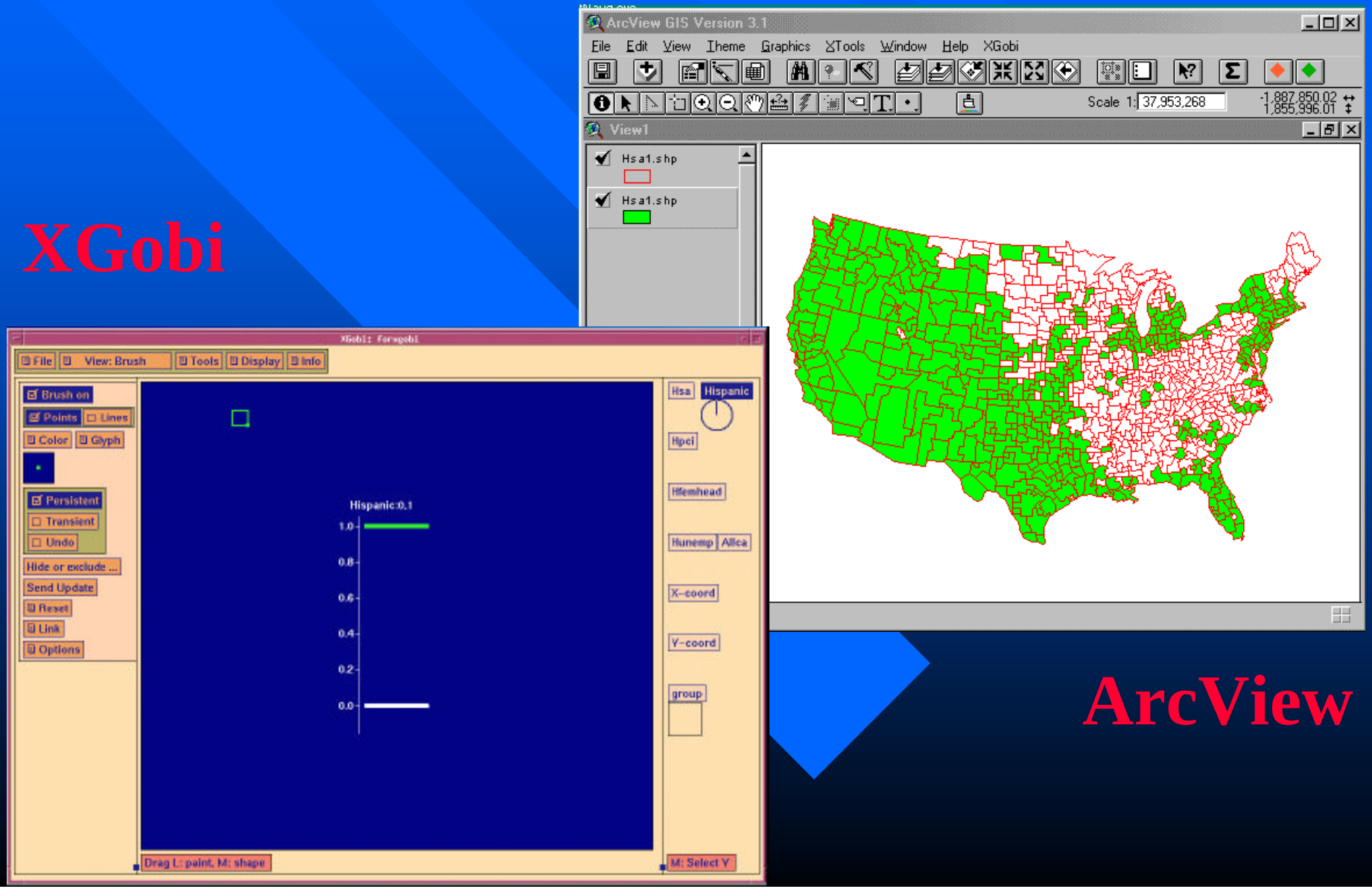
- Variables:
  - Proportion of the hispanic population
  - Per capita income
  - Proportion of female heads of households
  - Proportion of unemployed
  - Cancer Mortality Rate (in deaths per 100,000)

Initial Choropleth Map of Cancer Mortality

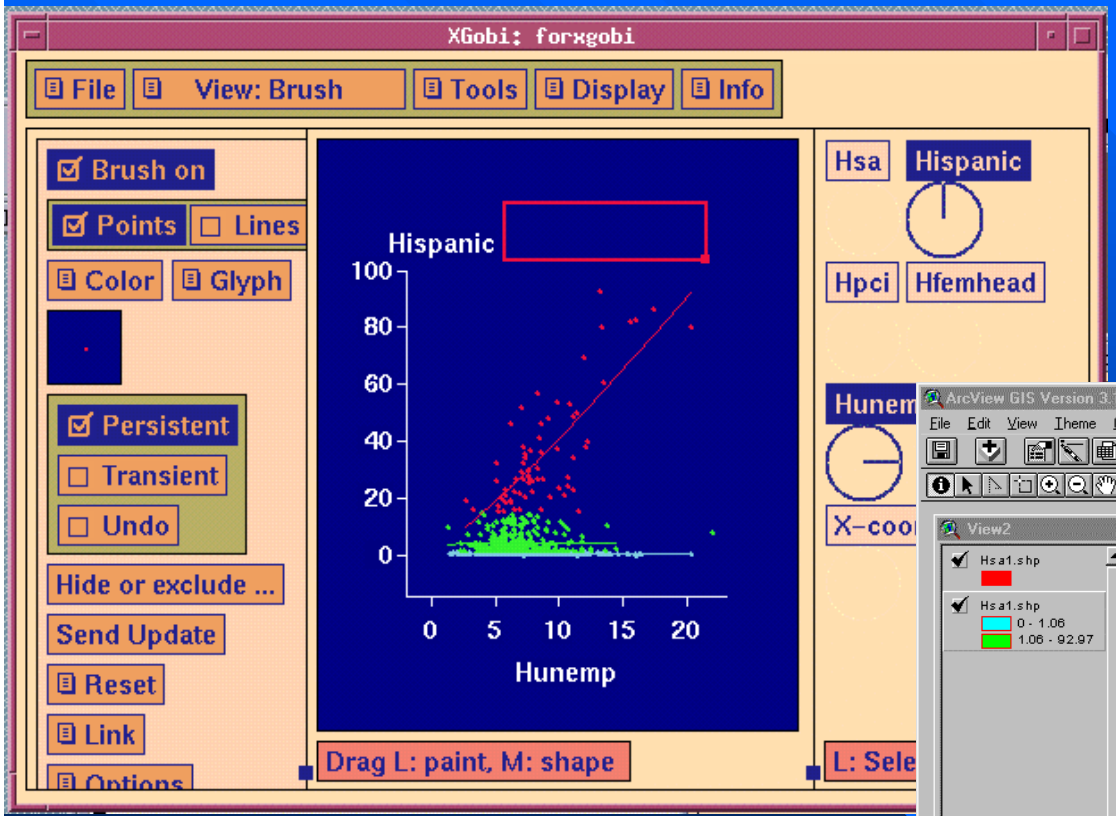
ArcView



After Brushing in XGobi

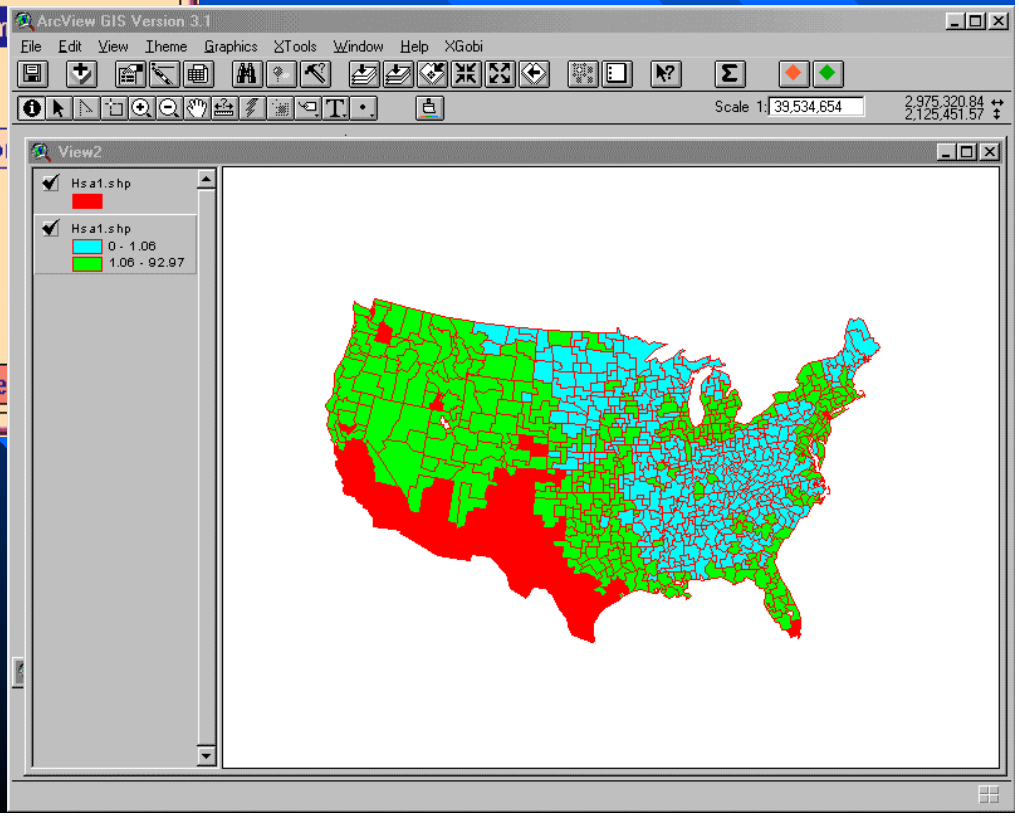


Further Data Exploration



XGobi

ArcView



Limitations of ArcView/XGobi/XploRe

- UNIX only
- RPC method passes strings among software
  - slow
  - small data sets only
- Limited linking options
- Difficult to maintain

Recent Development: GGobi

- http://www.ggobi.org
- Swayne, Temple Lang, Buja, Cook
- Freely available
- Follow-up to XGobi
- Based on GTK+ graphical toolkit
- Runs on UNIX, Linux, Windows, Macintosh
- XGobi data format & XML

GGobi and Other Software

- Designed as an open system
- Embedding GGobi within other applications (e.g., S and R)
- Extending GGobi with plugins
- Distributed/remote access via client/server architecture

R

- http://www.r-project.org/
- Gentleman, Ihaka
- Freely available
- Inspired by S/S-Plus
- External packages available, e.g.,
  - Splancs: Spatial and Space-Time Point Pattern Analysis (Rowlingson, Diggle - Bivand)
  - sgeo: Geostatistical Modeling (Majure - Gebhardt)

Live Demo: R/GGobi

- 736 records from NOAA TAO project
- Data from 5 buoys in Pacific Ocean
- 11/1/93 - 1/31/94 & 11/1/97 - 1/31/98
- SST, Air Temp, Humidy, Wind (E/W, N/S)
- Missing values
- 97/98 was El Nino year!

Achievements/Deficits

- R/GGobi more powerful than XGobi/XploRe
- Huge potential for ESDA
- Many spatial functions available for R
- Open environment that can be easily extended
- Link to GIS still missing
- Temporary solution: Map readers for R and GGobi

Questions ?