NetSurv & Data Viewer
Prototype space-time analysis and visualization software from TerraSeer

Dunrie Greiling, TerraSeer Inc.

TerraSeer

- Software sales
  - BoundarySeer for boundary detection and analysis
  - ClusterSeer for disease cluster detection
  - SpaceStat for spatial regression modeling
- Training
  - Short courses
- Custom development

www.terraseer.com
BioMedware

TerraSeer’s R&D partner
- developed BoundarySeer and ClusterSeer
- NIH/NCI SBIR funding

Selection from current projects
- NetSurv
  - distributed disease surveillance software
- Cancer Atlas Viewer
  - spatio-temporal visualization of the National Cancer Mortality Atlas
- DataViewer under construction

www.terraseer.com
**NetSurv Project**

- Provide decision support and monitoring tools that will enhance existing disease surveillance systems and support timely analysis, policy formulation, and public health actions

*www.terraseer.com*

---

**Surveillance**

- Continuous and systematic process of collection, *analysis*, and *interpretation* of information for monitoring health problems
- Ongoing monitoring of temporal and spatial disease trends

*www.terraseer.com*
NIH SBIR grants

- Small Business Innovation Research
  - Phase I
    - Evaluate scientific and technical merit and feasibility of an idea (6 months)
  - Phase II
    - Expand on the results and further pursue the development of Phase I (2 years)

NetSurv: Phase I

- Provide CuSum technique (Hutwagner et al 1997) for monitoring temporal trends, providing direct access to a surveillance database and graphical display of results
  - access to single dataset
**CuSum Technique**

- Cumulative sum over time, of the differences between observed case counts and a reference/baseline value
- Differences are added together and plotted on graph over time
- Magnifies small, abrupt change which are too small to be visible in conventional graphical plots of a fluctuating series of data

---

**NetSurv: Phase I**

- CuSum technique (Hutwagner et al 1997)
- Distributed system
  - Web browser interface “thin client”
Multi-Tier Distributed Apps

- Windows
- Mac
- Unix
- Browser
- Thin Clients
- NetSurf Components
- CuSum Method
- Security Manager
- Database Broker
- DBMS Server
- census
- surveillance
- Application Server
- Web Server

Interface screenshot

Grouping variables:
- Age (years)
  - 1-5
- Sex
  - Male
  - Female
- County
  - Area
  - Region

Parameters:
- CuSum
  - h
- k

Average disease frequency:
- mean
- median

Run Analysis, View Map, View Result Table, Logout

www.terraseer.com
**Net Surv phase I results**

- Web-based interface difficult, not user friendly
  - difficult: interface complex, difficult to implement
  - not user friendly: mapping, graphing slow, interface static not dynamic

www.terraseer.com

**BioMedware**

- TerraSeer's R&D partner
  - developed BoundarySeer and ClusterSeer
  - SBIR funding
- Selection from current projects
  - NetSurv
    - distributed disease surveillance software
  - Cancer Atlas Viewer
    - spatio-temporal visualization of the National Cancer Mortality Atlas
  - DataViewer under construction

www.terraseer.com
Motivation for Cancer Atlas Viewer

- Provide real-time visualization of the National Cancer Mortality Atlas Data
- Provide statistics for spatial, temporal, and space-time evaluation of Atlas data
- Explore general STIS specifications with a specific example

www.terraseer.com
Real Time Interaction

Avoid the “world wide wait”

Map being produced...

Please wait for map creation.

Map creation may take up to a minute, depending upon your connection speed.
Real Time Interaction

- Provide more flexible access to the data.
- Concurrency issues

www.terraseer.com
Downloading Data

Real Time Interaction
- Provide linked views that you can brush for interactive data exploration
  - Map
  - Scatterplot
  - Box plot
  - Histogram
  - Table

www.terraseer.com
Space-Time Viz

- Slideshow
  - Group of maps with a common legend

Provide Statistics

- Standardization
  - Z-score

- LISA
  - Univariate
    - spatial contagion
  - Bivariate
    - space-time contagion
  - Cluster persistence
Moran’s I

- Global statistic – 1 value for entire dataset
- Spatially weighted correlation coefficient
- Range ~ (-1, 1)

Calculation of LISA’s

1. Standardize data as z-score
   \[ z_i = \frac{(x_i - \mu_x)}{\sqrt{\text{var}(x)}} \]
2. Calculate LISA statistics (Anselin, 1995)
   - local statistic, 1 value for every location
     \[ I_i = z_i \sum w_{ij} z_j \]
3. Evaluate significance of LISA statistics via Monte Carlo randomization
The Moran Scatter Plot

- Graphs the values \((z_i)\) of each area versus the average of its neighbors
- \[ \sum w_{ij} z_j \]
- Has four quadrants that display high-high and low-low clusters, and high-low and low-high outliers

Local Clustering (LISA)
Mask Sparse Data

- Count < 6

Analyze Masked Datasets
Provide Statistics

- Standardization
  - Z-score
- LISA
  - Univariate
    - spatial contagion
  - Bivariate
    - space-time contagion
- Cluster persistence

Long Term

- Include other statistics
  - ClusterSeer
    - temporal, spatial, spatio-temp, & surveillance methods
  - BoundarySeer
    - edge detection (wombling), classification (fuzzy, spatially-constrained)
  - Other
    - change detection
- Provide open interface for user-scripted methods
  - Python

www.terraseer.com
Long Term

- Open to other data (more general product)
  - Currently - Adding visualization of points moving through time
    - modeling individuals’ movements
  - Interested in applying to infectious disease spread
    - humans
    - plant pathogen
    - amphibians

Back to NetSurv

- Replace static web-based interface with more interactive Atlas/Data Viewer like interface
NetSurv phase II

- Retain attention to data concurrency
  - web access to download data
  - check for updates
- Retain attention to permissions/privacy concerns
- Pull down data and then do analysis on local machine
  - avoids world-wide-wait for mapping, graphing

Long term plans for NetSurv

- Atlas-like interface
- Custom statistics for surveillance applications
  - User-programmed in Python
- Interact with existing web data repositories
  - DataWeb
  - Census
  - Geographic data
  - plus provide room for custom/non-public data repositories
Acknowledgments

- NetSurv was funded by a grant from the National Cancer Institute and the National Library of Medicine to BioMedware, Inc.
- The Cancer Atlas software was funded by a grant from the National Cancer Institute to BioMedware, Inc.

www.terraseer.com