Human-Centered Design of Geovisualization Tools for Cancer Epidemiology

Anthony C. Robinson, Jin Chen, Hans G. Meyer, Alan M. MacEachren
GeoVISTA Center
The Pennsylvania State University
University Park, PA 16802, USA
{arobinson, jxc93, hmeyer, maceachren} @psu.edu

New methods of geovisualization hold great promise for a number of domain-specific applications. Geovisualization tools provide the capability to visually explore and analyze complex spatial and spatio-temporal data. Implementing these kinds of tools in a usable and applicable manner is a particular challenge facing the GIScience community today (MacEachren and Kraak 2001; Muntz et al. 2003). We will present our multi-faceted approach to human-centered design through a case study of an exploratory geovisualization toolkit designed specifically for application in cancer epidemiology.

ESTAT, the Exploratory Spatio-Temporal Analysis Toolkit, has been developed by the Penn State GeoVISTA Center under a contract with the National Cancer Institute (NCI). NCI requested a geovisualization application that could accomplish a number of specific tasks as well as fit within their existing computing architecture and long-term development goals. NCI's interest in an application for their needs stems from prior work in developing geovisualization tools for exploring health data. In the past, maps have promoted the creation of new hypotheses related to health predictors and outcomes (Pickle 1996). New methods of geovisualization are likely to enhance this capability through their dynamic and interactive nature (Carr et al. 2000; MacEachren et al. 1999).

ESTAT features four primary visual analysis elements. We have combined a scatter plot, bivariate map, time series plot, and parallel coordinate plot in a dynamically-

linked interface (Fig. 1). In each element, brushing, selection, classification, and color scheme choice is linked in real-time.

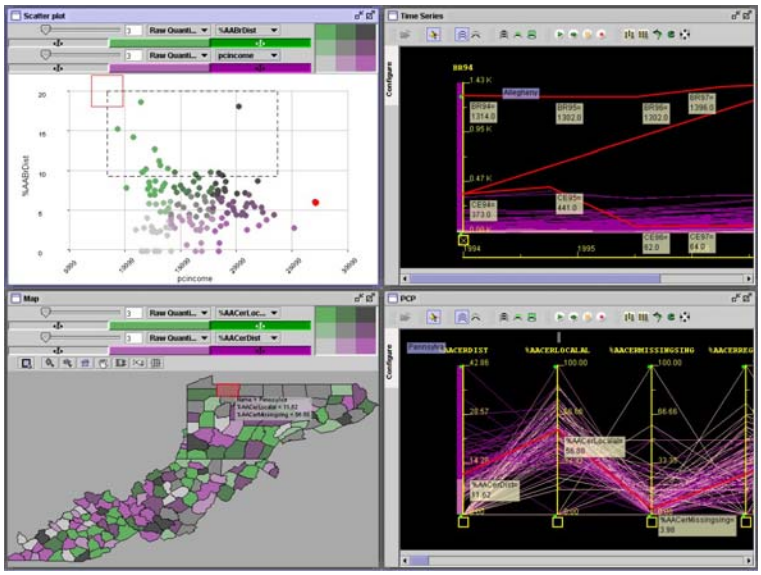

Figure 1: The ESTAT Toolkit

These individual geovisual elements are not new themselves. Used in combination, however, these dynamically linked tools support flexible visual exploration and analysis of geospatial health data and covariates across space and time. The parallel coordinate plot (PCP) in ESTAT has been developed specifically to support a range of analysis tasks determined by cancer data analysts to be relevant for cancer surveillance and control (e.g., display of box plots on each axis, alignment of axes on median values, generation of summary plots depicting state averages from county data or medians for each data class). Additionally, considerable effort has been made to integrate these tools into an interface that works for our users in epidemiology. Specifically, we have built a customized data loading and selection mechanism, incorporated new statistical measures, and modified configuration options; all as a result of iterative usability testing.

ESTAT is based on our geovisualization tool platform GeoVISTA *Studio* (http://www.geovistastudio.psu.edu/jsp/index.jsp), a Java-based visual programming tool (Takatsuka and Gahegan 2002). In *Studio*, applications are built by connecting individual components (Java Beans) in a codeless environment. *Studio*'s inherent flexibility has enabled us to quickly modify our tools and rectify bugs when necessary.

The current version of ESTAT is the result of a continuing effort to design the toolkit in a human-centered manner. We have conducted an array of usability studies, both formal and informal, in order to assess and refine our tools. At every step we have made a concerted effort to understand our users and the ways in which they intend to interact with these tools. Prior efforts to create usable geovisualization tools, including some of our own (Edsall et al. 2001), have tried to incorporate user input, but have limited this input to specific portions of the software design process (Andrienko et al. 2002; Slocum et al. 2003). Often, geovisualization tools are created before a user group is identified, and it is challenging to adapt them for a specific context of use. Our efforts have specifically targeted the task of providing tools tailored for the tasks of cancer epidemiology.

We will present the process we have undertaken and illustrate the evolution of our toolkit over time into its present form. In addition, we will provide insight into the benefits and detriments of the usability assessment techniques that we have employed so far. Specifically, we will emphasize techniques that appear to be well-suited for assessing the utility of exploratory geovisualization tools. Finally, we will describe how we have involved our colleagues in cancer epidemiology in each step of the software development process.

Acknowledgements

This material is based upon work supported by the National Institutes of Health under

Grant no. R01 CA95949-01.

References

Andrienko, G. L., N. V. Andrienko, H. Voss, F. Bernardo, J. Hipolito, and U. Kretchmer. 2002. Testing the usability of interactive maps in common GIS. *Cartography and Geographic Information Science* 29:325-42.

Carr, D., J. Wallin, and D. A. Carr. 2000. Two new templates for epidemiology applications: Linked micromap plots and conditioned choropleth maps. *Statistics in Medicine* 19:2521-38.

Edsall, R. M., A. M. MacEachren, and L. W. Pickle. 2001. Case study: Design and assessment of an enhanced geographic information system for exploration of multivariate health statistics. Proceedings of IEEE Symposium on Information Visualization 2001, San Diego, CA, October 22-25.

MacEachren, A. M., and M.-J. Kraak. 2001. Research challenges in geovisualization. *Cartography and Geographic Information Science* 28:3-12.

MacEachren, A. M., M. Wachowicz, R. Edsall, D. Haug, and R. Masters. 1999. Constructing knowledge from multivariate spatiotemporal data: Integrating geographical visualization with knowledge discovery in database methods. *International Journal of Geographical Information Science* 13:311-34.

Muntz, R. R., T. Barclay, J. Dozier, C. Faloutsos, A. M. MacEachren, J. L. Martin, C. M. Pancake, and M. Satyanarayanan. 2003. *IT roadmap to a geospatial future, report of the committee on intersections between geospatial information and information technology*. Washington, DC: National Academies Press.

Pickle, L. W. 1996. *Atlas of United States mortality*. Hyattsville, MD: National Center for Health Statistics, Centers for Disease Control and Prevention.

Slocum, T., D. Cliburn, J. Feddema, and J. Miller. 2003. Evaluating the usability of a tool for visualizing the uncertainty of the future global water balance. *Cartography and Geographic Information Science* 30:299-317.

Takatsuka, M., and M. Gahegan. 2002. GeoVISTA *Studio*: A codeless visual programming environment for geoscientific data analysis and visualization. *Computers and Geosciences* 28:1131-44.