OpenGeoDa, Open Source Software for the Exploration and Visualization of Geospatial Data

Luc Anselin
GeoDa Center
Arizona State University
Tempe, AZ 85287-5302

Luc.Anselin@asu.edu

Mark McCann
GeoDa Center
Arizona State University
Tempe, AZ 85287-5302

Mark.Mccann@asu.edu

ABSTRACT

OpenGeoDa is the open source successor to GeoDa, a software package designed to introduce non-experts to spatial data analysis. GeoDa was developed under the auspices of the NSF funded Center for Spatially Integrated Social Science (CSISS) [4]. Since its release in 2003, "Legacy GeoDa" has gained over 40,000 individual users and has become a standard to teach introductory spatial data analysis. GeoDa (and Open GeoDa) was conceived as enabling researchers to move from geovisualization to exploratory data analysis, the study of spatial autocorrelation and ending up with spatial regression modeling.

Categories and Subject Descriptors

Algorithms

General Terms

Algorithms

Keywords

GIS and Spatial Analysis, Exploratory Spatial Data Analysis, Spatial Autocorrelation, GeoDa.

1. INTRODUCTION

OpenGeoDa is the open source successor to GeoDa, a software package designed to introduce non-experts to spatial data analysis. GeoDa was developed under the auspices of the NSF funded Center for Spatially Integrated Social Science (CSISS) [4]. Since its release in 2003, "legacy GeoDa" has gained over 40,000 individual users and has become a standard to teach introductory spatial data analysis. GeoDa (and Open GeoDa) was conceived as enabling researchers to move from geovisualization to exploratory data analysis, the study of spatial autocorrelation and ending up with spatial regression modeling. Moving GeoDa to an open source and cross-platform paradigm required total recoding of the program, using the wxWidgets toolbox to implement native look and feel on each platform. At this point, OpenGeoDa runs on Mac

OSX, Win XP and Vista, and Linux Ubuntu, in both 32 bit and 64 bit versions. The development of OpenGeoDa is being carried out at the GeoDa Center for GeoSpatial Analysis and Computation at Arizona State University.

OpenGeoDa is designed to provide advanced geovisualization and spatial analytical functionality not typically found in commercial GIS. This includes unique features such as outlier mapping, cluster mapping (based on local indicators of spatial autocorrelation), visualization of spatial autocorrelation (using the Moran scatter plot) and spatial regression (including advanced algorithms to deal with very large data sets). In addition, standard exploratory data analysis methods such as histograms, box plots, scatterplots, parallel coordinate plots, 3-D scatter plots, conditional plots, etc. are included as well.

This demonstration illustrates the functionality of OpenGeoDa on Mac OSX, Windows Vista and Ubuntu Linux and briefly reviews the core functionality of the program.

2. FUNCTIONALITY

The functionality of OpenGeoDa closely follows that of Legacy GeoDa and can be classified into six major groups: data manipulation, mapping, spatial weights creation, multivariate (non-spatial) exploratory data analysis (EDA), spatial autocorrelation analysis and spatial regression.

Data manipulation allows the creation and manipulation of spatial information into standard industry formats, such as ESRI shape files. For example, this allows simple text files of point coordinates to be turned into shape files.

The mapping functionality in OpenGeoDa includes standard choropleth maps as well as so-called outlier maps, such as a box map, which highlight the location of extreme values on the map. In addition, map animation, cartograms and conditional maps are included. In addition, specialized functionality is included to deal with maps for rates (proportions) which implements a range of smoothing algorithms.

The special niche of OpenGeoDa is how it deals with spatial data exploration. It includes extensive functionality to create spatial weights [3], compute local and global spatial autocorrelation coefficients and visualize these in maps and graphs [2].

There is also standard multivariate EDA, such as histograms, box plots, scatterplots, parallel coordinate plots and 3-D scatterplots. Finally, the spatial regression functionality implements several specialized algorithms to carry out estimation by means of maximum likelihood methods [5]. It also includes a battery of

diagnostic tests for spatial effects in classical regression models [1].

OpenGeoDa (as GeoDa) is designed to allow for full linking and brushing across all "windows." This allows for the selection of subsets of observations in all windows simultaneously, a powerful feature to facilitate the exploration of patterns in spatial data.

The port from Legacy GeoDa to OpenGeoDa was implemented using the wxWidgets open source GUI library. This allows for a single code base to support a wide range of platforms. Currently we have an alpha release of versions that run on Windows XP, Windows Vista, Mac OSX and Linux Ubuntu (Figures 1- 4).
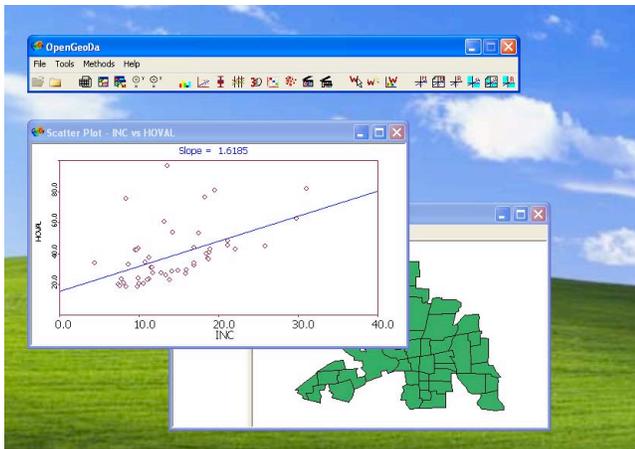

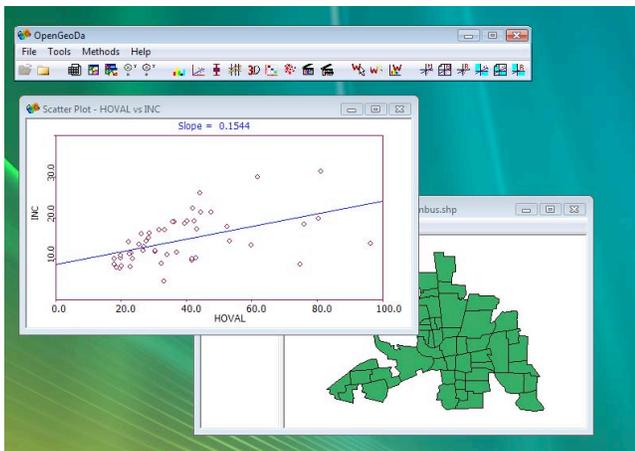**Figure 1. OpenGeoDa interface in Windows XP**
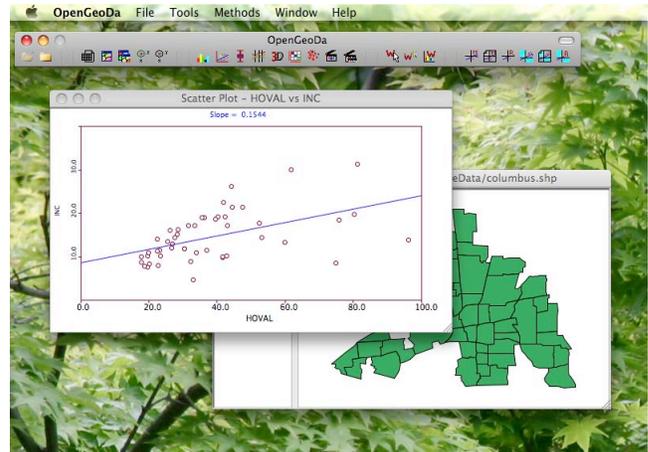

Figure 2. OpenGeoDa interface in Windows Vista
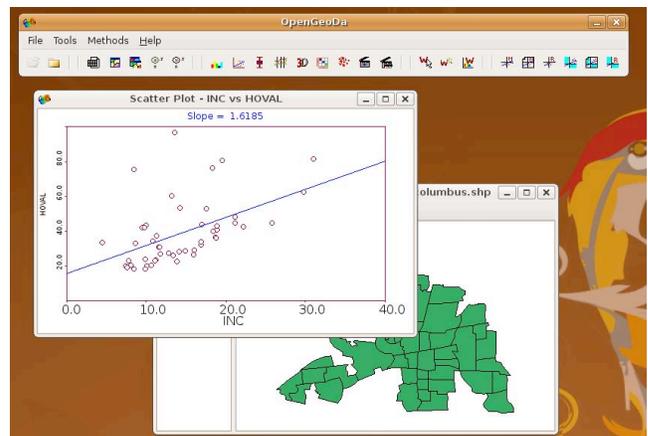

Figure 3. OpenGeoDa interface in Mac OSX

Figure 4. OpenGeoDa interface in Linux Ubuntu

3. ACKNOWLEDGMENTS

4. REFERENCES

[1] Anselin, L. (1988). Spatial Econometrics, Methods and Models. Dordrecht, Kluwer Academic.

[2] Anselin, L. (1995). Local Indicators of Spatial Association – LISA. Geographical Analysis 27, 93-115.

[3] Anselin, L. and O. Smirnov (1996). Efficient Algorithms for Constructing Proper Higher Order Spatial Lag Operators. Journal of Regional Science 36, 67-89.

[4] Anselin, L., I. Syabri and Y. Kho (2006). GeoDa: An Introduction to Spatial Data Analysis. Geographical Analysis 38, 5-22.

[5] Smirnov, O. and L. Anselin (2001). Fast Maximum Likelihood Estimation of Very Large Spatial Autoregressive Models: A Characteristic Polynomial Approach. Computational Statistics and Data Analysis 35, 301-319.