

GIST: A NEW TOOL FOR VISUALIZING GEOGRAPHIC DATA

Alison M. Eyth, Prashant P. Pai*
 MCNC Environmental Modeling Center, RTP, NC
 e-mail: eyth@emc.mcnc.org
 Web address: <http://www.emc.mcnc.org>
 Voice (919) 248-9256 Fax (919) 248-9245

1. INTRODUCTION

This paper describes the motivation behind the development of, and the features included in, a pure Java Geographic Information System (GIS) tool for geospatial data analysis and visualization. The use of existing GISs is often limited by issues such as platform-dependence, cost, and inability to support time-varying and/or 3-dimensional data.

With the GIS Tool (GIST), we seek to combine traditional GIS features with additional features, such as animating data, extracting slices of data from three-dimensional data sets, and computing and plotting the results of mathematical functions. As a result, GIST will be portable, powerful, and useful for many different types of applications that use geospatial data – such as the analysis of multimedia model, air quality model and emissions inventory data.

GIST will have the flexibility to be deployed as both a desktop application for powerful interactive analysis, and as an applet to facilitate information dissemination over the Web. Basic features include setting and changing map projections, adding and deleting layers, and probing. More advanced features are also implemented such as probing attributes, thematic mapping, query building and execution. Future directions such as animation and formulas are described.

2.0 PAVE AND ITS LIMITATIONS

The MCNC Environmental Modeling Center has a great deal of experience with air quality, meteorological, and emissions models. The datasets produced by these models are large (multiple gigabytes), time varying, and on three-dimensional grids. In the mid 1990s, MCNC developed the Package for Analysis and

Visualization of Environmental data (PAVE) to help model developers and users interpret the data sets used and produced by these models. PAVE was (and still is) very popular in the air quality modeling community. It can subset the data to the user's specifications, and interactively display colorful animations of the data plotted in conjunction with standard map backgrounds.

PAVE incorporates a unique formula syntax that allows it to perform mathematical computations using data from multiple datasets, and then display the results (e.g. the percent difference between two model outputs). PAVE plots are customizable and can be saved to several common graphics formats. PAVE supports a scripting language that automates analyses, thus making them repeatable.

But, PAVE has its limitations. For example:

- PAVE's graphics are written in X-windows and Motif, which are not very popular in today's PC desktop environment.
- PAVE must be compiled separately for every computer platform (e.g. Sun, HP, Windows 2000).
- Only a few maps are provided with PAVE (i.e. states, counties, rivers, and roads), and its ability to render these maps is very limited (e.g. all rivers are drawn with the same line thickness).
- PAVE is designed to work with gridded data for which all grid cells are uniform rectangles, therefore new data and plot types are very difficult to add and display.
- PAVE plots are accessible over the Web only as static images, although they are sometimes animated with the help of a Java applet.

3.0 APPLYING GIS CONCEPTS

After using PAVE for several years, Alison Eyth of MCNC recognized its limitations and began working on a concept for a new

* *Corresponding author addresses:* Alison Eyth or Prashant Pai, MCNC Environmental Modeling Center, P.O. Box 12889 RTP, NC 27709

visualization tool. Conceptually, a GIS addresses many of PAVE's limitations in that it allows many types of data to be integrated within a single display. According to <http://www.gis.com>, GIS software provides the functions and tools needed to store, analyze and display information about places.

The basic requirements of a GIS include displaying geographic data (e.g., points, lines, and polygons) on a *map* according to a map projection, being able to view the data as *layers*, and having the ability to display many different layers in the same space. Note: A map projection enables the 3-dimensional surface of the earth to be mapped to a plane surface using the assumption that the earth is an ellipsoid. A GIS should be able to convert the coordinates of points, lines, and polygons from one map projection to another (e.g. from Latitude-Longitude to Universal Transverse Mercator (UTM)).

In addition to the basic features, most GISs provide thematic mapping capabilities to color the data objects according to some attribute, and many have a programming and/or scripting language to make maps reproducible.

Within the Geographic Information System Tool (GIST), we seek to combine traditional GIS features with some of the functions that made PAVE so popular – such as animating data, extracting slices of data from three-dimensional data sets, and plotting the result of mathematical formulas. GIST is a portable, powerful application that is useful for many different types of applications that involve geospatial data.

4.0 GIST FEATURES

4.1 Java and JLoox

GIST is written in Java. This makes GIST portable across PC, Unix, and Mac platforms without any recompilation.

To provide basic GIS functions, we have purchased a commercial off-the-shelf Java library called JLooxGIS [1]. JLooxGIS is a flexible system with built-in implementations for important features such as

- plotting points, lines, and polygons at specified coordinates to create a map,
- coordinate transformations;
- zooming;
- printing;
- panning (via a map overview);
- toggling layer visibility, and

- layer control (i.e. controlling the order in which layers are drawn).

The use of JLoox has saved a lot of development time for GIST. However, a disadvantage is that it prevents the possibility of GIST becoming an open source tool. This is because to give away the source code would show people how to use JLoox and essentially be giving away development licenses. The good news is that there are no run-time license fees for JLoox, so GIST could be made available in a low- or no-cost form similar to PAVE.

4.2 GIST Main GUI

GIST adds a number of additional features on top of JLooxGIS, including a standalone application and applet, a layer library, probing, queries, thematic mapping, and animation. The GIST Main GUI consists of four primary components:

- a layer controller panel for specifying the ordering and rendering style of each layer,
- a query panel to support querying based on data attribute values,
- a map viewer panel to display the data as a map,
- and an overlay panel provided by JLooxGIS to support panning of the map display.

Our layer controller enables the user to change the selected map projection. Each object shown on the map is reprojected to its new location when the map projection changes. Layers can be added to the map from the layer library (see section 4.2). They can be hidden or shown, and raised or lowered. Raising a layer above another layer causes its objects to be plotted on top of the objects in the lower layers.

As stated above, JLooxGIS provides the basic concepts of layers and rendering style for geographic objects. The controls for rendering the various types of objects are provided on a Layer Options panel. From this panel, the user can set the layer's properties such as its object's colors, size, symbol, and line thickness. Polygon shapes can have both an outline color as well as a fill color. Also, the polygon outline can be of different styles and thickness. Line shapes have a color, thickness and style associated with them. Points have a color and thickness, and can be of various shapes such as circle, square, diamond, triangle, star or cross.

The map viewer panel contains a toolbar that allows the user to zoom, print, pan, set a region of

interest, display mouse pointer coordinates, etc. The user can also see the position of the mouse on the map in either latitude-longitude coordinates or projected x, y coordinates. The region of interest for the map can be set in world coordinates, such that only shapes that lie in this region will be shown on the map.

4.2 Layer Library

In GIST, the layers that have been loaded by the user are stored in a “layer library”. The layer library has multiple purposes. First, it saves the user time because it maintains the list of layers between sessions. A second purpose is to create a concept of a layer independent of the file format. The currently supported data formats are: ESRI Shapefiles [2], and an ASCII format that allows the user to plot text, points (with optional labels and symbols), lines, and polygons at specified latitude-longitude coordinates. In the future, we hope to add many air quality model related formats such as netCDF and IDA.

The layer library has its own window in the GUI. When a file is added to the layer library, it is given a “friendly name”, and that becomes the name that the user sees on the main GIST GUI. If GIST is running as an applet, the layer library allows the user to load the layers of interest from user-friendly names, without having to find the data files on the remote server.

4.3 Probing Data Attributes

Each object in a layer includes both coordinates that allow it to be placed on a map and fields of attribute data that provide additional information about it. All of the objects in a layer have the same attribute fields, but each object has its own values for each field (e.g., each city would have *its own* population and area attributes). By pressing the right mouse button, the user can see attribute names and values for the data object on the top-most layer under the mouse pointer. The name-value pairs appear as a popup window on the screen. In the future, this feature could be expanded to allow the user to view the attribute information for the objects from all the layers under the mouse pointer.

4.4 Thematic Mapping

In a thematic map, the objects in a layer are placed into discrete bins according to their value for a given attribute. Each bin is then rendered in

a different way. Examples of ways to differentiate the objects in the bins are by color, shape, and size. For example, you may want to draw points that represent cities with small populations as little black circles and cities with large populations as big red diamonds, with a few gradations in between. Thus, thematic mapping allows us to visually classify the objects based on their values for a particular attribute.

To create a thematic map in GIST, the user first opens the Layer Options Panel for the desired layer. From this panel the user checks the Thematic Map checkbox, then chooses a numeric attribute available for the layer and the number of bins to use. GIST then computes the highest and lowest values for the attribute, scales the values of the attribute linearly into each bin, assigns each object in the layer to a bin and renders them accordingly. An example of a thematic map for polygons created using GIST is shown in Figure 1.

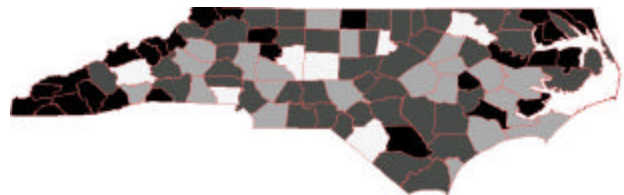


Figure 1. Thematic Map of Volatile Organic Compound Emissions in North Carolina.

Currently the only aspect of rendering that can be set for a thematic map in GIST is the object’s color. But one can see that it would also be useful to adjust all the rendering aspects of each object type, such as the size or line thickness, and the symbol type or fill pattern. Other expansions to the thematic mapping feature will support non-linear scaling of the bins to be used, and thematic mapping for non-numeric attributes that have a discrete set of values (e.g. soil type).

4.5 Querying Attribute Values

Queries can be performed on the attributes for the currently selected layer (e.g. find all the cities with a population > 100,000). A Structured Query Language (SQL)-like approach is used for creating and executing queries [3]. More complex queries can be created using the AND and OR functions. In the future, even more complex queries could be handled through the addition of grouping delimiters and query functions.

The objects that match the query specifications can be shown on the map or viewed in a table. The query mapping feature hides the shapes that do not match the query. Thus, the

user can visually determine which shapes matched a query. The results table presents the attributes of the objects that satisfied the query. Either all the attributes can be shown, or the user can select a subset of attributes to be displayed. The table can be sorted according to the values of a selected attribute (e.g. from the highest to lowest NOx emission values). The results of a query can be saved to a file, and could then be loaded into a spreadsheet or other software.

5.0 FUTURE DIRECTIONS

There are many possible future directions for GIST, but the tool cannot be improved until one or more funding sources are found. Most of the initial development of GIST was funded by MCNC, but MCNC does not plan to develop the tool further. The parts that were not funded by MCNC are the Map Projection Dialog, some parts of the map viewer panel, and the ASCII file format. These were developed under contract to the United States Environmental Protection Agency (EPA) Office of Research and Development via GSA task #4TAG68002718. These components are available as part of a plug-in to EPA's Multimedia Integrated Modeling System (MIMS).

If one or more funding sources can be identified, some of the enhancements we have envisioned include:

- a scripting language;
- a formula capability similar to that in PAVE;
- saving images to various formats;
- support for additional file formats (e.g. emission inventory files, and model inputs and outputs);
- improved probing tool to show attributes from all layers;
- improved overview control (e.g. synchronization with map viewer display);
- saving the configuration of the current map and restart GIST in the same state;
- support for more complex queries;
- viewing satellite images and digital terrain elevation data;
- labeling objects;
- measuring distances on the map;
- selecting objects within a specified radius of a coordinate; and
- data subsetting and improved memory management (i.e. not reading entire data sets into memory).

Another major potential enhancement is animation. For this, we propose a generic animation capability that will allow the user to select a set of attributes over which to animate from a single dataset or from multiple datasets. Controls will be provided to step forward and backward through the animation, go to specific frames of the animation, and to continue looping through the animation.

This generic animation capability will allow us to emulate PAVE's capability of animating through time, but it is even more flexible. Suppose you have seasonal data by county, then you could choose those attributes and animate through the values for each season. If the attributes selected corresponded to layers of an air quality model grid, this feature would support animating through the grid in the Z direction. If the user is allowed to select a set of attributes from multiple layers to animate over, situations for which boundaries change over time could be supported (e.g., agricultural areas and variable grid data sets).

6.0 CONCLUSIONS

We believe that GIST would provide the environmental community with an easy-to-use, flexible, and portable tool for the analysis of environmental data such as emission inventories, air quality model data and observations, and multimedia model input and output data. We hope that we can find a way to provide the environmental community with access to this tool at a low cost.

7.0 REFERENCES

- [1] JLooxGIS, <http://www.loox.com>
- [2] ESRI Shapefile Technical Description, <http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf>
- [3] Celko, Joe. SQL for Smarties: Advanced SQL Programming, Morgan Kaufman, San Francisco, 2000.