Dynamic Graphics in a GIS: A Link between ARC/INFO^{TM} and XGobi

Jürgen Symanzik[1], **James Majure**[2], **Dianne Cook**[1], **Noel Cressie**[1]

[1] Department of Statistics, Iowa State University, Ames, IA 50011, USA

[2] GIS Support and Research Facility, Iowa State University

symanzik@iastate.edu

Abstract. This paper describes a link between a Geographical Information System (GIS), ARC/INFO^{TM} , and an interactive dynamic graphics program, XGobi. GISs provide a user with a standard and convenient software for spatial geographical data. In particular, the GIS ARC/INFO is a combination of two systems: ARC maintains the spatial information of map features and provides tools for spatial analyses while INFO maintains the thematic or attribute information associated with the map features. XGobi is an interactive dynamic graphics program for data visualization in the X Window System^{TM}. It is designed for the exploration of multivariate data, primarily by manipulating and displaying scatterplots in arbitrary dimensions.

The motivation for the work is to link the dynamic, interactive strengths of XGobi for visualizing high–dimensional data with the exhaustive map handling tools of ARC/INFO, specifically to explore spatial data. This paper presents information about the technical realization of the link between ARC/INFO and XGobi as well as an introductory example of its use.

1 Introduction

Interactive and dynamic graphics for high–dimensional data have proved useful for exploring relationships among multiple variables. Incorporating similar tools in the context of spatial data promises to be a valuable aid in exploring spatial dependencies. Geographical Information Systems (GISs) have developed sophisticated capabilities for managing multivariate spatial data bases but limited capabilities for conducting interactive exploratory data analysis. The combination of a GIS and a dynamic graphics system for multivariate data comprises a potentially powerful tool for interactive exploratory spatial data analysis.

In Section 2 we describe general features that should be available for an interactive dynamic graphics tool that operates on data available in a GIS data base. As one particular application, the interface between the GIS ARC/INFO^{TM} and XGobi, an interactive dynamic graphics program for data visualization in the X Window System^{TM}, is described in Section 3. An example is given in Section 4. We conclude this paper by describing possibilities for future work.

2 Integration of Interactive and Dynamic Graphics Tools into a GIS

The inclusion of spatial location in data analyses can be addressed in different ways. Using a GIS, we maintain a geographic context of spatial location relative to landcover, streams, roads, and other relevant information. It would also be feasible to include the spatial coordinates as two (or d) additional variables into the analysis, but this approach by itself does not exploit the considerable spatial capabilities of GISs.

Emphasis in GIS development has been on the input of data, its management (storage, retrieval), and the display of maps, graphs, and tables. GISs have some capability to allow statistical analyses but it is generally limited. A number of recent suggestions have been made (e. g., Openshaw, 1991; Anselin and Getis, 1992; Ding and Fotheringham, 1992; Fotheringham and Rogerson, 1993) to redress this imbalance. Still others have incorporated some dynamic graphical tools into systems that lack the full features and flexibility of a GIS (e. g., Haslett et al., 1991).

Our research addresses the extremely important problem of multivariate exploratory spatial data analysis in a GIS. GIS data structures allow the representation of areal features (e. g., for the storage of information reported at an aggregated spatial level, such as counties or census tracts), linear features (e. g., for the storage of information collected from a stream or a transportation network), and point features. The topological data structure of a GIS makes it possible to determine spatial relationships between sampling locations, such as stream sites, that would be difficult to determine otherwise. The display capabilities of a GIS allow the spatial variables to be overlaid on a background of hydrography, transportation, population, land use, or other information relevant

---

^{TM} *ARC/INFO* is a trademark of Environmental Systems Research Institute, Inc.

^{TM} *X Window System* is a trademark of MIT.

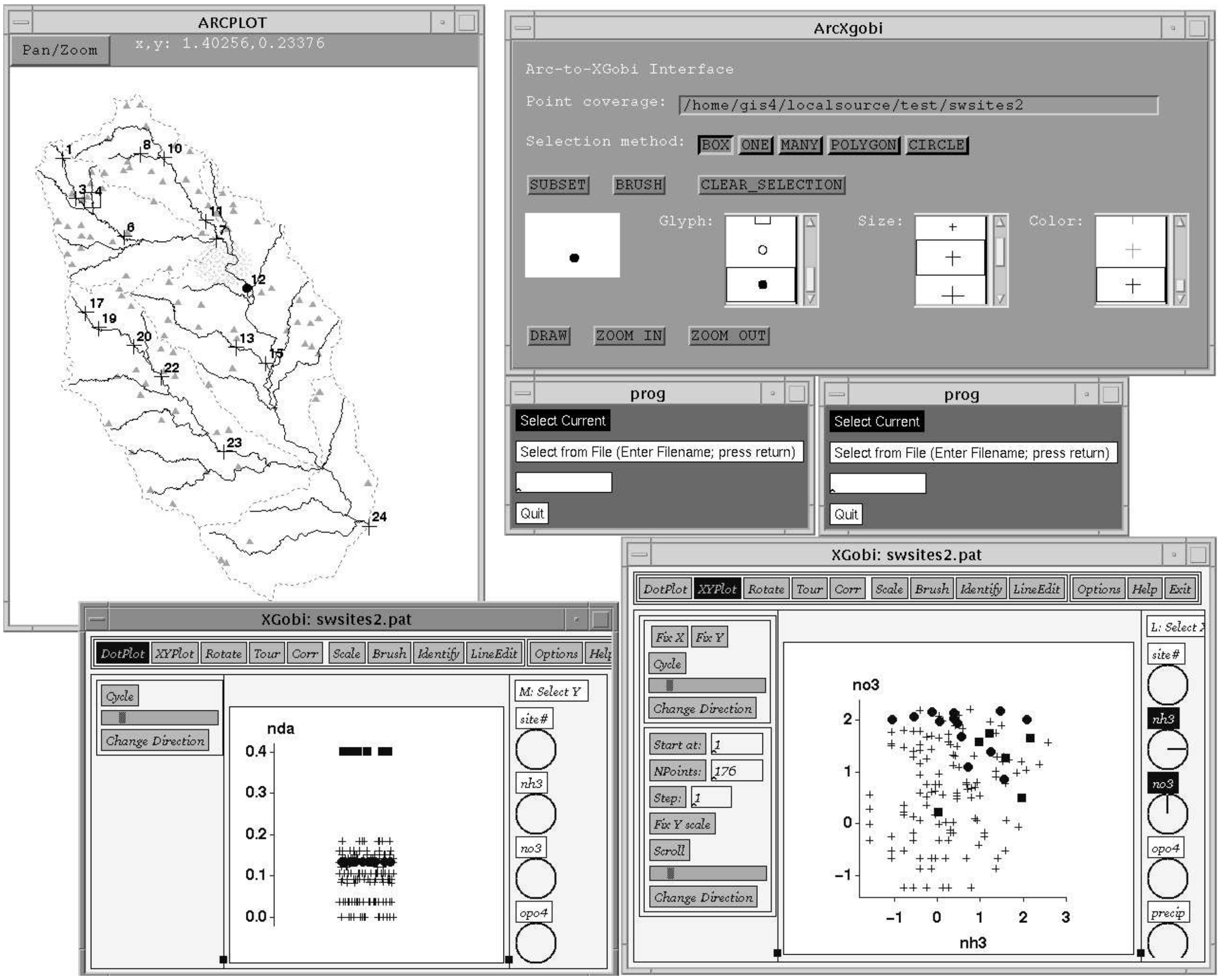

Figure 1: *ARC/INFO control panel and example map view linked to two XGobi views.*

to the attributes[1] being considered. For example, in Figure 1 the map view shows sampling sites along streams in Erath County, Texas. Information about the topography or land use near a sample site can give valuable insights into the values of attributes (e. g., ammonia concentration) collected at the site.

A GIS is intrinsically multivariate and yet this is ignored by the largely univariate statistical analyses currently available. By building an interface between a GIS and software for dynamic graphics, we will also provide a platform for developing new spatial graphical methods for spatial data sets available in the GIS (e. g., Cook et al., 1994).

[1]In the context of GISs the expression *attribute* is used instead of the statistical expression *variable*.

3 The ARC/INFO to XGobi Interface

Our efforts have focused on interfacing the GIS software ARC/INFO with XGobi (Swayne et al., 1991). ARC/INFO has been chosen because it is one of the most frequently used GIS systems and because it is extensible through its macro language, allowing the development of menus and programs to carry out ARC/INFO tasks. XGobi provides interactive and dynamic graphical tools in the X Window System environment for exploring multivariate data through the manipulation of scatterplots. ARC/INFO is used to maintain the GIS data base and to display the geography, while XGobi is used primarily to explore the relationships within and between the attributes. Figure 2 shows how the communication between these two programs is established.

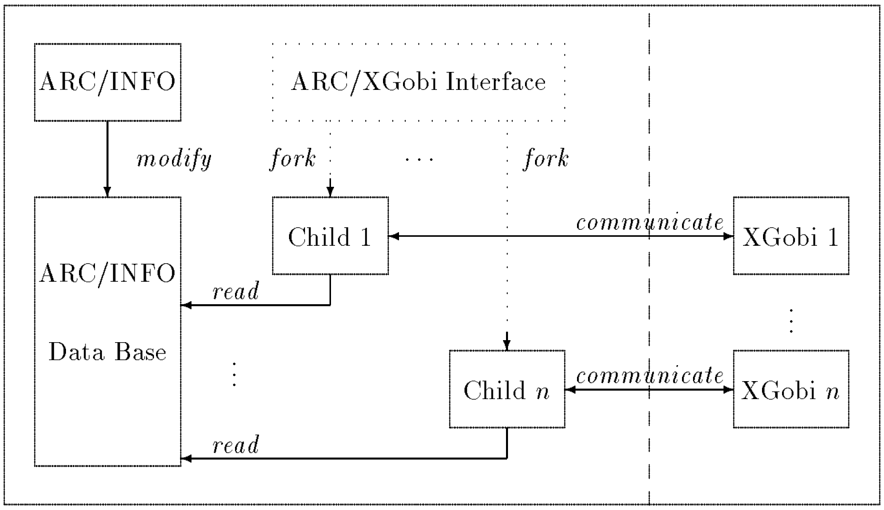

Figure 2: *Interface linking ARC/INFO with XGobi.*

3.1 The ARC/INFO Part

ARC/INFO is used to display the location of sampling sites in a graphics window. The sampling sites can be displayed on a background of roads, streams, or any other relevant geographic data sets available. A control panel, the upper right window shown in Figure 1, allows the user to brush or subset the sampling sites interactively. The term "brush" refers to changing the symbol used to represent the specified points and "subset" refers to choosing a subset of the points for further analysis, disregarding (temporarily) the other points. The brushed (or subsetted) sites are redrawn with the specified glyph, size, and color, and the ARC/INFO data base is modified. These changes are detected and passed to XGobi by the intermediate process, as described in subsequent sections.

The ARC/INFO portion of the application is implemented with AML (Arc Macro Language) and works as follows. An ARC/INFO data set consists of a set of spatial features, in this case points, each of which has a record in a data base table. When the application starts, a column in the table is initialized with a default value which represents the symbol, i. e., glyph, size, and color, with which to draw each point. As the user interactively queries the points, the values in this column are updated to reflect the user's actions. The changes to this column are detected by the intermediate ARC/XGobi server process and sent to XGobi.

Pseudo code for the ARC/INFO part is given below. In this pseudo code, the control panel is represented by the repeat loop.

```
set current symbol to default symbol
repeat
  wait for user action
  case user action {
    when "identify ARC/INFO data set"
      initialize the symbol column to current symbol
    when "brush"
      spatially select points to brush
      set symbol column of selected points to current symbol
    when "subset"
      spatially select points to subset
      set symbol column of selected points to current symbol
      set symbol column of other points to 0
    when "clear selection"
      reset the symbol column of all points to default symbol
    when "change color"
      reset current symbol to reflect changed color
    when "change glyph"
      reset current symbol to reflect changed glyph
    when "change size"
      reset current symbol to reflect changed size
  }
until (forever)
```

3.2 The Intermediate Process

The intermediate process, denoted as ARC/XGobi Interface in Figure 2, has to serve the requests of the XGobi clients by reading information from the ARC/INFO data base. The interprocess communication between this server and the XGobi clients is based on Stevens' (1990) concurrent server example, and uses a Transmission Control Protocol (TCP) socket, i. e., an Internet stream socket. Upon receiving a connection request from an XGobi client, the intermediate process forks an identical child process. Each child process communicates with one XGobi client; thus, one–to–one connections between server processes and XGobi clients are established.

Obviously, the forking of child processes is a heavy weight mechanism to provide a concurrent server. However, we assume that this mechanism is available for all hardware environments that support ARC/INFO. An alternative for some workstations (e. g., DEC™) is the use of multithreads which are light weight processes, but this approach is not available on all systems (e. g., Sun™/Sparc™ workstations).

The main task of the child processes is the following: If the XGobi client indicates that it wants the currently selected ARC/INFO data set and future updates of this selection, the related ARC/XGobi server (child) has to check continually whether the ARC/INFO data base has been changed. If so, the modifications, such as new brushed or subsetted points, are immediately passed to the corresponding XGobi clients.

A child process is terminated by a QUIT command of its XGobi counterpart, or if it detects the unexpected termination of the client process or the breakdown of the communication channel. The intermediate (parent) process will operate until it is explicitly terminated by the user. The pseudo code for the ARC/XGobi interface follows.

Parent:

```
init ARC/INFO defaults
init sockets
repeat
  accept connection from XGobi client
  fork child process
until (forever)
```

Children:

```
repeat
  wait for input from XGobi client or for Timeout
  if (input received = SEND Filename)
     then {send data from file Filename; Update = false}
  else if (input received = SEND current)
     then {send data from current selection; Update = true}
  else if (Timeout and Update)
     then if (current selection modified since last send)
             then send update of current selection
until (input received = QUIT or abort of client
                               or channel down)
```

3.3 The XGobi Part

There are several methods that we considered when initially contemplating a link from ARC/INFO to XGobi: directly writing new functionality into XGobi, accessing the XGobi data structures by calling XGobi as a subroutine, or using the linked brushing protocols existing in XGobi. The first method is feasible because the code for XGobi is available, but it is undesirable because it would require maintaining updates with new releases of the XGobi code. The third option strictly limits the interaction to the data structures available in the XGobi linked brushing code. Calling XGobi as a subroutine from a small control panel was chosen as the method that best suited our needs. Almost all the data structures used in XGobi are available for modification using this approach.

The structure of the calling program is based on the subroutine template code provided with the XGobi source code. (The subroutine approach also has been used by Littman et al., 1992, for the implementation of the XGvis software system.) A control panel is initiated for each instance of XGobi (see Figure 1), from which the user has the option of selecting data from an ARC/INFO data base file or to receive the data set that is currently selected within ARC/INFO. Once the data source has been determined and the data received, the XGobi window is initialized.

Internally, an additional working procedure, namely a routine that runs once whenever the X Window System event loop finds no events, has been added to XGobi to check for incoming data from the ARC/XGobi server. If this routine receives updates of the data, the attribute values currently visible in XGobi linked to the brushed or subsetted coordinates in ARC/INFO will instantaneously be set to the same glyph, size, and color.

Otherwise, the entire functionality of XGobi has been maintained. The XGobi part can be described via the following pseudo code.

```
init sockets
connect to ARC/XGobi server
init XGobi defaults
init startup window
repeat
  wait for user input
  send input to ARC/XGobi server
  wait for data from ARC/XGobi server
  if (XGobi not invoked)
     then invoke XGobi
     else update XGobi structures and data sets
until (user input = QUIT)
```

3.4 Usage

ARC/INFO and the ARC/XGobi (parent) interface process must be activated on the host where the ARC/INFO data base is located. Then, XGobi client processes can connect to the ARC/XGobi server. Clients can reside on the same host or anywhere else on the Internet. Internet addresses, ports, and communication protocols are encoded into the program. So, the user does not have to worry about common setups for server and clients. If the user only wants to use XGobi to analyze the attribute data in an ARC/INFO data set, the invocation of ARC/INFO is not required.

™*DEC* is a trademark of Digital Equipment Corporation.
™*Sun* is a trademark of Sun Microsystems, Inc.
™*Sparc* is a trademark of Sun Microsystems, Inc.

4 An Example

As an example of how the link between ARC/INFO and XGobi can be used to explore data we show a data set containing water–quality data collected during several weeks at seventeen surface–water sampling sites in Erath County, Texas (see Figure 1). The pollutants are being modelled inter alia through explanatory variables, such as the number of dairies per acre or the number of head of cattle per acre, to account for large–scale variability.

As well as the sampling sites (numbered from 1 to 24 with some numbers missing), the ARC/INFO mapview shows streams (continuous lines), boundaries of large basins (dashed lines), dairies (triangles), and a town (shaded area). After an initial examination of the map, two of the sampling sites, numbered 4 and 12, have been brushed in order to see if the data collected there is anomalous.

Site 4 has been brushed because it is at the outlet of a very small basin containing four dairies. Thus, the response variables (i. e., pollutants) might be expected to be unusually high. The XGobi view on the left shows that the explanatory variable "nda" (the number of dairies per acre) is extremely large relative to the other sites. The XGobi view on the right shows that the responses "no3" (standardized nitrate) and "nh3" (standardized ammonia) at this site are high, though not outlying.

Site 12 has been brushed because it is located just below the town of Stephenville, Texas; the waste water treatment plant of Stephenville discharges into the stream above the sampling site. The XGobi view on the right shows that standardized nitrate is consistently high at this site, but nothing remarkable can be said about standardized ammonia. Based on this result, an analyst doing an exploratory data analysis probably would be interested in how strong the nitrate concentration is further downstream of Stephenville and, therefore, the next site to be brushed in the ARC/INFO view might be site 24.

This example demonstrates briefly how geography can give an analyst insight into the exploration of a data set and, thus, why a link between ARC/INFO and XGobi is a useful tool.

5 Future Work

We have presented an interface linking ARC/INFO with XGobi. This interface allows the user to link ARC/INFO and XGobi views interactively such that modifications of the ARC/INFO view automatically change the other views in the different XGobi clients. So far, this link is only unidirectional. Future work will focus on the other direction of the link, that is, the update of the ARC/INFO view according to one (or several) XGobi view(s). However, this direction is more problematic since substantial questions, such as security (Who is allowed to modify the data?), concurrency (What if two XGobi clients send different update information at the same time?), and technical issues (How can events be incorporated into a primarily non–event–driven FORTRAN program?) have to be resolved. We are also looking towards a new release of ArcView™ . According to preliminary announcements from the distributors of this software, this new version seems to be more suitable to facilitate the inverse link from XGobi to ARC/INFO.

Acknowledgements

Symanzik's research was partially supported by a German "DAAD–Doktorandenstipendium aus Mitteln des zweiten Hochschulsonderprogramms". Cressie's research was supported by the National Science Foundation, the National Security Agency, and the Office of Naval Research. We thank Deborah Swayne for her valuable assistance in integrating XGobi into our application.

References

Anselin, L. and Getis, A. (1992). Spatial Statistical Analysis and Geographic Information Systems. *Annals of Regional Science*, 26:19–33.

Cook, D., Cressie, N., Majure, J., and Symanzik, J. (1994). Some Dynamic Graphics for Spatial Data (with Multiple Attributes) in a GIS. In *COMPSTAT '94–Proceedings (To appear)*.

Ding, Y. and Fotheringham, A. S. (1992). The Integration of Spatial Analysis and GIS. *Computers, Environment and Urban Systems*, 16:3–19.

Fotheringham, A. S. and Rogerson, P. (1993). GIS and Spatial Analytical Problems. *International Journal of Geographical Information Systems*, 7:3–19.

Haslett, J., Bradley, R., Craig, P., Unwin, A., and Wills, G. (1991). Dynamic Graphics for Exploring Spatial Data with Application to Locating Global and Local Anomalies. *The American Statistician*, 45(3):234–242.

Littman, M., Swayne, D. F., Dean, N., and Buja, A. (1992). Visualizing the Embedding of Objects in Euclidean Space. *Computing Science and Statistics*, 24:208–217.

Openshaw, S. (1991). Developing Appropriate Spatial Analysis Methods for GIS. In Maguire, D. J., Goodchild, M. F., and Hind, D. W., editors, *Geographical Information Systems: Principles and Applications, vol. 1*, pages 389–402, London. Longman.

Stevens, W. R. (1990). *UNIX Network Programming*. Prentice–Hall, Englewood Cliffs, NJ.

Swayne, D. F., Cook, D., and Buja, A. (1991). XGobi: Interactive Dynamic Graphics in the X Window System with a Link to S. In *ASA Proceedings of the Section on Statistical Graphics*, pages 1–8, Alexandria, VA. American Statistical Association.