THE USE OF GIS TECHNOLOGY IN SOURCE WATER ASSESSMENT

Paul DiGirolamo[1] and Cindy Daniel[2]

AUTHORS: [1]GIS Intern and [2]Hydrologist and Senior Environmental Planner, Atlanta Regional Commission, 40 Courtland Street, 2nd Floor, Atlanta, GA 30303.
REFERENCE: Proceedings of the 2001 Georgia Water Resources Conference, held March 26-27, 2001, at the University of Georgia. Kathryn J. Hatcher, editor, Institute of Ecology, the University of Georgia, Athens, Georgia.

Abstract. Geographic Information Systems (GIS) technology has been an invaluable component of the Atlanta Regional Commission's efforts in the Source Water Assessment Project. In fact, GIS has been involved in nearly every stage of the project, including watershed delineation, mapping of potential pollutant sources, generation of working maps for field verification, and visual and spatial analysis. By utilizing the data integration, visualization, and analysis capabilities of GIS to complete the Metro Atlanta SWAP, a foundation is being created for future watershed protection efforts in the Atlanta region.

INTRODUCTION

The Environmental Planning Division of the Atlanta Regional Commission is currently working on the Metro Atlanta Source Water Assessment Project (SWAP). Extensive use of Geographic Information Systems technology has been an invaluable element of this effort. GIS work for this project was done using ESRI's ArcView v3.2 and ArcInfo v7.2.1 software.

BACKGROUND

Source Water Assessment Program

The Metro Atlanta Source Water Assessment Project is part of a larger, nation-wide inventory and risk assessment of potential threats to drinking water supplies. Following the 1996 Amendments to the Safe Drinking Water Act, a SWAP was required of each state in the U.S. by the Environmental Protection Agency (EPA). The Metro Atlanta SWAP is currently being conducted by ARC for twenty-eight watersheds in the metro region.

METHODS - GIS and SWAP

GIS Components of SWAP

There are four main components to the SWAP Program: 1) Delineation of Assessment Area; 2) Inventory of Potential Pollutant Sources and Contaminants of Concern; 3) Susceptibility Determination; and 4) Public Communication Plan. Application of GIS technology is an integral part of each of these components.

Delineation of Assessment Area

This component consists of several steps: 1) Delineation of water supply watersheds; 2) Determination of priority areas; and 3) Construction of base maps.

A source water watershed is the contributing land area above an intake from which water is withdrawn for a public water system. To delineate these areas, public surface water intakes were first spatially located, primarily from analysis of aerial photography (DOQQs), with some additional data provided by participating water utilities. The contributing water supply watersheds were then delineated to the intake using USGS topographical maps in combination with Georgia EPD Geological Survey Branch (GAGS) GIS coverages of 8- and 12-digit HUCs (hydrologic unit codes). These delineated watershed boundaries were digitized in ArcInfo and converted to polygon shapefiles. Attribute information for each boundary polygon was added including basin, source, intake managing entity, perimeter and area. A total of twenty-eight water supply watersheds were delineated for the Metro Atlanta SWAP.

Counties included in the delineated water supply watersheds consist of the ten counties in the Metro Atlanta region (Cherokee, Clayton, Cobb, Dekalb, Douglas, Fayette, Fulton, Gwinnett, Henry, and Rockdale) as well as fourteen peripheral counties partially contained in the watersheds (Bartow, Carroll, Coweta, Dawson, Fannin, Forsyth, Gilmer, Hall, Lumpkin, Paulding, Pickens, Spalding, Union, and Walton).

Based on the state's minimum guidance requirements, the watersheds were divided into management zones. The basis of management zones is distance from the intake. The Inner Management Zone (IMZ) is the watershed area within a 7-mile radius from the water supply intake. The Outer Management Zone (OMZ) is the watershed area located outside the Inner Management Zone but within a 20-mile radius of the water supply intake or the outer watershed boundary. Outside of this area is referred to as the Non-Management Zone (NMZ). The Inner and Outer Management Zones for each watershed were defined in ArcView and created as a series of polygon layers.

The final step of the delineation of the SWAP assessment area was the construction of base maps of the study area. These maps were created in ArcView using a combination of point, line, and polygon data themes including linear and polygonal hydrography, roads, railroads, and county and municipal boundaries. Data was obtained from the Georgia GIS Data Clearinghouse as well as ARC's Atlanta Region Information System (ARIS) CD. The water supply watershed boundaries and Inner and Outer Management Zones were then added to create a complete picture of the SWAP study area. Additional cartographic considerations were given at this point to visual hierarchy of base map elements and the appearance of legends and graphics.

Inventory of Potential Pollutant Sources and Contaminants of Concern

Use of GIS was involved in this component of SWAP in the following ways: 1) Collection and display of available data for Potential Pollutant Sources (PPS) for the entire region of study; 2) Determination of the spatial location of facilities through geocoding; 3) A query of all facilities to determine which lie within the boundaries of the watersheds; 4) Creation of Working Maps to be used in a field verification process; and 5) Revision of existing data to reflect field observations, thereby resulting in an accurate and current inventory of Potential Pollutant Sources within the study area.

ARC developed a list of Potential Pollutant Sources covering all of the states' minimum requirements as well as other facility types deemed important to the study. Data was obtained from many sources, including existing ARC datasets, EPA and GA EPD databases, GIS point coverages downloaded from the Georgia EPD website (www.ganet.org/dnr/environ/) and the Georgia GIS Data Clearinghouse (http://gis.state.ga.us), and hard copy records. Every attempt was made to find the most recent, best-documented, and most accurate data available. At this stage, a database template was created as a composite of relevant data fields from each dataset and each dataset was formatted to it. Due to limitations within ArcView's data management functions, much of the data management and manipulation was done in MS Excel and Access.

From a mapping standpoint, available facilities data was of two types: data that was obtained as an existing GIS point coverage with spatial reference information included, and data obtained in spreadsheet or other form with no geographic or spatial component other than a facility address. In order to map these latter datasets, it was necessary to 'geocode' them. Geocoding is an operation in which specially designed software processes facility address information, references it to a base map, plots a point on the map, determines the coordinates (latitude and longitude in unprojected decimal degrees) and returns the spatial information appended to the original spreadsheet. The dataset is then brought into ArcView as a table, and the spatial data is used to add it to a view as an 'Event Theme' and then create a shapefile from it. For this project ARC used a combination of ArcView's geocoding function and an external private geocoding service. The average success rate between the two programs was approximately 70%. In other words, of some 10,000 facilities geocoded 3200 were unable to be successfully located by either program. These unmatched facilities were set aside for later comparison to new facilities found in the field.

After spatially locating as many of the facilities within the study area as possible, it was then necessary to change the spatial reference point for many of the shapefile datasets in order to make them compatible with the basemap data. After "reprojecting" each dataset into a 'State Plane 83, Georgia West' projection, an ArcView script was used to extract the (X,Y) coordinates from each point and add these coordinates to the database attribute tables.

The next step was to determine which facilities fall within the watersheds of concern. Each facilities dataset was brought into ArcView as a point theme. Then the polygons representing the boundaries of the water supply watersheds of concern were used to 'clip' the points, resulting in a data layer containing only those facilities that are within the watersheds of concern and therefore a potential contamination source. These clipped themes were then combined into one comprehensive data layer, resulting in a single point theme containing all available spatial and attribute information for every potential contaminant source within the study area. In this way, it was possible to narrow the focus of the study from 10,300 facility points within the 24-county region to 3,234 facilities within the watersheds of concern.

Following the creation of the comprehensive dataset, it was necessary to verify the existence and correct location of mapped facilities and identify additional potential pollutant sources not included in the data sets used thus far. This was accomplished by thorough fieldwork performed by ARC staff and staff of the water supply utilities. Working maps were created for use as reference in the field. The facilities data for each watershed was overlaid onto the base maps, with each facility type given a unique legend symbol and each mapped facility labeled with a unique Map ID. Due to the large extent of many of the watersheds it was necessary to create a numbered index system so that each watershed could be broken up into manageable units that were mappable at a scale allowing easy identification and visual analysis of data points. A field form was generated in MS Access to accompany the working maps, consisting of a series of questions regarding each facility as well as space for additional comments. Forms for recording information on 'new' or unmapped facilities were also provided. Based on this field survey process, corrections to existing facility information have been made and additional facilities listed and plotted on the maps. Approximately 600 facilities were added to the database based on field observations, bringing the total number of facilities in the watersheds of concern to almost 4000.

In assessing the accuracy of the datasets and the success of the field verification, it should be noted that the quality of locational information provided in the datasets varied widely, and the effort was further complicated by inherent shortcomings in the geocoding process. For example, self-reported locations tended to be the least accurate, while GPS coordinates were consistently correct. However, as a general observation, it could be concluded that in older, stable areas the datasets were more accurate and there were not as many revisions to be made, while in newer, fast growing areas in the suburban fringe the datasets were not keeping up with changes and numerous revisions were necessary.

The end result of the inventory component of this project is the most current listing of Potential Pollutant Sources within the study area.

Susceptibility Determination

The main goal of this component of the SWAP (still in progress) is to rank each potential pollutant source and water supply watershed as high, medium or low priority on the basis of both risk and potential. This phase of the project combines the facilities inventory with other resources to develop an understanding of how likely it is that the water supply could be affected by these pollutant sources. The State criteria for susceptibility determination requires insight into a variety of physical characteristics of the sources. ArcView GIS with Spatial Analyst and Network Analyst extensions was used to determine such criteria as distance to water supply, distance to water intake, and average watershed slope. 1999 land cover data created by ARC from aerial imagery was used to calculate the percentage of impervious surface area for each watershed, and this data was used in a non-point source susceptibility analysis. A combination of these analyses will contribute to the overall watershed rankings.

Communication Plan

This component of the Metro Atlanta SWAP is still in the planning stage, and several options for sharing the results of this effort are being discussed. GIS technology is an integral part of two strong possibilities. One option being researched involves setting up a server at ARC and making the data and maps available online through the use of an internet-based GIS package such as ESRI's ArcIMS (Internet Map Server) program. This would give interested parties online access to the data through limited GIS functions such as query building, data display options, zoom in/out, pan/scan, etc. Another option being considered is the packaging of all relevant data and maps in project form on a CD-ROM for each separate watershed and then distributing these to the appropriate water management utilities. This would put all the data in the hands of the interested parties, and give them the option for more advanced data management and GIS use. Potential drawbacks of this approach would be the assumption of a greater familiarity with GIS by the utilities' staff, and preclusion of any further management or revision of existing data by ARC. These and other options are still being explored in an attempt to best address the informational needs of the utilities and the general public.

CONCLUSION

Application of GIS technology has played a major role in nearly every phase of the Metro Atlanta SWAP and has been an invaluable tool in the effort thus far. Utilization of the data management, visualization, and analysis capabilities of GIS is central to ARC's efforts to build a comprehensive database of Potential Pollutant Sources and determine susceptibility rankings for the watersheds of concern. There is a wealth of data available for Potential Pollutant Sources that may impact drinking water quality; however, such data often varies widely in format, scale, and quality. The ability to integrate data in many different formats and scales is a significant strength of GIS applications and has made it possible for ARC to construct a comprehensive database from diverse data sources. In addition, the ability to quickly visualize and analyze data based on type, location, ownership, proximity to stream, and other attributes has resulted in tremendous time savings and increased accuracy of the resulting inventory. Indeed, one could go so far as to say that a project of this scope would be virtually impossible to accomplish without access to GIS technology, and this is particularly true in an urban area where the number and density of potential threats is enormous. One result to date of the Metro Atlanta SWAP is a database of approximately 11,000 Potential Pollutant Sources in the 24-county study area, nearly 4000 of which are within the boundaries of watersheds feeding public water supply intakes. This database provides a crucial first step in future watershed protection efforts for this rapidly growing metropolitan area, establishing a baseline of potential pollution threats and determining criteria for susceptibility rankings. Communicating this information to water utilities and affected communities will provide a foundation of knowledge on which to build an effective and focused protection plan. It is hoped that this paper will provide a useful approach to similar projects in other areas.

LITERATURE CITED

State of Georgia Environmental Planning Division, Department of Natural Resources, Water Resources Group, 2000. EPD Source Water Assessment and Protection Plan.

Environmental Systems Research Institute, Inc. (ESRI), 1996. Using ArcView GIS.

Atlanta Regional Commission, 2000. Atlanta Region Information System (ARIS) Data CD

Georgia GIS Data Clearinghouse, http://www.gis.state.ga.us/

Georgia DNR Environmental Protection Division, Technical Guidance Documents: http://www.ganet.org/dnr/environ/