Improving Regional Water Quality Assessment

Geodatabase improves data management and analysis capabilities

By Jon Becker, U.S. Environmental Protection Agency Region 4

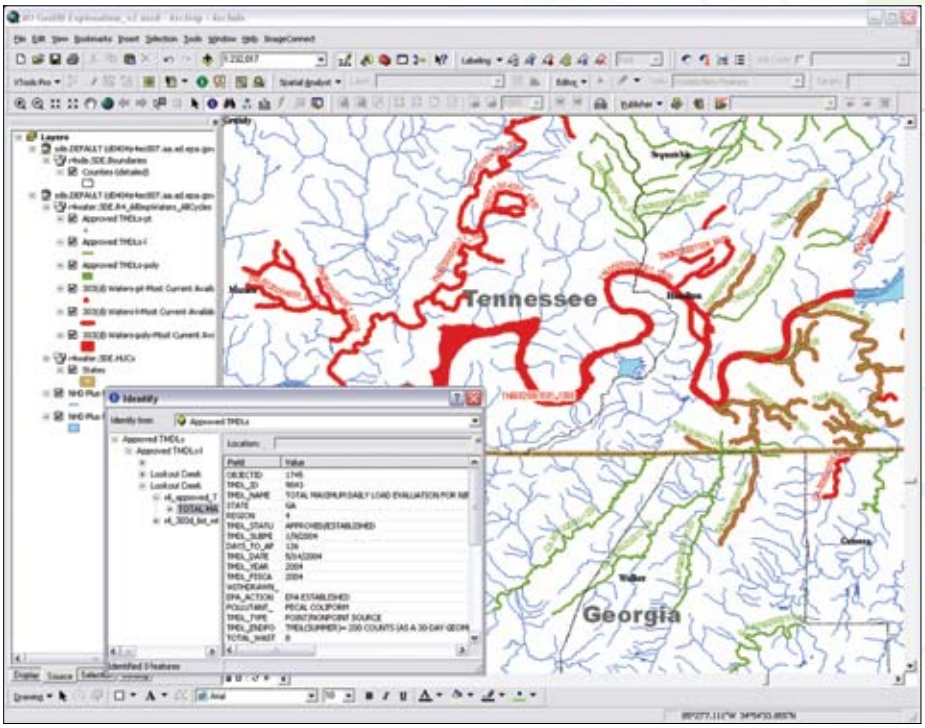
A geodatabase developed by the U.S. Environmental Protection Agency (EPA) Region 4 Water Management Division (WMD) manages georeferenced water quality assessment data from the region's eight southeastern states for various years far more efficiently than the previous system.

In accordance with the Clean Water Act, each state environmental agency tracks the status of water quality for water bodies located within its boundaries. Every even-numbered year, an assessment is made of each water body to determine whether it is meeting its designated uses. Each state agency tracks the status assessments in a local database and generates GIS datasets that depict the locations of these water bodies. Each lake, estuary, or river reach is called an assessment unit (AU) and assigned a unique identifier. While the EPA suggests recommended formats for assessment data and complementary GIS data, states often deviate somewhat from these formats to meet local needs.

Eight southeastern states submit assessment data and GIS files to EPA Region 4. These files are then forwarded on to EPA national headquarters for eventual input into the national Assessment and TMDL Tracking and Implementation System (ATTAINS). *[The Total Maximum Daily Load (TMDL) program determines the safe level of loading for a pollutant.]* The GIS files are addressed by river reach to the 1:100,000-scale National Hydrography Dataset (NHD) and incorporated into EPA's Watershed Assessment, Tracking and Environmental Results System (WATERS). This system integrates various EPA water-related databases via reach-addresses of the NHD.

Over the years, the staff of EPA Region 4 have obtained numerous assessment databases and GIS datasets from its eight southeastern states that were furnished in assorted and incompatible shapefiles, tables, and Microsoft Access databases. Using these datasets to analyze water body status for more than one state or year was difficult and time consuming even for experienced GIS users.

In 2007, EPA Region 4 staff explored the possibility of using a geodatabase to better manage water quality assessment data. Although data is eventually reformatted by contractors at EPA headquarters into consistent formats for inclusion in ATTAINS and WATERS, the obvious roadblock in creating a geodatabase for Region 4 was the variety of assessment data formats and GIS data used by the states.



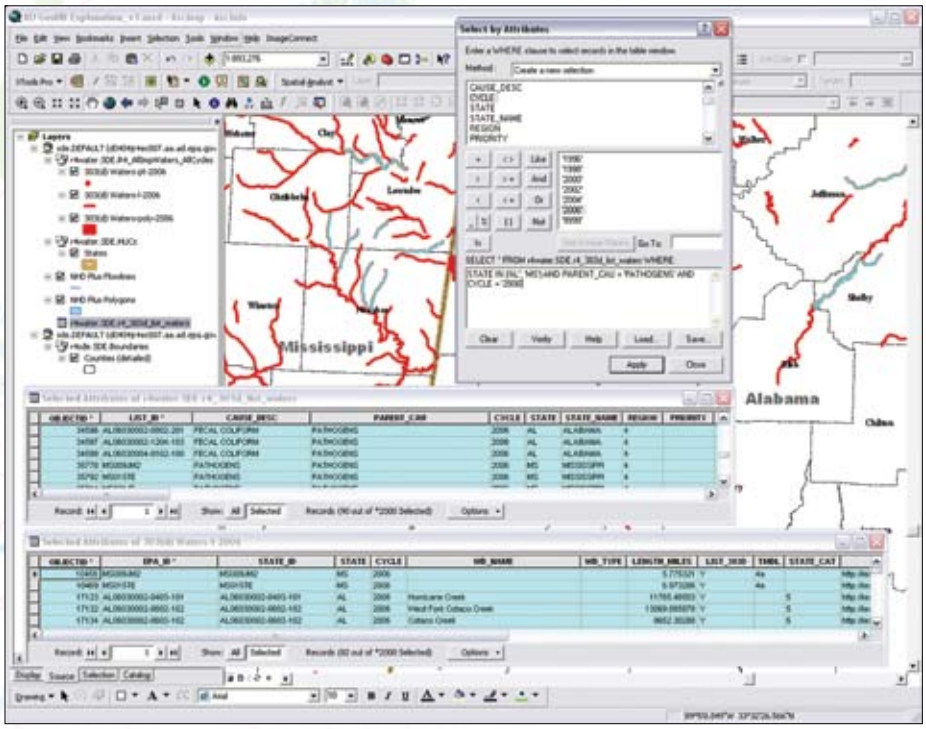
Users can use the Identify tool to click on a feature and view records about those features stored in related tables or activate the hyperlink to the EPA database record for the feature.

Region 4 contracted with Research Triangle Institute to develop a prototype geodatabase and populate it with reach-addressed versions of the states' assessment datasets. Region 4 staff then generated tables detailing the status of each AU from ATTAINS (then called the National Total Maximum Daily Load Tracking System). Relationship classes were created to relate the point, line, and polygon AU features to the assessment status tables using the AU identifier field. This was a one-to-many relationship because the AU is sometimes listed multiple times in the assessment status table if it has been tracked for several cycles and/or multiple water quality standard impairments (e.g., failures to meet swimming designated use because of high levels of pathogens).

Although this geodatabase format functioned fine, Region 4 staff soon realized it might be advantageous to populate it with the original state of GIS shapefiles rather than reach-addressed versions. While many states use NHD as the basis for their assessment units, some have reasons for not doing this.

For example, because Florida has mostly flat topography, and complex, highly modified hydrography is not well depicted at the NHD 1:100,000 scale, the Florida Department of Environmental Protection prefers to use its own water body identifiers (WBIDs) that depict small contributing drainage areas. If Region 4 staff are reviewing assessment data for Florida in a GIS, it is helpful to be able to view and use WBIDs instead of a reach-addressed approximation of those features.

However, Region 4 recognizes the value of having all state assessment GIS data in a consistent NHD-based format and continues



Complex queries may be made in the detailed tables, then the relationships to the point, line, and polygon feature classes may be activated to display the locations of the results.

encouraging states to adopt the NHD format as the framework for AUs. Some states, such as Florida, are exploring using a higher-resolution version of NHD to meet this need.

Because the region wanted to populate the geodatabase with GIS features from the states, the tables from each state GIS dataset had to be reformatted to a consistent format. The essential component for each AU record was an AU identifier because it was used to relate back to the ATTAINS tables. Features were also dissolved by the AU identifier so that there was only one feature per AU. Water body name and type fields for the AU were populated if the information was readily available in the state GIS dataset. A hyperlink field provides a link to the EPA Web page describing that water body. Length in miles was calculated for linear features and area in acres for polygon features.

The assessment data feature classes, tables, and relationship classes have been loaded into the ArcSDE server for Region 4. This provides a stable, fast-drawing platform. Having all state data and cycles in one feature dataset allows staff at the region to build various layer files pointing to these feature classes, which are just different definition queries. A layer depicting locations of approved TMDLs was also developed. Using the Identify button in ArcMap or ArcReader, a user can click on an impaired water body and see its impairments for all cycles and see the TMDLs that have been established for it. The built-in relationships give users the ability to query 303(d) listings or approved TMDLs by certain parameters, such as all pathogen-impaired waters, and activate the relationship classes to display those features in the GIS. *[A 303(d) listing identifies those lakes, wetlands, streams, rivers, and portions of rivers that do not meet all water quality standards for that state.]*

The geodatabase relationships between various assessment unit features and detailed tables provide WMD GIS staff with powerful and versatile analytic capabilities. By integrating disparate data formats from different states into a common structure, the geodatabase makes multiyear, multiple-state spatial analysis much easier than before. Answers to complex management requests or environmental progress reports can be generated in far less time than when georeferencing was contained in unmatched shapefiles.

Now new GIS data files are modified to fit the table format and loaded into the appropriate feature class. The tables from ATTAINS with the water body assessment status are periodically updated. Serving the data via ArcSDE also gives the region the ability to include these layers in ArcReader projects or other applications.

Although the Region has made great progress in managing its assessment GIS data, there is still much work to be done verifying the quality of the data and populating the ATTAINS database with assessment decisions. New organizational restructuring and procedures should help address these issues.

Reformatting the historical GIS files into this new consistent tabular framework involved a lot of work, but the resulting product has proved well worth the effort. GIS users can perform complex queries in related tables and activate the related features or simply use the Identify tool, click on AU features, and see the related tabular information. Relationship classes linking the ATTAINS tables of impaired waters and approved TMDLs have created a much more powerful and versatile analytic dataset that can generate useful information in a fraction of the time previously required.

For more information, contact Jon Becker of U.S. EPA Region 4, Watershed Management Division, at becker.jon@epa.gov.

About the Author

Jon Becker is an environmental protection specialist in the Water Quality Analysis Branch of U.S. EPA Region 4 in Atlanta, Georgia. He obtained a master's degree in applied geography in environmental and resource studies from Texas State University.

See the ESRI instructor-led course Data Management in the Multiuser Geodatabase *for more information on this topic.*