

Sound Data Management as a Foundation for Natural Resources Management and Science

Introduction—A Business and Scientific Case for Data Management

Effective decision making is closely related to the quality and completeness of available data and information (Burley and Peine, 2009). Data management helps to ensure data quality in any discipline and supports decision making. Managing data as a long-term scientific asset helps to ensure that data will be usable beyond the original intended application. Emerging issues in water-resources management and climate variability require the ability to analyze change in the conditions of natural resources over time. The availability of quality, well-managed, and documented data from the past and present helps support this requirement (fig. 1).

Sound data management encompasses the activities associated with planning, acquisition, dissemination, documentation, and ongoing maintenance of data. Included in these activities are the planning, implementation, documentation, and administration of the systems necessary for maintaining data in a usable and accessible format. Most if not all environmental, scientific, and natural resources-management organizations collect and create data in some form (fig. 2). Such data are valuable scientific assets resulting from dedicated resources (time, money) associated with their creation.

Electronic data collection and management, resulting in increasingly larger datasets and data-management costs, have grown more complex with the emergence of new formats and technology. The emergence of geographic information systems (GIS), along with geospatial data, has created additional complexity (Burley and Peine, 2010). The U.S. Department of the Interior estimates that about 80 percent of all government information has a geospatial data component, such as an address or latitude and longitude coordinates; in addition, up to 80 percent of GIS costs are related to the collection and management of geospatial data (U.S. General Accounting Office, 2003) (fig. 3). The considerable growth of data acquisition tools and geospatial technology has made data management a moving target. Front-end planning is necessary to minimize lost data utility resulting from data mismanagement.

What Can Databases Do for You?

Database technology provides a sound infrastructure for managing spatial and nonspatial data in support of data integration, cutting-edge science, and data and information

delivery. Databases provide considerable advantages over spreadsheets and other means of storing raw data, particularly for long-term data management. These advantages include the following:

- Facilitation of data integration from multiple sources
- Data queries and manipulation by using Structured Query Language (SQL)
- Data integrity support
- Scalable methods for data updating

In addition, the emergence of spatially enabled relational databases (geodatabases) provides additional functionality.

A geodatabase is a spatially enabled database that is capable of handling spatial and nonspatial data. A geodatabase provides a framework and an interactive tool to aid in the understanding of spatial and temporal trends in data. Also, geographically referenced data can be visualized and manipulated by using a GIS to produce maps and relational database queries and provide a foundation for various types of spatial analysis.

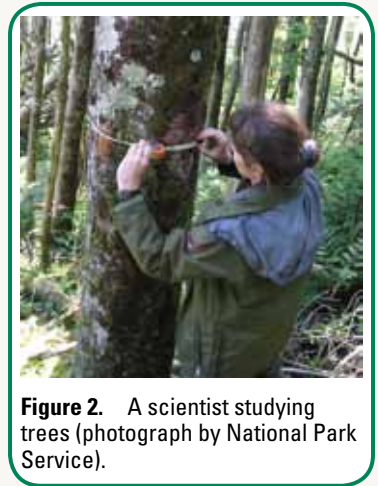


Figure 2. A scientist studying trees (photograph by National Park Service).

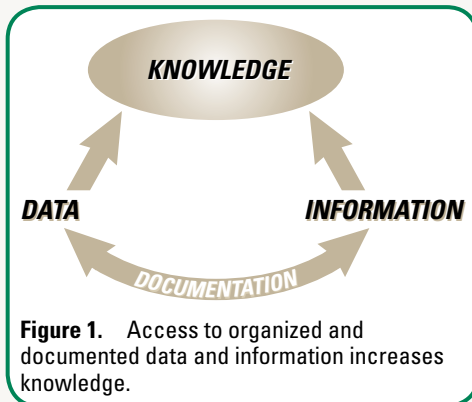


Figure 1. Access to organized and documented data and information increases knowledge.

Data Management Capabilities of the USGS Texas Water Science Center

The U.S. Geological Survey Texas Water Science Center (USGS TXWSC) supports the USGS mission by providing reliable, impartial, timely data and information that are needed to understand the Nation's resources. The staff of the TXWSC comprises geographers, hydrologists, geologists, and information technology specialists with data management and delivery capabilities and delivery method expertise.

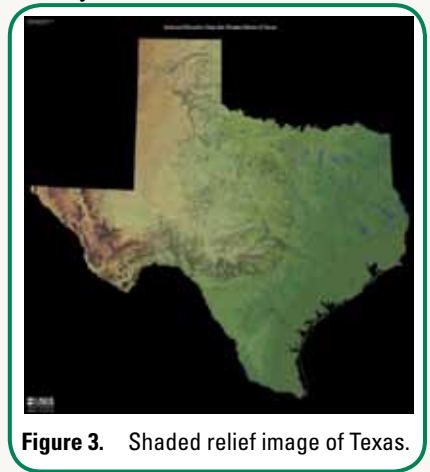


Figure 3. Shaded relief image of Texas.

Data and Information Management Capabilities:

- Data inventory and assessment/reconciliation
- Data integration

- Database and geodatabase design
- Database administration
- Programming/custom tools
- Quality assurance/quality control
- Virtual and/or on-site training
- Tutorial and standards development
- Standardized metadata (design and training)
- Web mapping/data visualization
- Web services
- Publication-quality cartography
- Database hosting
- Direct transfer to client

Recent Activities and Reports

Design and Compilation of a Water Resources Geodatabase for the Rio Grande Basin—San Acacia, New Mexico, to Fort Quitman, Texas

The USGS TXWSC, in cooperation with the New Mexico Interstate Stream Commission, has built a geospatial water-resources database, or geodatabase, to meet the need for an integrated and comprehensive data compendium for the Rio Grande Basin. Well-organized and accessible data increase abilities to identify management priorities and project areas, develop and evaluate management strategies, and monitor and document water-quality improvement and degradation. The end products of this effort are a water-resources geodatabase that enables the visualization of primary sampling sites for surface discharges, groundwater altitudes, and water-quality and associated data for the study area as well as repeatable methods and tools for refreshing the geodatabase with 24 static and updatable data sources.

For more information: <http://pubs.usgs.gov/of/2010/1331/>

Historical Data Recovery for a Streamflow-Gaging Station with More Than 100 Years of Data

In addition to streamflow measurements, field notes, photographs, slides, and other physical records exist as historical records of how the landscape has changed over time. A data rescue and recovery project is underway to improve and increase accessibility to ancillary USGS water data stored in Federal archives. This project also addresses the risk of data and information loss from obsolete media such as deteriorating slides, punch cards, and floppy disks. More than 100 years of photographs, measurements, reports, charts, maps, and other data for one USGS streamflow-gaging station are now stored in a database and available online. The methods developed and lessons learned during this effort have been documented and are now being applied to other USGS streamflow-gaging stations.

For more information: <http://txpub.usgs.gov/lrm/>

Texas Commission on Environmental Quality Watershed Characterization

The USGS TXWSC, in cooperation with the Texas Commission on Environmental Quality (TCEQ), has developed a library of geodatabases that contain basin delineations and characteristics for more than 3,000 water-quality monitoring stations in Texas. The TCEQ collects and analyzes water-quality

data at surface-water monitoring stations throughout Texas to ensure compliance with applicable State environmental standards. When the water quality at any given station does not meet State standards, a plan for remediation is developed. Such plans require detailed basin characterization information, including average rainfall, altitude, slope, land-use or land-cover statistics, and total drainage area. The databases will facilitate displays of geographic information and streamline data analysis, which will help the TCEQ make remediation, policy, and resource allotment decisions based upon consistently characterized watershed information.

For more information: <http://pubs.usgs.gov/of/2012/1077/>

Spatially Pooled Depth-Dependent Reservoir Storage, Elevation, and Water-Quality Data for Selected Reservoirs in Texas

The USGS TXWSC, in cooperation with Texas Tech University, constructed a selected storage, elevation in feet above datum (hereinafter referred to as reservoir elevation), and water-quality dataset for selected reservoirs in Texas. Data were acquired from existing databases, spreadsheets, delimited text files, and hard copy reports. Primary data sources include the USGS National Water Information System, the TCEQ Surface Water-Quality Information Management System (SWQIMS), and the Texas Water Development Board (TWDB) monthly reports on reservoir conditions. Lastly, USGS Texas Annual Water Data Reports, containing data that do not currently (2012) exist within the NWIS database, were used. The data product for this report is a collection of delimited text files of source-aggregated, spatially pooled, depth-dependent, water-quality data as well as instantaneous, daily, and monthly storage and elevation reservoir data.

For more information: <http://pubs.usgs.gov/ds/594/>

References

- Burley, T.E., and Peine, J.D., 2009, NBII-SAIN data management toolkit: U.S. Geological Survey Open-File Report 2009–1170, 96 p. (Also available at <http://pubs.usgs.gov/of/2009/1170/>.)
- Burley, T.E., and Peine, J.D., 2010, The future of geospatial data management—A natural resource perspective: *GeoWorld*, v. 23, no. 7, p. 20–23. (Also available at <http://www.geoplace.com/>.)
- Ferderer, David, Skinner, Chris, and Gunther, Gregory, 2001, Central Energy Resources Team data management—A data management life-cycle: U.S. Geological Survey Fact Sheet 2000–0163. (Also available at <http://pubs.usgs.gov/fs/2000/fs163-00/>.)
- U.S. General Accounting Office, 2003, Geographic information systems—Challenges to effective data sharing, statement of Linda D Koontz, Director, Information Management Issues: GAO-03-874T, 20 p. (Also available at <http://www.gao.gov/new.items/d03874t.pdf>.)

—*Thomas E. Burley*

For additional information, contact

Director, USGS Texas Water Science Center
<http://tx.usgs.gov/>
gs-w-txpublic-info@usgs.gov

Publishing support provided by
 Lafayette Publishing Service Center