U.S. Department of the Interior
U.S. Geological Survey

Digital Mapping Techniques '05—Workshop Proceedings

Edited by David R. Soller

April 24-27, 2005
Baton Rouge, Louisiana

Convened by the
Association of American State Geologists
and the
United States Geological Survey

Hosted by the
Louisiana Geological Survey

U.S. GEOLOGICAL SURVEY OPEN-FILE REPORT 2005-1428
2005

CONTENTS

Introduction

By David R. Soller

U.S. Geological Survey
926-A National Center
Reston, VA 20192
Telephone: (703) 648-6907
Fax: (703) 648-6977
e-mail: drsoller@usgs.gov

The Digital Mapping Techniques '05 (DMT'05) workshop was attended by more than 100 technical experts from 47 agencies, universities, and private companies, including representatives from 25 state geological surveys (see Appendix A). This workshop was similar in nature to the previous eight meetings, held in Lawrence, Kansas (Soller, 1997), in Champaign, Illinois (Soller, 1998), in Madison, Wisconsin (Soller, 1999), in Lexington, Kentucky (Soller, 2000), in Tuscaloosa, Alabama (Soller, 2001), in Salt Lake City, Utah (Soller, 2002), in Millersville, Pennsylvania (Soller, 2003), and in Portland, Oregon (Soller, 2004). This year's meeting was hosted by the Louisiana Geological Survey, from April 24-27, 2005, on the Louisiana State University campus in Baton Rouge, Louisiana. As in the previous meetings, the objective was to foster informal discussion and exchange of technical information. It is with great pleasure I note that the objective was successfully met, as attendees continued to share and exchange knowledge and information, and to renew friendships and collegial work begun at past DMT workshops.

Each DMT workshop has been coordinated by the Association of American State Geologists (AASG) and U.S. Geological Survey (USGS) Data Capture Working Group, which was formed in August 1996, to support the AASG and the USGS in their effort to build a National Geologic Map Database (see Soller and Berg, this volume, and http://ngmdb.usgs.gov/info/standards/datacapt/). The Working Group was formed because increased production efficiencies, standardization, and quality of digital map products were needed for the database—and for the State and Federal geological surveys—to provide more high-quality digital maps to the public.

At the 2005 meeting, oral and poster presentations and special discussion sessions emphasized: 1) methods for creating and publishing map products (here, "publishing" includes Web-based release); 2) field data capture software and techniques, including the use of LIDAR; 3) digital cartographic techniques; 4) migration of digital maps into ArcGIS Geodatabase format; 5) analytical GIS techniques; 6) continued development of the National Geologic Map Database; and 7) progress toward building and implementing a standard geologic map data model and standard science language for the U.S. and for North America.

ACKNOWLEDGMENTS

I thank the Louisiana Geological Survey (LAGS) and their Director and State Geologist, Chacko John, for hosting this meeting and for arranging for corporate sponsorship. In the tradition of past DMT meetings, the attendees were given a very informative, productive, and enjoyable experience. I especially thank Robert Paulsell (LAGS), who coordinated the events; Robert provided excellent support for the attendees, designing the website, arranged for corporate sponsorship, and, in the Louisiana tradition, organized the social activities (for example, a crawfish boil). Thanks also to Jeanne Johnson for managing the registration; Reed Bourgeois, John Johnston III, Rick McCulloh, Riley Milner, and Lisa Pond for all their help with the meeting's logistics; and a special thanks to John Snead, Cherri Cowen, and Ethan Killet for designing and providing to each attendee a bottle of DMT'05 Digital Ya-Ya hot sauce. I also thank Louisiana State University for providing an excellent venue and support for our meeting. Regarding the effects of Hurricane Katrina upon this fine State, later that year, I extend my deepest sympathies and hopes for a full recovery.

The meeting was greatly improved through the generous donations of the Baton Rouge Geological Society, the Louisiana chapter of the American Association of Petroleum Geologists, the Louisiana Oil Spill Coordinators Office (LOSCO), and Navigation Electronics Inc., of Lafayette, Louisiana.

I also, with gratitude, acknowledge Tom Berg (Chair, AASG Digital Geologic Mapping Committee) for his friendship and his help in conducting the meeting, and for his continued support of AASG/USGS efforts to collaborate on the National Geologic Map Database. Thanks of course also are extended to the members of the Data Capture Working Group (Warren Anderson, Ken-

tucky Geological Survey; Rick Berquist and Elizabeth Campbell, Virginia Division of Mines and Geology; Rob Krumm and Barb Stiff, Illinois State Geological Survey; Scott McColloch, West Virginia Geological and Economic Survey; Gina Ross, Kansas Geological Survey; George Saucedo, California Geological Survey; and Tom Whitfield, Pennsylvania Geological Survey) for advice in planning the workshop's content.

I warmly thank Lisa Van Doren (Ohio Geological Survey) for typesetting the Proceedings. Numerous software and hardware vendors attended the meeting and made significant contributions, and they are acknowledged below. I also thank Sheena Beaverson (Illinois State Geological Survey) for moderating the discussion session on large-format plotters. Finally, I thank all attendees for their participation; their enthusiasm and expertise were the primary reasons for the meeting's success.

PRESENTATIONS

The workshop included 29 oral presentations. Most are supported by a short paper contained in these Proceedings. The papers describe technical and procedural approaches that currently meet some or all needs for digital mapping at the respective agency. There is not, of course, a single "solution" or approach to digital mapping that will work for each agency or for each program or group within an agency; personnel and funding levels, and the schedule, data format, and manner in which we must deliver our information to the public require that each agency design their own approach. However, the value of this workshop and other forums like it is through their roles in helping to design or refine these agency-specific approaches to digital mapping, and to find applicable approaches used by other agencies. In other words, communication helps us to avoid "reinventing the wheel."

Several vendors participated in the workshop, by giving presentations and answering many questions from attendees. Their presence was greatly appreciated by all. Presentations included:

1. Technical discussion of ESRI products for creating, managing, and serving geoscience map information, by Brig Bowles and Veronica Schindler, ESRI,;
2. Technical discussion of Adobe products for creating geoscience map information, by Mike Bennett and Lynn Grillo, Adobe Systems, Inc.;
3. Technical discussion of Avenza products for creating geoscience map information, by David Andrec and Doug Smith, Avenza Systems, Inc.;
4. Discussion of LIDAR technology, by Kevin Lim, Optimal Geomatics, Inc. (formerly Atlantic Tech.);
5. "Building 3D geological models directly from the data? A new approach applied to Broken Hill, Australia" by Philip McInerney, Intrepid Geophysics, Australia (see paper in these Proceedings);
6. "Digital Mapping at Noranda-Falconbridge Exploration" by Pierre St-Antoine, Noranda-Falconbridge Exploration.

POSTERS AND COMPUTER DEMOS

More than 20 posters were exhibited and several computer demonstrations were provided throughout the workshop. These provided an excellent focus for technical discussions and support for oral presentations. Many are documented with a paper in these Proceedings, following those for the oral presentations; the other posters generally provided material in support of oral presentations, and so are not documented here.

DISCUSSION SESSIONS

ESRI Geodatabase

Most geological surveys use ESRI GIS products, and are in the process of migrating files and techniques from the ArcInfo Coverage and/or the ArcView Shapefile format to the ArcGIS Geodatabase format. For the past two years, we have held a discussion session with ESRI personnel in order to obtain technical information and tips, and to convey our needs to them. To prepare for the session, DMT attendees submitted questions, which I compiled and forwarded to ESRI prior to the meeting. These questions were addressed during the discussion session, and served as the basis for additional discussion; this format seemed to work well, and will be used in future discussion sessions with ESRI.

Adobe / Avenza

Adobe Illustrator, Photoshop, InDesign, and other software are used by most geological surveys, to prepare maps for publication. The Avenza MaPublisher plug-in to Illustrator provides a useful means of managing and exporting georeferenced maps, which can then be converted to GIS format. Technical and sales representatives from Adobe and Avenza participated in a joint discussion session, which proved to be very informative. To prepare for the session, DMT attendees submitted questions, which I compiled and forwarded to Adobe and Avenza prior to the meeting. These questions were addressed during the discussion session, and served as the basis for additional discussion; this format seemed to work well, and will be used in future discussion sessions with Adobe and Avenza.

Lidar

To provide the opportunity to consider a topic in some detail, informal discussion sessions are held at the DMT workshops. This year there were two: 1) large-format plotters, and 2) digital cartographic techniques and how we can share information on this subject. Session 1 began with a presentation by Randy Heilbrunn (Hewlett-Packard) followed by extensive discussion that was moderated by Sheena Beaverson (Illinois State Geological Survey). The discussion session's outline is available at http://ngmdb.usgs.gov/Info/dmt/docs.html.

Large-format Printing at Geological Agencies

This session focused on issues related to the use of large-format plotters for publication of geoscience maps. Before the meeting, Sheena Beaverson (Illinois State Geological Survey) asked for attendee's input on the following topics, which were discussed in the session: What technical hurdles have you overcome in the past year? What large-format plotter brands and models do you use? Are you planning a major hardware purchase in the near future? Do you use onboard an RIP, or a separate software RIP? If separate, what is the software and do you like it? Are you having problems choosing the appropriate media for different purposes? Do you prefer standard or UV inks? What other issues will your agency be facing, with respect to large format plotting? A summary of the session, including responses to the questions, is available at http://ngmdb.usgs.gov/Info/dmt/docs/beaverson05.ppt.

THE NEXT DMT WORKSHOP

The tenth annual DMT meeting will be held June 11-14, 2006, on the campus of The Ohio State University, in Columbus, Ohio. Please consult the Web site (http://ngmdb.usgs.gov/Info/dmt/) for updated information. While planning for that event, the Data Capture Working Group will carefully consider recommendations for meeting content and format offered by DMT'05 attendees.

REFERENCES

Soller, D.R., editor, 2004, Digital Mapping Techniques '04—Workshop Proceedings: U.S. Geological Survey Open-File Report 2004-1451, 220 p., accessed at http://pubs.usgs.gov/of/2004/1451/.

Soller, D.R., editor, 2003, Digital Mapping Techniques '03—Workshop Proceedings: U.S. Geological Survey Open-File Report 03-471, 262 p., accessed at http://pubs.usgs.gov/of/2003/of03-471/.

Soller, D.R., editor, 2002, Digital Mapping Techniques '02—Workshop Proceedings: U.S. Geological Survey Open-File Report 02-370, 214 p., accessed at http://pubs.usgs.gov/of/2002/of02-370/.

Soller, D.R., editor, 2001, Digital Mapping Techniques '01—Workshop Proceedings: U.S. Geological Survey Open-File Report 01-223, 248 p., accessed at http://pubs.usgs.gov/of/2001/of01-223/.

Soller, D.R., editor, 2000, Digital Mapping Techniques '00—Workshop Proceedings: U.S. Geological Survey Open-File Report 00-325, 209 p., accessed at http://pubs.usgs.gov/of/of00-325/.

Soller, D.R., editor, 1999, Digital Mapping Techniques '99—Workshop Proceedings: U.S. Geological Survey Open-File Report 99-386, 216 p., accessed at http://pubs.usgs.gov/of/of99-386/front.html.

Soller, D.R., editor, 1998, Digital Mapping Techniques '98—Workshop Proceedings: U.S. Geological Survey Open-File Report 98-487, 134 p., accessed at http://pubs.usgs.gov/of/of98-487/.

Soller, D.R., editor, 1997, Proceedings of a workshop on digital mapping techniques: Methods for geologic map data capture, management, and publication: U.S. Geological Survey Open-File Report 97-269, 120 p., accessed at http://pubs.usgs.gov/of/of97-269/.

The Kentucky Geological Survey's Online Geologic Map and Information System

By Gerald A. Weisenfluh, Douglas C. Curl, and Matthew M. Crawford

Kentucky Geological Survey
228 Mining and Mineral Resources Bldg.
University of Kentucky
Lexington, KY 40506-0107
Telephone: (859) 257-5500
Fax: (859) 257-1147
e-mail: {jerryw, doug, mcrawford}@uky.edu

INTRODUCTION

With the completion of the digital conversion of all of its 1:24,000-scale geologic maps, the Kentucky Geological Survey is developing a Web site that will allow users to create a highly customized geologic map for any project area in the state and then view related information, including well information, geotechnical and hazards data, related publications, photographs, illustrations, and a variety of other descriptions about Kentucky geology. A prototype was released to the public in April 2005 to demonstrate its capabilities and to solicit public comment. Because Kentucky has been completely mapped at a scale of 1:24,000, a detailed geologic map can be made for any area without concern about mismatches along map borders, or missing information. The process of converting these maps to digital format included edgematching of adjoining maps to minimize discontinuities at the edges of quadrangles (Anderson and others, 1997). All of the map data will be stored in a seamless spatial database so that rendering of map units is uniform in all locations. The online geologic map system is integrated with another Web service—the KGSGeoPortal—to facilitate locating the user's area of interest and linking to other useful data sources.

DESIGN CONSIDERATIONS

The original objective of this initiative was simply to make data from existing geologic maps available in an online system. In the early stages of development though, it became clear that users desired much more. They wanted to see other related data in the context of the geologic maps. For example, land-use planners needed to view sinkhole locations or landslide potential in the context of the geologic base. Coal companies were interested in viewing mapped coal beds and related site measurements and sample locations. It was also evident that users of this online mapping system would have diverse geologic backgrounds and differ in how they use geologic maps. The system needed to be flexible for this diverse audience, and easy to use. In order to address these challenges, the service was implemented as a customized ArcIMS application using the ActiveX connector with ASP technology.

It was also thought that most users of the system would be looking for information about a specific geographic area—a project area, a property, or a state park, for example. The system should allow the user to quickly find the area of interest, then provide the necessary map information and the ability to print it or save it for future use. Finding specific geographic areas on regional maps provided through the internet has always been a challenge, and adding full-featured search capability to a service also adds complexity. Because of this, KGS developed a separate geographic search function, called the KGSGeoPortal, that can be used to link to a variety of Kentucky data sources, including the new geologic map service described in this paper.

THE KGSGEOPORTAL

The KGSGeoPortal (http://kgsmap.uky.edu/website/KGSGeoPortal/KGSGeoPortal.asp) has two functions. The first is to allow users to search for common Kentucky geographic names and then view a map that encompasses the extent of the chosen geographic feature. Although many geographic search functions use a central coordinate as a proxy for the location of geographic features, KGS developed a database of Kentucky geography that stores the minimum and maximum coordinate extents for features. Because each extent is a custom area that fully encompasses the geographic feature, the initial map zoom should represent the user's area of interest more accurately than centroid-based systems, saving the user further map zooms. A wide variety of standard geographic feature types are presently available (see pull down menu in Figure 1), and nonstandard areas, such as a research project

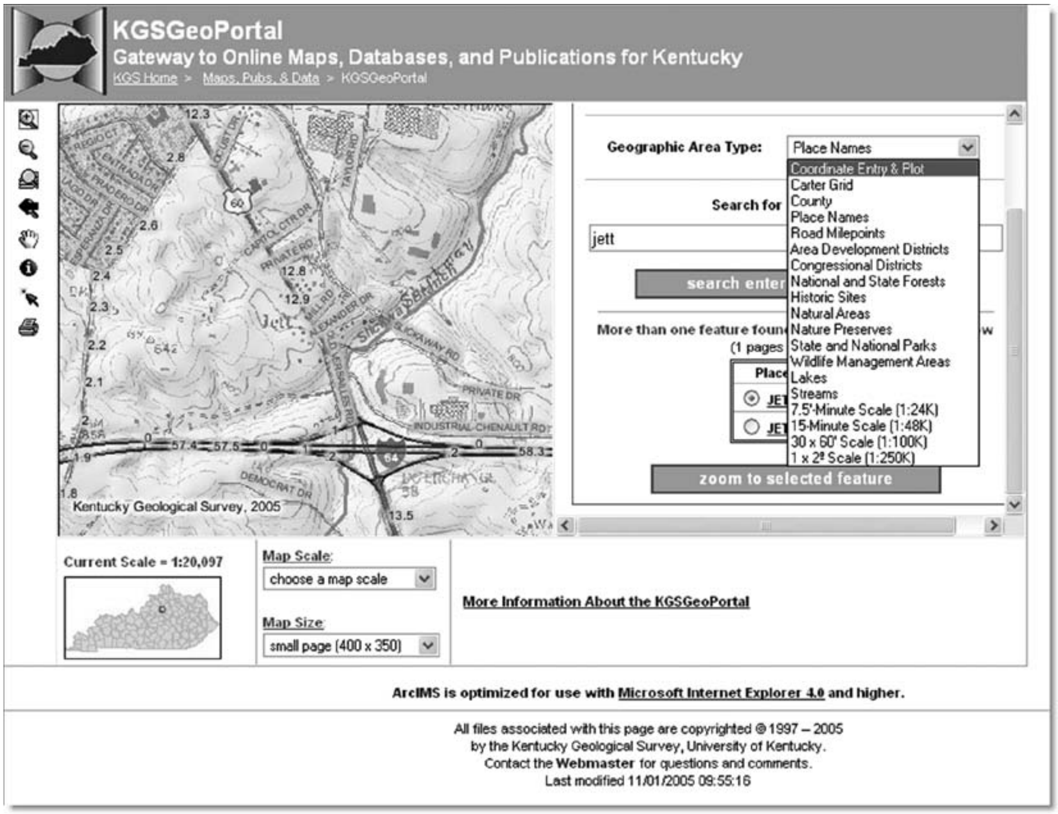


Figure 1. Upper part of KGSGeoPortal site showing map and geographic search types.

area, can be added. Once the initial extent is reached, the map can be adjusted with standard zoom and pan controls to refine the area of interest.

The second function of the portal is to link users to other Web data sources using the map's coordinate bounds as search criteria. Any Web service that accepts a URL with coordinates as arguments can be linked. The linked services open in another window, with the same coordinate extent as the portal. Tabular databases, such as wells or sample locations, can also be queried with sql-based language. A table of links to 30 maps and databases for Kentucky are presently provided on the KGSGeoPortal (Figure 2). In the same way that this service can link to other data sources, those sites can link back to the portal to take advantage of its features. For example, the U.S. Geological Survey's National Geologic Map Database's Map Catalog Product Description Pages for Kentucky maps contain a back-link to the portal so that users can view other data for the same area as that of the published map (e.g., http://ngmdb.usgs.gov/Prodesc/proddesc_52383.htm).

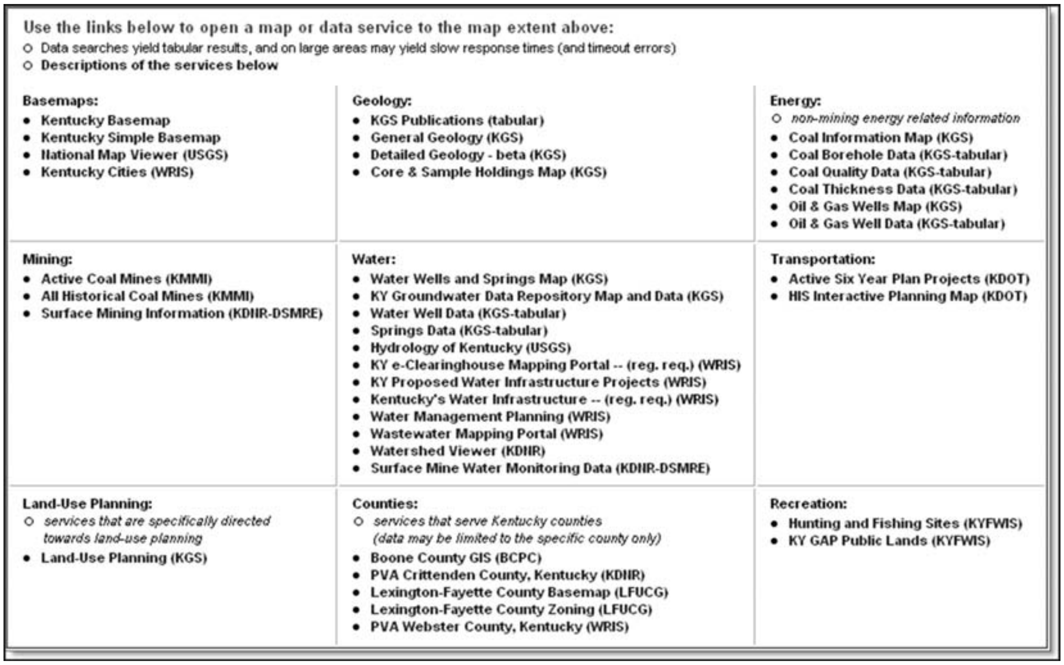
Although the KGSGeoPortal provides efficient geographic searching and the ability to compare a variety of data derived from the internet for the same area, it does not have the capability to overlay those data in a single map view. For this reason, the geologic map service was designed to allow that functionality, with an interface that is simple and intuitive to use.

GEOLOGIC MAP INTERFACE

The design of the geologic map service is shown in Figure 3. The Web page http://kgsmap.uky.edu/website/KGSGeology/viewer.asp is divided into three frames. The map frame with standard navigation tools is in the upper left. It is set to a fixed pixel dimension, but can be resized to fit other standard dimensions using the "Map Size" control in the lower frame. Setting the map frame to "full page," for example, results in map dimensions that print exactly on 8.5 by 11 inch paper. The map can also be set to an exact scale by selecting from the "Map Scale" pull down menu in the lower frame. The right hand frame serves three separate functions indicated by the links at the top of the frame: a map legend (the active function on Figure 2), layer control, and a geologic information page. Each function is selected by clicking its link at the top of the frame.

CUSTOMIZED MAPS AND DATA OVERLAYS

The "Map Layers" tab on the right frame (Figure 4) provides the function of adding or removing thematic and base layers from the view. Because geologic map information can be used for a wide variety of applications, many different styles of maps can be created.



Use the links below to open a map or data service to the map extent above:

- Data searches yield tabular results, and on large areas may yield slow response times (and timeout errors)
- Descriptions of the services below

Basemaps:
- Kentucky Basemap
- Kentucky Simple Basemap
- National Map Viewer (USGS)
- Kentucky Cities (WRIS)

Geology:
- KGS Publications (tabular)
- General Geology (KGS)
- Detailed Geology - beta (KGS)
- Core & Sample Holdings Map (KGS)

Energy:
- *non-mining energy related information*
- Coal Information Map (KGS)
- Coal Borehole Data (KGS-tabular)
- Coal Quality Data (KGS-tabular)
- Coal Thickness Data (KGS-tabular)
- Oil & Gas Wells Map (KGS)
- Oil & Gas Well Data (KGS-tabular)

Mining:
- Active Coal Mines (KMMI)
- All Historical Coal Mines (KMMI)
- Surface Mining Information (KDNR-DSMRE)

Water:
- Water Wells and Springs Map (KGS)
- KY Groundwater Data Repository Map and Data (KGS)
- Water Well Data (KGS-tabular)
- Springs Data (KGS-tabular)
- Hydrology of Kentucky (USGS)
- KY e-Clearinghouse Mapping Portal -- (reg. req.) (WRIS)
- KY Proposed Water Infrastructure Projects (WRIS)
- Kentucky's Water Infrastructure -- (reg. req.) (WRIS)
- Water Management Planning (WRIS)
- Wastewater Mapping Portal (WRIS)
- Watershed Viewer (KDNR)
- Surface Mine Water Monitoring Data (KDNR-DSMRE)

Transportation:
- Active Six Year Plan Projects (KDOT)
- HIS Interactive Planning Map (KDOT)

Land-Use Planning:
- *services that are specifically directed towards land-use planning*
- Land-Use Planning (KGS)

Counties:
- *services that serve Kentucky counties (data may be limited to the specific county only)*
- Boone County GIS (BCPC)
- PVA Crittenden County, Kentucky (KDNR)
- Lexington-Fayette County Basemap (LFUCG)
- Lexington-Fayette County Zoning (LFUCG)
- PVA Webster County, Kentucky (WRIS)

Recreation:
- Hunting and Fishing Sites (KYFWIS)
- KY GAP Public Lands (KYFWIS)

Figure 2. Lower part of KGSGeoPortal site showing available data sources for Kentucky.

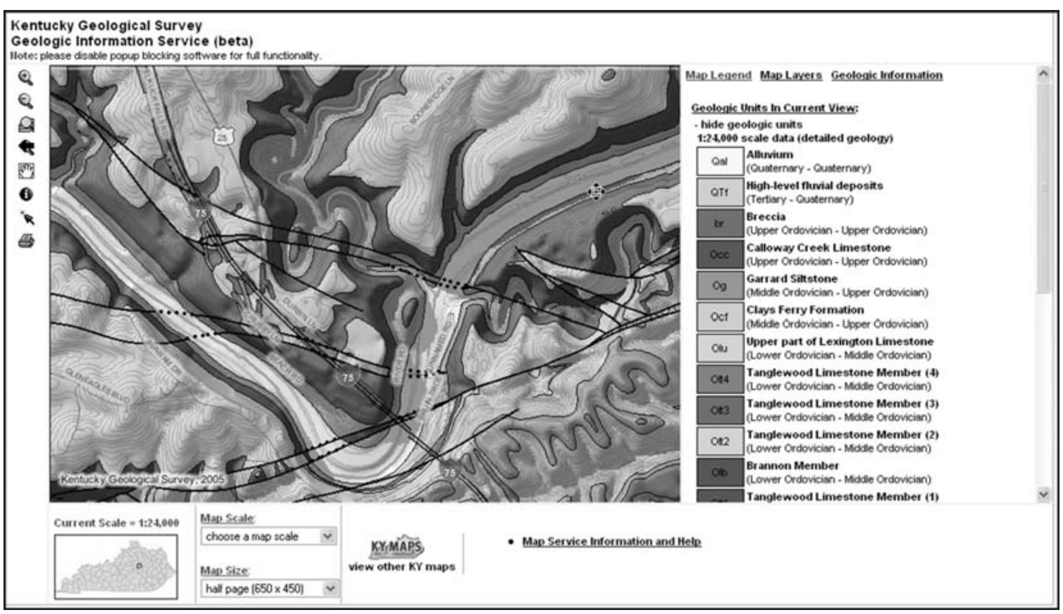


Figure 3. A map display from the KGS geologic map information Web site.

All available themes are found under the "Customize" heading. They are grouped according to function: base maps, geology, imagery, derivative classifications, etc. The standard base map is a hillshade topographic image. Selected, or "active" geologic unit themes can be draped over this base to simulate a three-dimensional effect and are in turn overlain by vector elevation contours (see Figure 3). Derivative classifications of the geologic units, such as karst potential or primary lithology, are available as an alternative to standard stratigraphic symbolization. This is accomplished by constructing tables that translate the geologic units to other symbols according to a set of rules determined by KGS geologists. Most themes have assigned scale dependency to prevent rendering them at inappropriate scales. Theme names that are not visible at a given scale are shown in gray type, but can still be selected for inclusion in the layout.

Many users will want to compare geologic units to

other kinds of map information. For example, sinkholes or quarry outlines can be superimposed on the map for analysis. Locations of a variety of point data, such as oil wells, water wells, coal measurements, sample locations, measured sections, and photographs, can also be added to the map. Data pertaining to those locations will be accessible through search tools. The site data included on this map service are maintained in a separate relational database, and these data may change on a daily basis. This has been a challenge for ArcIMS maps that require the points to be converted to spatial themes (shapefiles or SDE layers), because the service must be stopped while the theme is updated, in order to avoid corruption. The KGS geologic map service circumvents the problem by sending queries directly to the tabular databases to add graphic overlays of the point locations. The data shown on the geologic map are always current with respect to the tabular database. The disadvantage of this method is that the points cannot be queried directly to obtain attribute information. Custom query tools have been designed that simulate an identify tool. Rather than searching a spatial layer, the tool sends a coordinate-based query to the database for the attributes. This method turns out to be more efficient than querying spatial themes.

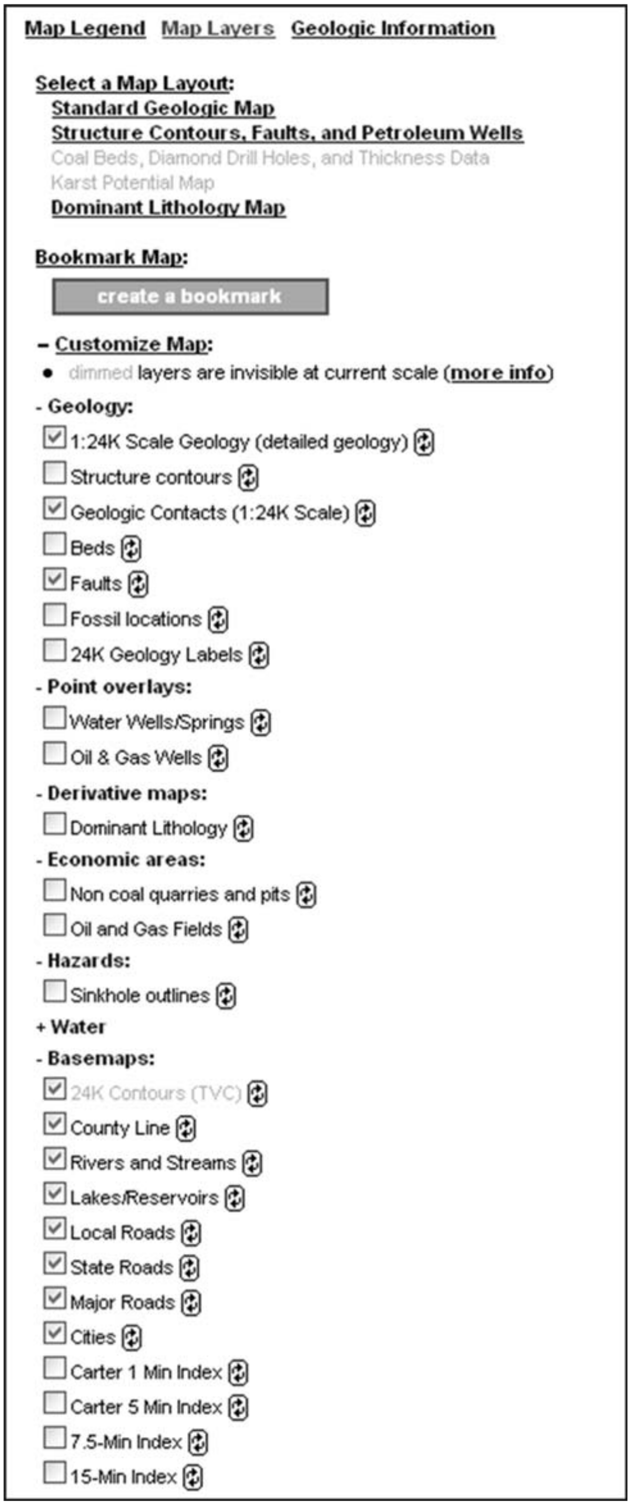


Figure 4. Map layers tab, with predefined and custom map layouts.

The map layers tab also provides a quick method of customizing the view—that is, standard layouts. These links, found at the top of the frame, represent commonly used, predesigned layouts that save users the time needed to browse through the “customize” theme list.

Once users have a map design that suits their purpose, the map can be bookmarked as a browser favorite for future viewing or for sending to a colleague. Two kinds of bookmarks are provided: (1) the map layout and its coordinate extent or (2) only the layout. Each unique bookmarked layout is stored in a KGS database and assigned an ID number. To retrieve a layout, the map’s base URL need only contain an additional variable with the layout ID, and, optionally, the bounding coordinates of the view. An example bookmark would be:

http://kgsmap.uky.edu/website/KGSGeology/viewer.asp?LayoutID=13&QueryZoom=Yes&startLeft=5272839.22544311&startRight=5292301.9485407&startTop=3941629.60683334&startBottom=3928155.41391963

GETTING INFORMATION

Identify Tool

Each published geologic quadrangle map includes a variety of descriptions for geologic units, as well as economic activities and other related subjects. These textual descriptions are being cataloged in a database to provide easy access with a map query tool (the black circle with white “i” on Figure 1). Users can click a location on any part of a map, and will receive a report of all available descriptions for that geologic unit, along with part of the stratigraphic column for context (Figure 5). Links are also provided to an online version of the original published

map and a separate image of the full stratigraphic column. The stratigraphic columns are especially important, because the digital conversion of the geologic maps resulted in some changes to nomenclature and some mapped units were combined on the digital map. As a result, stratigraphic names contained in our digital map database and accessed through the Web map may not match those of the original, printed map. The unit descriptions are cataloged according to the original nomenclature, and the information report shows the hierarchical relationships between map units and descriptions. Original unit descriptions have also been subdivided where formal or informal parts or facies were described, but not mapped. Because of the sometimes complex relationship between unit descriptions and their spatial representation, the database contains the unit name associated with the description as well as the names of parent units on the published map and digital compilation.

Most of the original geologic quadrangle maps included a section called "Economic Geology." These paragraphs actually contain a diverse collection of information about economic activity (at the time of mapping), engineering, paleontology, archeology, land use, hydrology, geophysics, and structural geology. These descriptions have been cataloged according to topic, and those that pertain to specific geologic units will appear in the "identify report" for the unit that was selected.

In the same Web page report, the identify tool also returns information about other visible thematic features in the view area. For example, if there is a quarry and

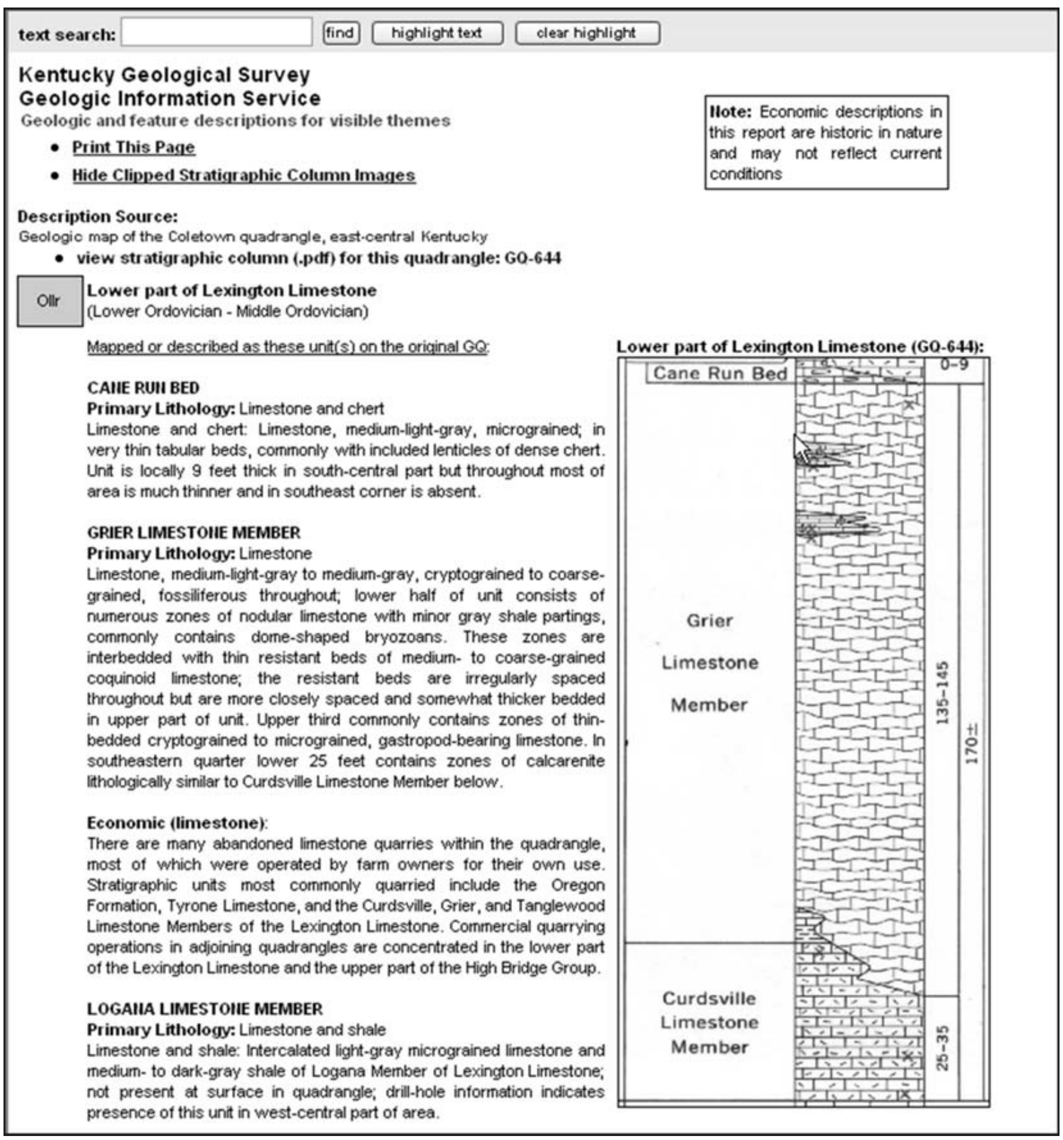


Figure 5. Lithology information results page, showing relationship between descriptions and map units.

measured section in close proximity to the selected location, summary information will appear for those features at the bottom of the report. Most features will also contain links to extended data, such as the interval descriptions for a measured section, electric logs for an oil well, or commodity information for a quarry.

Geologic Information Tool

Whereas the map query or “identity” tool provides descriptions for a single location, the geologic information tab in the right frame (see Figure 3) provides a more comprehensive search for all information pertaining to the viewable map area (Figure 6). The “geologic information” functions work by searching the KGS description database by map coordinate extent, rather than by map attributes. Two kinds of searches are performed simultaneously by this “extent” tool. Data represented by point locations (e.g., wells or photos) can be searched by their coordinate values. Other geologic descriptions and images are identified by their association with a published or unpublished source that has an assigned map extent overlapping the user’s view. The geologic information tool provides individual links to each kind of information rather than a single report. This is because of the potentially large amount of data that can be returned to a user in a single request.

Extent queries may return descriptions from multiple sources, if the user’s view includes more than one geologic map, or because there are descriptions in the database from other kinds of sources, such as published reports or unpublished observations. Consequently, the results are initially sorted by source type and scale, then individual sources. As an alternative, lithology reports can be sorted by stratigraphic unit so that users can compare descriptions on adjacent maps for a single unit. All reports also provide a text search function to locate occurrences of keywords of interest.

FUTURE DEVELOPMENT

The new KGS Geologic Map and Information site was released in beta version in April 2005. All of the functions described in this paper work for limited data sets in central, eastern, and western Kentucky. The next phase of development is to add the remainder of the spatial data to the ArcSDE database, including geologic themes, as well as additional related data, such as karst groundwater dye-trace results. A number of point feature types, including coal information, photographs, measured sections, and Natural Resource Conservation Service observation points, must be added to the layer list. A derivative classification for karst potential index has been developed and will be added in the near future. The most significant amount of future work will be preparing and loading text descriptions from the remaining 643 published geologic maps. This is expected to take an additional year.

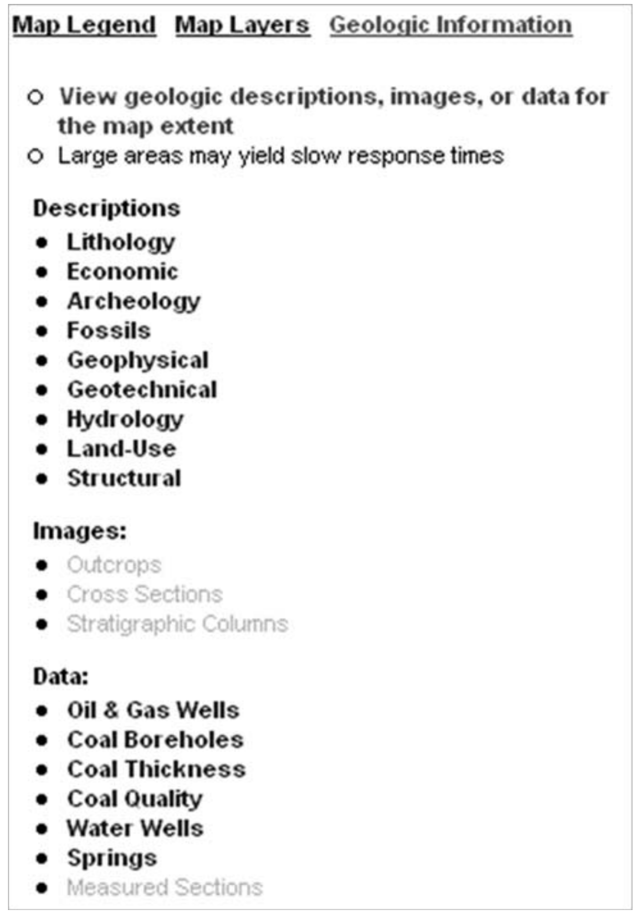


Figure 6. Subject and feature categories accessed from the geologic information tab. Grayed items not yet implemented.

Another related effort is a Web application that permits geologists to submit unpublished descriptions, observations, and images to the geologic information system. This system, currently under development, will allow approved users to catalog their knowledge about Kentucky rock units in the same database designed for published geologic maps. They will also be able to add and annotate photographs that they have taken of geologic features, and all these data will become available to users of the geologic map site.

REFERENCES

Anderson, W.H., Morris, L.G., and Sparks, T.N., 1997, Semi-automated data capture for vectorizing geological quadrangle maps in Kentucky, in Soller, D.R., ed., Digital Mapping Techniques 1997, Workshop Proceedings, U.S. Geological Survey Open File Report 97-269, p. 9-13, available at http://pubs.usgs.gov/of/of97-269/anderson.html.

Non-Survey, Non-Digital Completed Geologic Maps in File Drawers and Theses: How Can They Be Transformed into Useful Available Digital Geo-spatial Data?

By Robert D. Hatcher, Jr.

Department of Earth and Planetary Sciences
and Science Alliance Center of Excellence
University of Tennessee
306 Earth & Planetary Sciences Building
Knoxville, TN 37996-1410
Telephone: (865) 974–6565
Fax: (865) 974–9326
e–mail: bobmap@utk.edu

INTRODUCTION

Many reasons exist for making detailed geologic maps, including addressing basic research problems, mineral and hydrocarbon exploration and development, location of low value resources, and for foundations of engineered structures. Some use geologic maps at 1:24,000 or larger scale to combine small areas into islands of knowledge that permit addressing larger-scale problems. With reconnaissance mapping we frequently can extrapolate between islands of knowledge, in order to compile smaller-scale geologic maps. The important point, however, is that detailed geologic maps at 1:24,000 scale or larger form the basis for high quality geologic basic or applied uses.

Large numbers of detailed geologic maps the were not produced by State or Federal agencies are stored (not formally archived) in geological survey and university faculty file drawers, in theses and dissertations in university libraries, in engineering reports, mining and petroleum companies, and elsewhere. These are mostly not digital geologic maps. Numerous archived detailed geologic maps are of areas that now are inaccessible through urbanization or through concealment beneath various kinds of engineered structures, or are in flooded mines and abandoned oil fields. These maps constitute valuable data sets that should be preserved and made available to the geologic community. Conversion of these maps to digital geo-referenced GIS maps and databases is possible, with adequate time and funding. At the very least, they could be scanned and made available as georeferenced images. The map categories identified here should be subjected to a quality filter before the digitization process begins.

There are several categories of non-federal, non-state, non-digital geologic maps. Some are published at small scale in journals and survey publications. In addition to the maps listed above, many detailed geologic maps were made during the 1970s and 1980s when nuclear power plants were being constructed, and these reside in the archives of the Nuclear Regulatory Commission. They also should be digitized and made more widely available. Detailed geologic maps of dam, building, and non-nuclear power plant foundations and parts of other engineered structures should similarly be digitized.

The primary utility of digital geologic maps is virtually the same as that of paper geologic quadrangle maps. They contain the primary geometric, spatial, and resources data useful for crustal and surficial geologic research, and mining, petroleum, engineering, and environmental applications. In addition however, digital geologic maps provide the ability to quickly and easily add data, and to revise maps while they are being constructed. Editorial changes also are readily incorporated and, in addition, there is greater ease in integrating geologic maps into local or regional compilations if they are in digital format. Moreover, computer systems that allow geologists to record attribute and spatial (GIS) data in the field provide a more effective means of migrating field observations and mapping into a formal, published map database.

CONVERTING NON-DIGITAL TO DIGITAL GEOLOGIC MAPS

Our process of converting paper to digital geologic maps requires scanning the paper version and re-compiling it on a geo-referenced base in Adobe Illustrator™ or another graphics program. Adobe Illustrator is preferred because the add-on program MaPublisher™ permits geo-referencing maps at the beginning of digitizing. The

map can be printed as a draft and edited, then the editing incorporated into the digital file. The map explanation can most conveniently be assembled in the graphics program. The geo-referenced map can be brought into a GIS for addition of other attributes and data.

A digital raster graphic (DRG) or digital line graph (DLG) file of the base map is obtained and opened in Adobe PhotoShop™. The basic colors in the map file are converted to black or some other background color, and the PhotoShop document is saved. The base map file can then be opened in Adobe Illustrator and geo-registered using MaPublisher. An advantage to approaching geologic mapping this way is that it permits geologic contacts to be compiled daily, on-screen, in Adobe Illustrator layers, and lithologic, resources, and structural data to simultaneously be compiled onto the map and into a spread sheet. Once mapping is complete and contacts are drawn, polygons can be constructed to complete the geologic map. Finally, the title, explanation, scale, coordinates, etc., can be added to the margins of the map.

The digital geologic map can then be converted to GIS format by using MaPublisher Adobe Illustrator to create georegistered shape files of the geology and the base map. The shape file can then be converted to an ArcGIS file. The geologic attribute data for points, lines, and polygons can be created as a separate file and incorporated in ArcGIS.

EXAMPLES OF CONVERSION TO DIGITAL MAPS

The examples below are taken from my and my students' work, because of ready access to our detailed geologic map data accumulated over >40 years of geologic mapping (Figure 1). Many of these maps remain in paper or mylar format in file drawers and theses, or have been published on paper (e.g., Hatcher, 1980; Hatcher and Acker, 1984; Ausburn et al., 2000), but a decision was made by myself during the mid-1990s to begin making digital geologic maps. In addition, we have been converting older non-digital maps as opportunities arise. As a result, close to 50 percent of the detailed geologic maps we have made over the past 40 years are now digitized, and this process continues (Figure 2).

Importance of Geology in Part of the Columbia, Tennessee, Quadrangle

The central part of the Columbia 7.5-minute quadrangle, Tennessee (Wilson et al., 1964; Figure 3) was mapped during 1962 as one of the first mapping projects in my career. It contains a topographic high in the central part of the quadrangle that preserves some locally complex and regionally important geologic relationships that were not appreciated until around 2000. Because of this, the geologic map of the central part of the quadrangle has been digitized (Figure 3). The geologic data remain good: contacts were correctly mapped and structural data were correctly measured and plotted. So, despite the fact that the data were collected in the early 1960s, understanding their regional geologic significance did not occur until recently. Beneath the Fort Payne chert (early Mississippian) is an unconformity that terminates a faulted syncline. This structure provides important evidence that the middle Paleozoic Neoacadian (360-350 Ma) orogeny affected both the southeastern Appalachians and adjacent craton. In addition, truncation of faults, synclines, and anticlines here and elsewhere (e.g., Wilson, 1971) beneath the unconformity provides a new model for hydrocarbon plays and exploration in the Middle Ordovician Nashville - Stones River Groups (Trenton - Black River ages) farther east beneath the Cumberland Plateau in Tennessee and southern Kentucky.

Prentiss Quadrangle, North Carolina

Geologic mapping of the Prentiss 7.5-minute quadrangle, North Carolina, was completed during the 1970s. The Prentiss quadrangle is published on paper (Hatcher, 1980), and was recently digitized from scanned raster images of the original paper maps (Figure 4). Bedrock and Quaternary geology has been systematically recompiled in MaPublisher georegistered Adobe Illustrator files. Once the digital compilation of contacts and structural and resource data was completed, a preliminary digital geologic map was printed and edited, permitting complete conversion of these maps to digital geologic maps.

Contacts in the Prentiss quadrangle were correctly located during geologic mapping, but at least one contact, the Soque River fault (southeastern part of the quadrangle), has been reinterpreted as a fault. This and the fault to the northwest are now known to be tectonostratigraphic terrane boundaries, with three tectonostratigraphic terranes represented here. Compare the digital map (Figure 4) with the Hatcher (1980) version. While the geometry and location of contacts on a properly constructed geologic map should be correct, interpretation of contacts may change through time. In addition to the terranes represented in the Prentiss quadrangle, there are several small massive sulfide deposits, sub-economic sillimanite deposits, and one or more rock units that could serve as sources of dimension stone. The U.S. Forest Service Coweeta Hydrologic Research Laboratory is also located in the Prentiss quadrangle (e.g., Hatcher, 1988; Swank and Crossley, 1988). The detailed geologic map of this quadrangle thus has considerable utility both from an academic and an applied geoscience perspective, as well as use by non-geologists for basic and applied research in forest ecology.

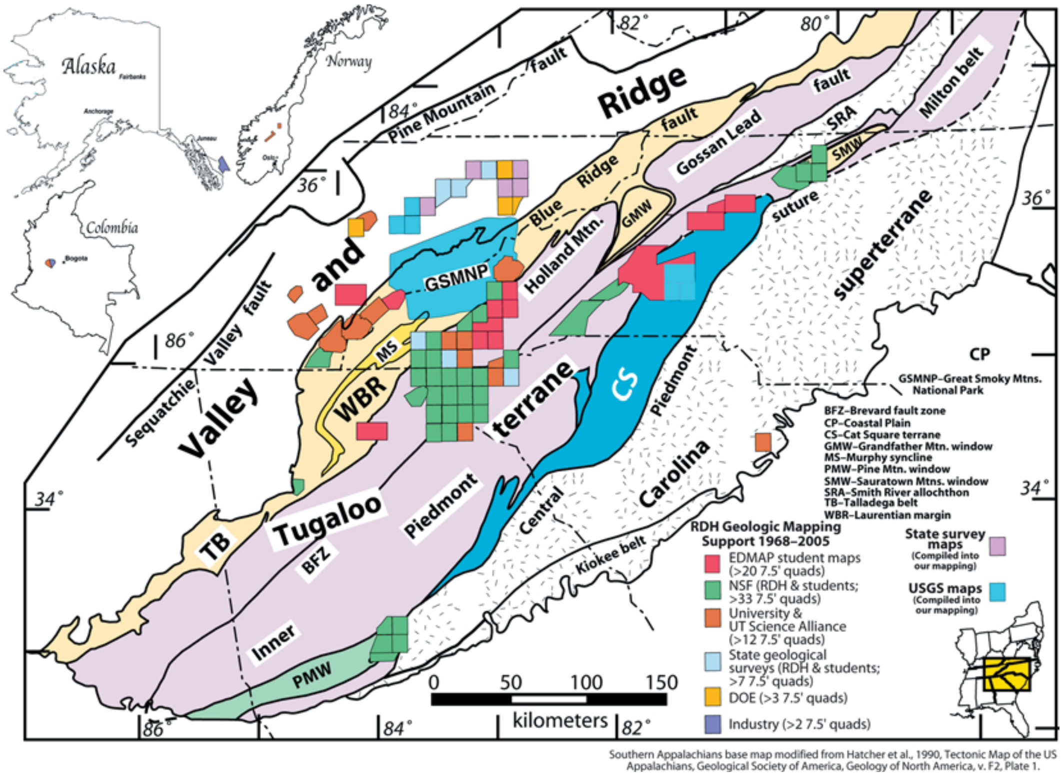


Figure 1. Index map of geologic mapping by RDH, undergraduate, and graduate students since the late 1960s, and funding sources [a more legible color version of this figure is available at http://ngmdb.usgs.gov/Info/dmt/docs/hatcher05.html]. EDMAP–Educational component of the USGS-managed National Cooperative Geologic Mapping Program. NSF–National Science Foundation. UT–University of Tennessee–Knoxville. DOE–U.S. Department of Energy.

Tugaloo Lake and Adjacent Quadrangles, Georgia-South Carolina

The detailed geologic map of the Tugaloo Lake quadrangle (Figure 5) was completed during the 1960s and early 1970s. It preserves a record of major tectonic events ranging from the 1.1 Ga Grenville orogeny, and several Paleozoic orogenies, through Mesozoic extension prior to opening of the Atlantic Ocean, and Tertiary-Quaternary drainage development. The map area is astride the eastern Blue Ridge and western Inner Piedmont geologic provinces in the internal parts of the southern Appalachians, in the Tugaloo tectonostratigraphic terrane (Hatcher, 2002). These provinces are separated by the Brevard fault zone, but several other major structures are also present. In the northwestern part of the geologic map is the southeastern flank of the Tallulah Falls dome, which is rimmed by several 1.15 Ga Grenville basement bodies (Hatcher et al., 2004). The multiply-reactivated Brevard fault zone (e.g., Hatcher, 2001) trends northeast-southwest across the central part of the map. Detailed geologic mapping has revealed that the entire Brevard fault zone is repeated by one or more large faults. In addition, late Brevard faults cut klippens that are remnants of the Alto allochthon in the Six Mile thrust sheet to the southeast (Hopson and Hatcher, 1988) providing critical evidence supporting the reactivation history of the fault zone. All rock units were later crosscut by Jurassic diabase dikes that preclude subsequent movement on the Brevard or other faults in this area. Rock units northwest of the Brevard fault zone contain subcommercial grade quantities of kyanite that have been prospected. Ordovician granitoids suitable for quarrying, and small amounts of sulfide minerals occur in the Poor Mountain Amphibolite southeast of the Brevard fault zone.

Graduate Student EDMAP Digital Geologic Maps

My graduate students and I have been producing digital 1:12,000 or 1:24,000 scale detailed geologic maps directly from field data since the mid-1990s. This permits daily compilation and revision of geologic maps as they accumulate data toward completion of a detailed

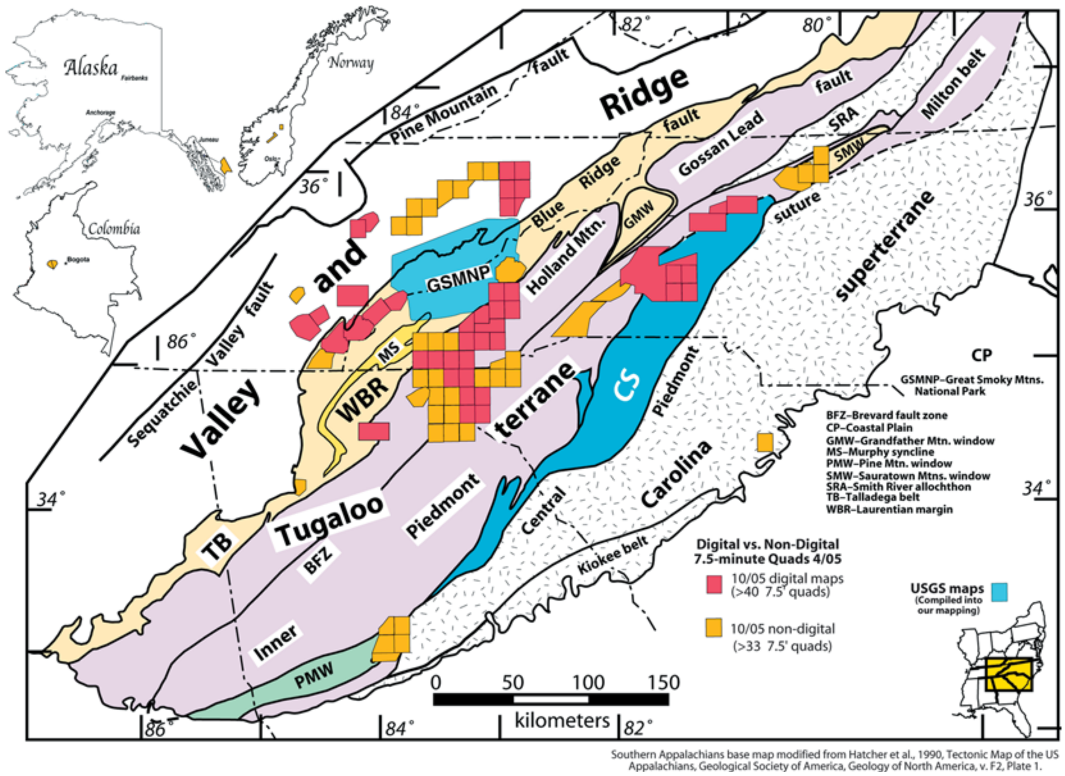


Figure 2. Index map showing digital (red) vs. non-digital (yellow) compiled geologic maps [a more legible color version of this figure is available at http://ngmdb.usgs.gov/Info/dmt/docs/hatcher05.html].

geologic map. Scott D. Giorgis (1999) mapped portions of four 7.5-minute quadrangles in the Appalachian Inner Piedmont near Morganton, North Carolina (Figure 6), and recognized a major fault—now considered a tectonostratigraphic terrane boundary, with supporting state-of-the-art geochronologic data. This and subsequent mapping by 10 more graduate students in that area has been supported by the EDMAP component of the USGS-administered National Cooperative Geologic Mapping Program. All produced high quality detailed geologic maps, but the geologic maps completed under this program throughout the United States exhibit a wide range of quality for a variety of reasons. These maps regardless of quality presently have no outlet into the community except for a few that are published or placed into open files by state geological surveys. Many EDMAP geologic maps remain in non-digital format, and many of those judged to be high quality maps should be scanned or converted to digital maps and made available through major databases, e.g., the National Geologic Map Database or GEON, or other outlets.

Building Islands From Detailed Geologic Maps

An important use of detailed geologic maps is compilation into maps of islands of knowledge, and tying these islands together using reconnaissance geologic mapping into more regional, small-scale maps useful for interpretation of regional geology and tectonic synthesis. Some of these islands consist of maps that have been scanned and redrawn in Adobe Illustrator and composited using MaPublisher into maps of larger areas (Figure 7). These geologic maps become very useful sources of information for compiling tectonic, resource, and other derivative maps (Figure 8).

CONCLUSIONS

1. The large numbers of non-federal, non-state, non-digital detailed geologic maps that exist in state geological surveys, faculty file drawers, theses, dissertations, engineering reports, in the files

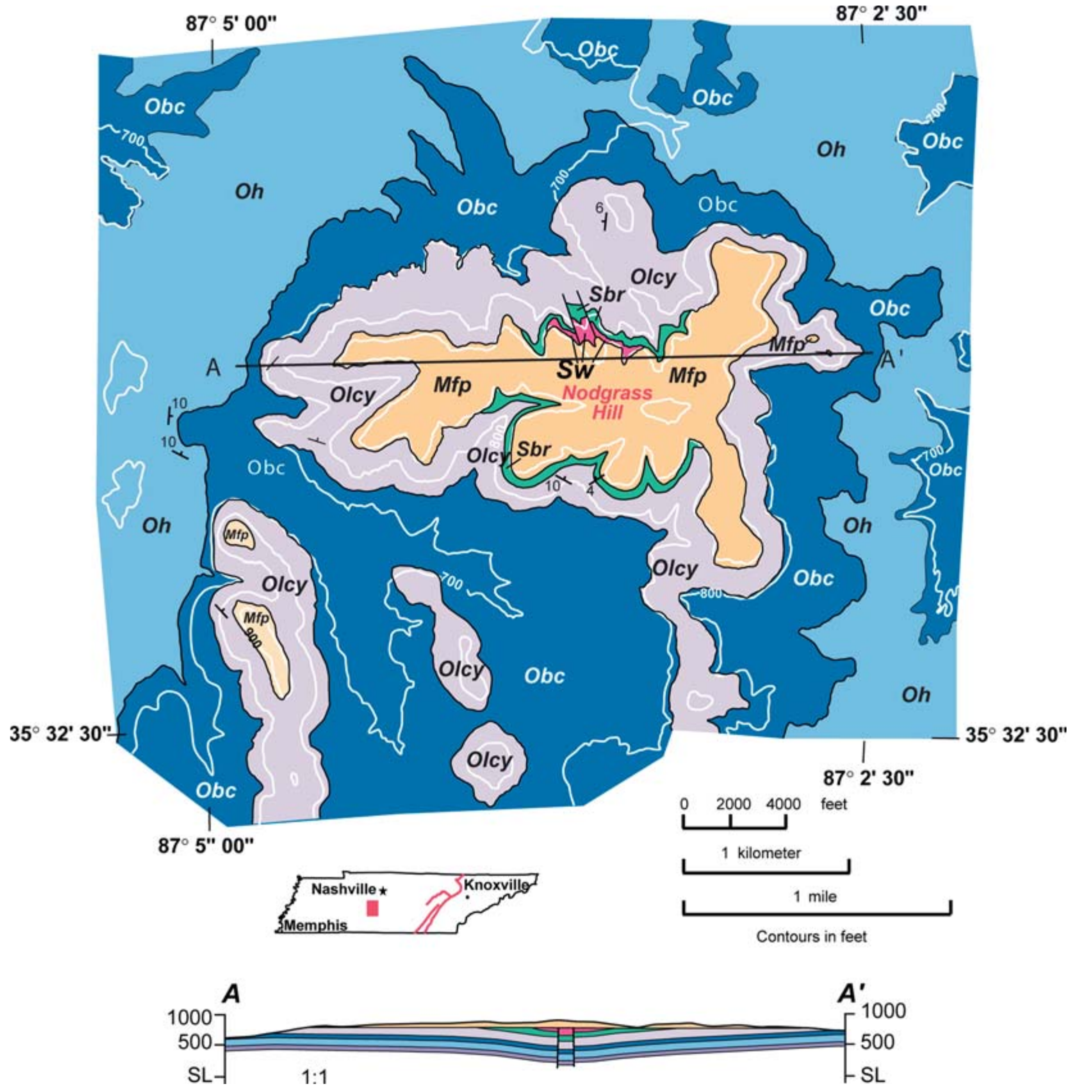


Figure 3. Geology of the central part of the Columbia, Tennessee, quadrangle (after Wilson et al., 1964) [a more legible color version of this figure is available at http://ngmdb.usgs.gov/Info/dmt/docs/hatcher05.html]. Oc – Carters Limestone (shown in cross section only; Middle Ordovician). Oh – Hermitage Formation (Middle Ordovician). Obc – Bigby-Cannon Limestone (Middle Ordovician). Olcy – Leipers Formation and Catheys Limestone, undivided (Middle and Upper Ordovician). Sbr – Brassfield Limestone (Lower Silurian). Sw – Wayne Group, undivided (Lower Silurian). Mfp – Ft. Payne Formation and Chattanooga Shale, undivided (lower Mississippian).

of mining and petroleum companies, should be made available in a digital format or as rasters in the large geospatial databases now being compiled.

2. Conversion of these maps to truly digital geo-referenced maps, then to well-attributed GIS databases, is possible, but is labor-intensive and requires substantial funding.
3. Digital geologic maps currently being made should be constructed in a format that permits ready conversion to GIS databases.

ACKNOWLEDGMENTS

U.S. National Science Foundation grants GA–1409, GA–20321, EAR–810852, EAR–8206949, EAR–8417894, EAR–9004604, and EAR–9814800 supported much of the detailed geologic mapping in the Carolinas and northeast Georgia by RDH and graduate students during the 1970s, 1980s, 1990s, and early 2000s. Additional detailed mapping support during the late 1960s and early 1970s was provided by the South Carolina Geological

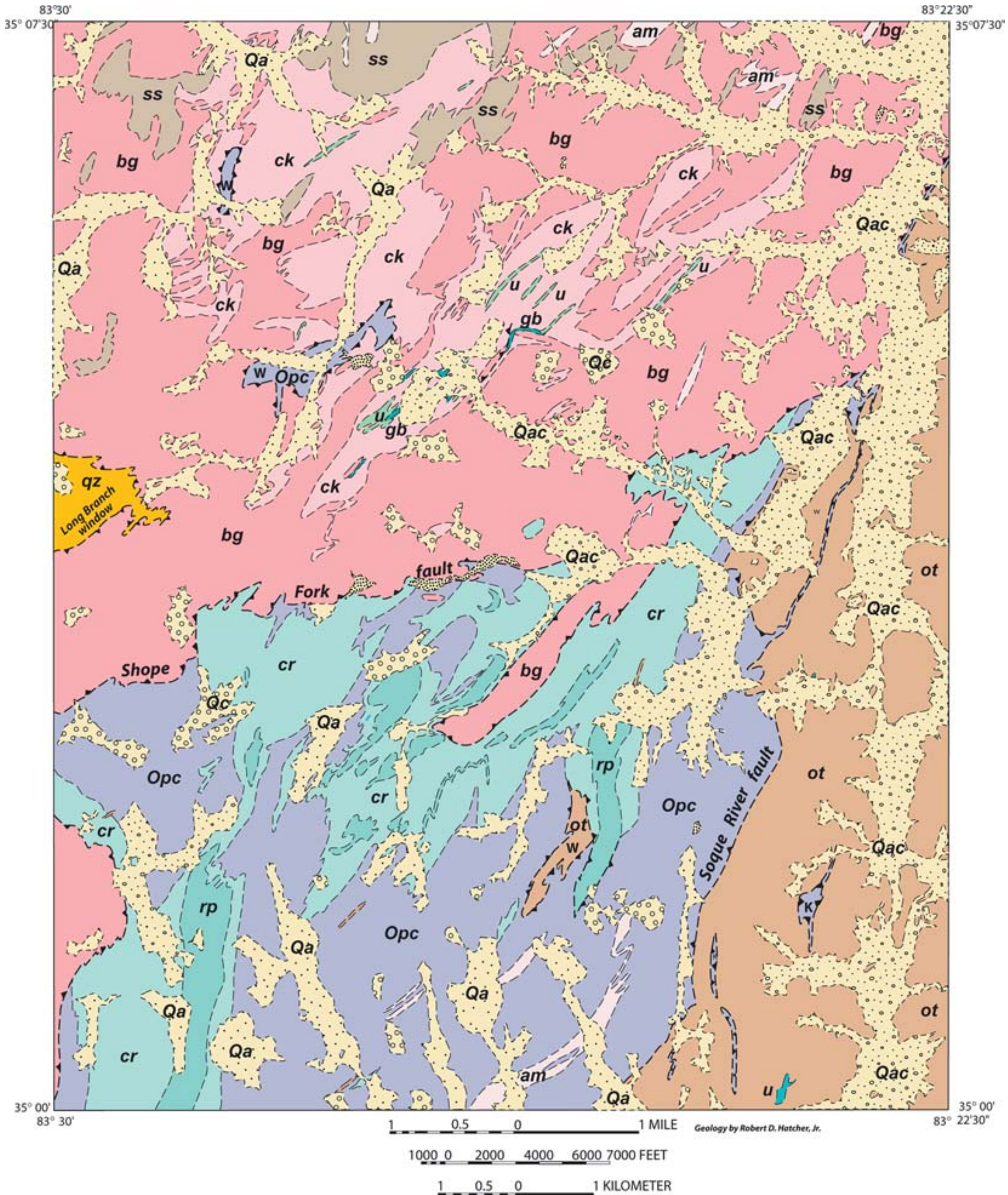


Figure 4. Digital geologic map of the Prentiss 7.5-minute quadrangle, North Carolina (after Hatcher, 1980) [a more legible color version of this figure is available at http://ngmdb.usgs.gov/Info/dmt/docs/hatcher05.html].

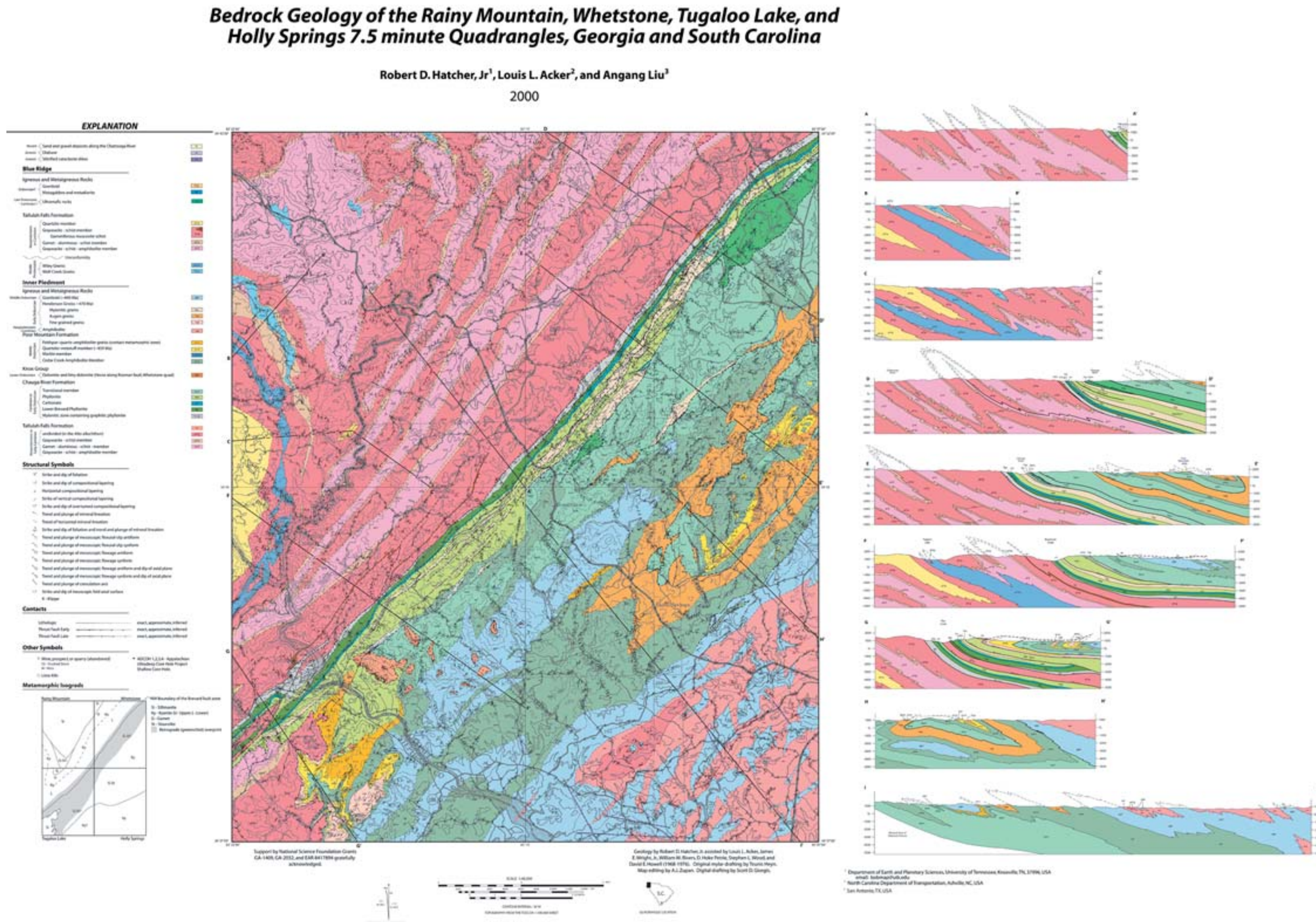


Figure 5. Bedrock geology of the Rainy Mountain, Whetstone, Tugaloo Lake, and Holly Springs 7.5 minute Quadrangles, Georgia and South Carolina (Hatcher, Acker, and Liu, 2000). Mapping assistance between 1968-1976 was provided by Clemson University undergraduates James E. Wright, Jr., William M. Rivers, D. Hoke Petrie, and Stephen L. Wood, and by David E. Howell (SC Geological Survey). Original mylar drafting by Teunis Heyn. Map editing by A.J. Zupan (SC Geological Survey). Digital graphic created at 1: 48,000 scale by S.D. Giorgis. This figure shows a reduced-scale version of the map; the full-sized, color map is available at http://ngmdb.usgs.gov/Info/dmt/docs/hatcher05.html.

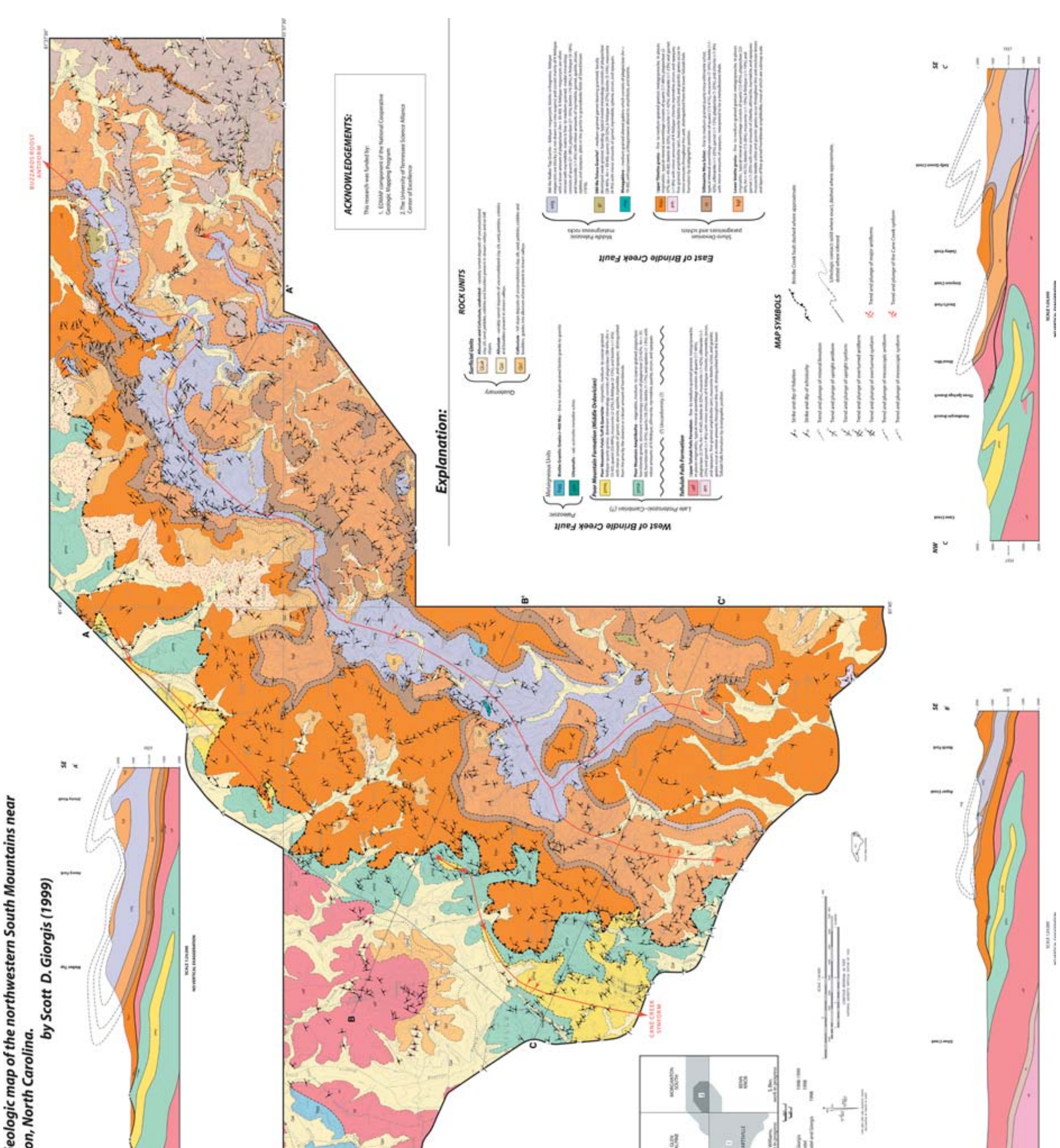


Figure 6. Geologic map of the northwestern South Mountains near Morganton, North Carolina (Plate 1 in Giorgis, 1999). This figure shows a reduced-scale version of the map; the full-sized, color map is available at http://ngmdb.usgs.gov/Info/dmt/docs/hatcher05.html.

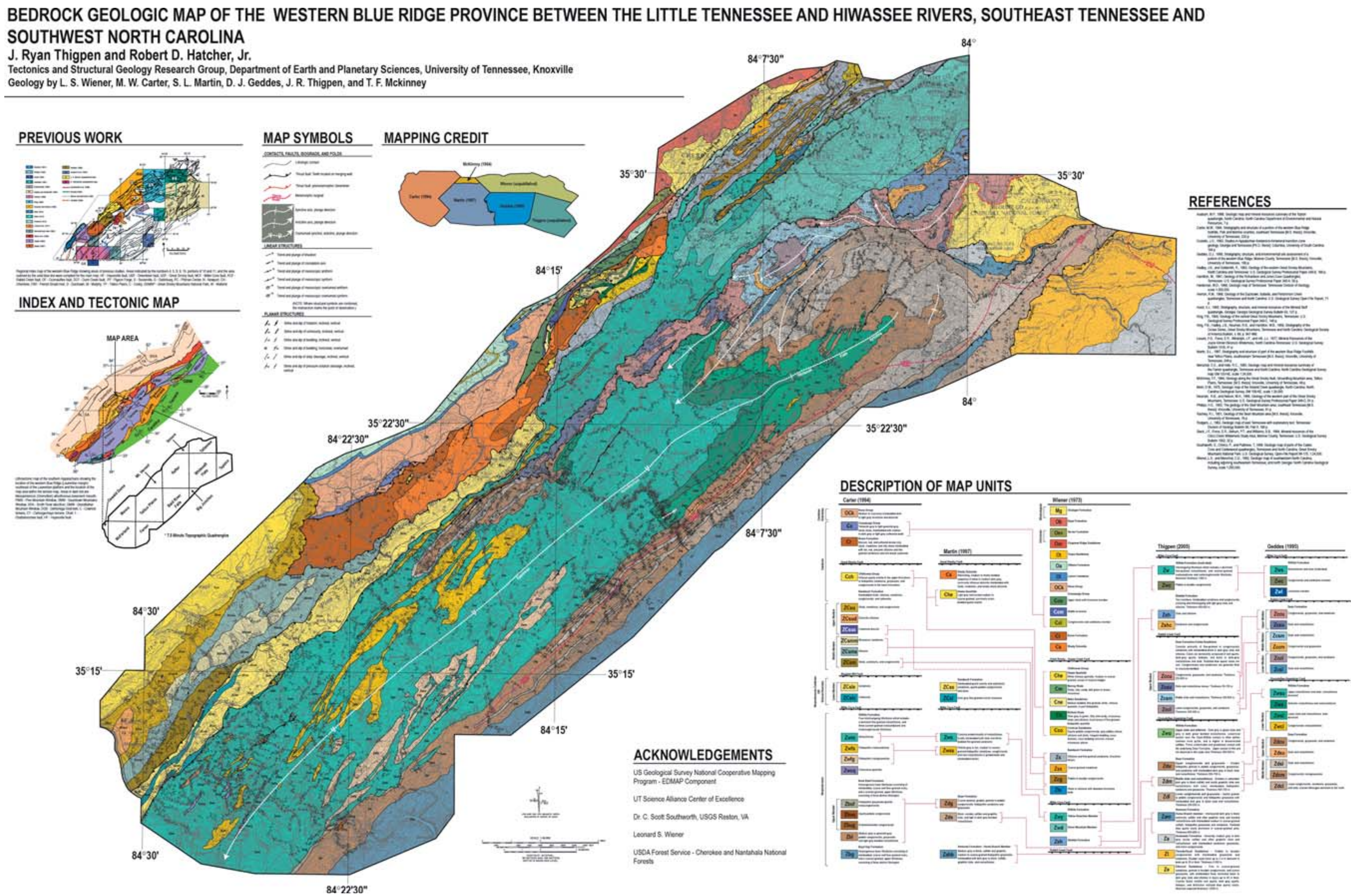


Figure 7. Bedrock geologic map of the western Blue Ridge Province between the Little Tennessee and Hiwassee Rivers, southeast Tennessee and southwest North Carolina (from Thigpen, 2005). Compiled from University of Tennessee theses by Phillips (1952), McKinney (1964), Carter (1994), Geddes (1995), Martin (1997), Whisner (2005), and Thigpen (2005); and Lesure et al. (1977), Merschat and Hale (1983), and Slack et al. (1984). This figure shows a reduced-scale version of the map; the full-sized, color map is available at http://ngmdb.usgs.gov/Info/dmt/docs/hatcher05.html.

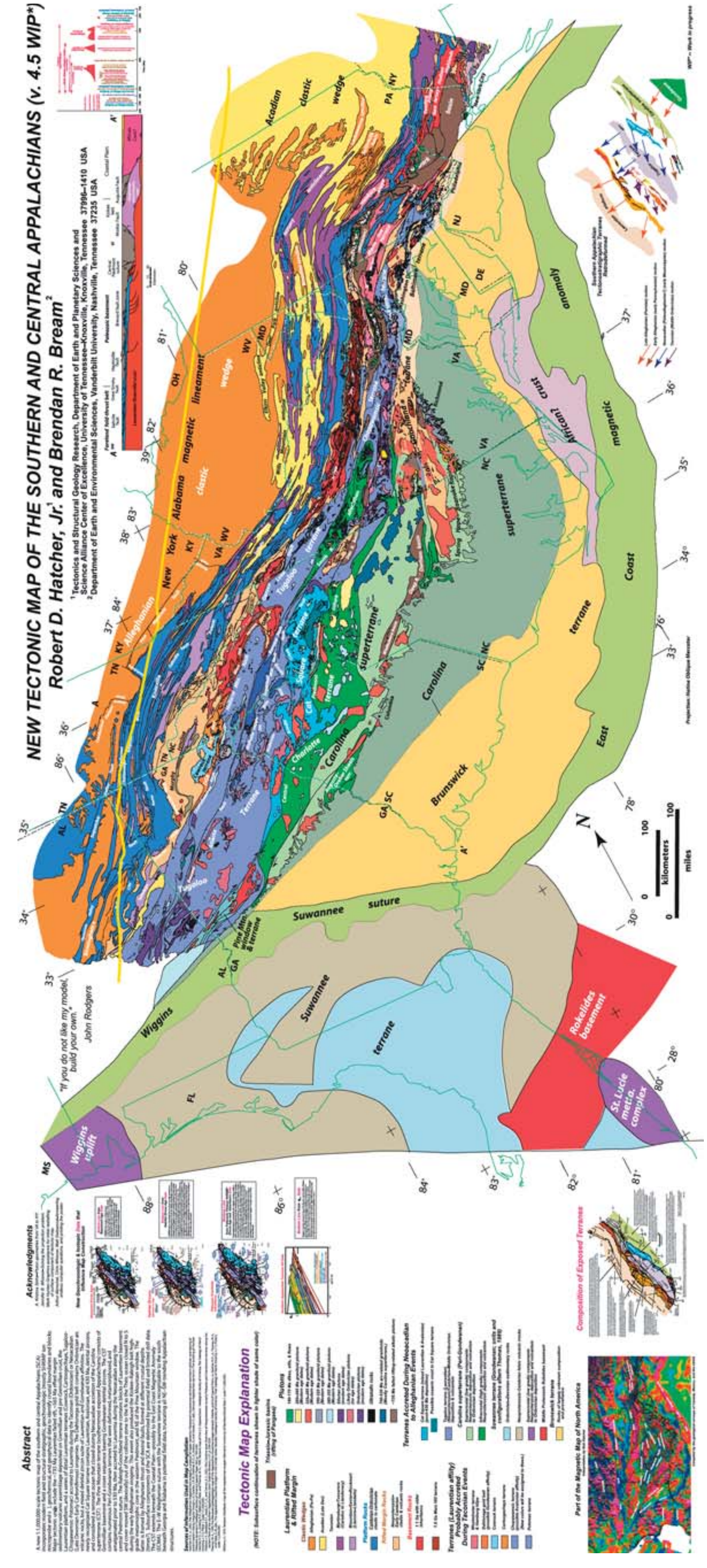


Figure 8. New tectonic map of the southern and central Appalachians (modified from Hatcher and Bream, in preparation, 2006). This figure shows a reduced-scale version of the map; the full-sized, color map is available at http://ngmdb.usgs.gov/Info/dmt/docs/hatcher05.html.

Survey, Henry S. Johnson, Jr., and Norman K. Olson, state geologists. National Science Foundation Grants GA-1409, GA-2032, and EAR 8417894, and the South Carolina Geological Survey, supported the mapping in Figure 5. Detailed geologic mapping in North Carolina, Georgia, and Tennessee during the late 1990s and early 2000s has been supported by grants from the EDMAP component of the National Cooperative Geologic Mapping Program administered by the U.S. Geological Survey, with cooperation from the geological surveys of North Carolina, Tennessee, and Georgia. The University of Tennessee Science Alliance Center of Excellence also has provided considerable support for RDH and students since 1986. Review by David Soller resulted in significant improvement of the manuscript, but I remain culpable for any errors of fact or interpretation.

REFERENCES

Ausburn, M.P., Hatcher, R.D., Jr., Bryan, J.G., Godfrey, S.E., and Acker, L.L., 1998, Geologic map and mineral resources summary of the Topton quadrangle, North Carolina: North Carolina Geological Survey, GM 159–NW, scale 1:24,000.

Carter, M.W., 1994, Stratigraphy and structure of part of the Blue Ridge Foothills, Polk and Monroe Counties, Tennessee [M.S. thesis]: Knoxville, University of Tennessee, 233 p.

Geddes, D.J., 1995, Stratigraphy, structure, and environmental site assessment of a portion of the western Blue Ridge, Monroe County, Tennessee [M.S. thesis]: Knoxville, University of Tennessee, 195 p.

Giorgis, S.D., 1999, Inner Piedmont geology of the northwestern South Mountains near Morganton, North Carolina [M.S. thesis]: Knoxville, University of Tennessee, 186 p.

Hatcher, R.D., Jr., 1980, Geologic map and mineral resources of the Prentiss quadrangle, North Carolina: North Carolina Geological Survey, GM 167-SW, scale 1:24,000.

Hatcher, R.D., Jr., 1988, Bedrock geology and regional geologic setting of Coweeta Hydrologic Laboratory in the eastern Blue Ridge, with some discussion of Quaternary deposits and structural controls of topography, in Swank, W.T., and Crossley, D.A., Jr., eds., Studies in ecology, Coweeta Symposium Volume 66: New York, Springer–Verlag, p. 81–92.

Hatcher, R.D., Jr., 2001, Rheological partitioning during multiple reactivation of the Paleozoic Brevard Fault Zone, Southern Appalachians, USA, in Holdsworth, R.E., Strachan, R.A., Macloughlin, J.F., and Knipe, R.J., eds., The nature and significance of fault zone weakening: Geological Society of London Special Publication 186, p. 255–269.

Hatcher, R.D., Jr., 2002, An Inner Piedmont primer, in Hatcher, R.D., Jr., and Bream, B.R., eds., Inner Piedmont tectonics focused mostly on detailed studies in the South Mountains and the southern Brushy Mountains, North Carolina: Carolina Geological Society Guidebook, North Carolina Geological Survey, p. 1–18.

Hatcher, R.D., Jr., and Acker, L.L., 1984, Geology of the Salem quadrangle, South Carolina: South Carolina Geological Survey, MS–26, 23 p., scale 1:24,000.

Hatcher, R.D., Jr., Acker, L.L., and Liu, A., 2000, Bedrock geology of the Rainy Mountain, Whetstone, Tugaloo Lake, and Holly Springs 7.5 minute Quadrangles, Georgia and South Carolina: unpublished map funded by National Science Foundation Grants GA-1409, GA-2032, and EAR 8417894, scale 1:24,000.

Hatcher, R.D., Jr., and Bream, B.R., in preparation 2006, Tectonic map of the southern and central Appalachians: scale 1:1,000,000.

Hatcher, R.D., Jr., Bream, B.R., Miller, C.L., Eckert, J.O., Jr., Fullagar, P.D., and Carrigan, C.W., 2004, Paleozoic Structure of Southern Appalachian Blue Ridge Grenvillian Internal Basement Massifs, *in* Tollo, R.P., Corriveau, L., McLelland, J., and Bartholomew, M.J., eds., Proterozoic evolution of the Grenville orogen in North America: Boulder, Colorado, Geological Society of America Memoir 197, p. 525–547.

Hopson, J.L., and Hatcher, R.D., Jr., 1988, Structural and stratigraphic setting of the Alto allochthon, northeast Georgia: Geological Society of America Bulletin, v. 100, p. 339–350.

Lesure, F.G., Force, E.R., Windolph, J.F., and Hill, J.J., 1977, Mineral Resources of the Joyce Kilmer-Slickrock Wilderness, North Carolina-Tennessee: U.S. Geological Survey Bulletin 1416, 41 p.

Martin, S.L., 1997, Stratigraphy and structure of part of the western Blue Ridge foothills near Tellico Plains, southeastern Tennessee [M.S. thesis]: Knoxville, University of Tennessee, 249 p.

Merschat, C.E., and Hale, R.C., 1983, Geologic map and mineral resources summary of the Farner quadrangle, Tennessee and North Carolina: North Carolina Geological Survey Map GM 133-NE, scale 1:24,000.

McKinney, T.F., 1964, Geology along the Great Smoky fault, Groundhog Mountain area, Tellico Plains, Tennessee: [M.S. thesis]: Knoxville, University of Tennessee, 49 p.

Phillips, H.E., 1952, The geology of the Starr Mountain area, southeast Tennessee [M.S. thesis]: Knoxville, University of Tennessee, 61 p.

Slack, J.F., Force, E.R., Behum, P.T., and Williams, B.B., 1984, Mineral resources of the Citico Creek Wilderness study area, Monroe County, Tennessee: U.S. Geological Survey Bulletin 1552, 32 p.

Swank, W.T., and Crossley, D.A., Jr., eds., 1988, Studies in ecology, Coweeta Symposium Volume 66: New York, Springer–Verlag, 469 p.

Thigpen, J.R., 2005, Stratigraphic and structural relationships of the Ocoee Supergroup, southern Appalachians: Implications for Neoproterozoic rift basin architecture and Paleozoic collisional orogenesis [M.S. thesis]: Knoxville, University of Tennessee, 156 p.

Whisner, S.C., 2005, The Middle Ordovician Tellico–Sevier Syncline: A stratigraphic, structural, and paleoseismic investigation [Ph.D. dissertation]: Knoxville, University of Tennessee, 228 p.

Wilson, C.W., Jr., 1971, Geologic map of the Buffalo Valley quadrangle, Tennessee: Tennessee Division of Geology, GM 322-NE, scale 1:24,000.

Wilson, C.W., Jr., Kirby, J.P., and Miller, R.A., 1964, Geologic map of the Columbia quadrangle, Tennessee: Tennessee Division of Geology, GM 57-SE, scale 1:24,000.

The National Geologic Map Database Project: Overview and Progress

By David R. Soller[1], Thomas M. Berg[2] and Nancy R. Stamm[1]

[1]U.S. Geological Survey
926-A National Center
Reston, VA 20192
Telephone: (703) 648-6907
Fax: (703) 648-6977
e-mail: drsoller@usgs.gov; nstamm@usgs.gov

[2]Ohio Geological Survey
2045 Morse Rd., Bldg. C-1
Columbus, OH 43229-6693
Telephone: (614) 265-6988
Fax: (614) 268-3669
e-mail: thomas.berg@dnr.state.oh.us

The National Geologic Map Database (NGMDB) project continues to fulfill its mandate[1]. Some of its accomplishments are specific and tangible, and others are more general in nature—for example, the NGMDB contributes to advancements in digital mapping techniques and database design by agencies in the United States and internationally. However, without extensive collaboration from highly skilled and enthusiastic members of the state geological surveys and the Geological Survey of Canada, these accomplishments would not have been possible. Highlights of the past year include:

- the Geoscience Map Catalog continued to increase its content; it now contains bibliographic records for more than 70,000 map products published by about 300 organizations including the U.S. Geological Survey (USGS), 45 state geological surveys, universities, and scientific societies and organizations,
- the prototype Geologic Map Image Library, an extension of the Map Catalog, has evolved into a useful collection of more than 4000 high-resolution images of geologic maps,
- the websites for the NGMDB's principal databases (Map Catalog, Image Library, and GEOLEX [the U.S. Geologic Names Lexicon]) were visited about 100,000 times by 30,000 users each month. This is a dramatic increase (about 100%) from last year. NGMDB personnel responded to the many inquiries and requests from these users.
- the project contributed significantly to evolution of the North American standard data model, science terminology, and data-interchange format, and to the U.S. cartographic standard for geologic maps. The project also contributed to technical work under the aegis of the International Union of Geological Sciences (IUGS), designed to improve interoperability among map databases worldwide. Internationally, NGMDB staff participated as a council member of the IUGS Commission for the Management and Application of Geoscience Information, and as a member of the map standards committee for the Commission for the Geological Map of the World,
- the project coordinated the ninth annual Digital Mapping Techniques workshop, bringing together about 100 technical experts from 47 agencies, and
- work continued on design and implementation of the online map database, focusing on development of a data-entry tool and standardized science terminology.

[1]At each annual Digital Mapping Techniques workshop, this project offers a report of progress. For workshop attendees, a comprehensive overview of the project's numerous activities and databases is not necessary. However, because many readers of this volume are not familiar with the project's goals and long-term accomplishments, we felt it appropriate to update the previous year's report (Soller and Berg, 2004) in order to provide a comprehensive, up-to-date summary.

INTRODUCTION

This project provides an unusual if not unique opportunity to foster better relations and technical collaboration among all geological surveys in the nation. Given the nature of the issue—the creation and management of geoscience map information in digital format during a period of rapid technological evolution—collaboration is critically important. Perhaps more significant, these are changing times for all geological surveys—funding and staff seem to become more scarce each year—and through collaboration we can share our intellectual and computing resources and not "reinvent the wheel" within each agency.

Before describing the NGMDB components and progress, we wish to highlight the various mechanisms by which we define and accomplish our goals. Because advice, guidance, and technical collaboration are an integral part of this project, we discuss the project plan at numerous venues throughout the year. These include geoscience and related professional society meetings, the Digital Mapping Techniques workshop, and site visits to state geological surveys. Advice gathered at these venues serves to refine and, in some cases, to redirect the project's goals. Comments from users, generally via our Web feedback form, also provide us with valuable perspectives, and have prompted us to make numerous modifications, especially to our Web interface design.

Because the NGMDB's scope is so broad, its success relies on the many people and agencies that participate in its activities. Members of the committees and small working groups that have advised and contributed to the project's goals are listed in Appendix A. These committees are an important mechanism for coordinating with each agency, and they deserve noting:

- Digital Geologic Mapping Committee of the Association of American State Geologists (AASG)—charged with representing all state geological surveys in the NGMDB project, and with providing authoritative guidance to the project.
- Technical Advisory Committee—provided technical vision and guidance to the NGMDB, especially on the project's Phase Three.
- Map Symbol Standards Committee—oversees the completion, and then the maintenance, of the Geologic Map Symbolization Standard, which will become a Federal standard endorsed by the Federal Geographic Data Committee.
- AASG/USGS Data Capture Working Group—coordinates the annual Digital Mapping Techniques workshop, and provides through an email listserver a forum for exchange of technical information.
- AASG/USGS Metadata Working Group—summarized issues related to creating metadata, and identified useful software tools.
- AASG/USGS Data Information Exchange Working Group—created technical guidance for map publication guidelines.
- AASG/USGS Data Model Working Group—defined a draft version of a standard geologic map data model.
- North American Data Model Steering Committee—succeeded the Data Model Working Group, and is developing a standard data model, science language, and data-interchange format for the North American geoscience community.
- NGMDB contact-persons—within each state geological survey, several people work with us on various project databases and activities.

BACKGROUND

The National Geologic Mapping Act of 1992 and its reauthorizations in 1997 and 1999 (PL106-148) require a National Geologic Map Database to be built by the USGS in cooperation with the AASG. This database is intended to serve as a "national archive" of standardized geoscience information for addressing societal issues and improving our base of scientific knowledge. The Mapping Act anticipates a broad spectrum of users including private citizens, professional geologists, engineers, land-use planners, and government officials. The Act requires the NGMDB to include these geoscience themes: geology, geophysics, geochemistry, paleontology, and geochronology.

In mid-1995, the general stipulations in the Geologic Mapping Act were addressed in the proposed NGMDB design and implementation plan developed by the USGS and AASG. Summaries of this plan are listed in Appendix B. Because of the mandate's broad scope, we proposed a phased, incremental design for the NGMDB. A phased approach has two benefits: 1) it enables us to identify the nature and quality of existing information and quickly serve it to the public; and 2) it gives us time to build consensus and expertise among the database designers in the state geological surveys and the USGS. Furthermore, it enables us to more effectively consider and respond to evolving technology and user needs. These phases, and our progress, are shown in Figure 1.

In the first and most fundamental phase of the project, we are building a set of easy-to-use reference databases; for example, a comprehensive, searchable map catalog of all geoscience maps in the United States, whether in paper or digital format. The second phase of the project focuses on the development of standards and guidelines needed to improve the utility of digital maps. The third phase proposes to, in the long term, develop an online database of (mostly vector-based) geologic map information at various scales and resolution.

In late 1995, work began on Phase One. The formation in mid-1996 of several AASG/USGS Standards

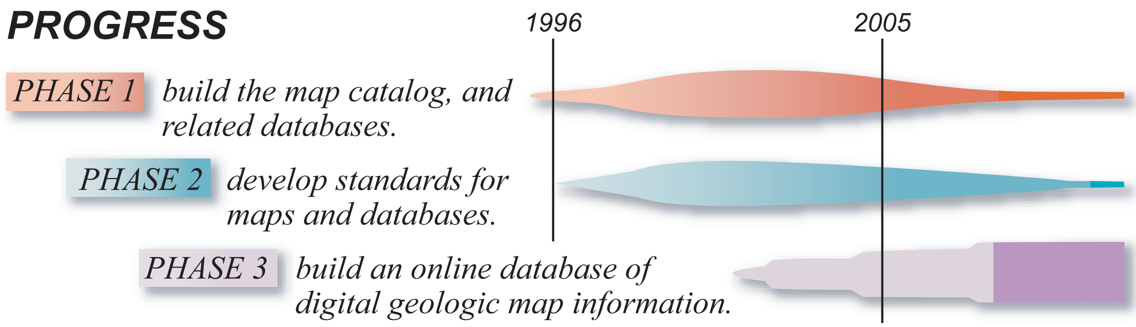


Figure 1. Diagram showing the three NGMDB Phases, and progress toward our goals (for example, documenting in the Geoscience Map Catalog all maps and related products for the United States and its territories and possessions).

Working Groups initiated work on Phase Two. The project opened its Web site to the public in January 1997, as a prototype intended to solicit comments on the Map Catalog. At the Digital Mapping Techniques '98 through '05 workshops, a series of presentations and discussion sessions provided updates on the NGMDB and, specifically, on the activities of the Standards Working Groups (see Appendix B). This report summarizes accomplishments since the project's inception, and therefore repeats material from previous reports, but it focuses on activities since mid-2004. Additional and more current information may be found at the NGMDB project-information Web site, http://ngmdb.usgs.gov/info. The searchable databases are available at http://ngmdb.usgs.gov.

To submit general comments about project scope and direction, please address the authors directly. For technical comments on the databases or Web page design, please use our Web feedback form; this form is linked from many of our search pages (see "Your comments are welcome", at http://ngmdb.usgs.gov/).

PHASE ONE

Through ongoing discussions with private companies, citizens, government officials, and research geologists, it is clear that first and foremost, we need to provide reference databases so that geoscience maps and descriptive information can be found and used. Many people want to better understand the geologic framework beneath their home, business, or town, and so we are building several databases that support general, "data-discovery" questions posed by citizens and researchers alike. These reference databases are: 1) the Geoscience Map Catalog and its extension, the prototype Geologic Map Image Library; 2) GEOLEX, the U.S. geologic names lexicon; and 3) Geologic Mapping in Progress, which provides information for ongoing National Cooperative Geologic Mapping Program (NCGMP) mapping projects, prior to inclusion of their products in the Map Catalog. Plans for the National Paleontology Database also are discussed below.

Figure 2 shows the number of people (actually, the number of unique IP addresses or computers) who have used the NGMDB, per month since it opened to the public in January 1997. These numbers indicate that the site has become a useful resource. Additional increases in use are expected as the Map Catalog, GEOLEX, and Image Library become fully populated.

The Geoscience Map Catalog and Image Library

"I want to know if a map exists for an area, and where I can get a copy of it…"
"I want to see a picture of this geologic map, online…"

Many organizations produce paper and digital geoscience maps and related products. Discovering whether a product exists for an area, and if so, where it can be purchased or obtained online, can be a time-consuming process. In the past, people found this information by contacting various agencies and institutions, and by conducting extensive library searches. To increase accessibility and use of these paper and digital products, we built the Geoscience Map Catalog as a comprehensive, searchable database of all maps and related products for the United States and its territories and possessions.

The Geoscience Map Catalog contains bibliographic records for more than 70,000 products from about 300 publishers (see our most current list of publishers at http://ngmdb.usgs.gov/ngmdb/pub_series.html). Most of these products are from the USGS and from 45 state geological surveys. Other publishers include state agencies, federal agencies, scientific societies, park associations, universities, and private companies. Products range from digital maps to books that don't contain maps but describe the geology of an area, and can be formal series products,

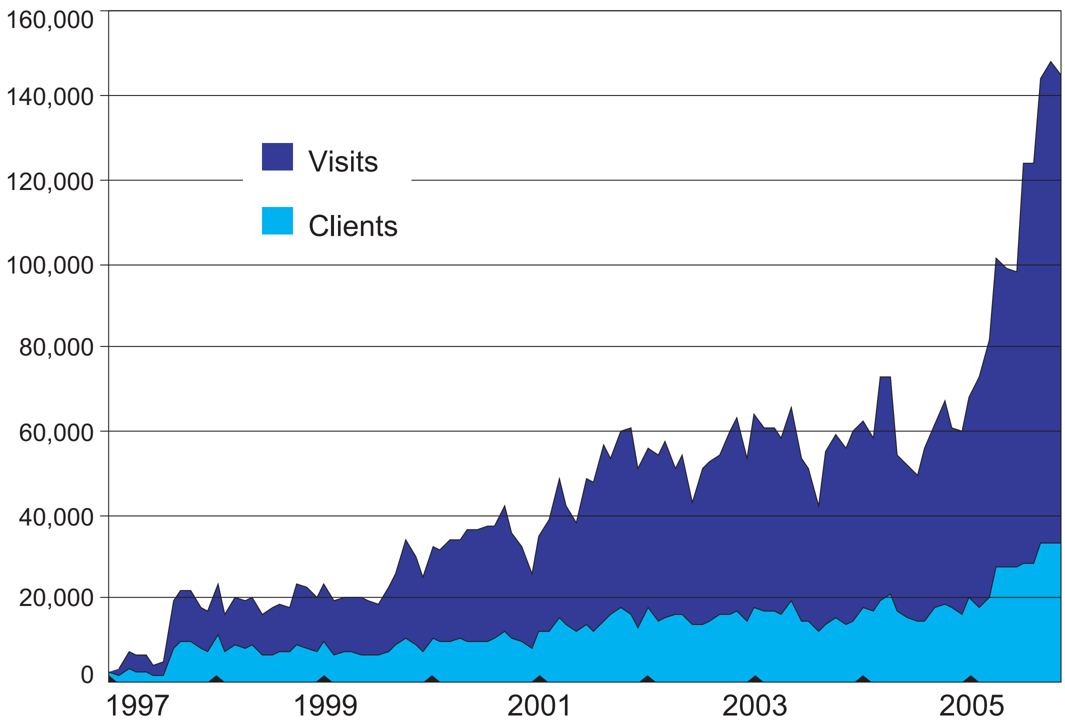


Figure 2. Web usage for the Geoscience Map Catalog, GEOLEX, Image Library, and Mapping in Progress Databases. This diagram shows that the number of people (actually, the number of unique IP addresses or computers) using the NGMDB has gradually increased as these resource databases become more widely known; this usage trend is punctuated by sharp increases after essentially all USGS maps were entered into the Catalog (ca. 2000), again after many state geological surveys began to enter map records (ca. 2001), and more recently presumably due to increases in natural commodity prices.

open-file reports, or unpublished dissertations (Figure 3). Because there are many types of geoscience maps and related products, we categorize them by theme (Figure 4).

The Geoscience Map Catalog provides links to more than 4000 published, downloadable products of the USGS and the state geological surveys. These links are established only to stable Web pages that provide the official copy-of-record for the publication—in the USGS, links are established only to the Publications Server and the NSDI Clearinghouse node.

The Geoscience Map Catalog identifies products that meet the user's search criteria, and provides links to the downloadable data and metadata, to a depository library, or to the appropriate organization for information about how to purchase the product (Figure 5). We address the diverse needs of our user audience through four search options. The easy-to-use Place Name Search is based on the USGS Geographic Names Information System (GNIS); it is designed mostly to address the needs of non-geologists who want to use a simple interface to find information about their home, town, or worksite. In contrast, other choices such as the Comprehensive Search offer more search criteria.

Through discussions with users, and from comments received via our Web feedback form, it became clear that many people are interested in viewing and/or obtaining maps "online." Interpretation of the phrase "providing maps online" varies widely—to some people, it implies access to fully attributed, vector-based map databases, whereas to other people, it implies access to map images. Regarding the vector-based map database, we address this large task in Phase Three, below. Regarding access to map images, we have begun to provide these to users via our prototype Geologic Map Image Library (Soller and Berg, 2003). The Image Library contains high-resolution (300 dpi) images that are compressed into MrSID format and served to the user via a standard Web browser. These MrSID-compressed images are easily and quickly viewed in detail, and in most cases can be downloaded. Upon request, we also provide access to the source image file, in non-georeferenced TIFF format.

The Image Library is a relatively new initiative, and its search interface and design are still under development. We anticipate that in the near future it will be integrated into the Geoscience Map Catalog because: 1) the Image Library's database is based on a subset of the

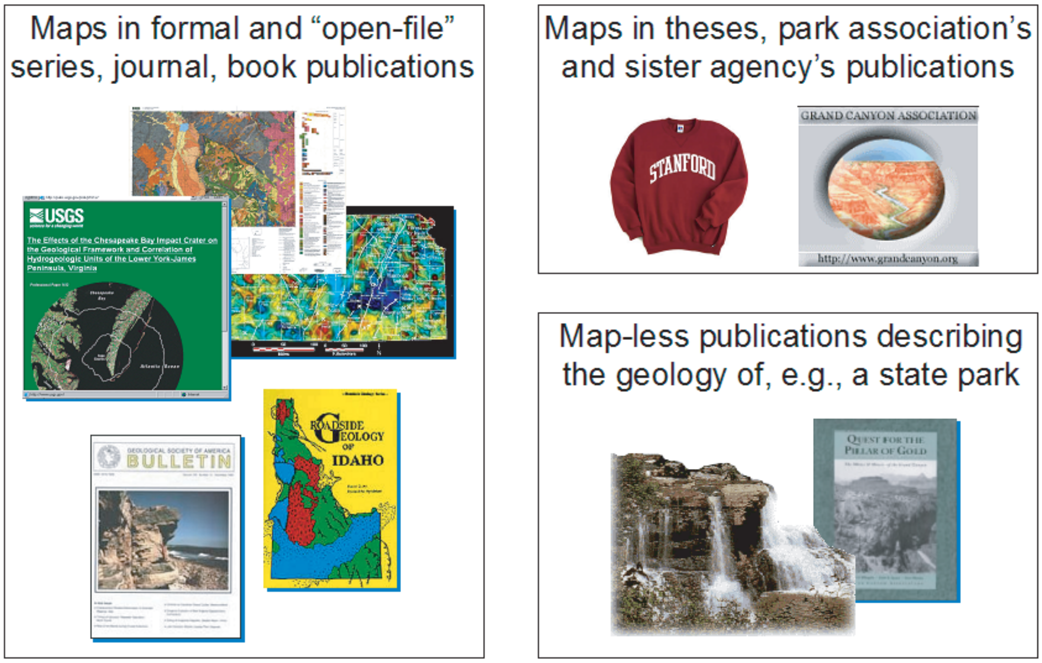


Figure 3. Bibliographic records in the Geoscience Map Catalog are drawn from a diverse group of about 300 publishers.

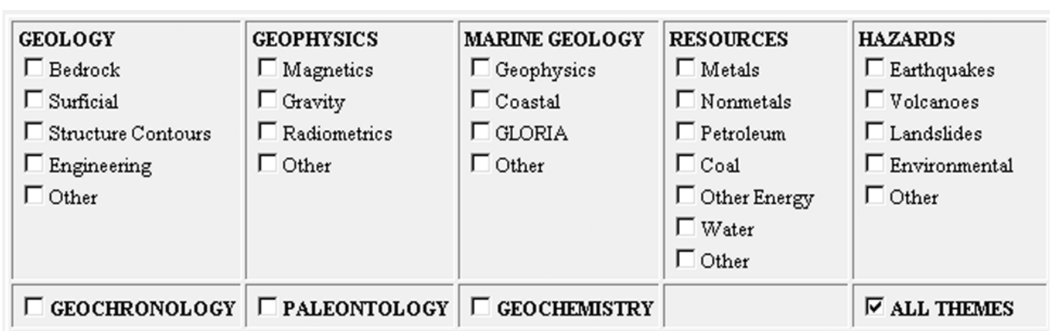


Figure 4. A portion of the Geoscience Map Catalog search page, showing the types of products included.

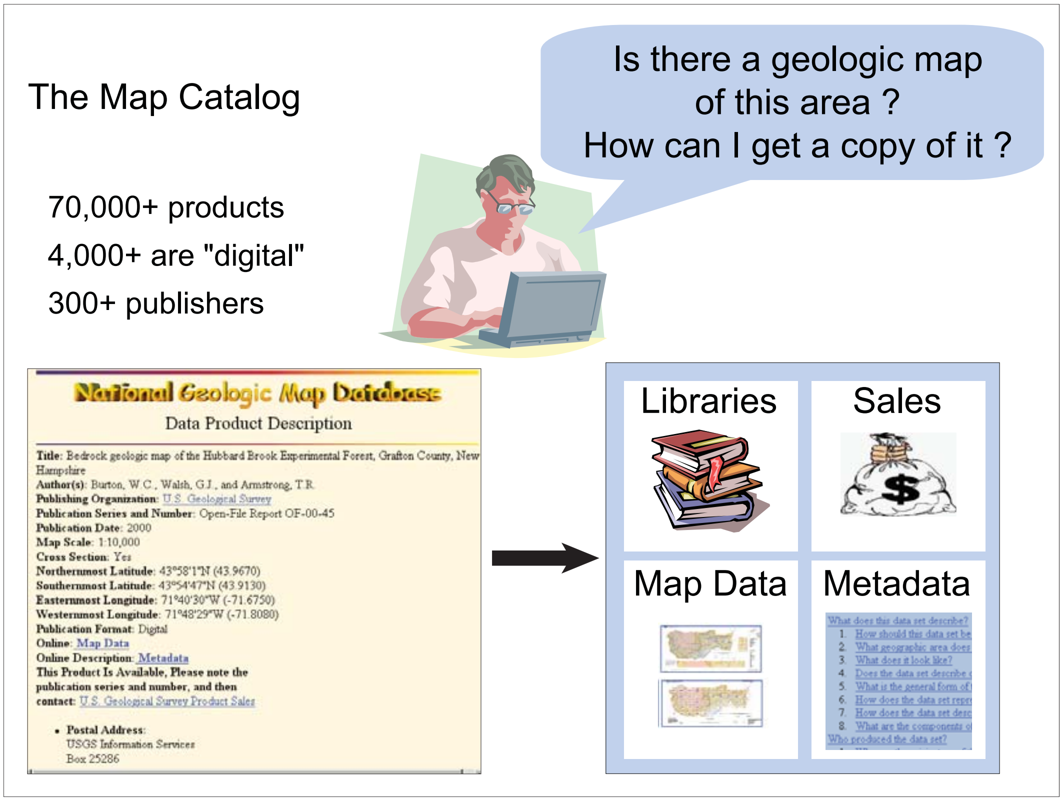


Figure 5. Interested in knowing something about the geology of an area (such as the land beneath their house), the user queries the Geoscience Map Catalog, which returns a hit list of possibly useful maps and related products. The user selects one of these and, from the Product Description Page (shown on left side of figure), obtains further information and can then choose to buy the product, view and download it, inspect the metadata, or find it at a depository library.

Map Catalog's bibliographic database, and 2) an integrated search of bibliographic information and images will benefit our users.

The U.S. Geologic Names Lexicon ("GEOLEX")

"I want to know more about the geologic units shown on this map…"

This is the nation's lexicon of geologic nomenclature. GEOLEX contains information for more than 16,000 geologic units in the U.S. (Stamm and others, 2000). It is an excellent resource for finding significant publications that defined and described geologic units mapped in the U.S. These publications can be critically important in field studies, enabling students and mappers to compare these published descriptions with what they see in the field.

GEOLEX includes the content of the four geologic names databases on USGS Digital Data Series DDS-6 (Mac Lachlan and others, 1996). Before incorporating into GEOLEX, those databases were consolidated, revised, and error-corrected. Our work continues to focus on:

1. resolving the name conflicts found in the four databases of Mac Lachlan and others (1996). This is done by consulting publications, previous U.S. geologic names lexicons (listed in Appendix A of Stamm and others, 2000), and the records of the U.S. Geologic Names Committee (GNC),
2. using the previous lexicons to incorporate type locality, publication history, geologic age, areal extent, and usage information for many geologic units listed in Mac Lachlan and others (1996),
3. adding geologic names not recorded in Mac Lachlan and others (1996) but found in the old USGS regional geologic names card catalogs, and
4. adding geologic names approved by the state geological surveys but not recorded in GEOLEX.

Many state geological surveys have been registering new geologic names with the USGS for decades, and are encouraged to continue this practice. In order to promote standardized geologic nomenclature within the U.S., we

are petitioning the USGS to re-establish the GNC. Formerly a committee that focused on nomenclature issues within the USGS, we propose that the new GNC should include members from each state geological survey. When a conflict arises, GNC members from the USGS and those states affected will resolve it, and any changes will be recorded in GEOLEX. Through this mechanism, we anticipate that GEOLEX will serve the entire U.S. geoscience community.

Geologic Mapping in Progress Database

"I see from the Map Catalog that a map hasn't been published for this area—is anyone mapping there now?"

Our Geologic Mapping in Progress Database provides users with information about current mapping activities (mostly at 1:24,000- and 1:100,000-scale, but at 1:63,360- and 1:250,000-scale in Alaska) that is funded by the National Cooperative Geologic Mapping Program (NCGMP). In 2005, we significantly updated this database with information provided by the NCGMP.

Paleontology Database

"I want to know if there is any fossil data from this area…"

The NGMDB project has designed, and will soon develop, a National Paleontology Database (see Wardlaw and others, 2001). As originally envisioned, we would build prototypes of this database in areas where geologic mapping is underway, so that we could work with mapping projects to design a National database useful to science as well as to the public. Plans for a prototype were delayed in order to assess new priorities and what could be accomplished with funding and personnel resources more modest than earlier anticipated. We now envision a system that: 1) includes unpublished USGS paleontology reports, and 2) archives and serves unpublished databases that have already been developed by USGS and other scientists.

PHASE TWO

Phase Two focuses on development of standards and guidelines needed to assist the USGS and state geological surveys in efficiently producing digital geologic maps, in a more standardized and common format. Our profession encourages innovation and individual pursuit of science, and so the question may be posed—why do we need these standards? Clearly, standards should not impede science but instead should help us efficiently communicate our science to the public. The need for communication was perhaps best articulated by former USGS Director John Wesley Powell, while planning for the new Geologic Atlas of the United States:

> *"… the maps are designed not so much for the specialist as for the people, who justly look to the official geologist for a classification, nomenclature, and system of convention so simple and expressive as to render his work immediately [understandable]…" (Powell, 1888).*

At that time, and throughout the early 20th century, Powell and others guided the USGS and the Nation's geoscientists toward a set of robust, practical standards for classifying geologic units and materials and representing them on maps. Those standards endured and evolved, and continue as basic guidelines for geologic mapping. Although today we commonly record in the field and laboratory far more complex information than during Powell's era, the necessity to provide it to the public in a standardized format remains unchanged. Newly evolving data formats and display techniques made feasible by computerization challenge us to revisit Powell's vision, and to develop standards and guidelines appropriate to today's technology and science.

In mid-1996, the NGMDB project and the AASG convened a meeting to identify the types of standards and guidelines that would improve the quality and utility of digital maps produced by the nation's geological surveys. From that meeting, Standards Working Groups were formed to address: 1) standard symbolization on geologic maps; 2) standard procedures for creating digital maps; 3) guidelines for publishing digital geologic maps; 4) documentation of methods and information via formal metadata; and 5) standard data structures and science terminology for geologic databases. The working group results will help provide a set of national standards to support public use of standard, seamless geologic map information for the entire country. In essence, Powell's pragmatic vision for the Geologic Atlas of the U.S. has been applied a century later to the National Geologic Map Database.

The tasks assigned to these Standards Working Groups are interrelated, as shown in Figure 6—when in the field, a geologist makes observations and (often, provisionally) draws geologic features on a base map; at that time, the accuracy with which these features are located on the map can be estimated. Further, the information may be recorded digitally in the field; if so, it can be structured similar to, or compatible with, the map database's structure (the "data model" in this figure). Returning to the office, the geologist commonly organizes and interprets field observations and prepares for map production—descriptions may be standardized according to an agency or project-level terminology or "science language," the map data may be structured according to the standard data model implemented by the agency, and procedures may be documented with metadata both in the office and when gathering data in the field. The descriptive information

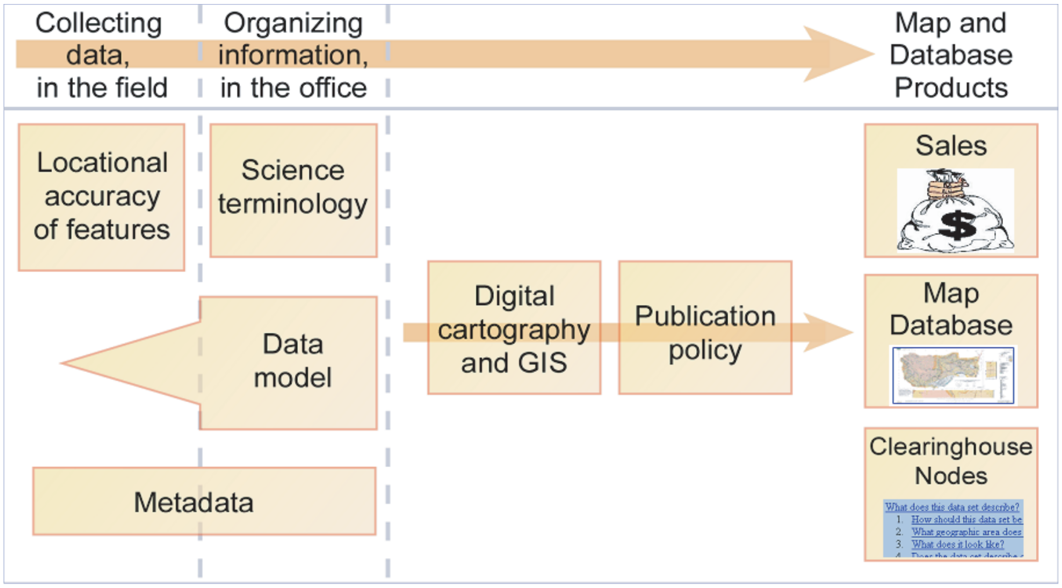


Figure 6. Diagram showing how the standards and guidelines under development by the NGMDB and related groups relate to the process of creating and publishing a map and database.

then is combined with the feature location information in a GIS, and digital cartography is applied to create a map that is published according to agency policies. Finally, the map is released to the public and accessed through various mechanisms including the NGMDB.

As described below, since 1996 these Working Groups and their successor organizations have made significant progress toward developing some of the necessary standards and guidelines. General information about the Working Groups and details of their activities are available at http://ngmdb.usgs.gov/info/standards/. Working Group members are listed in Appendix A.

Internationally, the NGMDB participates in venues that help to develop and refine the U.S. standards. These venues also bring our work to the international community, thereby promoting greater standardization with other countries. Examples include:

1. participation as a Council Member of the International Union of Geological Sciences’ Commission for the Management and Application of Geoscience Information (“IUGS CGI”; http://www.iugs.org/iugs/science/sci-cnfo.htm),
2. participation in the CGI Data Model Collaboration Working Group (http://www.bgs.ac.uk/cgi_web/tech_collaboration/data_model/data_model.html). This group is working on international standards for geologic information, to enable interoperability among national geological surveys, and
3. participation in “DIMAS”, the map standards committee of the Commission for the Geological Map of the World (see (Asch, 2003; and http://www.geology.cz/host/dimas.htm).

Geologic Map Symbolization

A draft standard for geologic map line and point symbology and map patterns and colors, published in a USGS Open-File Report in 1995, was reviewed in 1996 by the AASG, USGS, and Federal Geographic Data Committee (FGDC). It was revised by the NGMDB project team and members of the USGS Western Region Publications Group, and in late 1997 was circulated for internal review. The revised draft then was prepared as a proposed federal standard, for consideration by the FGDC. The draft was, in late 1999 through early 2000, considered and approved for public review by the FGDC and its Geologic Data Subcommittee. The document was released for public comment within the period May 19 through September 15, 2000 (see http://ngmdb.usgs.gov/fgdc_gds/mapsymb/ for the document and for information about the review process). This draft standard is described in some detail in Soller and Lindquist (2000). Based on public review comments, in 2002 a new section was added to the draft standard to address uncertainty in locational accuracy of map features. This section was presented for comment (Soller and others, 2002) and revised accordingly. With assistance from a Standing Committee to oversee resolution of review comments and long-term maintenance of the standard, the document has been prepared for submittal to FGDC, for final discussion and adoption as a Federal standard. This process is expected to conclude in 2006. Thereafter, the NGMDB with assistance from the

Standing Committee will maintain and, as needed, update the standard.

Digital Mapping

The Data Capture Working Group has coordinated nine annual "Digital Mapping Techniques" (DMT) workshops for state, federal, and Canadian geologists, cartographers, managers, and industry partners. These informal meetings serve as a forum for discussion and information-sharing, and have been quite successful. They have significantly helped the geoscience community converge on more standardized approaches for digital mapping and GIS analysis, and thus agencies have adopted new, more efficient techniques for digital map preparation, analysis, and production. In support of DMT workshops, an email listserver is maintained to facilitate the exchange of specific technical information.

The most recent DMT workshop, held in Baton Rouge, Louisiana, and hosted by the Louisiana Geological Survey, was attended by about 100 representatives of 47 state, federal, and Canadian agencies and private companies. Workshop Proceedings are published in paper format and online (see Appendix B and http://ngmdb.usgs.gov/info/dmt/). The website also provides: 1) a search mechanism for all Proceedings, by author, title, affiliation, and topic; and 2) downloadable presentations and posters from recent Proceedings. Copies of the printed Proceedings may be obtained from David Soller or Thomas Berg.

Map Publication Requirements

Through the USGS Geologic Division Information Council, the NGMDB led development of the USGS policy "Publication Requirements for Digital Map Products" (enacted May 24, 1999; see link under Map Publication Guidelines, at http://ngmdb.usgs.gov/info/standards/). A less USGS-specific version of this document was developed by the Data Information Exchange Working Group and presented for technical review at a special session of the Digital Mapping Techniques '99 workshop (Soller and others, 1999). The revised document (entitled "Proposed Guidelines for Inclusion of Digital Map Products in the National Geologic Map Database") was reviewed by the AASG Digital Geologic Mapping Committee. In 2002, it was unanimously approved via an AASG resolution, and has been incorporated as a guideline for digital map product deliverables to the STATEMAP component of the National Cooperative Geologic Mapping Program (see link under Map Publication Guidelines, at http://ngmdb.usgs.gov/info/standards/). The guideline also is recommended for participants in the Program's EDMAP component, which provides funding to university students to conduct geologic mapping.

Among the geological surveys there are many approaches to determining authorship credit and citation format for geologic maps, digital geologic maps, and associated digital databases. It is prudent for agencies to adopt policies that preserve the relationship of the geologist-authors to their product, the map image, and to identify the appropriate authorship (if any) and/or credit for persons responsible for creating the database files. A summary of this issue and a proposed guideline was outlined and discussed at the Digital Mapping Techniques workshop in 2001 (Berquist and Soller, 2001). This guideline stresses the importance of providing the suggested citation with each publication, and has proven useful to geological surveys as they attempt to balance responsibility and credit among field geologists, GIS specialists, and cartographers involved in creating a geologic map and database.

Metadata

The Metadata Working Group developed its final report in 1998. The report provides guidance on the creation and management of well-structured formal metadata for digital maps (see http:// ngmdb.usgs.gov/info/standards/metadata/metaWG.html). The report contains links to metadata-creation tools and general discussions of metadata concepts (see, for example, the metadata-creation tools, "Metadata in Plain Language," and other helpful information at http://geology.usgs.gov/tools/metadata/.

Geologic Map Data Model

In early 1999, with informal release of a draft version of a data model (Johnson and others, 1998), the Data Model Working Group concluded its work. The Group then was succeeded by the North American Geologic Map Data Model Steering Committee (NADMSC, http://nadm-geo.org, Figure 7). The NGMDB evaluated the draft data model, and developed in a prototype, object-relational database environment a data model that more effectively managed the geologic map information (Soller and others, 2002). This prototype was conducted in cooperation with the Kentucky Geological Survey, the Geological Survey of Canada, and the University of California—Santa Barbara.

Several prototypes, including the NGMDB (see "variants and implementations" at http://nadm-geo.org/dmdt/), provided the basis for the NADMSC to continue to refine its ideas. In 2004, this work produced two significant accomplishments: 1) a conceptual data model known as NADM C1 (NADM, 2004a; published simultaneously by the USGS and GSC), and 2) a draft standard terminology for earth materials (NADM, 2004b). State and USGS collaborators on the NGMDB continue to participate in the NADMSC, helping to further develop, refine, and test the NADM C1 and the science terminology that accompanies it.

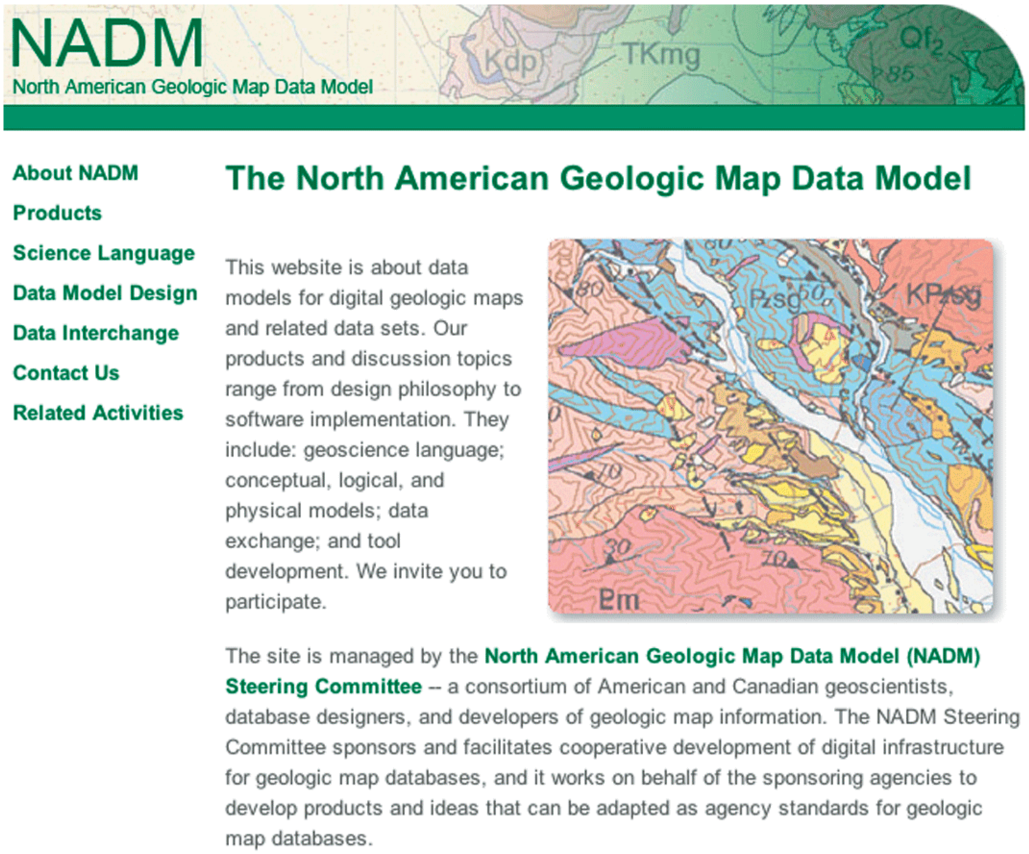


Figure 7. Website of the North American Geologic Map Data Model Steering Committee.

The NGMDB also is involved with the vendor community, for example through discussions with ESRI regarding their interest in defining an ArcGIS template or data model for geology, similar in concept to templates that ESRI has defined for other business sectors (see “geology” and other links at http://support.esri.com/index.cfm?fa=downloads.dataModels.gateway). We will continue to discuss this issue with ESRI, as we develop a database of map information for the NGMDB, (see discussion under “Phase Three”, below).

The NGMDB also contributes to development of international standards that will promote the management and interchange of geoscience information. This work is conducted under the aegis of the International Union of Geological Sciences’ Commission for the Management and Application of Geoscience Information (“IUGS CGI”; http://www.iugs.org/iugs/science/sci-cnfo.htm), specifically under its Data Model Collaboration Working Group (http://www.bgs.ac.uk/cgi_web/tech_collaboration/data_model/data_model.html). The NGMDB, and U.S. agencies in general, are benefiting significantly from this collaborative effort because:

1. our research and products are being tested and refined by numerous experts around the world, thereby improving their usefulness for the NGMDB, and
2. products and ideas developed by our Working Group colleagues can be directly applied to the NGMDB (e.g., concepts and technology developed by CSIRO Australia’s Exploration and Mining Markup Language project (“XMML”; http://www.seegrid.csiro.au/xmml).

PHASE THREE

Over the past few decades, significant advances in computer technology have begun to permit complex spatial information (especially vector-based) to be stored, managed, and analyzed for use by a growing number of geoscientists. At the beginning of the NGMDB project, we judged that computer-based mapping was not a sufficiently mature discipline to permit us to develop an online map database that addressed the scope mandated by the National Geologic Mapping Act. In particular, technology

for display and query of complex spatial information on the Web was in its infancy, and hence was not seriously considered by the NGMDB project as a viable means to deliver information to the general public. However, there now exists: 1) sufficient digital geologic map data; 2) sufficient convergence on standard data formats, data models, GIS and digital cartographic practices and field data capture techniques; and 3) sufficient technological advances in Internet delivery of spatial information to warrant a research effort for a prototype, online map database.

Before beginning to design this database, project personnel held numerous discussions with geoscientists and the general public to gauge interest in an online database and to define its scope. Based on these discussions, it was clear that this database should be:

1. built from edge-matched geologic maps at various scales;
2. managed and accessed as a coherent body of map information, not just as a set of discrete map products;
3. updated by mappers and/or a committee, "on the fly" when new information becomes available - it should be a "living" database;
4. standardized, adhering to a standard data model with standard scientific terminology; and
5. available to users via Web browsers and commonly available GIS tools.

This map database will integrate with other databases developed under the NGMDB project. For example, a user accessing the online map database might identify a map unit of interest, and then want to purchase or download the original published map product, or inquire about fossils found within that unit, or learn about the history of the geologic unit. Also, a user might access the Map Catalog and identify a map of interest, and then be linked to the online map database in order to browse and query it.

Prototyping

The NGMDB project is conducting a series of prototypes to advance our understanding of the technical and management challenges to developing the operational system; an introduction is given in Soller and others (2000). In 1999, we outlined some basic requirements for the prototype and tested them using map data for the greater Yellowstone area of Wyoming and Montana (Wahl and others, 2000). The second prototype (Soller and others, 2001) was conducted in cooperation with the Kentucky Geological Survey. In that prototype, we demonstrated in a commercial database system (GE-Smallworld; http://www.gepower.com/prod_serv/products/gis_software/en/smallworld4.htm) how the geologic database could be analyzed over the Web in concert with local datasets. The data model for the second prototype is described in Soller and others (2002), and was a significant contributor to the design of the new NADM Conceptual Data Model noted above.

Before proceeding further with plans for the online map database, we need to define a set of standardized terminology for the properties of earth materials (the "science language"). This must be sufficiently robust to accommodate terminology generated through today's field mapping, and terminology found in map unit descriptions on older and on smaller-scale maps, where descriptions tend to be highly generalized. In our current prototype we are creating richly-attributed map data with a standardized data model and science terminology. To achieve this, we have invested significantly in development of a data-entry software tool supported by science terminology derived from the NADMSC (see report by the NADM, 2004b).

The data-entry tool is being designed as a stand-alone application that will connect to a relational database that implements the NGMDB design (see Richard and others, 2004, and this volume). The tool will support: 1) development and editing of science vocabularies required by the NGMDB database implementation; 2) construction of formal descriptions for geologic units, earth materials, and geologic structure; and 3) the construction and editing of metadata to document the source and processing history of data. Because the NGMDB is envisioned as a distributed information system, with a variety of state and federal entities responsible for maintaining distinct bodies of data or repositories, the data-entry tool will include provisions for establishing data ownership and for maintaining access control based on user permissions for different repositories. Our priorities are to: 1) increase the number of science vocabularies developed or endorsed by the NGMDB and available through the data-entry tool; 2) develop an import and export functionality using the NADM GML interchange format; and 3) create an effective user interface.

What is a data model, and how does it apply to geologic maps?

A data model provides organization to the descriptive and spatial information that constitute a geologic map. The relations between a data model, science terminology, and the geologic map require some explanation. A data model may be highly conceptual, or it may describe the data structure for managing information within a specific hardware/software platform. In either case, it is a central construct because it addresses the database design for geologic maps in GIS format. In Figure 8, the data model is simplified to four locations, or "bins", where information can be stored, with each bin containing many database tables and fields:

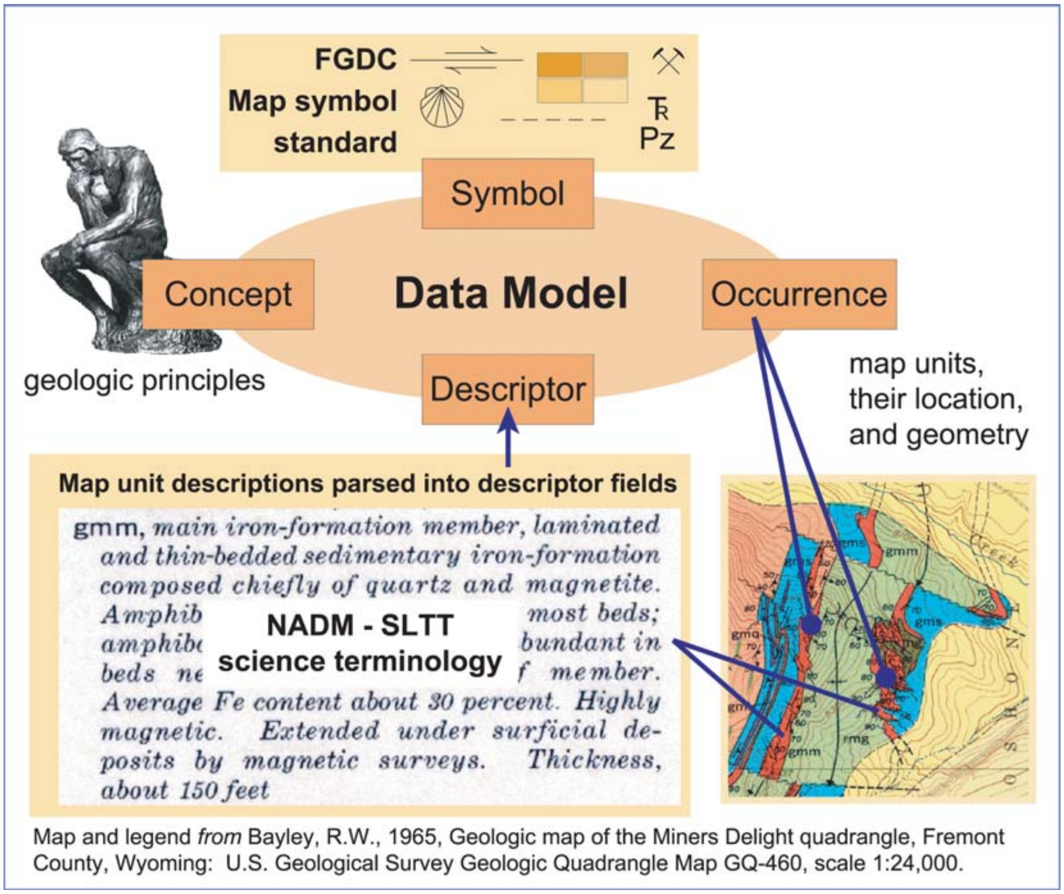


Figure 8. Simplified representation of the data model and its application to a typical, 2-D geologic map. The presence of a geologic unit on the map, referred to in the data model as an "occurrence" of that map unit, is described by: 1) its bounding contacts and faults, whose coordinates are stored as the unit's "geometry"; and 2) its physical properties, which are stored as the unit's "descriptors."

1. Occurrence—this bin contains the spatial geometry for each geologic feature in a map database. For example, the map unit identifier and the coordinates that define the outline of a map unit are included here.
2. Descriptor—this bin contains the wealth of descriptive information for each feature that occurs in the map database. This can include the full map unit description and simple attributes such as dominant lithology, color, and the nature of bedding.
3. Concept—this bin contains essential reference standards, such as geologic time scale(s) and science terminology. It also contains concepts and definitions essential for querying the database (for example, the concept that a rock can "intrude" another rock).
4. Symbol—this bin includes cartographic entities for symbolizing the map on-screen and in print form.

Will the U.S. have a single standard data model and science terminology?

The NGMDB online map database is envisioned as a distributed system that will provide seamless access to, and display of, map data served by many agencies. To achieve this vision, significant funding and time will be required. If all agencies used the same science terminology and exactly the same data model, and if it were implemented on the same hardware and software platform, building a functional system would be relatively straightforward. That, however, is not a realistic scenario.

Each agency has a unique history, set of objectives, and budget that will dictate the nature of their map database. (It should be noted that not every geological surveys in the U.S. can even afford to build such a system.) A more realistic approach is to assume a heterogenous computing environment, and to build software that can translate data structure and science terminology from one agency's system to another. This translation mechanism ensures "interoperability" between systems, and is the most realistic approach for the NGMDB. A prototype system developed by the U.S. GEON project (funded by the National Science Foundation) was discussed at DMT'03 (Ludascher and others, 2003).

To facilitate interoperability among systems, the NGMDB will define and maintain a set of reference standards (for data model, science terminology, geologic time scale) based in part on those produced by the NADMSC. Interoperability software that enables disparate systems to appear to the user as a single system is now being evaluated by groups including the NADMSC, NGMDB, GEON, and the IUGS's CGI. Through this technology, agencies should be able to correlate their unique data structure and scientific terminology to the reference standard, and translators (presumably GML-based) should enable us to display the information to the user in a single view.

Extending the data model to include three-dimensional (3-D) map information

The NADM C1 data model was designed for the typical geologic map, which provides a two-dimensional representation of the geologic framework. On most geologic maps, this framework is expressed generally, in cross-sections and map unit descriptions. The NGMDB project is exploring methods for incorporating a more complete depiction of geologic information in three dimensions, especially in raster (and voxel) format (Soller and Berg, 2003b). This 3-D information will be managed in the data model, which will require extensions to NADM.

National and regional map coverage

The online map database will be more useful if it includes some geologic map coverage for the entire nation. To that end, the NGMDB has supported compilation and GIS development of several regional maps. Most significant is the digital version of the "Geologic Map of North America". This map is the final product of the Geological Society of America's (GSA) Decade of North American Geology project. We provided funding and expertise for development of the digital files that were used to print the map, in order to develop the database for this map. With the map recently printed, we now will begin design of a prototype database, to be provided to GSA and the agencies that compiled the map (USGS, Geological Survey of Canada, and Woods Hole Oceanographic Institute) so that the plan for managing and serving the database can be developed.

OUTREACH TO LATIN AMERICAN GEOLOGICAL SURVEYS

We regularly meet with colleagues from other geological surveys, and especially with other Federal surveys in which similar work is being planned or conducted. The purpose of these meetings is to share information and to improve the design and quality of the databases and standards under development within the NGMDB and other agencies. To most geoscientists, the terminology and concepts of Information Technology and database design are relatively new and unfamiliar. Therefore, it can be especially difficult to convey the subtle meaning of these technical terms and concepts to colleagues who speak different languages. In an attempt to improve our communication with neighboring countries, the NGMDB project worked with the USGS/University of Arizona's Earth Surface Processes Research Institute (ESPRI), the University of Arizona's National Center for Interpretation, and the City of Tucson public schools to translate several reports from English to Spanish. The translated reports and a summary of the Spanish translation project are posted to http://ngmdb.usgs.gov/Info/reports/reports-esp.html. We hope that the translated reports will be of significant value. Before deciding whether to expand or discontinue this effort, we will evaluate the response to this website (Figure 9).

ACKNOWLEDGEMENTS

The authors thank the members of the NGMDB project staff and collaborators for their enthusiastic and expert support, without whom the project would not be possible. In particular, we thank: Ed Pfeifer, Alex Acosta, Dennis McMacken, Jana Ruhlman, and Michael Gishey (USGS, Flagstaff and Tucson, AZ; Website and database management), Chuck Mayfield (USGS; Map Catalog content), Robert Wardwell and Ben Carter (USGS, Vancouver, WA, and Reston, VA; Image Library), Steve Richard (Arizona Geological Survey, Tucson, AZ; data model and science terminology), and Jon Craigue (University of Arizona / USGS, Tucson, AZ; data-entry tool).

We also thank the many committee members who provided technical guidance and standards (Appendix A).

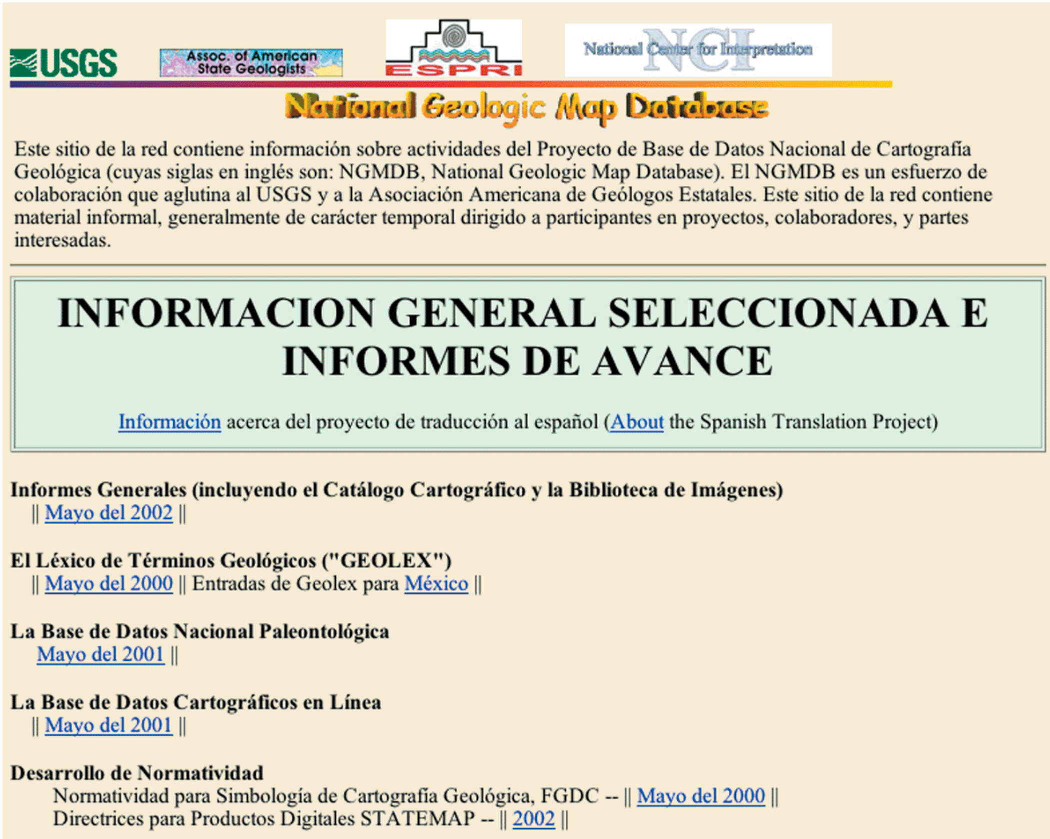


Figure 9. The NGMDB Spanish language website, containing translations of selected technical reports and geologic names of Mexico as found in GEOLEX.

REFERENCES

Asch, Kristine, 2003, CGMW Working Group on standards for digital geological data ("DIMAS"), *in* Soller, D.R., ed., Digital Mapping Techniques '03—Workshop proceedings: U.S. Geological Survey Open-file Report 03-471, p. 193-194, accessed at http://pubs.usgs.gov/of/2003/of03-471/asch/.

Berquist, C.R., Jr., and Soller, D.R., 2001, Map authorship and citation guidelines: Summary of a discussion session, *in* Soller, D.R., ed., Digital Mapping Techniques '01—Workshop proceedings: U.S. Geological Survey Open-file Report 01-223, p. 159-164, accessed at http://pubs.usgs.gov/of/2001/of01-223/berquist.html.

Johnson, B.R., Brodaric, B., Raines, G.L., Hastings, J., and Wahl, R., 1998, Digital geologic map data model v. 4.3: AASG/USGS Data Model Working Group Report, accessed at http://nadm-geo.org.

Ludascher, Bertram, Lin, Kai, Brodaric, Boyan, and Baru, Chaitan, 2003, GEON: Toward a cyberinfrastructure for the geosciences—A prototype for geologic map integration via domain ontologies, *in* Soller, D.R., ed., Digital Mapping Techniques '03—Workshop Proceedings: U.S. Geological Survey Open-File Report 03-471, p. 223-229, accessed at http://pubs.usgs.gov/of/2003/of03-471/ludascher/.

Mac Lachlan, M.E., and others, compilers, 1996, Stratigraphic nomenclature databases for the United States, its possessions, and territories: U.S. Geological Survey Digital Data Series DDS-6, Release 3, one CD-ROM.

North American Geologic Map Data Model (NADM) Steering Committee Data Model Design Team, 2004a, NADM Conceptual Model 1.0—A Conceptual Model for Geologic Map Information: U.S. Geological Survey Open-File Report 2004-1334, 60p., accessed at http://pubs.usgs.gov/of/2004/1334/.

North American Geologic Map Data Model (NADM) Steering Committee Science Language Technical Team, 2004b, Report on progress to develop a North American science-language standard for digital geologic-map databases, *in* Soller, D.R., ed., Digital Mapping Techniques '04—Workshop Proceedings: U.S. Geological Survey Open-File Report 2004-1451, p. 85-94, accessed at http://pubs.usgs.gov/of/2004/1451/nadm/.

Powell, J.W., 1888, Methods of geologic cartography in use by the United States Geological Survey, *in* Congrès Géologique International, Compte Rendu de la 3me Session, Berlin, 1885: A.W. Schade's Buchdruckerei, Berlin, p. 221–240.

Richard, S.M., Craigue, Jon, and Soller, D.R., 2004, Implementing NADM C1 for the National Geologic Map Database, *in* Soller, D.R., ed., Digital Mapping Techniques '04—Workshop Proceedings: U.S. Geological Survey Open-File Report 2004-1451, p. 111-144, accessed at http://pubs.usgs.gov/of/2004/1451/richard/.

Soller, D.R., and Berg, T.M., 2004, The National Geologic Map Database project: Overview and progress, *in* Soller, D.R., ed., Digital Mapping Techniques '04—Workshop Proceedings: U.S. Geological Survey Open-File Report 2004-1451, p. 15-31, accessed at http://pubs.usgs.gov/of/2004/1451/soller/.

Soller, D.R., and Berg, T.M., 2003, The National Geologic Map Database Image Library—General concepts, *in* Soller, D.R., ed., Digital Mapping Techniques '03—Workshop Proceedings: U.S. Geological Survey Open-File Report 03-471, p. 89-93, accessed at http://pubs.usgs.gov/of/2003/of03-471/soller2/.

Soller, D.R., and Lindquist, Taryn, 2000, Development and public review of the draft "Digital cartographic standard for geologic map symbolization", *in* Soller, D.R., ed., Digital Mapping Techniques '00—Workshop Proceedings: U.S. Geological Survey Open-File Report 00-325, p. 43-47, accessed at http://pubs.usgs.gov/of/of00-325/soller3.html.

Soller, D.R., Berg, T.M., and Wahl, Ron, 2000, Developing the National Geologic Map Database, phase 3—An online, "living" database of map information, *in* Soller, D.R., ed., Digital Mapping Techniques '00—Workshop Proceedings: U.S. Geological Survey Open-File Report 00-325, p. 49-52, accessed at http://pubs.usgs.gov/of/of00-325/soller4.html.

Soller, D.R., Brodaric, Boyan, Hastings, J.T., Wahl, Ron, and Weisenfluh, G.A., 2002, The central Kentucky prototype: An object-oriented geologic map data model for the National Geologic Map Database: U.S. Geological Survey Open-file Report 02-202, 38 p., accessed at http://pubs.usgs.gov/of/2002/of02-202/.

Soller, D.R., Duncan, Ian, Ellis, Gene, Giglierano, Jim, and Hess, Ron, 1999, Proposed guidelines for inclusion of digital map products in the National Geologic Map Database, *in* Soller, D.R., ed., Digital Mapping Techniques '99—Workshop Proceedings: U.S. Geological Survey Open-File Report 99-386, p. 35-38, accessed at http://pubs.usgs.gov/of/of99-386/soller2.html.

Soller, D.R., Lindquist, T.A., and Matti, J.C., 2002, Field description of the scientific and locational accuracy of geologic features (a part of the draft FGDC geologic map symbolization standard), *in* Soller, D.R., ed., Digital Mapping Techniques '02—Workshop Proceedings: U.S. Geological Survey Open-File Report 02-370, p. 33-40, accessed at http://pubs.usgs.gov/of/2002/of02-370/soller1.html.

Soller, D.R., Wahl, Ron, Weisenfluh, Jerry, Brodaric, Boyan, Hastings, Jordan, Laudati, Robert, and Fredericks, Roger, 2001, Progress report on the National Geologic Map Database, phase 3: An online database of map information, *in* Soller, D.R., ed., Digital Mapping Techniques '01—Workshop proceedings: U.S. Geological Survey Open-file Report 01-223, p. 71-78, accessed at http://pubs.usgs.gov/of/2001/of01-223/soller2.html.

Stamm, N.R., Wardlaw, B.R., and Soller, D.R., 2000, GEOLEX—The National Geologic Map Database's geologic names lexicon, *in* Soller, D.R., ed., Digital Mapping Techniques '00—Workshop Proceedings: U.S. Geological Survey Open-File Report 00-325, p. 31-42, accessed at http://pubs.usgs.gov/of/2000/of00-325/stamm.html.

Wahl, R.R., Soller, D.R., and Yeldell, Steven, 2000, Prototype Implementation of the NADMSC Draft Standard Data Model, Greater Yellowstone Area, *in* Soller, D.R., ed., Digital Mapping Techniques '00—Workshop Proceedings: U.S. Geological Survey Open-File Report 00-325, p. 57-63, accessed at http://pubs.usgs.gov/of/of00-325/wahl.html.

Wardlaw, B.R., Stamm, N.R., and Soller, D.R., 2001, The U.S. Geological Survey National Paleontological Database, *in* Soller, D.R., editor, Digital Mapping Techniques '01—Workshop proceedings: U.S. Geological Survey Open-file Report 01-223, p. 59-70, accessed at http://pubs.usgs.gov/of/2001/of01-223/wardlaw.html.

APPENDIX A

Principal committees and people collaborating with the National Geologic Map Database project.

Digital Geologic Mapping Committee of the Association of American State Geologists:
Tom Berg (Ohio Geological Survey and Committee Chair)
Rick Allis (Utah Geological Survey)
Larry Becker (Vermont Geological Survey)
Rick Berquist (Virginia Division of Mineral Resources)
Jim Cobb (Kentucky Geological Survey)
Ian Duncan (Texas Bureau of Economic Geology)
Rich Lively (Minnesota Geological Survey)
Jay Parrish (Pennsylvania Geological Survey)
Bill Shilts (Illinois State Geological Survey)
Nick Tew (Alabama Geological Survey)
Harvey Thorleifson (Minnesota Geological Survey)

Geologic Data Subcommittee of the Federal Geographic Data Committee:
Dave Soller (U.S. Geological Survey and Subcommittee Chair)
Jerry Bernard (USDA-Natural Resources Conservation Service)
Mark Crowell (Dept. of Homeland Security, Federal Emergency Mgmt. Agency)
Jim Gauthier-Warinner (U.S. Forest Service, Minerals and Geology Management)
Laurel T. Gorman (U.S. Army Engineer Research and Development Center)
John L. LaBrecque (National Aeronautics and Space Administration)
Lindsay McClelland (National Park Service)
Jay Parrish (State Geologist, Pennsylvania Geological Survey)
George F. Sharman (NOAA National Geophysical Data Center)
Dave Zinzer (Minerals Management Service)

Map Symbol Standards Committee:
Dave Soller (U.S. Geological Survey and Committee Coordinator)
Tom Berg (State Geologist, Ohio Geological Survey)
Bob Hatcher (University of Tennessee, Knoxville)
Mark Jirsa (Minnesota Geological Survey)
Taryn Lindquist (U.S. Geological Survey)
Jon Matti (U.S. Geological Survey)
Jay Parrish (State Geologist, Pennsylvania Geological Survey)
Jack Reed (U.S. Geological Survey)
Steve Reynolds (Arizona State University)
Byron Stone (U.S. Geological Survey)

AASG/USGS Data Capture Working Group:
Dave Soller (U.S. Geological Survey and Working Group Chair)
Warren Anderson (Kentucky Geological Survey)
Rick Berquist (Virginia Geological Survey)
Elizabeth Campbell (Virginia Division of Mineral Resources)
Rob Krumm (Illinois State Geological Survey)
Scott McCulloch (West Virginia Geological and Economic Survey)
Gina Ross (Kansas Geological Survey)
George Saucedo (California Geological Survey)
Barb Stiff (Illinois State Geological Survey)
Tom Whitfield (Pennsylvania Geological Survey)

DMT Listserve:
Maintained by Doug Behm, University of Alabama

North American Data Model Steering Committee:
Dave Soller (U.S. Geological Survey and Committee Coordinator)
Tom Berg (Ohio Geological Survey)
Boyan Brodaric (Geological Survey of Canada and Chair of the Data Model Design Technical Team)
Peter Davenport (Geological Survey of Canada)
Bruce Johnson (U.S. Geological Survey and Chair of the Data Interchange Technical Team)
Rob Krumm (Illinois State Geological Survey)
Scott McColloch (West Virginia Geological and Economic Survey)
Steve Richard (Arizona Geological Survey)
Loudon Stanford (Idaho Geological Survey)
Jerry Weisenfluh (Kentucky Geological Survey)

IUGS Commission for the Management and Application of Geoscience Information:
Dave Soller (U.S. Geological Survey, Council Member)

Conceptual model/Interchange Task Group (of the Data Model Collaboration Working Group of the IUGS Commission for the Management and Application of Geoscience Information):
Steve Richard (Arizona Geological Survey, Task Group Member)

DIMAS (Digital Map Standards Working Group of the Commission for the Geological Map of the World):
Dave Soller (U.S. Geological Survey, Working Group Member)

NGDMB contact-persons in each State geological survey:

These people help the NGMDB with the Geoscience Map Catalog, GEOLEX, the Geologic Map Image Library, and the Mapping in Progress Database. Please see http://ngmdb.usgs.gov/info/statecontacts.html for this list.

These groups have fulfilled their mission and are no longer active:

NGMDB Technical Advisory Committee:

Boyan Brodaric (Geological Survey of Canada)
David Collins (Kansas Geological Survey)
Larry Freeman (Alaska Division of Geological & Geophysical Surveys)
Jordan Hastings (University of California, Santa Barbara)
Dan Nelson (Illinois State Geological Survey)
Stephen Richard (Arizona Geological Survey)
Jerry Weisenfluh (Kentucky Geological Survey)

AASG/USGS Metadata Working Group:

Peter Schweitzer (U.S. Geological Survey and Working Group Chair)
Dan Nelson (Illinois State Geological Survey)
Greg Hermann (New Jersey Geological Survey)
Kate Barrett (Wisconsin Geological and Natural History Survey)
Ron Wahl (U.S. Geological Survey)

AASG/USGS Data Information Exchange Working Group:

Dave Soller (U.S. Geological Survey and Working Group Chair)
Ron Hess (Nevada Bureau of Mines and Geology)
Ian Duncan (Virginia Division of Mineral Resources)
Gene Ellis (U.S. Geological Survey)
Jim Giglierano (Iowa Geological Survey)

AASG/USGS Data Model Working Group:

Gary Raines (U.S. Geological Survey and Working Group Chair)
Boyan Brodaric (Geological Survey of Canada)
Jim Cobb (Kentucky Geological Survey)
Ralph Haugerud (U.S. Geological Survey)
Greg Hermann (New Jersey Geological Survey)
Bruce Johnson (U.S. Geological Survey)
Jon Matti (U.S. Geological Survey)
Jim McDonald (Ohio Geological Survey)
Don McKay (Illinois State Geological Survey)
Steve Schilling (U.S. Geological Survey)
Randy Schumann (U.S. Geological Survey)
Bill Shilts (Illinois State Geological Survey)
Ron Wahl (U.S. Geological Survey)

APPENDIX B

List of progress reports on the National Geologic Map Database, and Proceedings of the Digital Mapping Techniques workshops.

Soller, D.R., editor, 2004, Digital Mapping Techniques '04—Workshop Proceedings: U.S. Geological Survey Open-File Report 2004-1451, 220 p., accessed at http://pubs.usgs.gov/of/2004/1451/.

Soller, D.R., editor, 2003, Digital Mapping Techniques '03—Workshop Proceedings: U.S. Geological Survey Open-File Report 03-471, 262 p., accessed at http://pubs.usgs.gov/of/2003/of03-471/.

Soller, D.R., editor, 2002, Digital Mapping Techniques '02—Workshop Proceedings: U.S. Geological Survey Open-File Report 02-370, 214 p., accessed at http://pubs.usgs.gov/of/2002/of02-370/.

Soller, D.R., editor, 2001, Digital Mapping Techniques '01—Workshop Proceedings: U.S. Geological Survey Open-File Report 01-223, 248 p., accessed at http://pubs.usgs.gov/of/2001/of01-223/.

Soller, D.R., editor, 2000, Digital Mapping Techniques '00—Workshop proceedings: U.S. Geological Survey Open-file Report 00-325, 209 p., accessed at http://pubs.usgs.gov/of/of00-325/.

Soller, D.R., editor, 1999, Digital Mapping Techniques '99—Workshop proceedings: U.S. Geological Survey Open-file Report 99-386, 216 p., accessed at http://pubs.usgs.gov/of/of99-386/front.html.

Soller, D.R., editor, 1998, Digital Mapping Techniques '98—Workshop Proceedings: U.S. Geological Survey Open-File Report 98-487, 134 p., accessed at http://pubs.usgs.gov/of/of98-487/.

Soller, D.R., editor, 1997, Proceedings of a workshop on digital mapping techniques: Methods for geologic map data capture, management, and publication: U.S. Geological Survey Open-File Report 97-269, 120 p., accessed at http://pubs.usgs.gov/of/of97-269/.

Soller, D.R., Berg, T.M., and Stamm, N.R., 2004, The National Geologic Map Database project: Overview and progress, *in* Soller, D.R., ed., Digital Mapping Techniques '04—Workshop Proceedings: U.S. Geological Survey Open-File Report 2005-1451, p.15-31, accessed at http://pubs.usgs.gov/of/2004/1451/soller/.

Soller, D.R., and Berg, T.M., 2003, The National Geologic Map Database project: Overview and progress, *in* Soller, D.R., ed., Digital Mapping Techniques '03—Workshop Proceedings: U.S. Geological Survey Open-File Report 03-471, p. 57-77, accessed at http://pubs.usgs.gov/of/2003/of03-471/soller1/.

Soller, D.R., and Berg, T.M., 2002, The National Geologic Map Database: A progress report, *in* Soller, D.R., editor, Digital Mapping Techniques '02—Workshop proceedings: U.S. Geological Survey Open-file Report 02-370, p. 75-83, accessed at http://pubs.usgs.gov/of/2002/of02-370/soller2.html.

Soller, D.R., and Berg, T.M., 2001, The National Geologic Map Database--A progress report, *in* Soller, D.R., editor, Digital Mapping Techniques '01—Workshop proceedings: U.S. Geological Survey Open-file Report 01-223, p. 51-57, accessed at http://pubs.usgs.gov/of/2001/of01-223/soller1.html.

Soller, D.R., and Berg, T.M., 2000, The National Geologic Map Database--A progress report, *in* Soller, D.R., editor, Digital Mapping Techniques '00—Workshop proceedings: U.S. Geological Survey Open-file Report 00-325, p. 27-30, accessed at http://pubs.usgs.gov/of/of00-325/soller2.html.

Soller, D.R., and Berg, T.M., 1999a, Building the National Geologic Map Database: Progress and challenges, *in* Derksen, C.R.M, and Manson, C.J., editors, Accreting the continent's collections: Geoscience Information Society Proceedings, v. 29, p. 47-55, accessed at http://ngmdb.usgs.gov/info/reports/gisproc98.html.

Soller, D.R., and Berg, T.M., 1999b, The National Geologic Map Database—A progress report, *in* Soller, D.R., editor, Digital Mapping Techniques '99—Workshop proceedings: U.S. Geological Survey Open-file Report 99-386, p. 31-34, accessed at http://pubs.usgs.gov/of/of99-386/soller1.html.

Soller, D.R., and Berg, T.M., 1998, Progress Toward Development of the National Geologic Map Database, *in* Soller, D.R., editor, Digital Mapping Techniques '98—Workshop proceedings: U.S. Geological Survey Open-file Report 98-487, p. 37-39, accessed at http://pubs.usgs.gov/of/of98-487/soller2.html.

Soller, D.R., and Berg. T.M., 1997, The National Geologic Map Database—A progress report: Geotimes, v. 42, no. 12, p. 29-31, accessed at http://ngmdb.usgs.gov/info/reports/geotimes97.html.

Soller, D.R., and Berg, T.M., 1995, Developing the National Geologic Map Database: Geotimes, v. 40, no. 6, p. 16-18, accessed at http://ngmdb.usgs.gov/info/reports/geotimes95.html.

Assessing the Status of Geologic Map Coverage of the United States—A New Application of the National Geologic Map Database

By David R. Soller

U.S. Geological Survey
926-A National Center
Reston, VA 20192
Telephone: (703) 648-6907
Fax: (703) 648-6977
e-mail: drsoller@usgs.gov

INTRODUCTION

Systematic geologic mapping of the United States has been conducted for more than 125 years. In the period centered on 1895-1920, the USGS conducted the first such program, the Geologic Atlas of the United States. The Atlas included about 230 products, at scales ranging from 1:14,400 to 1:250,000 (Figure 1). It is notable that the scientific and cartographic standards developed to guide that mapping (Powell, 1888) have, with modest revision, endured to this day.

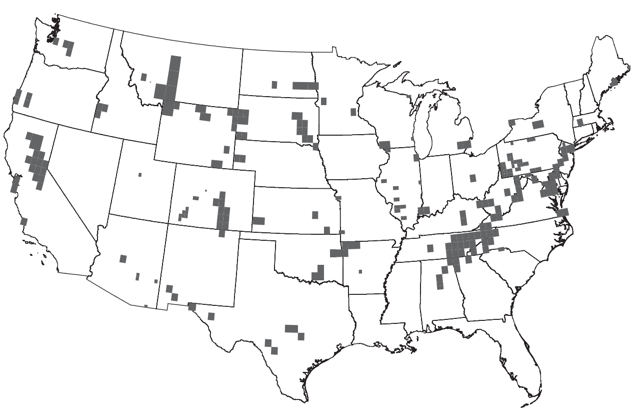


Figure 1. Index map of the USGS series "Geologic Atlas of the United States" (ca. 1895-1920). Information derived from the NGMDB Map Catalog.

In more recent times, geologic mapping has been conducted under many programs in Federal, State, and other agencies. These programs have differed in emphasis owing to funding source and time period; partly as the result, geologic maps vary significantly in content and format. These differences present a real challenge to the preparation of index maps that purport to show geologic map coverage – if all geologic maps are not alike in content, scale, detail, vintage, or currency, which then should be included in an index map? By what criteria should we differentiate or classify geologic maps for this purpose?

Purpose

The principal and most obvious purpose for index maps is to convey to any user, whether a practicing geologist or a homeowner, the availability of published geologic maps. Also, the Nation's geological surveys need to know the areas for which geologic maps have been made, and when those maps were published. Such information helps each agency to prioritize areas that should be mapped in the future (or remapped, usually in more detail). Furthermore, it demonstrates to legislators and oversight agencies such as the U.S. Office of Management and Budget (OMB) that funding has produced tangible results.

In response to requests from the Association of American State Geologists (AASG) and the USGS National Cooperative Geologic Mapping Program (NCGMP), the National Geologic Map Database project (NGMDB) in 2005 developed the capability to generate index maps showing geologic map coverage at different scales and for various time periods. The information supplied by the NGMDB (i.e., index maps and numeric summaries of the extent of intermediate and large-scale geologic mapping in the U.S.) was needed in order to fulfill two OMB requirements for NCGMP "performance metrics". These metrics served as partial documentation of AASG and USGS performance in addressing the goals of the Geologic Mapping Act (http://ncgmp.usgs.gov/ncgmpabout/ngmact/). It is anticipated that such information will be required annually.

Ideally, these index maps would be created from a database designed specifically for the task. However, no such database existed. For the AASG and NCGMP,

therefore, the logical choice was to extract information from the NGMDB Geoscience Map Catalog (available at http://ngmdb.usgs.gov/). The Catalog is a general-purpose database containing bibliographic records on more than 75,000 geologic maps and other geoscience maps and reports for the U.S., published by more than 350 organizations. Based on new AASG and NCGMP requirements for this information, and on the general informative value of index maps, the NGMDB project is now endeavoring to revise its database to accommodate agency requests for such information, as described below.

METHOD

The request from AASG and NCGMP was to provide index maps showing the location of "modern", general-purpose geologic maps of intermediate and detailed scale. As noted in the Introduction, geologic maps are not all alike in content, scale, detail, vintage, or currency. Which maps should be considered modern, which are general purpose, and what is an intermediate or detailed scale? It was decided for this purpose that "modern" maps would be somewhat arbitrarily defined as those published since 1959. Because the NCGMP was created to address the goals of the Geologic Mapping Act of 1992, a secondary objective was to identify maps produced since that date. General-purpose geologic maps are those that include all geologic units in the map area and that focus on geologic history and the characteristics of the materials (for example, the typical geologic map of a quadrangle or county). The decision regarding map scales appropriate to portray was a difficult one, and so for the initial set of index maps, all maps of scale 1:250,000 and more detailed were included. Upon inspection of these index maps it became clear that the scale should be more restricted, to intermediate-scale (defined for this project as 1:100,000) and more detailed maps. Decisions such as this, necessary but arbitrary in nature, resulted in the omission of numerous useful geologic maps (e.g., county maps at scales approximately 1:125,000).

Step 1 – Selecting records to evaluate

With these definitions providing a constraint, the process of creating the index maps began with a query of the Map Catalog's "theme" field, in order to identify all bibliographic records that contain bedrock or surficial geologic information. The problem with this approach was immediately evident – although each of these products contain geologic map information, many could not legitimately be described as general-purpose geologic maps. This problem was unavoidable because no field more relevant than "theme" was available.

Each publication in the Map Catalog is assigned one or more geologic themes that describe its content (see Figure 2 and http://ngmdb.usgs.gov/ngmdb/define.html). For example, consider a publication that addresses landslide hazards – the landslide hazard potential, or the surveyed landslides, commonly are shown on a geologic map in order to provide context for these features. The geologic map may be newly-developed by the landslide-mapping project, or it may have been reproduced in full or reduced detail from a map originally released in another publication. In the Map Catalog database, the landslide map would be assigned the geologic themes "Landslides" and either "Bedrock Geology" or "Surficial Geology" because it contains a geologic map. The purpose of the "themes" field is to assist the user in finding the type(s) of maps they need without omitting from the database search any publications that could be useful. When this database was under development (ca. 1995) this was recognized as an important feature – it returns to the user a list of all maps that might possibly be useful, thereby giving the user the opportunity to choose from a larger set of products than would be possible if only the principal theme of the product were recorded in the database.

The search for "Theme=bedrock or surficial geology" yielded 28,100 publications. Bibliographic information and bounding coordinates for each publication were exported from the Map Catalog's Oracle database to a .DBF (v. 4) file. To simplify the handling of these records,

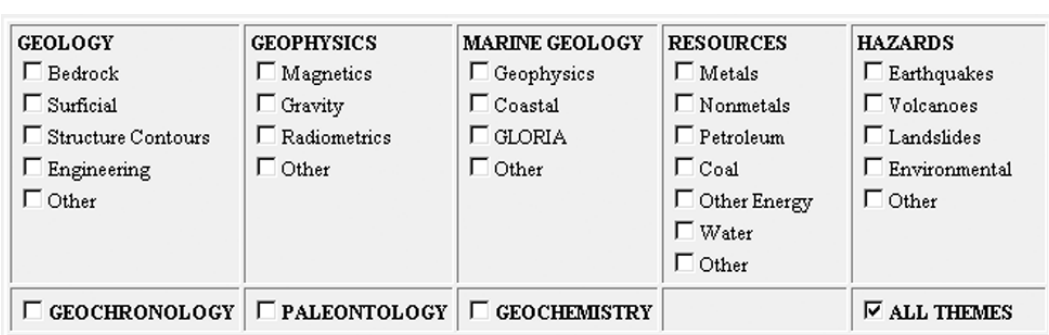


Figure 2. Geoscience themes in the Map Catalog.

the publications were divided into four files, by scale range (roughly, 1:24,000 and more detailed; 1:25,000 – 1:99,000; 1:100,000; and 1:101,00 – 1:251,000).

Step 2 – Creating the “footprint” of each map

To simplify the entry and management of bibliographic records, and to minimize data-entry errors, the Map Catalog database contains the bounding coordinates (actually, the NW and SE corners) of each map publication rather than the actual extent of the mapped area. For maps that fill a rectilinear area (e.g., a quadrangle), the bounding coordinates accurately represent the mapped area. However, for a map of an irregular area (e.g., a “strip map” of the geology along the Chesapeake and Ohio Canal in D.C., WV, VA, and MD), the bounding coordinates can drastically over-depict the area mapped (Figure 3). Therefore, it was decided that maps of such irregular areas should not be shown by the index maps; differentiating these maps from more rectilinear maps proved difficult, as explained in Step 5, below.

To prepare the selected records for evaluation and display, each .DBF file was imported into ArcView, using an Avenue script written by Chris Garrity (USGS) that converted each bounding coordinate pair to a polygon. Each of the four ArcView shapefiles were visually evaluated for obvious errors in the data-entry of bounding coordinates or map scale (e.g., a map of a large area that, while specified as 1:100,000, is obviously of a smaller scale, such as 1:1,000,000). This visual check was found to be a valuable supplement to the automated, logical checks performed on each newly-added Catalog record, and will therefore be included in database error-checking procedures. Erroneous coordinates and map scales were corrected, and the file then was reimported to ArcView, to be rechecked until no errors were visually detected.

Step 3 – Exploring the data

As may be apparent from Steps 1 and 2, index maps could not simply be generated by selecting the map theme “bedrock geology” or “surficial geology” and then displaying the outline of each map’s bounding box. Instead, each publication had to be evaluated for suitability. This evaluation focused on the most informative field for this purpose – publication title. From the title, I hoped to differentiate:

1. general-purpose versus more specialized geologic maps (e.g., “Revision of Middle Proterozoic Yellowjacket Formation, Central Idaho”), and
2. maps of quadrangles and counties (and parishes and boroughs) versus maps of more irregular outline.

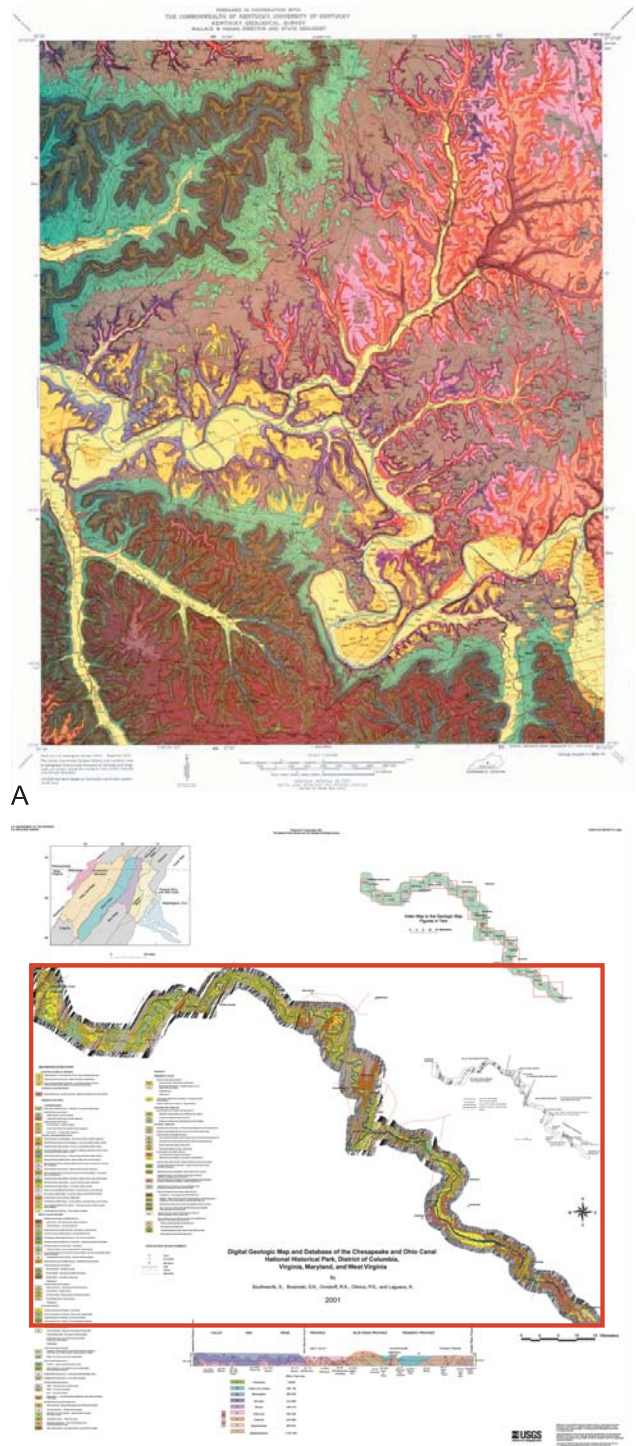


Figure 3. Comparison of the area that was geologically mapped versus the map area as described by its bounding coordinates.
A) a typical quadrangle map (Kepferle, 1973); here, the mapped area and the bounding box are the same, and so the box is a valid estimate of the mapped area.
B) a geologic map of an irregular area (Southworth et al., 2001); the bounding box (thick, dark line) is not a valid estimate of the mapped area.

Before attempting to evaluate the titles, several days were allotted to peruse the data, seemingly at random, in order to become more familiar with the content. The importance of this can scarcely be over-emphasized – if I had immediately "waded into" the evaluation process, relying on my assumptions about how to efficiently proceed, the result would have been far less accurate in portraying the status of geologic map coverage.

Step 4 – Defining the evaluation criteria

From Step 3 it was verified that useful information regarding map content and shape (i.e., general-purpose versus specialized geologic map; quadrangle versus irregular shape) could be interpreted from the publication titles. Because titles vary greatly in format and information content, and because any such interpretation inherently carries some degree of uncertainty, this process represents a necessary compromise dictated by the reality that time and effort could not be expended in order for each publication to be reviewed by the person or agency most familiar with its content.

Two database fields were added to the Map Catalog: "mapshape" and "include"; the latter considered map shape and, further, indicated whether a map would be included in the index maps. The mapshape field included these values:

- 4 = quadrangle
- 3 = county, parish, or borough
- 2 = irregular (rarely used; evaluated by finding and then inspecting the map)
- 1 = probably irregular (evaluated based on the publication title)
- 0 = not evaluated, because not a geologic map.

For the "include" field, titles pertaining to general-purpose maps of quadrangles and counties were identified and assigned to be included on the index maps. Specialized geologic maps and general-purpose geologic maps of irregular areas were assigned to be not included, with an exception: maps of islands, large coastal areas, and major parks, where quadrangle and county maps were unavailable, were assigned a value that enabled them to be considered for inclusion on the index maps. As noted in the concluding section of this paper, because selected irregularly-shaped maps were included, the area that has been geologically mapped was somewhat overestimated; this error will be minimized in future versions of these index maps.

The "include" field allowed these values:

- 4 = general-purpose geologic map of a quadrangle or county
- 3 = general-purpose geologic map of an irregular area such as an island, a coastal region, or a major park
- 2 = general-purpose geologic map of an irregular area
- 1 = specialized geologic map, or not a geologic map but containing geologic content
- 0 = does not contain geologic map information.

Examples of maps assigned to these two categories are provided in Figure 4.

Step 5 – Evaluating each map

The most efficient method for identifying appropriate maps was found to be a systematic query of map titles, in ArcView; the query searched for keywords associated with general-purpose geologic quadrangle maps or more specialized geologic maps. For example, a query might search for titles containing text strings such as "geologic map" and "quadrangle", and include these maps. Conversely, a query might search for terms that commonly are applied to specialized maps or those of irregular outline, such as "formation" or "range" or "district", and omit these. All maps selected by a query then were assigned values for the two new fields described above.

To a significant degree this systematic and logical approach identified the appropriate maps, and it greatly expedited the assignment of this new information to each publication. However, through typical error-checking procedures it became clear that this approach caused to be omitted many maps that should have been included (e.g., those that used the term "sheet" or "folio" rather than "quadrangle", or those whose quadrangle name included an omitted term such as "range"). Conversely, the process caused to be included many maps that should have been omitted (e.g., a map with a title something like "Geologic map of gold-bearing rocks within the XYZ quadrangle"). Because I could not be assured that all titles had been correctly interpreted by these queries, each title then was inspected in order to identify publications that likely had been assigned to the incorrect "include" or "mapshape" category. Corrections were made, and the ArcView files were then ready for preparation of index maps. Of the 28,100 publications originally selected, 15,026 were identified according to the criteria listed above. Because it was later determined that we would show only the maps of scale 1:100,000 and larger (i.e., more detailed), the number of relevant publications was further reduced, to 13,597.

Step 6 – The index maps

Following discussions with the AASG and NCGMP, maps classified with an "include" value of 3 or 4 were

A

Include?		Title
	2	Geology and ore deposits of the Afterthought mine, Shasta County, California
	4	Geology and mineral resources of Jones Ranch School quadrangle, McKinley County, New Mexico
	4	Geology and coal resources of Vanderwagen quadrangle, McKinley County, New Mexico
	4	Geologic map of the Lake Helen quadrangle, Big Horn and Johnson counties, Wyoming
	4	Preliminary geologic map of the Lillis Ranch quadrangle, California
NO	2	Geology of the Tejon Hills area - Arvin and Tejon Hills quadrangles, Kern County, California
NO	1	Stratigraphy, structure, and mineral deposits in the Oro Grande series near Victorville, California
YES	4	Bedrock geologic map of the Woodford quadrangle, Bennington and Windham Counties, Vermont
YES	3	Bedrock geologic map of Yosemite Valley, Yosemite National Park, California
	4	Geologic map of the Vidal, California, and Parker SW, California-Arizona quadrangles
	4	Geologic map of the Vidal NW, Vidal Junction, and parts of the Savahia Peak SW and Savahia Peak

B C

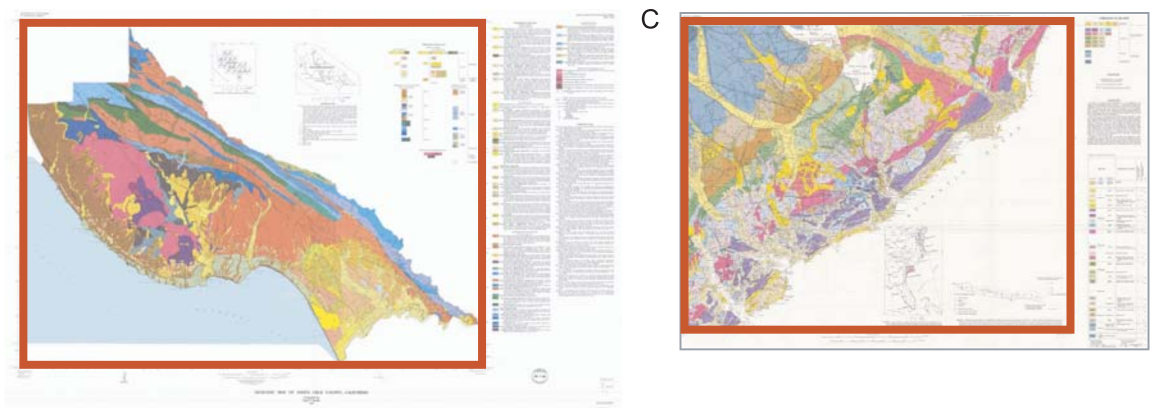


Figure 4. Examples of products included on the index maps.
A) inferring the type of geologic map, based on title ("yes" = included on index maps; refer to values for the "include" field in Step 4). Example of a general-purpose quadrangle geologic map is shown in Figure 3A.
B) general-purpose county geologic map (Brabb, 1989). Bounding box shown by thick, dark line.
C) general-purpose geologic map of irregular area (McCartan et al., 1984). Bounding box shown by thick, dark line. See Figure 3B for example of irregular-area geologic map that was not included.

incorporated into a variety of index maps. Clearly, many permutations of map scale and vintage can be displayed in such maps; three examples are provided here (Figure 5). Commentary on the extent and distribution of geologic map coverage across the Nation, although feasible and desirable, is not within the scope of this methods-oriented paper. Such commentary, if it were to be provided at some later date, would most appropriately come from the NCGMP and AASG.

CONCLUSIONS AND RECOMMENDATIONS

The NGMDB project has long recognized the need to produce, for the practicing geologist, decisionmaker, and general public, a set of accurate and detailed index maps of geologic map coverage. The requirement for performance metrics provided the impetus to focus on this need, thereby giving it a much higher priority than was previously justified. In the process of creating the index maps, an important ancillary benefit was realized – by viewing the size and shape of the bounding box of each map, and by reading the titles and perusing the bibliographic information, previously undetected errors were found and corrected. Although a careful review of each record is time-consuming, it clearly improves the quality of the database.

In order to facilitate a more efficient and routine production of index maps, the data structure for the NGMDB Map Catalog will be revised to incorporate the new fields defined in this study. In cooperation with the agencies that provide new bibliographic records to the Map Catalog, all new entries will include information for these fields. Also, a more automated process of index-map creation will be developed; this process will include the error-checking procedures used here.

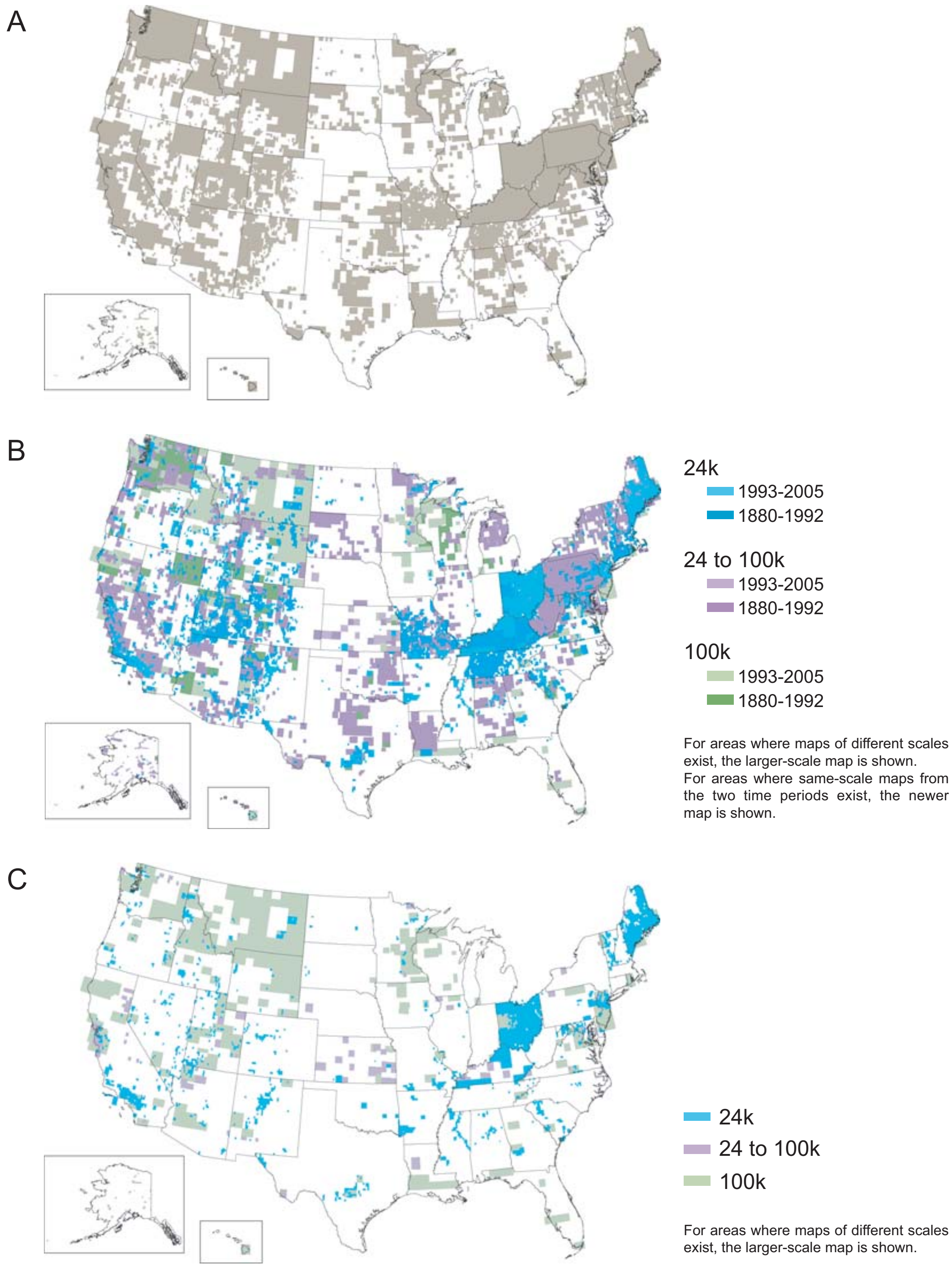


Figure 5. Index maps provided to the NCGMP and AASG, in partial fulfillment of the OMB request for performance metrics.

A) geologic map coverage, at scales 1:100,000 and more detailed, published from 1880 to 2005.

B) geologic map coverage, differentiated into three scale ranges (1:100,000, 1:24,000, and scales between), and published from 1880-1992 or 1993-2005.

C) geologic map coverage, at scales 1:100,000 and more detailed, published from 1993-2005.

Because these index maps show the bounding box for each geologic map there is, for some maps, an overestimation of the area that was actually mapped. In Step 4, it was noted that certain irregularly-shaped geologic maps were included in the index maps. Also, a quadrangle map may cover parts of two states, but the geology may have been mapped for only one state. Therefore, in the future these index maps should more precisely show map boundaries, for example by intersecting certain geologic maps with shoreline and state boundary GIS files, and by representing county and park maps by their true boundaries rather than by bounding box.

I conclude with a somewhat tongue-in-cheek comment. In this study I struggled to infer, from each publication title, certain basic characteristics for each map. My struggle begs the question – do our users, especially the non-scientists, select the most appropriate products from the (sometimes extensive) list of publications shown by a Map Catalog search or an agency's publication list? I suggest that our products might be more readily used if their titles were more succinct and standardized in format and terminology. At this time, I don't have specific recommendations, but as an author I recognize that my own titles could benefit by suggestions for clarity!

REFERENCES

Brabb, E.E., 1989, Geologic map of Santa Cruz County, California: U.S. Geological Survey Miscellaneous Investigations Series Map I-1905, scale 1:62,500, available at http://ngmdb.usgs.gov/Prodesc/proddesc_47.htm

Kepferle, R.C., 1973, Geologic map of the Raywick quadrangle, central Kentucky: U.S. Geological Survey Geologic Quadrangle Map GQ-1048, scale 1:24,000, available at http://ngmdb.usgs.gov/Prodesc/proddesc_10593.htm.

McCartan, Lucy, Lemon, E.M., and Weems, R.E., 1984, Geologic map of the area between Charleston and Orangeburg, South Carolina: U.S. Geological Survey Miscellaneous Investigations Series Map I-1472, scale 1:250,000, available at http://ngmdb.usgs.gov/Prodesc/proddesc_9253.htm.

Powell, J.W., 1888, Methods of geologic cartography in use by the United States Geological Survey, in Congrès Géologique International, Compte Rendu de la 3me Session, Berlin, 1885: A.W. Schade's Buchdruckerei, Berlin, p. 221–240.

Southworth, C.S., Brezinski, D.K., Orndorff, R.K., Chirico, P.G., and Lagueux, K., 2001, Geology of the Chesapeake and Ohio Canal National Historical Park and Potomac River Corridor, District of Columbia, Maryland, West Virginia, and Virginia: U.S. Geological Survey Open-File Report 01-188, scale 1:24,000, available at http://ngmdb.usgs.gov/Prodesc/proddesc_42277.htm.

Cartographic and GIS Activities in the U.S. Fish and Wildlife Service

By Douglas L. Vandegraft

Chief Cartographer
U.S. Fish and Wildlife Service
NWRS - Division of Realty
4401 North Fairfax Drive, Room 622
Arlington, VA 22203
Telephone: (703) 358-2404
Fax: (703) 358-2223
e-mail: Doug_Vandegraft@fws.gov

ABSTRACT

The U.S. Fish and Wildlife Service (FWS) is a federal agency whose mission, working with others, is to conserve fish and wildlife and their habitats for the continuing benefit of the American people. This mission is partially fulfilled through the establishment and maintenance of wildlife refuges. Under the management of fish and wildlife professionals, the National Wildlife Refuge System has become the world's premier network of wildlife habitats. The FWS is making use of modern cartographic methods and implementing Geographic Information Systems to more effectively manage the lands and resources entrusted to them.

THE NATIONAL WILDLIFE REFUGE SYSTEM

One hundred years in the making, the National Wildlife Refuge System is a network of habitats that benefit wildlife, provide unparalleled outdoor experiences for all Americans, and protect a healthy environment. Since President Theodore Roosevelt designated Florida's Pelican Island as the first wildlife refuge in 1903, these lands have grown to encompass approximately 100 million acres within more than 540 wildlife refuges (http://www.fws.gov/refuges/). Refuges are special places where the FWS and its partners restore, protect, and manage habitat for America's wildlife. Wildlife refuges occur across the breadth of the U.S. States, Territories, and Possessions, from the Virgin Islands to Guam, north to Alaska and south to the American Samoa. Mapping refuges across a region that extends across almost half of the planet makes the use of automated systems a necessity. The FWS has been using computers for refuge mapping, land management, and habitat analysis since the mid-1980s.

MAPPING AND GIS

Presently, Geographic Information Systems (GIS) and spatial data development efforts are coordinated within seven management "regions" and nationally through a FWS GIS Steering Committee (Figure 1). At the local level, GIS is used by refuge managers, biologists, and outdoor recreation planners for planning, operating, and evaluating refuge projects and programs. At the regional or landscape level, GIS is used to identify and implement refuge goals that support larger ecosystem goals and partnership efforts. At the national level, GIS is used to summarize information from a group of refuges within an ecosystem, flyway, or political boundary; or to develop regional or local wildlife management objectives from the analysis of information with broad geographic content (e.g., national or continental coverage of migratory birds and their habitats). Finally, GIS provides another avenue to communicate and share information with partner groups and the public.

GIS DATA WEBSITE

The FWS maintains a website, Geographic Information Systems and Spatial Data (http://www.fws.gov/data/gishome.html) devoted specifically to sharing the spatial data that is being generated in support of its mission. Available on this site is an interactive map-server that allows the user to view many of the digital refuge boundaries and download the boundary data and the associated metadata. The website uses the ArcIMS internet map-server software, made by Environmental Science Research Institute (ESRI, Inc.). The digital boundaries are in the shapefile format with a standard attribute scheme that includes the name of, and contact information for, the particular refuge. Other available layers are the Wetland

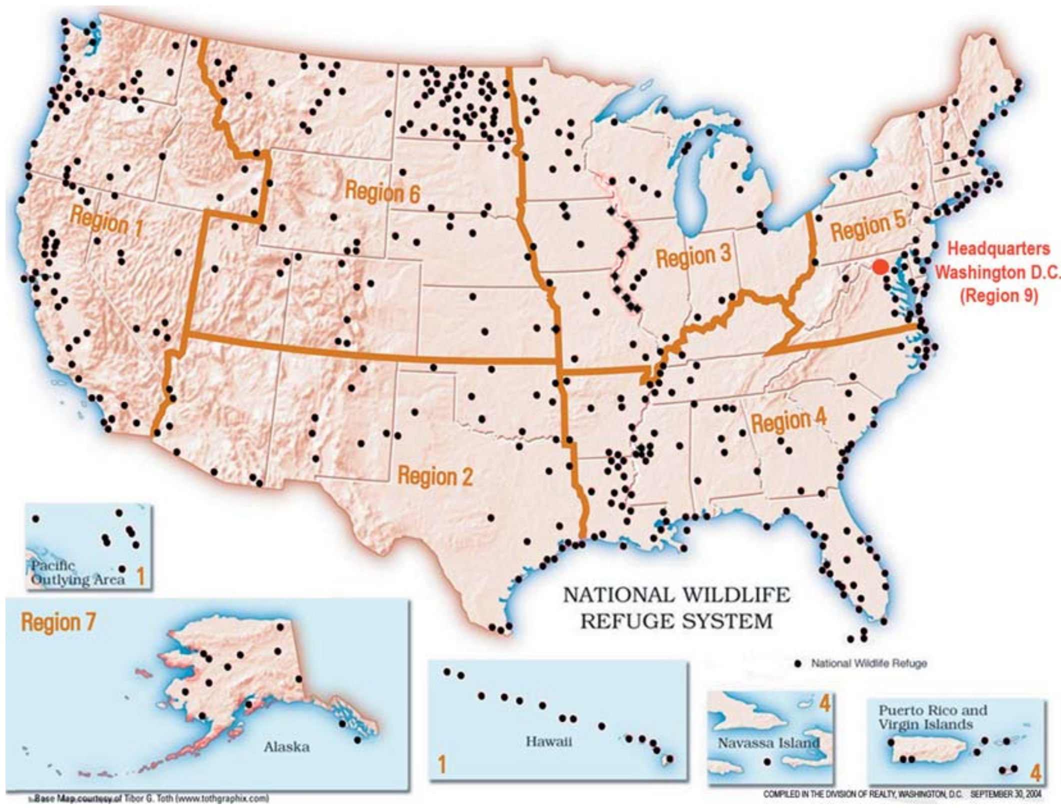


Figure 1. The Management Regions of the National Wildlife Refuge System.

Management Districts, Interstate Highways, and major water bodies. The main site has spatial data for download from other federal agencies, State agencies, educational groups, non-government groups, as well as commercially sponsored sites. Also at the main site is information describing GIS tools, FWS data standards (such as vegetation), as well as links to several other unique GIS sites (such as the Environmental Conservation Online System) that provide viewing and downloading of spatial data.

NEW MAP OF THE NATIONAL WILDLIFE REFUGE SYSTEM

Cartographers from the FWS have worked with the U.S. Geological Survey on a new map of the National Wildlife Refuge System, which was released in December 2002. The map is a combination of FWS and USGS data and will be included as part of the National Atlas (http://www.nationalatlas.gov). This collaborative effort marks the first time a national map for the refuge system has been produced entirely in a digital format. The 1:7.5-million scale of the map is consistent for all areas of the U.S., with the exception of the Pacific Islands area, where a scale of 1:29 million was used to reflect an area extending from Hawaii to Guam. The consistent scale allows the viewer to appreciate Alaska's tremendous size and the true expanse of the wildlife refuges located there.

NEW FORMAT FOR REFUGE MAPPING

The boundaries of the National Wildlife Refuges are intended to protect specific wildlife habitats and natural resources, as identified by FWS biologists. Private lands within these boundaries are purchased only from willing sellers. Maps that reflect these land ownership patterns have used a consistent format since 1936. In May 2002, cartographers from all seven management regions gathered to update the existing "Mapping" chapter in the FWS manual, which describes Service-wide standards for mapping (http://www.fws.gov/policy/343fw3.html). A new cartographic format for all maps produced for the FWS was designed and has been implemented since that time. The new mapping standards are based on existing graphic standards for FWS publications, which is similar to the standards for the other agencuies within the Department of the Interior. The maps will use USGS digital orthophoto quads and digital raster graphs as a base, with color shading to represent the various land status categories (see Figures 2 and 3).

In October 2004, the same group of cartographers gathered to revise the Standard Operating Procedure for generating digital refuge boundaries and digital land status. The new SOP document will implement the use of the ArcGIS "geodatabase" format to store boundaries, FWS-owned parcels, and private lands and their associated attributes (see Figure 4).

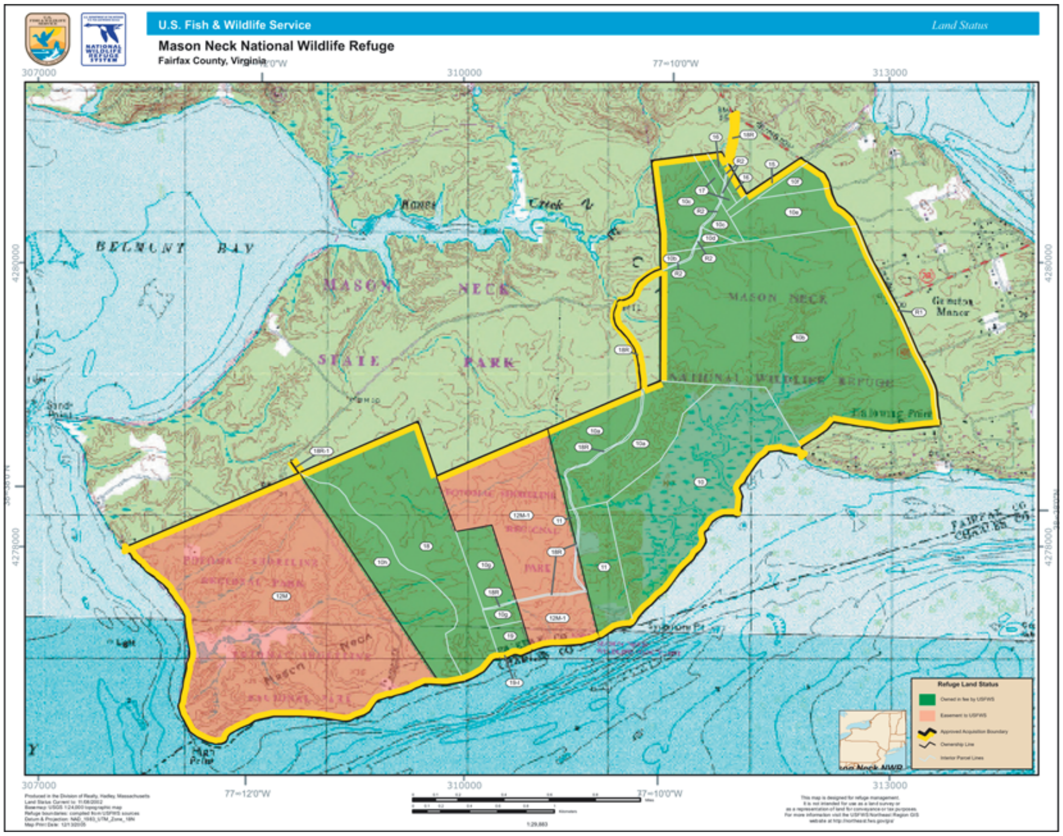


Figure 2. Example of the new map format, based on the existing FWS graphic standards.

FUTURE GOALS

A GIS coordination team at FWS Headquarters in Washington, D.C., has recently completed (1/06/2006) an extensive report to the FWS Directorate entitled “Recommendations for GIS in the National Wildlife Refuge System.” The report contains recommendations on national, regional, and field office needs for spatial data acquisition, development, and management in support of the National Wildlife Refuge System programs. This will help identify programs with common data needs, provide access to data systems and existing databases, reduce duplication of efforts, and provide technical support for science-based decision-making. The report will soon be available at the main FWS GIS website (http://www.fws.gov/data/gishome.html). The major recommendations are as follows:

1. Promote the Development and Implementation of a Consistent Workstation GIS for Refuge Field Stations.
2. Prioritize the Development and Acquisition of High-Resolution Digital Imagery for Vegetation and Land Cover Mapping.
3. Increase Technical Support to Field Stations.
4. Improve field station access to hardware and network connectivity.
5. Promote Data Standards in Preparation for an Enterprise GIS.

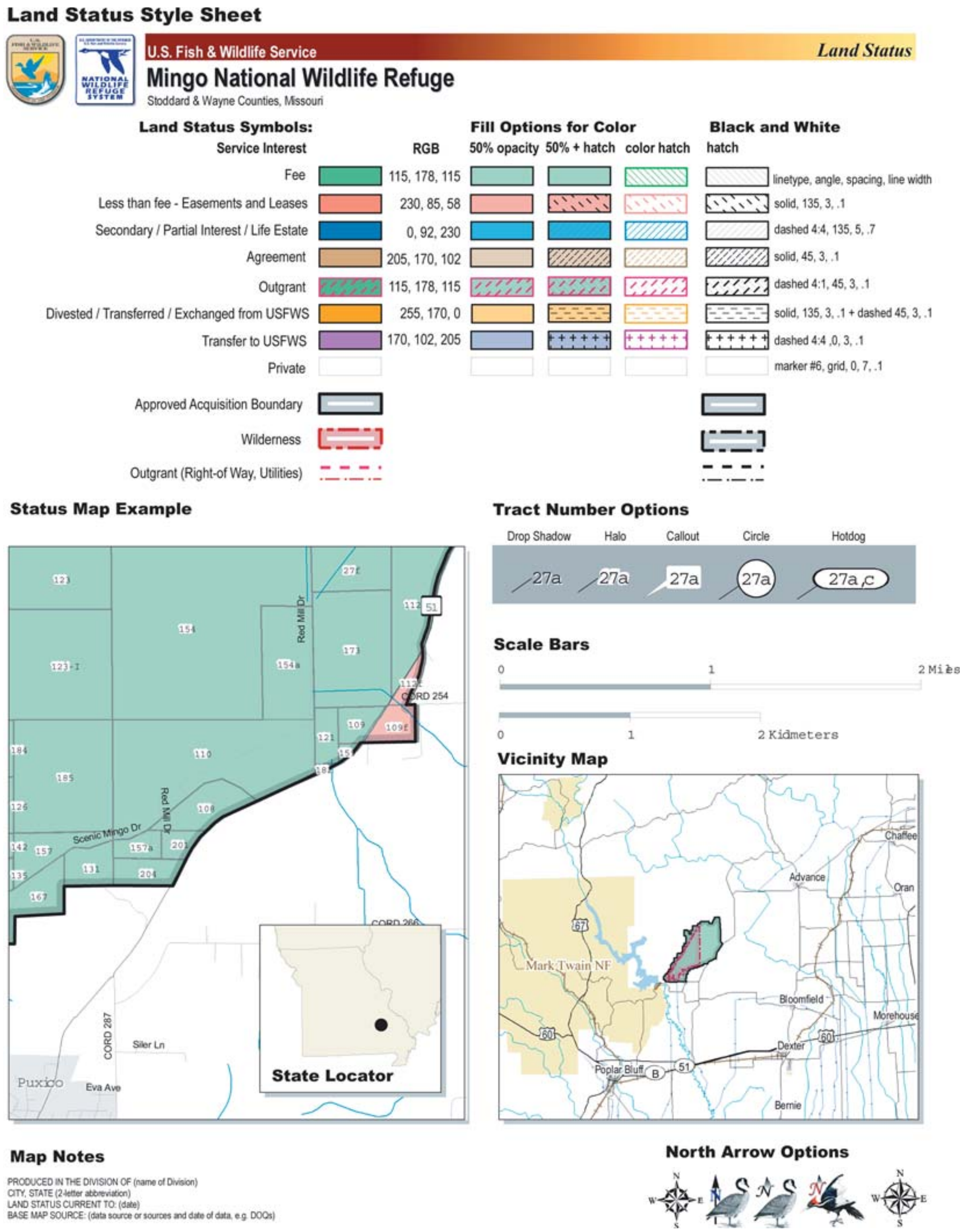


Figure 3. Guidance for making land status maps for the FWS are shown on the "style sheet."

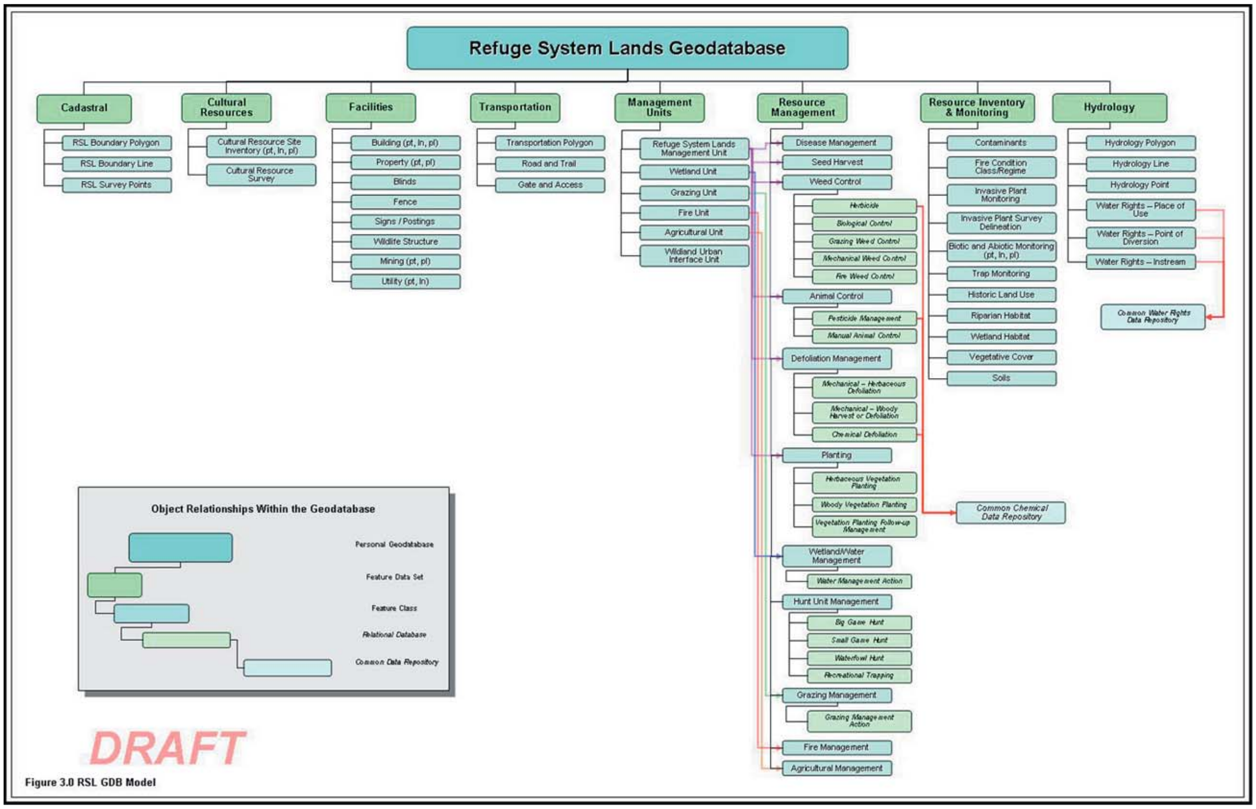


Figure 4. Proposed design of the geodatabase for lands within the National Wildlife Refuge System.

Map Production and Data Distribution the Idaho Way: An Update

By Jane S. Freed and Loudon R. Stanford

Idaho Geological Survey
University of Idaho
Moscow, ID 83844-3014
Telephone: (208) 885-7479
Fax: (208) 885-5826
email: jfreed@uidaho.edu, stanford@uidaho.edu

INTRODUCTION

In 1997, at the first Digital Mapping Techniques workshop, in Lawrence, Kansas, the Idaho Geological Survey (IGS) described their map production techniques. Since then, the IGS has published about one hundred geologic maps. This report will update how those maps, and the data behind the maps, are produced, published, and delivered online.

About the Idaho Geological Survey

The Idaho Geological Survey is housed on the University of Idaho campus in Moscow, 300 miles from the capital, Boise. There currently are 10 full-time state-funded staff, half of whom are geologists. Geologic mapping is the IGS's largest mission. The IGS has published nearly 400 geologic maps, many of which are page-sized figures in bulletins or pamphlets as part of mining reports. The IGS in 1989 began operating a full-time map production facility, which from the start was a "digital" shop that used inexpensive Computer Aided Design (CAD) software. In 1992, IGS began to use AutoCAD software to capture geologic map data in a GIS-compatible format. In 1992 the IGS published its first full color, press-run geologic map using digital processes to make map separates. Our methods of publication have changed dramatically in the last ten years (see Figures 1 and 2). At the time of our first DMT presentation, in 1997, most maps were published in black-and-white, Xerographically. To publish a map in color before 1998 required a printing press. Then and now, there was little money in the budget for expensive press runs. Beginning in 1998, all maps published by the IGS were produced in color using methods discussed in this paper. Maps were design to be printed-on-demand using ink-jet, large-format print technology. Beginning in 2001, all newly published maps became available online free as Acrobat Reader files (PDF format).

About the Idaho Geological Survey's Digital Mapping Lab

Over the past 10 years the IGS Digital Mapping Lab has prepared more than 100 geologic maps for publication, which is more than one-fourth of all maps produced since establishment of the Idaho Geological Survey (Figure 3). Currently the Digital Mapping Lab has two full-time and two part-time employees. It is responsible for producing published maps, managing all data generated from mapping projects, producing digital geologic map databases, managing the local network and file server, and maintaining the IGS web site.

DATA CAPTURE

The staff at the Idaho Geological Survey find it more productive and efficient for the geologists to do the science and the cartographers to make the maps. The geologists do all the fieldwork and draw lines on Mylar overlays or on greenlines and prepare all other relevant materials. The IGS's Digital Mapping Lab takes these materials and captures the data for publication as a geologic map and as a database.

Digitizing

Geologic data capture is done in AutoCAD r 14 with CADmappr, an AutoCAD plug-in developed in-house. Data is digitized or captured in real coordinates, specifically Idaho state plane feet. This "real units" map is the database version of the map in which all geologic objects are captured, geologically attributed, and given a source identifier. All subsequent changes are made to the database map, which is ultimately released as part of a GIS data set (see Figure 2). Digitizing is usually done on large digitizing tablets with programmable 16-button cursors. To begin the process, the map is aligned on the

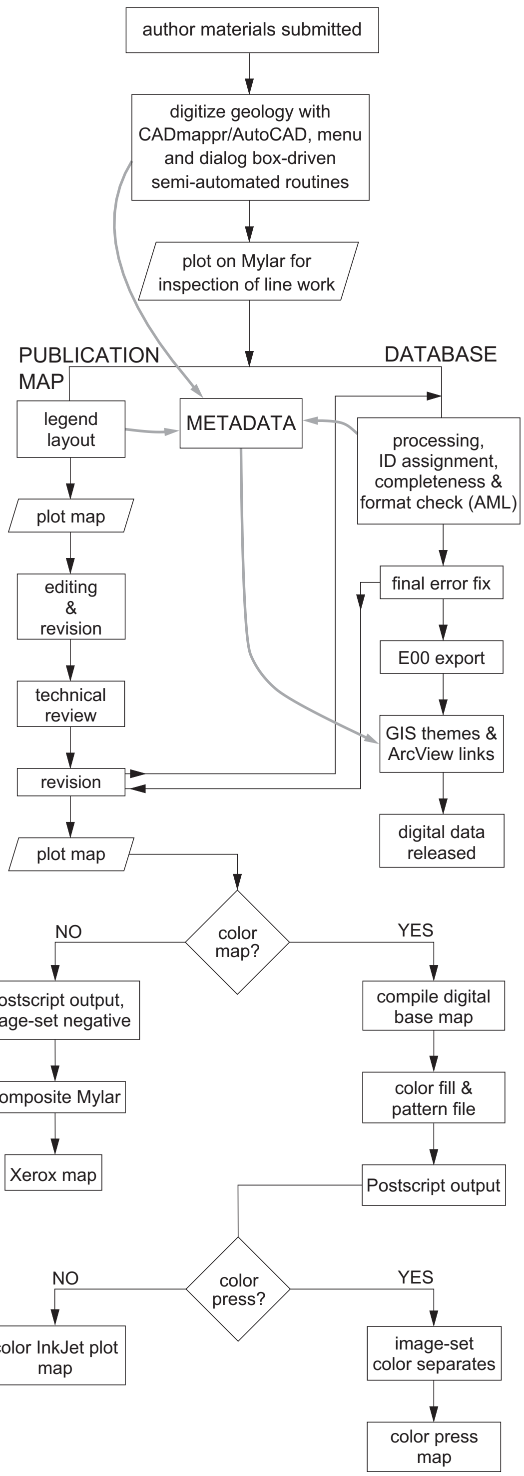


Figure 1. Chart showing map production workflow in 1997.

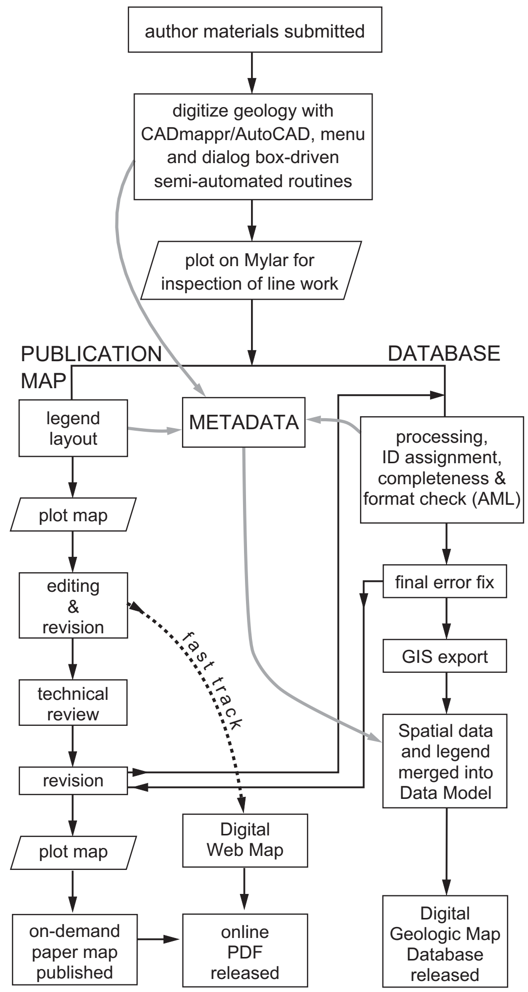


Figure 2. Chart showing map production workflow today.

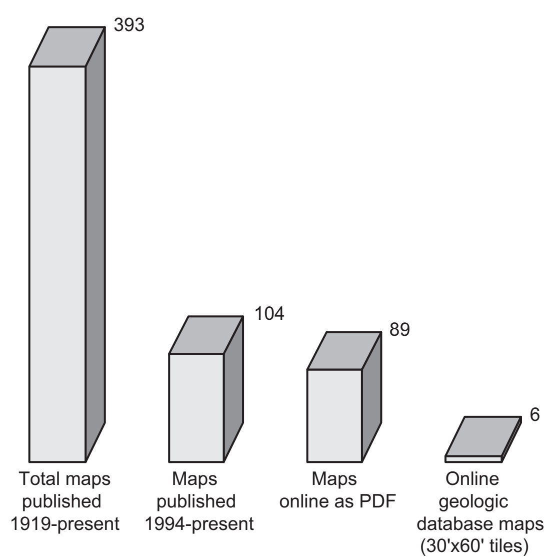


Figure 3. Chart showing geologic map publication statistics for the Idaho Geological Survey.

digitizing tablet. The mapping to be captured is either on a stable base with a Mylar overlay registered to it, or on a greenline. The map is taped to the digitizing table, then the corner tics of the map are aligned to the on-screen tics. The on screen tics for a 7.5-minute quadrangle are placed at 2.5-minute intervals for the latitude and longitude extent of the map. Once the map is aligned with sufficiently low error (RMS 10 or less), we begin by digitizing the geologic contacts.

Once a contact is digitized, a dialog box pops up with the list of units, line types, and whether it's a contact, fault, or hidden contact. After picking the units that lie on either side of the line, the screen returns to the map and the topology is chosen for the two units by pointing to the correct contact side for each segment. Map unit topology is stored with each contact (AutoCAD extended entity data) and also in the layer-naming format for each contact. To make PostScript polygon fill files, all similar map unit layers are combined automatically. For example, all "Qal" polygons can be shown by turning on all layers with "Qal" in their name. GIS polygons are only generated after exporting a specialized ArcInfo interchange file (E00) from AutoCAD to ArcGIS using routines that combine layers as above with a centroid label for each polygon. This process is repeated until all contacts and faults are captured. Contacts and faults that also serve as geologic contacts are digitized first, then "dangling" fault lines are added. Each contact segment is digitized once. If the contact happens to be a fault, it is still only digitized once. Separate geologic layers or themes (e.g., all faults) can easily be generated at any time by running a query or simply turning AutoCAD layers on or off using custom tools.

Next, bedding attitude and other symbols are digitized. Again, a series of custom tools along with the programmed 16-button cursor make this task relatively simple and efficient. Finally, text labels for units and objects (e.g., fault names) are added to the map database.

When digitizing is complete a check plot is made. Check plots are done on the lab's beloved 15-year-old HP pen plotter to make sure everything that should be on the map is indeed on the map and spatially correct. Lab staff checks the plot against the geologist's line work and makes any corrections. Geologists are encouraged to also use this check plot to find errors. Geologic source attributes are then added to all geologic objects on the map (Table 1). Next, a series of quality control checks verify the map's completeness. A series of manual checks and software routines check for polygon closure, fault type and movement direction, and topology of contacts. Polygon topology need not be checked, because if all contacts are correct then polygons have to be correct. Polygons are only generated outside of the AutoCAD drawing for two purposes; the first is when PostScript color file files are generated for FreeHand or Illustrator, and the second is to check for missing unit labels in ArcGIS. The map is then ready to save as the "pub" version.

Table 1. Typical attributes captured for a geologic object digitized in AutoCAD.

Geologic Object Attribute	Description
Line type	Certainty and relationship
Unit to Left and Right of line	Topology
Fault type	Normal, thrust, etc.
Movement	Direction of fault movement relative to first vertex
Source	Geologic reference
Map code	Quadrangle or map name
Tile code	30' x 60' minute tile name
Object name	Fault name or contact name

To expedite map layout in FreeHand or Illustrator, the CadMappr plug-in includes an entire module for exporting PostScript files from AutoCAD. First, the "pub" version is made by reducing the database map to the appropriate publication scale using a custom tool. Once at publication scale the map is ready for final steps before export. Color and pattern selection for each map unit is the most time intensive of these steps. A lookup table holds the values for color and pattern type. This table is used to generate color fills and patterns when the AutoCAD unit polygons are exported as PostScript files. Another lookup table is used to control line weights and the PostScript output order for line and symbol objects when they are exported. In PostScript, one map object, then the next object, and so on, are written to the print file; if two objects occupy the same map area, the object written to the file first will be physically covered by the object written later. If it is desired that a red symbol be shown under black geologic contacts, the PostScript coding for the symbol must be written before the contact information. In this way, PostScript files for various layers of geology are generated for layout (Figure 4).

AutoCAD to GIS

AutoCAD and CADmappr allows the Idaho Geological Survey's Digital Mapping Lab to be productive while attributing each geologic object (Table 1). These geologic attributes are designed for export to ArcInfo and ArcGIS. For example, the movement direction for faults is stored while digitizing. This attribute follows ArcInfo rules for right-left side relative to the beginning node or first vertex. Separate geologic layers are easily created for export by running a query.

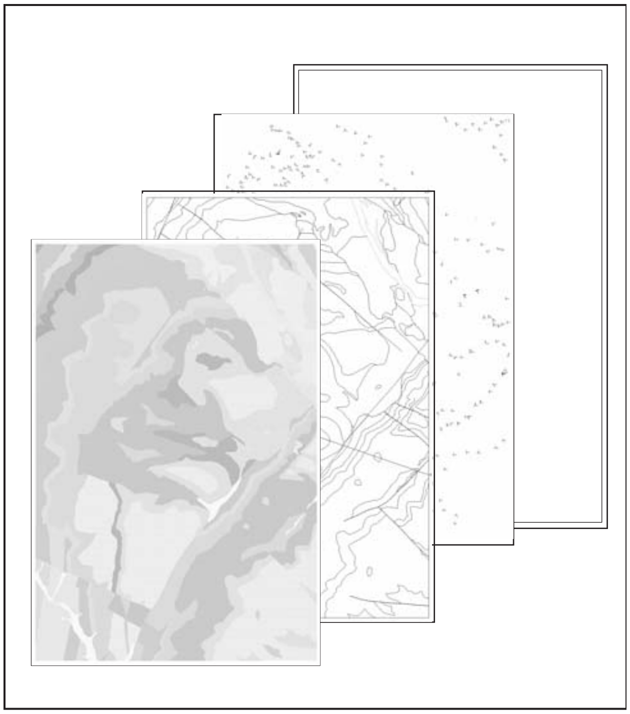


Figure 4. EPS/PS layers from AutoCAD, for layout in FreeHand or Illustrator. From front to back, the layers are: color fill, geologic contacts, geologic symbols, and alignment boxes.

LAYOUT

Either FreeHand or Adobe Illustrator can be used for the final geologic map layout. PostScript (PS) files can be imported directly into Illustrator. Before import into FreeHand, the PostScript layers are processed by a conversion program called Transverter Pro and turned into EPS files; EPS files work better in FreeHand, whereas PS files work better in AI. Once the PostScript files from AutoCAD are imported (or placed), the layout procedures described below are essentially the same for either FreeHand or Illustrator.

The first step in map layout is to open a neat line template. This template also contains various "boilerplate" information including cartographic and author credits, scale, declination, and title blocks, which can be modified. Then, the PostScript layers generated in AutoCAD are imported and aligned to a specially created alignment layer in FreeHand (the "backmost" layer in Figure 4). Next, the scanned image of the stable film base map is rectified and imported where it is also aligned. In Illustrator, all layers are brought in and aligned to the base map visually, and patterns are added manually from the USGS pattern set (U.S. Geological Survey, 1999). Once all map layers are imported and aligned, all other parts of the map are imported and manipulated to create an aesthetically pleasing map. These remaining parts include cross-sections and correlation charts which are created in AutoCAD using the same procedures as for the geologic map. If the map sheet includes tables, these are imported from Microsoft Excel and formatted to a template. Text is imported from WordPerfect and formatted to a standard style. If a booklet is required, it is laid out in FreeHand.

Once laid out, the completed map is printed in color on an ink-jet printer. The authors check this first plot, and corrections are made accordingly. Necessary corrections to the geology are made in the database version of the map and are exported again. At this point the map author decides whether the map is either "fast-tracked" (IGS Open-File) and published online as a Digital Web Map, or sent on for technical (scientific) review. Finally, after technical review and review by the IGS publications editor, the map is edited as necessary and prepared for publication in one of the IGS's map series (Table 2). All geologic map data captured is eventually compiled, with other mapping, into a 30' x 60' tile and released as a geologic map database.

FreeHand vs. Illustrator

In our experience, FreeHand has been easier to learn, is easier to use, and has better text editing and search capabilities. FreeHand can handle booklet layout. It is easier to precisely align layers in FreeHand. On the other hand, Adobe Illustrator is generally a more powerful map layout environment. It is more compatible with other software, especially Microsoft products. The USGS standard patterns can be imported and used in Illustrator. Superior transparency control means that Illustrator is preferred for making shaded relief illustrations and maps.

MAP DISTRIBUTION

Web

The Idaho Geological Survey maintains a sales office at its Moscow offices on the University of Idaho campus. All IGS publications plus many map publications pertinent to Idaho are sold at the office. Historically, the biggest sales items have been topographic maps. As many of these products have become more readily available on the Web, the number of walk-in customers has decreased. At the same time, the IGS's web site has seen a steady increase in the number of downloads of PDF maps, reports, and geologic map databases. While it is important to maintain a sales office to manage the distribution of traditional publications, the Web will increasingly be the best avenue for reaching the public.

Since 2001, when the IGS began posting PDF maps, nearly 90 maps have been placed online. Most of these maps are in full color and all are available free of charge. A new publication category, Digital Web Map (DWM),

Table 2. Geologic map products at the Idaho Geological Survey.

Publication Series	Description	Online free (pubs since 1998)	Paper Print Available	Average Time to Publish
Geologic Map	External technical review; review by IGS editor; PDF and print-on-demand	yes	yes	2.5 years
Surficial Geologic Map	External technical review; review by IGS editor; PDF and print-on-demand	yes	yes	2.5 years
Technical Report	Possible external technical review; possible review by IGS editor; PDF and print-on-demand	yes	yes	1.5 years
Digital Web Map	Review by authors; open-file PDF and print-on-demand	yes	yes	60 days
Digital Geologic Map Database	External technical review; review by IGS editor; GIS data set with data model	yes	no	3.5 years

was created in 2003 especially for web distribution of geologic maps. Digital Web Maps are an open file or preliminary version of the geology designed to make the science available in the most timely manner possible.

Print-on-demand

Paper copies of most older (pre-1998) IGS map publications are still only available as black and white photocopies. However, nearly all IGS maps produced after 1997 are available by print-on-demand. The printing of these maps is contracted to an on-campus imaging service.

DATA DISTRIBUTION

As described earlier in this paper, IGS collects spatial data and attributes about geologic object on geologic maps. When the spatial components of a map are linked or combined with map legend data from the same geologic map, in a systematic framework, a data model is created. The IGS developed a geologic map data model design to provide a framework for storing the spatial, legend, and metadata components of its geologic maps (Stanford and MacKubbin, 2000; also see http://www.idahogeology.com/Lab/datamodel.htm). Since 2003, six data sets have been created and released in the IGS's Digital Geologic Map Database series.

Map information in our standard data model are released in two formats, and are designed to be used together in a GIS. The spatial components are in ArcInfo coverage format, and the legend and metadata tables are in Microsoft Access format. These data sets are available for free download or can be purchased on CDROM. If properly designed and maintained, this data model framework also provides a mechanism to manage geologic map data for the long-term. Currently, the IGS data model is being redesigned to work with the ArcGIS Geodatabase model. The hope is that data will be easier to distribute and manage over time in this environment.

CONCLUSIONS

The Idaho Geological Survey divides the labor of capturing and publishing geologic maps along traditional lines while using new and old digital tools to get the job done. The IGS mapping lab continues to use AutoCAD as its chosen geologic map capture software. The easy editing environment combined with a series of customized tools and plug-ins make this platform a productive tool for map data capture and publication. FreeHand and Adobe Illustrator are used to lay out the digitized geology and print the final map publication (a summary of hardware and software used by the IGS is provided in Tables 3 and 4).

The Web is a vital to the distribution of IGS products. All newly published maps are available as PDF files, at no cost, and can also be ordered as paper prints.

REFERENCES

U.S. Geological Survey, 1999, Public Review Draft—Digital Cartographic Standard for Geologic Map Symbolization (PostScript Implementation): U.S. Geological Survey Open-File Report 99-430, available at http://pubs.usgs.gov/of/1999/of99-430/.

Stanford, L.R., and MacKubbin, V.T., 2000, Application of a Digital Geologic Map Data Model in ArcView GIS, in D.R. Soller, ed., Digital Mapping Techniques '00—Workshop Proceedings: U.S. Geological Survey Open-File Report 00-325, p. 55-56, available at http://pubs.usgs.gov/of/2000/of00-325/stanford.html.

Table 3. Hardware used to make geologic maps.

Hardware	Specs	Comments
Digitizing Table	GTCO 36 x 48 back lighted/ 16 button puck	2 years old
Digitizing Table	GTCO 36 x 48 /16 button puck	5 years old
Digitizing Table	40 x 60 Calcomp/16 button puck	13 years old /Works only with NT
Pen Plotter	HP Draft Master I	15.5 years old and still working
Ink Jet Plotter	HP 2500 /36 inch	IGS contracts out color plotting to campus media center
Large Format Scanner	54 inch 400 dpi optical	IGS buys access to this device
XP Work Stations	1-1.5 GB RAM/large monitors	
NT Work Station	Large monitor	Old machine for old digitizing tablet
IGS File Server	Small server for map production file serving and backup	
IGS Web Server	Development and production machines	

Table 4. Software used to make geologic maps.

Software	Purpose	Comments
AutoCAD r14	Capture/design/layout	Good for digitizing, editing/large learning curve
Cadmappr	Capture/Postscript/GIS	AutoCAD plug-in, developed in-house
AutoDesk Map 5	Export/import	
ArcGIS 8.x	Distribution/import/export/GIS	
IGS tools /ArcInfo 8	Conversion tools for Acad r14 map data	Developed in-house
Adobe Acrobat	PDF making/converting	
Transverter Pro	Pre-Press/PostScript conversion software	Good RIP/Plot file pre-viewer
Adobe Photo Shop	Image editing	
Paint Shop Pro	Image editing	
Adobe Illustrator 10	Map layout	More compatible than FreeHand/ harder to use/cannot do booklets
FreeHand 10	Map layout/booklet layout	Easier to use than AI but limited features
WordPerfect	Word processing/pre-layout	Use as little as possible
Microsoft Word	Word processing/pre-layout	Use as little as possible
Microsoft Access	Databases/metadata/data model	
Microsoft SQL Server	Databases	Future SDE/IMS?
Microsoft Excel	Spread sheet/table import	
Text Pad	Text editing/simple programming/ data editing	

Analysis of Errors Occurring in the Transfer of Geologic Point Data from Field Maps to Digital Data Sets

By Elizabeth Campbell[1], Ian Duncan[2], and Harry Hibbitts[1]

[1]Virginia DMME-Division of Mineral Resources
900 Natural Resources Dr. – Suite 500
Charlottesville, VA 22903
Telephone: (434) 951-6343
Fax: (434) 951-6366
e-mail: elizabeth.campbell@dmme.virginia.gov

[2]Texas Bureau of Economic Geology
John A. & Katherine G. Jackson School of Geosciences
The University of Texas at Austin
University Station, Box X
Austin, TX 78713-8924

ABSTRACT

An analysis was performed to evaluate the frequency of blunders that resulted from the use of multiple transfer sheets to compile a printed geologic map from field maps. Point data from field maps were compared to the corresponding final published map for five quadrangles. The blunder rate ranged from 0.46% to 15.99%. The most frequent blunder was using the wrong symbol for bedding measurements, followed by an error in the dip number. Blunder analysis yielded insight regarding the approximate percentage of measured geologic points that are possibly misrepresented on any given published geological map that was created using traditional publication procedures and techniques.

THE PROBLEM

The Virginia Division of Mineral Resources has made a digital compilation of the geology of fifteen quadrangles along Interstate 81 in the central Shenandoah Valley of Virginia (Figure 1) as part of the USGS STATEMAP program. The goal of the project was to create a digital mosaic of the geologic information that could be used by counties, state agencies, and other stakeholders for planning and management purposes. A comprehensive digital collection of geologic field data, particularly point data, facilitates the integration of geologic information into the decision making process. Additionally, the compilation enables the study of large geologic structures that are not fully expressed in the area of a single 7.5-minute quadrangle.

Geologic maps of twelve of the quadrangles had been previously published in the traditional manner and three existed only as mylar-based draft maps. All of the maps were scanned and digitized. For five of the published maps, one or more original field maps were available. It was observed that many field maps had more structural measurements data than appeared on the final, published maps. Consequently, in an effort to increase data density, data were digitized from the original field maps when available.

During the process of digitizing point data from field maps, differences were noted between the published and original field maps for the same geographic location or field station data point. These ranged from different dip numbers to different symbols. For example, the published paper map might show a bedding measurement where the field map shows a cleavage measurement (Figure 2) or the field map shows an overturned bedding symbol where the published map shows a normal bedding symbol (Figure 3). Because an error in transcribing a symbol can affect the structural interpretation of the geology (Figure 3), we wanted to quantify the error rate of data misrepresentation on maps published using the traditional pre-digital map production procedures.

BACKGROUND

An understanding of the methods of pre-digital map production is necessary in order to understand the sources of possible error. Generally, the field geologist recorded observations and field measurements in his notebook and on a field map. This information then was

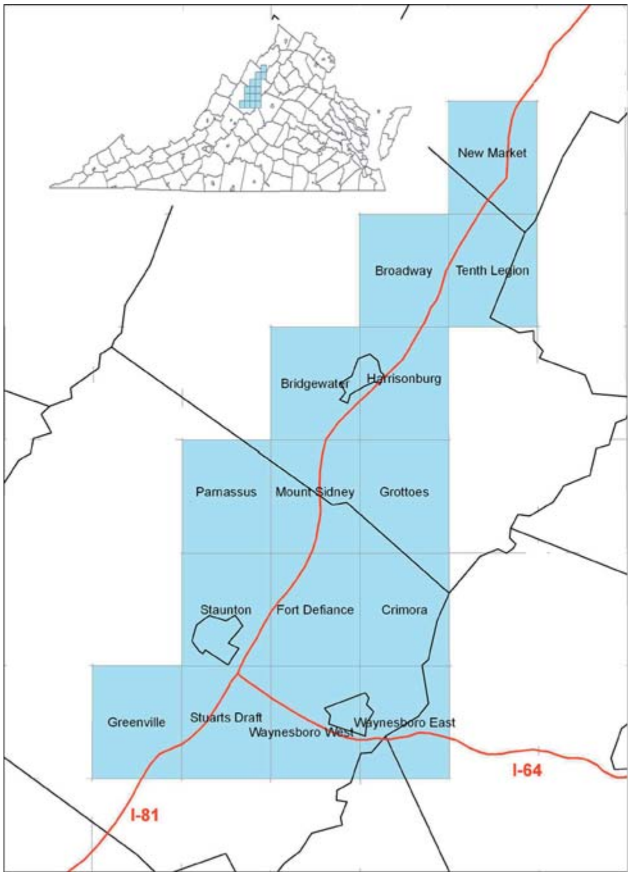


Figure 1. Location map showing the area of study in Virginia.

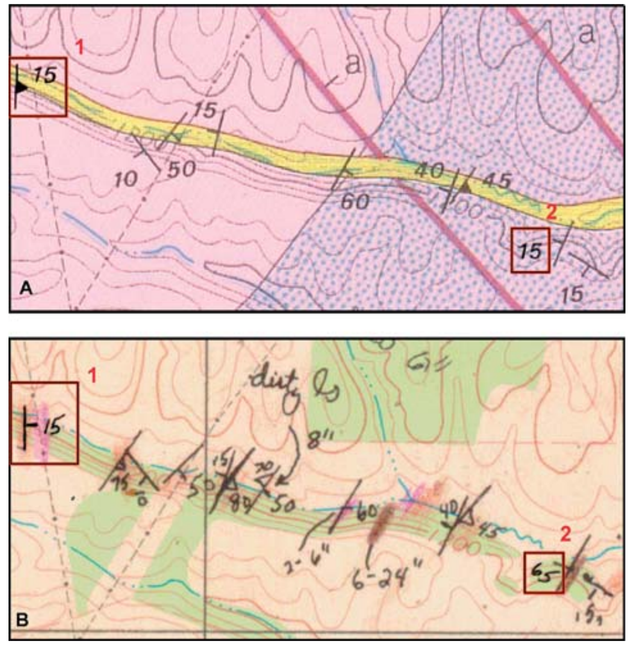


Figure 2. Published map (A) and the original field map (B) for the same area illustrating two types of discrepancies found. In example 1, on the left edge of the maps, the symbol changed from bedding measurement on the field map (B) to cleavage measurement on the published map (A). In example 2, on the right edge of the maps, the dip measurement was 65 on the field map (B) and 15 on the published map (A).

compiled onto a clean map. If more than one geologist mapped in an area, each geologist had his own field and compilation maps which he would ultimately give to the senior geologist who was charged with compiling the final manuscript map.

The senior geologist determined what data to place on the final map. In order to produce a readable, not overly crowded, printed map, it was common to not use all the field measurement data collected, but instead to place on the map representative data that supported the structural interpretation. For example, multiple data points at nearby sites could be represented by one generalized point on the map. The senior geologist compiled the map on paper and had a draftsman transfer it to a mylar base.

When it had been reviewed and approved for publication, the map was prepared for the printer. A draftsman transferred the line data and structural measurement to mylar overlays. For the point data, the draftsman used pre-printed rub-on symbols and numbers. The mylar topographic base map and mylar overlays were sent to the printer, where scribers traced the line work and area colors onto "scribe coats", which were then used to create printing plates. A separate printing plate was created from

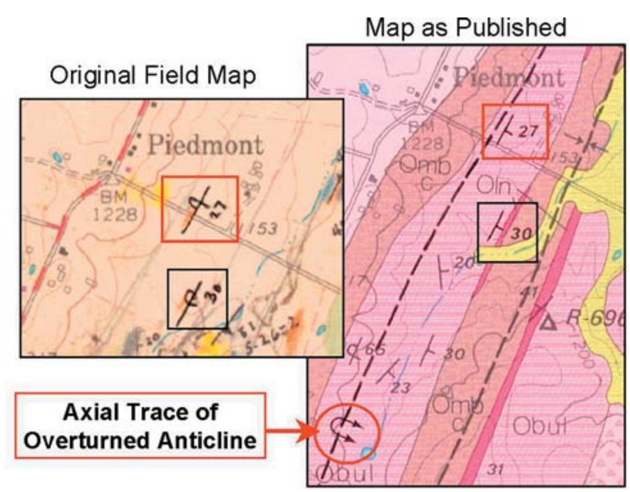


Figure 3. Field map and published map for the same area, illustrating the effect on structural interpretation arising from possible errors in transcribing. Bedding measurements marked as overturned on the field map are shown on the published map as normally oriented. If the points had remained overturned, the axial trace of the overturned anticline would have a different alignment.

the point data overlay. During the printing process, the paper was run through each of the color printing plates and the black point data plate in order to produce the final, printed map. Obviously, alignment of the plates was critical in this operation.

Every time the map data were transferred, there was a chance that the data would be copied incorrectly. Many of these errors were caught during the reviews that occurred at multiple points in the traditional map creation process.

SOURCES OF ERROR

Error has three components—blunders, random error, and systemic error. Of all the error components that affect digital geologic data sets, blunders have received the least study. Although blunders may commonly be thought of as large errors, they can be of any size or magnitude. Blunders are gross errors or mistakes related to carelessness or some other lapse in the system. Blunders are typically related to human mistakes such as reversing coordinates during data entry. Blunders, like systematic errors, are in theory avoidable. However, blunders are not predicable or regular in occurrence. Consequently, they are hard to detect and correct. Blunders can be revealed by comparison of the same data in different datasets.

The discovery of undescribed blunders undermines end-user confidence not only in the data but also in the agency that publishes or distributes the data (Duncan and Campbell, 2005). As noted by Openshaw (1989) "what many applications…need is not precise estimates of error but rather some confidence that error and uncertainty levels are not so high as to render in doubt the validity of the results." To maintain the end user's confidence in the data set, the reliability assessment should seek to identify and characterize blunders and systemic error in the data.

How do blunders occur in geologic mapping? Several possible sources of data error exist. The first is mistakes in data collection due to poor training or inattentiveness in the field. Could the field mapper (data collector) recognize field evidence for overturned bedding or the difference between cleavage and bedding? Proper training and testing prior to being allowed to do field work should be standard. Capturing the "collector" as an attribute for each point allows some discrimination should it be later determined that a collector did indeed have a flawed field technique. The data could then be targeted for re-sampling.

A second source is transcription mistakes—the misreading of numbers and symbols as data are copied from, for example, field sheets to compilation maps prepared in the office. The increasing use of portable devices allowing the collector to digitally record the data in the field should greatly reduce the likelihood of a transcription error by reducing the number of times the data is transferred from one media to another, and by reducing the need to read handwritten numbers and symbols.

A final source of difference between the field map and the final published version may be the editing done by the geologist who is in charge of finalizing the map. Some changes may be the result of reinterpretations by the senior geologist while in the office, or after subsequent field visits where the field notes and conclusions are either not recorded or the records have been misplaced.

PROCEDURE TO CHARACTERIZE BLUNDERS

The type of mistakes revealed by comparing field maps to published maps were largely human error—blunders. The errors were divided into two categories based on the perceived severity of the mistake to geologic interpretation.

A '**Blunder**' was recorded if one or more of the following conditions existed for any given point:

- If the symbol on the field map was not the same as the symbol on the published map
- If the difference in the dip of the bed on the field map and the corresponding point measurement on the published map was greater than or equal to 10 degrees.
- If the difference in the strike of the bed on the field map and the corresponding point measurement on the published map was greater than or equal to 10 degrees.

A '**Conflict**' was recorded if one or more of the following conditions existed for any given point:

- If the difference in the dip of the bed on the field map and the corresponding point measurement on the published map was less than 10 degrees but greater than 0 degrees.
- If the difference in the strike of the bed on the field map and the corresponding point measurement on the published map was less than 10 degrees but greater than 0 degrees.

The digitizing of point data from the original field maps occurred after the published/mylar point data had been digitized. Paper copies of scans of the field maps were printed and compared to the final published map. When a specific point on the original field map was correctly represented on the published/mylar map, then that point on the original field map was marked to indicate that it had been correctly represented on the published map.

If a point on the original field map was not on the published map or was represented differently on the published map, then the following details were recorded in tabular format for that particular point:

- Quadrangle name and scale abbreviation
- Geologist of the original field map
- Symbol type (i.e. vertical bedding, foliation showing strike and dip)
- Strike azimuth, manually measured directly from the oriented original field map
- Dip, as scribed on field map
- Yes/No—Is this original field map data point in conflict with a corresponding published/mylar map data point? (The answer would be “No” for a point that was on the field map but not on the published map.)
- Yes/No—If this point is different from the published map, is it a symbol error?
- Yes/No—If this point is different from the published map, is it a dip error?
- Yes/No—If this point is different from the published map, is it a strike error?
- Conflict/Blunder—Was the difference a conflict or a blunder?

The latitude and longitude of the field map data point locations were digitized using a DOS based computer program. The locations each field map data point were combined with the attributes above. When all points of a particular original field map were digitized, the field map point data were printed on a base map and visually inspected for accuracy.

RESULTS

Original field maps and notebooks were available for the following five quadrangles: Waynesboro West, Fort Defiance, Crimora, Mount Sidney and Broadway. Fort Defiance and Mount Sidney quadrangles had multiple field maps because a team of geologists mapped these quadrangles.

Table 1 shows the results of analysis of the quantity of *conflicts* and *blunders* present. The location blunder rate (number of published data point locations with one or more blunders divided by the total number of published point locations) ranged from 0.46% to 15.99%. The location conflict rate (number of published data point locations with one or more conflicts divided by the total number of published point locations) ranged from 0.46% to 5.81%. In Table 1, locations are counted only once. More than one error could be associated with a single point. A point could have a symbol blunder, a dip blunder, and a strike blunder; therefore a maximum of three blunders could be associated with one point, but the point would only count once in the location blunder rate.

Four of the five quadrangles have points with multiple errors. Examining the blunders more closely, Table 2 shows the total number of each type of blunder by quadrangle. Using the sum of blunders instead of the total number of locations with blunders raises the blunder rate (0.46% to 16.95%).

The most frequent blunder was symbol error followed closely by a difference in the dip number. Looking more closely at the type of symbol error (Table 3), the most common mistake was changing a symbol indicating overturned bedding on the field map to right-side-up bedding symbol on the published map. The converse is the next most prevalent mistake. Changing symbols indicating cleavage measurements on field maps to symbols indicating bedding measurement on the published map is the third most common mistake.

DISCUSSION AND CONCLUSIONS

Past studies of error in digital geological databases have been largely concerned with the nature and magnitude of location error. The data presented in this study indicates that the occurrence of blunders, particularly symbol substitution, during the map creation process have

Table 1. Number of locations with one or more errors (either blunder or conflict) in the five quadrangles studied.

Quadrangle	BWAY24	MSID24	FDEF24	WYNW24	CRIM24	Total
Number of field maps	1 field map	3 field maps	2 field maps	1 field map	5 field maps	–
Total Number of published points	216	519	499	934	407	2575
Number of locations with Blunders	1	83	60	14	13	171
Number of locations with Conflicts	1	25	29	9	6	70
Location Blunder Rate	**0.46%**	**15.99%**	**12.02%**	**1.50%**	**3.19%**	**6.64%**
Location Conflict Rate	**0.46%**	**4.82%**	**5.81%**	**0.96%**	**1.47%**	**2.72%**
Total Number of Errors	2	108	89	23	19	241
Location Total Error Rate	**0.93%**	**20.81%**	**17.84%**	**2.46%**	**4.67%**	**9.36%**

Table 2. The total number of blunders by type of error in the five quadrangles studied.

Blunders Types	BWAY24	MSID24	FDEF24	WAYNW24	CRIM24	Total Blunders	Number of Blunders/ Total pub pt
Number of Symbol Blunders	0	51	36	6	5	98	**3.80%**
Number of Dip Blunders	1	33	26	9	8	77	**2.99%**
Number of Strike Blunders	0	4	4	2	3	13	**0.50%**
Sum of all blunders	**1**	**88**	**66**	**17**	**16**	**188**	–
Published points	216	519	499	934	407	2575	–
Blunder rate	**0.46%**	**16.95%**	**13.22%**	**1.82%**	**3.93%**	**7.30%**	–

Table 3. Breakdown of the symbol error blunders.

Field symbol	Published symbol	Number of points
overturned bedding strike and dip	bedding strike and dip	31
bedding strike and dip	overturned bedding strike and dip	19
bedding strike and dip - top known	bedding strike and dip	6
cleavage strike and dip	bedding strike and dip	8
bedding strike and dip	cleavage strike and dip	3

been a source of error in geologic maps produced using the traditional method.

A debate arose during the course of this study—when is an overturned bed fact, and when is it interpretation? Field determination of overturning is based on primary, field observable evidence such as graded bedding or cross-bedding. Properly done, such a determination is factual. In the office, the map compiler may determine that a fold limb is entirely overturned based on structural analysis. This determination is an interpretation. It is important to differentiate between field-observed overturned bedding and an interpretation of overturning that is based on analysis in the office. In this study, an overturned symbol on a field map was considered to be the correct symbol if the printed map showed a normal bedding symbol at that location and the symbol in the final data set was consequently changed. Such a change can have a profound effect on the structural interpretation (Figure 3).

Because field maps could be located for only a few quadrangles, this is not a rigorous regional investigation. However, it does bring to light some interesting aspects of data collection and map creation. Because each map has its own unique history, each map has a characteristic error rate. Some maps are more error prone than others. This is an intrinsic characteristic of the data and gives the user valuable information about the overall reliability of data from one map relative to data from another. If there was a discrepancy between two maps, which one should be reexamined? Even when it is not possible to create blunder rates for every quadrangle due to the lack of a comparison dataset, characterizing the error of a sample of quadrangles gives some measure of the reliability of the overall dataset.

Blunders occur in almost all datasets, whether or not they are recognized. The tendency is to not document the number of blunders found during editing. Yet the blunder rate is a characteristic of the dataset - a function of variables such as authorship and methodology. Characterizing the blunder rate of a dataset gives the user a better understanding of the limitations of that dataset.

REFERENCES

Duncan, I., and Campbell, E., 2005, Reliability in Geoinformatics: Theory and Practical Applications in Digital Geologic Map Compilations; Geological Society of America Abstracts with Programs, Vol. 37, No. 7.

Openshaw, S., 1989, Learning to live with errors in spatial databases in: Goodchild, M.F., and Gopal, S., ed., Accuracy in Spatial Databases: Taylor and Francis, Bristol PA, p. 263-276.

Three-Dimensional Geological Mapping for Groundwater Applications: Recent Activities

By H.A.J. Russell[1], L.H. Thorleifson[2], and R.C. Berg[3]

[1]Geological Survey of Canada
601 Booth Street
Ottawa, ON, K1A 0E8
e-mail: hrussell@nrcan.gc.ca

[2]Minnesota Geological Survey, 2642 University Ave. W., St Paul, MN, USA, 55114-1057

[3]Illinois State Geological Survey, Department of Natural Resources,
615 Peabody Drive, Champaign, Il, USA, 61820-6964

INTRODUCTION

In 2001 a workshop on three-dimensional (3-D) geological mapping was convened in conjunction with the Geological Society of America North-Central Sectional meeting in Normal, Illinois (Berg and Thorleifson, 2001). The purpose of the workshop was to provide a reference point for state, provincial, and federal government geologists involved in 3-D mapping for groundwater applications and to broadly discuss a host of geological issues pertaining to urbanization. The initial workshop was followed on an 18-month schedule by three workshops (Thorleifson and Berg, 2002; Berg et al., 2004; Russell et al., 2006a).

This document provides an overview of the workshop contents as well as issues and progress that participants have made in 3-D geological mapping (Table 1). It attempts to synthesize the 97 presentations made in the four workshops within five thematic groups: (i) basin analysis, (ii) data sets and data management, (iii) modeling, (iv) groundwater, and (v) communication (Table 2). Information on the workshops is available at the workshop website (http://www.isgs.uiuc.edu/3DWorkshop/nu_3d_welcome.html), where all the abstracts and some of the PowerPoint presentations can be found.

General Setting

The objective of the four workshops involved the development of interagency collaboration and improved methodologies. This parallels recommendations by the U.S. National Research Council (2000) for *Investigating Groundwater Systems on Regional and National Scales*. The National Research Council review highlights two important needs that connect geology and hydrogeology.

Table 1. Range of societal issues dealt with by workshop presentations.

Issue	Location	Sample reference
Transportation	Illinois	(Berg et al., 2002)
Vulnerability	South Carolina	(Shafer et al., 2002)
Resource assessment	Michigan	(Stone et al., 2002)
Hydrogeological framework model	Nevada & California	(Sweetkind et al., 2002)
Landuse planning	Ontario, Canada	Logan et al., 2004
Groundwater exploration	Ontario, Canada	(Russell et al., 2004)
Urban hazard mitigation	Seattle, Washington	(Troost et al., 2001)
Aquifer recharge	Finland	(Artimo et al., 2002)
Brownfields development	Manchester, Great Britain	(Bridge et al., 2004)
Geothermal resources	Poland	(Malolepszy, 2006)
National framework	Netherlands, England,	(Gunnink, 2006; Kessler et al., 2006)

Table 2. Thematic groupings of workshop presentations.

Theme	2001	2002	2004	2005
Basin Analysis	4	5	5	4
Data Sets and Data Management	2	1	2	2
3-D modeling	17	12	11	16
Groundwater studies	1	3	3	1
Communication	2	4	4	2
Miscellaneous	1	1	0	3

- Collaboration that permits geological information to be used in scaling up the results of local groundwater studies in areas where hydrogeological data are sparse or nonexistent.
- Characterization of heterogeneous aquifers at large and small scales through an understanding of links between geology and hydrogeology.

Regional Aquifer Systems Analysis (RASA) studies completed in the USA are one example, on a national scale, of a multi-agency approach to investigating groundwater (Sun and Johnson, 1994). On a more regional scale, the Central Great Lakes Geologic Mapping Coalition was formed between the geological surveys of four upper Midwestern states and the U.S. Geological Survey, with the intent of developing 3-D geologic models to address groundwater considerations (Berg et al., 1999). In Canada, a more recent attempt to foster and develop such an approach is illustrated by Rivera et al. (2003) in the *Canadian Framework for Collaboration on Groundwater*.

WORKSHOP SERIES

The organizers of the initial workshop (Berg and Thorleifson, 2001) recognized that there was a wealth of 3-D geological modeling knowledge developed by the mining and petroleum industries (e.g. Hughes, 1993; Houlding, 1994; Yarus and Chambers, 1994), but that there was considerably less such experience and knowledge reported for shallow glacial basins and comparable settings. As highlighted by Anderson (1989), it was apparent that many geologists working in glaciated basins were facing unique challenges in the construction of numeric 3-D models, and that much of the data was synthesized in geological models that are difficult to integrate into numeric groundwater flow models. This problem was exacerbated by the lack of geological collaboration and lack of integration of geological complexity in many groundwater models, a problem succinctly highlighted by Fogg (1986).

The first workshop was convened by the Illinois State Geological Survey and the Geological Survey of Canada, to encourage broader collaboration and knowledge transfer between geological surveys, industry, and academia. During the workshop it was apparent that many participants, while actively engaged in developing 3-D conceptual models, were doing so independently and were not aware that colleagues in other organizations were working to develop similar models. Successive workshops maintained the central theme of encouraging collaboration and knowledge transfer; however, the focus changed in response to the demands and expectations of the participants, as well as new technological innovations for managing data, portraying 3-D geology, and constructing derivative products tailored to specific needs of users.

Attendees at the workshops have been predominantly government survey scientists (e.g., state, provincial, federal) involved in geological mapping of surface and subsurface glacial sediment. To foster improved collaboration and information transfer between application (government surveys) and research (universities), each workshop included a number of academic researchers. Attendance has been predominantly from the northern American states and central Canada. Successive workshops, however, have had progressively more participation by Europeans (Finland, Great Britain, The Netherlands, Poland). The 2005 workshop also included non-presenting attendees from Korea, Pakistan, and China.

WORKSHOP CONTENT

The separate workshop presentations are grouped under five thematic headings (Table 2). First, basin analysis provided the framework for guiding the context of many of the workshop presentations. Presentations addressed data collection using geophysics, hydrochemistry, sedimentology, and the role of geological models. Second, although data management was an important element of many talks, and is considered absolutely essential for a viable 3-D mapping program; it was nevertheless the focus of relatively few presentations. Third, and dominating the workshop proceedings, were studies that focused on the generation of 3-D stratigraphic and hydrostratigraphic models. Workshop organizers emphasized to presenters the importance of discussing the building of internally consistent and fully integrated 3-D geologic solids models that represent the geometry, stratigraphy, hydrostratigraphy, and sedimentology of aquifer and aquiclude units. Proportionally fewer studies emphasized volume models and their population with physical parameter data. Fourth, to maintain connection with the needs of the groundwater community, there were a number of groundwater modeling presentations, and most emphasized that improved 3-D geological conceptual models were a necessity prior to groundwater modeling. Lastly, many presentations advocated the importance of 3-D models as visualization tools to improve communication between inter-disciplinary scientists, and most importantly between geologists and the user-community.

Basin Analysis

The importance of basin analysis as a framework for 3-D model construction was enunciated in workshop introductions (e.g. Thorleifson and Berg, 2002) and by Sharpe et al.(2002). A key theme of the presentation by Sharpe et al. was the need for new data to provide reliable benchmarks for understanding less reliable archival data. The core message, however, was the need for data that permit the geological history of the basin to be interpreted. It is only through understanding of the depositional environments and processes that acted in the basin that any predictive ability can be developed. This hypothesized framework is then available to guide additional data collection and to provide verification of the stratigraphic model. Sharpe presented a system for applying basin analysis techniques to ground water studies that progresses stepwise from database development to quantitative understanding of the flow system (Figure 1). Intermediate steps in this progression involve development of three stages of models - geological, hydrostratigraphic, and groundwater flow. This progression is similar to the workflow model of an ASTM standard for groundwater investigations (ASTM, 1998). It also echoes comments of LeGrand and Rosen (1998), who stress the need for development of first order, exploratory geological models before embarking on expensive data collection programs for site investigations.

Data Sets and Data Management

Most presentations addressed the exploitation of common data sets that included digital elevation models (DEM), geological maps, borehole information, and geophysical data. Only a few presentations, however, (Table 2) really focused on this issue. In a few studies, geochemical, hydrochemical, and hydraulic data were incorporated. In many cases, the geologic data were archival. Archival data generally refers to data previously collected by persons other than trained geologists, such as: water well records, geotechnical data, and data from hydrogeological studies. Archival data generally do not contain adequate geological detail to infer geological processes. This is also commonly true of core descriptions from hydrogeological studies that commonly use the Universal Soil Classification, which does not capture salient,

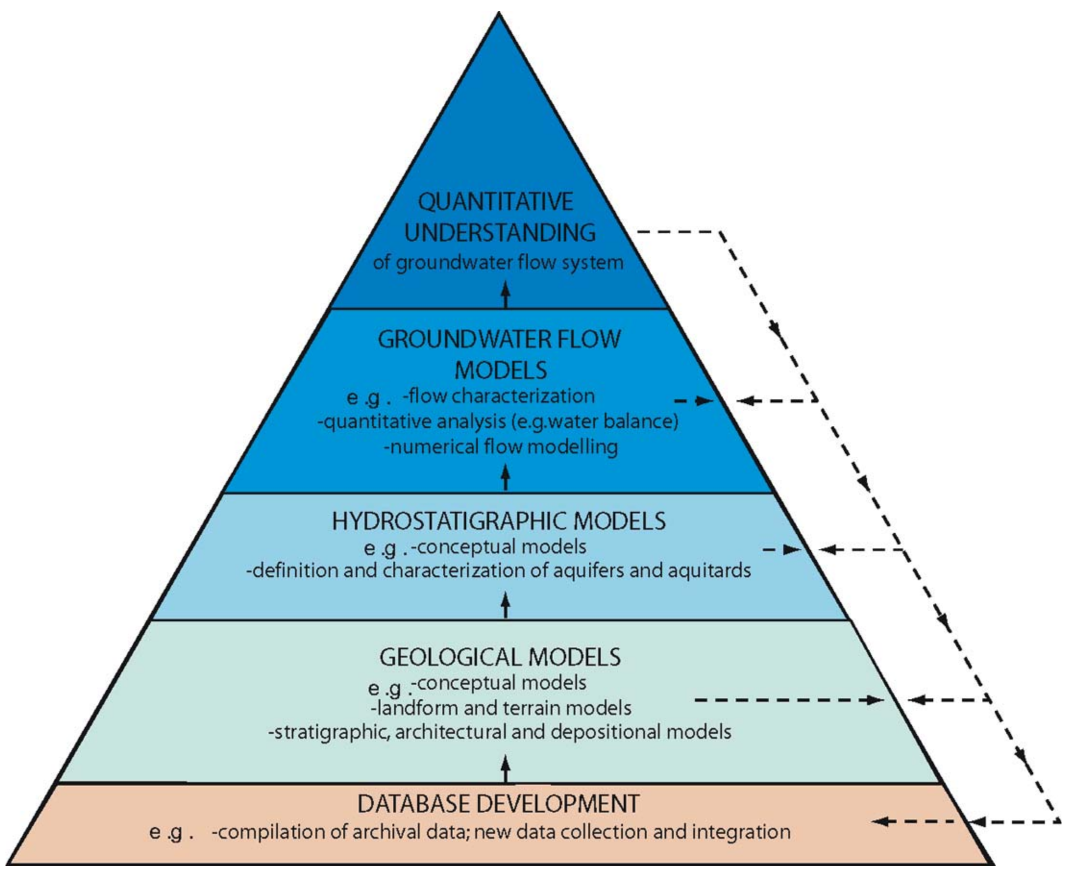


Figure 1. Simplified basin analysis approach used for regional hydrogeology analysis (Sharpe et al., 2002). The approach leads progressively from database development early in a study (base of figure) to quantitative understanding of groundwater flow systems as the study matures (top of figure).

essential geological information (e.g. Fogg et al., 1998). By comparing archival data to key stratigraphic control borings and/or seismic data, particularly in areas of sparse data, some inferences and correlations can be made (Russell et al., 2006a). This is essential for improving data coverage of large study areas. Table 3 summarizes key data sets used by various studies, and illustrates how these data were used to enhance model development.

Data management and standardization issues are central to the development of 3-D models, and more critically, to institutional progress in 3-D model development. Much of the effort in model construction goes toward data collation, preparation, standardization, and error checking (including data location verification). Development of robust, efficient data handling protocols is, consequently, essential to the cost-effective and timely delivery of individual models and institutional objectives (Gunnink, 2006; Kessler et al., 2006). A number of presentations alluded to these issues and the general workflow descriptions of model development (e.g., Figure 2).

The issue of data management is highlighted by the number of borehole records used in some of the regional models. For example, the Oak Ridges Moraine study has data from >100, 000 water well, geotechnical, hydrogeological, and shallow field mapping sites. In addition, continuous cored boreholes of up to 200 m can have more than 2000 logged units, 100's of grain-size and water content samples, complemented by 100's of digital images (e.g. Logan et al., 2004). Geophysical surveys, particularly in lacustrine or marine environments, can generate 10's to 1000's of kilometers of digital and analogue seismic reflection data (e.g. Thorleifson et al., 2001). At national levels, the scale of data management issues can be proportionally greater (Gunnink, 2006; Kessler et al., 2006). Fortunately, most national, state and provincial geological surveys are involved in various data standards exercises that are establishing standards from which individual studies are able to benefit (e.g. Soller, 2004; Soller and Berg, 2006).

Once data are formatted in a data structure, the work of data assessment, verification and standardization remains. In many cases, the most problematic data are the ubiquitous water well records. Issues of location accuracy are a common problem, and verification is a first step in using the data. Three different approaches were discussed, reflecting different reporting protocols (cadastral, geographic, postal address) in different jurisdictions - Illinois (Barnhardt et al., 2001), Manitoba (Thorleifson et al., 2002), and Ontario (Kenny et al., 1997).

Part of the process of "cleaning" data sets is the standardization of material descriptions to a common scheme with a limited number of unique descriptions. Once a standardized material coding scheme has been applied, all data records can be stratigraphically coded manually or programmatically. This issue was addressed in a variety of ways (Arnold et al., 2001; Ross et al., 2001; Logan et al., 2004).

3-D Modeling

Presentations at the four workshops have undergone a transition from a focus on data management issues to

Table 3. Model input data.

Category	Data set	Comments
Map or surface	DEM Bedrock geology maps Surficial geology maps Stratigraphic	Surface relief, model elevation datum Along with outcrop locations this data control structural surfaces Surface control on model units, length and width scales of geological features Subsurface units, pre-interpreted model layers
Archival	Waterwell records Geotechnical Hydrogeological	Stratigraphy, screen intervals Shallow stratigraphy, physical characteristics Physical characteristics, accompanied by hydrogeological data
Borehole	Continuous core Petroleum wells	Sedimentological detail, physical characteristics Bedrock stratigraphic picks
Hydro-geological	Hydrochemistry Hydraulic	Major and trace element, CFC, stable isotopes (Tritium, oxygen, …)
Geophysics	Downhole Seismic reflection Marine seismic	Gamma, conductivity, magnetic susceptibility, p and s-wave, neutron, … High resolution shallow seismic Digital and analogue high resolution

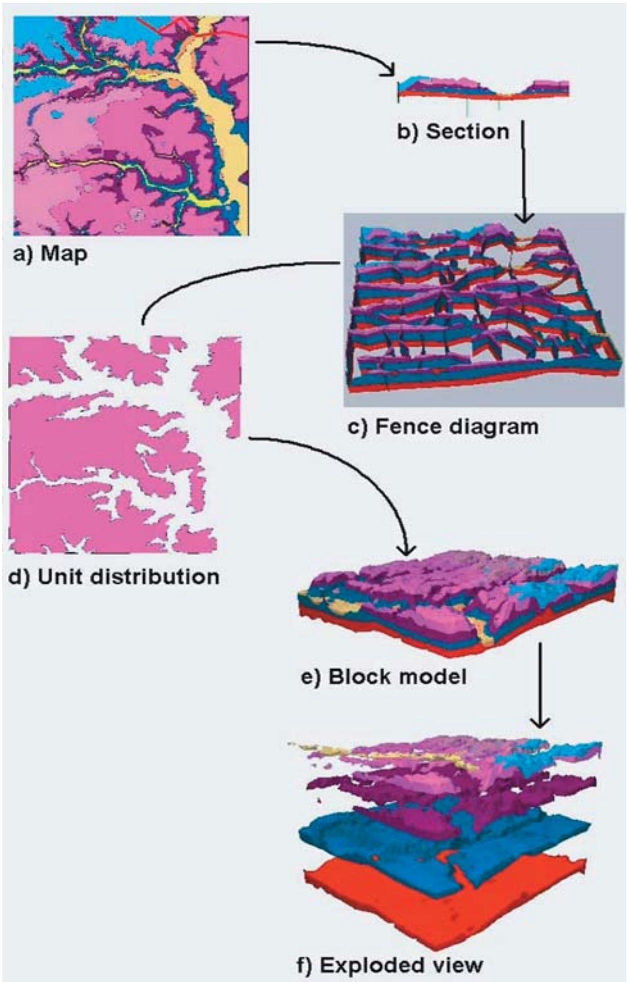


Figure 2. GSI3D (modeling software name) workflow (from Kessler et al., 2006).

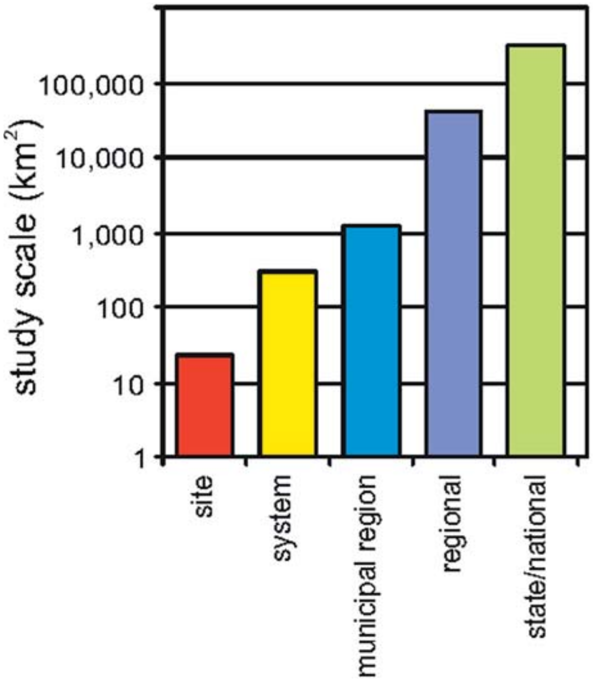


Figure 3. Summary of scale of models presented in the 2005 workshop.

presentations on 3-D models, 3-D modeling software, and applications of 3-D modeling (e.g. Datamine, ; gOcad, ; Rockworks, ; Shafer et al., 2006). There was also an increase in the number of presentations that focused on the construction of small-scale regional models of large jurisdictional areas (Figure 3). These regional models address an important issue raised by LeGrand and Rosen (1998), regarding the need for regional context prior to embarking on site-specific data collection.

3-D models have been presented for a range of study area scales and a variety of geological basins and depositional environments. Most of these studies have been regional in nature. Only a few studies have focused on modeling bedrock units (Arthur et al., 2002; Sweetkind et al., 2002) or constructing a model depicting both sediment and undeformed sedimentary rock strata (Ross et al., 2001; Thorleifson et al., 2001). A few studies have focused on specific depositional settings or landforms, for example eskers (e.g. Artimo et al., 2002; Bolduc et al., 2006).

Many studies have not clearly elucidated the approach to geological model development. Nevertheless a number of papers raise various philosophical and procedural issues regarding modeling. Philosophical issues of modeling pertain to the need for improved geological realism in models that lack adequate data support. This illustrates the difference between purely data-driven models (Logan et al., 2006) and intermediate models with an interpretive element imbedded via a synthetic data set (Soller et al., 1999; Kassenaar et al., 2004; Keller et al., 2006).

Kassenaar et al. (2004) presented one perspective on the capture of geological interpretation by blending hand drawn interpretations with on-screen stratigraphic assignments. This is an approach that has similarities with the method employed in the Red River area of Manitoba and Minnesota (Thorleifson et al., 2001; Thorleifson et al., 2002; Keller et al., 2006; Thorleifson et al., 2006). In these studies, synthetic stratigraphic picks were generated from vector surfaces along a series of cross-sections. By contrast Logan et al. (2006) used an iterative rules-based approach that stratigraphically coded archival data in a two-step process using a training data set and training surfaces. This approach has similarities to that employed in the Netherlands. Gunnink (2006) discussed an iterative process of model development where control surfaces and synthetic data are integrated into smaller scale models based on scaling of modeled features from larger scale models or by sampling earlier iterations of the model.

Many studies discussed the construction of 2.5-D models, with data management in a relational database

and 2-D GIS software such as ArcInfo or MapInfo. These models build structural surfaces of unit tops and then calculate isopachs of unit thicknesses. In some cases these 2.5-D models were imported into 3-D software and converted to volume models for visualization and analysis. In 2005, many studies were using 3-D software packages to generate complete volume models (e.g. Keller et al., 2006; Malolepszy, 2006; Ross et al., 2006; Thorleifson et al., 2006; Figure 4). An important aspect of modeling is the geostatistical realization of model surface or unit heterogeneity (Ritzi, 2002; Weissemann, 2006). This was most clearly enunciated by Weissemann in a presentation on the integration of geological map data and the subsequent conditional simulation of lower resolution subsurface data.

Groundwater

The application of geological models to hydrogeology is the *raison d'etre* of all four workshops. The workshop organizers stressed the importance of integrating increased geological knowledge in groundwater models. A 3-D geological conceptual model based on the best available data, predictive sedimentological models, and representing the geometry and stratigraphy of all materials, can only result in improved groundwater models. Presentations on the subject of hydrogeology described methods to regionalize physical property data, transform geological model grids to hydrogeological grids, integrate geological models, and develop database-supported flow mapping. These presentations touched on a range of societal issues dealing with groundwater: potable water supply, industrial extraction, energy supply, contaminant mitigation, and radioactive waste disposal. An integration of geology, hydrostratigrahy, flow modeling and database management was presented by Shafer et al. (2006) for Coastal Plain sediment in Beaufort, South Carolina.

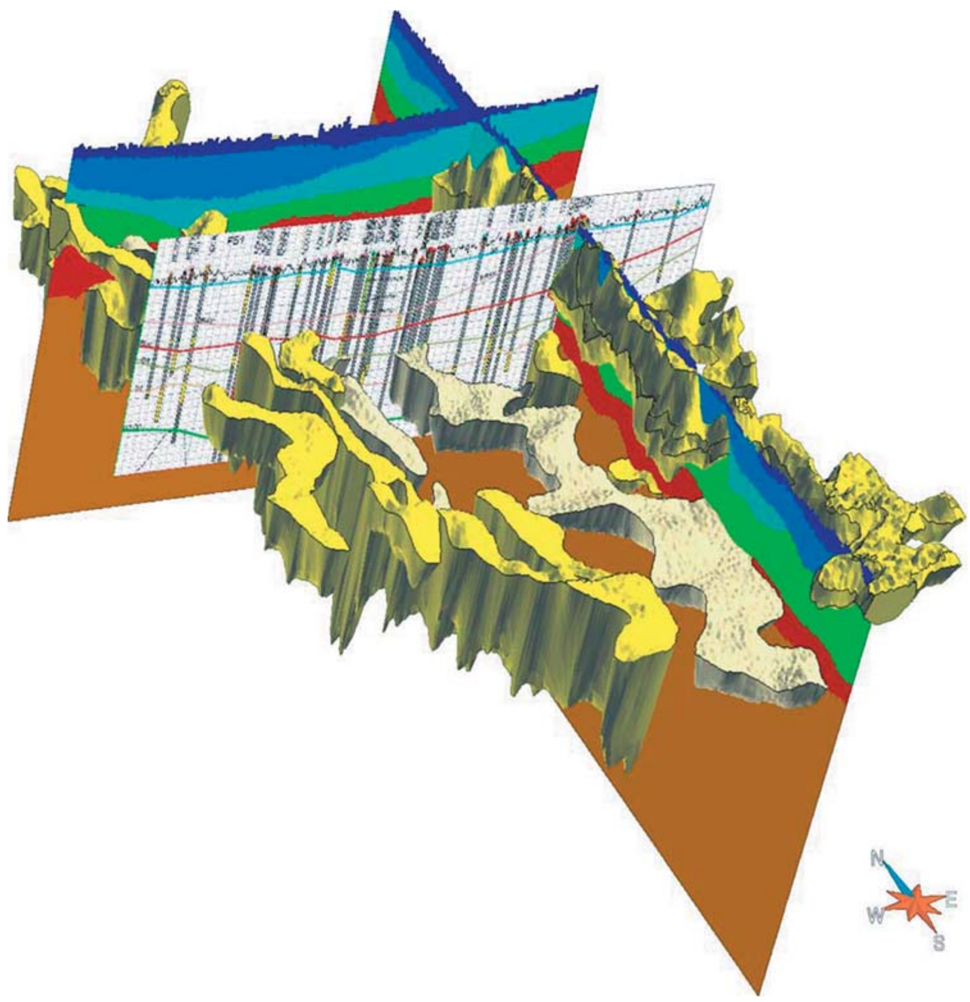


Figure 4. 3-D geology was interpreted on cross sections that were scanned in 3D space and utilized to construct 3-D solids; this example (Thorleifson et al., 2006) shows thick glacial lake clay in blue in the Fargo, North Dakota, area, and a thinner clay cover above a major aquifer in the eastern portion of that area, to the right. Note the north-south valley fills of sand interpreted as tunnel valleys.

Communicating the Vision and Value

The value of geological mapping has only rarely been quantified (Bhagwat and Ipe, 2000). A number of presentations addressed different aspects of the societal value and economic benefit of geological mapping (Hanmer, 2004; Jackson, 2004; Troost and Booth, 2006). Troost and Booth provided a perspective on the economic cost and value of geological mapping in the Seattle area, where they illustrated the need for renewed mapping as technology changes and societal needs and expectations within densely populated urban areas increase. They highlighted, the difference between the 1962 and 2005 geological maps of the area. Differences in surface sediment types and thicknesses can have significant implications for landuse planners, zoning regulations, and for framing site investigations. To illustrate the cost effectiveness of geological mapping, Troost and Booth compared the cost of mapping a 1:12,000-scale quadrangle with other activities of similar cost ($500,000) undertaken by municipal authorities. From this analysis, they showed that the cost of 3-D mapping is comparable to a single lawsuit, 3 ground water modeling studies, or 2 reconnaissance-level route selection studies.

DAWN OF A NEW ERA (WEB BASED MAP DELIVERY)

Recently there have been enormous developments in the promotion of geographic data through online programs such as Google Earth (http://earth.google.com). Also the British Geological Survey's national Lithoframe model (Table 4) and The Netherlands 3-D modeling program, illustrate approaches for 3-D geological mapping on national scales. Google Earth, and perhaps examples of large-scale 3-D models, will hopefully capture the imagination of viewers and serve valuable public outreach purposes. Early developments in the web-enabling of 3-D models are illustrated by links provided in Table 4. Berg and Soller (2006) demonstrated a web-enabled borehole tool for querying stacked-layer maps online. As highlighted at the third workshop by Jackson (2004) of the British Geological Survey, the fundamental purpose of national geological surveys (and state and provincial surveys as well) is serving public need. Google Earth has demonstrated that there is an overwhelming public interest in geospatial data when linked to issues or themes with which the public can identify, and in an interface that is easy to manage.

SUMMARY

Conflicting land use and urbanization problems stimulate interest for generating 3-D models for groundwater studies. Often in jurisdictions with abundant water, the resource is not equally distributed or accessible. Furthermore, dwindling aggregate resources, increasing practice of industrial agriculture, and the demand for alternative energy supplies (geothermal, coal bed methane, heavy oil) will ensure that groundwater resource and protection issues remain a central theme. Geologists need to embrace other disciplines (geostatistics, hydrochemistry, hydrogeology) as they attempt to move beyond first-order stratigraphic models to more complete models of the basin subsurface. It is evident that geologists have responded to a central recommendation of Anderson (1989), because many are now indeed familiar with 3-D numeric model construction, and can synthesize data in conceptual geological models and integrate data into flow models. It is also apparent that it is not only geologists in the USA and Canada who are grappling with this issue; it is of interest in many countries (e.g., Culshaw, 2005). Through the development of GIS-based digital geological models, geologists are becoming better positioned to support hydrogeological flow modeling exercises.

ACKNOWLEDGMENT

Reviews at the GSC by R. Knight and at the ISGS by J. Goodwin are much appreciated.

Table 4. Examples of some online 3-D models presented at the workshops.

Jurisdiction	Site Information	URL
Great Britain	Lithoframe; national model of Great Britian	http://www.bgs.ac.uk/science/3Dmodelling/lithoframe1M.html
Netherlands	National model of Netherlands	http://www.nitg.tno.nl/ned/projects_new/pdf_s/2_09eng.pdf
Manitoba	3-D Model of Manitoba	http://www.gov.mb.ca/iedm/mrd/geo/3dmodel/index.html#introduction
Illinois	East-central 3-D model	http://ngmdb.usgs.gov/ecill/
Seattle	GeoMapNW	http://geomapnw.ess.washington.edu/

REFERENCES

Anderson, M.P., 1989, Hydrogeologic facies models to delineate large-scale spatial trends in glacial and glaciofluvial sediments: Geological Society of America Bulletin, v. 101, p. 501-511.

ArcInfo, http://www.esri.com/.

Arnold, T.L., Friedel, M.J., and Warner, K.L., 2001, Hydrogeological inventory of the upper Illinois River Basin—Creating a large database from well construction records, *in* Berg, R.C., and Thorleifson, L.H., eds., Geological Models for Groundwater Flow Modeling, 35th annual meeting, north-central section GSA: Bloomington-Normal, Illinois, Illinois State Geological Survey, Open File Series 2000-1, p. 1-5.

Arthur, J.D., Cichon, J., Baker, A., Marquez, J., Rudin, A., and Wood, A., 2002, Hydrogeological mapping and aquifer vulnerability modeling in Florida: 2D and 3D data analysis and visulaization, *in* Thorleifson, L.H., and Berg, R.C., eds., Three-Dimensional Geological Mapping for Groundwater Applications: Workshop Extended Abstracts: Ottawa, Ontario, Geological Survey of Canada, Open File 1449, p. 1-4.

Artimo, A., Berg, R.C., Abert, C.C., J, and Salonen, V.P., 2002, Three-dimensional geological modeling of the Virttaankangas aquifer, southwestern Finland, *in* Thorleifson, L.H., and Berg, R.C., eds., Three-Dimensional Geological Mapping for Groundwater Applications: Workshop Extended Abstracts: Ottawa, Ontario, Geological Survey of Canada, Open File 1449, p. 5-7.

ASTM, 1998, Standard guide for conceptualization and characterization of ground-water systems, D 5979-96, American Society for Testing Materials (ASTM) standards on design, planning, and reporting of ground water and vadose zone investigations, 2nd edition (Second Edition ed.): West Conshohocken, PA, ASTM, p. 278-284.

Barnhardt, M.L., Hansel, A.K., and Stumpf, A.J., 2001, Developing the database for 3-D modeling: aquiring, assembling, verifying assessing, interpreting and integrating source data, *in* Berg, R.C., and Thorleifson, L.H., eds., Geological Models for Groundwater Flow Modeling, 35th annual meeting, north-central section GSA: Bloomington-Normal, Illinois, Illinois State Geological Survey, Open File Series 2000-1, p. 6.

Berg, R.C., Bleuer, N.K., Jones, B.E., Kincare, K.A., Pavey, R.R., and Stone, B.D., 1999, Mapping the glacial geology of the central Great Lakes region in three-dimensions—a model for State-Federal cooperation: U.S. Geological Survey, Open-File Report 99-349.

Berg, R.C., McKay, E.D., Keefer, D.A., Bauer, R.A., Johnstone, P.D., Stiff, B.J., Pugin, A., Weibel, C.P., Stumpf, A.J., Larson, T.H., Su, W.J., and Homrighous, G.T., 2002, Three-dimensional geological mapping for transportation planning in central-northern Illinois: Data selectioin, map construction, and model development, *in* Thorleifson, L.H., and Berg, R.C., eds., Three-Dimensional Geological Mapping for Groundwater Applications: Workshop Extended Abstracts: Ottawa, Ontario, Geological Survey of Canada, Open File 1449, p. 13-17.

Berg, R.C., Russell, H.A.J., and Thorleifson, L.H., 2004, Three-dimensional geological mapping for groundwater applications: Illinois State Geological Survey, Open File Series 2004-8.

Berg, R.C., and Soller, D.R., 2006, Toward a goal of a "Super" three-dimensional geological map, *in* Russell, H.A.J., Berg, R.C., and Thorleifson, L.H., eds., Three-dimensional geological mapping for groundwater applications, Geological Survey of Canada, Open File 5048, p. 11-16.

Berg, R.C., and Thorleifson, L.H., 2001, Geological Models for Groundwater Flow Modeling: Illinois State Geological Survey, Open File Series 2000-1.

Bhagwat, S.B., and Ipe, V.C., 2000, Value of geological quadrangle maps of Kentucky: Illinois State Geological Survey, Special Report 3.

Bridge, D.M., Hough, E., Kessler, H., Lelliott, M., Price, S.J., and Reeves, H.J., 2004, Integrating modelling of geoscience information to suport sustainable urban planning, Greater Manchester Area, Northwest England, *in* Berg, R.C., Russell, H.A.J., and Thorleifson, L.H., eds., Three-dimensional geological mapping for groundwater applications, Illinois State Geological Survey, Open File Series 2004-8, p. 16-19.

Culshaw, M.G., 2005, From concept toward reality: developing the attributed 3D geological model of the shallow subsurface: Quarterly Journal of Engineering Geology and Hydrogeology, v. 38, p. 231-284.

Datamine, http://www.datamine.co.uk/.

Fogg, G.E., 1986, Groundwater flow and sand body interconnectedness in a thick multiple-aquifer system: Water Resources Journal, v. 22, p. 679-694.

Fogg, G.E., Noyes, C.D., and Carle, S.F., 1998, Geological based model of heterogeneous hydraulic conductivity in an alluvial setting: Hydrogeology Journal, v. 6, p. 131-143.

gOcad, 2005, http://www.earthdecision.com/.

Google Earth, http://earth.google.com.

Gunnink, J.L., 2006, Geological mapping in the Netehrlands: The case of the shallow subsurface (-550 m), *in* Russell, H.A.J., Berg, R.C., and Thorleifson, L.H., eds., Three-dimensional geological mapping for groundwater applications: Workshop Extended Abstracts, Geologial Survey of Canada, Open File 5048, 104 p, p. 26-29.

Hanmer, S., 2004, Cooperative geological mapping strategies across Canada: What relevance for groundwater?, *in* Berg, R.C., Russell, H.A.J., and Thorleifson, L.H., eds., Three-dimensional geological mapping for groundwater applications, Illinois State Geological Survey, Open File Series 2004-8, p. 21-22.

Houlding, S.W., 1994, 3D Geoscience Modeling, Computer Techniques for Geological Characterization, Springer Verlag, 309 p.

Hughes, J.D., 1993, GEOMODEL; an expert system for modeling layered geological sequences applied to the assessment of coalfields, *in* Kirkham, R.V., Sinclair, W.D., Thorpe, R.I., and Duke, J.M., eds., Mineral Deposit Modeling, Geological Association of Canada, Special Paper 40, p. 707-734.

Jackson, I., 2004, Communicating the Vision?, *in* Berg, R.C., Russell, H.A.J., and Thorleifson, L.H., eds., Three-dimensional geological mapping for groundwater applications, Illinois State Geological Survey, Open File Series 2004-8, p. 32-34.

Kassenaar, D., Gerber, R., and Holysh, S., 2004, Merging conceptual insight and secondary indicators into the hydrogeological modeling process: Examples from the Oak Ridges Moraine, southern Ontario, *in* Berg, R.C., Russell, H.A.J., and Thorleifson, L.H., eds., Three-dimensional geological

mapping for groundwater applications, Illinois State Geological Survey, Open File Series 2004-8, p. 35-39.

Keller, G.R., Matile, G.L.D., Thorleifson, L.H., and Malolepszy, Z., 2006, 3D geological model of the Red River Valley, central North America, *in* Russell, H.A.J., Berg, R.C., and Thorleifson, L.H., eds., Three-dimensional geological mapping for groundwater applications, Geological Survey of Canada, Open File 5048, p. 34-37.

Kenny, F.M., Chan, P., and Hunter, G., 1997, Quality control of the Positional Accuracy of Records in Ontario's Water well database using automated GIS techniques, *in* Conference Proceedings Geomatics in the Era of RADARSAT GER 97, Ottawa, May 24-30, 1997, http://gsc.nrcan.gc.ca/hydrogeo/orm/pdf/kenny_et_al_1997.pdf.

Kessler, H., Lelliott, M., Bridge, D.M., Ford, J., Sobusch, H.-G., Mathers, S., Price, S., Merritt, J., and Royse, K., 2006, 3D Geoscience models and their delivery to customers, *in* Russell, H.A.J., Berg, R.C., and Thorleifson, L.H., eds., Three-dimensional geological mapping for groundwater applications, Geological Survey of Canada, Open File 5048, p. 38-41.

LeGrand, H.E., and Rosen, L., 1998, Putting hydrogeological site studies on track: Ground Water, v. 36, p. 193-194.

Logan, C., Russell, H.A.J., and Sharpe, D.R., 2004, A stratigraphic database approach to 3-D modelling in the Oak Ridges Moraine Area, southern Ontario, *in* Berg, R.C., Russell, H.A.J., and Thorleifson, L.H., eds., Three-dimensional geological mapping for groundwater applications, Illinois State Geological Survey, Open File Series 2004-8, p. 45-47.

Logan, C., Russell, H.A.J., Sharpe, D.R., and Kenny, F.M., 2006, The role of expert knowledge, GIS and geospatial data management in a basin analysis, Oak Ridges Moraine, southern Ontario, *in* Harris, J., ed., GIS Applications in the Earth Sciences, Geological Association of Canada, p. 523-545.

Malolepszy, Z., 2006, Three-dimensional geological model of Poland and its application to geothermal resource assessment, *in* Russell, H.A.J., Berg, R.C., and Thorleifson, L.H., eds., Three-dimensional geological mapping for groundwater applications, Geological Survey of Canada, Open File 5048, p. 45-48.

MapInfo, http://www.mapinfo.com/.

National Research Council, 2000, Investigating groundwater systems on regional and national scales: Committee on USGS Water Reources Research, Water Science and Technology Board: Washington, DC., National academy Press, National Research Council.

Ritzi, R.W.J., 2002, Three-dimensional aquifer mapping using indicator geostatistics, *in* Thorleifson, L.H., and Berg, R.C., eds., Three-Dimensional Geological Mapping for Groundwater Applications: Workshop Extended Abstracts: Ottawa, Ontario, Geological Survey of Canada, Open File 1449, p. 51.

Rivera, A., Crowe, A., Kohut, A., Rudolph, D., Baker, C., Pupek, D., Shaheen, N., Lewis, M., and Parks, K., 2003, Canadian framework for collaboration on groundwater: Ottawa, Ontario, Geological Survey of Canada, 55 p.

Rockworks, 2006, http://www.rockware.com/set.html.

Ross, M., Aitssi, L., Martel, R., and Parent, M., 2006, From geological to groundwater flow models: an example fo inter-operability for semi-regular grids, *in* Russell, H.A.J., Berg, R.C., and Thorleifson, L.H., eds., Three-dimensional geological mapping for groundwater applications, Geological Survey of Canada, Open File 5048, p. 62-64.

Ross, M., Parent, M., Michaud, M., Boisvert, E., and Girard, F., 2001, On the construction of 3D geological models for applicatioins in regional hydrogeology in complex Quaternary terrains of eastern Canada, *in* Berg, R.C., and Thorleifson, L.H., eds., Geological Models for Groundwater Flow Modeling, 35th annual meeting, north-central section GSA: Bloomington-Normal, Illinois, Illinois State Geological Survey, Open File Series 2000-1, p. 34-37.

Russell, H.A.J., Berg, R.C., and Thorleifson, L.H., 2006, Three-dimensional geological mapping for groundwater applications: Geological Survey of Canada, Open File 5048.

Russell, H.A.J., Pullan, S.E., Hunter, J.A., Sharpe, D.R., and Holysh, S., 2004, Buried-valley aquifers: Delineation and characterization from reflection seismic and core data Caledon East, Ontario, *in* Berg, R.C., Russell, H.A.J., and Thorleifson, L.H., eds., Three-dimensional geological mapping for groundwater applications, Illinois State Geological Survey, Open File Series 2004-8, p. 73-76.

Shafer, J.M., Covington, E., Rine, J.M., and Berg, R.C., 2006, Incorporation of groundwater flow into traditional sensitivity analysis of a southeast coastal plain layered aquifer system, *in* Russell, H.A.J., Berg, R.C., and Thorleifson, L.H., eds., Three-dimensional geological mapping for groundwater applications, Geological Survey of Canada, Open File 5048, p. 70-74.

Shafer, J.M., Rine, J.M., Waddell, M.G., and Berg, R.C., 2002, 3D hydrogeological characterization of the marine corps air station at Beaufort, South Carolina for aquifer vulnerability analysis and groundwater flow and transport modeling, *in* Thorleifson, L.H., and Berg, R.C., eds., Three-Dimensional Geological Mapping for Groundwater Applications: Workshop Extended Abstracts: Ottawa, Ontario, Geological Survey of Canada, Open File 1449, p. 58-61.

Sharpe, D.R., Hinton, M.J., Russell, H.A.J., and Desbarats, A.J., 2002, The need for basin analysis in regional hydrogeological studies: Oak Ridges Moraine, Southern Ontario: Geoscience Canada, v. 29, no. 1, p. 3-20.

Soller, D.R., 2004, Facillitating 3-D analysis with standardization Information - The role of the National Geological Map Database, *in* Berg, R.C., Russell, H.A.J., and Thorleifson, L.H., eds., Three-dimensional geological mapping for groundwater applications, Illinois State Geological Survey, Open File Series 2004-8, p. 85-88.

Soller, D.R., and Berg, R.C., 2006, The national geological map database project: Overview and progress, *in* Soller, D.R., ed., DMT'03 - Digital Mapping techniques Proceedings, U.S. Geological Survey, Open-file Report 03-471.

Soller, D.R., Price, S.D., Kempton, J.P., and Berg, R.C., 1999, Three-dimensional geological maps of Quaternary sediments in east-central Illinois, USGS I-Map 2669, Three atlas sheets.

Stone, B.A., Kincare, K.A., and O'Leary, D.W.O., 2002, Glacial geology mapping in Barrien County, Michigan: Resolving the third dimension for increasing the accuracy of resource assessment, *in* Thorleifson, L.H., and Berg, R.C., eds., Three-Dimensional Geological Mapping for Groundwater Applications: Workshop Extended Abstracts: Ottawa, Ontario, Geological Survey of Canada, Open File 1449, p. 67-70.

Sun, R.J., and Johnson, R.H., 1994, Regional Aquifer-Systems Analysis Program of the U.S. Geological Survey, 1978-92:

U.S. Geological Survey, Circular 1099.

Sweetkind, D.S., Belcher, W.R., and Faunt, C.C., 2002, Hydrogeological framework construction, using an example from the Death Valley ground-water flow system, Nevada and California, USA, *in* Thorleifson, L.H., and Berg, R.C., eds., Three-Dimensional Geological Mapping for Groundwater Applications: Workshop Extended Abstracts: Ottawa, Ontario, Geological Survey of Canada, Open File 1449, p. 71-74.

Thorleifson, L.H., Berg, J.A., Tipping, R., G, Harris, K.L., Lusardi, B.A., Setterholm, D.R., Malolepszy, Z., and Anderson, F.J., 2006, 3D geological modeling in support of groundwater inventory in the Fargo-Moorhead region, Minnesota and North Dakota, *in* Russell, H.A.J., Berg, R.C., and Thorleifson, L.H., eds., Three-dimensional geological mapping for groundwater applications, Geological Survey of Canada, Open File 5048, p. 92-95.

Thorleifson, L.H., and Berg, R.C., 2002, Three-Dimensional Geological Mapping for Groundwater Applications: Workshop Extended Abstracts: Geological Survey of Canada, Open File 1449.

Thorleifson, L.H., Matile, G.L.D., Keller, G.R., and Pyne, D.M., 2002, Construction of a geological model for southern Manitoba for groundwater modelling, *in* Thorleifson, L.H., and Berg, R.C., eds., Three-Dimensional Geological Mapping for Groundwater Applications: Workshop Extended Abstracts: Ottawa, Ontario, Geological Survey of Canada, Open File 1449, p. 75-78.

Thorleifson, L.H., Matile, G.L.D., Pyne, D.M., and Keller, G.R., 2001, Construction of a geological model of the Winnipeg region for groundwater modeling, *in* Berg, R.C., and Thorleifson, L.H., eds., Geological Models for Groundwater Flow Modeling, 35th annual meeting, north-central section GSA: Bloomington-Normal, Illinois, Illinois State Geological Survey, Open File Series 2000-1, p. 52-54.

Troost, K.G., Booth, D.A., Shimel, S.A., and O'Neal, M.A., 2001, The Seattle-area geological mapping project and the geological framework of Seattle, *in* Berg, R.C., and Thorleifson, L.H., eds., Geological Models for Groundwater Flow Modeling, 35th annual meeting, north-central section GSA: Bloomington-Normal, Illinois, Illinois State Geological Survey, Open File Series 2000-1, p. 58-61.

Troost, K.G., and Booth, D.B., 2006, Cost of 1:12,000-scale geological map; $500,000: Cost of 3D data, priceless, *in* Russell, H.A.J., Berg, R.C., and Thorleifson, L.H., eds., Three-dimensional geological mapping for groundwater applications: Workshop Extended Abstracts, Geologial Survey of Canada, Open File 5048, p. 96-99.

Weissemann, G., 2006, Application of transition probability geostatistics in a detailed stratigraphic framework, *in* Russell, H.A.J., Berg, R.C., and Thorleifson, L.H., eds., Three-dimensional geological mapping for groundwater applications: Workshop Extended Abstracts, Geologial Survey of Canada, Open File 5048, p. 100-103.

Yarus, J.M., and Chambers, R.L., eds., 1994, Stochastic modeling and geostatistics: principles, methods, and case studies: AAPG Computer Applications in Geology, no 3, The American Association of Petroleum Geologists, 379 p.

Glacier inventory and temporal database of glacier change in the U.S. West, exclusive of Alaska

By David Percy and Andrew Fountain

Portland State University
Department of Geology
1721 SW Broadway
Portland, OR 97201
Telephone: (503) 725-3373
e-mail: percyd@pdx.edu

ABSTRACT

A comprehensive database of glacier extents for the western United States is being compiled in order to document glacier change over the period of the 20th century. Sources for the database include printed historical maps, aerial and oblique photographs, digital data from USGS 1:24000-series topographic maps, and satellite data from NASA. Included in the online interactive database will be scanned versions of all primary source maps. The URL for this project is http://glaciers.us

MOTIVATION

Our motivations for this study range from applied science to testing new remote sensing image-processing techniques. For example, we are seeking to understand the hydrologic inputs that glaciers provide during the dry months of the summer. We postulate and hope to quantify the extent to which glaciers provide a buffer to the hydrologic system during the nearly rain-free months of August and September. It should also be possible to estimate the increased amount of runoff that has been available to the fish-bearing streams over the last century as glaciers have receded and given up their water more rapidly during these hot, dry months.

If we are correct and glaciers are shown to be a significant source of late summer water to the west's streams, this will be important information for local planners in terms of water budgets. We anticipate showing that glacier input affects both the quantity and the temperature of stream water, with concomitant effects on viability of streams for fish habitat.

Another motivation and product expected from this work is simply a full inventory of where perennial ice is, or has been, located in the western U.S. We will also include paleo-glaciers as documented by moraines.

We hope this database will provide "ground-truthing" for climate modelers. As a first attempt at this, we have had students use the database in conjunction with precipitation, temperature, and elevation data to perform logistic regression. This should predict the presence or absence of glaciers based on these physical parameters. So far, results have been mixed for various physiographic areas. For example, Washington's regression analysis had reasonable results, whereas Montana's did not. This leads us to look for other factors to include in the regression analysis.

EXPECTED PRODUCTS

Based on this team's experience in compiling similar databases for the Northern Cascades and several stratovolcanoes in the Pacific Northwest, we expect to find three timeslices for each mapped area. Usually there will be available some early historical data such as old photographs and/ or maps from the early 20th century. The second time slice of data will be from the USGS 1:24,000-series topographic mapping effort of the 1950s and 60s. The third data set will be satellite data from the 1970s and later, collected by NASA and the EROS data center. These will primarily be ASTER, 15m panchromatic images.

SOURCES OF DATA

The primary data set that ties all other data together is the 1:24,000-scale topographic series. In order to assemble these data, we initially queried the USGS hydrographic data, but unfortunately the data compilation is quite incomplete. Next we turned to the U.S. Forest Service (USFS) and were given a digital set of GIS data of glacier outlines that had been digitized from 1:24,000-scale USGS maps. These data ended up being the starting point for our compilation. The National Park Service (NPS) has been a reliable source for many data sets, and luckily many glaciated regions occur within parks. Historical maps and photos are available from many sources, including local hiking and climbing clubs.

METHODS

For completing the coverage of glaciers found on USGS topographic maps we overlaid the lines received from the USFS on either Digital Raster Graphics from the USGS or a similar data product from the USFS know as PBS (Portable Base Series). We then checked each line against the reference map to verify that it was indeed a glacier. In some cases, the lines were actually lakes, whereas in other cases they were moraines. In the former case we deleted the lines, while in the latter we saved them for a future study of rock-covered glaciers. Whenever we encountered glaciers on the map sheet that were not depicted by lines within the database we manually digitized them. By hand-scanning the maps at a scale of 1:12,000 we were able to ensure ourselves that we were not missing any glaciers, and that all lines were actually representing perennial ice features.

After compiling these data as a set of lines, we ran a “lines to polygons” conversion with a tolerance set to weed out densely-distributed points created by over-zealous digitizers. We met USGS National Map Accuracy standards for 1:24,000-scale data—one vertex for every 1.7 meters on the ground. In cases where the area near polygon closures did not have sufficient vertices to meet the standard, we were required to digitize more vertices.

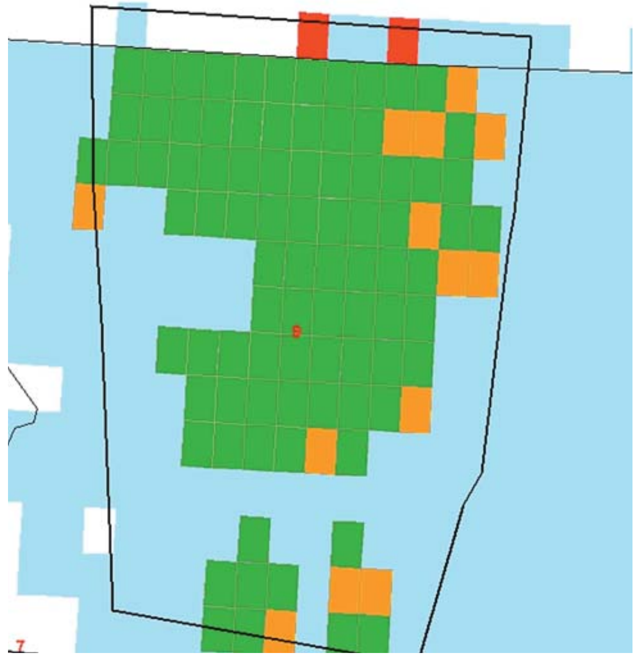


Figure 1. Coverage of quads in the North Cascades, WA. Different colors represent different agency sources (USGS, USFS); yellow (or light gray) quads are USGS, green (or medium gray) quads are USFS, red (or dark gray) quads are not available and blue (or pale gray) quads are not needed for this study.

PROBLEMS ENCOUNTERED

A variety of problems could be expected in any large scale undertaking such as this, and we encountered our share. To begin with, the incomplete coverage of quad maps from any central source required us to seek base-maps from many sources, including state agencies, BLM, USFS, and USGS (Figure 1).

Not uncommonly, lines ended at quadrangle boundaries although the features actually spanned quadrangles. Also, in some cases the lines did not match across quadrangle boundaries. As already noted, there were glaciers missing from the source data sets that needed to be digitized, incorrectly coded lines, and donut polygons and shifted polygons (Figure 2).

Additionally we have encountered the problem of determining which dated set of photos were used to digitize the glaciers depicted on a map. For example, many maps have an initial date of photo acquisition with additional dates of update. In many cases, the glacial features have not been updated at the same time as the cultural features. In most cases we must assume that the earliest date of photographic acquisition is the same date for the glaciers shown on the map.

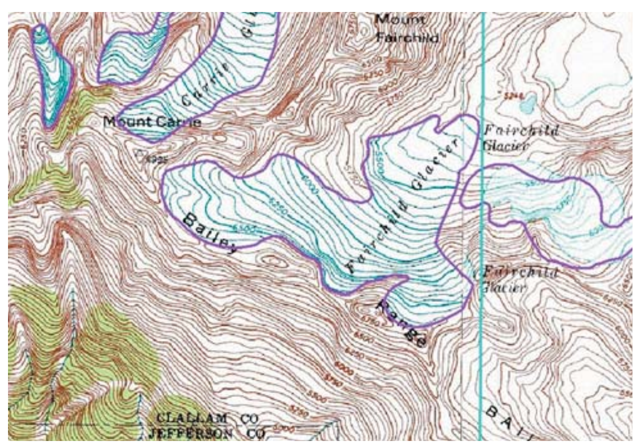


Figure 2. Offset glacier in the Olympic Range. Note that four polygons are in place but one, near the eastern edge, is offset to the west (to left in figure).

A conceptual problem encountered is that of identifying a glacier through time. It would be a phenomenal task to go through and individually manage in the database each glacier as it changes over time. This would require assigning every glacier object a unique identification and then relating each change in size back to that object. This is particularly problematic in a GIS sense, as a glacier breaks into smaller ice patches. We have chosen to solve this problem by managing the glaciers at the watershed level. Since this project was funded as a hydrology issue, this is consistent with our charter. Thus we sum all glacier volumes in the watershed, and then calculate the volume changes in the watershed over time.

CURRENT STATUS

We can proudly announce that after one year of this study, all glaciers found on 1:24000-scale maps have been digitized. There are still issues related to attributes for the individual glaciers, such as the above-mentioned photo date. We also struggle with generating accurate zonal data from the underlying elevation data such as slope and aspect.

Many historical time-slices are available from various agencies and clubs, and digitizing is in progress. In addition, various data sets have been acquired from the various Parks Service agencies. These need to be converted to our data format, in some cases this means converting terminus positions to polygons.

An assets database that links the actual glacier timeslices to the underlying photographic, map or satellite data is under development. In this system the user will be able to click on an individual glacier outline and see the sources that were used for determining the outlines of the glaciers.

ACKNOWLEDGMENTS

We are indebted to our hard-working team of graduate student research assistants for their efforts on this project. These include: Chad Delaney, Amy Ebnet, Hassan Basagic, Matt Hoffman, Keith Jackson, Thomas Nylen, Peter Sniffen, Frank Granshaw, and Jon Ebnet. See the “About Us” page at http://glaciers.us for more information. The various public agencies have also been a joy to work with, and finally we acknowledge the funding from NASA and NSF to pay for many parts of this work.

LASED Geodatabase: A Tool to Manage, Analyze, Distribute, and Archive Geologic Data from the Louisiana Coastal Zone

By Shawn V. Dadisman[1], James G. Flocks[2], and Karynna Calderon[3]

[1]U.S. Geological Survey
Florida Integrated Science Center
600 4th Street South
St. Petersburg, FL 33701
Telephone: (727) 803-8747 x3029
Fax: (727) 803-2032
e-mail: sdadisman@usgs.gov
[2]U.S. Geological Survey, e-mail: jflocks@usgs.gov
[3]Computer Sciences Corporation, e-mail: kcalderon@usgs.gov

ABSTRACT

The Louisiana Sedimentary and Environmental Database (LASED) is the result of combined efforts of the U.S. Geological Survey (USGS) and State and academic cooperators to manage decades of geologic data gathered from the Louisiana coastal zone. The database incorporates a wide range of data types: sediment-sample descriptions and analyses, geophysical profiles, raster-image basemaps, logbooks, etc. The data is integrated with spatial and attribute information to provide processing and visualization capabilities using standard Geographic Information System (GIS) and Internet-browsing tools. The data types are linked so that complex queries and analyses can be performed across datasets. Decades-old to recently acquired analog data are included through the use of new technology and processing techniques. The integrated geodatabase is quickly and easily expanded and serves as a digital archive of almost any type of data. The development of the geodatabase is in response to a growing need for the USGS and collaborators to efficiently access coastal geologic data for shoreline management issues. Full access to LASED data is available to registered users via the Intranet, and public access to view map products and data is available over the Internet.

INTRODUCTION

The coastal region of Louisiana is currently under stress; human impact and natural processes have resulted in severe erosion of the shoreline. Wetland deterioration and land-loss rates due to subsidence, manmade alteration, and tropical storms exceed those found elsewhere in the Gulf of Mexico (Morton and others, 2004). A management plan is needed for long-term shoreline protection of the Louisiana coastline (Louisiana Coastal Wetlands Conservation and Restoration Task Force and the Wetlands Conservation and Restoration Authority, 1998). This plan must be based on a scientific understanding of the geologic history and processes involved in shoreline change, which includes an efficient and comprehensive use of all available scientific information.

The river deltas and barrier-island shorelines that protect wetlands and developed regions of Louisiana are the result of a complex series of fluvial and marine depositional events that have occurred over the past 4,000 years (Frazier, 1967). Identifying the stratigraphic framework of these past events is important to define the region's coastal evolution and to locate adequate resources for coastal management. In response to these needs, the USGS, in cooperation with the Louisiana Geological Survey, Louisiana Department of Natural Resources (LDNR), and University of New Orleans, is actively collecting scientific information associated with near-surface (upper 100 meters) geology and geomorphology of the coastal zone. This information is integrated with geospatial data to produce a geodatabase that includes various forms of georeferenced basemaps, such as digital elevation models (DEMs) and satellite imagery (Figure 1). The geodatabase contains geographic positions and links to attribute information for sediment cores, seismic-reflection profiles, sidescan sonar mosaics, bathymetry, and numerous basemap features. The geodatabase system is dynamic and flexible; a wide variety of data is included on an ongoing basis from both historical and recent sources. The system is highly interactive in that it allows for cross-referencing of different types of data as well as links to displays of the data. LASED is inclusive of the Louisiana coastal zone and a component of a much larger geodatabase system developed by the USGS Florida Integrated Science (FISC), St. Petersburg. Database management and data visualization are actively maintained to keep pace with developing technology. The

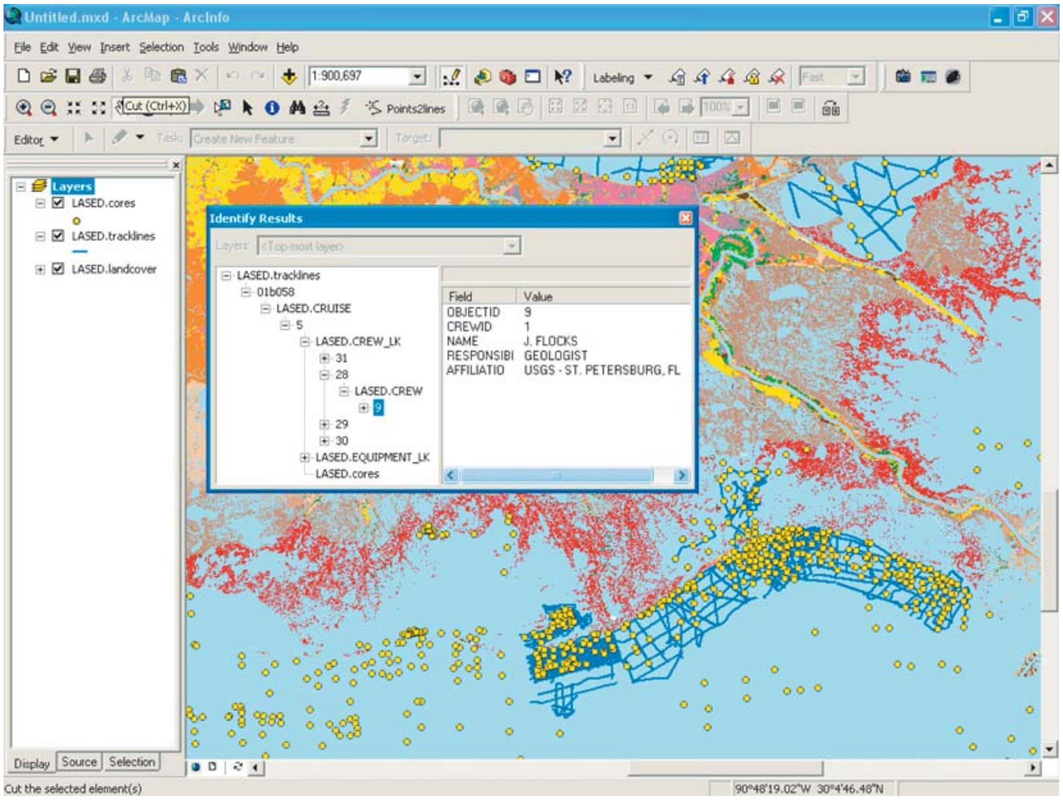


Figure 1. Display of LASED geodatabase using ArcMap. Data layers include seismic-reflection profile locations (lines) and sediment-core sites (dots) along the Louisiana coastal zone. Raster basemap feature is landcover produced by the Louisiana Gap Analysis Project of the USGS Biological Research Division’s National Wetlands Research Center (NWRC) and based partly on Thematic Mapper (TM) satellite imagery. Inset shows an example of the linked crew attribute table information.

geodatabase also serves as the infrastructure for a permanent online digital data-archive system and it controls access to all associated information, such as logbooks, photographs, analyses, and publications (Figure 2).

DATA COLLECTION AND MANAGEMENT

Although most data currently collected by scientists are in digital format, there is a large repository of information and literature that exists in paper archives. Conversion of this information into standard digital formats accessible by computer applications is often difficult and results in loss or degradation of data. A system developed at the St. Petersburg Office converts traditional paper sediment-core description sheets into digital spreadsheets that capture all of the symbols, text, and quantitative attributes of the sample (Flocks, 2004). Because scientific data are commonly collected in a repetitious, concise matter (i.e., forms, tables, graphs, etc.), these data can be recovered digitally using a translation process that relates the position of an attribute in defined two-dimensional space to the value that the attribute represents (Figure 3). The relations between position and data value is dependent on a key table that is customized to the type and format of data being converted. This system allows the incorporation of a vast amount of previously non-digital scientific data into the geodatabase.

Recent advances in technology, like large format scanners and image processing software, now allow shared access through LASED to decades-old legacy datasets. Digital scans of thousands of feet of analog seismic-reflection profiles and new image-processing software permit conversion of the resulting digital images into industry-standard format for further processing and enhancement. Additional software converts the processed data into an interactive seismic-reflection profile webpage that allows the viewer to find a geographic location and depth for a cursor position on the interactive profile using a web browser (Figure 4). These new tools provide improved visualization and use of these legacy datasets, and LASED serves as the “search engine” for locating these data spatially or by attribute.

Finally, the development of standardized logs, acquisition parameters, and naming conventions has enabled

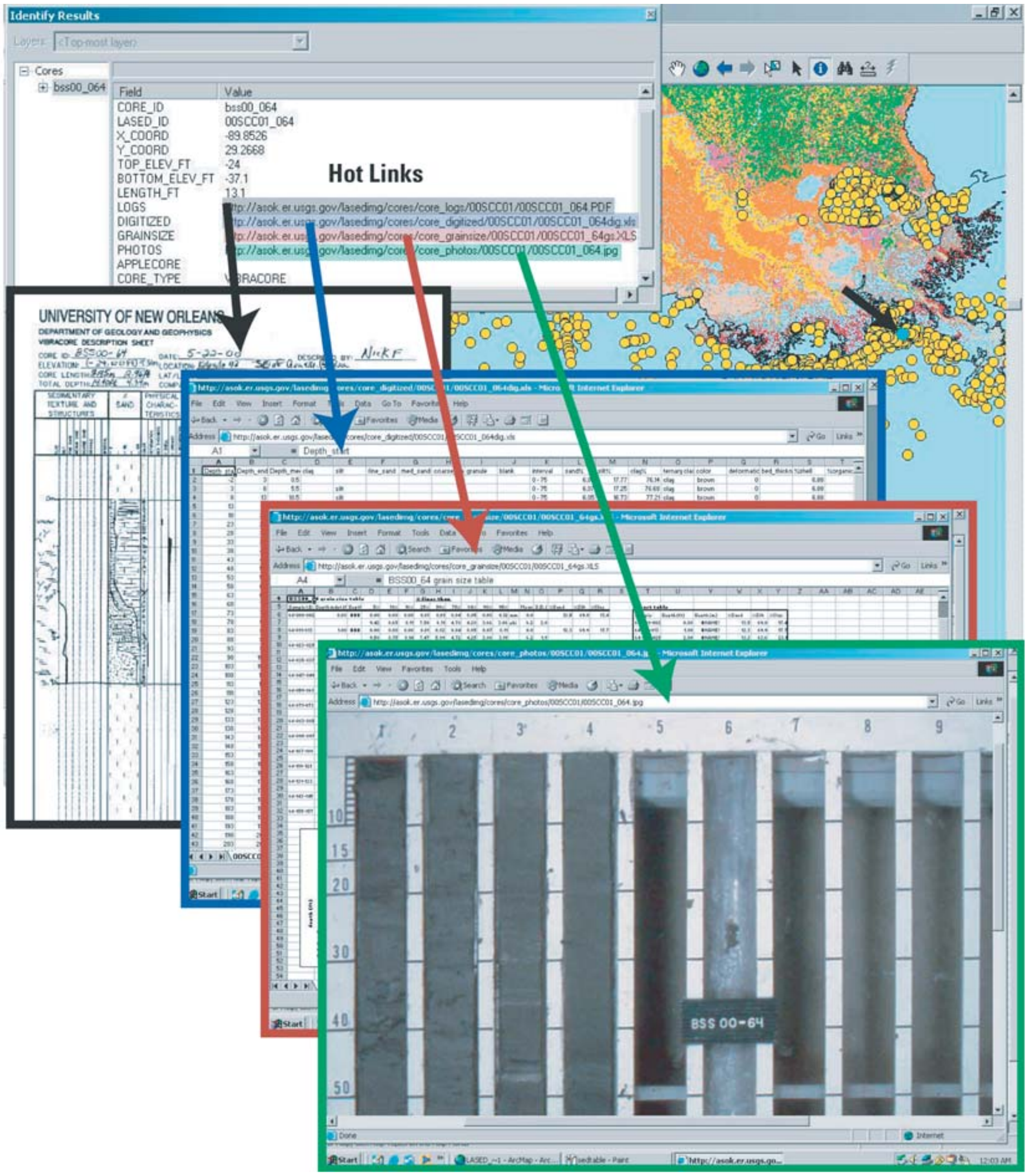


Figure 2. Example of sample information stored in the LASED geodatabase and hot linked to website locations (overlaid boxes) for associated sample data. Hot linked data are (from left to right) scanned core description log, digitized core log, grain size analysis results, and digital core photograph. Together, this information forms a permanent online digital archive of these data. Small arrow on map points to core location.

rapid processing of newly acquired digital data. Use of standardized methods and formats makes population of the geodatabase and associated published data archives routine. The rapid population of the geodatabase provides instant feedback to project planners. For example, a comprehensive display of all available data shows a roadmap of where data are lacking. Since all data stored or linked to the geodatabase are in digital format, distribution of the data can be accomplished electronically. The geodatabase is a powerful analysis tool. Using standard ArcGIS software, complex multivariate statistical analysis can be applied to produce new datasets and customized map products.

GEODATABASE SYSTEM ARCHITECTURE

Two Oracle 9i databases managed by Environmental Systems Research Institute's (ESRI) ArcSDE (Spatial Database Engine) 9.0 server software form the core of the

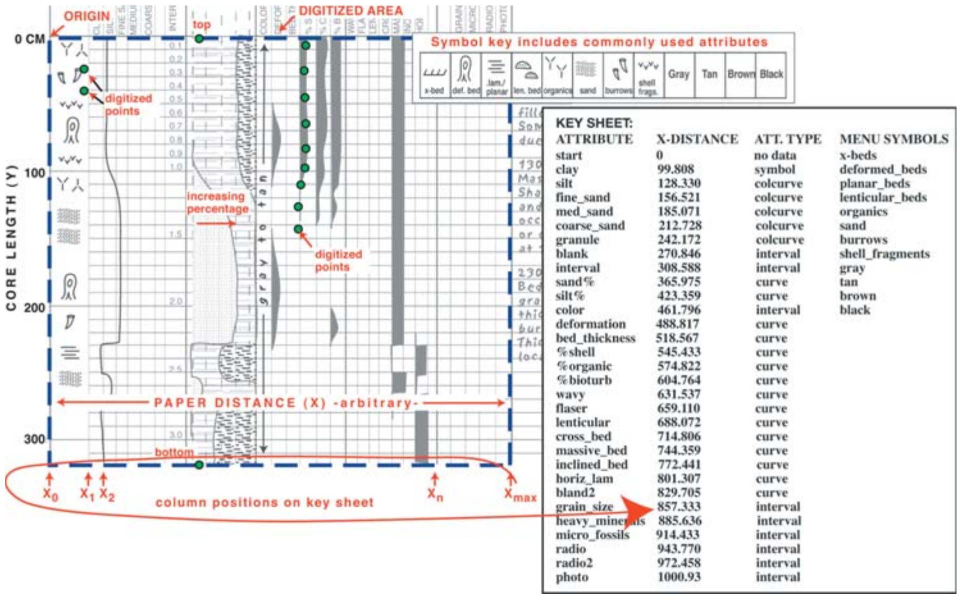


Figure 3. System to convert sediment core information from form-style description sheets into quantitative spreadsheet format: description type (e.g., percent sand) is categorized by the style of data (symbol, curve, etc.). This style determines which subroutine is used to convert the digitized data based on a key sheet annotated by column position. Abundance or presence of attribute is calculated by converting a digitized point to a magnitude value within the column. For more detail refer to Flocks (2004).

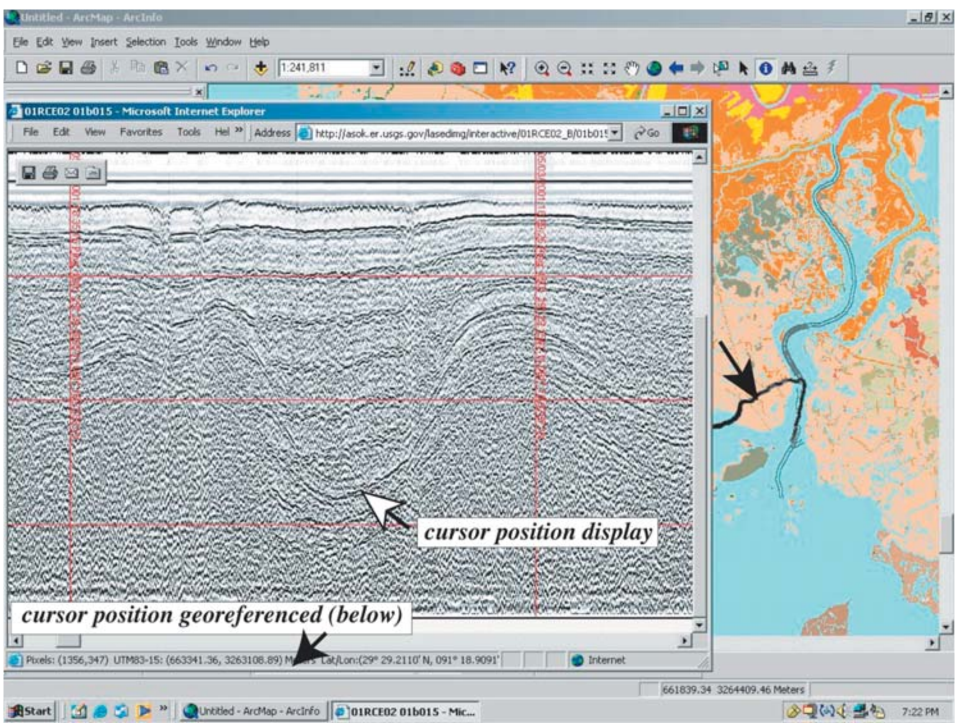


Figure 4. Display of an interactive seismic-reflection profile webpage (produced by Chesapeake Technology Inc., SonarWeb software) with ArcGIS. Seismic-profiles (double gray lines) are displayed atop the NWRC's landcover basemap. Selected seismic-profile line (bold black) is displayed in a web browser inset. Position of cursor on profile image is tracked with geographic coordinates (latitude/longitude and Universal Transverse Mercator (UTM)) and depth (meters and milliseconds) in status bar at base of browser window.

FISC-St. Petersburg geodatabase system. The databases reside on a Sun Enterprise E250 running Solaris 8 with a directly-attached Sun D2 Array. Detailed summaries of the computing platform and disk architecture are shown in Table 1. The other main components of the geodatabase system are an Apache web server and an ArcIMS (Internet Map Server) (Figure 5).

ArcSDE management of the Oracle databases is similar to a personal or "unshared" geodatabase in that there is straightforward access by means of standard ArcGIS desktop software. However, ArcSDE management offers several advantages including multi-user editing, versioning, raster storage, and an unlimited database size (Environmental Systems Research Institute, 2002).

The St. Petersburg Office has two geodatabase instances so that different data types can be stored optimally; they are the Spatially United Data Sets (SUDS) and the Spatially United Raster Features (SURF). SUDS stores exclusively vector data (survey tracklines, laboratory analyses, etc.) whereas SURF stores raster data (remote sensing imagery, elevation maps, etc.). Vector and raster data are divided into two separate databases for organization, speed, tuning, and to facilitate backups. Oracle database software was chosen because it is compatible with ArcSDE software and the FISC-St. Petersburg Solaris/Linux computing environment, is easily expanded, and is

Table 1. Oracle database server architecture.

Geodatabase Server Platform	
Hardware:	Sun Enterprise E250 with a directly attached Sun D2 Array
Operating System:	Solaris 8
Memory:	2 GB RAM
CPU Speed:	2 x 300 MHz UltraSparc-II
Disk Space:	378 GB raw disk capacity
Server Disk Management	
Management Software:	Solaris Disk Suite v4.2.1
Operating System:	Mirrored on 2 x 9 GB disks
Oracle & ArcSDE software:	Mirrored on 2 x 36 GB disks
Oracle Data:	RAID 5 system on 5 x 36 GB disks, with 1 hot spare 36 GB disk

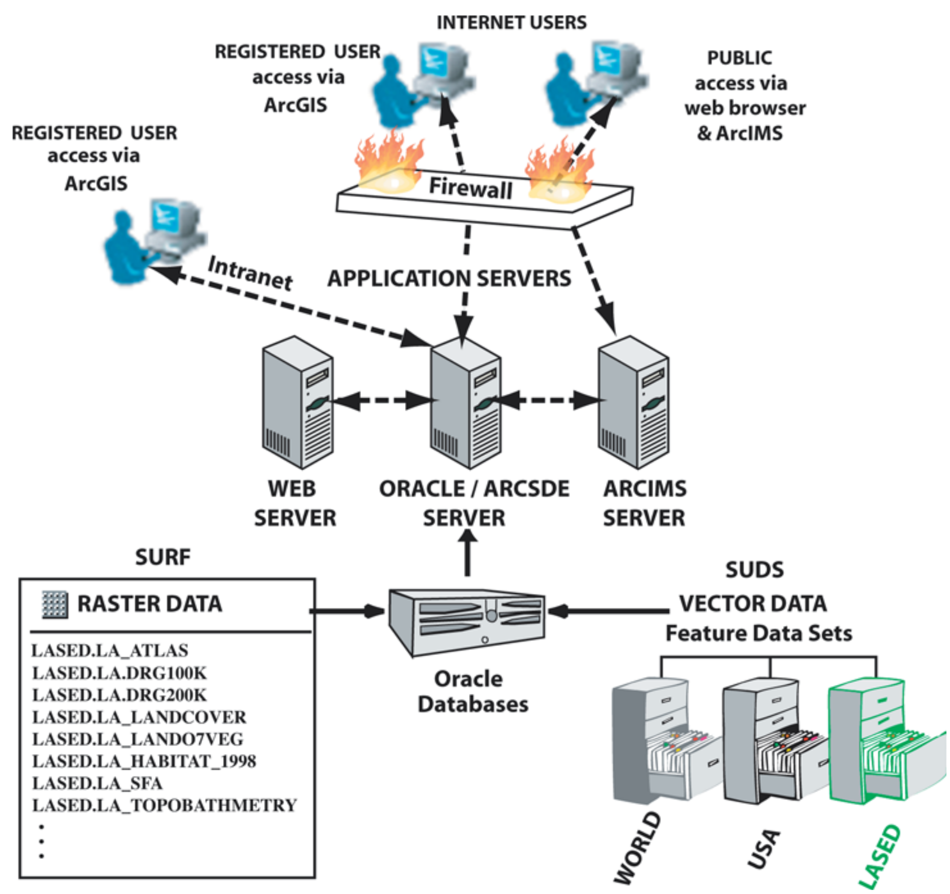


Figure 5. Schematic of the USGS FISC-St. Petersburg geodatabase system showing the relationships between the users, applications servers, databases, and the LASED raster and vector components.

reasonably priced through government contracts. ArcSDE software was selected because it is the standard USGS geodatabase solution and is widely deployed.

Within the St. Petersburg geodatabase system, LASED encompasses a schema made up of a feature dataset, tables, relationship classes, and raster datasets that together store, organize, link, and serve Louisiana coastal data to USGS scientists and collaborators (Figure 5). LASED data is incorporated within both Oracle databases (SUDS/vector and SURF/raster). Within SUDS, the LASED feature dataset currently consists of 29 feature classes and 4 relationship classes. Locating data into one feature dataset facilitates data organization and read-and-write permissions. There are also 7 LASED tables and 4 LASED relationship classes outside of the feature dataset (Figure 6). In addition, there are 70 Louisiana basemap feature classes within SUDS to which LASED users have access. Within SURF, naming conventions are used to organize approximately 40 LASED raster datasets, and read-and-write permissions are managed individually for each layer. All SUDS and SURF layers have Federal Geographic Data Committee (FGDC) compliant metadata.

The LASED schema is based on ESRI's Marine Data Model, or ArcMarine (Wright and others, in press), which attempts to spatially integrate the many data types that are unique to the marine realm (Figure 7), and considers how coastal and marine scientific data can be more effectively integrated into 3-D space and time series. Currently, LASED contains a wide variety of data including bathymetry, sediment-core data, seafloor-change images, seismic-reflection tracklines, and sidescan-sonar mosaics. Building on an established data model like ArcMarine has two major advantages: 1) for the GIS user, the model provides a template for the geodatabase structure that promotes networking and data sharing through established standards, formats, and relationships; and 2) for the developer, the model provides a basic framework for writing program code that can be used by a wider audience (ArcMarine Working Group, written comm. 2005, Figure 7). As more users build on this data model, additional tools to analyze or visualize these unique data types should evolve.

GEODATABASE ACCESS AND DATA INTEGRITY

The FISC-St. Petersburg geodatabases exist on an internal network protected by commercial firewall products. Direct access to the geodatabases through standard ArcGIS products, such as ArcCatalog and ArcMap, is allowed only to registered users. All accounts are password protected, and permission to load data into the geodatabases is limited to privileged users. LASED data and products can also be viewed over the Internet via an ArcIMS located at http://coastal.er.usgs.gov/lased/.

The St. Petersburg geodatabase backup plan consists of running both databases in ARCHIVELOG mode and using Oracle's Recovery Manager (RMAN) utility to do monthly full backups and daily incremental backups. The backups are copied to tapes which are then rotated offsite. RMAN has many advantages over user-managed backups that use operating system utilities: for example, the ability to perform incremental backups, a simplified procedure for backing up an open database, corrupt block detection, automated backups, backup catalogs (metadata), backup validation, automated management of backup files, and automated recovery (Oracle, 2005). The St. Petersburg Office also maintains on a second server duplicate development databases that can quickly be pressed into service to ensure minimal user interruption should the production databases require downtime for any reason. The development databases are also used for testing purposes.

The disk management system is shown in Table 1. The UNIX operating system, Oracle software, and ArcSDE software are mirrored on separate drives. If a disk failure occurs, the mirrored drive will take its place with no downtime. The data is stored in a RAID5 configuration with one hot swap disk. If one disk fails, a new disk can be brought online and the RAID5 system will repopulate the new disk automatically.

ONLINE DIGITAL DATA-ARCHIVE SYSTEM

LASED is more than just a geodatabase; it also forms the backbone of a permanent online digital data-archive system. Seamless interaction between published data archives (e.g., Calderon and others, 2003) and the geodatabase allows for rapid dispersion of data to collaborators. The system takes full advantage of ArcGIS information tools and a web server. The data files are linked either spatially or by attribute and provide to the remote user digital representations of almost any combination of data. Data types, such as sediment-core analyses, descriptions, and photographs, are dynamically linked to cruise information, equipment information, scanned field logbooks, processed and analog seismic-reflection profiles, and related publications and websites. As an example, the linked tables allow any user with standard ArcGIS tools and a web browser to identify on a map where a core was taken, see a digital image of the original core log, view a core photograph (Figure 2), query grain-size analyses (Figure 8), and find out when the core was taken, by whom, and where it is stored. LASED currently holds data for 20 cruises that include 750 seismic-reflection profiles covering about 6,500 kilometers (4,000 miles) and 1,150 cores.

SUMMARY

A wealth of geologic data exists for the Louisiana coastal zone. There is a critical need for a long-term data

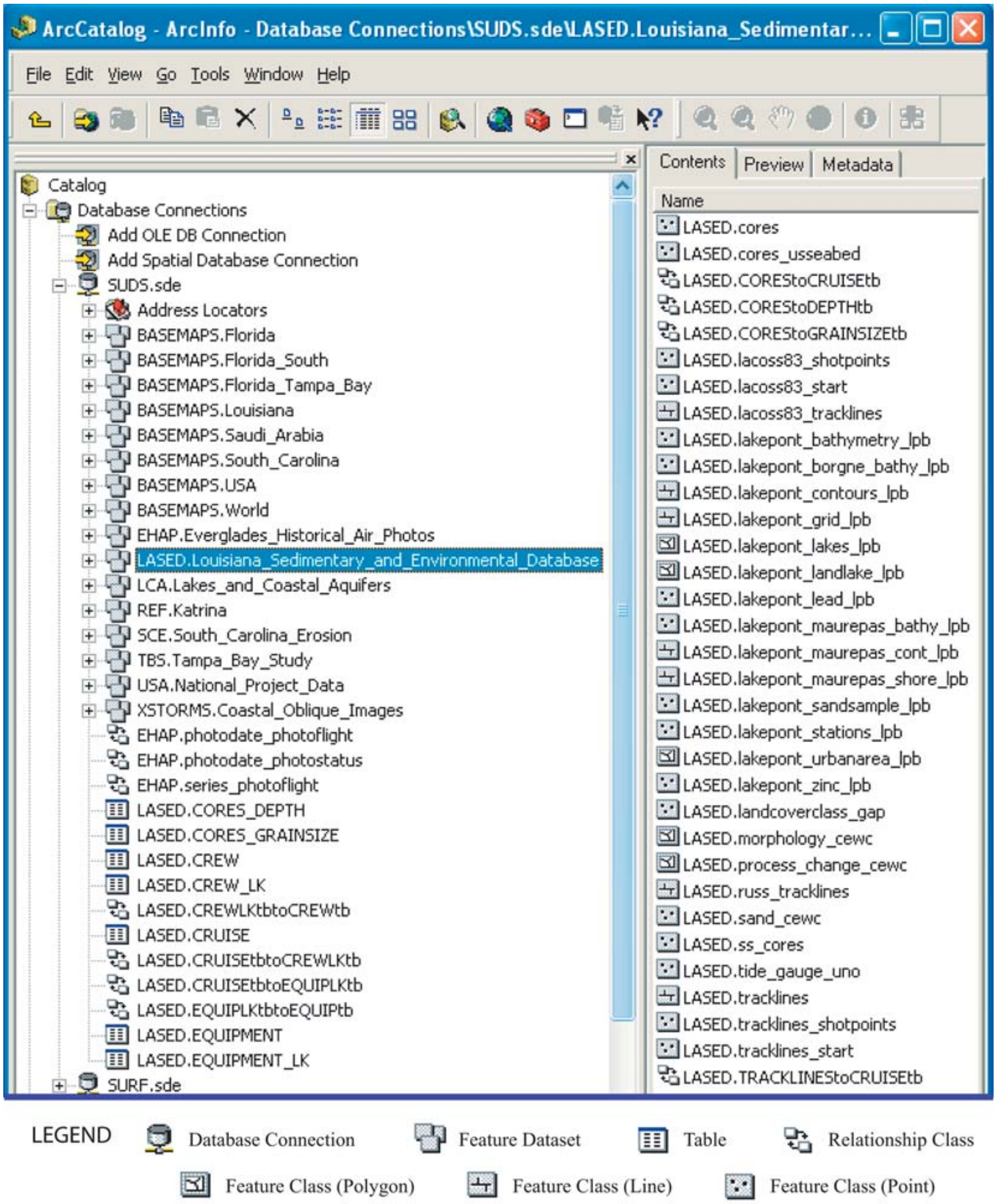


Figure 6. ArcCatalog display of the hierarchical structure of LASED data in the SUDS vector geodatabase. Icons are defined in legend.

management plan that provides data access for researchers and planners who are combating wetland loss and shoreline change. The USGS and Louisiana State cooperators have joined forces to address this need and have launched an aggressive effort to assemble all existing geologic and geomorphic data from the Louisiana coastal zone.

A substantial repository of information and literature exists in paper archives. Systems developed at the St. Petersburg Office convert traditional paper descriptions, logs, and profiles into digital format and allow the distribution of a vast amount of previously non-digital scientific data. New technology and software permit processing and computer-assisted visualization of data in ways that allow direct spatial interaction with the information. Finally, the use of standardized logs, acquisition parameters, and naming conventions permits rapid processing of newly acquired digital data and prompt distribution to researchers and project planners.

The St. Petersburg Office has organized the stockpile of data and all associated metadata into a comprehensive geodatabase system. LASED is a component of the St. Petersburg geodatabase system and is geographically

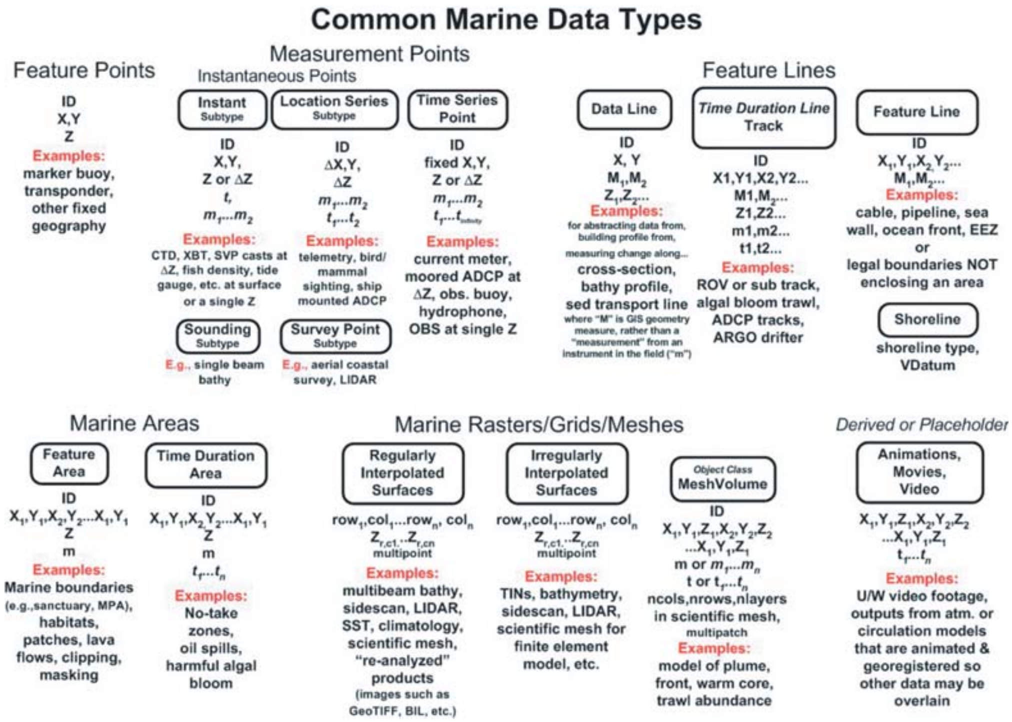


Figure 7. Diagram of common marine data types included in the Marine Data Model (from Wright and others, in press). Figure shows examples of common types of marine data and the relationships they share with geospatial information and features. For detailed explanation refer to http://dusk2.geo.orst.edu/djl/arcgis/index.html.

focused on the Louisiana coastal zone. Two Oracle databases form the core of the geodatabase system and are managed by ArcSDE software. Advantages for managing the data with ArcSDE include multi-user editing, versioning, raster storage, and an unlimited database size. The other main components of the geodatabase system are a web server and an Internet Map Server. The geodatabase is highly interactive in that it allows the display of a wide variety of data that can be cross-referenced by geographic position or attribute. Direct access to the geodatabases is allowed only to registered users via the Intranet and permission to read or load data into the geodatabases is limited to certain users. Access to LASED data is also available over the Internet via an ArcIMS.

LASED is based on ESRI’s Marine Data Model, or ArcMarine, which attempts to spatially integrate the many data types that are unique to the marine realm. Currently, LASED contains a wide variety of data including bathymetry, sediment-core data, seafloor change images, seismic-reflection tracklines, and sidescan-sonar mosaics. LASED also forms the backbone of a permanent online digital data-archive system that currently holds information for 20 research cruises.

ACKNOWLEDGMENTS

The Coastal and Marine Geology Program of the USGS, the U.S. Army Corps of Engineers, the University of New Orleans, and the LDNR provided funding and/or support for this project. We thank Jack Kindinger for his support and guidance and Heather Mounts for implementing SUDS and SURF and the prototype of LASED. This document was improved by the reviews of Robert Wertz and Barbara Poore of the USGS in St. Petersburg, Florida.

REFERENCES

Calderon, K., Dadisman, S.V., Kindinger, J.L., Perry, G.M., Flocks, J.G., Wiese, D.S., Kulp, M., Penland, S., Britsch, L.D., and Brooks, G.R., 2003, Archive of boomer seismic-reflection data collected during USGS cruises 00SCC02 and 00SCC04, Barataria Basin, Louisiana, May 12-31 and June 17-July 2, 2000: U.S. Geological Survey Open-File Report 03-402, DVD (1 disc), available at http://pubs.usgs.gov/of/2003/of03-402/.

Environmental System Research Institute (ESRI), 2002, Working with the Geodatabase: Powerful multiuser editing and sophisticated data integrity: ESRI white paper, 9 p. available at http://www.esri.com/library/whitepapers/pdfs/arcgis_geodb_multiuser.pdf.

Flocks, J., 2004, Converting analog interpretive data to digital formats for use in database and GIS applications: U.S. Geological Survey, Open-File Report, 2004-1070, 21 p.

Frazier, D., 1967, Recent deltaic deposits of the Mississippi River: Their development and chronology: Gulf Coast

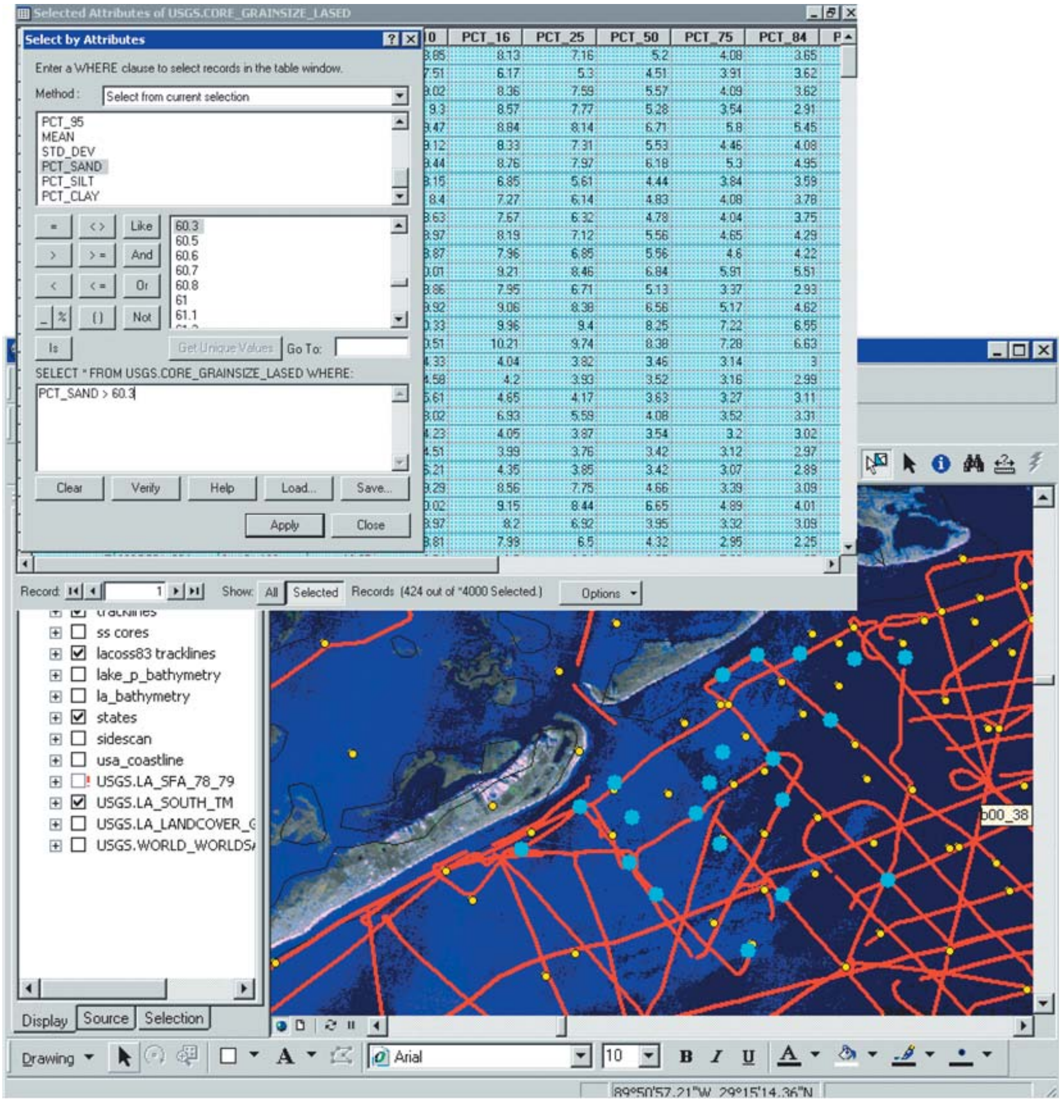


Figure 8. Example of a data query of sediment-core grain-size analyses with greater than 60 percent sand. Background map shows query results highlighted as large dots. Raster basemap features LANDSAT Thematic Mapper (TM) satellite imagery.

Association of Geological Societies, Transactions, v. 17, p. 287-315.

Louisiana Coastal Wetlands Conservation and Restoration Task Force and the Wetlands Conservation and Restoration Authority, 1998, Coast 2050: Toward a Sustainable Coastal Louisiana: Baton Rouge, La., Louisiana Department of Natural Resources, 161 p., available at http://www.coast2050.gov/2050reports.htm/.

Morton, R., Miller, T., and More, L., 2004, National assessment of shoreline change: Part 1 Historical shoreline changes and associated coastal land loss along the U.S. Gulf of Mexico: U.S. Geological Survey Open-File Report 2004-1043, 42 p.

Oracle, 2005, Oracle 9i backup and recovery concepts: Oracle software manual, available at http://www.stanford.edu/dept/itss/docs/oracle/9i/server.920/a96519/toc.htm/.

Wright, D.J., Halpin, P.N., Blongewicz, M.J., and Breman, J.B., in press, ArcMarine: GIS for a Blue Planet: Redlands, California, ESRI Press, http://dusk2.geo.orst.edu/djl/arcgis/index.html.

SOFTWARE CITED

Apache—Apache HTTP Server project, http://httpd.apache.org/.

ArcGIS, ArcIMS, ArcMap, ArcCatalog, ArcSDE—Environmental Systems Research Institute (ESRI), Inc., 380 New York St., Redlands, CA, 92373-8100 USA, (909) 793-2853, http://www.esri.com/.

ArcMarine—ArcMarine: The ArcGIS Marine Data Model, http://dusk2.geo.orst.edu/djl/arcgis/index.html.

Oracle—Oracle Corp., 500 Oracle Parkway, Redwood City, CA 94065 USA, (800) ORACLE-1, http://www.oracle.com/.

SonarWeb—Chesapeake Technology Inc. Chesapeake Technology, Inc. (CTI), 1146 Kathy Way, Mountain View, CA 94040, USA, (650) 967-2045, http://www.chesapeaketech.com/.

A Conceptual Approach to the Development of Digital Geological Field Data Collection

By Guy Buller

Northern Division
Geological Survey of Canada
255-601 Booth Street
Ottawa, Ontario K1A 0E8
Telephone: (613) 947-1889
Fax: (613) 947-9518
e-mail: gbuller@nrcan.gc.ca

ABSTRACT

Members of the Geological Survey of Canada (GSC) have for several years been successfully developing digital systems that aid geologists in the capture of field data. In the past, development has been completed, or driven, by an individual researcher on a per project basis and, therefore, systems have been specific to that geologist's work. This sort of application development has often meant that the work takes place in virtual isolation and the resulting application can be very limited in scope or usability for other researchers.

Due to the demands of business re-alignment in the GSC over the past few years, there has been an attempt to work toward a single system that could be used by a variety of researchers for the collection of field information. To accomplish this broad-spectrum development, the work has been conducted in coordination with many mapping projects; this has proven advantageous in coordinating development between many projects across the organisation. By following this strategy, the GSC is attempting to bring consistency to the data-gathering efforts, and thereby also minimize the isolation of projects, which was a problem in the past.

THE VIEW

Fieldwork and data gathering processes that are carried out using pencil and paper are in no way flawed, but the raw data from the fieldwork can appear cryptic because of an individual's unique techniques or terminologies, or the specific goals of the project. Furthermore, because there are repetitive aspects to fieldwork, the mapper commonly develops an individualized note-taking style that includes various abbreviations and other "short hand" techniques that provide short cuts to limit the amount of writing that is necessary. Often, the data collection location is not as idyllic as shown in Figure 1, and the amount of short hand can be dependent on the amount of time available at each site or proportional to the number of biting insects (Figure 2) or the temperature. These short cuts are easily understood because they are in context with what was recorded in previous day's work (e.g., "SOS" may mean "Same Old Stuff"), but this ambiguous information, over time, loses its meaning.

Short cuts are most used by people who are trying to solve problems or make progress under tight time constraints (Shalloway and Trott, 2004). However, they are impossible to process electronically, because the context cannot necessarily be captured and the number of short cuts and their meanings are unlimited. The use of the term "24-7" has come to mean "all the time"; although the context is not present, it is 'understood' by nearly everyone. On the other hand, project-specific short cuts are seldom as widely understood. Attempts to interpret the meaning of such short cuts may result in information or field observations that may not have been the intent of the original researcher. This may not be a problem if the individual is available to resolve any ambiguity, but with the passage of time, the researcher will become unavailable, to put it gently. As a result, attempts to convert old information into a database can introduce significant errors in data and scientific interpretations.

In the past, the principal developers of field data collection systems were the individuals who conduct the scientific research, and they addressed the data collection issues of their own project. By using a variety of software applications, data gathering systems have been developed, with inherent, project-specific short cuts. This sort of development has been very effective, because the person who controls how the information is to be collected or interpreted also can make any changes to the application that may be required. In some cases, however, these applications can imbue the data with a regional or research specific flavour that may be rather unique, even though geologic principles and observations are the same for any project.

Figure 1. Quaternary data collection in a pipeline trench (photograph by A. Plouffe, 2005).

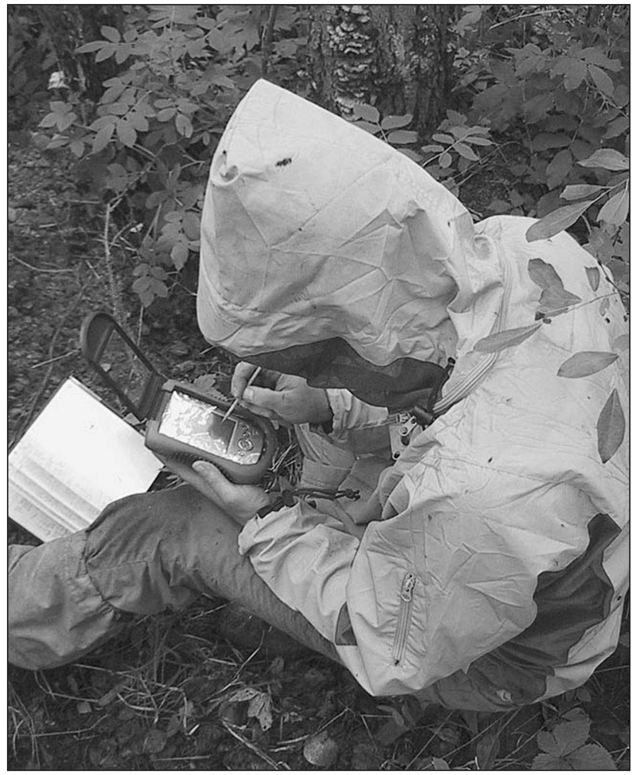


Figure 2. Data collection in North Western Alberta (photograph by A.Plouffe, 2005).

These unique approaches to data capture can also be a product of application development, as a researcher is faced with a short preparation time prior to the field season. If a suitable beta application is developed and successfully meets the immediate needs of the project in the first season, it is probable that with each successive year the application will gain more functionality. With each subsequent year of use, the application becomes more entrenched into a specific data collection format and subsequently becomes less accessible by researchers outside of that specific project.

The building of applications to meet individual project needs has worked well in the past but more multidisciplinary, cooperatively-driven projects with diverse expertise have come into vogue, and systems that are developed as described above do not easily transfer to these larger, more complex projects. In order to facilitate another group's use of an application, a redesign of some sort must take place and the existing application is often patched to address the needs of the project. Changes to the application commonly are determined by decomposing it into its functional parts and if problems are found, those specific parts are modified. This functional decomposition (Shalloway and Trott, 2004) is a natural way for people to understand very large systems, but sometimes these modifications cause other parts of the system to be adversely affected. Further changes can mean that the modified parts become even more bound to certain field-specific functions and there is no "graceful evolution" (Shalloway and Trott, 2004) toward a deployable solution or to any new requirements that may occur.

Due to this project-specific development, there now exist many different applications that do not communicate well with one another. Furthermore, the maintenance of such systems over the long term becomes onerous for the researcher; time is spent tending the application rather than concentrating on their science. In some situations the entrenchment of specific systems is so strong that there is reluctance to change to newer systems, which in turn creates a certain unwillingness toward sharing and storing data within a corporate system. Yet, no matter the process of collection, all of the data that researchers accumulate is important to both the organisation as well as the scientist.

MOVING FORWARD

Many individuals recognize that much of the geological information gathered in different projects is virtually

the same in terms of content, although the style of reporting the observations, including format and terminology used, may differ greatly. This recognition has been an impetus to develop field data-capture and data-storage standards that are based on broadly accepted international standards. Some standards development had begun at the GSC (Buller 2004) but was very closely tied to the activities of an individual division at the GSC rather than being generic to the discipline of geology. To achieve a better model and to adhere to international standards it was decided to follow the work coordinated by the Open GIS Consortium (OGC). The OGC's release of a paper addressing Observations and Measurements (Cox, 2003) has helped to advance the conceptual modelling for electronic data collection, and the OGC model fits well with the way the GSC scientists collect information. Work is being done to formulate a physical model that can be applied to actual field activities and will form the foundation for the new modular approach of the latest version of the digital field collection application.

As projects move to an electronic system of data capture, they must rely more heavily on IT professionals, who serve as the bridge between the final corporate database and the users of field systems. The ability to understand the needs of the researcher is paramount as they are, for our field system, the end users; as such, they provide essential guidance on how the system should be designed. At the same time, objectivity during development must be maintained because there is a strong tendency to focus on the needs of a single (perhaps dominant) client, thereby risking the possibility of making development too project-specific. With lack of objectivity there is no change from the existing development process but rather simply a change in who does the development work.

The advantage of a non-project-specific development group is that there is no vested interest in any existing system; instead, the group focuses on the needs of all researchers. By being the 'interpreters' of the existing diverse collection systems, the development group is able to mingle concepts together and develop a unique, customized view of the data collection system that is based on an all-encompassing generic framework. This approach is similar to having a personalized desktop on an office computer while running on a common network. For any new system development however, both the developers and end-users must be able to approach the process in a cooperative manner and accept that the application should extend beyond the specifications of any one project. It must be recognized that a complete analysis of field data gathering techniques needs to be undertaken in order to understand how to develop systems that meet the needs of the whole organisation.

REQUIREMENTS

"One thing you will never hear (from developers) is, 'not only were our requirements complete, clear and understandable, but they laid out all of the functionality we were going to need for the next five years!"

As Shalloway and Trott (2004, p. 6) pointed out in this quote, initial requirements are not written in stone. When clients are presented with an application having some broad level of capabilities, they quickly can envision many other possible uses for the device, and so demands for future editions are soon developed. This means that requirement analysis is an on-going activity, and everyone can expect that changes are inevitable. To help consider these design changes, systematic business analysis using a commonly accepted approach, such as the Zachman Framework (Hay, 2003), allows programmers and users a long-term view of the development life cycle that clearly demonstrates the steps needed to meet existing requirements of a project. Business analysis is an iterative approach, and such an approach can yield better design criteria and flexibility to adapt to changing requirements. By using this approach it is expected that activities that are overly project-specific will be winnowed out and only common categories will remain to allow for generic object modelling. This modelling will be important in the development of precise object classes to facilitate the transfer of data to corporate databases.

Though the analysis approach is complete in its understanding of systems, it can sometimes run counter to project objectives that have specific mandates to produce something tangible in a limited time frame. There is always an implicit desire for a development team to have a final product that will be useful for many years, but in order to meet the short-term goals of a project these long-range plans often are sacrificed. To further exacerbate development barriers, resources often are extremely limited and yet the expectations of end users are very high.

The barriers to meeting long-term corporate goals can be overcome, but it must be understood by managers that system analysis in many organizations has not reached maturity and the learning curve for understanding and implementation is steep. The analysis activity simply produces the blue prints to the application and only models how a solution will be developed based on the requirements discovered. For people who require "real answers" and are not familiar with requirements analysis, this stage of development can sometimes be thought of as non-work, as it only gives a path toward the solution, rather than the solution itself. There is a distinct need for application architects and software developers to be able to muster management support and understanding for this critical process of application planning. With proper analysis, the final product will be better suited to the needs of the user and will be developed through fewer iterations. Over the long term, a well-designed application will be easier to maintain, will be expandable and will cost less in development for any future modifications.

A SOLUTION

The first iterations of the collection system organized all the field observations into one or two Shapefiles. The linked dbf files held the collected information that was supported by a single multi-line (1000+ lines of code) script that controlled the user interface. The single, large script quickly became unmanageable, as developers, in two different areas of the country, were required to make rapid user-defined functional changes to the application. In a very short time, multiple iterations of the same application were available, with no concrete way of addressing the variety of "wish lists" that were being submitted by users.

To solve the problems, steps were taken to reformulate the script and make the system more modular by relating individual field activities to individual Shapefiles. This objectification simply models common activities of a researcher that take place at the various stations that are visited during a day of fieldwork. Activities for most geologists are very similar (Figure 3) where certain activities are followed by other dependent activities. For example, all activities must be related to a station, and a sample must be related to an earth material.

By functionally decomposing the work, the information-gathering process becomes compartmentalized. In terms of a final data product, rather than a traditional spreadsheet comprised of sixty or more columns, we consider the information to be the attributes of geo-referenced points placed directly on an electronic map. By dividing up the different activity sets (as shown in Figure 3) into distinct layers within a GIS, independent information can be collected in a loose relational format that leverages the GIS capabilities of the platform application. The extra number of layers does increase the number of files to be handled by the system, but it is a design trade off to allow for more functionality for the user. This latest development also tries to implement the recent OGC specification for field data capture, but does so without the burden of specifying any particular relational database. This 'distributed' spreadsheet flat file data holding can be easily transferred into a relational database system where the power of the database can be brought to bear on the collected information. This transfer is made feasible by the fact that the flat files have built-in relationships to the associated activity that has been previously captured (e.g., samples must have an earth material and earth material must have a station).

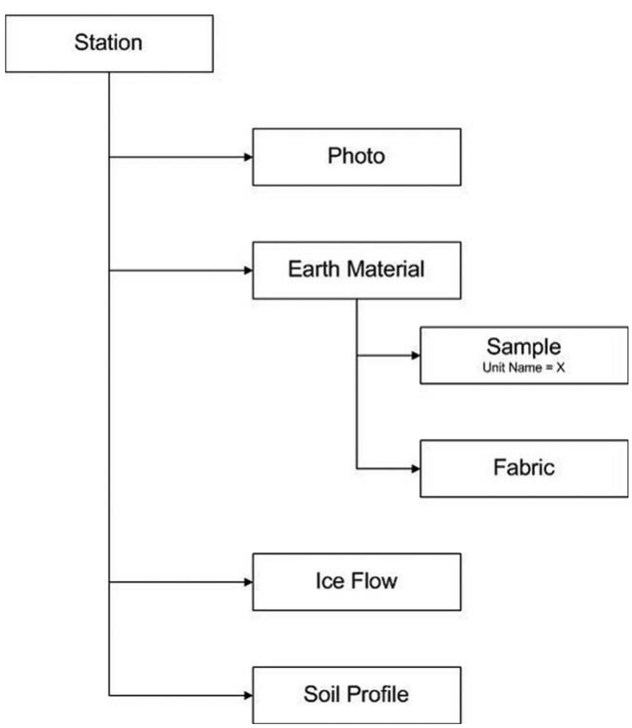


Figure 3. Field activity model.

By applying requirements analysis during a planning stage and by relying on the experienced gained from previous year's development, changes to the application were kept in line with the goal of data transfer to a corporate system. With solid communication between developers and good team procedures along with individual component development that was tightly focused, changing requirements to suit our user's needs were easy to administer. In this way, the various parts of data collection are treated as discrete objects having specific attributes and properties and a single platform is able to have a multitude of functionality that can easily evolve over time. Furthermore, compartmentalized coding has helped to isolate any glitches within specific components, thus making them easier to discover and correct. Also, if one of the data collection modules does not operate properly, then only that component is unavailable, rather than breaking the entire application.

The development team has found that the length of development has become shorter since the implementation of the modular approach. The focus of development is on the module to be added, rather than on determining how it will fit with or affect the rest of the application. What used to take a couple of weeks of intense coding can now be shortened to a few days, depending on the complexity of the requirements for the new module. Most importantly, this means that a successful application module is not patched together, but instead is built upon the existing standardized business format and maintains the common end-user interface. Since the field system (which is called Ganfeld; see Buller, 2004) is a visual interface that leverages the GIS functionality of ArcPad, there was the need to develop a system from the ground up using a different design approach than had been used by designers of data capture systems in the past. By not focusing on past design, the development team was able to let go of the old applications and allow the new design processes to

advance more freely. This has resulted in a more flexible system that is easily adaptable to a variety of different foci of research and may even possibly be extended beyond geo-science projects, because the field activity model can be applied to any spatially related fieldwork.

IN CONCLUSION

It has become clear that the development of data collection applications cannot proceed in a non-systematic fashion. The ability to step back from an existing design and examine all possibilities allowed the development of a set of interrelated components. Also, the ability for developers to write and modify these components of the system without interfering with the whole application allowed for parts of the application to be delivered in time for the 2005 field season.

Business planning activities such as requirements analysis have not yet become mainstream; however, by following best practices in design we have been able to complete many of the goals that we set for ourselves. The need for an easily maintained system that can contain much of the scientific data collected is intrinsic to the many goals of an organisation. Also, the development of a field data repository that follows internationally accepted standards is required to ensure the preservation and access to all the information collected in the field.

It is expected that over time the various application modules that have been developed for field data capture will be altered to the point where certain modules will become a standard set of modules while more unique items turn into interchangeable components. As there is an existing modular design, these changes will be easily accomplished while still maintaining a level of base functionality.

More work is planned that will smooth out the rougher edges of the application, and procedures will be put in place to more easily consider changes in requirements. Further development of the field objects as per the OGC specification will continue, as will the realignment of data storage systems to contain this information.

REFERENCES

Buller, Guy H.D.P., 2004, Ganfeld: Geological Field Data Capture, in D.R. Soller, ed., Digital Mapping Techniques '04—Workshop Proceedings: U.S. Geological Survey Open-file Report 2004-1451, p. 49-53, accessed at http://pubs.usgs.gov/of/2004/1451.

Cox, Simon ed., 2003, Observations and Measurements Version 0.9.2; Open GIS Interoperability Program Report: Open GIS Consortium Inc., 4 April 2003, reference number OGC 03-022r3.

Hay, D.C., 2003, Requirements Analysis From Business Views to Architecture: Prentice Hall PTR, Upper Saddle River N.J.

Shalloway, Alan, and Trott, J.R., 2004, Design Patterns Explained: A New Perspective on Object Oriented Design, 2nd Edition: Software Patterns Series, Addison-Wesley Professional.

MAP IT: a GIS/GPS Software Solution for Digital Mapping

By Mauro De Donatis, Lorenzo Bruciatelli, and Sara Susini

LINEE – Laboratory of Information-technology for Earth and Environmental Sciences
Universita' degli Studi di Urbino "Carlo Bo"
Campus Scientifico
Loc. Crocicchia
61029 Urbino – Italy
Telephone and FAX: int.+ 39. 0722.304295
e-mail: dedonatis@uniurb.it

ABSTRACT

LINEE (Laboratory of Information Technology for Earth and Environmental Sciences) has developed a new system for geological mapping with a tablet PC computer in the field using GIS software. In this software several new tools have been integrated using Visual Basic in the Microsoft SDK environment for programming for Windows XP. It includes support for GPS. Data and observations (in spreadsheets, notes, sketches, photos, etc.) can be georeferenced. In addition, day/hour and position (latitude, longitude and altitude) are automatically associated with the photos managed by this system. All documents and files can be recorded in a database with Microsoft Access and ESRI format database. This software is designed for use by field geologists without GIS or other advanced computer experience.

INTRODUCTION

The idea to create a field data-capture instrument that takes advantage of the possibilities offered by Tablet PCs was born from several years of experience and hundreds of kilometres of geological and geomorphologic surveying using traditional methods for creating geological and "derivative" or applied geological maps at different scales, for regional and national mapping projects. We have found that the invaluable notebook of the geologic mapper may be substituted in most cases by the Table PC.

Because we must use our computers in the field, conventional laptop and handheld computers are found to be not sufficiently durable; however, more "rugged" computers have now entered the commercial market, for use in the field. Battery life, screen display and the reduction in weight are the factors to be considered, though almost every month new models are available which improve performance, reduce operative problems, and are less expensive.

Our original idea was to create a tool to be efficiently and simply used in the field by geologists who may have a limited knowledge of GIS and who want to minimize the learning time for new technologies. Moreover, the information gathered and immediately entered in the field drastically reduces the loss of information. Often data, sketches, and interpretations, which lead to a final synthesis, remain in the geologist's notebook or in their mind. In many cases, the geologists, at the end of the field project, present their work (which often consists of a piece of paper, a legend and, perhaps, illustrative notes and sections) to the director or coordinator of the study; this person certainly has not been at each location in the mapped area visited by the field geologist, and therefore does not have the ability to "weigh" the data or interpretations. For example: does a field observation have a regional value or is it merely linked to an incidental location? If one knows the geographic path and thought processes of the geologist, it is easier to answer this question.

THE SOFTWARE

Map IT is developed in *MS Visual Basic*. Its principal functions are: editing cartography, georeferencing, coordinate conversion, database association, topological operations (topological clean-up, union, intersect, identity, clipping, erase, buffer), raster operations, construction of advanced and/or personal symbology, and import/export in the most common formats (shp, dwg, dxf, mif, mid, and the cadastral ntf). Map IT therefore offers all the essential tools of a Geographic Information System, yet is integrated in a fast and practical interface suitably familiar to people who utilize *Digital-Ink* technology to add or modify graphic elements in a digital map.

In Map IT, the user can utilize the digital pen to select commands and write notes, sketch designs, and highlight areas directly on the map, thereby connecting the information to a geographic position. Map IT makes

it possible to easily create and manage point, line, and polygonal topological elements using functionality similar to CAD for drawing, such as: snap functions, lock angles, automatically close an arc to the initial node; the possibility to insert or modify coordinates via the keyboard, and distance and area measuring tools. Also, proprietary functions have been developed (positioning via GPS, custom form creation, multimedia data collector, the ability to add annotations to the cartography or to images, etc.) which make Map IT a complete and valuable tool for those who wish to directly apply the technology of Geographic Information Systems on-site.

GPS IN MAP IT

Map IT connects to any GPS (Global Positioning System) device (Figure 1). While conducting a field mapping project, the data collector can display the current position on the Tablet PC's map image, in real time. This functionality is fundamental when a geologist is working in an area without any direct points of reference which would help pinpoint the exact position on the map. When direct points of reference are available, immediate comparison and validation of data may be made using the data furnished by the GPS.

The GPS settings window of Map IT conveys information relative to the displacement of the data collector (speed and direction) and the status of the satellites. Map IT permits the user to set the interval in seconds upon which to base the positioning information from the receiver. In this way, Map IT calculates the average fix received from the GPS in a time interval specified by the user, considerably improving the data precision.

With *Map IT*, a *coordinate system* different from the WGS84 used as the standard by all GPS data collectors may be defined. The data regarding the position obtained from satellites is immediately compatible with that being used by the mapping project.

Once a GPS connection is established, Map IT offers the user an array of functions. For example, you can center the current display on the receiver coordinates, or, while walking, maintain the data collector's position at the center of the view while moving the map image to follow the position of the GPS. It is possible to set the Map IT plotting function to capture a path (capture the coordinates of the GPS) traversed by the data collector. In this case, Map IT offers the possibility to choose whether to capture the path data as points or arcs and whether to set a spatial tolerance (in meters) and a time period (in seconds), for capturing the spatial data. This function allows the creation, for example, of a path map of a park, simply by walking along its pathways.

Selecting the *Draw Point on the Map* command, a point is drawn on the map in the exact position of the data collector. In this way, a point feature is instantly created with which information may be associated. This function is extremely useful when surveying the position on precise points, such as wells, manholes, etc. Moreover, it is possible to automatically label such points with a preselected cartographic symbol.

Another useful tool offered by Map IT is to

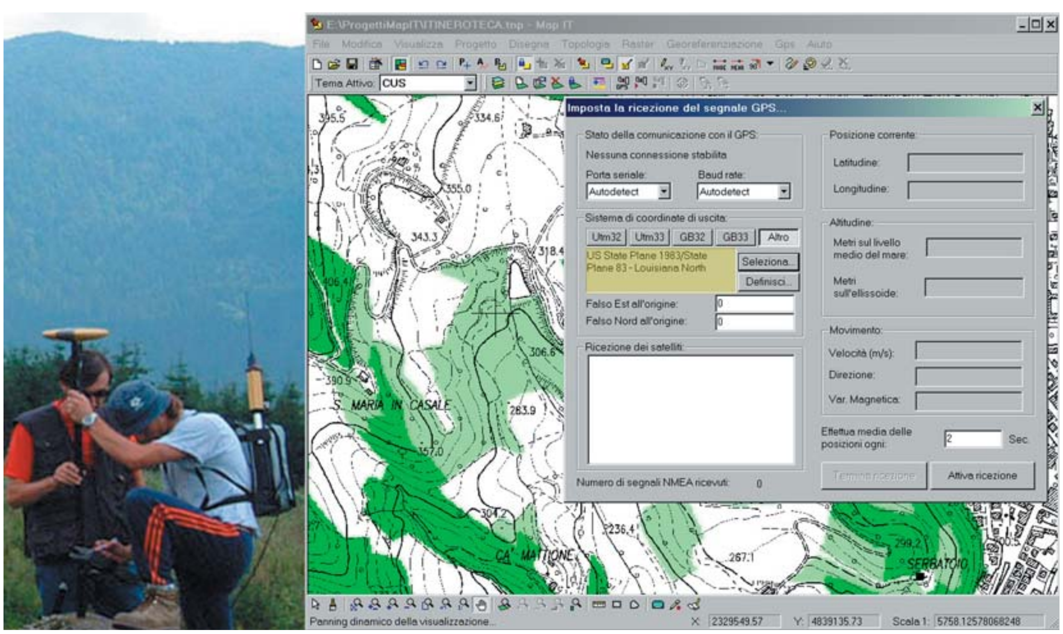


Figure 1. Rugged Tablet PC connected to a GPS antenna via USB or Bluetooth port and windows of Map IT for GPS connection (although the Italian-language version of the software is shown here, note the U.S. State Plane coordinate system being used).

georeference documents that were created during a geologic mapping project. Imagine having walked over a nature path taking photos with a digital camera, writing a few annotations in a text file, or recording voice notes. Map IT automatically positions the documents in the exact position on the map in which they were created, using the date/time information of the creation of the file along with the data recorded by the GPS receiver. Thus, all activities and observations conducted during a mapping project (photographs, written notes, drawings, designs, etc.) may be automatically linked to a geographic position.

EASY NOTE—MULTIMEDIA COLLECTOR OF GEOREFERENCED INFORMATION

Map IT can associate various geographical objects with textual information or the link to any document readable in Windows (.doc, .xls, .dbf, images, html files, etc.) via *Easy Note*, a tool for taking impromptu notes (Figure 2). Easy Note is activated with a rapid click on the features already present on the map or is automatically displayed when a new graphic object is added. When Easy Note appears on the screen, it is possible to write the note freehand using the digital pen or by typing text on the keyboard. Simply dragging a document inside Easy Note's yellow note window creates a link to the file from any application present on the PC. It is therefore possible to insert not only textual information, but also photos, sketches, movies, spreadsheets, HTML pages, internet links, and recordings. By hovering the cursor over a graphic object which already has associated information, the user may display the previously stored text or open the documents linked to the object.

This "slip of paper" created by Easy Note becomes a multimedia collector of all the information that the geologist intends to associate with a graphic object. In this way, when ending a mapping project, the geologist commands a complete registration of the activities recorded and is able to profitably utilize the gathered information.

Right-clicking on a file's icon, a pop-up menu appears which offers the user the possibility to open the file, view the file's properties, or remove the file from Easy Note. If the file is an image file, the pop-up menu also contains the choice to open the file in the Map IT image display (Figure 3). This simple tool was designed to instantaneously display all image files and manage them quickly and practically. The image display toolbar offers the user several functions. The zoom command enlarges (with a left-click) or reduces (with a right-click) the current display. The pan command allows the user to move the image within the window. The zoom and pan commands are dynamic; by clicking on a point in the

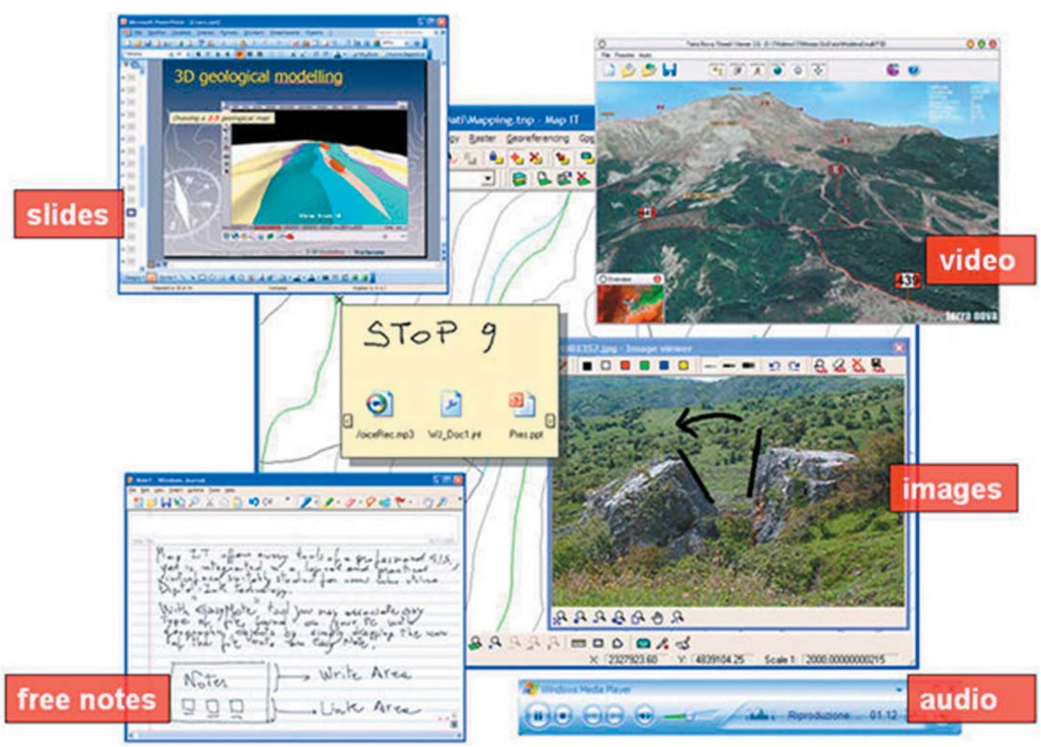


Figure 2. Screen shot of Map IT with Easy Note and Image Display windows. In the upper part of the Easy Note window you can write with a digital pen or the computer's keyboard; in the lower part of the window you can insert links to any kind of file.

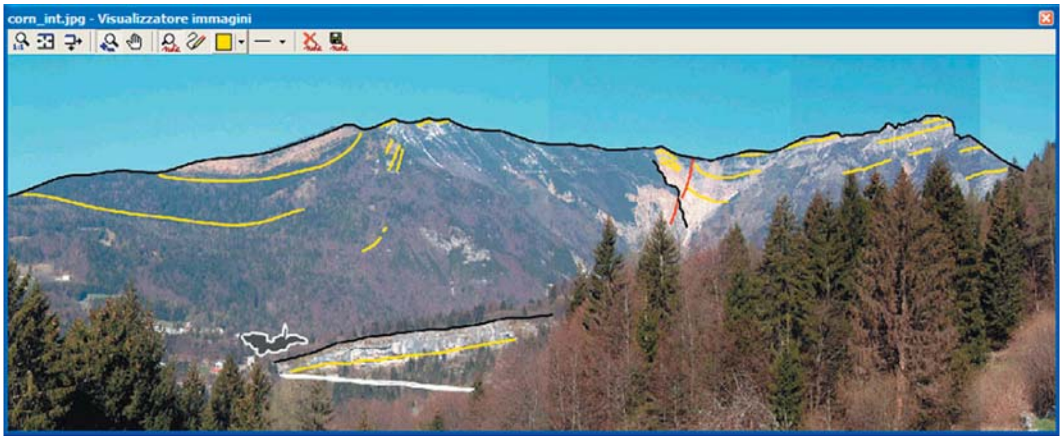


Figure 3. Image Display: you can sketch and write notes using the digital pen.

image, the desired effect is obtained and contemporaneously the clicked area is moved to the center of the screen. The image also may be adjusted to the window size and vice versa, or may be reset to the original size. Moreover, it is possible to write comments or make sketches on the image, and to choose the colour and thickness of the sketch pen. The image is not modified since the notes are inserted as a glossy transparency that may be displayed or hidden from view. In each case, it is possible to save a copy of the image with the notes and the renderings made by the data collector.

FORM EDITOR—CREATION OF CUSTOM FORMS

Via the Map IT *Form Editor* tool, custom forms may be prepared that will guide and facilitate data entry during field work (Figure 4). Each time a new feature (for example, an outcrop observation) is inserted in a project theme, the appropriate form created for that feature appears, so that the data collector immediately sees the information that must be gathered.

When creating a form, it is possible to insert labels, text fields, frames, check boxes, option buttons, combo boxes, list boxes, and action buttons. The forms and buttons can be configured; for example, *text* (the text displayed in the control), *font* (the character format of the text which appears in the control), *backcolour* (the background colour of the control), and *left*, *top*, *width*, *height* (the size and position of the control) can be set by the user. Attribute properties allow the user to stipulate which attributes of the feature should be associated with an object. In this way, the data collector defines the feature attributes, updating the database table records of all attributes associated with the feature. This results in a simpler and faster system: it is analogous to using a template in MS Access to fill a database table.

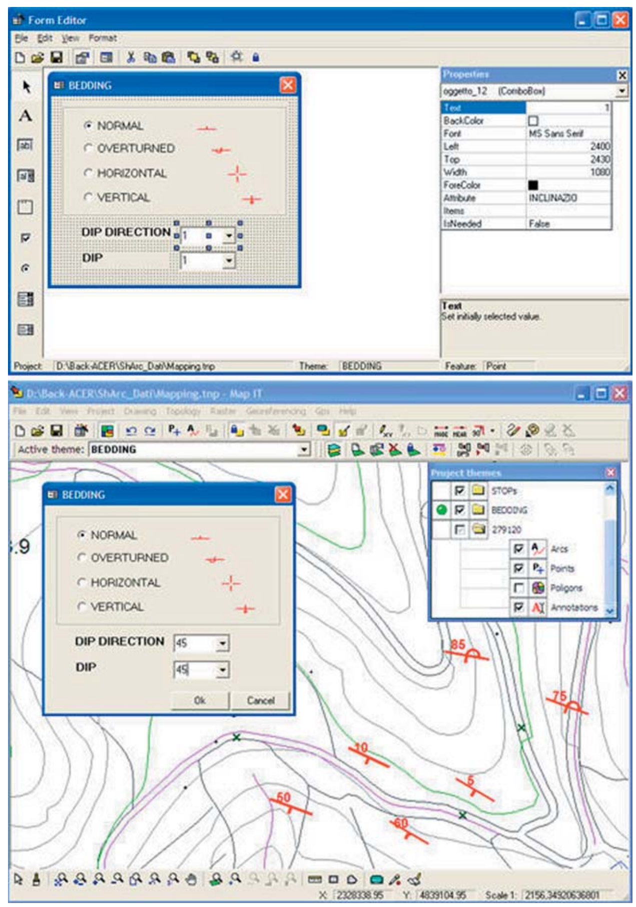


Figure 4. Form Editor shown in the upper part of the image. When you insert an element in a selected layer where you have a form, this will appear immediately (simple example of a form for bedding measurement).

DOUBLE LEVEL OF USERS

Prior to beginning a mapping project, it is necessary to properly design it. Map IT provides the possibility to build a project on two levels.

- The **first level** is for the *manager*, who sets all of the characteristics of the project which the data collector will use during field work. This phase consists of the preparation of a series of geologic and base cartographic themes that will serve as the basis for the new mapping, and for defining the features and their attributes that likely will be encountered during mapping. Moreover, the manager defines the symbology of the graphic objects that the data collector must utilize.
- The **second level** is that of the *mapper* who physically carries out the field work. The mapper may be the manager or a different person. The mapper will work within the simplified bounds of the interface prepared by the manager for the quick insertion of information. Each time that a cartographic element is added in a theme, the appropriate form appears for the insertion of the data that must be gathered for the associated theme.

CONCLUSIONS

Map IT brings together the "rugged field geologist" and the "GIS desk specialist", creating the role of the digital field mapper. The benefits of this new capability include more efficient and standardized workflow, and a reduced chance of losing data during the course of processing. You can find more information and contacts at http://www.uniurb.it/ISDA/MAPIT/index.htm.

REFERENCES

Akciz, S.O., Sheehan, D.D., Niemi, N.A., Nguyen, H., Hutchison, W.E., Carr, C.E., Hodges, K.V., Burchfiel, B.C., and Fuller, E., 2002. What does it take to collect GIS data in the field? (abstract): Geological Society of America Annual Meeting, Abstracts with Program, p. 185-29, http://www.geosociety.org/meetings/2002/.

Brimhall, G.H., and Vanegas, A., 2001, Removing Science Workflow Barriers to Adoption of Digital Geological Mapping by Using the GeoMapper Universal Program and Visual User Interface, in D.R. Soller, ed., Digital Mapping Techniques '01—Workshop Proceedings: U.S. Geological Survey Open-file Report 01-223, p. 103-114, available at http://pubs.usgs.gov/of/2001/of01-223/brimhall.html.

Briner, A.P., Kronenberg, H., Mazurek, M., Horn, H., Engi, M., and Peters, T., 1999, Fieldbook and geodatabase-tools for field data acquisition and analysis: Computers & Geosciences, v. 25, no. 10, p. 1101-1111.

Brodaric, B., 2004, The design of GSC FieldLog: ontology-based software for computer aided geological field mapping: Computers & Geosciences, v. 30, no. 1, p. 5-20.

Carver, S., Heywood, I., Cornelius, S., and Sear, D., 1995, Evaluating field-based GIS for environmental modelling, characterisation and decision support: International Journal of Geographical Information Systems, v. 9, no. 4, p. 475-486.

Howard, A.S., 2002, Capturing digital data in the field -The British Geological Survey's SIGMA Project: digital field data capture in a corporate context, in Proceedings of Capturing digital data in the field Workshop 2002, available at http://www.bgs.ac.uk/dfdc/details.html.

Maerten, L., Pollard, D.D., and Maerten, F., 2001, Digital mapping of three-dimensional structures of Chimney Rock: Journal of Structural Geology, v. 23, p. 585-592.

Niemi, N.A., Sheehan, D.D., Akciz, S.O., Hodges, K.V., Nguyen, H.L.Q., Carr, C.R.E., Hutchison, W.E., Studnicki-Gizbert, C., Burchfiel, B.C., and Whipple, K.X., 2002, Incorporating handheld computers and pocket GIS into the undergraduate and graduate field geology curriculum; Geological Society of America Annual Meeting, Abstracts with Program, p. 133-136, http://www.geosociety.org/meetings/2002/.

Soller, D.R., ed., 2000. Digital Mapping Techniques '00—Workshop Proceedings: U.S. Geological Survey Open-File Report 00-325, available at http://pubs.usgs.gov/of/2000/of00-325/.

Walker, J.D., and Black, R.A., 2000, Mapping the outcrop: Geotimes, v. 45, no. 11, p. 28-31.

Dynamic Digital Maps: A Means to Distribute Maps and Associated Media via Web and CD

By Christopher D. Condit

Department of Geosciences
University of Massachusetts, Amherst
Amherst, MA 01003-9297
Telephone: (413) 545-0272
e-mail: ccondit@geo.umass.edu
URL: http://ddm.geo.umass.edu

ABSTRACT

Dynamic Digital Maps (DDMs) are computer programs that provide a way to distribute and maximize the use of map products in an easily accessible digital format. High quality color maps, digital images, movies, analytical data, and explanatory text, including collar text and field guides, can be integrated in this cross-platform web enabled format that is intuitive to use, easily and quickly searchable, and requires no additional proprietary software to operate. Maps and photos (saved as jpeg files) and movies are stored outside the program, which acts as an organizational framework and index to present these data. Analytical data are uploaded and stored as tab-delimited text within the program, and can be saved as text documents, for use out of the program, or for inclusion in traditional databases. An open source program, the “DDM-Template” into which you can insert your data, and an accompanying “Cookbook” on how to do this are available at http://ddm.geo.umass.edu, along with numerous DDMs that demonstrate this potential. Making a DDM from the Template requires the use of the multi-platform programming environment Revolution (www.runrev.com), which has a low learning curve. Once your data have been added to the “DDM-Template” and stored in specified directories, a single short step allows you to create stand-alone programs for numerous Unix, and all Windows and Macintosh, operating systems. The correct stand-alone DDM program for a given user’s operating system can be made available for download from http sites. The DDM program can then access its associated data directly from that site with no browser needed. Alternatively, the entire package can be distributed and used from CD, DVD, or from flash-memory storage. The Office of the Massachusetts State Geologist is experimenting with the production of the Marlborough Quadrangle as a Dynamic Digital Map (http://ddm.geo.umass.edu/ddm-marl).

INTRODUCTION

With the advent of the digital age, and the public’s demand and dependence on it as a primary source of information, geologic map publishers face the difficult problem of how most effectively to distribute their products in this new medium. The problem centers around how best to integrate the associated text, analytical data, and media (images and movies) demanded by the public, into an easily usable map-based package that can be distributed via both the web and CD/DVD. Ideally, this product must be compatible with a variety of computer operating systems, not be dependent on browsers, and should not require the user to own any specialized software to use. With its inception an outgrowth of the U.S. Geological Survey’s map modernization program, the Dynamic Digital Map of the Springerville Volcanic Field (below, “DDM-SVF”), published by the Geological Society of America on CD (Condit, 1995a, 1995b, 1999, and Condit and others, 1999) was a first step toward answering this challenge. Subsequent funding by the National Science Foundation resulted in the creation of a DDM Template that others can use to make their own DDMs. This Template is a cross-platform open-source computer program into which map authors, editors, or compilers insert their own map data using the Revolution programming environment. The map data to be inserted includes file names, which are used by the program to link external data files (for example, maps and images in jpeg format, QuickTime movies, animated gifs, etc.), and text files (for example, map collar text, field trip guides, geologic settings) and any analytical data, such as geochemistry. When modifications to the Template are complete, the map creator instructs Revolution to, in a single step, make numerous stand-alone programs (DDMs) for different operating systems. These DDMs can then be compressed and saved as zip or dmg (disk image) files, and made available for download from a web page (Figure 1). This manner of distribution

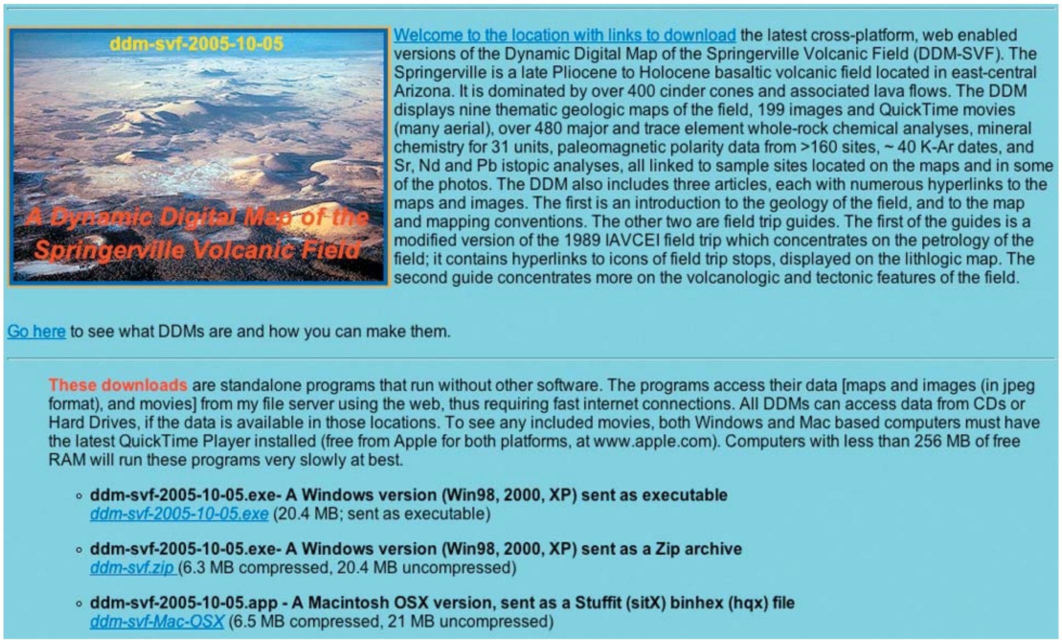


Figure 1. A typical web page of a Dynamic Digital Map (of the Springerville Volcanic Field, in this case, and referred in some other figures as "DDM-SVF"). From this page the viewer can download the DDM that is compatible with their operating system.

requires the user to be linked to a fast web connection, so that associated files can be displayed by the program as it asks for them from the DDM's http address. Alternatively, the DDMs can be distributed on CD, DVD, or flash-memory drives.

In this paper I describe DDMs and introduce you to how you can make them. DDMs are not, it should be understood, a substitute for the analytical capabilities of data base programs, or for GIS such as ArcGIS, which are superb tools in the own right. Instead, DDMs provide a tool for anyone who wishes to publish color maps, images, movies, animations and analytical data in a universal format requiring no proprietary software to use. They integrate the digital publication of high quality color maps and associated media in a way that can reach an audience ranging from the research specialist to the interested layman.

A LOOK AT A DDM

The start-up page of the DDM (Figure 2) gives the user the option to access associated media (the DDM's maps, image, and movies) using "Local Access" data sources (for example, CD, DVD, or a hard drive) or remote "WEB Access" sources. The link to "Program Status Notes" opens a page with text that may be aimed at the programmer, or may be aimed at the user, giving plans for on-going improvements or additions to the DDM program. After the user chooses the data source (web or local disk), the program opens the DDM's "Home Screen".

The "Home Screen" page features an index map that provides links to open detailed maps (Figure 3) from graphic illustrations outlining the areas of these map segments. Because the maps that a DDM displays are jpeg images, and the size limit for these raster files is 4000 x 4000 pixels, any map larger than those dimensions must be divided into smaller segments. For maps that have the option to display each of these areas as different thematic types (for example, a topographic map and an orthophoto map), selection buttons across the bottom of the screen give that choice. On the left, beneath the "Quit" button, a series of buttons open indexes that contain lists of the DDM's content. These index categories include (Figure 4) images (photos and movies), articles (for example, map collar text, guide articles), maps, correlation charts, cross sections (not shown for this map), and available analytical data.

Index lists contain hyperlinked text lines, each linking to the data described. These lists can be sorted by index number (leftmost column), or by various column headings (Figure 4). Lists can likewise be searched by keywords found in the comment field.

All DDMs have menu selections across the top of the screen, most of which, in the "Controls-Access" menu list (Figure 5), mimic the functions of the buttons found on the "Home Screen". One notable addition is the first menu selection, which allows the user to close all windows and return to the "Home Screen". The first two "Map" menu items ("Map Explanation" and "Map Features Access") open palettes explaining map unit symbols and colors

Figure 2. Start-up page of the Dynamic Digital Map of the Springerville Volcanic Field. In addition to giving the user a choice of accessing the program's data from local or remote sources via the "Local Access" and "WEB Access" buttons, the page displays acknowledgements and gives access to the web page where program updates are stored, and to notes describing recent changes made to the program.

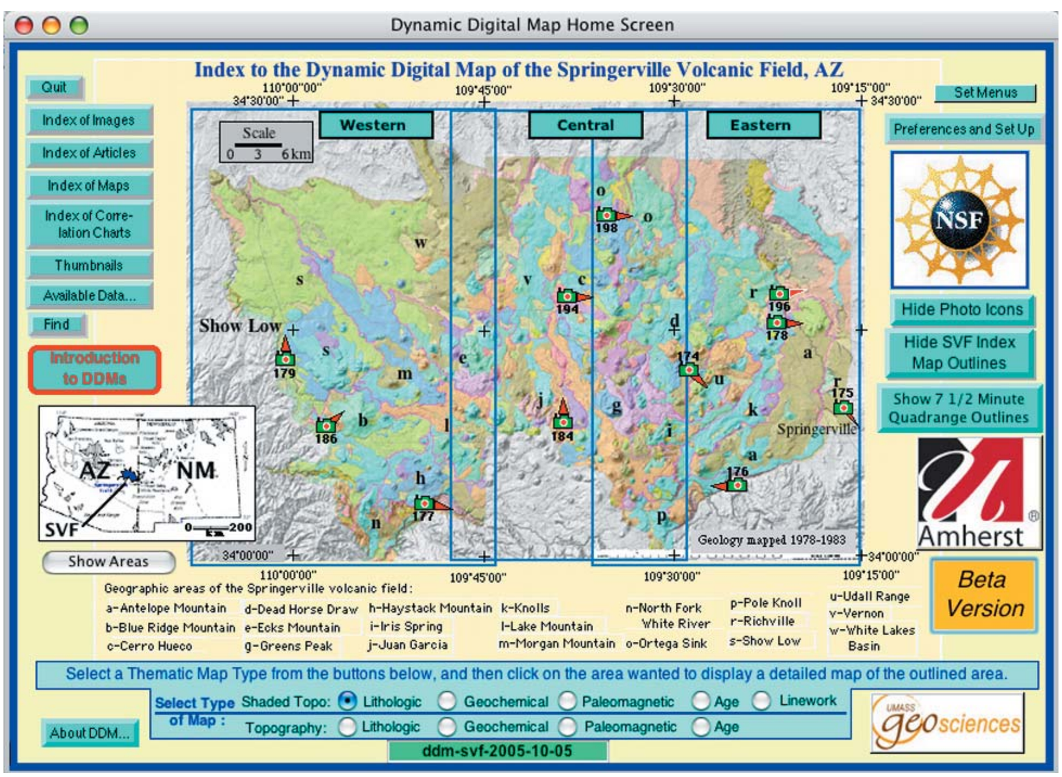


Figure 3. Home Screen of the Dynamic Digital Map. Access to all components of the DDM (maps, images, articles, correlation charts, and analytical data) can be opened from buttons on this page. Across the bottom, radio buttons allow the selection of thematic map type.

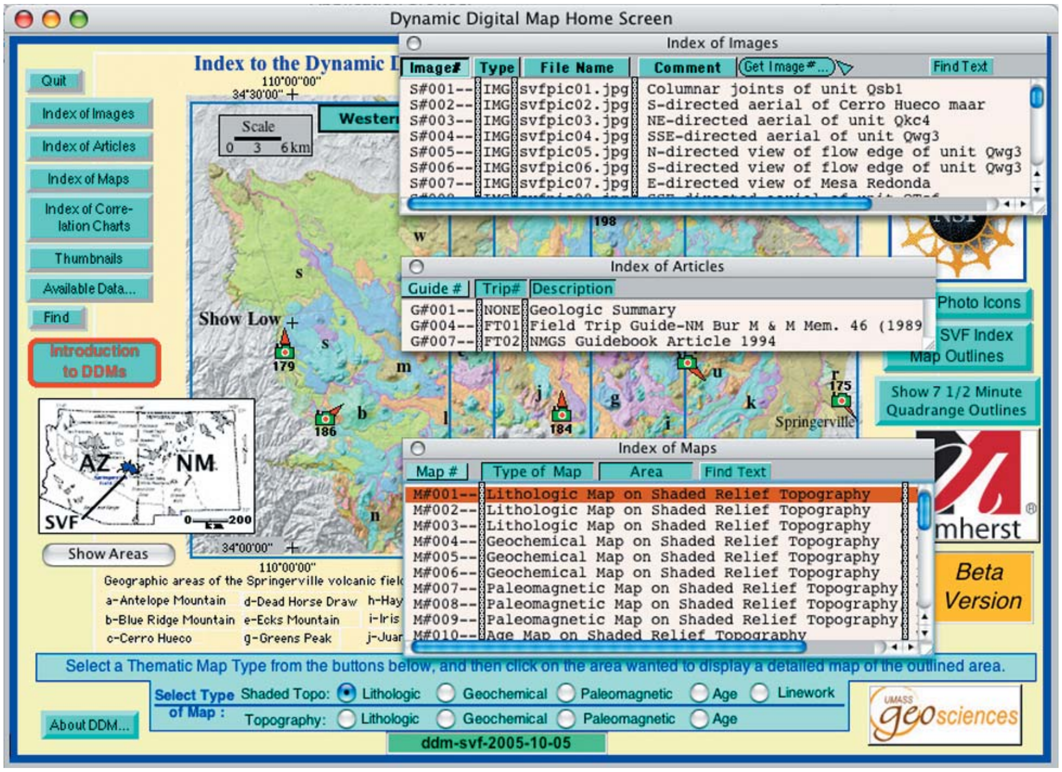


Figure 4. Examples of three types of index lists in the DDM-SVF. All lists can be sorted by column, by clicking on the column headings. For example, in the Index of Images, column sorts include (from left to right) by “Image Number”, by “Type” (Image, QuickTime movie, QT panoramic movie), by “File Name”, and alpha-numerically by first word in the ”Comment” or key words column.

Figure 5. An example of some of the Menu Selections available in the DDM-SVF. The first four “Index” menus provide linked “Tables of Contents” for the DDM. The “Index of Maps and Images With Feature” menu selection will dynamically search for and compile a list of all maps and images (including the image’s captions) that contain the asked-for words. Not shown is a “File-Save Data” menu selection used to save analytical data to tab-delimited text files. Over 480 samples of whole-rock major and trace-element data are included in this DDM, along with mineral chemistry for 31 units, magneto-polarity for 160 sites, K-Ar data for 40 sites, and Sr, Nd, and Pb isotopic data for 35, 22, and 20 sites, respectively.

for a given thematic map, and provide access to features associated with the open map segment (top center and left, respectively, of Figure 6). In the top right portion of the map view window (Figure 6), a map explanation for symbols common to all thematic maps is displayed.

All map unit symbols and sample site labels can be selected with a mouse click to open a new window or palette showing information pertaining to that label (the label is stored in the program as a named text vector object, see below). On Figure 6, a lithologic thematic map of the eastern part of the Springerville field, two labels for lava flow units Qkc6 (near the lower center of the map) and Qae2 (near the center right of the map) have been clicked on, to display floating palettes containing unit descriptions (top left of Figure). Sample site labels near these units have also been clicked to display the major-element chemistry associated with each site. Additional sample data can be displayed by clicking on the sample ID found in the Unit's Description palette, or by clicking one of the buttons across the top of that palette (as can additional Unit Description palettes, with similar clicks on map unit symbols).

Figure 7 shows the same area as Figure 6, but instead displays a different thematic map, the paleomagnetic polarity of the lava flows. Sample sites associated with paleomagnetic data replace those associated primarily with chemical and petrologic information on this map, and several sites have been clicked on to display their associated data. In the lower right side of the maps in Figures 6 and 7, note the latitude and longitude read-out in the "Map Window Control Palette", which gives the cursor location. Measurements for straight and curved line distances and areas, along with the cursor location, can be made using the "Tools" button and its associated palette (left side of Figure 7). The "Measurement Collection Palette" (upper right) records the data for map measurements, and can be saved as a text file. The drawings of these measurements (M#1 just above the palette showing magnetic polarity data, and M#2 just below the left side of

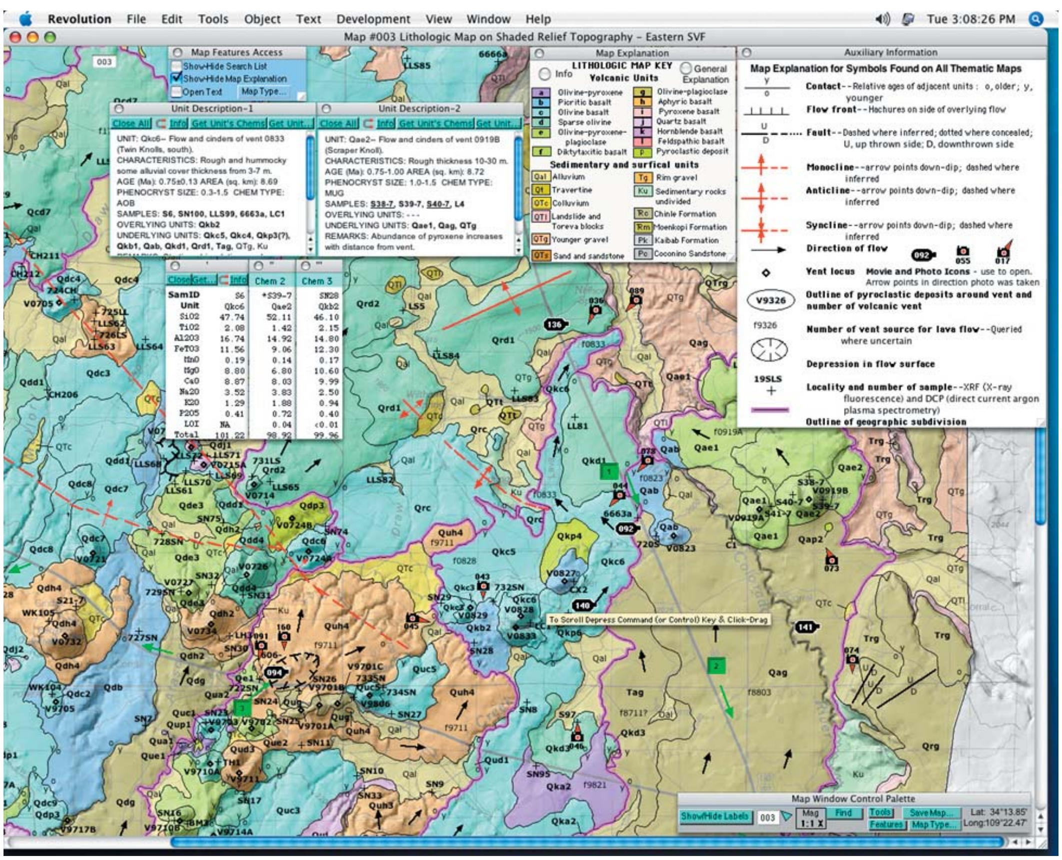


Figure 6. The easternmost of three map segments of the DDM-SVF, displaying one of nine thematic types; here, the map units reflect the lithologic variations (see map key in upper center) of the lava flows. Compiled on a 1:100K topographic map base, the geologic map is draped over shaded relief. Several of the bold symbols on the map have been selected by mouse-click, causing the data associated with them to be displayed (here, we see map unit descriptions and lava flow chemistry). Map scrolling is via scroll bars or by using a click-drag with the control (or command) key depressed (see the popup "Tool Tip", near bottom-center of map).

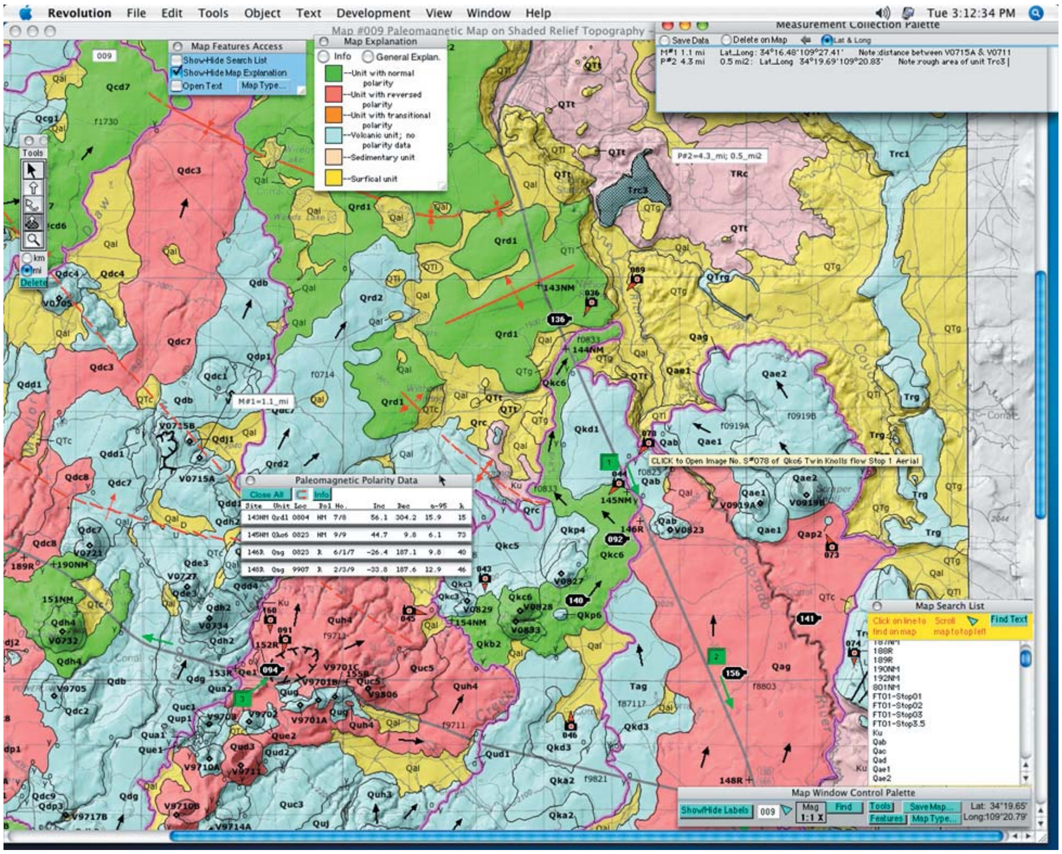


Figure 7. A paleomagnetic thematic map of the same area shown in Figure 6. The sample site symbols selected by a mouse-click display the data associated with this new thematic map type, and the units are colored to highlight the magnetic polarity of the lava flows. Note the latitude-longitude read-out of the cursor location, and the "Map Search List" palette (lower right).

the "Measurement Collection Palette") can also be saved as an image, along with the underlying map, using the "Save Map" button on the right side of the "Map Window Control Palette".

Camera icons are placed at the location of photos, and the arrows denote the direction the camera was aimed. When the user holds the cursor over a camera icon (see the camera icon "078", Figure 7, right center, just above the rollover text "Click to open…), the key words describing that image, found in the "comments" column of the "Index of Images" palette are shown as a "Tool Tip" in the rollover. Movie icons (see southwest of icon "078") likewise display key words, as do field trip stop icons (the green rectangle numbered "1" slightly closer to the image icon "078").

A click on field trip icon "1", (Figure 7, right center) opens a floating palette containing an associated article, in this case a Field Trip Guide (Figure 8). The program centers the text at the point in the guide that describes that stop. A click on the hyperlinked text in the guide (in bold font and underlined, S#078) opens the associated digital photo. An "alt-click" on that same text would scroll the map and center it on the camera icon, alternatively, an "alt-click" on the line containing the image number in the "Index of Images" palette does the same. Likewise, a click on the turquoise image number in the upper left side of an open image finds the camera icon on the map. Additional clicks on other camera icons on the map or other images will open up to nine images; a click on the movie icon "092" in image "078" will open a QuickTime panorama of the lava flow top. Any one of these images or movies may have a Figure caption associated with it, displayed by selecting the "Caption" button on the "Image Windows Control Palette" (bottom, Figure 8).

The text for each article or image/movie caption may be displayed at one of three levels of sophistication (or in one of three languages), if the DDM complier chooses to include this capability. In this way a single DDM can reach audiences ranging from the research scientist to the interested layman or student, thus maximizing its outreach potential. Changes to the "User Level" settings can be made in the "Preferences and Set Up" palette (not shown), along with changes to various other settings (for example, thematic map type, "Tool Tip" displays, desk top visibility, and which ancillary palettes to automatically open when a map segment is displayed).

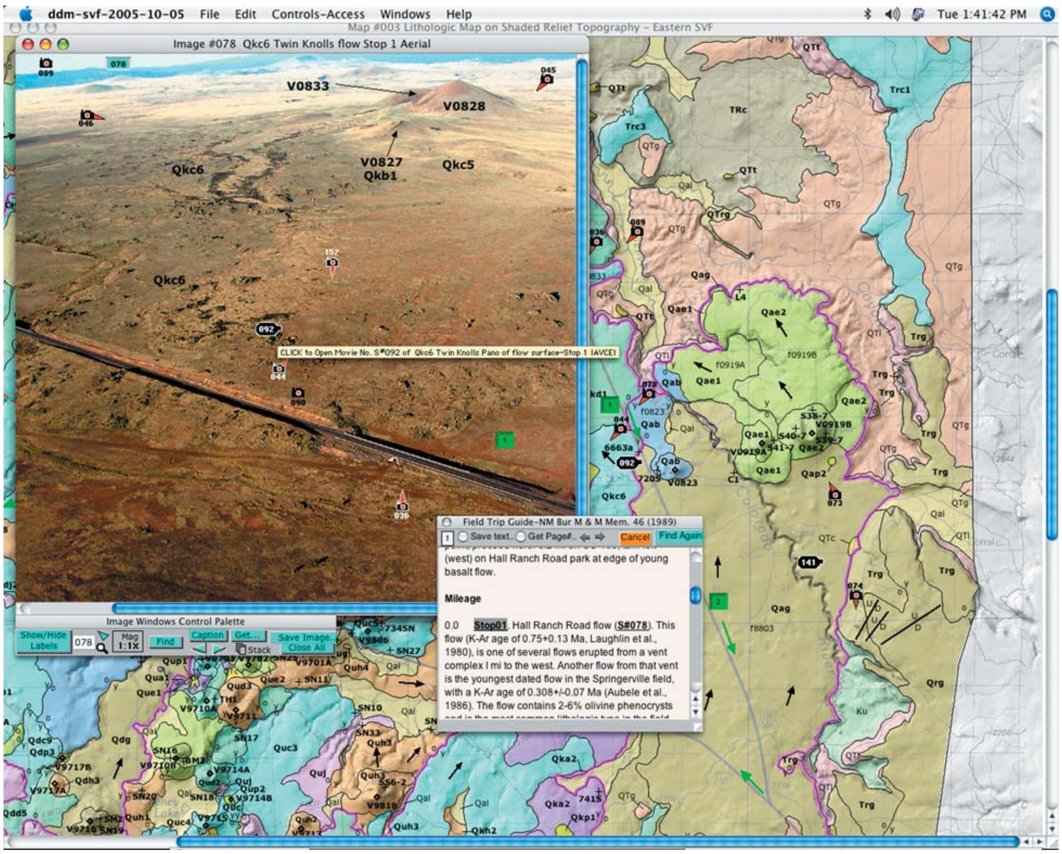


Figure 8. Example of the hyperlink-accessed image and Field Guide. This image is opened by clicking on camera icon "078" of the lithologic map. When the icon on a map that denotes a field trip location or "stop" (for example, the green 4 on this map) is selected with a mouse-click, the description of the field trip stop, as reported in the guidebook article, is displayed. The image window takes its title from the key words associated with the image number in the "Index of Images" palette. Key words also are displayed in the "Tool Tip" for movie image "092", as shown in the center of the figure.

DDMs have built into them the capability to search all the analytical data sets included in them for a sample's (or site's) data or to determine if a geologic unit has been analyzed for a given data type (for example, is there any trace element data for unit Qab?). The search will either highlight the found sample in a list of each type of data, or both highlight and display the data (Figure 9). Selecting the "Controls-Access / Available Data…" menu (or the "Available Data…" button on the "Home Screen") opens a dialog (Figure 9) that gives the user a choice of search options; a subsequent dialog box provides the user a place to enter the sample, site, or unit for which information is needed. The result (top right, Figure 9) is a series of lists of all samples for each type of analytical data. In this case I've asked the program to search for sample 717MR. In 7 of the 11 lists, sample 717MR was found, and is highlighted in the second ("Chem") and third ("Traces") lists but not in the fourth ("Min Chem") list, etc. In lists containing the selected sample the background color of the title for that data type changed to green. In those lists where sample 717MR is not found, the background title for that data type is colored orange. If that sample is found, it is highlighted in the list; clicking on the sample's entry displays the data, whereas an alt-click locates the sample on the map.

Searches for the location of objects on maps and images (for example, samples or map units), initiated from sample lists as described above, or by using the menu selection "Control-Access / Find Feature" recenters the map on the found feature. A red box surrounds the feature and flashes several times. When the feature is not found on the open map, the user is asked if they want to look for it on other maps or images. If so, the DDM searches all maps, images and image captions, and presents the user with a palette (Figure 9, lower right) containing two click lists. The upper list displays all maps with that sample, the lower list all images (and image captions) containing that sample. A click on the list opens the map or image in the line clicked on, and as above, finds the object.

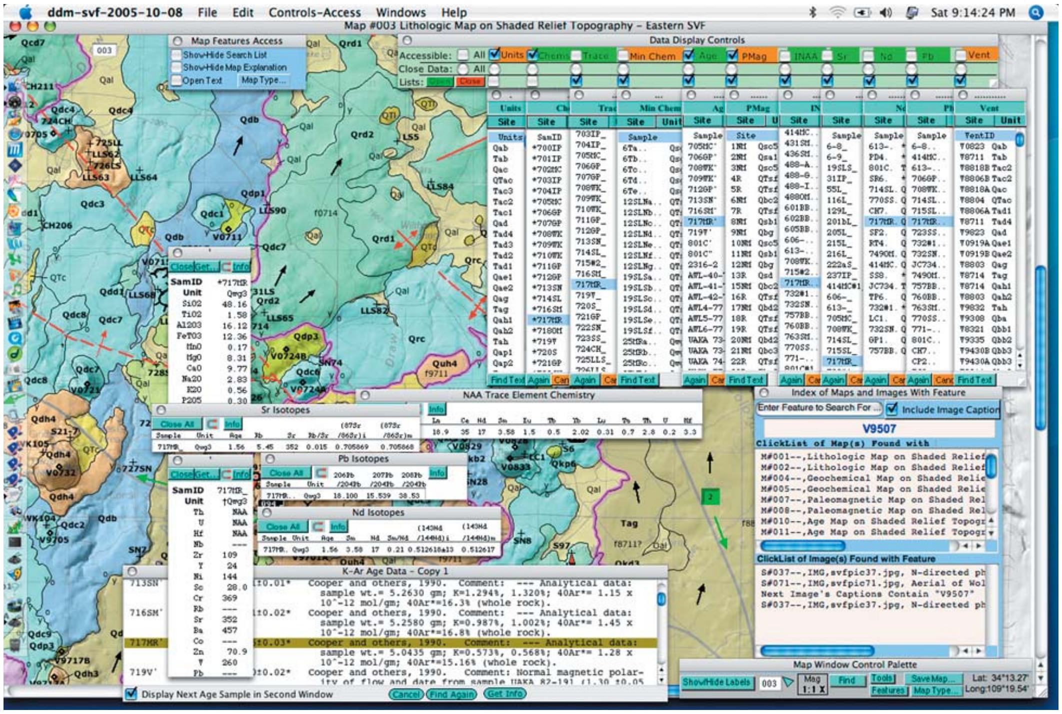


Figure 9. Screen showing the result of searching for analytical data using the DDM-SVF. Here, the "Available Data" search shows any occurrences of a sample in all data sets, and displays that data. In the lower right part of the map window, a palette displays the result of searching for a feature (a volcanic vent, number V9507) that was not found on the displayed map. All maps, images, and captions are searched for the occurrence of the feature, and a list of where it was found (maps, top part of palette; images, bottom part of palette) is presented to the user. A click on a record will display that map or image.

A "Map Search List" can be opened from a menu item, or a button in the "Map Features Palette" (Figure 8). This function compiles a list of all objects on the open map segment (for example, unit symbols, sample site designators, cultural feature names, image site icons by image number, and field trip stop icons). A click on a line in the "Map Search List" finds the object on the map, as described above.

BUILDING A DYNAMIC DIGITAL MAP

A Dynamic Digital Map is composed of two major components. The first is an open source computer program (the DDM-Template), written in a programming environment called Revolution. The second is a system of file structures into which various components that make up the individual DDM are placed. These include maps and photos (jpeg file format), and movies (QuickTime format). The DDM-Cookbook (Condit and Albrecht, 2005) contains detailed step-by-step descriptions, complete with screenshots, of how to make a DDM (http://ddm.geo.umass.edu/Cookbook2005-08-12.pdf).

The Program and Programming Environment

The DDM-Template program is essentially a shell into which the mapmaker inserts their data (for example, map collar text, text for field trip guides, associated analytical data) and metadata about external files the program must access when a user asks for them to be displayed.

The programming environment, Revolution, was chosen because it is a graphical, high-level, object oriented "language" that executes quickly, has a low learning curve, and is multi-platform. The graphic orientation is important, since mapmakers or compilers must visualize the map layout and produce a graphic product. The Revolution programming environment (Figure 10) organizes its data into packages known as "stacks" that loosely correspond to windows, with each stack containing one or more cards, on which objects can be displayed. These objects may be text, buttons, images, or vector graphics to name a few; each stack can display only one card at a time. Revolution is conceptually very similar to the Macintosh HyperCard programming environment, but

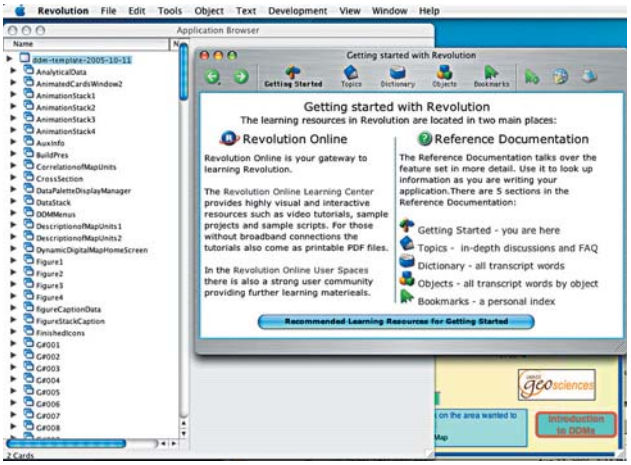


Figure 10. Example of the Revolution programming interface. The left side lists the "stacks" included in the DDM-Template shell (see text for details). On the right is an example of the documentation and programming help available.

was originally written (as MetaCard in 1990) for use on various UNIX-based operating systems, and later ported to the Windows and the Macintosh OS. The MetaCard engine technology was in 2003 acquired by Runtime Revolution, who had built a sophisticated user interface with expanded capabilities, based on that engine. (For details see http://www.metacard.com/.)

Revolution is a high-level language, meaning most of the command-line programming steps that create objects such as windows or buttons are hidden from the user, and these objects can be simply called and fit together to make the map. The objects, once created, can be duplicated (or "cloned"), and these duplicates retain all the characteristics of their "parent" object, thereby making programming more efficient. Revolution executes very quickly because it compiles the code (or scripting) associated with any object.

Revolution has a low learning curve; the essentials for creating DDMs from the DDM-Template can be learned in a day or two. The company's site (http://www.runrev.com) has web access to about a dozen excellent, free, 8-to-10 minute videos, of which 6 are directly applicable to making DDMs. The scripting is in a format familiar to those who use the English language. Built into the Revolution programming interface is an excellent suite of "Help" items, including Introduction to Revolution, Quick Reference guides, a Dictionary, and Frequently Asked Questions. The online exercise, "Getting Started with Revolution" takes a few hours and introduces you to using Revolution. I have co-authored a "DDM-Cookbook" which assumes you have taken the time to view the videos, and have worked through the online tutorial exercise. My tutorial, which takes an additional few hours, is entitled "A Quick Start to Making DDMs." It gives step-by-step instructions in how to make a DDM from images and maps already prepared. The DDM-Template that the mapmaker uses to create their own DDM was designed to minimize the need for the mapmaker to do any programming. The mapmaker's job instead is to substitute into the Template their own data and metadata to complete their own DDM.

Revolution is a multi-platform programming environment, which means one can make DDMs on Windows, Macintosh, and Linux (and other Unix) platforms. Many of my students at UMass work in my lab on Macintoshes, copy their work to USB flash-memory drives, transfer it to their Windows machines and work on it at home. They then bring their work back to the classroom, transfer it to their Mac, and continue their work. I have accommodated differences between Windows and Macintosh systems almost completely in the Template; most irksome are occasional font compatibility issues. For this reason, I recommend checking the DDM on various platforms.

Once a DDM-maker has completed their DDM in the programming environment, they can turn it into a "stand-alone" program that requires no other program to run. Because the old Macintosh OS 9 has file characteristics that cannot be created in the Windows environment, to make a Macintosh OS 9 file at the same time you make a Windows stand-alone DDM, you must be working on a Macintosh computer. You can make both a Macintosh OS X and a Windows DDM stand-alone on either Windows or Macintosh OS X machines. You make these stand-alones by checking the settings for each type of operating system on a separate page, and then simply clicking a button to start the stand-alone creation process, and (depending on the complexity of the DDM) have a cup of coffee. These executable programs can then be posted on the web, or compressed and posted; the DDM-Template has built into it the capability to access maps, images, and movies from http-capable file servers, or locally from CD, DVD, or hard disk.

The File Directory Structure

The second major component of a DDM is a system of file directories that the DDM uses to access the many associated data and image files (Figure 11). If the access mode is "Local" (i.e., not from the web), the program assumes that these directories are in the "home" directory where the other directories are located. When accessing files over the web, the program looks for the same directory structure, but with reference to a "home" directory on a file server. Metadata, which consists of file names associated with map segments, correlation charts, cross sections, and figures is entered by the DDM-maker into text fields on a card in the DDM's "dataStack". Image and movie file names are entered directly into the scrolling field on the first (and only) card of the "Index of Images" stack.

The DDM-maker places map segment files into the

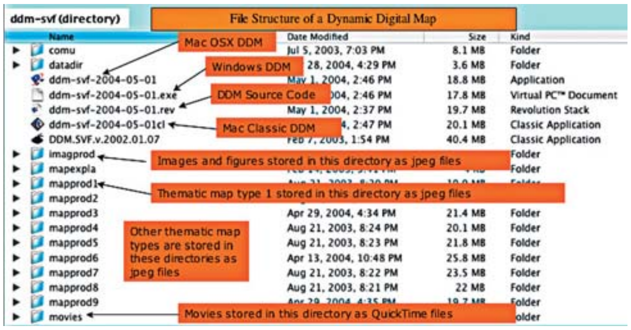


Figure 11. An example of the file structure for the DDM-SVF, showing directories and files in the "main" directory. The DDM-SVF application calls data from these designated directories and opens it into various windows designated for each.

correct directory for that thematic map type (in Figure 11, "mapprod1" to "mapprod9"). For example, let's suppose you had a geologic map divided into four tiles, one for each quadrant (NW, NE, SE, and SW), that were saved as the files named nw,jpg, ne.jpg, se.jpg and sw.jpg, respectively. These files, since they are all of the same thematic type (geology), would be stored in the directory "mapprod1". Next, suppose you also had an orthophoto map of the same area, divided into these same four tiles (essential, so that each pixel on both thematic types corresponds to a pixel of the identical location). Because the tiles in the orthophoto thematic map type encompass identical areas, you also name their files nw,jpg, ne.jpg, se.jpg and sw.jpg *but place them in the directory "mapprod2"*. Additional thematic map-type tiles would be placed in sequentially numbered "mapprod" directories (for example, mappprod3, mapprod4, etc.). Files for images and movies are both placed in the "imagprod" directory), those for the Correlation of Map Units are placed in the "comu" directory), and so forth for several others not shown in Figure 11 (such as figures, map explanations ["mapex"], image thumbnails ["thumbdir"], etc.). Maps, images, correlation charts, figures, etc. are saved as jpeg files because they are a universal format across computer platforms, are reasonably compact, and transmit well across the internet. DDMs can magnify (zoom in on) map images and figures, and display them at half or quarter resolution, although at more than 2X, pixelation occurs quickly (except in QuickTime Virtual Reality movies).

Analytical Data, and Map and Image Overlays

Analytical data are imported by the DDM-maker, from tab delimited text files into text fields in the DDM's "AnalyticalData" stack (Figure 12). Buttons (top, Figure 12) contain scripts that format the data so that it can be displayed in pop-up palettes, when the user clicks on a sample site or unit label. In most cases, up to nine analyses per data type can be displayed at one time. Unit descriptions can be displayed in either a single scrolling text field (like a word processor window), or in pop-up palettes (see Figure 6, top left), depending on the DDM-maker's preference. Each analytical data set (or all) can be saved to disk as text or tab-delimited text files using the "File / Save Data" menu selection.

DDMs make use of text field labels (for example, as sample site or unit labels), placed next to locations of map images or photo images to both open corresponding data and locate these labels on the map or image. These labels are stored in the program as part of an "overlay" group described in the next paragraph. The "Map" display stack and each of the nine image display stacks ("Image1" to "Image9") contain script that intercepts a mouse click on a text field label, and gets that object's name (an attribute that can be set to be the same as its text label). It then sends this name through script in the "AnalyticalData" stack that queries each type of analytical data field to see if an analysis with that site or sample ID is found. If so, it copies the data to a data display palette dedicated to that data type and displays it, and passes the query on to the next data type. Similarly, each of these stacks contain script that can locate the x and y coordinate of a selected text label and then scroll the map or image to center the object in the window. It also resizes an invisible red box to fit around the label and flashes it several times.

Each map and image file has a corresponding "overlay" (a group of labels), placed on a card in either the "MapOverlay" or "ImageOverlay" stack, one card per map or image (Figure 13). When a map (or similarly a digital photo) is opened by the user, the "Map" stack (or window) opens, and the map's jpeg file is called in from disk to fill an empty "image" object. Then the corresponding map's label overlay, containing the unit labels, sample sites, etc., is placed on top of that map's image object. Overlay labels are added to the DDM using DDM-Template's "ProjectModfier" palette, which, along with the Revolution interface, is the main set of tools for assembling a DDM. A map is displayed, and the DDM-maker creates and places labels in the correct locations on the map and uses a button in the "ProjectModifier" palette to add the object to the map's overlay group, and then saves it back into the program. Similarly, the DDM contains a palette (Figure 13, bottom right) that enables the DDM complier to create camera (or movie) icons with a specific number corresponding to a given image (found in the "Index of Images" palette), and add them into the map's overlay group.

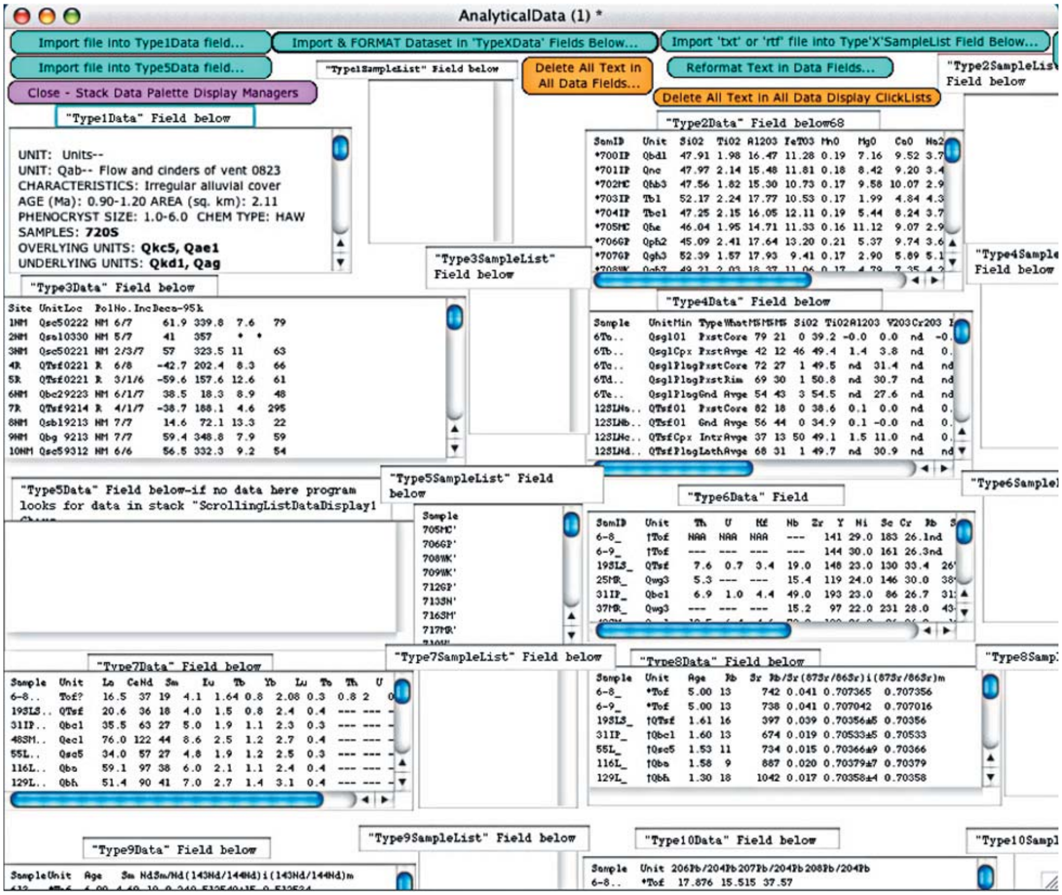


Figure 12. A DDM stores analytical data in text fields in a card in the 'AnalyticalData' stack, examples of which are shown here. The data are imported and formatted into the DDM using the buttons shown here. The input data comes from tab or comma delimited text files, such as you might save from Excel or other database programs. When a user selects a sample symbol on a map (or the sample click list), the data displayed is obtained from these fields.

Adding Text Components and File Lists into the DDM

Three types of text components can be added to DDMs. The first involves modifying the DDM-Template's Indexes to reflect the content of the DDM under construction. The other two involve adding "Articles" (for example geologic summaries, map collar text, or guidebook articles), and captions (one for each image file).

Adding text to all indexes (except for the maps) involves simply following the format of the existing lines in the Template for a given index, and replacing them with your own content. First, the editor needs to use a button in the "Project Modifier" palette (not shown) to reset the index's text field so they can type into it. For example, each image you include is assigned a new, sequential "S#", entered in the first column. You enter the type of image (jpeg image or "IMG", QuickTime Movie or "QTM", or QuickTime Virtual Reality panoramic movie or "VRM") in column two, and in column three, enter the file name. All file names adhere to the "8.3" character convention, to ensure files that won't be garbled if you are using the web for access. The last column contains key words describing the image; this automatically becomes the title of the image window when the image is opened, and fills "Tool Tips" when a camera icon is "moused over". Other indexes can be filled the same way; when finished filling the index, you again use a button in the "Project Modifier" palette to reset the index's text field so you can use it to open files.

Captions can be written in your favorite word processor, and either pasted directly into one of the three different "user level" fields. Alternatively, and especially useful for heavily formatted text (for example, with numerous

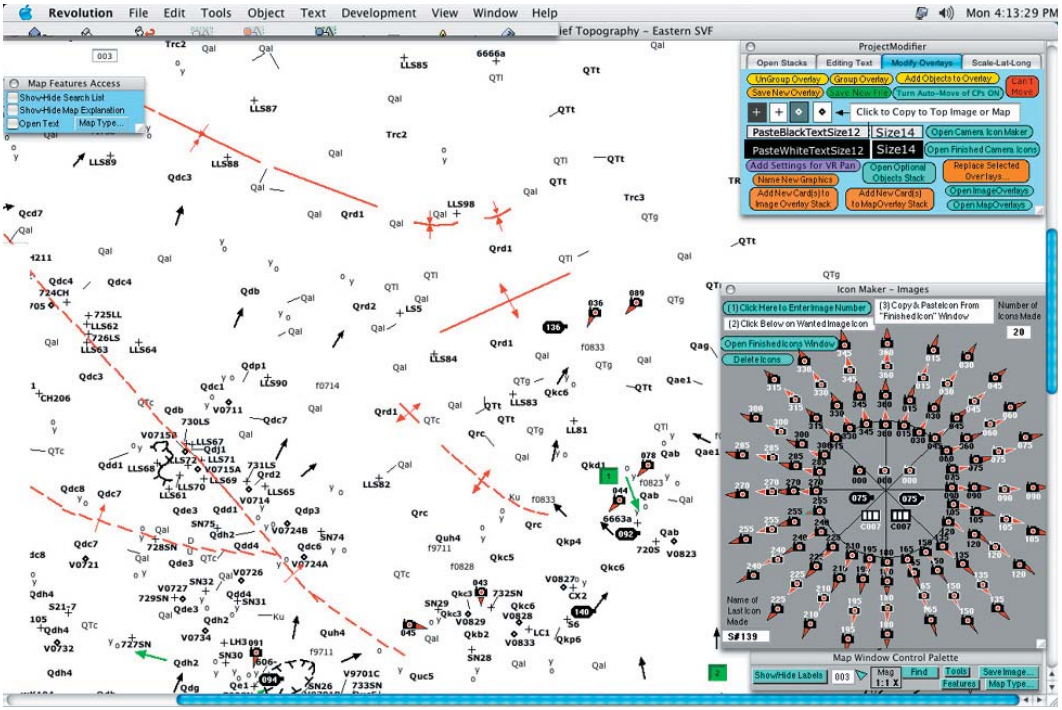


Figure 13. A “Map Overlay” composed of a group of map units and sample labels for the geologic map shown in Figure 6. A DDM stores map symbols, sample sites, and other graphic objects that the DDM-maker wants to have associated with locations on the map, as named vector objects. These are overlain on the map (or image) when it is opened by the DDM program. The DDM-maker can use the “Icon Maker-Images” palette (lower right) to place camera icons in the proper locations on the maps or images in the program.

super- and sub-scripts), text files can be saved in “rtf” or, better yet, in “html” format out of your word processor, and imported, using the “Insert Formatted Text” button in the “imageCaptionData” stack. Each caption is preceded by the file name and two dashes, with the caption ending in a carriage return, as you would a paragraph. Depending on the “user level” chosen by the user, the caption, when requested, is copied out of one of these fields and inserted into the floating “Image Caption Palette” that can be opened when an image is displayed.

Likewise, text for Articles can be imported from word processor files that were saved in html or rtf format. Each article can be entered into from one to three different fields, corresponding to different user levels. The ‘Article” palette (lower left, Figure 14) entitled ‘Field Guide to the SVF”, contains some underlined words in bold type that, when selected, will automatically invoke a response. The underlining shows that I have set these letters in Revolution to be “linked” text, to start this chain of events. For example any linked text starting with “http://” will automatically open your default web browser to that URL, as will links starting with “www”. Likewise, a link that starts with the symbols “S#” will automatically open that image or, if “alt-clicked”, will search for that image icon in an open map or image (as will a link starting with “Stop” and ending in a two digit number). Other symbols will likewise open other objects not described in this article (F# for figures, CC for Correlation Charts, XC for cross sections). The Cookbook describes how, with some simple scripting applied to the “Article” palette, you can do other things with these links (like open animations, etc.).

DISTRIBUTION OF QUADRANGLE MAPS VIA DDM

The Office of the Massachusetts State Geologist faces the task of making maps easily accessible to citizens in the most efficient and expeditious manner possible. Many of these potential users have neither the expertise nor software to use a GIS product, but have the basic computer skills to use an integrated stand-alone product such as the DDM. To accommodate that need, the Office is experimenting with the release of a preliminary version of the Marlborough Quadrangle of Massachusetts as a DDM (Figure 15).

A click on the proper link at the DDM-Marl web site (Figure 16, http://ddm.geo.umass.edu/ddm-marl/) will

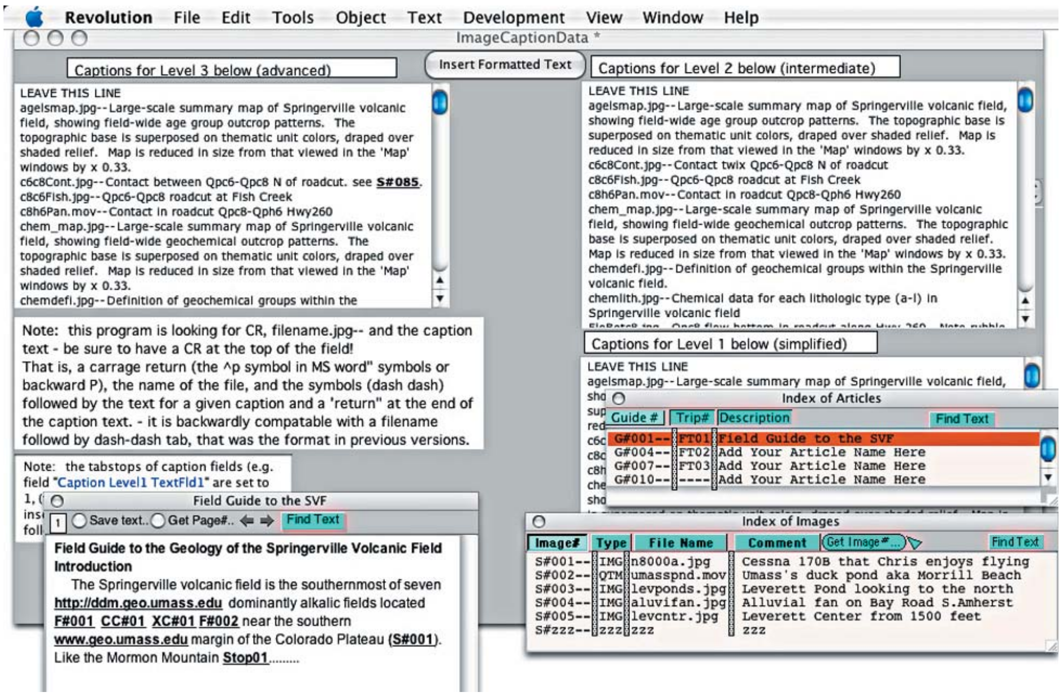


Figure 14. Examples of text data components of DDMs, shown in four stacks. Here, for example, figure or image captions and geologic summaries or guidebook articles are stored in the DDM-Template. They are put there by either cutting them from word processor documents and pasting them in specified places in the Template, or by importing formatted text (in html format) using buttons installed in the Template.

download a version of the software appropriate for the user's computer. Provided the user has a fast web connection, they can then use the "Web Access" button to open maps, image, figures, etc. from the web server located at the University of Massachusetts. The DDM can also be distributed on CD, DVD, or USB flash-memory device, for use without web access. All maps, data, images, and figures can be saved out of the program to the local computer's hard drive, and printed for hard copy use.

The "DDM-Marl" contains three thematic maps (bedrock geology [Figure 17], surficial geology, and surficial material, with simplified and detailed unit explanations), a cross section (for the bedrock map), two correlation charts (for the bedrock and surficial maps) and three articles, one describing each map type. The map is referenced with latitude-longitude read-outs of the cursor location, and has both linear and area measurement capabilities. It also contains 136 images (mostly outcrop photographs) with captions describing each, 77 figures (scans of borehole data sheets), and fracture data for over 3057 stations, linked to 32 sites plotted on the maps. Each map contains labels of all units, and major geographic and cultural map features that can be searched for and located, using the "Map Search List" palette (lower right, Figure 17).

SUMMARY

DDMs provide a self-contained way to distribute map products that can be tightly integrated with associated data, text and multimedia products. DDMs are made using the programming environment Revolution, from an open source "DDM-Template" or shell program. Using this shell, into which the map builder inserts their own text, analytical data, and metadata, the final DDM application can then access the media from a file structure on the local computer or via the web. Once assembled in Revolution, the mapmaker can, in a single step, create these cross-platform stand-alone applications for Windows, Macintosh, Linux and other Unix operating systems. These DDMs are web-enabled, intuitive to use, easily searchable, and require no additional proprietary software to operate. Examples of more than 10 DDMs can be downloaded from links at the web site, http://ddm.geo.umass.edu, where you can also download the DDM Cookbook in pdf format, and the DDM-Template. Trial versions of Revolution can be found at www.runrev.com.

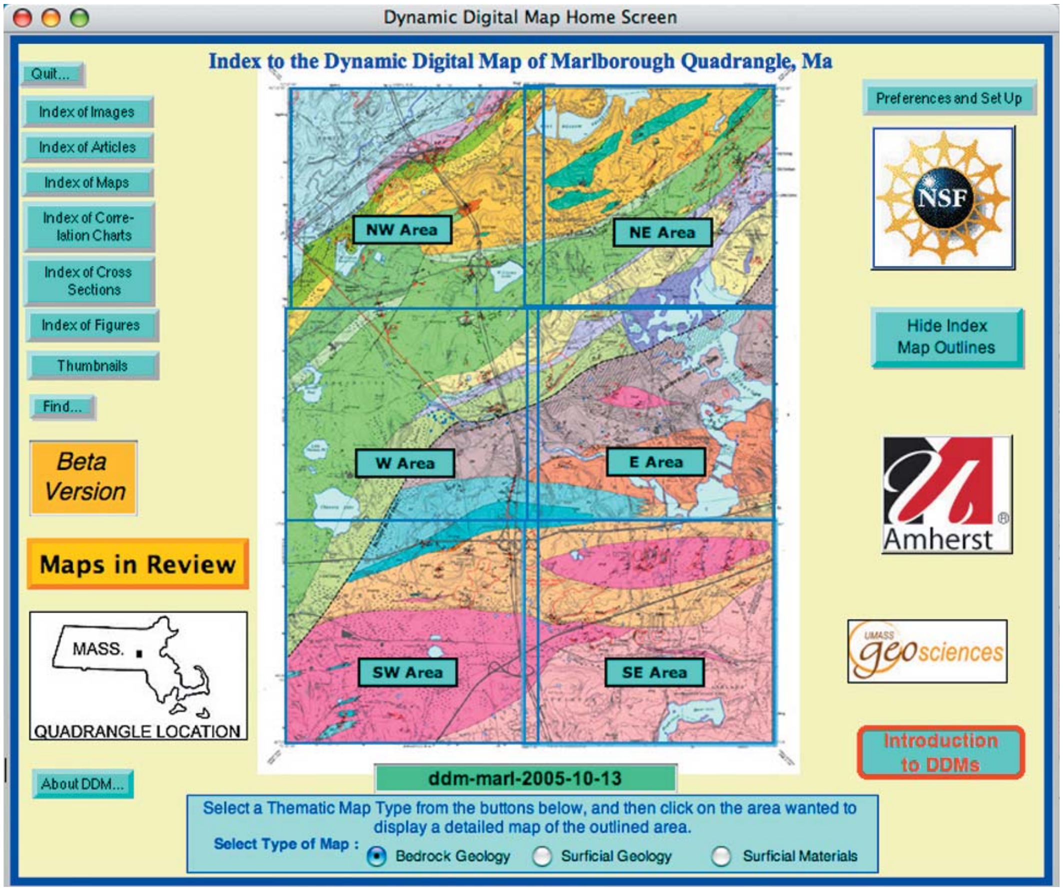


Figure 15. Home screen of the Dynamic Digital Map of Marlborough (DDM-Marl). The maps in this DDM are saved as one continuous jpeg image; the map view simply scrolls to the selected part of the map. The jpeg images are 2821 pixels wide by 3836 pixels high. Comfortable transmission rates of reasonably good quality images, with file sizes of about 2.5 MB, are possible via the web. Better quality maps, that is, with less lossy compression, are supplied via CD or DVD.

REFERENCES CITED

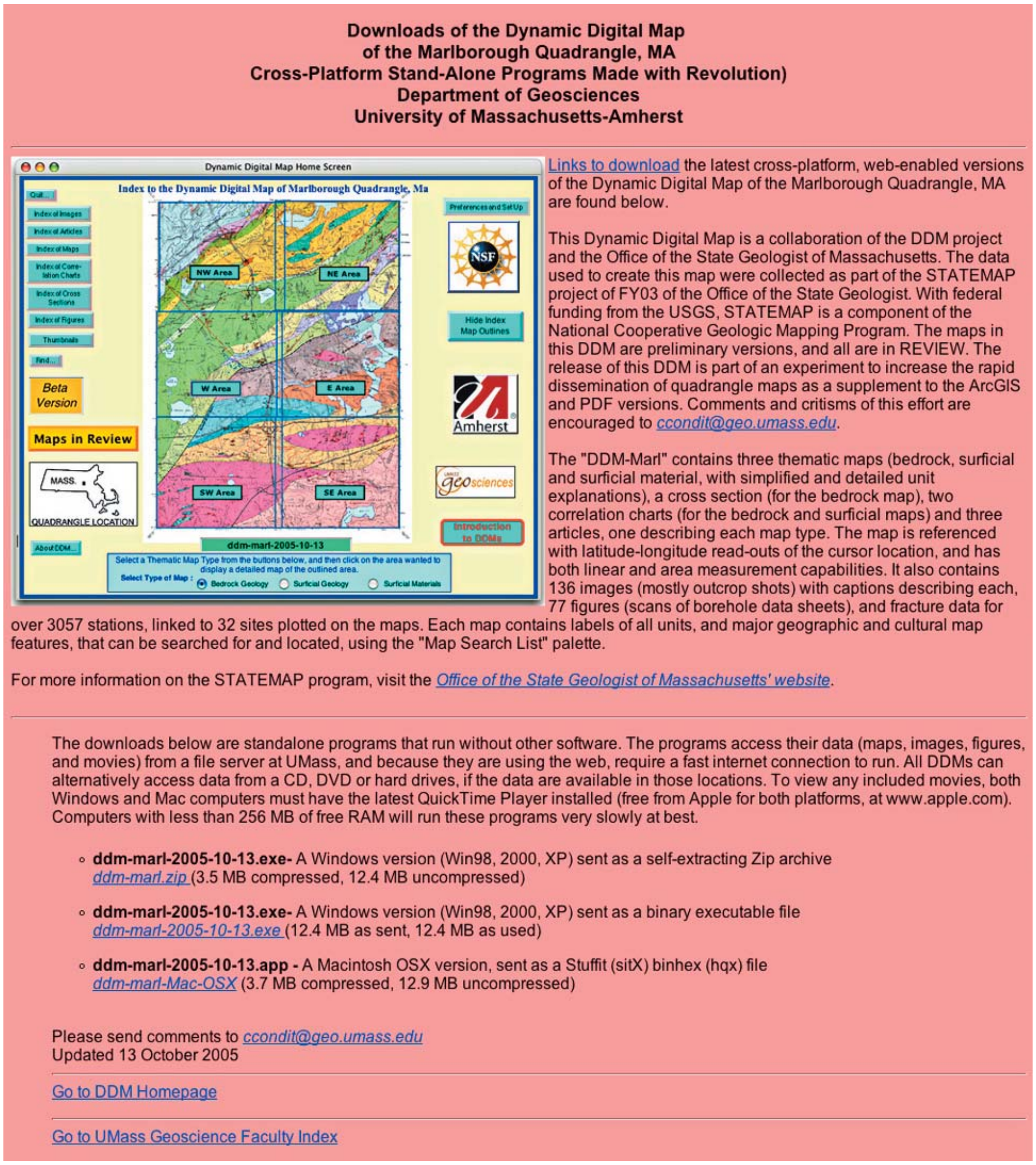
Condit, C.D., 1999, Components of Dynamic Digital Maps: Computers & Geosciences, v. 25, p. 511-522.

Condit, C.D., 1995a, DDM.SVF: A prototype Dynamic Digital Map of the Springerville volcanic field, Arizona: GSA Today, v. 5, p. 69, 87-88.

Condit, C.D., 1995b, Dynamic Digital Map: The Springerville Volcanic Field: Prototype color digital maps with ancillary data: Geological Society of America Digital Publication Series DPSM01MC (CD-ROM for the Macintosh), v. 4.10.95, 36.7 megabytes.

Condit, C.D., and Albrecht, T.R., 2005, A Cookbook for Dynamic Digital Maps, 68p., available at http://ddm.geo.umass.edu/Cookbook2005-08-12.pdf.

Condit, C.D., Crumpler, L.S., and Aubele, J.C., 1999, Lithologic, age-group, magnetopolarity and geochemical maps of the Springerville volcanic field, east-central Arizona: U.S. Geological Survey Miscellaneous Investigation Series Map I-2431, scale 1:100,000, 5 sheets, available at http://geopubs.wr.usgs.gov/i-map/i2431/.



Downloads of the Dynamic Digital Map
of the Marlborough Quadrangle, MA
Cross-Platform Stand-Alone Programs Made with Revolution)
Department of Geosciences
University of Massachusetts-Amherst

Links to download the latest cross-platform, web-enabled versions of the Dynamic Digital Map of the Marlborough Quadrangle, MA are found below.

This Dynamic Digital Map is a collaboration of the DDM project and the Office of the State Geologist of Massachusetts. The data used to create this map were collected as part of the STATEMAP project of FY03 of the Office of the State Geologist. With federal funding from the USGS, STATEMAP is a component of the National Cooperative Geologic Mapping Program. The maps in this DDM are preliminary versions, and all are in REVIEW. The release of this DDM is part of an experiment to increase the rapid dissemination of quadrangle maps as a supplement to the ArcGIS and PDF versions. Comments and critisms of this effort are encouraged to *ccondit@geo.umass.edu*.

The "DDM-Marl" contains three thematic maps (bedrock, surficial and surficial material, with simplified and detailed unit explanations), a cross section (for the bedrock map), two correlation charts (for the bedrock and surficial maps) and three articles, one describing each map type. The map is referenced with latitude-longitude read-outs of the cursor location, and has both linear and area measurement capabilities. It also contains 136 images (mostly outcrop shots) with captions describing each, 77 figures (scans of borehole data sheets), and fracture data for over 3057 stations, linked to 32 sites plotted on the maps. Each map contains labels of all units, and major geographic and cultural map features, that can be searched for and located, using the "Map Search List" palette.

For more information on the STATEMAP program, visit the *Office of the State Geologist of Massachusetts' website*.

The downloads below are standalone programs that run without other software. The programs access their data (maps, images, figures, and movies) from a file server at UMass, and because they are using the web, require a fast internet connection to run. All DDMs can alternatively access data from a CD, DVD or hard drives, if the data are available in those locations. To view any included movies, both Windows and Mac computers must have the latest QuickTime Player installed (free from Apple for both platforms, at www.apple.com). Computers with less than 256 MB of free RAM will run these programs very slowly at best.

- **ddm-marl-2005-10-13.exe-** A Windows version (Win98, 2000, XP) sent as a self-extracting Zip archive *ddm-marl.zip* (3.5 MB compressed, 12.4 MB uncompressed)
- **ddm-marl-2005-10-13.exe-** A Windows version (Win98, 2000, XP) sent as a binary executable file *ddm-marl-2005-10-13.exe* (12.4 MB as sent, 12.4 MB as used)
- **ddm-marl-2005-10-13.app -** A Macintosh OSX version, sent as a Stuffit (sitX) binhex (hqx) file *ddm-marl-Mac-OSX* (3.7 MB compressed, 12.9 MB uncompressed)

Please send comments to *ccondit@geo.umass.edu*
Updated 13 October 2005

Go to DDM Homepage

Go to UMass Geoscience Faculty Index

Figure 16. Web page for the DDM-Marl. The Office of the Massachusetts State Geologist is experimenting with using DDMs as one mode of distributing geologic maps and associated data. This map can be accessed from the URL http://ddm.geo.umass.edu/ddm-marl/. The user can download the application that matches their operating system, and access the maps, images, etc. via a fast web connection using this method.

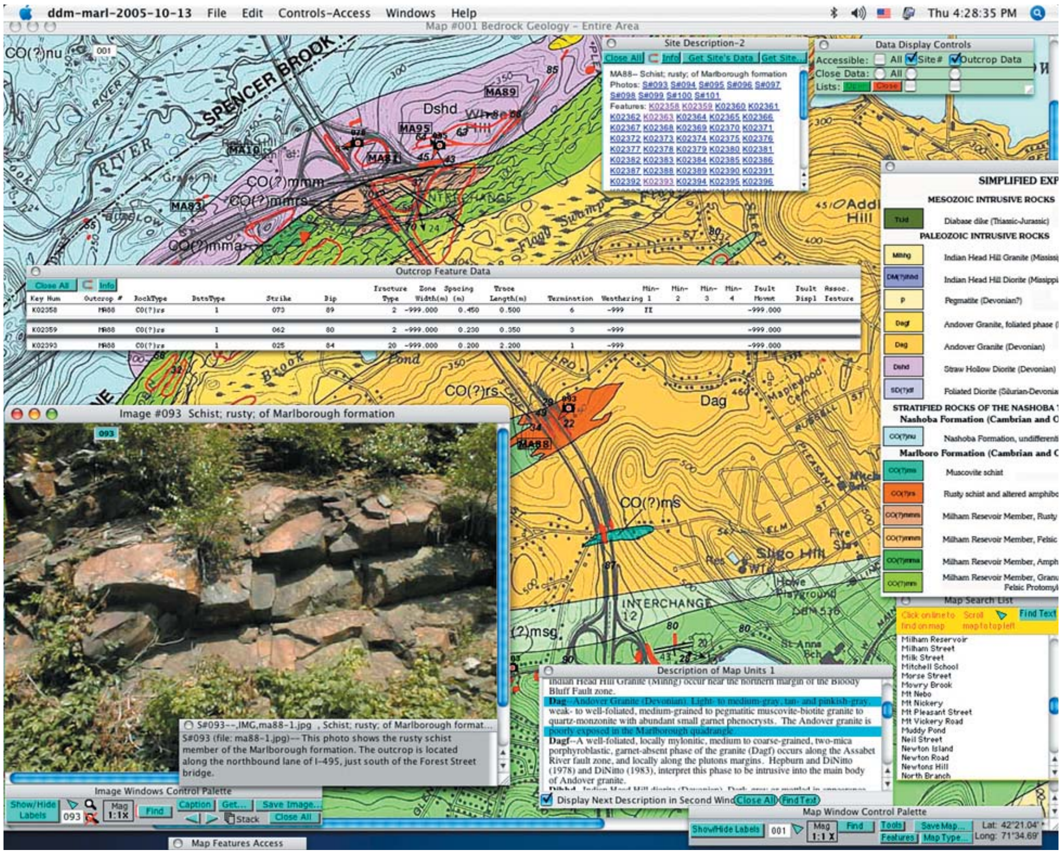


Figure 17. Bedrock geology map of the Marlborough Quadrangle, one of three thematic maps comprising the Dynamic Digital Map of Marlborough, MA. Note the latitude-longitude read-out capability (lower right) for the cursor location, and the “Map Search List”. A click on a feature’s name in the List will center the map on it. A simplified Explanation of Map Units for the bedrock map (right, center) and the complete text describing units (lower right) are displayed. A click on a unit’s box in the Simplified Explanation will center the map on that unit symbol. An alt-click on it will open the more detailed description. Fracture data for given sites also are shown (center). The “Data Display Controls” palette lets the user open and close these palettes quickly, or select which types of data to display when a site is selected on the map. A photograph of one of the sites (camera icon 093, center) is shown in the lower left, with its figure caption.

Building 3D Geological Models Directly from the Data? A new approach applied to Broken Hill, Australia

By Philip McInerney[1], Antonio Guillen[1], Gabriel Courrioux[2], Philippe Calcagno[2], and Terry Lees[3]

[1]Intrepid Geophysics
Unit 2, 1 Male Street
Brighton, Vic 3186
AUSTRALIA
Telephone: +61 3 9593 1077
Fax: +61 3 9592 4142
e-mail: {phil, antonio}@intrepid-geophysics.com

[2]BRGM
3 avenue Claude Guillemin
BP 6009, 45 060 Orléans
FRANCE
Telephone: +33 2 38 64 34 34
Fax: +33 2 38 64 35 18
e-mail: {g.courrioux, p.calcagno}@brgm.fr

[3]Copper Strike Limited
Level 9, 356 Collins Street
Melbourne, Vic 3000
AUSTRALIA
Telephone: +61 3 9640 0955
Fax: +61 3 9642 0698
e-mail: terry@copperstrike.com.au

ABSTRACT

With 3D geologic modelling, it is frequently difficult to incorporate new data, and to revise the geologic model. The potential field geologic modelling method described here automates the task of model building, and computes a model directly from data (the geologic observations). A geological interface (e.g., the upper surface of a geologic unit) is modelled as an iso-surface of a scalar potential field which is defined in 3D space. Structural data are treated as the gradient of the field. The interpolation of the field uses cokriging to take into account both contact and structural data, and generates surfaces that honour all of these data.

Since the model is computed directly from the observations, when new data are added to a project, a revised model can be quickly regenerated to take into account the new information. The method also exploits the regular structure of layered geology, by using a single potential to provide a set of sub-parallel surfaces to model a corresponding series of closely related horizons. For more complex geology, several potential interpolators are used; one for each different series of geologic strata. In this case, a unique geological model can be generated provided that the order of the stratigraphic succession of geologic units (the stratigraphic 'pile' or 'column'), and the onlapping or cross-cutting relationships between series are defined.

Several practical implementation issues designed to produce improved 3D models are presented. Faults can be taken into account. A network of faults—with some faults stopping on other faults—can be used. The regular geometry of fold structures can be described to improve the shape of interpolated folded surfaces. Gravity and magnetic data can be integrated with the model via inversion. These features are illustrated by application to the Broken Hill district.

INTRODUCTION

The traditional method of recording and communicating an understanding of the geological structure of a region has been to create a map of the geology (first done in 1801 by William Smith with his 'map that changed the world' (Winchester, 2001)). Geologic maps often include a cross-section to provide some insight into the third dimension. More recently there has been a growing interest in constructing complete three-dimensional models of geology, and indeed such three-dimensional models are very sophisticated in areas where extensive drilling and 3D seismic mapping provide a wealth of data.

Much more commonly, however, we never have 'enough' data, and yet we require a defendable 3D model of a project area—for a range of environmental, hazard and resource exploration and development studies. The challenge, then, is:

- to build a 3D model—often with quite sparse data due to sparse sampling of the geology as a consequence of cover, or the expense of acquiring data at depth.
- then revise the 3D model as new data are progressively added, or our interpretive understanding of the geology evolves. New data are often slowly acquired over periods of years—during which the model should evolve.

It is this latter point—the need to *revise the model*—which has driven much of the development presented here. Depending on how a model has been constructed, it can be an onerous task to make changes. The solution that is proposed here is to automate the task, and *compute* a model *directly from data* (the geologic observations). A revision, then, implies (1) adding the new data, and (2) re-computing the model from the updated database. This new approach has been implemented in a new 3D geology modelling software package—3D GeoModeller (http://www.geomodeller.com).

In this paper, the 3D methodology is discussed in the context of a modelling project completed at Broken Hill, in western New South Wales, Australia (Figure 1). Broken Hill is a world-class silver-lead-zinc resource which has been mined for over 100 years. A model (20 x 20 x 5km deep), centred on the mining district, was developed from the existing published geology, with further interpretation by the authors.

THE METHODOLOGY

Interpolation Requirements

A geologist who interprets the geology of an area typically is interpolating a line (in 2D) or a surface (in 3D) such that the interpolated shape—which represents a geological boundary - fits some set of geologic observations. In order to automatically compute a model directly from data, then, we need an *interpolator* to compute surfaces which represent geological boundaries or faults. The interpolator must be able to work with practical geologic data that can be observed in standard field-mapping practice as itemised below. Poor outcrop, and the expense of drilling, typically imposes constraints on the number and type of field observations that can be obtained.

Figure 1. Location of Broken Hill, western New South Wales, Australia.

Requirements for an interpolator include:

1. The position of a geologic contact or boundary is known at some (few ?) locations; the surface must be fitted *through* such points,
2. The attitude of the geology may also be measured, often at different locations. These orientation data (strike, dip and facing) can be represented as vectors, locally orthogonal to the geology. Since these orientation data may be recorded somewhere above (or below) the contact, and rarely *on* the contact itself, we need an interpolator which can take the orientation data into account … but *not* necessarily fit a surface *through* those data, and
3. We may also know other geologic data that was obtained from, or *within*, the unit (*not* on the contact). These data are more difficult to use, since they involve *uncertainty* … but nevertheless these data do define limits which the ideal interpolator must honour.

The complexity and unpredictability of geology make the task of interpolation challenging! It is also true, however, that geologic structures can be well-ordered and

predictable, and so it is important to use an interpolation process that can exploit any regularity that may be present in the geology. Given the *layered* nature of geological strata-forming processes:

4. There will be circumstances where we would like to fit a *series of surfaces*, all of which are sub-parallel by virtue of their shared geological history; the interpolator should be capable of generating a *set* of geologic contact surfaces which have a layered (stratified) geometrical relationship to each other.

There is one further requirement:

5. There can be discontinuities (faults) in geological horizons; thus the interpolator must be able to model such breaks along arbitrary fault surfaces.

The problem, then, is to find (a set of) surfaces which respect the overall configuration of the 3D geologic framework. Specific surfaces must pass through known sets of contact points, they must honour the directional vectors of orientation data points, and they must accommodate discontinuities at known faults. The interpolator must be general enough to model the surfaces of any arbitrarily complex 3D shape.

Interpolation Method

There are several computational algorithms designed to fit a surface to position and orientation data; many are unsuitable for our purpose since they typically require all of the relevant data to be *on* the surface being fitted; we have noted that some of our data—which we must take into account—may be *above* or *below* the geologic contact surface that we want to generate.

The interpolator method that we have developed is based on **potential field theory**. A set of smoothly curving, sub-parallel geologic surfaces in 3D space can be seen to be analogous to a set of iso-potential surfaces of a scalar (potential) field. A unique solution for the 3D geometry of the interfaces between formations is obtained by assuming that:

- contact data for each interface lie on a potential field surface (an iso-potential),
- orientation vectors are orthogonal to a local tangential plane to the potential field.

On this basis, the field <u>*increment*</u> (i.e. the *change* in potential) between any two points belonging to the same geologic interface is null. Orientation data represent the gradient or <u>*derivative*</u> of the field. The scalar field is then interpolated by **cokriging** the (null) <u>increment</u> data and their <u>derivatives</u> (Lajaunie *et al*., 1997). Interfaces (e.g., geologic contacts) are drawn as iso-values of the interpolated scalar field; iso-lines in 2D (Figure 2) or iso-surfaces in 3D.

An overview of the potential field method and the cokriging of the potentials is presented in the following sections, but for a more complete discussion see Chilès *et al*. (2004).

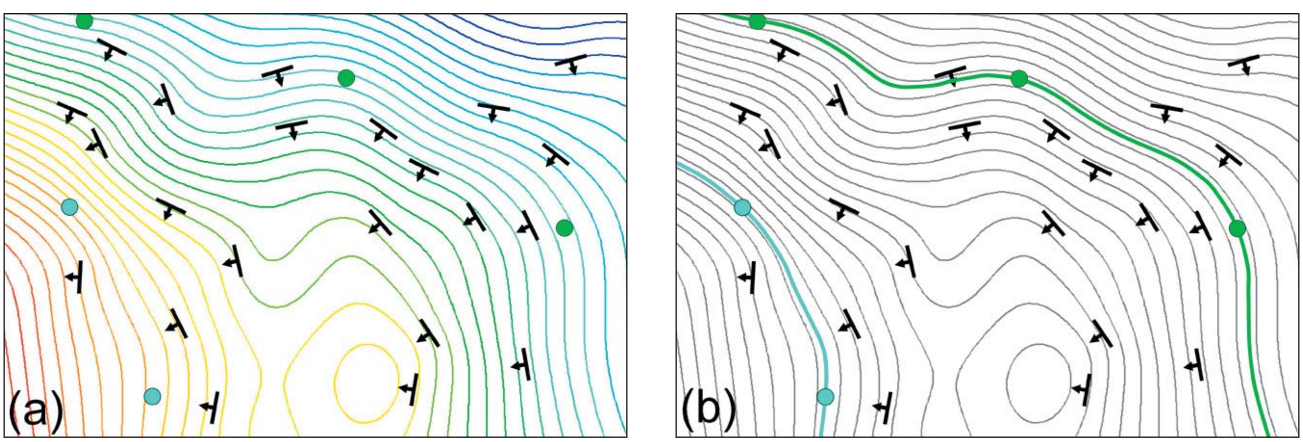


Figure 2. Map showing known geologic contacts (see black dots) for formations belonging to a single series, and also structural data. In (a) the potential field (interpolator) has been computed; note that structural data are all taken into account, with the field always orthogonal to the (structural) orientation vector. In (b) two iso-potentials of the field are plotted such that they pass through the two sets of geologic contact points. Note that the interpolator has proposed a geologic model that honours the contact data, but also takes full account of orientation data which are both above and below the geologic contacts.

Advantages of the Potential Field Interpolation Method

This solution is ideal for the case of layered geology. A series of surfaces—each at different iso-values of the interpolated scalar field—can be derived from the one interpolator. We refer to these layered strata as belonging to a geologic *series*[1]. If the rock relationships do support the premise of a shared geological history, then combining layers together into a single series has the big advantage that data from one horizon can influence the shape of other nearby horizons, and vice versa. The interpolator—being constructed from additional, relevant geological observations—is therefore an improved predictor of the shape of all geological boundaries in the series.

Measurements of strike and dip recorded *anywhere within the series* will all be taken into account at the point of their measurement. Whilst specific iso-surfaces, representing geologic boundaries, do not necessarily pass through any orientation data, nevertheless, all of the orientation data do exert an influence on the local attitude of those (nearby) iso-surfaces.

A potential field ensures smooth boundaries; the method provides a surface which is sufficiently curved to fit to the data, but has no more curvature than required. If the geologic structure is known to be complex, then the sampling of the geology must be high-frequency; the interpolated surface will honour the high-spatial-frequency signal content.

A potential also ensures no self-crossing. The premise for layered geology being combined into a single series is that the strata have all shared a common geological history; this precludes the possibility of significant erosional breaks and unconformities within a series. In the simplest case, each geological boundary within a series would represent a time-line, which cannot cross other time lines. The potential interpolator ensures this.

Finally, the physics and mathematics of potentials are well understood. The mathematical form of a potential is an implicit function; it can be expressed in the form f(x, y, z) = 0. The potential function allows us to immediately know 'which formation' is present at any arbitrary point *p* in 3D space. This is achieved simply by computing the value of the potential at the point *p,* and comparing it to the iso-values representing the various geologic interfaces. With reference to Figure 3:

- Let $v_1, v_i \ldots v_n$ be increasing values of potential corresponding to the different iso-values of *n* geologic interfaces in a series, being the 'tops' of geologic formations $f_1, \ldots, f_n$. Assume also that there is a cover formation f_{n+1}

[1]The term *series* is used here with a conventional English meaning viz. a 'sequence' or 'set' of geologic layered strata; it should not be confused with the chronostratigraphic usage of the term.

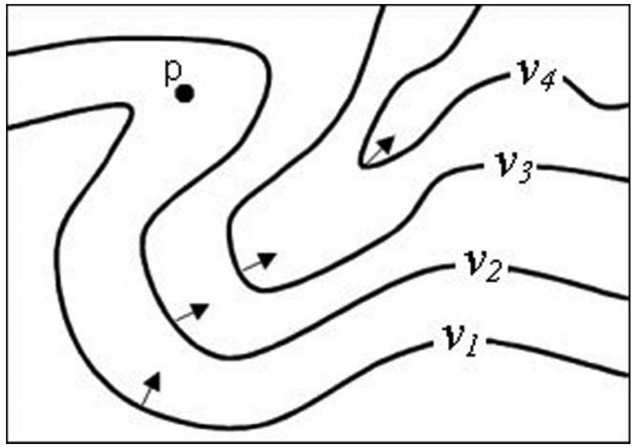


Figure 3. Determining the formation by using the interpolated potential field. For the point *p*, the potential has a value *V(p)*; by comparing this value with the iso-potential values used to model individual geologic formations, the geology at *p* is determined. In this case, $v_1 < V(p) < v_2 \ldots$ so *p* must be in formation f_2 (iso-potential v_2 represents the top of formation f_2).

- Let *V(p)* be the value of the potential at some point *p*

Then:

- If $V(p) <= v_1$ then p is in formation f_1
- If $v_i < V(p) <= v_{i+1}$ then p is in formation f_{i+1}
- If $v_n < V(p)$ then p is in formation f_{n+1}

Advantage of Cokriging in the Interpolator

The interpolation uses cokriging, which is the best unbiased linear estimator, and provides a means of dealing with error in geoscience data. Error may be simple observational or spatial errors, but the term error must also be considered in the context of geological *signal* and *noise* … and these are typically scale-dependent. When mapping, a geologist must make a decision about the mapping-scale; a 1:250,000-scale map is very different from the 1:25,000 scale maps over the same area. With 3D geological modeling, the same decision must be made; essentially the process is one of defining the relationship between the dimension of a project, and the 'wavelength' of geological structures to be modelled. In our 3D modelling, this decision is quantified through the setting of the cokriging parameters.

Thus, for detailed *mapping* a geologist would include data which define the geology in detail, and draw interpretive boundaries showing the geological complexity. The same boundary and orientation data would be included in a 3D *modelling* project, and one would expect to produce a complex model which accurately honoured all available data.

The purpose of a regional map, however, is to generate an overview of the geology at the broader scale. Thus, the mapping geologist might 'average' the data, and draw a simplified interpretive contact. In 3D modelling the same outcome is achieved. When an area of detailed data are included in a regional modelling project, the interpolator (using the default cokriging parameters) *broadly* honours an *averaged* value of the observed data, and may *not* accurately honour the individual data points. The method can also provide an estimate of the error or uncertainty at all points. (Chilès *et al*., 2004)

Modelling Complex Geology

The preceding discussion considered the case of simple, layered geology, and proposed the advantage of being able to model several horizons with a single potential. For the case where the geological history is more complex, and geologic horizons are not sub-parallel, separate potential interpolators must be used - one for each series of strata. For this case it is necessary to define the stratigraphic column, which records the chronological order of the strata, and also the series relationships (either 'onlap' or 'erode'). Where two geologic surfaces from different potential interpolators intersect, an 'erode' surface cuts across any stratigraphically older horizons, whereas an 'onlap' surface would 'stop' against the older surface (Figure 4). This coded information in the stratigraphic column is sufficient to ensure that a *unique* geological model is constructed from several overlapping potentials.

It is worth also noting that, from a topological viewpoint, the cross-cutting relationships of an eroded contact are no different from the cross-cutting nature of an intrusive contact; thus the 'erode' case is also used to model an intrusive.

Faults

Faults are taken into account by (a) defining the location of the fault surface, and the limits of the fault's region of influence, and then (b) introducing discontinuous drift functions into the cokriging equations. The method, documented more fully in Chilès *et al*. (2004), is based on the work of Maréchal (1984), who used drift functions to model faults in 2D seismic data.

For each fault these discontinuous spatial functions model the shape of the influence of the fault. For a finite fault, with limits to the region of influence of the fault defined, the function has a value 0 on one side of the fault and decreases from 1 to 0 on the other side, scaled according to distance from the fault and distance from the edge-extents of the fault (Figure 5); for this case the relative displacement on the fault gradually decreases towards the edges (Figure 6c). Where no limits are defined, the shape of the function is a simple step (an infinite fault, Figure 5).

The fault surface itself is modelled in the same manner as any geologic interface; one or more data points define the location of the fault, and one or more orientation data define its attitude; the fault surface is then modelled using a potential interpolator which is constructed from these data.

In the modelling of faults there are two further practical details:

- A fault is typically restricted to affect only specified parts of the stratigraphic succession. This information is recorded in a table, which links faults with (geologic) series.
- It is possible to have a fault stopped by some other fault. Such a *network of faults* is also managed in a

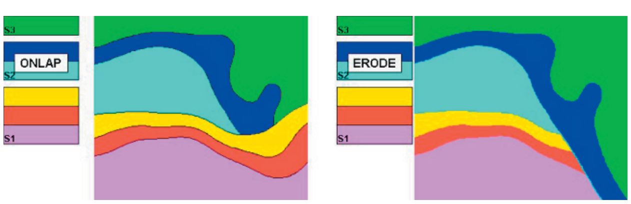


Figure 4. A 2D view of a geologic map or section consisting of three different *series* of geologic formations. Three interpolators (one for each series) can produce a unique geologic model only with reference to the model's stratigraphic column, which records the chronological order of formations and series, and the relationships between the series. On the left, series S2 'onlaps', and stops against the older S1 series. For the 'erode' case (right) the series S2 cuts across older formations.

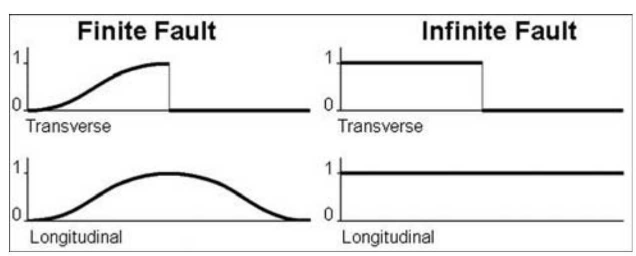


Figure 5. Profiles of the drift functions used in cokriging, in order to model faults. For a fault of finite extent, it is necessary to define limits to the region of influence for the fault; the drift function steps from 0 to 1 as it crosses the fault (*transverse*), but tapers back to 0 at the limit, some distance from the fault. Similarly, in the *longitudinal* direction (*along* the fault) the drift function approaches 0 towards the fault limits. In the simplest case there are no limits to the extent of a fault; an infinite fault.

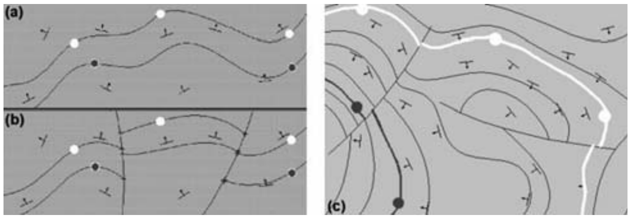


Figure 6. Examples of how faults can be modelled. *Contacts* for two formations belonging to a single series, and also *structural data*, are shown in (a) and (b). These data are modelled with no faults in (a), and with two faults added in (b); note that—on the basis of the two contacts being in a single series—the interpolator can reasonably predict a position for both contacts within the central fault block, despite the limited data available. Faults of finite extent are modelled in (c); the relative displacement decreases towards the edges.

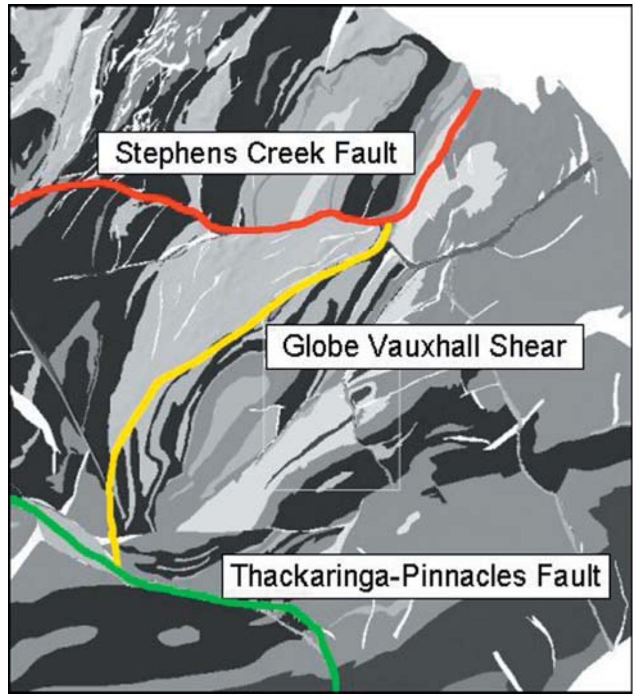


Figure 7. The geology of Broken Hill, showing three of the major shears in the district. The Globe Vauxhall Shear terminates against major faults to the north and south. This *network of faults* is defined in a table which shows the relationships between each pair of faults in a project.

table, in which the relationship between every pair of faults in the model is specified (Figure 7).

Folds

The potential field method does not require any special treatment of folds. Fold structures, however, commonly are regular and predictable shapes, and it is useful to exploit any aspect of geology which can assist the process of interpolation. The structure of folds is used as follows:

- A fold axial surface is defined (Figure 8). As for any geologic interface, one or more data points define the location of the fold axial surface, and one or more orientation data define its attitude; the axial surface is then modelled using a potential interpolator which is constructed from these data.
- A section is constructed along this axial surface.
- A hinge line can be defined on this axial surface section view. By definition, a *hinge line* is the intersection between a (folded) geologic horizon and the fold's axial surface.
- The shape of the fold is also recorded; anticline or syncline, and additional parameters including the inter-limb angle. On the basis of these parameters, additional orientation vectors are constructed which define the shape of the fold, and must be taken into account when the model is re-computed from the data.

BUILDING THE BROKEN HILL 3D MODEL

Scope of the Broken Hill 3D Modelling Project

The Broken Hill 3D Geological Modelling Project was designed as a demonstration of a new technological approach to geological modelling, to be completed within a six-month project life. The model covers an area of 20km x 20km centred on the Broken Hill mining district. It is a regional scale model, developed using the *group* level stratigraphic classification for the district, as defined by mapping by the Geological Survey of New South Wales (NSW). Detailed mine-scale stratigraphic sub-divisions were not incorporated into the model. Even at regional group-level scale the geological structure is complex, however, and this complex geology was captured into a coherent, fully 3D model during the short time of the project.

The model was developed using existing data from

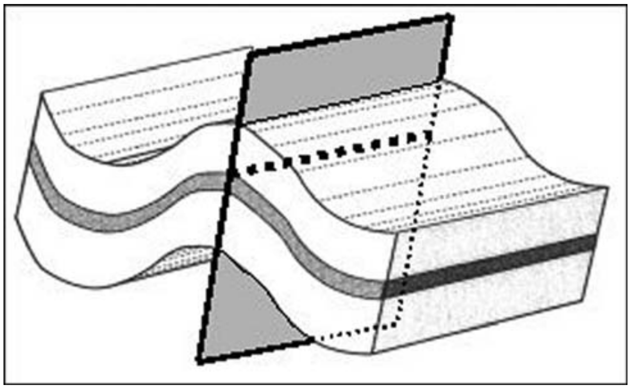


Figure 8. The regularity of fold structures can assist with interpolation. This diagram shows the fold axial surface, and the hinge line where a geologic formation intersects the axial surface. The hinge line, together with other parameters describing the shape of the fold structure, are then used in re-computing the model.

government and industry sources (see below). No additional mapping was done. Nevertheless, the model is an interpretation of the Broken Hill geology by the authors, since the process of GeoModeller model-building—working in three dimensions as it does—requires the user to *interpret the data* being drawn together from different data sources in order to create a coherent 3D model of the geology.

Inputs to the 3D Model

The principal inputs (Figure 9) to the Broken Hill model were:

- the stratigraphy and mapping of the Geological Survey of NSW (mainly Willis, 1989),
- interpreted regional geologic cross-sections (author T. Lees: interpreted from published mapping and personal field-mapping and mine drill-hole logging by author Lees, under the auspices of the Predictive Mineral Discovery CRC's C1 research project),

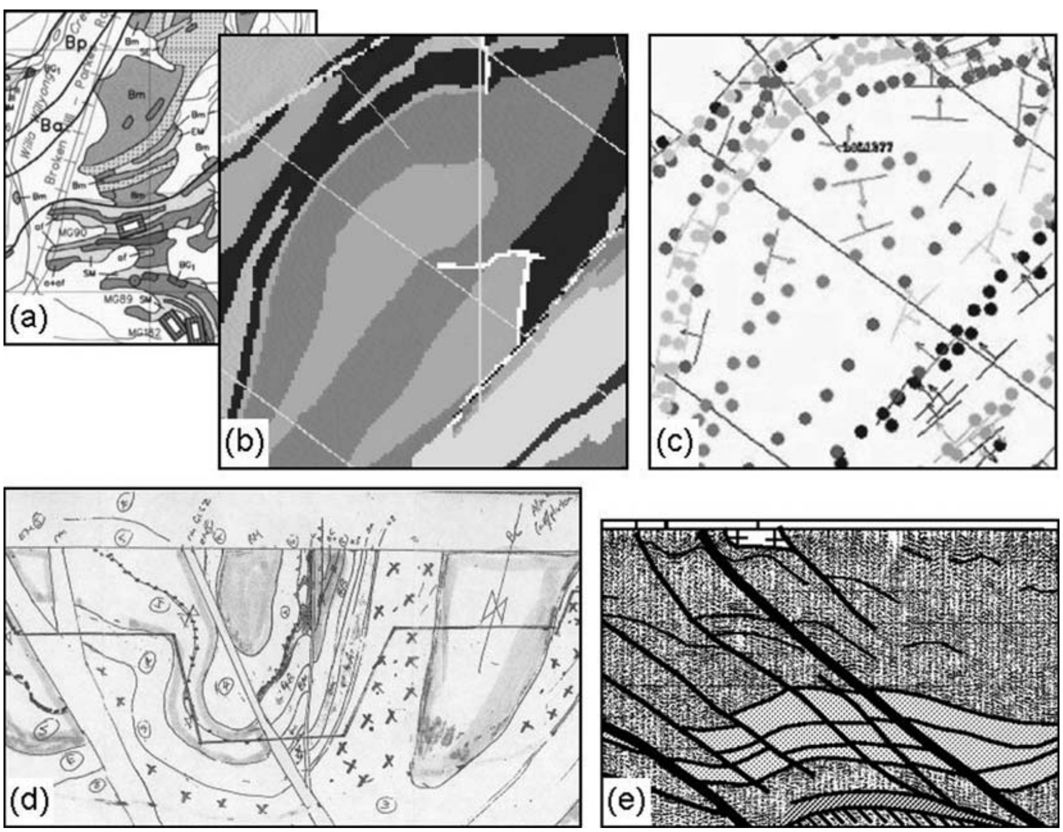


Figure 9. Inputs to the 3D model included published geology at various scales (a) from which geologic observations were digitised at the 'group' level of the stratigraphy (b, c). Digitised contact points are shown as spots in (c); dip and strike data, also derived from published geology maps, are shown with a strike-line and facing vector symbol. Input was also taken from the geologist's interpretive regional cross-sections (d) and seismic interpretation (e).

- interpretation from the Pasminco-Fractal study (Archibald *et al*., 2000, and Mason *et al*., 2003),
- unpublished data from the mine line-of-lode,
- the interpreted Geoscience Australia seismic profile (Gibson *et al*., 1998), and
- geological syntheses of the area (Stevens, 1980; Noble, 2000, Gibson and Nutman, 2004).

Given the planned scope of the modelling project—to produce a *regional* scale model - *selected* data were digitised from the regional maps and interpreted sections.

Sampling Geology

Drawing a geological map or section is a process of *interpolation*, attempting to predict from *sampled* observations (e.g., field mapping, or logging drill-core) where some geological contact is expected to occur. In 3D model building, as in any field mapping exercise, the ability to predict or interpolate is wholly dependent on the quality and *frequency* of the sampling of the geology. In our experience the best result is achieved by a combination of just-enough points to define the geological boundary position, together with strategically located orientation data to guide the orientation of the geological surface that will be fitted through the observations. As the geology becomes more complex, more points are needed to define the geological structure; in other words, the sampling of the geology must be done at a closer sample spacing.

Building the 3D Model

The building of a 3D model is partly a process of ‘sampling the geology’ as discussed above … but almost always it *also requires an interpretive process by the geologist*. This continual need to be ‘interpreting the geology’ is significant. There is no expectation that some computer software will successfully and automatically ‘build a model’! The reality is that interpretive input from a skilled geologist is essential to build a model; the software is simply a tool to facilitate the model-building process.

The interpretive process is encapsulated in the *input—compute—plot - review* cycle described below. Having defined the stratigraphic pile for the project, and also the faults, the basic process of creating the Broken Hill geological model was an iterative cycle of:

- Input: In the map-view, or any of the section-views, digitise points at intervals along a geologic contact (thus capturing *geologic contact data*). Likewise selected *orientation data* may be input. (Note that there are options in GeoModeller for *importing* data from digital sources; this was not done for the Broken Hill study),
- Compute the model,
- Plot the modelled geology on the map or a section: Sections can be generated anywhere in the project area, and the model plotted to assist the geologist’s assessment of the model, and
- Review: The geologist must review the model, and compare the model against known data—or against his/her expectations.

This cycle—*compute the model, and then review*—tests the model against the geologist’s expectations, and is essentially an interpretive process. If the model contradicts some known data, then the geologist must add those additional observations, in order to take them into account when the model is recomputed. Frequently, however, the geologist does not have additional data, but does have an understanding of the geology, which is a valid basis for proposing that the current model cannot be correct, and needs to be adjusted. The geologist imposes his/her interpretation on the model simply by adding (hypothesised) contact data or orientation data. When the model is recomputed and replotted in the section-view where the geologist has proposed this interpretation, the geologist can again review the model, and can observe how the shape of the model has been adjusted as a consequence of his/her interpretation. The geologist can also review the implications of this revised model in any other section view. Note that the geologist can test different ideas about the geological structure of the project area, and so can evaluate alternative interpretations.

It is significant that by far the most ‘geologist time’ spent on the Broken Hill project was spent doing this cycle of ‘input-compute-draw-review’ … with the geologist continually working *as a geologist*, trying to fathom the complexity of Broken Hill geology in three dimensions, and continually adding further ‘observations’ to the GeoModeller model; these observations were either additional samples from original maps and interpretive sections, or the geologist’s hypotheses based on his/her evolving interpretation of that complex 3D geology.

THE BROKEN HILL MODEL—OUTPUTS AND INVERSION

The Broken Hill model produced in this project was developed from the inputs described above, *as interpreted by the authors* (principally T. Lees). The building of the model in three dimensions raised questions about earlier interpretations presented in various generations of published maps; the need to honour all the data inputs but also achieve a 3D integrity meant that several revisions of the regional geology were proposed during this model-building interpretive process. Notable revisions proposed by T. Lees are in the area along the eastern side of the Broken Hill Synform and the western edge of the Sundown Group.

It is worth noting that a model in this software is not a set of shapes or surfaces, but rather a mathematical function in three dimensions. By interrogating this model-equation in various ways, a variety of visualisation outputs can be generated. Thus the model can be presented in full three-dimensional form (Figure 10), but it is also easily presented as 2D views. This flexibility is important. Building a model in 3D can expose the flaws of a simple 2D interpretation; it forces the interpreter to develop a more robust and coherent understanding of the geology. In practical terms, however, the actual process of working with the developing 3D model is often best achieved through a series of conventional—and simpler—2D views of the model (the map, and sections). Certainly all of the interpretive input in this project was done in 2D views—but then reviewed in various other 2D and 3D views.

Outputs from the 3D Model

The (mathematical) model of Broken Hill was used to generate several outputs:

- Maps and Sections. Any surface that intersects the model-space is a *section*. The DTM (Digital Terrain Model, a topographic surface) is thus a section, and is used to create a conventional geologic map. Any arbitrary section can also be created, allowing the 3D model to be examined in any 2D view. Maps and sections were used throughout the (interpretive) model building phase, and section-plots were a standard output,
- 3D Views: the software has a 3D viewer, within which the geology data points, the orientation data (displayed as small ‘discs’), and the 3D shapes of geological formations (see below) can be visualised from any angle,
- 3D Shapes: shapes for each geological formation, defined by triangulated surfaces, were generated and exported in T-Surf[2] file format, suitable for import and visualisation in Gocad, FracSIS, etc.,
- VRML files: for 3D visualisation using a VRML plug-in to a web browser, and
- Voxels: a 3D voxel model, with geology assigned to voxels, was generated and exported in Voxet[2] format, suitable for import and visualisation in Gocad.

A final component of the Broken Hill project was to demonstrate the application of gravity inversion to further refine and test the accuracy of the model. The voxel model was an important input to the inversion processing.

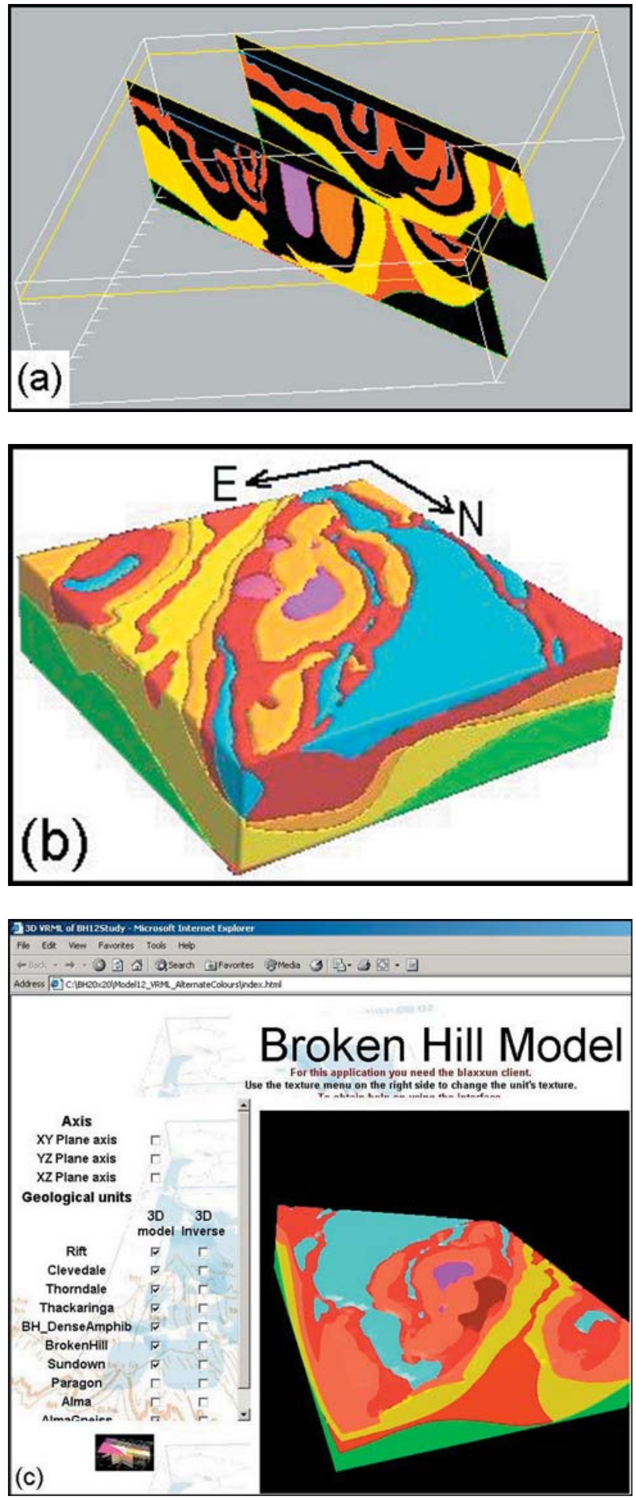


Figure 10. Outputs of the Broken Hill geological model. Conventional maps and sections can be drawn, and can be presented in perspective views (a). Full 3D models can be constructed and visualised, or exported in standard exchange format files suitable for import to other packages such as Gocad (b). The 3D shapes can also be used to create VRML files, suitable for viewing in a standard web-browser (c).

[2]T-Surf is an ASCII exchange file format, defined by Gocad software (http://www.gocad.com/), which describes surfaces and closed volumes in terms of the 3D coordinates of the vertices of triangles fitted to the surface. The Voxet format is also defined by Gocad.

Gravity Inversion of the Broken Hill Model

The purpose of generating realistic 3D geological models is often to provide a basis for further physical modelling and analysis. This might include investigation of ground-water characteristics, seismic hazard assessment, or thermal energy resource potential.

In Australia there are vast areas with little or no outcrop, and so the geology is often poorly understood. At the same time, these same areas often have good gravity coverage, and high quality magnetics coverage. Thus there is a strong interest in maximising the utilisation of these potential field data to improve geological understanding. Inversion of potential field data is often flawed by not having adequate models with which to begin the interpretation. Thus there is an interest to use an approach of (a) generating realistic models from all available sources of geological information (often not much!), and then (b) to use these models as a starting point for potential field inversion.

It is not the purpose of this paper to discuss inversion, but a summary is included here since the Broken Hill model was used to demonstrate an innovative approach to inversion which has been implemented in the GeoModeller software. For a more thorough treatment, see Guillen *et al*. (2004).

The inversion uses as a starting point what is expected to be a *realistic model of the geology*. On this basis there is an expectation that the misfit between the computed (gravity) forward model response and the field data will decrease relatively quickly, yielding a set of (inversion) models for which the computed geophysical response reasonably matches the field data. Practical comparisons can be made between a 'realistic' starting geologic model and the progressively revised voxel models generated by the inversion.

Inversion is performed on a voxel model of the *geology* rather than a model of some physical property, such as density. The *geologic unit* for each voxel is initially assigned from the starting model built by the project geologist; this may change during inversion. Physical property values are assigned to voxels using the parameters and statistical law which describe the distribution of that property for the given geologic unit.

Each inversion iteration makes a modification to one voxel only, or, optionally, to a small selection of voxels. The revised geophysical response due to each small adjustment of the model is computed very efficiently, and naturally the overall impact from a single iteration is small.

The inversion process is based on a Markov Chain Monte Carlo formulation, which is solely used to accept/reject each candidate model. The single voxel to be adjusted in each iteration is selected randomly. The assigned geologic unit for the selected voxel may be changed to match that of an adjacent voxel—on a random basis. The assignment of a density difference value is by random selection according to the probability function defined for the relevant geologic unit.

Whereas many inversion processes are designed to reduce the global misfit between the observed and the computed response, and then stop when the misfit has reached some specified low limit … the GeoModeller inversion continues to iterate. Rather than simply finding *one model* which matches the observed data, the approach is to explore a wide range of possible models—all of which have a computed response which have a known likelihood based on how well it matches the observed data; thus, potentially *<u>many millions of possible models are examined</u>*, and the inversion results are presented in terms of the probabilities … for example, the probability that a voxel *<u>v</u>* is stratigraphic unit *<u>s</u>*.

The inversion algorithm may be summarised as follows:

- for each inversion iteration, it randomly selects a voxel,
- it optionally changes the geologic unit and/or assigns a revised density value,
- it re-computes the model response,
- it compares the model with the field gravity data,
- if the misfit improves, the revised model is retained, and
- if the misfit is worse, the revised model *may* be kept or rejected (see below).

The last point—viz. keeping a model even though the misfit is worse—is designed to allow the inversion to move beyond local minima, and look for further solutions that might improve the fit. With some millions of iterations, the global misfit typically decreases to some small error between the computed model response and the field data. By *continuing the inversion* for many more millions of iterations beyond this point, the models that are 'kept' are all *models which reasonably match the gravity data* … and these many millions of models can be used to report the inversion outcome in terms of probability.

The inversion trials for Broken Hill were inconclusive. Early inversions yielded poor results and required some revision of the model, and reassessment of the true density value of some formations. *<u>All later inversions achieved a good match between the model gravity response and the field data</u>*. These results must be qualified, however, by the reality that density values for the Broken Hill formations are not well known. For some formations the distribution of density values is bimodal due to local variations in the percentage of either dense amphibolite or less density quartz pegmatites.

An outcome from this work has been that we have

recognised a need for inversion processing to be able to effectively manage these bimodal distributions of density, and have initiated experimental studies to implement an inversion option that allows for this.

3-D MODELLING AND DATA MANAGEMENT

The 3D modelling application presented in this paper is fundamentally designed as an interpretive tool to be applied by the project geologist. The software has been applied to a spectrum of tasks on a range of scales, and a small range of digital data input/output capabilities have been developed. We have a clear vision that this type of software must seamlessly integrate with an organisation's geologic data management to do the following:

- read geologic observations from databases,
- write back attributed data to those databases, and
- export lines and surfaces into the databases of GIS and presentation software.

Some of these data I/O requirements already exist in GeoModeller, and more are planned. It is worth commenting further about the input of digital data for 3D model building. In our experience to date, building 3D models needs an intelligent approach to selecting the data to be used. Simply importing *all available data* is often unsatisfactory. There are a variety of reasons for this. There can be quite trivial reasons, such as incompatibility between the stratigraphic nomenclature in the database compared to the modelling project, or the database may contain many micro-structural observations that are not immediately applicable for a regional modelling project. It is worth noting two other points:

1. When constructing a 3D model, it is often the case that there is little or no actual data in the third dimension. As with any uneven sampling problem, an abundance of data in one area cannot compensate for a lack of data in another, and it is—in our experience—unrealistic to generate 'high-frequency' models in zones of sparse data! In an area of good outcrop it may be possible to generate a high resolution map, but not necessarily a high-detail model beneath that. Thus it is not always possible to effectively use all of the mapping observations that are available; and we will be seeking to develop filters to assist the geologist in filtering the data, to select some, and reject others.
2. The GeoModeller software uses discrete *points* of geology. Some geologic databases record these, but there are also now vast repositories of GIS-geologic data recorded as *lines* of data. In many cases, a line in a GIS database is a combination of observation points and interpretation lines. In the GeoModeller software we would like to use the point (observation), but let the software (re)generate the line! To make best use of existing GIS data, we plan to develop tools to intelligently re-sample lines, and again give the geologist a filtering capability such that choices can be made about keeping or rejecting portions of these imported data.

3-D MODELLING AND DATA QUERYING/ PRESENTATION

In the future, we will provide some fundamental data presentation capabilities in the GeoModeller software. Already there is simple screen visualisation, and a capacity to produce a 3D VRML file—with little more than a click of a button! And presentation-quality printing of maps and sections is proposed. However, we see model-building as a process which must be integrated with other styles of data manipulation, querying and presentation, and so the export of standard interchange formats is a high priority. Several export formats are supported already, and more are planned.

ACKNOWLEDGMENTS

The 3D GeoModeller software was developed by BRGM, and is being commercialised and further developed by Intrepid Geophysics in collaboration with the BRGM. The Broken Hill geological modelling work was funded by an Innovation Access Programme grant under the Australian Government's innovation statement, Backing Australia's Ability, together with support from three sponsors: Geoscience Australia, the NSW Department of Mineral Resources and the Predictive Mineral Discovery CRC. Their support is gratefully acknowledged. The ground gravity data used in the inversion were obtained from the National Gravity Database maintained by Geoscience Australia.

REFERENCES

Archibald, N. J., Holden, D., Mason, R., and Green, T., 2000, A 3D Geological Model of the Broken Hill 'Line of Lode' and Regional area: Unpublished report, Pasminco Exploration.

Chilès, J.P., Aug, C., Guillen, A., and Lees, T., 2004, Modelling the Geometry of Geological Units and its Uncertainty in 3D From Structural Data: The Potential-Field Method: Workshop Proceedings: Orebody Modelling and Strategic Mine Planning: Uncertainty and Risk Management, Perth, WA, 22-24 November 2004; organised by the WH Bryan Mining Geology Research Centre (BRC), University of Queensland, Brisbane (http://www.minmet.uq.edu.au/~bryan/).

Gibson, G., Drummond, B., Fomin, T., Owen, A., Maidment, D., Gibson, D., Peljo, M., and Wake-Dyster, K., 1998, Re-evaluation of Crustal Structure of the Broken Hill Inlier through Structural Mapping and Seismic Profiling: AGSO Record 1998/11.

Gibson, G. M., and Nutman A. P., 2004, Detachment faulting and bimodal magmatism in the Palaeoproterozoic Willyama Supergroup, south-central Australia: keys to recognition of a multiply deformed Precambrian metamorphic core complex: J. Geol. Soc. London, 161, p. 55-66.

Guillen, A., Courrioux, G., Calcagno, P., Lane, R., Lees, T., and McInerney, P., 2004, Constrained gravity 3D litho-inversion applied to Broken Hill: *in* Conference Abstracts, ASEG 17th Geophysical Conference and Exhibition, Sydney.

Lajaunie, C., Courrioux, G., and Manual, L., 1997, Foliation fields and 3D cartography in geology: principles of a method based on potential interpolation: Mathematical Geology, v. 29, no. 4, p. 571-584.

Maréchal, A., 1984, Kriging seismic data in presence of faults, *in* Verly, G., David, M., Journel, A.G. and Maréchal, A., *eds*., Geostatistics for Natural Resources Characterization, Part 1: Reidel, Dordrecht, p. 271-294.

Mason, R.A., Archibald, N.J., Green, T.J., Broome, J.N.M., and Haydon, R.C., 2003, Resolution of the structural geometry of the Line of Lode and the Broken Hill Block: A 3D model, *in*: Peljo M., comp., Broken Hill Exploration Initiative: *Geoscience Australia Record* 2003/13, abstracts from the July 2003 conference.

Noble M. P., 2000, The Geology of the Broken Hill Synform, N.S.W. Australia: MSc Thesis, Monash University, Melbourne.

Stevens, B.P.J., ed., 1980, A guide to the stratigraphy and mineralisation of the Broken Hill Block: New South Wales Geological Survey Records 20(1), 153 p.

Willis, I. L., 1989, Broken Hill Stratigraphic Map: New South Wales Geological Survey, Sydney.

Winchester, S., 2001, The Map That Changed the World: William Smith and the Birth of Modern Geology: HarperCollins Publishers, New York, 329 p.

The Digital Geologic Map of Afghanistan

By Ronald R. Wahl and Robert G. Bohannon

U.S. Geological Survey
P.O. Box 25046, MS980, Denver Federal Center
Denver, CO 80225
Telephone: (303) 236-1320
Fax: (303) 236-5349
e-mail: rwahl@usgs.gov, bbohannon@usgs.gov

ABSTRACT

The 1:500,000-scale Russian-published geological map of Afghanistan (Abdullah and Chmyriov, 1977) was digitized and attributed with ArcGIS. Topology was created and maintained in ArcGIS and then the spatial data were converted to ESRI shapefiles. A lookup table for the map units was created from the Arc GIS Workstation coverages and then manually edited. The edited result is a one-table database that is the basis for drawing the map units for geologic maps as large as 1:250,000-scale, 1-degree x 2-degree sheets. The shape files were imported into Adobe Illustrator with MAPublisher from Avenza Systems, Inc. Custom graphic styles were linked to a map unit in the joined spatial attributes and the lookup table through MAPublisher stylesheets. Faults and contacts were assigned graphic styles separately. The map units, faults and contacts were combined to produce the geologic map. The current database is being reverse-engineered into a modern relational database. This final database will be useful in making mineral assessments, oil and gas assessments, hydrogeologic studies and as base information for road construction and environmental restoration.

INTRODUCTION

Project Framework

The U.S. Geological Survey (USGS) works in several science activities in Afghanistan in support of the Afghanistan Reconstruction Project (ARP). This project is carried out under the auspices of the United Nations (UN) with the United States as the major contributor. "The ARP carries out research and public education about selected issues related to the rebuilding of Afghanistan's institutions, society, and economy. The project supports efforts by the Afghan government, Afghan civil society, the United Nations Assistance Mission in Afghanistan, and donors to carry out a more effective reconstruction mission." (CIC, 2005).

The USGS Project

The U.S. Department of State's Agency for International Development (AID) and the Trade and Development Agency (TDA) facilitate USGS involvement in the ARP. AID funds geologic mapping and other USGS earth science efforts. TDA primarily funds energy resource studies.

Geographic and Geologic Setting

Afghanistan is in Central Asia. On the north, its neighbors are Turkmenistan, Uzbekistan, and Tajikistan. On the east the largest part of the border is with Pakistan. China borders Afghanistan on the east end of the narrow arm (the Wakhan Corridor), extending to the east from the northern part of the country. Pakistan wraps around eastern and southern borders of Afghanistan. The country's western border is shared with Iran.

The country ranges from low desert in the northwest and south, to high mountains and deep valleys in the central and eastern sections of the country. The climate is arid. The main source of moisture is melt water from ice and snow in the mountains that reaches the lowlands in rivers and groundwater.

Afghanistan is at the western edge of the Himalayas between the Indian/Asian and Arabian/Asian collision zones. Rocks exposed at the surface range in age from Archean to Quaternary. Alluvial and fluvial deposits cover large portions of the west and south. Eolian sand forms prominent dune fields in the south and northwest.

History of Regional Geologic Mapping

The nations mentioned above have had recent geologic maps compiled and interpreted in light of modern thought concerning plate tectonic theory, particularly as concerns the collision of the many plates that make up the Afghanistan of today. Afghanistan, however, has had little benefit from recent geologic studies because of persistent and extensive warfare and unrest throughout

the country. The general consensus among project participants is that geologic investigations and analysis in Afghanistan are ten to fifteen years behind those conducted in neighboring countries.

Numerous efforts have been made to map the geology of Afghanistan. The most detailed countrywide map is the Russian 1:500,000-scale geologic map (Abdullah and Chmyriov, 1977). German geologists (Wittekindt and Weppert, 1973) made a map of the south and central parts of Afghanistan, which USGS participants in Flagstaff, AZ have digitized and attributed. A series of fifteen 1:100,000-scale maps by French geologists has been scanned but not yet digitized. This paper reports only on our work with the Russian geologic map.

THE CURRENT STATUS

We started with digitized geology: the attributed vectors and polygons for the 1:500,000-scale Russian geologic map. The attribute information was summarized into a "lookup" table using the FREQUENCY option under ANALYSIS and STATISTICS in ArcGIS Toolbox. Attribute items were added to this table for geologic time, lithology, map unit labels, and map unit descriptions. Some map units are rendered by patterned polygons on the original map. For GIS purposes, we identified these as separate map units. When we had sufficient information to proceed, we identified large lithologic groupings or composition changes by geographic region or tectonic province.

In addition to the geologic map data, we have acquired published and unpublished data for:

- Georeferenced images of the fifteen pieces of the original Russian geologic map
- Mineral locations
- Plutonic rock composition database
- Further explanations of the geologic map data
- Shuttle Radar Topography Mission (SRTM) digital elevation model—90 meter resolution
- UN AIMS (2005) data for cities, roads, provincial and international boundaries, and spelling for place names in Afghanistan
- Some scanned Russian topographic mapping
- LandSat 7 TM imagery, and Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) imagery.

Standards Issues

We followed where appropriate the relevant standards for geologic time, labeling of map units, line styles for boundaries between map units (contacts and faults), polygon fill colors, and patterns with color. As this map was compiled by the Russians (Abdullah and Chmyriov, 1977), we decided to use the geologic time scale from the International Commission on Stratigraphy (ICS, http://www.stratigraphy.org). The version of the ICS stratigraphic time scale chosen was that colored and organized according to colors traditionally used by the USGS (ICS, 2004).

The draft U.S. geologic line and symbol standards (USGS, 1999) were used, with the exception that the typical Russian symbol for foliation or layering in igneous rocks was used for the location of a strike and dip direction identified from aerial photography.

Because the map contains more map units than can be represented by the colors in the ICS chart, we chose a range of colors related to each major time period represented on the ICS chart. In addition, we chose colors and patterns for the plutonic rocks more for contrast with the group "layered" rocks than to designate age. For volcanic and some sedimentary rocks, we chose to add patterning so that the number of color choices would be held to a minimum. We deviated from the ICS standard by using "Quaternary" in the traditional sense, in order to remain consistent with the Russian source map. We show map unit labels with a custom font that has many subscripts to indicate epochs and stages.

Challenges

The Russian geologic map (Abdullah and Chmyriov, 1977) was produced on fifteen sheets, each 3 degrees in longitude by 2 degrees in latitude. Our copy of the original map is a set of scans of these fifteen sheets. Nowhere can we find information about the map projection scheme or any related information. The scans were georeferenced to WGS84 geographic coordinates and transformed to the project-adopted Transverse Mercator projection with the same datum and spheroid. From careful checking with data sets from AIMS (2005), we're confident that the map is registered to within national map-accuracy standards of 250 meters (at the map scale of 1:500,000).

Another simple but important problem was the scale at which to render the map. We chose a map scale of 1:850, 000 because the map fits on 54"-wide plotter paper. This results in a substantial savings when printing the map on large-format Hewlett-Packard plotters.

The Construction of the Geologic Map

We decided to use Adobe Illustrator with Avenza's MAPublisher to render the finished map. We recognize that a complete ArcGIS geodatabase would allow more and varied analyses to be done with the digital map; however, the Afghan Geological Survey (AGS) has few modern computer facilities and fewer personnel to use these computers or complex software like ArcGIS. Paper maps are the preferred medium for conveying geologic

information to those in the country. Secondly, data sharing in this environment, when possible, is done with ESRI shape files and DBF databases. The choices of Illustrator and MAPublisher allow a small GIS capability such as table joins and queries and a sophisticated way of presenting the results of such queries easily on paper maps. This choice has also allowed project participants to easily construct several 1:250,000-scale geologic maps that are almost ready for publication.

The first hurdle we overcame was how to represent map units with color and fill patterns in a consistent and reliable manner. MAPublisher uses a “Map View” to control such things as georeferenced feature layers and placement of map information on a printed page. In addition, MAPublisher uses a “Map Stylesheet” to control the association of Illustrator graphic styles, symbols, and text fonts with map information.

Graphic styles in Illustrator can be created for both polygon fill and complex line representations through the “Graphic Styles palette” in Illustrator. For polygon fills, graphic styles are composed of one stroke layer (for the fill boundary) and (or) one or more fill and pattern layers, and a transparency and mode. This method allows us to control the total appearance of polygon fills on a style-by-style basis. In the Map Stylesheet, one can examine attributes associated with each of the polygons or lines. Each Stylesheet has a column that can be chosen from one of the attribute columns associated with feature types in the Stylesheet, and a column with the graphic styles that are available in the current drawing. If one desires to fill polygons based on chronostratigraphic divisions, an attribute column might contain such designations as “Cambrian, undivided” and “Cambrian limestone” and if one labeled newly created graphic styles with the same names, then all that remains to be done is to associate the graphic style with the appropriate attribute in the Stylesheet; MAPublisher will then fill with the graphic style “Cambrian” all polygons that have the attribute “Cambrian”.

We have imported geologic symbols from the U.S. draft standard (USGS, 1999) into Illustrator, as “Symbols” in a symbols library; these can be associated in a MAPublisher Stylesheet to data in an imported point shapefile. Symbols can be rotated counterclockwise based on a numeric column attribute.

In all of the above cases, we chose to put a minimum of information in the attribute fields of the feature shapefiles, and to keep the repetitive attributes in a look-up table. The lookup tables were exported to DBF format and then changed to Microsoft Excel files. We added to (and deleted from) the attribute columns originally present to arrive at a lookup table that contains all of the attributes by which we might want to query this table. After converting back to DBF format, we used MAPublisher to join the feature shape file attribute data to the lookup tables, in order to connect attributes like map unit name with a geometric object. We make the assumption (and try to edit the data so) that every entry in a feature file has a corresponding value in the look-up table. We then proceed to graphically represent the feature shape file data as described above. Errors in attributing both the polygon or line feature data can be rectified using the MAPublisher facilities to edit attribute data and then have the Stylesheet function refill the polygons with the proper style.

The Russians labeled map units in a complex fashion that to our knowledge cannot be reproduced in any common word processing software. The map unit labeled “N11” on the Russian map would have the first “1” as a superscript and the second “1” as a subscript directly under the superscript. Illustrator has no direct mechanism to allow this construct in a text string without considerable label-by-label work. To keep the map unit symbol as close to the one on the original map, we designed a True Type font with subscripts of needed characters and symbols for Paleogene, Cambrian, and Proterozoic. In international usage, the Triassic period is represented by a “T”. This font works both on a PC and on a Macintosh running OS X, and allows map unit labels with many subscripts to be placed as a text string in Illustrator without manual intervention to make the subscripts. In addition to working in Illustrator, the font also works in Microsoft Word and Excel, and in ArcGIS.

We used the latter application to label the polygons with ArcGIS (as an Illustrator layer), because MAPublisher has no way to easily plot polygon labels that fit entirely within a polygon, as does ArcGIS. We developed a workaround for this MAPublisher limitation,creating a map of the polygon shapefile in ArcGIS at a scale of 1:850,000 (the scale of the final map), labeling the polygons using the custom font (the polygon boundaries of the polygons were turned to light gray with no polygon fill color), and writing the map to a PDF file which was then imported into an Illustrator layer. This layer was then manually aligned with the contact layer, and then all of the polygon borders were erased, leaving the labels.

Some other limitations of working in MAPublisher are:

- No queries involving more than one attribute column directly in the Stylesheet.
- One look-up table per layer.
- No Open DataBase Connectivity (ODBC) capability.
- The current version of MAPublisher will not handle more that 50 graphic styles without manually editing a preferences file with a text editor.
- No multi-column primary keys for the database.
- Clearing of a table join is done by deleting the joined attribute columns one by one.

Some of the limitations of this look-up table are:

- No good way of handling hierarchical data. Every hierarchy must be explicitly entered into this table.

- The number of attribute columns can become quite unwieldy.
- Data integrity is difficult to maintain.
- Numeric data must be entered as real (or floating point) data for column joins to work successfully without manually editing the DBF files to ensure compatibility.

ACCOMPLISHMENTS

As a result of this work, some of the last of the 1:250,000-scale geologic maps to be compiled took only hours instead of days to complete. Editing and reviewing of these maps has become much simpler since color fills, patterns, and line weights, patterns, and colors are standardized for this project. The maps are therefore a much more consistent product even though they were created by different project participants.

Current Products

We have generated from these data sets:

1. A preliminary geodatabase in ArcGIS 9.1: this geodatabase was derived from the ArcGIS Workstation coverage made from the original digitization of the Russian geologic map. It uses the lookup table built in ArcGIS Workstation and modified as described above.
2. Maps including geologic maps clipped from the above database for the creation of 1-degree by 2-degree, 1:250,000-scale geologic maps. The project participants have generated 32 of these geologic maps. They are being published as USGS Open-File Reports in cooperation with the Afghanistan Geological Survey. SRTM data was used for a shaded relief background on each of these maps.
3. Clipped and corrected LandSat 7 TM data for each 2-degree quadrangle. Project participants at the USGS office in Flagstaff, AZ office have provided LandSat 7 TM data that has been corrected for the atmosphere and vegetation.
4. Geologic maps combined with geophysical mapping (mostly seismic survey data) for a preliminary petroleum analysis of the Ama Darya basin in northwestern Afghanistan.
5. A national geologic map of Afghanistan at a scale of 1:850,000, for which this work was primarily done.
6. A national geologic and minerals location map produced in cooperation with Jeff Doebrich (USGS, Reston) and Craig Wandrey (USGS, Denver) (Doebrich and Wahl, in press), which uses the map developed in (5) above.

WORK TO BE DONE

The remaining work on the look-up table entails three steps. First we will ensure that every attribute we could want to use in a query is in the table. Second, Wahl intends to reverse-engineer these look-up tables into a relational database that can, with a simple query, re-create the look-up table and perhaps other “views” of the data, but still have the reliability and data integrity of a true database. Third, we want to expand the database to allow for larger scale (1:100,000-scale) geologic mapping in phase two of this project. Most of the geologic data will be collected from the analysis of Landsat 7 TM and ASTER imagery, because the cost and effort needed to conduct field work in Afghanistan now is extreme.

REFERENCES

Afghanistan Information Management Service (AIMS), 2005, Various GIS data sets for Afghanistan, accessed at http://www.aims.org.af/.

Abdullah, Sh., and Chmyriov, V.M., eds., 1977, Map of mineral resources of Afghanistan: Kabul, Afghanistan, Democratic Republic of Afghanistan, Ministry of Mines and Industries, Department of Geological and Mineral Survey, V/O Technoexport USSR, scale 1:500,000.

Center for International Cooperation (CIC), 2005, Conflict Prevention, Recovery and Peace building, accessed at http://www.cic.nyu.edu/conflict/conflict_project4.html - overview.

Doebrich, Jeff L., and Wahl, Ronald R., in press, Geologic and mineral location map of Afghanistan: U.S. Geological Survey Open-file Report, scale 1:850,000.

International Commission on Stratigraphy (ICS), 2004, International stratigraphic chart (using the color codes according to the United States Geological Survey), accessed at http://www.stratigraphy.org/chus.pdf.

United States Geological Survey (USGS), 1999, Public review draft—Digital cartographic standard for geologic map symbolization (Postscript implementation): U.S. Geological Survey Open-file Report 99-430, available at http://pubs.usgs.gov/of/1999/of99-430/.

Wittekindt, H., and Weppert, D., 1973, Geological map of central and southern Afghanistan: German Geological Mission to Afghanistan, Geological Survey of the Federal Republic of Germany, scale 1:500,000.

NADM-H2O and H2O-GML: Enabling Decision Support by Extending NADM for Groundwater Information Interoperability

By Eric Boisvert[1], Jean Brodeur[2], and Boyan Brodaric[3]

[1]Geological Survey of Canada – Québec Division
Laboratoire de cartographie numérique et de photogrammétrie
Earth Sciences Sector of Natural Resources Canada
490 rue de la Couronne
Québec City, (Québec), G1K 9A9 Canada
Telephone: (418) 654-3705
Fax: (418) 654-2615
e-mail: eboisver@nrcan.gc.ca

[2]Geomatic of Canada – Centre for topographic information – Sherbrooke
Earth Sciences Sector of Natural Resources Canada
2144 King Street West
Sherbrooke, (Québec), J1J 2E8 Canada
Telephone: (819) 564-5600
Fax: (819) 564-5698
e-mail: brodeur@nrcan.gc.ca

[3]Geological Survey of Canada – Central Division
Earth Sciences Sector of Natural Resources Canada
601 Booth Street
Ottawa, (Ontario), K1A 0E8 Canada
Telephone: (613) 992-3562
Fax: (613) 992-9273
e-mail: brodaric@nrcan.gc.ca

ABSTRACT

This paper reports on extensions made to the North American Data Model or NADM-C1 conceptual model (http://nadm-geo.org/) and to the related NADM-C1 GML schema, for the modeling of groundwater concepts related to water quantity assessment. The extensions are described in detail, and a use case is presented to demonstrate their usefulness in delivering groundwater information from the National Groundwater Database of Canada.

BACKGROUND

Nearly 10 million Canadians rely on groundwater for their fresh water supply and yet the extent of the resource is poorly known. Knowledge about groundwater resources is not only key for water supply but also has ramifications for energy production, industry and community development. As this information is required in many types of decision making it is important to improve access to it, and thereby ensure that the Canadian government's goals in sustainable development are met.

The Groundwater program of the Earth Sciences Sector of Natural Resources Canada has funded a series of projects to improve knowledge of key Canadian aquifers and created a specific project to build an infrastructure to improve access to the resultant information. The National Groundwater Database (http://ntserv.gis.nrcan.gc.ca/gwp/ngwd/exploration) project is implementing a series of tools and procedures to connect heterogeneous data and to distribute them to the community. Full connectivity between the data providers and the data users is enabled via partnerships with other projects, such as with the PATHWAYS project (Brodaric et al., 2005, in this volume), which provides mechanisms for transforming groundwater information into forms useful for decision makers.

Non traditional users are also being reached through partnership with other federal and provincial departments, such as Environment Canada and Health Canada, through the RésEau project. RésEau is building a larger infrastructure for water (surface and groundwater) to create a single access point for all water related information in Canada.

IMPORTANCE OF GROUNDWATER INFORMATION INTEROPERABILITY

A large portion of the data used in groundwater projects comes from provincial sources. The bulk of this data comes from water well databases built incrementally from the well logs collected from various sources (well drillers, municipalities, other agencies). Each agency has its own motive for collecting such data. These are either legal, because the agency is legally bound to keep this information, or operational, because it needs the information to support its activities. The databases have different requirements, hence different structures and nomenclature. Furthermore, the databases are not static, since more and more information is being keyed in as new wells are being dug. Centralizing this information into a single national database is not possible because of practical concerns (we simply don't have the resources to keep this information up to date), technical reasons (addressing a large set of requirements within a single information system), and legalities (the data is owned by the provinces). The bottom line is: the data must stay where they are, and structured as they are. The solution to reach those data lies in interoperability technology.

Interoperability amongst data producers and data consumers is realized by us through the implementation of technologies promoted by the CGDI (Canadian Geospatial Data Infrastructure; http://cgdi.gc.ca/CGDI.cfm). To attain the CGDI vision, Geoconnections, a federal government arm of the CGDI, has for the last 5 years supported significant development efforts to implement Open Geospatial Consortium (OGC) standards. The OGC standards are themselves closely related to the ISO TC/211 standards, emphasising that technologies developed from OGC specifications have solid international credentials.

OGC standards only provide an interoperability framework, which must be adapted to domain specific data such as hydrogeological information. Therefore, for this technology to work, the community of users requiring interoperability must go through a supplemental round of standardisation that is specific for the domain. An important activity in this standardisation effort is the development of a common GML-based interchange format that can be shared (and served) by data providers (OGC, 2004). Several geoscience initiatives have elected to implement GML standards: e.g., XMML (eXploration and Mining Mark-up Language, http://xmml.arrc.csiro.au/), GeoSciML (IUGS Commission for the Management and Application of Geoscience Information, or "CGI"; https://www.seegrid.csiro.au/twiki/bin/view/CGIModel/GeoSciML), and NADM (Boisvert et al., 2004). H2O is the next step: it is a groundwater interchange standard based on NADM, XMML, and GeoSciML.

H2O: just add water (to NADM)

Our GML encoding for groundwater data is derived from the NADM-C1 GML effort (Boisvert et al., 2004), XMML, and the new international effort from IUGS (GeoSciML, 2005; Sen and Duffy, 2005). NADM provides the geoscience framework from which we could derive hydrogeological concepts, and XMML and GeoSciML provides the human artifacts (borehole, observations patterns, etc.). Note that GeoSciML is already a fusion between large portions of NADM-C1 and XMML. H2O is the sum of work carried in several projects within our departments (such as PATHWAYS; Brodaric, et al., 2005) and abroad. If we could put it in a single line, it would read as follows:

```
H2O = NADM + GeoSciML + XMML + NGWD +
    PATHWAYS + RésEau
```

(NGWD is the Canadian National Groundwater Database.)

The H2O model is still a work in progress: it addresses about half of the concepts required to successfully exchange groundwater data. The qualitative aspect is being worked on with our Environment Canada colleagues (through RésEau) while we have concentrated on the quantitative aspect.

Figure 1 shows the main classes we derived from NADM-C1 and XMML/GeoSciML. Most of the top level concepts shown there are drawn from NADM-C1, and a single concept (**`Waterwell`**, a specialisation of **`Borehole`**) is from XMML/GeoSciML. The contribution of XMML/GeoSciML is more in terms of the Observation and Measurement modules (the human artifacts).

We derived **`HydrogeologicUnit`** from **`GeologicUnit`** to provide a home for concepts such as **`Aquifer`** and **`Aquitard`**. We created **`HydrogeologicProperty`** from **`GeologicProperty`**, to provide properties specific to **`HydrogeologicUnit`**. We also had to create a new property under **`GeologicProperty`** called '*porosity*', which is truly a property of the

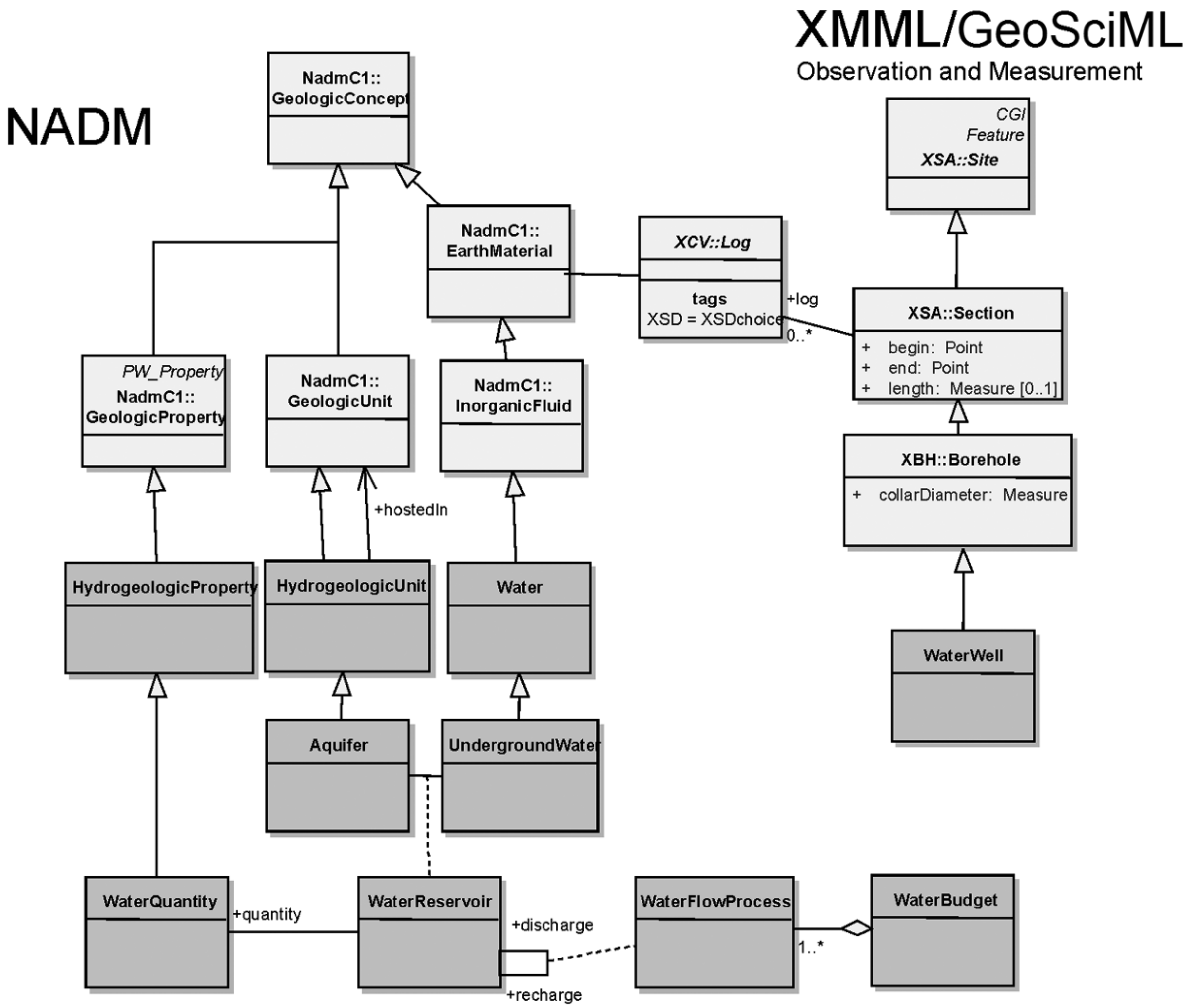


Figure 1. General model of H2O and derivation from NADM-C1 and XMML. The bottom part, shaded gray, represents extensions defined for groundwater.

`EarthMaterial` and not of the **`HydrogeologicUnit`**. The most interesting relation between **`HydrogeologicUnit`** and **`GeologicUnit`** is the '**`hostedIn`**' relation, enforcing the fact that an **`Aquifer`** (a **`HydrogeologicUnit`**) is hosted in **`GeologicUnits`**.

Finally, after much debate about water being a fluid or a mineral (water in the form of ice meets all the requirements of a mineral), we decided it was, for the purposes of exchanging groundwater data, an **`InorganicFluid`**. The relation between water and hydrogeologic unit is done through **`Reservoirs`**, and the flow of water between reservoirs is a water budget, which is at the heart of the quantitative model.

Figure 2 is a more detailed view of the Water Budget structure and related concepts. The **`WaterBudget`** is the aggregation of all inputs and outputs in a given reservoir (discharge and recharge depends on which reservoir you are considering) through a series of flow processes. One might point out that we missed an opportunity to derive those concepts from **`GeologicProcess`**, but most (if not all) of those processes are physical processes that are not restricted to the geological realm.

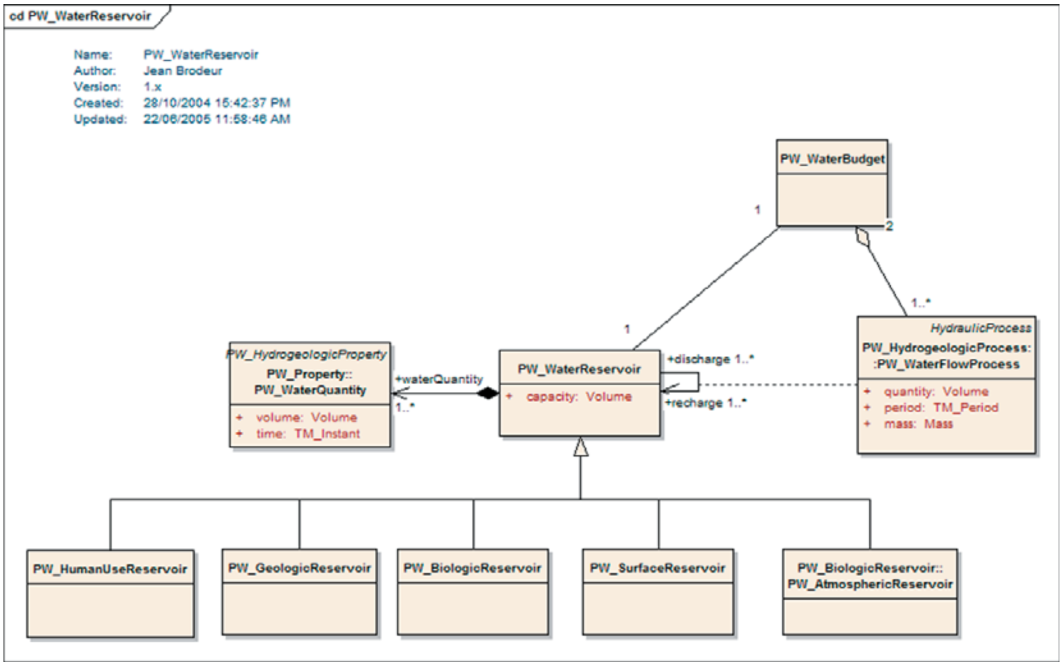


Figure 2. The water budget model. A budget is the sum of all flows that enter and exit a reservoir. The debit or credit is assigned depending on the flow direction (discharge or recharge).

The logic becomes clearer when we go through an example representing an instance of the model of Figure 2. For an introduction to GML, we refer the reader to OGC (2004), Lake et al. (2004), and Boisvert et al. (2004):

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<nadm:Nadm xmlns="http://gwp.nrcan.gc.ca/ngwd"
xmlns:gml="http://www.opengis.net/gml"
xmlns:nadm="http://geology.usgs.gov/dm/NADM/v1.0"
xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:xmml="http://www.opengis.net/xmml"
xmlns:xsi=http://www.w3.org/2001/XMLSchema-instance>
      <gml:featureMember>
     [1]  <GeologicReservoir gml:id="EskerAbitibi">
            <capacity uom="m3">0.0</capacity>
            <waterReservoirBudget>
     [2]      <WaterBudget gml:id="B1">
                <gml:description>
            Calculation of complete budget of the St-Mathieu esker water budget
                </gml:description>
                <gml:name>Complete budget 2004</gml:name>
     [3]        <waterFlowComponent>
                  <Precipitation gml:id="P1">
                    <quantity uom="m3">21870360</quantity>
                    <gml:timeInterval unit="year">1</gml:timeInterval>
                    <recharge xlink:href="http://gwp.nrcan.gc.ca/ngwd/
                           Reservoirs#Atmosphere"/>
```

```
          </Precipitation>
        </waterFlowComponent>
        <waterFlowComponent>
          <Pumping gml:id="Pu1">
            <gml:name>Pumping from Amos old wells</gml:name>
            <quantity uom="m3">1193795.28</quantity>
            <gml:timeInterval unit="year">1</gml:timeInterval>
            <discharge xlink:href="http://gwp.nrcan.gc.ca/ngwd/
                   Reservoirs#Municipal
            Facilities"/>
          </Pumping>
        </waterFlowComponent>
[more waterFlowComponent removed for readability ]
      </WaterBudget>
    </waterReservoirBudget>
[4] <groundWaterContainer>
      <Aquifer gml:id="EskerAbitibiAquifer">
        <!-- the aquifer is hosted in an Esker -->
[5]     <hostIn xlink:href="#E1"/>
      </Aquifer>
    </groundWaterContainer>
    <groundWaterContent>
[6]   <GroundWater gml:id="E1.W">
            <gml:description>Water contained in the esker, water properties
                   should be added at this point to characterise this
                   particular groundwater</gml:
            description>
      </GroundWater>
    </groundWaterContent>
  </GeologicReservoir>
</gml:featureMember>
<gml:featureMember>
[7] <nadm:GeomorphologicUnit gml:id="E1">
    <gml:description>Large N-S sand and gravel body</gml:description>
    <gml:name>Esker St-Mathieu/Berry</gml:name>
    <nadm:geologicUnitMember>
      <nadm:GeologicUnitPart gml:id="E1.P1">
        <nadm:proportion uom="pct">100</nadm:proportion>
        <nadm:gupMaterial>
          <nadm:UnconsolidatedMaterial gml:id="E1.P1.M1">
            <gml:description>Thick beds of coarse sand and gravel, poorly
                   sorted</gml:
            description>
            <gml:name>sand and gravel</gml:name>
          </nadm:UnconsolidatedMaterial>
        </nadm:gupMaterial>
        <nadm:guRole>composition</nadm:guRole>
      </nadm:GeologicUnitPart>
    </nadm:geologicUnitMember>
  </nadm:GeomorphologicUnit>
</gml:featureMember>
</nadm:Nadm>
```

This document describes a water budget for an aquifer in the Abitibi area of Québec, Canada (preliminary data from Riverin, in preparation). Points of interest are marked by a number in the left column. The line marked as [1] is the beginning of the description of a Reservoir (a **`GeologicReservoir`**) for which a budget has been calculated. The **`WaterBudget`** at [2] contains the list of all the **`waterFlowComponents`**. Each waterflow component [3] contains a specific process (Precipitation, Pumping, etc), the direction of the flow (**`discharge`** or **`recharge`**), and the reservoir the water comes from or goes to. The destination (or the origin) of the water is useful if we need to balance several budgets, like surface-groundwater interaction. The groundwater container (the unit that acts as the reservoir) is described in [4] (an **`Aquifer`** is a groundwater reservoir) and note in [5] that this **`Aquifer`** is hosted in a **`GeomorphologicalUnit`** that is described further down at [7] (and pointed to by an **`xlink:href`**, which is the mechanism employed in GML to point to other sections of the document or to elements in another document). In [6], we define the water that is contained in the aquifer. This looks superfluous at this point, but you might see this as a placeholder where water properties can be attached. Finally, in [7], the host unit (referred in [5]) is described with its components according to NADM-C1 model (Boisvert et al., 2004).

USE CASE

In this section we describe a use case that demonstrates the usefulness of interoperability for groundwater and related domains. In the use case, the water level in an aquifer is required to assess the sustainability of housing developments in certain communities where groundwater is the sole or principal source of water. Combining this information with other socioeconomic variables and providing local government with decision making tools then allows calculations to be made about current and future trends for water supply and demand. Using groundwater information in this way is the goal of the PATHWAYS project (http://sdki.nrcan.gc.ca/path/index_e.php).

The simplest use case would allow PATHWAYS modelling tools to access the water level information stored in various provincial water well databases, without any prior knowledge of how the data are actually structured or how to access them. The process, demonstrated to some extent during the DMT'05 presentation (Brodaric, et al., 2005), requires a series of intermediate pieces of software to handle the request from one step to the next. Figure 3 is a sketch of the process.

- First, a tool designed by the PATHWAYS project team (the Phoenix browser) sends a request to the National Groundwater Database (NGWD) for a specific theme (Water Level) using a common schema: H2O. The request is made using the Web Feature Service standard protocol (OGC, 2002);
- NGWD receives the request and determines which database holds this information. Once it locates the provincial service that might have this information, it translates the H2O request into a schema the provincial service can understand (it might be an OGC standard, or it might not);
- NGWD sends the translated request to the provincial database, which proceeds to extract the information. This may involve another translation step that turns the web based query into a database query– e.g., XML into a SQL statement;
- The information from the province is streamed back to NGWD in either XML, HTML, or another specified format. NGWD performs the reverse translation to turn this into the H2O public schema and sends it over to PATHWAYS, which is unaware of the provincial schema; and
- PATHWAYS receives the H2O document and turns it into the internal format required by the modelling tool.

Many variations of this scenario might exist. If the province follows OGC WFS standards, much less work is required by NGWD to translate it, because WFS is using GML (Boisvert et al., 2004). If the province follows the H2O public schema, NGWD does no translation at all. On the other hand, if the service is based on any other technology, a specific solution must be devised for this particular service. In any case, the goal of NGWD is to shield PATHWAYS from those details also that it is exposed only to data accessed using H2O and WFS.

CONCLUSION

NADM-C1 GML provided a good starting point for our groundwater data interchange format, called H2O. In H2O we leveraged the fact that hydrogeology is essentially an extension of geology (at least for its quantitative aspects), allowing us to reuse many of the concepts in NADM-C1. We showed how H2O is developed from NADM-C1, how it is structured, and how it is implemented in the National Groundwater Database. Future work involves extending H2O to include water quality concepts, so that it can be used as an interchange format for both water quantity and quality information.

ACKNOWLEDGMENTS

This work has been financed by the *Groundwater* program, the *Sustainable Development through Knowl-*

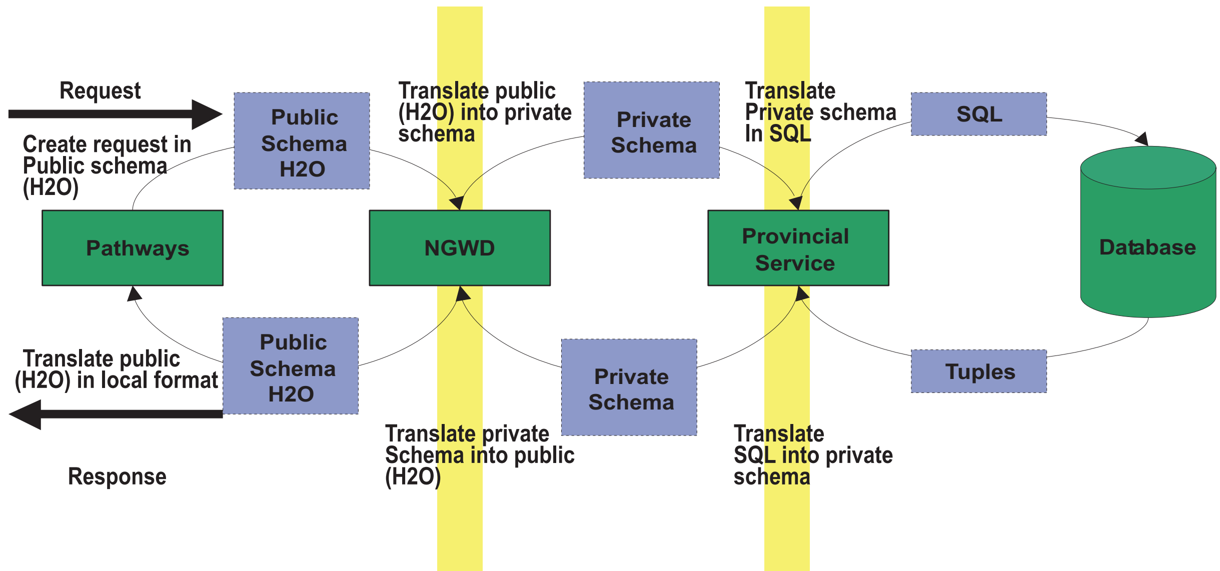


Figure 3. General view of the components involved in the DMT '05 demonstration of interoperability between PATHWAYS and the National Groundwater Database.

edge Integration program of the Earth Science Sector of Natural Resources Canada, and the ResEau project of Environment Canada (ESS contribution number 20060072).

A special thought for the people of Louisiana, in light of the terrible events that occurred a few months after they welcomed us to Baton Rouge.

REFERENCES

Boisvert, E., Johnson, B., Brodaric, B., and Cox, S., 2004, GML Encoding of NADM C1.0, Connecting geoscience to international standards, in D.R. Soller, ed., Digital Mapping Techniques '04—Workshop Proceedings: U.S. Geological Survey Open-file Report 2004-1451, p.95-103, available at http://pubs.usgs.gov/of/2004/1451/boisvert/index.html.

Brodaric, B., Van Ulden, J., Grant, R., Krauth, O., Journeay, M., Talwar, M., and Denny, S., 2005, Pathways-Phoenix: A Web Application for Decision Support using OGC Standards, in D.R. Soller, ed., Digital Mapping Techniques '05—Workshop Proceedings: U.S. Geological Survey Open-file Report 2005-1428.

CGDI Architecture Working Group, 2001, Canadian Geospatial Data Infrastructure Architecture Description Revision, version 1: CGDI Architecture Working Group, http://www.geoconnections.org/architecture/architecture_description.pdf.

Sen, M., and Duffy, T., 2005, GeoSciML: Development of a generic GeoScience Markup Language, Computers and Geosciences, v.31, no.9, p.1095-1103.

GeoSciML, 2005, https://www.seegrid.csiro.au/twiki/bin/view/CGIModel/WebHome.

Lake, R., Burggraf, D., Trninic, M., and Rae, L., 2004, Geography Mark-Up Language, Foundation for the Geo-Web: John Wiley and Sons, 432 p.

North American Geologic Map Data Model Steering Committee, 2004, NADM Conceptual Model 1.0, A Conceptual Model for Geologic Map Information: Geological Survey of Canada, OPEN FILE 4737, 1 CD-ROM, and U.S. Geological Survey Open-File Report 2004-1334, 58 p; available at http://pubs.usgs.gov/of/2004/1334 and at http://nadm-geo.org.

NGWD, 2005, National Groundwater Database, accessed at http://ntserv.gis.nrcan.gc.ca/gwp/ngwd/exploration.

OGC, 2004, ISO/TC 211/WG 4/PT 19136, Geographic information – Geography Markup Language (GML), accessed at http://portal.opengeospatial.org/files/?artifact_id=4700.

Rivera, A., 2003, Canadian Framework of Collaboration on Groundwater: Government of Canada internal paper, accessed at http://gwp.nrcan.gc.ca/pdf/cadre_canadien_collaboration_eau_souterraine_e.pdf.

Riverin, M-N, in preparation, Caractérisation et modélisation de la dynamique d'écoulement dans le système aquifère de l'esker Saint-Mathieu / Berry, Abitibi, Québec, Mémoire présenté pour l'obtention du grade de Maître ès science (M.Sc.) en Science de la Terre, Université du Québec, Institut national de la recherche scientifique, Eau,Terre et Environnement. 157 p.

XMML, 2004, https://www.seegrid.csiro.au/twiki/bin/view/Xmml/WebHome.

NGMDB Geologic Map Feature Class Model

By Steve Richard[1], Jon Craigue[2], and Dave Soller[3]

[1]Arizona Geological Survey
416 W. Congress #100
Tucson, AZ 85701
Telephone: (520) 770-3500
Fax: (520) 770-3305
e-mail: Steve.Richard@azgs.az.gov

[2]U.S. Geological Survey
Earth Surface Processes Research Institute
Tucson, AZ
e-mail: jcraigue@espri.arizona.edu

[3]U.S. Geological Survey
926-A National Center
Reston, VA 20192
e-mail: drsoller@usgs.gov

BACKGROUND

This document describes the spatial data feature classes in the prototype design for the National Geologic Map Database project (NGMDB; http://ngmdb.usgs.gov). The implementation of thematic data table in the NGMDB prototype is described in a separate document (Richard et al., 2004). The design presented here has been submitted as an ESRI geology data model, and is available from the ESRI support web site (see http://support.esri.com/index.cfm? fa=downloads.dataModels.filteredGateway&dmid=30, and find the NGMDB database design documents link).

Geoscience entities of interest

As a precursor to defining feature datasets and feature classes in an ESRI geodatabase, this section enumerates the geologic entities of interest and the spatial relationships between them. The term "entity" is used here to denote phenomena of interest observed in the 'real world', as opposed to features, which are the database objects implemented to represent our understanding of those phenomena. A feature in ESRI geodatabase usage is explicitly required to have a geometry property that specifies a location and shape. This discussion uses terminology and basic definitions from the NADM C1 model (NADMSC, 2004).

Geologists are concerned with the three dimensional arrangement of material within the Earth. The entities of interest are bodies of material (geologic units) and surfaces that bound or cut them (geologic surfaces). The 2-D map view that is the framework for a GIS represents the geometry of the intersection of these entities with some map horizon—typically the Earth's surface, but possibly some abstract surface like a mine-level, cross section surface, or some buried surface (e.g. top of Precambrian rock). The basic features that may be implemented in a 2 (or 2.5)-D GIS are points, lines, and polygons (and composite features aggregated from these simple features). Points represent the intersection of a line with the map horizon (e.g. a borehole collar), or an observation location on the map horizon (a station). Lines represent the intersection of a surface with the map horizon (surface trace), the projection of some buried line beneath the map horizon (e.g. the cutoff of a contact at a fault, an inclined borehole, a channel axis), or a line defined within the map horizon (e.g. sand dune crest, geomorphic escarpment). Polygons represent one of several situations, including the intersection of bodies with the map horizon (i.e. the outcrop of geologic units), patches defined on the map horizon, the projection of patches on a surface other than the map horizon into the map plane, or the projection of 3-D bodies that do not intersect the map horizon into the map plane.

The following discussion elaborates on this basic framework to define the entities of interest that need to be mapped into feature classes, feature datasets, and topology rules in an ESRI geodatabase implementation.

- **Genetic boundary surface**. Boundaries of geologic units related to the genesis of the unit. Includes depositional contacts, facies changes (in igneous, sedimentary or metamorphic rocks…), and intrusive contacts. Genetic boundaries are surfaces that are either truncated by other younger genetic boundary surfaces, by faults, or by the Earth surface. The identity of a genetic boundary surface is associated with the identity of one of the geologic units juxtaposed at the surface, for example 'base of Escabrosa Limestone', 'boundary of fine-grained facies of Oracle granite', 'top of Cambrian strata'.
- **Fault**. A surface across which there has been shear displacement significant at the scale of observation. Fault surfaces are truncated by younger genetic boundary surfaces, other faults, or by the Earth surface. Fault surfaces may also end because fault displacement may decrease to the point that the fault is no longer identifiable/mappable. Identity of a fault surface is defined by physical continuity at the time the fault is active. Inactive faults that are truncated by younger faults may have segments defined by continuity between truncating faults; these are correlated with other segments based on interpretation of movement history (timing, direction, and magnitude of displacement). Active faults may have discrete segments, separated by recognizable boundaries, which tend to rupture independently. Groups of faults may be interpreted to operate together as a fault system considered to be a single tectonic entity.
- **Fold hinge surface**. A surface defined as the locus of points that occupy the hinge of a single fold structure; the surface itself does not necessarily have a material manifestation, but is locatable. Fold hinge surfaces are truncated at younger genetic boundary surfaces, faults, or the Earth surface, or may terminate where a fold loses definition. Identity is defined by physical continuity of the hinge surface at the time of fold formation.
- **Geologic unit**. An identifiable part of the earth based on some geologic criteria. Typically, a body of material (rock or nonconsolidated). Identity criteria are variable; ideally defined by lithologic properties, but may be defined by identity of bounding surfaces, or interpreted properties like age, depositional environment, alteration history, or P-T conditions. Geologic units are bounded by genetic boundary surfaces, by faults, or by the Earth surface. They are grouped in various part-hierarchies used at different levels of generalization (e.g. member, formation, group, supergroup).
- **Dike**. A geologic unit of igneous origin that is very thin relative to its lateral extent. This generalized definition does not consider relationship to layering of host rock or orientation of body because dikes and sills are depicted the same way on maps. In detail, a dike has two genetic boundary surfaces (one on each side) but, in general, dikes are considered as a surface-like entity. Identity of an individual dike is defined by physical continuity at time of formation, but groups of dikes are typically classified as a unit based on lithology and interpretation of a single magmatic source. Dikes are truncated at genetic boundaries, faults, or the Earth surface, or may simply end where the intruded crack ends.
- **Vein**. Similar to dike, but groups of veins are classified into units based on lithology and interpreted relationship to hydrothermal events.
- **Escarpment**. Abrupt change of slope (from more gentle to very steep) on the Earth surface (exposed or buried), related to erosional or tectonic processes. Fault scarps are coincident with a fault. Identity based on physical continuity. Scarps are classified based on interpreted history.
- **Fissure**. A crack in the earth's surface, generally with dilatational deformation. Fissures may be: 1) intrabasinal in active sedimentary basins, related to desiccation, ground water withdrawal, or compaction, 2) surface collapse due to subsurface dissolution (evaporite or karst), or 3) related to slope failure.
- **Borehole**. A human-engineered hole drilled into the earth to obtain subsurface resources or information. Has an associated point (collar) from which the hole was drilled, typically the Earth surface, but may be a subsurface point from an underground mine or other working. Multiple boreholes may be associated with a single collar location. Identity of a bore hole is defined by a single 'drilling event'. Boreholes may be reentered to drill deeper or to produce a new borehole (sidetrack).
- **Station**. A point location at which data or samples were acquired. Identity defined by observer who locates the station. Stations are not necessarily associated with any particular identifying phenomenon. It is simply where the geologist stopped to measure bedding, record some observations, or perhaps take a picture. A station has a 3-D location that may be inherent in its association with a map horizon (e.g. X,Y coordinated on Earth surface), or borehole (location reported by depth below collar), or explicitly recorded as an XYZ coordinate (a location in an underground mine).

GEODATABASE FEATURE CLASSES

Table 1 summarizes ESRI geodatabase feature classes used to specify location in the NGMDB implementation. All spatial data tables include fields to specify a default text label and symbol to use in map displays if no other symbolization is specified. This is to simplify the rapid display of spatial data. OutcropBoundaryTrace and GeologicUnitOutcrop are line and polygon feature classes whose locations represent observable geologic phenomena in or on the Earth. Station is a point feature class that specifies a location at which observations were made, and does not (inherently) represent the location of some phenomena. The term observation is used in the

Table 1. Location specification tables.

Table	Description
OutcropBoundaryTrace	Line features that represent the intersection of geologic surface that bounds mapped rock bodies with the depicted map horizon. Subtypes include fault and geneticBoundary. These traces may bound GeologicUnitOutcrop polygons.
DikeVeinMarkerTrace	Line features that represent the intersection of a dike, vein, or marker bed with the depicted map horizon; subtypes differentiate these cases.
HingeSurfaceTrace	Line features that represent the intersection of the hinge surface of a fold with the depicted map horizon.
EarthFissureTrace	Line feature that represents a fissure in the depicted map horizon.
ConcealedBoundaryTrace	Line feature that represents the trace of a geologic surface in a map horizon different from the depicted map horizon. Subtypes include concealed faults and genetic boundaries, and structure contours.
GeologicUnitOutcrop	Polygons representing the intersection of a geologic unit with the map horizon.
Station	Point location at which one or more observations are made, or samples are collected.
BoreholeCollar	Point location at which a borehole section intersects a map horizon (typically the Earth surface).

sense of GML Observation and Measurement (Cox et al., 2004). These feature classes are organized according to their semantics, associated properties, and implementation issues for using topology rules in the ESRI geodatabase environment.

Notes on schema notation and implementation

The schema included in this paper use a UML profile defined by ESRI to build UML models for geodatabase physical implementation using CASE tools. Classes with names in *italics* represent abstract classes, i.e. there is no feature class in the geodatabase with the same name. Attributes of abstract classes are included in all subtype classes. UML subtype links have an open triangle at the parent class end of the link. Attributes in the UML classes represent fields in the geodatabase tables. In the models here, attribute names that end in 'GUID' represent database fields populated by 36-character string GUIDs (Globally Unique IDentifiers). ESRI geodatabase subtypes are linked to their parent class (which represents the physical table implemented in the database) using dashed lines. ESRI subtypes represent subsets of records in a particular relational database table that are differentiated based on an integer value in one of the fields identified as the 'subtype' field. In the NGMDB implementation, ESRI subtype field names always begin with a prefix "ESRI-Subtype". The undecorated solid lines linking classes in the diagrams represent associations implemented as ESRI relationship classes in the geodatabase. These associations are navigable in the ArcMap attribute browser.

Figure 1 shows the top level hierarchy of feature classes in the geodatabase implementation model. The ESRIClasses::Feature and *NGMDBFeature* classes define fields shared by all spatial data classes (see Richard et al., 2004 for discussion of standard NGMDB fields). The feature classes are divided into broad groups represented by abstract classes beneath *NGMDBFeature* in Figure 1. *ObservationLocation* represents features located based on human, observation factors—e.g. where access is possible, what part of a mountain could actually be seen, where the airplane flew. Their location is typically related to geologic phenomena of interest, but their location is not determined by the location of the phenomena. *MappedOccurrence* includes features whose location is determined by phenomena inherent in the earth—contacts between rock bodies, fault zones, fold hinges, etc. *GeologicRoute* features aggregate *MappedOccurrences* that are interpreted to represent the traces of extended geologic surfaces identified based on multiple observations. *AnnotationFeatures* position annotation in map displays. The location of these features is related to a *MappedOccurrence* or *ObservationLocation*, but the actual positioning is determined by cartographic considerations. BoreHoleCollar is in a sense a sort of *MappedOccurrence*, but because the properties of interest are different, they have been implemented as a distinct feature class. The following discussion first treats the subtypes of *MappedOccurrence*, followed by *GeologicRoutes*, *ObservationLocations*, and *AnnotationFeatures.*

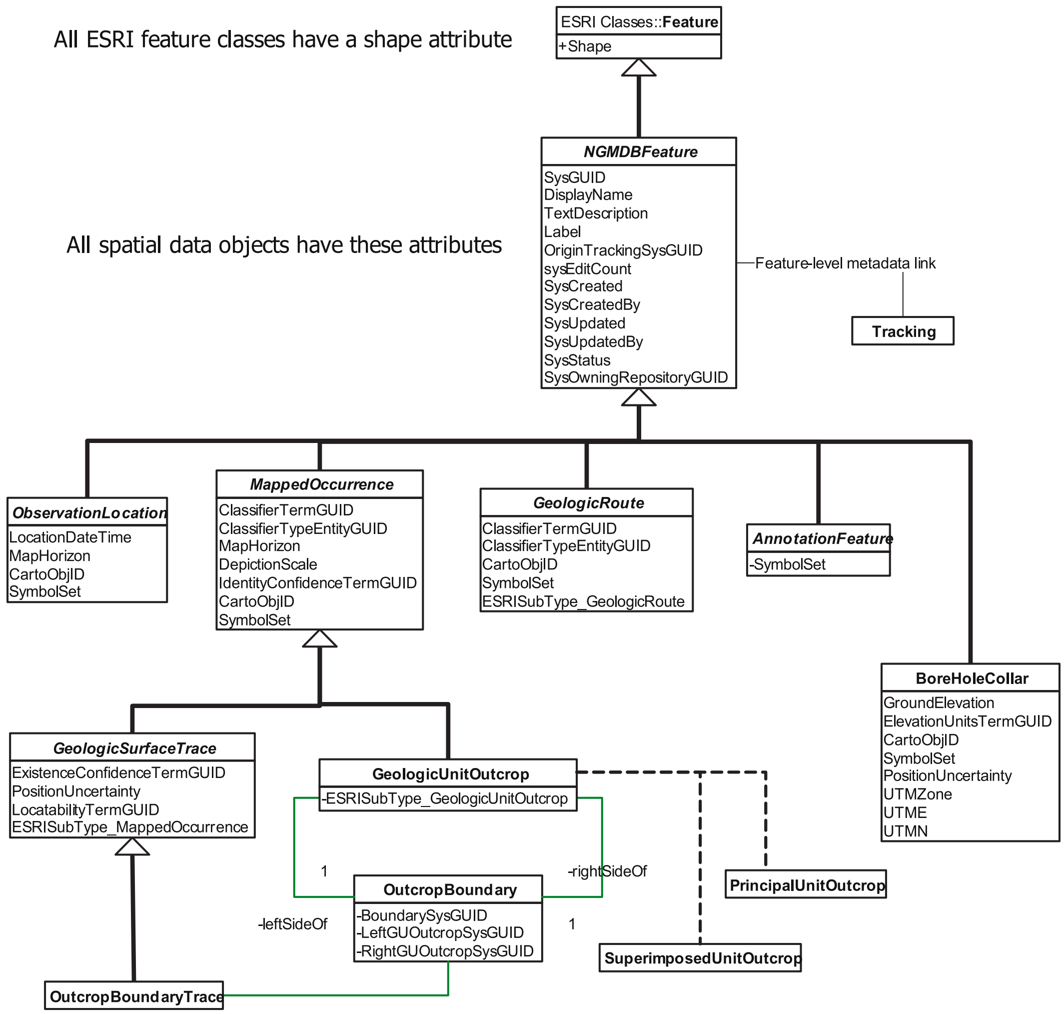


Figure 1. Top level features in the NGMDB implementation. GeologicUnitOutcrop polygons are associated with bounding OutcropBoundaryTrace arcs (LPoly and RPoly of Arc/Info coverages) explicitly through the OutcropBoundary correlation table. PrincipalUnitOutcrop and SuperimposedUnitOutcrop are ESRI subtypes of GeologicUnitOutcrop, discriminated using the ESRISubtype_GeologicUnitOutcrop field. See Figure 3 for some other MappedOccurrence subtypes.

Surface traces

Surface traces are the lines in a GIS that represent intersection of 3-D surfaces with a map horizon. Thus a surface trace always has an explicit or implicit defining map horizon (MapHorizon in *MappedOccurrence*, Figure 1), and a classifier that specifies the kind of surface that intersects the map horizon to form the trace. The surface trace has associated location uncertainty related to how discretely the mapped surface may be located (e.g. sharp or gradational contact), how precisely the location can be determined (good or poor exposure), and how accurately that location can be specified in map coordinates (Figure 2).

Surface traces that are in the depicted map horizon

A geologic map is assumed to portray surface traces and outcrops on some particular map horizon. The depicted map horizon may be different in different parts of the map. For example the current Earth surface may be the map horizon except in the area of a large mine, where the pre-mining surface may be depicted. All of these surface trace types may have elevation values (Z values in geodatabase) associated with vertices along the trace; these elevations represent the elevation of the map horizon.

- **OutcropBoundaryTrace**. This feature class contains

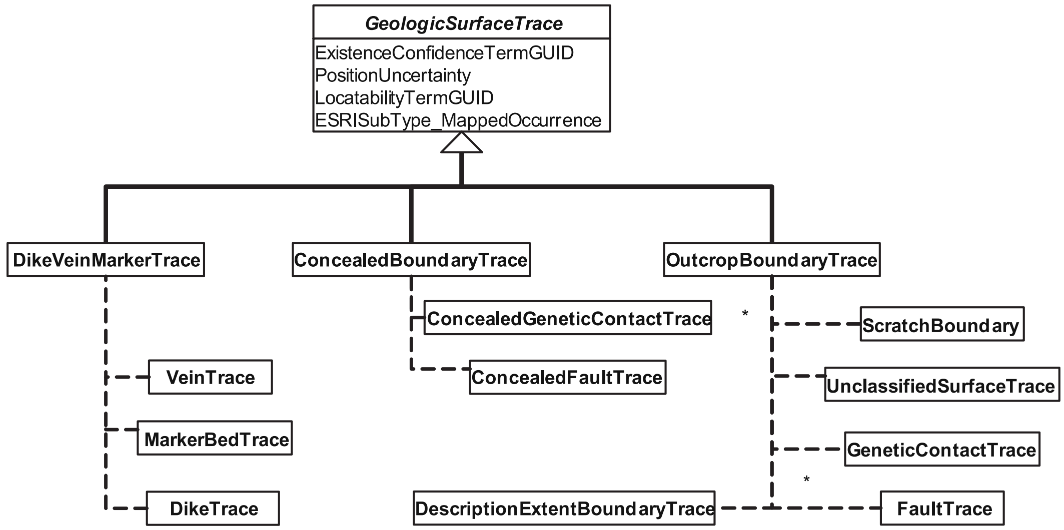


Figure 2. Geologic surface trace feature classes. These feature classes represent the intersection of geologic surfaces with a map horizon. The OutcropBoundaryTrace features participate in polygon topology with the GeologicUnitOutcrop.PrimaryUnitOutcrop features (Figure 1).

surface traces that participate in topologic rules between polygons and the geologic surface traces that bound them. Subtypes identify two broad categories (Figure 2):

- ◦ GeneticSurfaceTrace. Represents intersection of genetic boundary surface with a map horizon. Classifier term property identifies different kinds of geologic body boundaries (intrusive contact, depositional contact, facies boundary, etc.) using terms from a controlled vocabulary (see Vocabulary tables, above). Topologic constraints: no dangles—must connect to other genetic contacts or to faults. No self intersection, may self connect (to form closed loop). GeneticSurfaceTrace may not intersect FaultTrace or GeneticSurfaceTrace (one or the other must be younger, and break the older surface trace). GeneticSurfaceTrace may not intersect VeinTrace or DikeTrace (one or the other must be younger, and break the older surface trace), but may be covered by DikeTrace or VeinTrace where the dike or vein is intruded along the contact. Genetic contacts must be covered by the boundary of a geologic unit polygon.
- ◦ FaultTrace. Represents intersection of a fault surface with a map horizon. Classifier term property identifies different kinds of faults (thrust fault, normal fault, strike slip fault, detachment fault, etc.) using terms from a controlled vocabulary. Topologic constraints: dangles allowed, no self intersection, may self connect (to form closed loop – e.g. window through flat fault). FaultTrace may not intersect FaultTrace or GeneticSurfaceTrace (one or the other must be younger, and break the older surface trace). FaultTrace may not intersect VeinTrace or DikeTrace (one or the other must be younger, and break the older surface trace), but may be covered by DikeTrace or VeinTrace where the dike or vein is intruded along the fault. Has Z—determined by elevations on map horizon. Faults may cover the boundaries of outcrop polygons.

- **DikeVeinMarkerTrace**. Represents intersection of a dike, vein, or marker bed, considered as a surface, with a map horizon. Concealed traces are depicted using the same feature class. Subtypes differentiate dike trace, vein trace, and marker trace. Same rules as FaultTrace.
- **HingeSurfaceTrace** (Figure 3). Represents intersection of a fold hinge surface with a map horizon. Concealed traces are depicted using the same feature class. Topological constraints: dangles allowed, must not self intersect, may self connect (trace of recumbent fold on steep terrain). May intersect GeneticSurfaceTrace, FaultTrace, DikeTrace, VeinTrace (where the surface represented by the trace is folded).
- **EarthFissureTrace** (Figure 3). Represents intersection of a fissure with the map horizon. Must not self intersect.
- **GeomorphicFeatureTrace** (Figure 3). Lines representing the trace of a linear feature defined by the morphology of the map horizon surface. Subtypes identify various kinds with different topology rules and associated classifier/symbol domains.
  - ◦ Escarpment. Line representing the uphill edge of a geomorphic escarpment; line physically resides in a map horizon (e.g. Earth surface, top of bedrock). Topology rules: must not self intersect.
  - ◦ FaultScarp. Line representing the uphill edge of

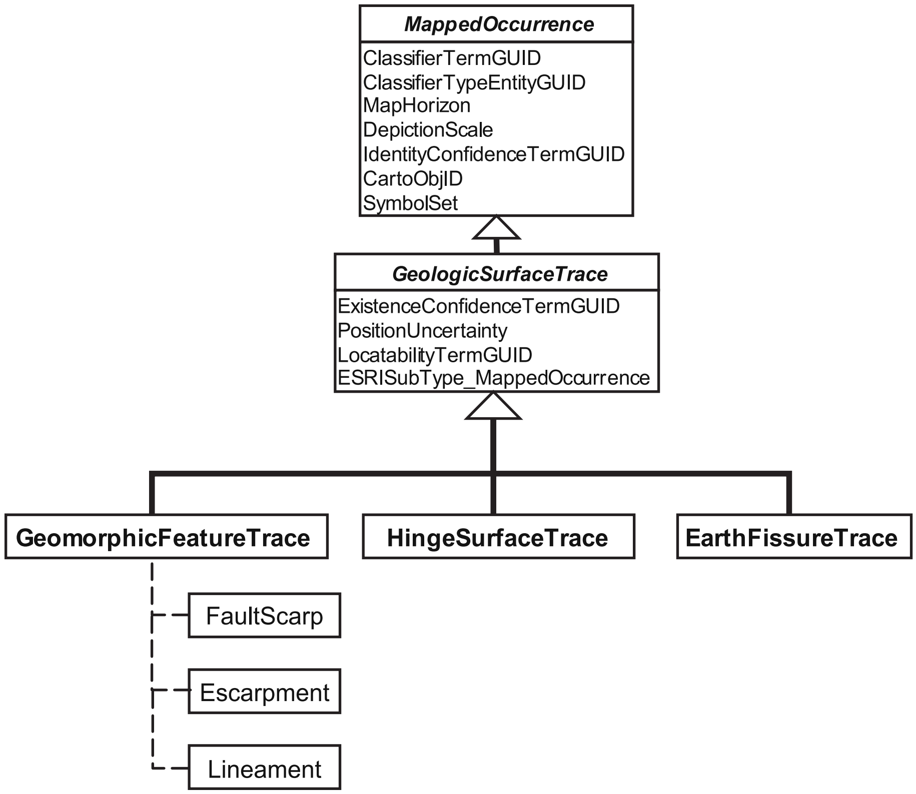


Figure 3. Other mapped occurrence feature classes. These are less frequently encountered mapped occurrences not shown in Figure 1.

a fault scarp. Topology rules: must be covered by FaultTrace or Concealed FaultTrace.
- ◦ Lineament. Line representing a geomorphically expressed lineament of uncertain nature.

Surface traces not in the depicted map horizon

- **ConcealedBoundaryTrace**. This feature class contains mapped features that represent traces that do not participate in topologic rules between polygons and the geologic surface traces that bound them (outcrop boundary traces). Subtypes identify three categories (Figure 2):
  - ◦ Concealed GeneticSurfaceTrace. Dotted contacts on geologic maps are used in different ways:
  - ◦ To represent the trace of a genetic boundary surface on a buried map horizon (e.g. top of bedrock beneath Quaternary cover).
  - ◦ To connect GeneticSurfaceTraces of surfaces that have been intruded by igneous rocks, and thus have no associated map horizon—they are purely to indicate inferred pre-intrusion continuity of some surface.
  - ◦ To show genetic boundaries on a pre-existing but related map horizon, for example the location of an active river channel based on several generations of air photos.

In either case the topology rules are similar: may not self intersect, no dangles. Concealed GeneticSurfaceTraces are not required to be covered by the boundary of a GeologicUnitOutcrop—they do not participate in geologic unit polygon topology on the map horizon depicted. Concealed GeneticSurfaceTraces may intersect other GeneticSurfaceTraces (e.g. boundaries of mapped covering geologic unit outcrops), but may not intersect FaultTrace, DikeTrace, VeinTrace or Concealed GeneticSurfaceTrace, Concealed FaultTrace, Concealed DikeTrace or Concealed VeinTrace.

  - ◦ Concealed FaultTrace. Trace of a buried fault or connection of intruded fault (see Concealed GeneticSurfaceTrace, above). Rules same as for FaultTrace, except Concealed FaultTrace may intersect GeneticSurfaceTrace.
  - ◦ StructureContour. The trace of a geologic surface on a horizontal surface of some given elevation. This is essentially an abstract map horizon, and the StructureContour is a surface trace on that horizontal map horizon. Thus, structure contours may be modeled as concealed surface traces with the contour elevation identified by the map horizon property of the trace , or using the Z values for the line in the geodatabase feature class. Structure contour maps contrast with other geologic maps in that they do not portray a

single map horizon, but rather display surface traces from a collection of map horizons in a single portrayal. Topology rules: must not self intersect.

Outcrop

Outcrop is used here in the very specific sense of the intersection between a geologic unit and a map horizon, whether or not the geologic unit is exposed on that horizon. The term exposure is used to refer to places where the geologic unit is visible on the map horizon. Outcrops are represented by polygons in the GIS. An outcrop always has an explicit or implicit defining map horizon, and a classifier that specifies the geologic unit that intersects the map horizon to form the outcrop.

- **GeologicUnitOutcrop** (Figure 1). Polygon that represents the outcrop of a geologic unit. Topology rules: boundary must be covered by GeneticSurfaceTrace or by FaultTrace on same map horizon. Must not overlap other GeologicUnitOutcrop on same map horizon. Must not have gaps (map horizon must be covered over the extent of the map). Subtypes:
    - PrincipalUnitOutcrop. Outcrops of units of principal interest.
    - SuperimposedUnitOutcrop. Polygon representing outcrop of a geologic unit superimposed on the principal mapped units, for example alteration zones, zones of crushing, metamorphic zones. These outcrops are on the same map horizon as the associated principalUnitOutcrop. Overlaps between these intersections of contacts must be treated as topology rule exceptions. If there are a sufficient number of superimposed unit outcrop polygons, a new GeologicUnit FeatureDataset should be created with its own lines, polys, and topology rules.

Routes

Collections of surface traces classified to belong to a single ‘broader’ classification entity. The individual surface trace instances in the SurfaceTrace feature classes are differentiated based on their classification (depositional contact, thrust fault, intrusive contact, facies boundary…), and observation-related properties (classification confidence, location confidence, observer, depiction scale, observation method…). These may be considered together to represent the trace of some geologic entity. These aggregated features are treated separately from the MappedOccurrences because they are fundamentally interpretive, and have different metadata properties.

- **Fault routes** (Figure 4). Fault arcs (individual surface trace instances) may be aggregated to form fault segments or faults, and these may be further aggregated into fault systems of different scales. Fault segments are outcrop traces of fault surfaces bounded by their intersection with other fault surfaces or genetic contact surfaces. Faults are aggregations of fault segments, fault systems are aggregations of faults. Topology rules (using fault as example): segments must be covered by surface trace instances, faults must be covered by fault segments, fault system must be covered by faults. In this context ‘covered’ means that there must one or more coincident data instances in the covering feature class that together completely match the covered class. Fault segments may not intersect or overlap. No self intersections.
- **Fold routes** (Figure 4). Similar to fault route, individual folds, fold systems (anticlinorium, synclinorium…).
- **Contact route** (Figure 4). Similar to fault route— represents outcrop of a particular geologic boundary surface.

Observation location features

Observation location features have geometry that records the location at which some data or sample was collected. The NGMDB implementation allows structural measurements, text notes, images, and samples to be associated with observation locations. Stations are point features that locate observation sites, SectionLine lines and AreaOfInterest polygons represent extended observation sites. SectionLines may represent traverses, which are observation sites located in the depicted MapHorizon. BoreHoleProjection and FlightlineGroundTracks represent observation sections that do not lie in the depicted map horizon; they are projected into the map plane to depict in a 2-D image. The implementation revolves around the concept of an observation section as a linear analog of a map horizon surface. The observation section is a line in three dimensional space that provides a reference frame in which three dimensional points may be located using a single coordinate, measured along the observation section line from some defined origin (e.g. depth below kelly bushing in a bore hole). In the NGMDB implementation, this reference frame is described informally (as text) in the CoordinateReferenceSystem field (*ObservationSection*, Figure 5).

- **Station**. Point represents an observation location. Has related structural measurements, text notes, images, samples…
- **SectionLine**. Line features that depict 3-D section line in a 2-dimensional (or 2.5-D if they have Z values associated with vertices along their length) view. Each subtype of SectionLine has an associated geodatabase object class that represents the full, three dimensional observation section. FlightlineGoundPath is the vertical projection of a flight line down to the map horizon. TraverseTrace is the trace of a surface traverse in the map horizon. BoreholeProjection is the vertical projection from a borehole to the map horizon.
- **AreaOfInterest**. Polygons that associate observation

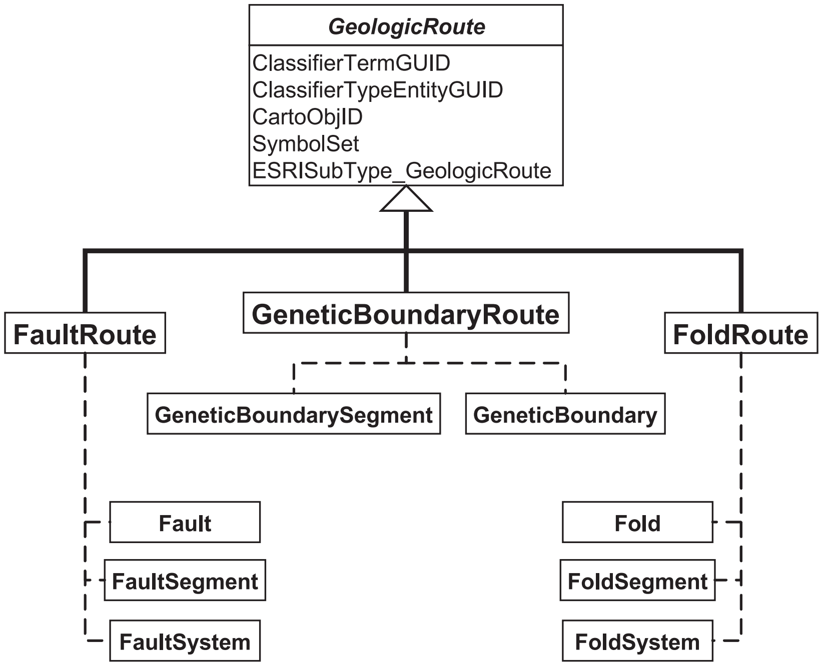


Figure 4. Geologic route feature classes. These feature classes are used to aggregate line segments (arcs) differentiated based on observation-related properties (information source, locatability, location uncertainty) into traces of surfaces that have identity and are bounded by intersection with other surfaces (Fault, Fold). These may then be aggregated into systems that represent compound structures that include multiple faults or folds.

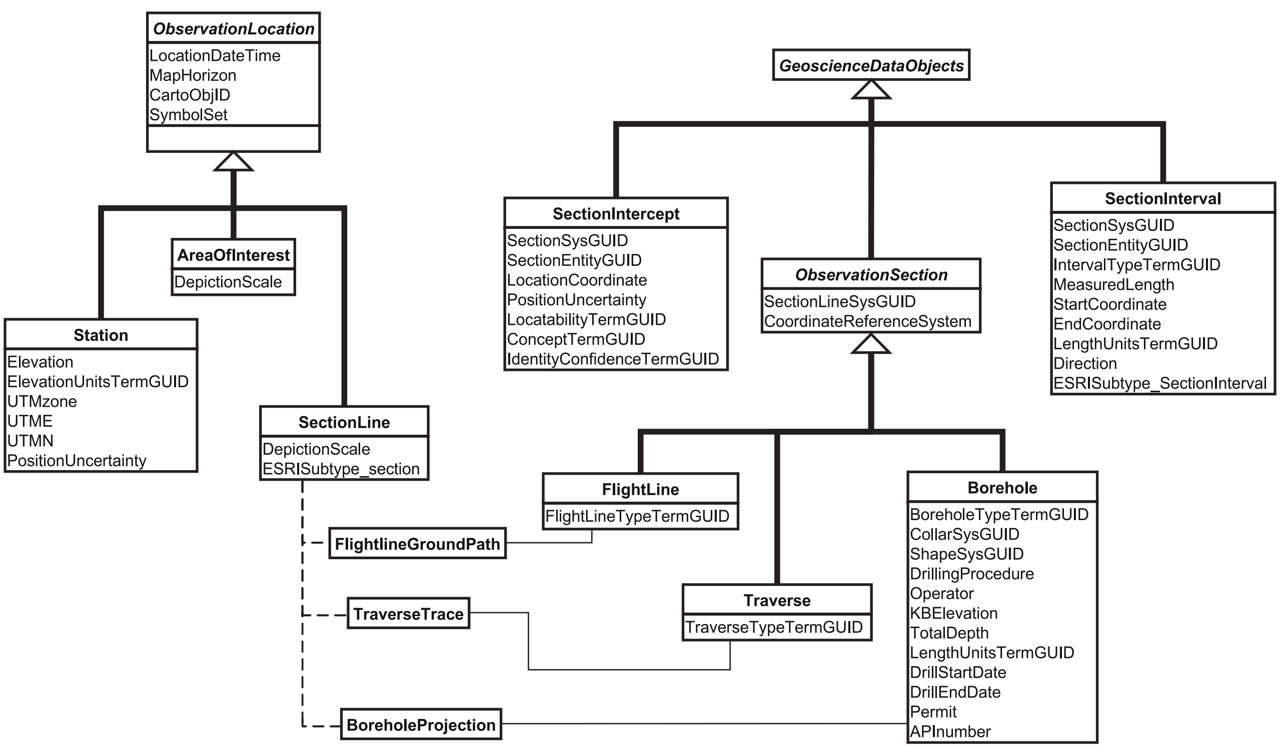


Figure 5. Observation location features. These (the classes on the left side of the diagram) are mapped features that have geometry that records the location at which some data or sample was collected. Stations are point features that locate observation sites. AreaOfInterest polygons represent an extended observation site in some map horizon. SectionLine lines are the projection of 3-D observation sections into a map horizon.

data with some area, for instance to represent that a bedding orientation value or some particular weathering character is consistent over some outcrop area.

Borehole Representation

Figure 6 shows the principal geodatabase object classes associated with BoreHoleCollar points to construct a representation of subsurface geologic information derived from boreholes. This model uses the same pattern to represent data from surface traverses (e.g. measured sections), and flight lines (or ship tracks…), but only the treatment of borehole sections is discussed here.

- **BoreholeCollar**. Point represents collar location from which a borehole was drilled. Generally is intersection of borehole with the Earth surface. Has one or more associated borehole object(s), which are not geodatabase feature classes.

One or more boreholes may be associated with a single BoreHoleCollar point, because a collar may be reentered to drill a splay or side track, which is considered a separate borehole. Individual boreholes (or any observation sections) have collections of associated intervals (SectionInterval table) and intercepts (SectionIntercept table) that may have associated description data. Intervals have a top and bottom, and represent rock volumes, while intercepts have a single coordinate location in the observation section, and represent the intersection of a surface with the section, or a point location in the earth. The NGMDB implementation allows classification of an interval (or an intercept, link not shown in Figure 6) through an ObservationRelationship instance, analogous to mapped occurrence alternative classification (Figure 7). This allows assignment of intervals to geologic units, lithology classes, or any other sort of classification (e.g. aquifer), and association of intercepts with multiple surface classifications. StructureObservation (orientation) data, free text descriptions, and samples may also be associated with intervals or intercepts.

Classification And Descriptions Associated With Features

The NGMDB implementation includes a link to a classifier term in each geodatabase feature class for MappedOccurrences (Figure 1). This primary classification records the original intention with which the mapped feature was delineated, and is thus considered to inhere more strongly with the geometric description than other possible classifications. For GeologicUnitOutcrops, the default or primary classifier will be a GeologicUnit term. Outcrop trace and observation features will typically be classified using terms from a science vocabulary (ScienceLanguage table). Alternative classifications are used for derivative maps or analytical processes, and are produced by grouping the primary classifiers to define a different classification scheme. Such derivative or alternative classifications are represented using classification-typed data instances in the ObservationRelationship correlation table (Figure 7).

Any of the various kinds of descriptions implemented in the NGMDB database may have an associated feature that specifies a geographic extent over which the properties in the description are asserted to hold (Figures 8-10). The feature assigned as the extent for a description may be an ObservationLocation feature or any of the various MappedOccurrences.

Annotation

These are locations in a map view used to position graphical elements for cartographic display. The location of the symbol is related to some geologic feature, but its actual positioning is based on cartographic consideration. In the ArcMap v.9 implementation, annotation that is a text string used to provide supplemental information for spatial objects (point, line, polygon) is best represented using feature-linked annotation feature classes that are built into the Geodatabase. For symbols that are located by points, and identified by a symbol identifier (CartoObjID), the implementation includes a PointAnnotationFeature class. This is subtyped for different kinds of annotation based on relationships to other feature classes.

- **MapAnnotationPointAnno**. Text label indicates dip or plunge value of linked StrikeDip symbol, fault dip symbol, contact dip symbol, or hinge surface orientation symbol. ESRI annotation feature class, related to MapAnnotationPoint features.
- **GeologicUnitOutcropAnno**. Text label annotation. Text string, value is linked to label field in GeologicUnitOutcrop. ESRI annotation feature class.
- **MapAnnotationPoint**
    - StrikeDipPoint. Point feature that positions symbols representing orientation measurements not associated with a surface trace (may cover a Station feature). Has associated structural observation data (not a geodatabase feature class). StructureObservation instances are located by association with a station; this is a one (station) to many (observations) association, and the station may have many other related data objects (samples, images, text notes, physical property measurements). Positioning of the symbol on the map is dictated by cartographic considerations in addition to the actual location at which the measurement was made. Symbol rotates according to value in azimuth field, and has dip or plunge value text label from the label field in the associated structural observation. In the database the symbol information is maintained separately from the location of the station data. Has linked dip (or plunge) label text. See the cluster of three measurements made in the Xp (Pinal Schist)

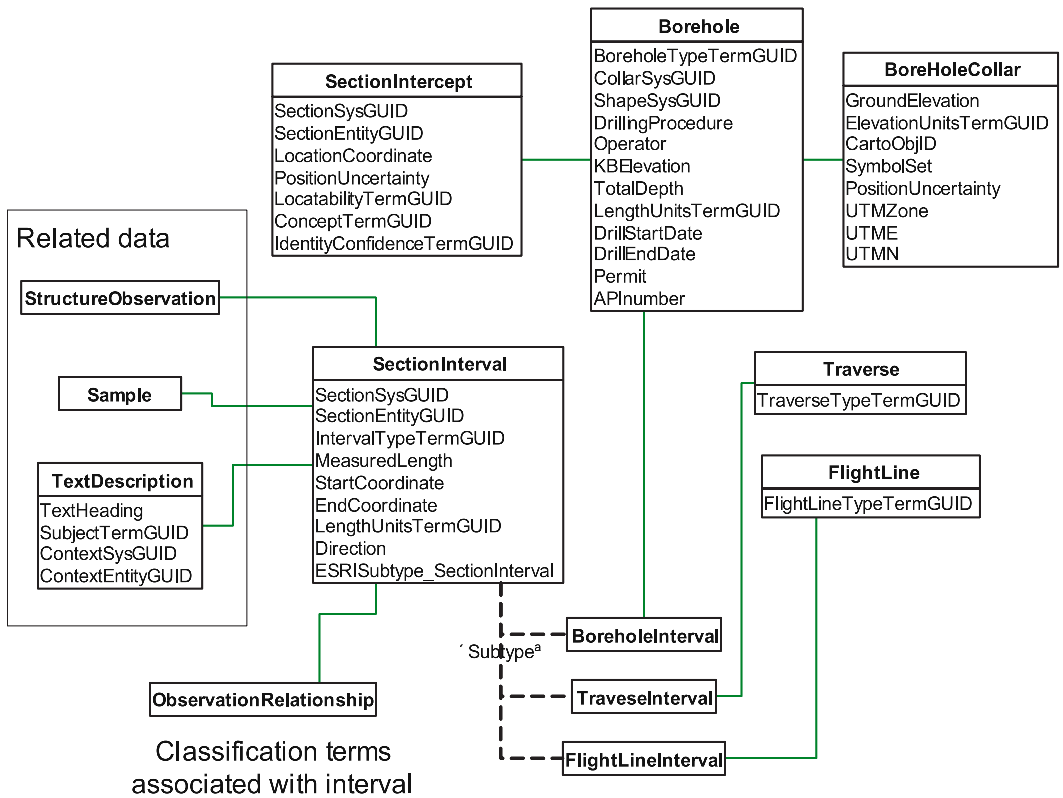


Figure 6. Observation sections: borehole implementation. BoreHoleCollar is a point feature class in the geodatabase that represents the intersection of a borehole with a map horizon. Borehole is a geodatabase object class. The ShapeSysGUID field in Borehole is an NGMDB-defined link to a three dimensional description of the borehole geometry. This is a hook to provide compatibility with three-dimensional models. Borehole is a kind of ObservationSection (Figure 5, see also https://www.seegrid.csiro.au/twiki/bin/view/Xmml/ BoreHole, and discussion in text here).

unit in the demo data map (Figure 11). When the structure data are displayed as an event theme on the StructureObservation table, the symbols are on top of each other. The StructureObservation events were exported to a shape file, then loaded into the MapAnnotationPoint feature class, and the point locations adjusted so that all three symbols are discernible on the map.

- FaultDisplacementPoint. Points that position displacement symbols along a fault trace according to cartographic considerations. Examples of symbols located by these points include bar and ball symbol, right and left arrows, or 'U' and 'D' used to indicate slip or separation on a fault. These should move with fault trace if its geometry is edited, and rotation of symbol is dictated by local trend of fault trace.
- SurfaceOrientationPoint. Points that position orientation symbols along a geologic surface trace, for which symbol is oriented perpendicular to surface trace at the point location. Dip arrows show orientation of a geologic surface, location of point is dictated by observation location and cartographic considerations. Linked to surface trace, should move if geometry of trace is modified. Rotation of symbol is determined by trend of surface trace at point location. Has linked dip label text.
- SurfaceTraceLinkedStrikeDipPoint. Points that position orientation symbols along a geologic surface trace, for which symbol is oriented according to a StructureObservation associated with the surface (not the local trend of the surface trace). The orientation of a gently dipping geologic surface may be annotated by an arrow for which the azimuth of the arrow is determined by strike measurement on the surface (from a structural observation), which may be discordant to the local surface trace trend because of topography. The orientation of the fold hinge for a map-scale fold may vary along a hinge surface trace, and may be discordant to the local trend of the surface trace; thus hinge line measurements linked to a fold hinge surface may be rotate according to structure observation data, not the local trend of the

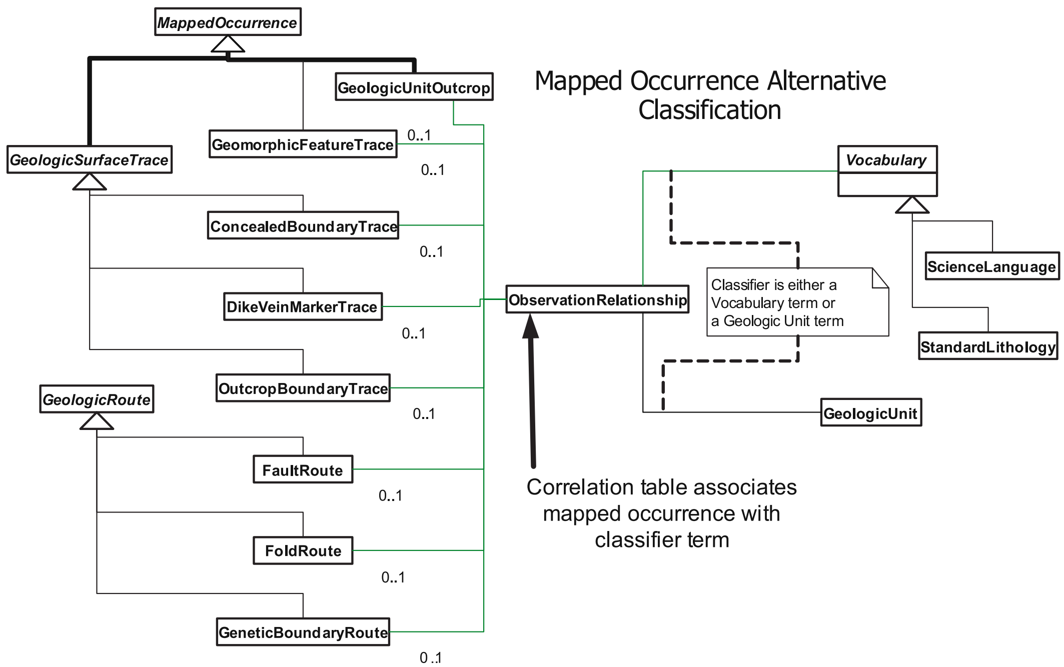


Figure 7. Observation relationship links for alternative classification of mapped features. Default (primary) classification is assigned by ClassifierTermGUID/ClassifierEntityGUID attribute tuple in the feature class table. The primary classification records the original intention with which the mapped feature was delineated. Alternative classifications are used for derivative maps produced by grouping the primary classifiers to define a different classification scheme. The ObservationRelationship correlation table uses sysGUID/entityGUID pairs to identify the type of feature class and actual data instance for the source and the type of classifier and particular classifier term for the target of the relationship. ObservationRelationship also has properties on the relationship (classification) instance to assign a confidence and other metadata information.

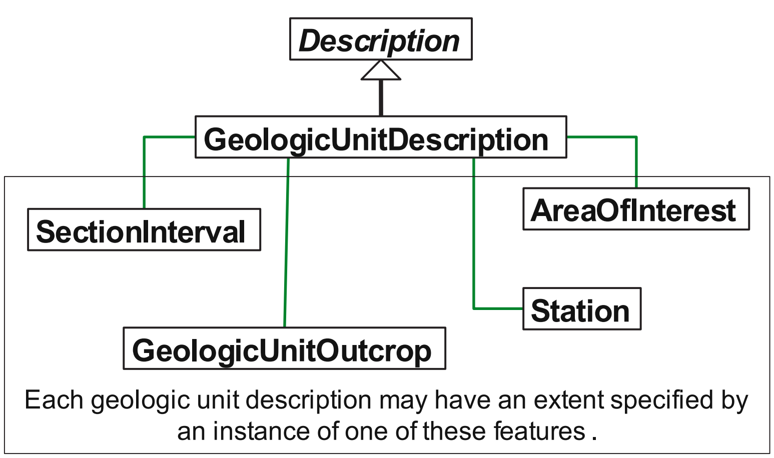


Figure 8. Geographic extents associated with geologic unit description. The extent associated with a description indicates the geographic region within which the description is asserted to be valid. The associations shown are implemented as geodatabase relationship classes, allowing navigation from mapped features (e.g. GeologicUnitOutcrop) to associated descriptions in the ArcMap interface. The foreign key from GeologicUnitDescription to the mapped feature is DescriptionExtentSysGUID, which links to the primary key (SysGUID) in the mapped feature. Details of GeologicUnitDescription are discussed in Richard et al. (2004).

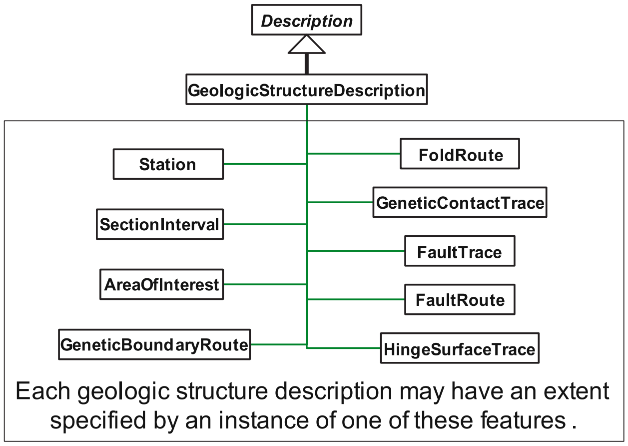


Figure 9. Geographic extents associated with geologic structure description. The extent associated with a description indicates the geographic region within which the description is asserted to be valid. The associations shown are implemented as geodatabase relationship classes, allowing navigation from mapped features (e.g. FaultTrace) to associated descriptions in the ArcMap interface. The foreign key from GeologicStructureDescription to the mapped feature is DescriptionExtentSysGUID, which links to the primary key (SysGUID) in the mapped feature. Details of GeologicStructureDescription are discussed in Richard et al. (2004).

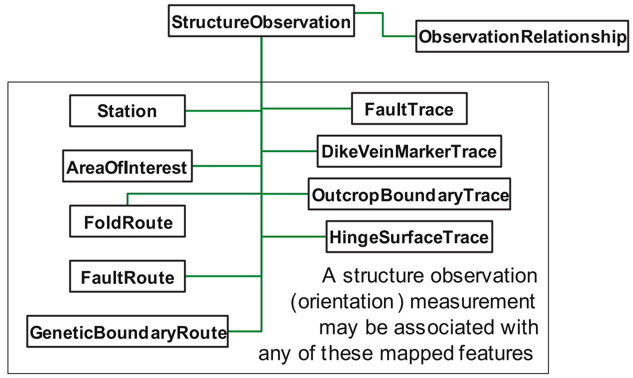


Figure 10. Structure observation description extent. Structure observation is a geodatabase ‘object class’ that is used to record orientation measurements for geologic structures. The extent associated with a structure observation indicates the geographic region within which the orientation measurement is asserted to be valid. The associations shown are implemented as geodatabase relationship classes, allowing navigation from mapped features (e.g. Station) to associated orientation measurements in the ArcMap interface. The foreign key from StructureObservation to the mapped feature is LocationSysGUID, which links to the primary key (SysGUID) in the mapped feature. Details of StructureObservation attributes are discussed in Richard et al. (2004). ObservationRelationship links to StructureObservation are used to correlate related observations (e.g. foliation in lineation, cleavage axial planar to fold).

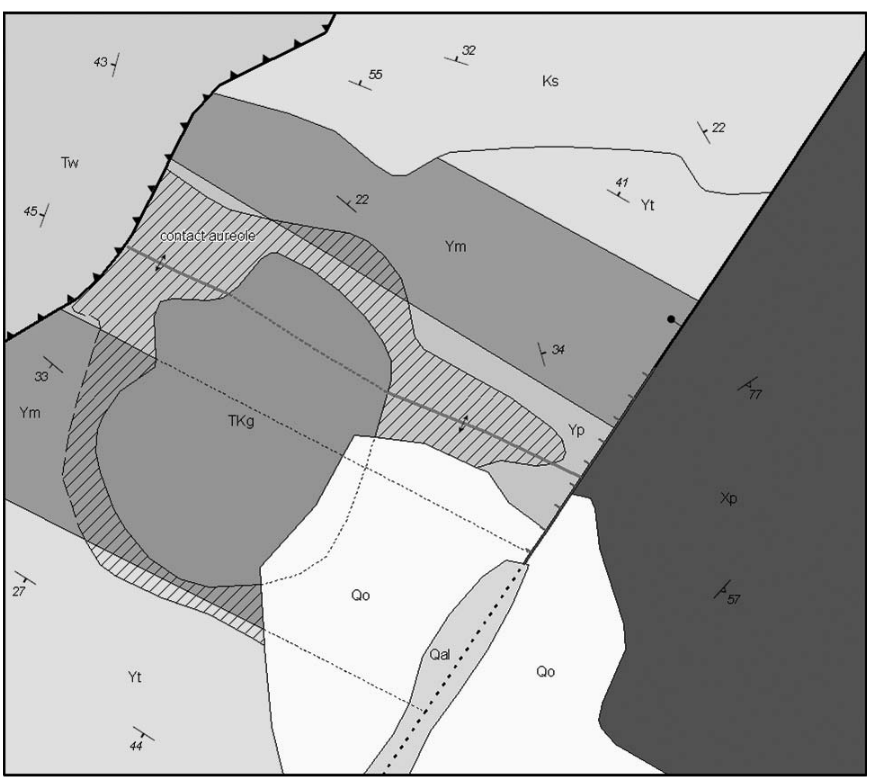


Figure 11. Geoland demonstration map. Original is in color, converted to grayscale for display in print.

hinge surface trace. Has linked dip (or plunge) label text. Location of symbol generally determined by an observation location (station), adjusted according to cartographic considerations, but should move if the associated surface trace is edited.

  - FoldGeometryPoint. Points that locate symbols placed along a hinge surface trace to indicate the geometry of the fold—syncline, anticline, overturned syncline, etc. Location determined by cartographic considerations. Should move with the surface trace if it is edited. Topology rule: must be covered by HingeSurfaceTrace.

Orientation data

Structure observation is a geodatabase 'object class' that is used to record orientation measurements for geologic structures. The extent associated with a structure observation indicates the geographic region within which the orientation measurement is asserted to be valid (Figure 10). The most common application is the association of common bedding or foliation measurements with a Station feature. This model makes a clear separation between the observation location—a station, and the data acquired at the location (a structure observation). For cartographic purposes, for example to display a strike and dip symbol, a MapAnnotationPoint is created using the station location and the azimuth value to rotate the symbol and text label for the symbol are determined from the related structure observation data. The symbol can be repositioned for cartographic purposes without losing information about the actual location at which the measurement was made.

EXAMPLE MAP

Figure 11 is a geologic map of an imaginary study area, based on a demonstration geodatabase designed to exercise various capabilities of the NGMDB design. Many of the geologic unit names and general character are based on units found in Arizona, but this is just a convenience to make easier the generation of descriptions. Figure 12 is the ArcMap table of contents for layers displayed in Figure 11; each layer is described below:

- The 'Map boundary' layer in the ArcMap table of contents (Figure 12) is based on OutcropBoundaryTraces, with a definition query for the layer set to '[CartoObjID] = 65' to select the lines in that feature class that are symbolized with the line symbol used for the map boundary line. This layer is just the neat line around the boundary of the map area to provide a reference frame. The map boundary layer is shown in the subsequent figures for reference.
- The MapAnnotationPointAnno layer (Figure 13) displays an ESRI Annotation feature class from the geodatabase that contains the strike and dip numbers for structure data on the map. This layer was generated by using the auto label feature in ArcMap, then converting labels to feature-linked annotation, and adjusting the position of the labels.
- MapAnnotationPoint is a point feature class that is used to position point symbols associated with other features but whose location is not determined by an observation location. In this map view (Figure 13), the two anticline symbols along the fold hinge surface trace and the bar-ball symbol on the fault trace are located by these points.
- Station is a point feature class (turned off in this view) that displays the locations at which observation data were collected. All strike and dip symbols shown in the map view have associated stations.
- The GeologicUnitOutcropAnno layer (Figure 14) displays an ESRI annotation feature class from the geodatabase, generated by the same procedure as MapAnnotationPointAnno.
- In the ArcMap table of contents (Figure 12), geology is a group layer that includes the mapped occurrences displayed in the map view.
- StructuralObservationEvents displays structure measurements from the StructureObservation table (Figure 13), located using the X, Y coordinates stored in that table. In a complex map view composition, these might be exported to the MapAnnotationPoint feature class to allow repositioning of the symbols for cartographic purposes without disrupting the observation location points. The symbols are rotated using the Azimuth field in the StructureObservation table. The symbols are identified by the integer CartoObjID field. This is a default symbolization scheme determined by the data originator.
- GeomorphicFeatureTrace (Figure 15) displays a line feature class with the same name. Only one feature is included—the scarp along a section of the NE trending fault in the east side of the map. This line was drawn coincident with the fault using snapping.
- HingeSurfaceTrace (Figure 15) displays a line feature class with the same name. Three arcs are included along the hinge surface trace of the upright anticline in the central part of the map. Two are solid, one part is dotted through the intruding pluton. A FoldRoute.FoldSegment object could be constructed to represent this entire hinge surface trace segment if there was a need to describe the fold in its entirety.
- The 'Superimposed unit Outcrop' layer (Figure 16) is based on the GeologicUnitOutcrop feature class, with a definition query for the layer set to '[ESRISubType_GeologicUnitOutcrop] = 1', to display only the superimposed geologic unit polygons. In this case, the only polygon in this layer represents the contact aureole mapped around the pluton. It is symbolized with a hatch pattern with transparent background so the underlying primary geologic unit polygons are visible.
- OutcropBoundaryTrace (Figure 17) displays a line

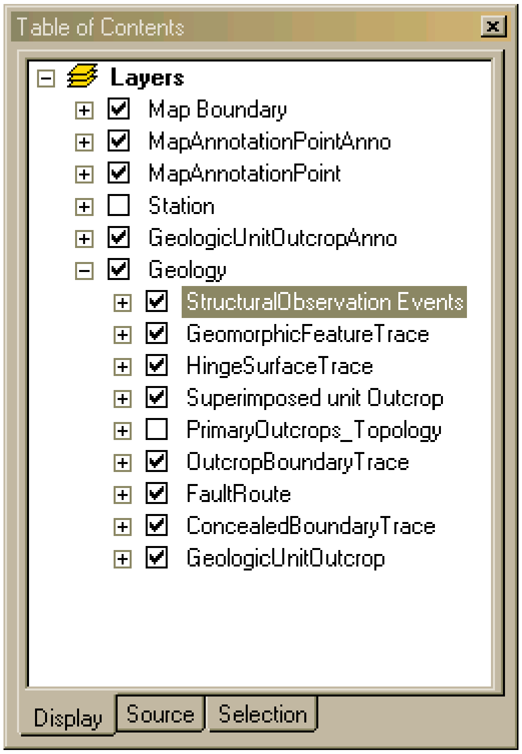


Figure 12. ArcMap table of contents showing layers displayed in map shown in Figure 11.

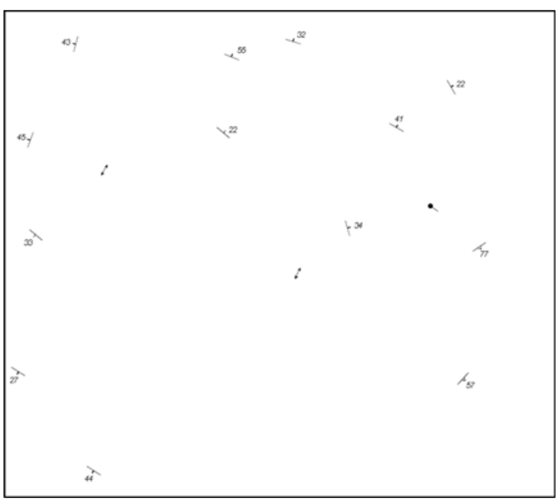


Figure 13. Map annotation points, StructureObservation events, and MapAnnotationPointAnno layers.

Figure 14. GeologicUnitOutcrop and GeologicUnitOutcropAnno layers.

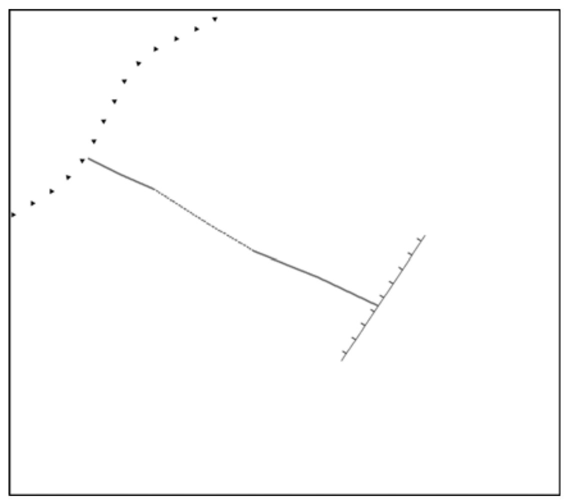


Figure 15. Geomorphic feature trace (fault scarp), hinge surface trace, and fault route layers. The fault route is used to place the triangular decorations along the thrust fault so that the spacing between triangles is regular, because the decorated line has no pseudonodes.

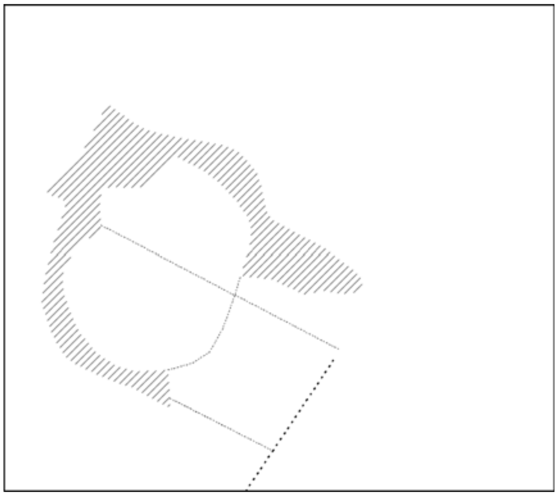


Figure 16. Superimposed geologic unit polygon and concealed boundary trace layers.

Figure 17. Outcrop boundary trace layer.

feature with the same name, symbolized on CartoObjID from a legend file that includes all line CartoObjIDs in the AZGSSymbolSet. These are the outcropping geologic contacts and faults in the map.

- FaultRoute (Figure 15) displays a line feature class with the same name. The features in this feature class aggregate the mapped parts of each fault trace into a single fault trace that may be associated with descriptions of the fault. The route features are also useful for symbolization. Labeling of the fault using the route feature avoids duplicate labels, and decorated faults (e.g. thrust, detachment, low-angle normal) can have the decorations applied using the fault route to obtain a much more cartographically pleasing effect. In this map view, the southeastern, high angle fault route is not symbolized. The thrust fault in the NW part of the map has no line stroke, but has filled triangle ornaments to place the teeth along the fault trace that is drawn by the OutcropBoundaryTrace layer.
- ConcealedBoundaryTrace (Figure 16) displays a line feature class with the same name. This feature class includes concealed faults and outcrop boundary contacts. They are included in a separate feature class because they do not participate in the polygon topology.
- GeologicUnitOutcrop (Figure 14) displays a polygon feature class with the same name that represents the outcrop of primary geologic units. This layer is based on the GeologicUnitOutcrop feature class, with a definition query for the layer set to '[ESRISubType_Geologic UnitOutcrop] = 0'. Colors in this map are assigned based on the Label field in the GeologicUnitOutcrop table. This is an expedient approach to symbolization during initial map compilation, when the final set of geologic units may not be known for sure.

OUTSTANDING ISSUES

Major implementation questions that must be addressed in developing a multiple map database, of the sort envisioned for a geological survey enterprise archive, include:

- how to implement multiple-scale representation. At different levels of generalization, a single geologic entity may be represented by different geometry. Topology rules must be applied to features that represent the same resolution.
- how to account for multiple map horizons (e.g. Earth surface, bedrock surface, top Precambrian…)
- how to account for different geologic perspectives that have related but loosely coupled geometric and topologic relationships. For example, geologic units on an alteration map may overlap some protolith geologic units and faults that cut them, while being cut by post-alteration structures and overlapped by post alteration geologic units that would share geometry with the stratigraphic geologic unit map.

One approach that we are experimenting with would group features into different feature datasets to represent different map horizons or resolution (generalization). Related, but loosely coupled geologic features like alteration map units would be represented using subtypes of the GeologicUnitOutcrop polygons, perhaps with a separate set of topologic rules from the principal unit outcrop boundary traces and outcrop polygons.

REFERENCES

Brodaric, B., and Hastings, J., 2001, Evolution of an Object-Oriented, NADM-Based Data Model Prototype for the USGS National Geologic Map Database Project [web page, abstract]: Annual Conference of the International Association for Mathematical Geology, IAMG2001, Cancun, Mexico, accessed at http://www.kgs.ku.edu/Conferences/IAMG/Sessions/I/brodaric.html.

Brodaric, Boyan, and Hastings, Jordan, 2002, An object model for geologic map information, in Richardson, D., and van Oosterom, P., eds;, Advances in Spatial Data Handling, 10th International Symposium on Spatial Data Handling: Heidelberg, Germany, Springer-Verlag.

Brodaric, B., Journeay, M., Talwar, S., and others, June 18, 1999, CordLink Digital Library Geologic Map Data Model Version 5.2 (Web Page), available at http://cordlink.gsc.nrcan.gc.ca/cordlink1/info_pages/English/dm52.pdf., accessed June 13, 2001.

Cox, Simon, Daisey, Paul, Lake, Ron, Portele, Clemens, Whiteside, Arliss, eds., 2004, OpenGIS® Geography Markup Language (GML) v. 3.1.0, Implementation Specification: OpenGIS® Recommendation Paper, Document OGC 03-105r1, ISO/TC 211/WG 4 Document 19136, 02-07-2004, 601 p., accessed at http://www.opengeospatial.org/specs/?page=recommendation.

Johnson, B.R., Brodaric, B., and Raines, G.L., 1998, Digital Geologic Maps Data Model, V. 4.3 (Web Page): unpublished report by AASG/USGS Geologic Map Data Model Working Group, accessed at http://www.nadm-geo.org/dmdt/.

NADMSC (North American Data Model Steering Committee), 2004, NADM conceptual model 1.0, A conceptual model for geologic map information: U.S. Geological Survey Open-File Repot 2004-1334, 60 pages text, 1 Adobe Acrobat (pdf) file, accessed at http://pubs.usgs.gov/of/2004/1334.

Richard, S.M., 2003, Geologic map database implementation in the ESRI Geodatabase environment, in Soller, D.R., ed., Digital Mapping Techniques ‘03—Workshop Proceedings, U.S. Geological Survey Open-File Report 03-471, p. 169-183, accessed at http://pubs.usgs.gov/of/2003/of03-471/richard2/index.html.

Richard, S.M., and Orr, T.R., 2001, Data structure for the Arizona Geological Survey Geologic Information System-Basic Geologic Map Data, in Soller, D.R., ed., Digital Mapping Techniques ‘01—Workshop Proceedings, U.S. Geological Survey Open-File Report 01-223, p. 167-188, accessed at http://pubs.usgs.gov/of/2001/of01-223/richard2.html.

Richard, S.M., Craigue, J., Soller, D.R., 2004, Implementing NADM C1 for the National Geologic Map Database, in Soller, D.R., Editor, Digital Mapping Techniques ‘04-Workshop Proceedings, U.S. Geological Survey Open-file Report 2004-1451, p. 111-144, accessed at http://pubs.usgs.gov/of/2004/1451/pdf/richard.pdf.

Soller, D.R., Brodaric, Boyan, Hastings, J.T., Wahl, Ron, and Weisenfluh, G.A., 2002, The central Kentucky prototype: An object-oriented geologic map data model for the National Geologic Map Database: U.S. Geological Survey Open-File Report 02-202, 39 p., available at http://pubs.usgs.gov/of/2002/of02-202/.

Ensuring Data Quality using Topology and Attribute Validation in the Geodatabase

By Heather I. Stanton[1], Stephanie O'Meara[1], James R. Chappell[1], Anne R. Poole[2], Gregory Mack[3], and Georgia Hybels[4]

[1]Colorado State University/National Park Service Cooperator
1201 Oak Ridge Drive, Suite 200
Fort Collins, CO 80525
Telephone: (970) 267-2159
e-mail: Heather_Stanton@partner.nps.gov,
Stephanie_O'Meara@partner.nps.gov, Jim_Chappell@partner.nps.gov

[2]United States Forest Service – Chippewa National Forest
e-mail: apoole@fs.fed.us

[3]National Park Service – Pacific West Region, Seattle Office
e-mail: Greg_Mack@nps.gov

[4]University of Denver/National Park Service Cooperator
e-mail: Georgia_Hybels@partner.nps.gov

INTRODUCTION

Data on paper geologic maps have certain, common-sense rules that govern how the data appear, such as spatial coincidence between certain features, and defined and limited attribution, as depicted in the map's legend (Figure 1). On a geologic map, faults are sometimes coincident with geologic contacts, and geologic contacts are always coincident with the boundaries of geologic units. A geologic unit is accompanied by a unit symbol appearing on the map, and units sharing the same symbol can also be assumed to share the same unit name, major lithology, and age, among other attributes. Features such as geologic contacts and faults have limited and defined lists of positional accuracy (for instance; known, approximate, or concealed).

Although these are common-sense rules, they are nonetheless important to retain and follow, particularly when translating the data on the paper map to a digital format. For instance, a small deviation from exact spatial coincidence between a fault and a geologic contact could have meaning – did the author intend to represent a fault that was 10 meters away from the geologic contact, or should one infer that they are coincident because at the printed map scale they appear to be coincident?

The National Park Service's (NPS) previous data model stored geologic GIS data in ESRI coverage and shapefile formats (O'Meara, et al, 2005a; O'Meara, et al, this volume). Spatial coincidence and attribute validity were ensured through the use of coverage topology (the manner in which geographic data are spatially interrelated), tables of appropriate values (domains) for certain attributes, and data capture methodology (including the use of Arc Macro Language (AML) programs designed to find and/or fix problems). However, over the life-cycle of creating digital geologic data through digitizing, editing, and quality checking (QC), it was difficult to maintain attribute validity and spatial coincidence where it was appropriate. It was not uncommon to find errors in coincidence and attribution after completion of a digital geologic map.

New methods of ensuring the validity of attributes and spatial coincidence, where appropriate, are now available with ESRI's latest software, ArcGIS, and its new format for storing geographic and tabular data, the geodatabase. Geodatabase topology can mimic the rules available with coverages and has additional rules that were previously unavailable, including those that relate data between different geographic layers, stored as feature classes within the geodatabase. Feature classes can also be subdivided using subtypes, or breaks in the feature class based on integer values stored in a field in its associated attribute table. These subtypes can be used to enforce different rules for attribution or topology for different parts of the feature class. In addition, attribution can be controlled by linking domains of acceptable values to selected fields in the feature class and by associating

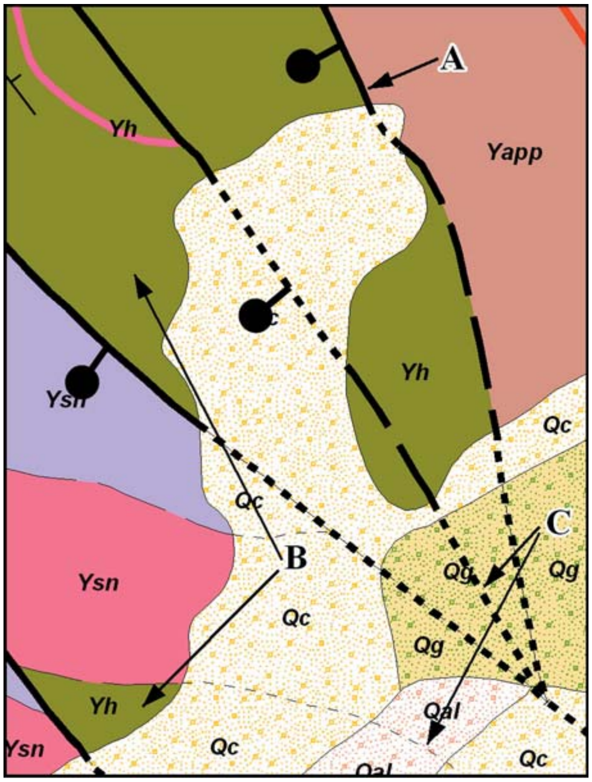


Figure 1. Excerpt of Digital Geologic Map of Glacier National Park and Vicinity (O'Meara, et al, 2003) illustrating common sense geologic map rules: A) Faults and Geologic Contacts are sometimes coincident; Geologic Contacts are always coincident with Geologic Unit boundaries. B) Geologic Units with the same unit symbol share defined ages, lithologies, etc. C) Faults and Geologic Contacts have limited and defined values for positional accuracy.

related tables of ancillary information using a key field through relationship classes, thereby reducing duplication of information throughout the geodatabase.

TOPOLOGICAL RULES IN THE GEODATABASE

To demonstrate the usefulness of ArcGIS geodatabase topological rules, we organize the rules into three classes: rules that mimic coverage topology, intra-feature class rules (within a feature class), and inter-feature class rules (between feature classes). All feature classes participating in the topology must be stored in a feature dataset, guaranteeing that they share the same spatial domain. Within the feature dataset, a topology is created where rules governing the interaction within and between feature classes are stored. Rules can then be validated in ArcCatalog or ArcMap. If any feature or features violate the topological rules, an error will be created. Errors can then be corrected using topological editing tools in an ArcMap edit session.

Examples of rules that mimic coverage topology include "Must Be Covered By Boundary Of", "Boundary Must Be Covered By" and "Must Not Have Dangles" (Figure 2). In coverages, polygonal features are stored together with their bounding arcs (lines), and the topology is inherent in the dataset. The attributes of the polygon features are stored in a .pat file (polygon attribute table) and the attributes of the associated arcs are stored in an .aat file (arc attribute table). An arc in a polygon coverage that ends without touching another arc (a dangle) is an error because its associated polygon(s) are not completely bounded by lines.

In the geodatabase, attributed line boundaries of polygons must be stored in a feature class separate from the polygon feature class, requiring that topological rules be created to maintain coincidence between the polygon features and their boundary lines. The "Must Be Covered By Boundary Of" rule ensures that all lines in a line feature class coincide with the boundaries of associated polygons,

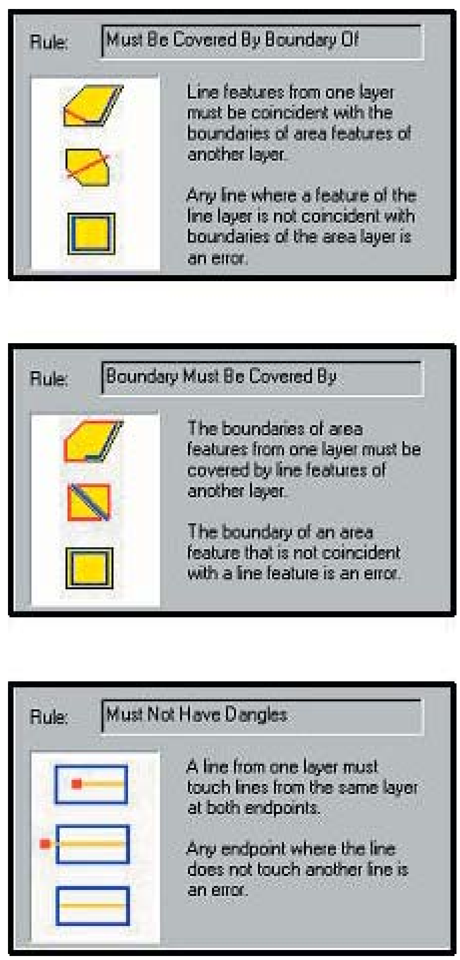


Figure 2. Screenshots from ArcGIS topology properties, illustrating the "Must Be Covered By Boundary Of", "Boundary Must Be Covered By" and "Must Not Have Dangles" rules. These are examples of topological rules that mimic the topology inherent in coverages.

and the “Boundary Must Be Covered By” rule ensures that all polygons have lines that coincide with their boundaries. The “Must Not Have Dangles” rule ensures that lines in the line feature class touch at least one other line at their endpoints, so that their associated polygons are completely bounded by lines. Following these rules is especially important when data are to be exported to coverages.

Intra-feature class rules include “Must Not Overlap”, “Must Not Have Gaps”, “Must Not Intersect or Touch Interior” and “Must Not Self-Intersect” (Figure 3). In order to maintain a complete mosaic of polygons in a polygon feature class where polygons do not overlap each other or have gaps between them, the rules “Must Not Overlap” and “Must Not Have Gaps” are used. “Must Not Intersect Or Touch Interior” ensures that lines do not intersect or touch other lines in the same feature class without the line being broken at the point of intersection, whereas “Must Not Self-Intersect” ensures that a line feature does not intersect itself. These rules can also be used to ensure that, when exported to coverages, polygons do not violate polygon coverage topology.

Inter-feature class rules demonstrate the real power of geodatabase topology, allowing feature classes within a feature dataset (or geographic grouping of feature classes) to maintain defined spatial relationships, such as coincidence (Figure 4). “Must Be Covered By Feature Class Of” is a rule that maintains coincidence between two feature classes, whereas “Must Not Overlap With” ensures that features from two different feature classes that should not be coincident are not. “Point Must Be Covered By Line” is used to ensure that point features from one feature class are coincident with lines in another feature class. While coincidence can be checked using other methods when data is stored in coverages and shapefiles, it cannot be enforced. The automated methods of topological validation and the tools provided by ArcMap to edit topological

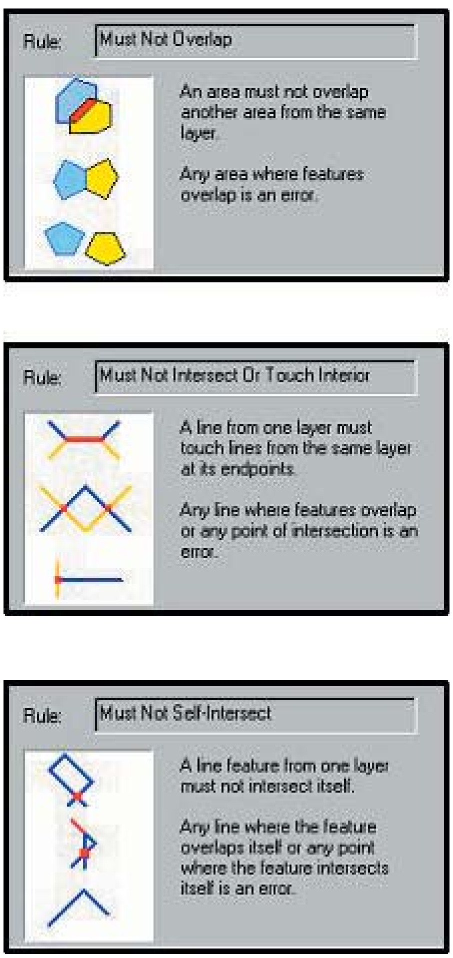


Figure 3. Screenshots from ArcGIS topology properties, illustrating the “Must Not Overlap”, “Must Not Intersect Or Touch Interior”, and “Must Not Self-Intersect” rules. These are examples of intra-feature class topological rules, applicable to features within a single feature class.

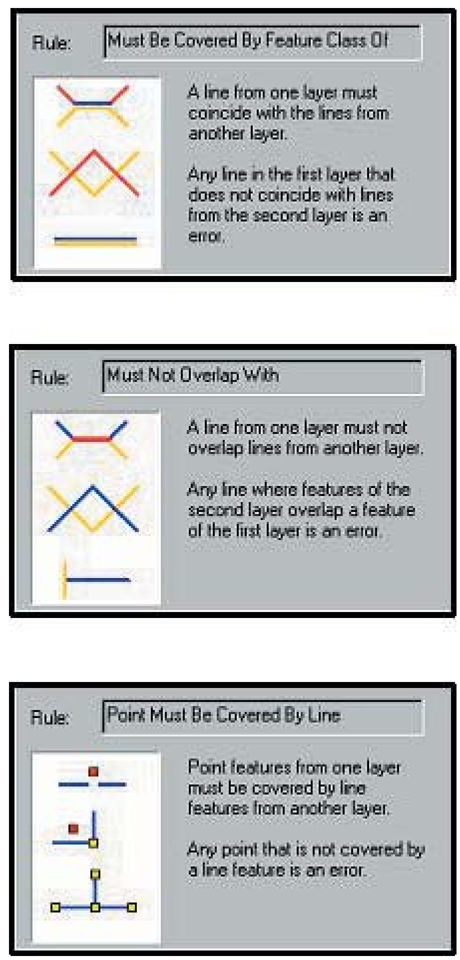


Figure 4. Screenshots from ArcGIS topology properties, illustrating the “Must Be Covered By Feature Class Of”, “Must Not Overlap With” and “Point Must Be Covered By Line” rules. These are examples of inter-feature class topological rules, making it possible to ensure spatial relationships between feature classes.

errors ensure that errors can be found and fixed without requiring user-created scripts and tools.

TOPOLOGICAL RULES IN THE NPS GEOLOGY-GIS GEODATABASE DATA MODEL

In the NPS Geology-GIS Geodatabase Data Model (O'Meara, et al, 2005b, O'Meara et al, this volume) topological rules are used to enforce the common-sense spatial coincidence rules that apply to data on geologic maps, as well as to enforce attribution in certain feature classes. A summary of the rules present in the data model can be found in Figure 5. Rules can be grouped based on:

1. whether a line feature class is associated with a polygon feature class (geologic contacts and geologic units; rules C, D, and H),
2. line feature classes overlap with other line feature classes (geologic contacts and faults; rules A and B),
3. line feature classes are spatially unrelated to other line and polygon feature classes (folds; rule G),
4. polygon feature classes are allowed to have gaps (surficial units, or recent deposits covering bedrock geologic units; rule G), or
5. not allowed to have gaps (bedrock geologic units, rule F).

This approach allows for the application of a set of topological rules to a feature class based on spatial/geometric relationships to other feature classes, rather than being specific to a single feature class. For instance, fold axes, structural contours, and glacial feature lines (such as moraine crests) have the same topological rules applied to them because they are spatially unrelated to other geologic features on the map; while surficial units, deformation areas, and dike swarm areas can have the same rules applied to them because they all have line feature classes that are coincident with their boundaries and they are allowed to have gaps between their individual polygons. Point feature classes, such as attitude measurements or symbology, have topological rules associated with them only if they are associated with line feature classes such as faults or folds.

Topological rules in the NPS Geology-GIS Geodatabase Data Model are often used in association with subtypes. Subtypes subdivide elements in a feature class into groups so that different rules for attribution or topology can be applied to those groups. For instance, faults in the data model are subdivided into either a "Fault" subtype or a "Fault/Contact" subtype. The "Fault" subtype refers to lines in the faults feature class that are not coincident

	Geologic Contacts	Faults	Surficial Contacts	Linear Dikes	Folds	Geologic Units	Surficial Units
Geologic Contacts	H	1-A-1, 0-B-0	-	-	-	C	-
Faults	1-A-1	G	-	-	-	-	-
Surficial Contacts	0-B, 2-A-0, 3-A-1	0-B, 1-A-0	H	-	-	-	C
Linear Dikes	0-B, 2-A-0, 3-A-1	0-B, 1-A-0	-	G	-	-	-
Folds	-	-	-	-	G	-	-
Geologic Units	D	-	-	-	-	F	-
Surficial Units	-	-	D	-	-	-	F

Rules between feature classes that are specific to the geodatabase
Rules that apply only within a single feature class; mimic coverage rules
Rules between feature classes that mimic coverage rules

A Must Be Covered By Feature Class Of
B Must Not Overlap With
C Must Be Covered By Boundary Of
D Boundary Must Be Covered By
E General Polygon Rules: Must Not Overlap
F General Polygon Rules 2: Must Not Overlap, Must Not Have Gaps
G General Line Rules: Must Not Intersect Or Touch Interior, Must Be Single Part, Must Not Self-Intersect
H General Line Rules 2: General Line Rules Plus Must Not Have Dangles

Line Subtypes:
Faults - 0: Fault, 1: Fault/Contact
Geologic Contacts - 0: Contact, 1: Contact/Fault
Others - 0: Self, 1: Self and Fault, 2: Self and Geologic Contact, 3: Self and Fault and Geologic Contact

Figure 5. Topological rules in the NPS Geology-GIS Geodatabase Data Model are represented in the table with the source or origin feature classes in rows and the destination feature classes in columns (a similar figure is present in O'Meara, et al, this volume). For instance, the Faults feature class (source) is related to the Geologic Contacts feature class (destination) using the rule 1-A-1, where the first "1" is the Fault/Contact subtype (see Line Subtypes below table) of the Faults, the "A" represents the topological rule (see list on right), and the second "1" is the Contact/Fault subtype of the Geologic Contacts. In other words, the Fault/Contact subtype "Must Be Covered By the Feature Class Of" or must coincide with the Contact/Fault subtype. If there is no number to represent a subtype for the source and/or destination feature class, the rule applies to the entire source and/or destination feature class regardless of subtype. Note that one or more rules may apply to a given source and destination feature class.

with geologic contacts, whereas the “Fault/Contact” subtype refers to the lines that are coincident with geologic contacts. The geologic contacts feature class is similarly subdivided into “Contact” and “Contact/Fault” subtypes.

Subtypes, in association with topological rules, can help to identify errors in attribution in the feature classes to which they are applied. For instance, a line in the faults feature class may be incorrectly coded as a “Fault/Contact” subtype because there is no associated coincident geologic contact (Figure 6). When the data model topological rules are validated, this line will violate the “Must Be Covered By Feature Class Of” rule because it does not coincide with a “Contact/Fault” line in the geologic contacts. The resulting topological error would alert the user that a change in attribution (coding the line as a “Fault”) is needed.

Topological rules can also be used with subtypes to identify features within one feature class that should be coincident with features in another feature class. For instance, at a glance, a fault and a geologic contact may appear to be coincident at map scale, but at larger scales it may become apparent that they are not (Figure 7). In the realm of geologic GIS data, these discrepancies are crucial to identify, and the errors generated by the validation of topological rules help to locate the discrepancies and fix them. In the NPS Geology-GIS Geodatabase Data Model, lines coded as “Contact/Fault” in the geologic contacts feature class and lines coded as “Fault/Contact” in the faults feature class that are not coincident with each other will be flagged as errors because they violate the “Must Be Covered By Feature Class Of” rule.

Another advantage of setting up a topology in the geodatabase is that coincident features from different feature classes can be moved as one entity using the Topology Edit tool in ArcMap (Figure 8). An example of this is editing the location of a fault and its associated point symbology that is also coincident with a geologic unit boundary and a geologic contact. With coverages and shapefiles, or feature classes not participating in a topology, each of the coincident features would have to be moved separately. Even with the snapping function set, there is a risk of losing coincidence because each individual vertex in a line has to be moved separately, and it is difficult to identify areas that are no longer coincident since often the lack of coincidence is not visible without examining lines in great detail.

ATTRIBUTE VALIDATION IN THE NPS GEOLOGY-GIS GEODATABASE DATA MODEL

In the NPS Geology-GIS Geodatabase Data Model, the consistency and quality of attribution are not just controlled by topological rules, but also in two other ways. Domains (lists of acceptable values) and default values are used along with subtypes to define attribution for various fields (including positional accuracy, fault types, and attitude measurement types) in the feature classes in the data model. Also, relationship classes between feature classes and tables are used to avoid duplication of data and to limit attribution for certain fields (for instance, a geologic unit information table related to the geologic units feature class).

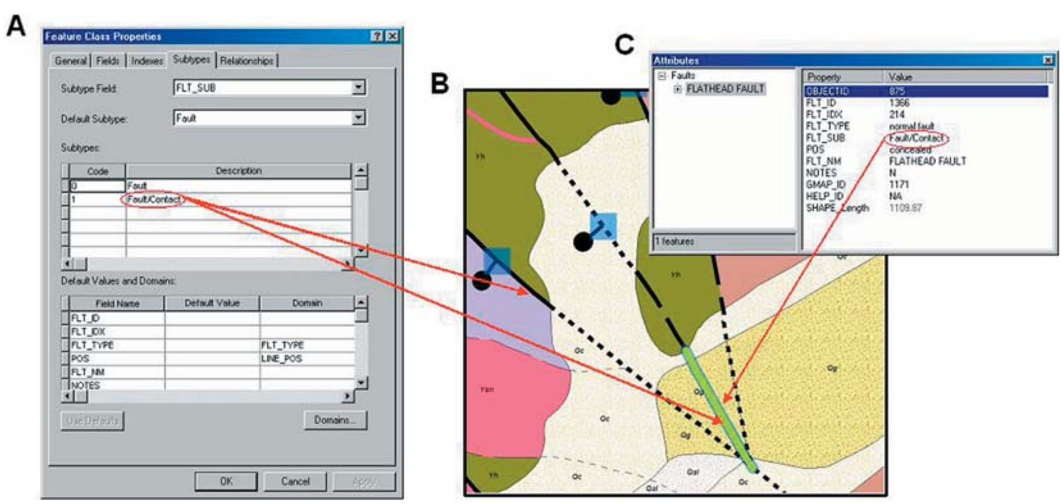


Figure 6. Excerpt of the Digital Geologic Map of Glacier National Park and Vicinity (O’Meara, et al, 2003). A) Feature class properties for the faults feature class, illustrating that “Fault” and “Fault/Contact” subtypes are stored in the FLT_SUB field. B) The fault highlighted on this map has violated the “Must Be Covered By Feature Class Of” rule because it is coded as a “Fault/Contact” and is not coincident with a geologic contact. C) The line should be coded as a “Fault” in the FLT_SUB field.

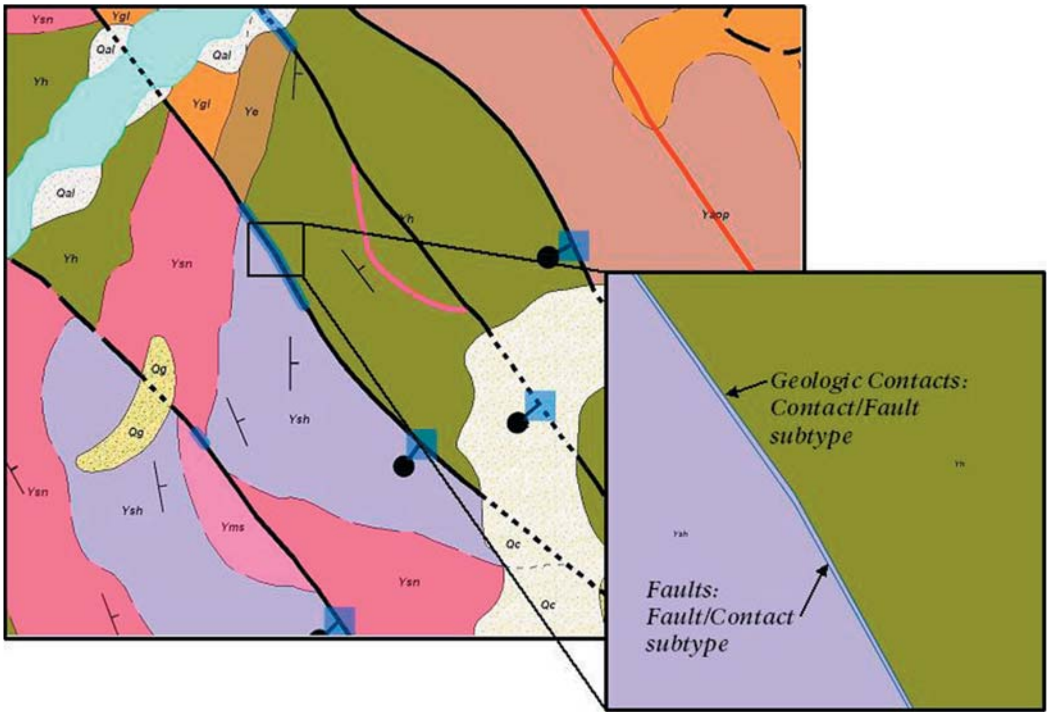


Figure 7. Excerpt of the Digital Geologic Map of Glacier National Park and Vicinity (O'Meara, et al, 2003). Both the fault and geologic contact present in this example have violated the "Must Be Covered By Feature Class Of" rule because, although they are coded correctly in their subtype fields, they are not spatially coincident. The problem can be fixed by editing the fault's vertices, snapping them to the geologic contact.

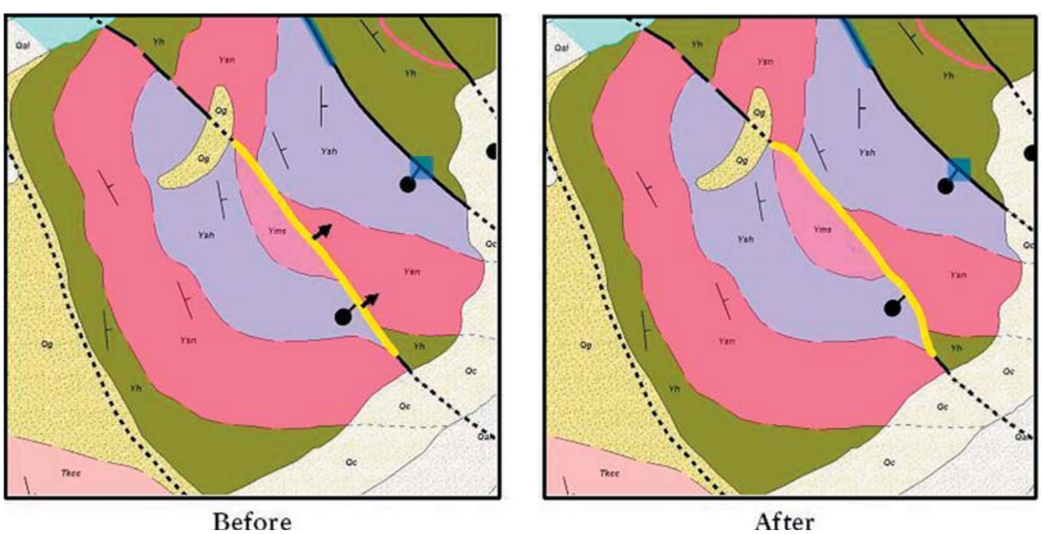


Figure 8. Excerpt of the Digital Geologic Map of Glacier National Park and Vicinity (O'Meara, et al, 2003). Without topology, in order to move the highlighted shared edge, each feature class (faults, geologic unit boundaries, geologic contacts and fault symbols) would have to be moved separately, risking lost spatial coincidence. Using the Topology Edit Tool, coincident features can be moved together.

Domains and default values are useful in a number of ways. In the NPS Geology-GIS Geodatabase Model, coded value domains contain an alias for each domain member, for instance, a positional accuracy of 1 is aliased to "known." A change in subtype in a given feature class changes the domains and default values that are accessed. For instance, when the subtype is changed during attribution to "Planar Measurements – Vertical" in the attitude measurements feature class, a domain restricted to vertical attitude types is accessed, the dip field is assigned a default value of 90 and a domain that has 90 (Vertical) as its only member is accessed. The default value provides automated attribution, and if a value other than 90 is later entered into the dip field, it will be flagged as an error during attribute validation because it is outside the domain for measurements with vertical dip.

For third-party users of the data, domains provide a built-in data dictionary in the geodatabase that is easily understood. Domains also provide ease of attribution by allowing the user to pick a description (strike and dip of beds) rather than a number referenced to an outside list of acceptable values. Default values increase this ease by automatically placing a commonly-used value in a field of the user's choice when a feature is created. During QC, domains are useful in yet another way. Using the built-in functionality of attribute validation in ArcGIS, values stored in a feature class's fields are compared to the domains assigned to those fields. If a value lies outside a domain because it was somehow misattributed, an error is generated and the feature(s) in error are identified (Figure 9) and can then be corrected.

In the NPS Geology-GIS Geodatabase Data Model, a geologic units table is used to store the name, age, description, and other information related to a given geologic unit. The unit information table is related to the geologic units feature class based on the geologic unit symbol using a relationship class. This avoids repeating multiple fields with the same information for every unit with the same map symbol in the feature class. During the initial attribution of geologic unit polygons, the table can be used to select which unit to apply to a series of polygons if desired. During attribute validation, the polygons in the geologic units feature class can be checked against the table for errors. If the geologic unit does not exist in the table, or if there are multiple entries in the table for a single unit on the map, an error will be generated, alerting the user to change attribution in the geologic units feature class or to edit the geologic units table (Figure 10).

CONCLUSION

The coverage/shapefile-based data model that was used prior to the development of the NPS Geology-GIS Geodatabase Data Model had a number of methods for ensuring attribute validity and coincidence. However, it was difficult to maintain these checks over the life cycle

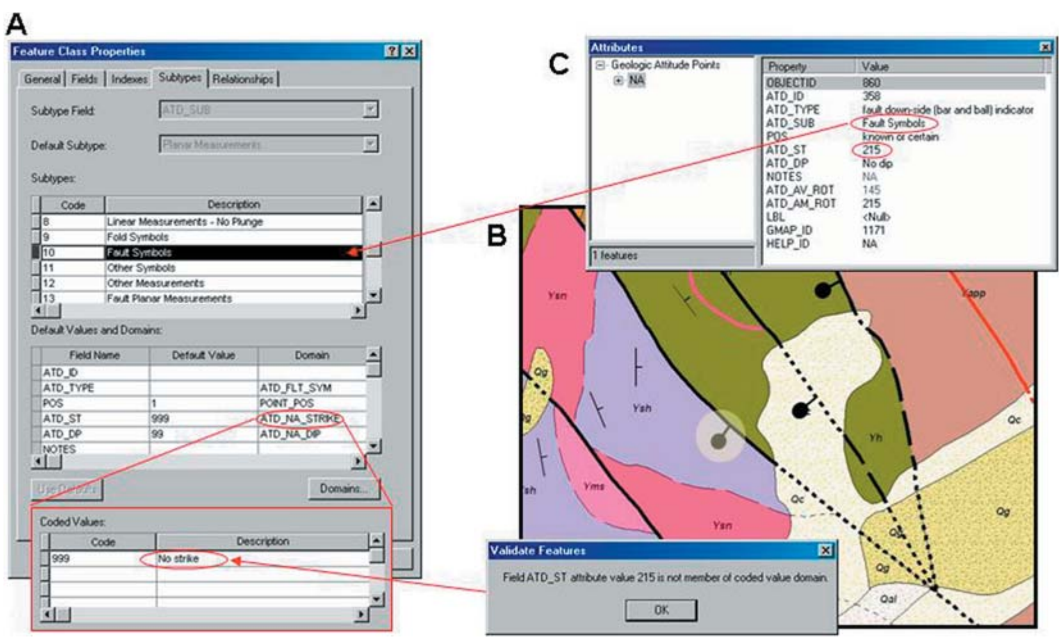


Figure 9. Excerpt of the Digital Geologic Map of Glacier National Park and Vicinity (O'Meara, et al, 2003). A) Feature class properties for bedding attitudes. Attribution for the "Fault Symbols" subtype is controlled by a number of domains, as well as default values. The value for ATD_ST (strike) for "Fault Symbols" should always be 999 or 'No strike'. B) When 'Validate Features' is carried out, the highlighted symbol, which is coded as a "Fault Symbols" subtype, is found to be in error because C) the ATD_ST (strike) value, 215, is not a member of the associated domain.

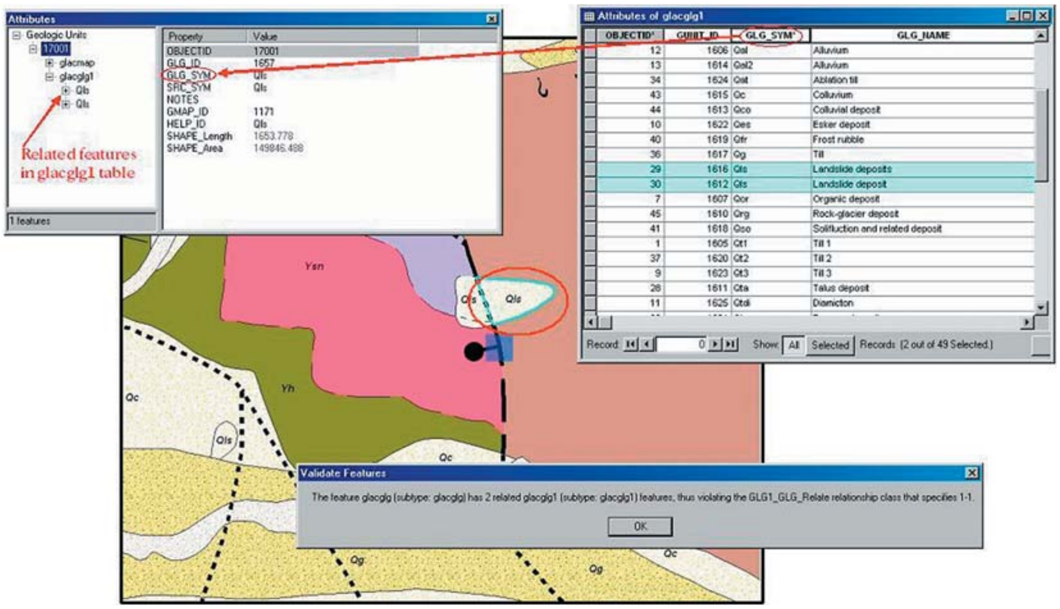


Figure 10. Excerpt of the Digital Geologic Map of Glacier National Park and Vicinity (O'Meara, et al, 2003). When 'Validate Features' is carried out on the circled geologic unit, an error occurs because there are two entries for the same GLG_SYM (unit symbol) in the related table. This violates the one-to-many simple relationship class set up between the geologic units and geologic unit information (glacglg1) table using GLG_SYM as the key field.

of a digital geologic map project. Geodatabases can not only reproduce coverage topology and programmatic methods for ensuring attribute validity, but they also have features that provide further assurances of spatial and attribute integrity. Attribute quality is improved through the use of subtypes in combination with domains and relationship classes that access records from related tables during feature capture. Spatial coincidence can also be ensured where appropriate, not only with rules mimicking coverage topology, but with added rules that interrelate different feature classes and the ability to apply rules to parts of feature classes using subtypes. Finally, as an added benefit, ArcGIS's automation for attribution and attribute and topology validation increase the ease and speed of attribution and QC for digital geologic maps.

In the future, we will continue to refine topological and attribution rules for the NPS Geology-GIS Geodatabase Data Model. We also will be creating programmatic methods for defining, and fixing, common attribution and topological errors. Finally, we will develop methods for updating and quality-checking our legacy data when migrating to the latest data model.

REFERENCES

O'Meara, S.A., Gregson, J., Poole, A.R., Mack, G., Stanton, H.I., and Chappell, J., 2005a, National Park Service Geologic Resources Evaluation Geology-GIS Coverage/Shapefile Data Model, available at http://science.nature.nps.gov/im/inventory/geology/GeologyGIS DataModel.htm.

O'Meara, S.A., Stanton, H.I., Chappell, J., Mack, G., Poole, A.R., and Hybels, G., 2005b, National Park Service Geologic Resources Evaluation Geology-GIS Geodatabase Data Model (Draft v. 1.2), available at http://science.nature.nps.gov/im/inventory/geology/GeologyGISDataModel.htm.

O'Meara, S.A., deWolfe, V., Johnson, R., and Schaeffer, M., 2003, Digital Geologic Map of Glacier National Park and Vicinity, Montana, digitized from Whipple (1992) and Carrara (1990): National Park Service Geologic Resources Evaluation (GRE) program, unpublished, 1:100,000 scale [digitized from Whipple, J.W., 1992, Geologic Map of Glacier National Park, Montana, U.S. Geological Survey Map I-1508-F, 1:100,000 scale; and Carrara, P.E., 1990, Surficial Geologic Map of Glacier National Park, Montana, U.S. Geological Survey Map I-1508-D, 1:100,000 scale].

SOFTWARE REFERENCES

ArcGIS 8.3 and 9.0 (ArcCatalog and ArcMap) – Environmental Systems Research Institute (ESRI) Inc., 380 New York St., Redlands, CA92373, http://www.esri.com.

Geodatabase Diagrammer – Developed by Michael Zeiler, Environmental Systems Research Institute (ESRI) Inc., 380 New York St., Redlands, CA 92373, see http://arcscripts.esri.com.

Geodatabase Designer v2 – Developed by Richie Carmichael, Environmental Systems Research Institute (ESRI) Inc., 380 New York St., Redlands, CA 92373, see http://arcscripts.esri.com.

Microsoft Office Visio Professional 2003, Microsoft Corporation, http://www.microsoft.com/.

Overview of Procedural Approach for Migrating Geologic Map Data and Related Processes to a Geodatabase

By James R. Chappell[1], Stephanie A. O'Meara[1], Heather S. Stanton[1], Greg Mack[2], Anne R. Poole[3] and Georgia Hybels[4]

[1]Colorado State University/National Park Service Cooperator
Colorado State University
Fort Collins, CO 80521
Telephone: (970) 491-5147
e-mail: Jim_Chappell@partner.nps.gov; Stephanie_O'Meara@partner.nps.gov; Heather_Stanton@partner.nps.gov

[2]National Park Service – Pacific West Region, Seattle Office
e-mail: Greg_Mack@nps.gov

[3]United States Forest Service – Chippewa National Forest
e-mail: apoole@fs.fed.us

[4]University of Denver/National Park Service Cooperator
e-mail: Georgia_Hybels@partner.nps.gov

INTRODUCTION

Geologic maps are an integral component of the physical science inventories stipulated by the National Park Service (NPS) in its Natural Resources Inventory and Monitoring (I&M) Guideline (http://science.nature.nps.gov/im/index.cfm). The NPS Geologic Resources Division (GRD) is currently developing a Geologic Resources Evaluation (GRE) that includes a geologic bibliography, the creation of summary reports of each park's geology, evaluation of existing geologic maps, and the development of a geology-GIS data model for implementation in the production of digital geologic-GIS data for each park (such as Rocky Mountain National Park or Great Sand Dunes National Park and Preserve). The current data model implemented by the GRE for digital geologic-GIS data is the NPS GRE Geology-GIS Coverage/Shapefile Data Model (O'Meara et. al., 2005).

Recently, Environmental Systems Research Institute (ESRI) released the personal geodatabase, a relational database management system (RDBMS) designed specifically for storing, updating and viewing spatial data. Compared to the coverage- and Shapefile-based GIS previously offered by ESRI, the personal geodatabase offers added functionality of attribute validation rules, relationship classes, and topological rules that maintain data integrity within and between data layers. The GRE has determined that the added functionality of the geodatabase will increase data quality and help stream-line the data production process. Migration of GIS data involves significant changes to the current GRE data model and the revision of existing data capture/conversion procedures. Additionally, the migration of GRE legacy data must also be addressed.

CURRENT DATA FORMATS AND PRODUCTION PROCESSES

Presently, all GRE digital geologic-GIS data are stored in both coverage and shapefile format. Completed digital geologic-GIS data for a specific park are comprised of a set of both shapefiles and coverages, with each set being a collection of data layers such as geologic contacts or faults, as defined by the NPS Geology-GIS Coverage/Shapefile Data Model (O'Meara et. al., 2005). The data layers included in each set can vary depending on the source maps from which the data was derived.

At present, the data capture process involves either hand digitization of hard-copy geologic maps or, less commonly, conversion of existing digital data. Digitization and conversion are primarily conducted using ESRI ArcInfo Workstation and ArcView 3.x software. The multi-step, modularized process (Figure 1) is segmented by Arc Macro Language (AML) scripts that aid in capture/conversion steps, provide some quality control and, most importantly, preserve topological relationships between geologic features on the map.

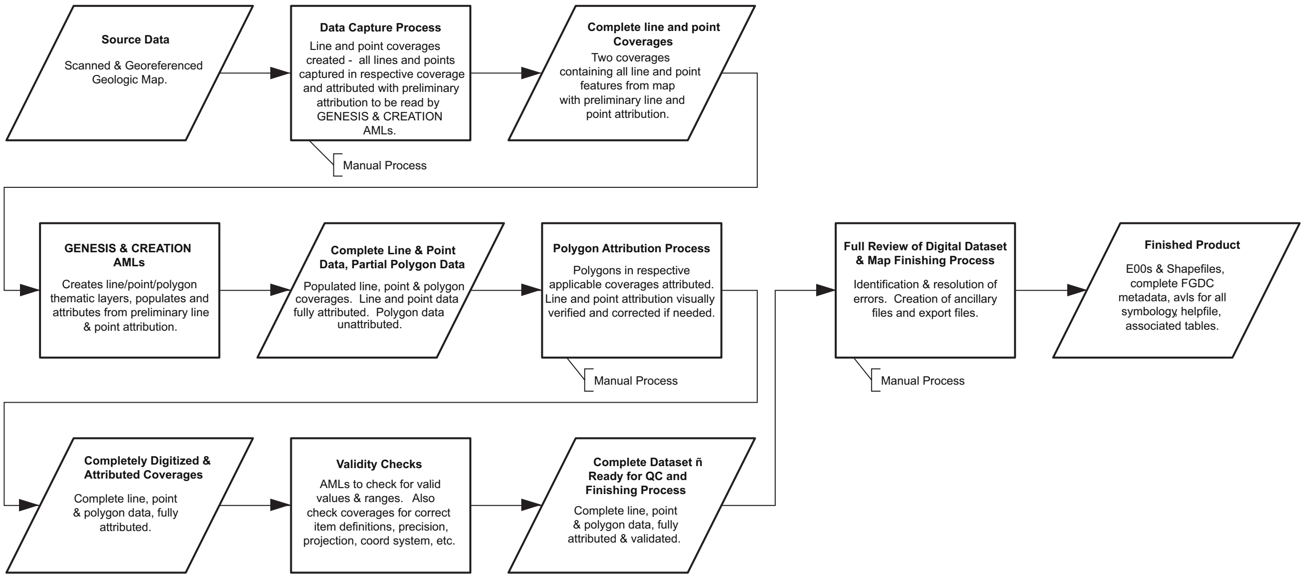


Figure 1. Schematic showing NPS GRE modular process and workflow for digitizing GIS data into the coverage/shapefile data model. Data is digitized and attributed in two datasets, one for point features, and one for line features. That data is parsed into one or more feature layers, according to the coverage/shapefile model. Polygons are attributed before the data enters the QC process. Process steps labeled as “Manual Process” require manual creation, editing or review of data. Note that in this schematic, only coverages are used in the digitizing process. Shapefiles can be substituted for coverages for manual capture and attribution tasks, however, all AML tools are written to run on coverages; shapefiles must be imported into coverage format for this reason.

The GRE’s current procedure for capturing hardcopy geologic map data involves digitizing all pertinent geologic features into two coverages, as defined by the coverage/shapefile data model. All line features are captured in a single arc/line coverage, whereas all points are captured in a single point coverage. This approach allows for topological relationships between features to be maintained in all data layers where they are coincident on the source map. Lines and points are then parsed to their respective data model layers (e.g., geologic contacts, faults, attitudes, etc.) using two AML scripts developed by the GRE—GENESIS.AML (lines) and CREATION.AML (points). Quality of captured data is ensured both manually and by employing AML scripts to check for data completeness according to the source map, and conformity to the data model.

Conversion of digital data is often more specialized depending on the source format, attribute structure, and the overall quality of the data. The GRE reviews digital data to assess the level of effort required to convert the data into the coverage/shapefile data model and to determine if additional editing will be necessary to bring the data up to GRE quality standards. Both quality and format are equally important and strictly enforced for all converted GIS data. Individual datasets (such as a single map) are converted to the coverage/shapefile data model manually. Multiple datasets (such as a collection of maps) from the same data model are often converted using AML scripts.

In order to ensure spatial coincidence between certain features such as faults and contacts, errors in existing data must be found and fixed. This can be problematic and time-consuming, considering that there are no readily available methods to ensure spatial coincidence between different coverages or shapefiles.

To supplement the GIS data, ancillary tables describing source map or source digital data and additional geologic unit information are generated. These tables are related to a data layer using a field common to both the ancillary table and the GIS data layer. Both shapefile and coverage formats do not allow for a permanent relationship between these tables and the GIS data; ancillary tables and GIS data must be temporarily joined when needed. Completed GIS data is combined with FGDC metadata, a Windows Helpfile containing source map information (legends, unit descriptions, cross-sections, etc.), and ArcView legend files to be used for symbolization, as part of the deliverable to each National Park.

MIGRATION APPROACH

When revising the coverage/shapefile data model, an iterative approach was adopted. Each data layer in the coverage/shapefile model was reviewed to determine how each layer would be defined in a geodatabase. The features and functionality in a geodatabase were discussed with regards to how they could best be employed for a

specific layer and its relationships with other layers. The resulting geodatabase schema, a detailed description of a data layer or layers, was implemented manually using dialogs and tools included in ArcCatalog. Data was loaded manually into the new schema and results were compared to data in the coverages/shapefile model. If any revisions were needed, the entire process was repeated until a final schema was agreed upon.

Implementing the data model in a geodatabase includes defining the data structure, setting up attribute value domains, and creating subtypes for use in topological validation. After evaluating Computer Aided Software Engineering (CASE) tools and other methods for implementing a data model, the GRE team decided to implement the geodatabase data model using an ESRI Developer Sample called Geodatabase Designer. The data model schema was stored in XML, as required by the Geodatabase Designer. The Geodatabase Designer is executed from ArcCatalog and, along with XML, provides the modular implementation necessary to load the layers needed for a specific geologic map. This is accomplished by designating an XML schema for each data layer in the geodatabase model. This aspect of modular implementation could not be accomplished using CASE tools, because they require one 'fixed' schema for all layers rather than individual schemas for specific layers. Although ESRI does not ensure long-term support for such a tool, the core functionality of the tool will most likely always work within the ArcCatalog architecture. Furthermore, ESRI supplies all source code for the Geodatabase Designer, thereby enabling users to modify and customize as needed.

It is of major importance to the GRE that production of digital geologic-GIS data not be interrupted by migration to the geodatabase data model. Equally important is the need to draw a well-defined distinction between data being produced in the old coverage/shapefile data model and data being produced in the new geodatabase data model. In order to reduce the impact of migration on geologic map production, this development work was done off-line until the new model had been designed, reviewed, revised, and released. This plan not only afforded time to properly develop the new model, but also enabled the continued use of AML scripts. Certain aspects of working in a geodatabase, such as topology validation, were inserted into the existing process to immediately improve the data being produced (Figure 2).

As part of the migration process, AML scripts that had previously been developed by the GRE to process data in coverage format must be redeveloped to work with data in a geodatabase. Currently, ESRI supports Python and COM languages, such as VBA, for automating processes within the ArcGIS framework. Any AML functionality that is not replaced by existing geodatabase functionality will have to be redeveloped using Python and/or COM. Until the new geodatabase model is implemented, however, our AML processes can remain in the digitization process alongside new geodatabase processes (Figure 2). Work has begun on using both Python and COM to replace AML processes (Figure 3).

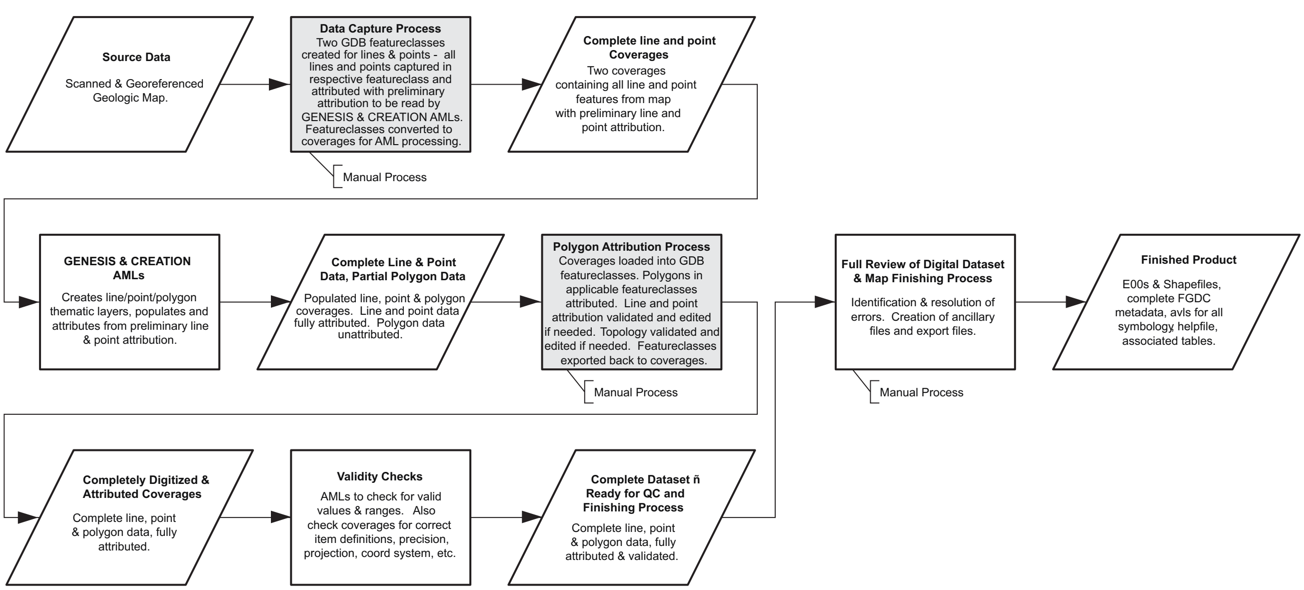


Figure 2. Schematic showing NPS GRE modular process for creating GIS data in the coverage/shapefile data model with geodatabase processes in gray. These processes were implemented immediately to enable use of the geodatabase's domain, topology and attribute validation functionality. Export is still in coverage/shapefile data model.

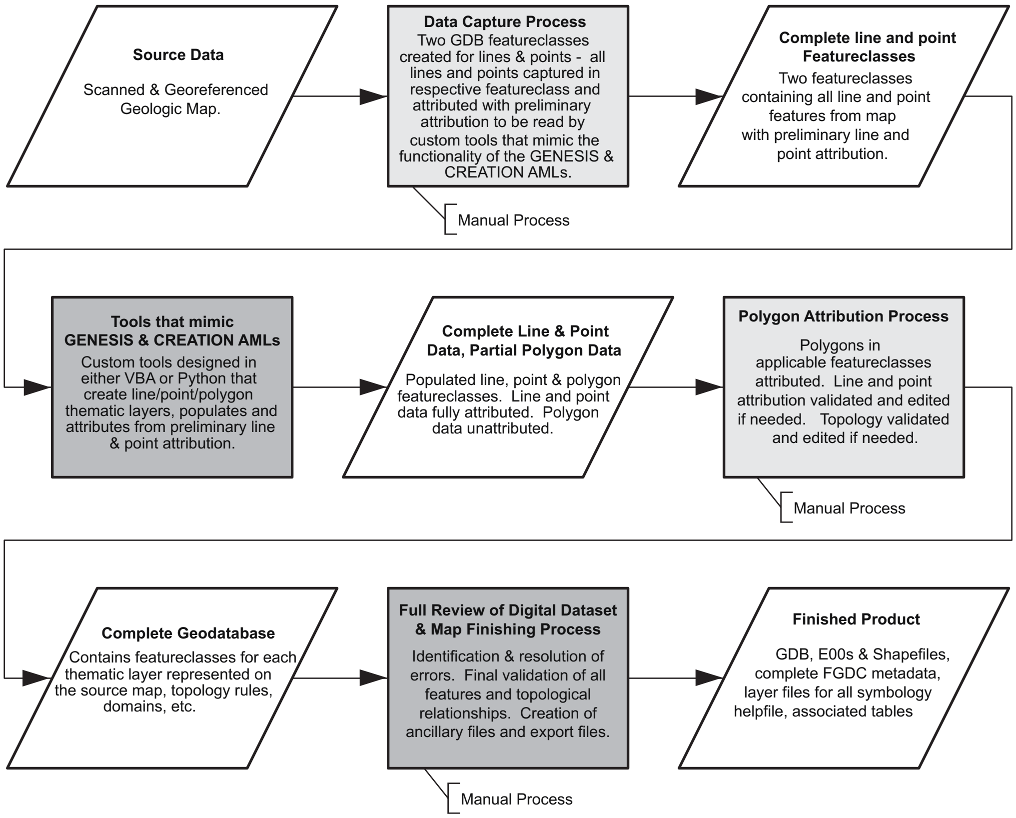


Figure 3. Schematic showing NPS GRE modular process for creating GIS data in the geodatabase data model with projected script development in darker gray. Work has begun on replicating the functionality of the GENESIS and CREATION AML scripts within the new geodatabase framework. Note all process steps in this schematic produce or require geodatabase featureclasses instead of coverages. Also note removal of 'Check Routines' step—compensated for by inherent geodatabase functionality and XML schema-loading process. Export includes geodatabase, coverages, and shapefiles.

CONCLUSION

The GRE recognizes the benefits offered by storing digital geology-GIS data in a personal geodatabase. The iterative process of revising data definitions to work within this new framework has begun, with good results. Topological validation has been successfully employed within existing digitization/conversion processes and new Python and COM scripts have begun to replace essential AML scripts. Conversion of legacy data to the new geodatabase data model will proceed, but only upon completion of digital geologic-GIS data for National Parks that have not yet been served by the GRE. The projected release date for the NPS Geology-GIS Geodatabase Data Model is the summer of 2005. The procedural approach developed for this process has worked well, with little to no interruption in production of GIS data.

REFERENCES

O'Meara, Stephanie, Gregson, Joe, Poole, Anne, Mack, Greg, Stanton, Heather, and Chappell, Jim, 2005, National Park Service Geologic Resources Evaluation Geology-GIS Coverage/Shapefile Data Model, available at http://science.nature.nps.gov/im/inventory/geology/GeologyGISDataModel.htm.

SOFTWARE REFERENCES

ArcGIS 8.3 and 9.0 (ArcCatalog, ArcMap, ArcEdit and Tables) – Environmental Systems Research Institute (ESRI) Inc., 380 New York St., Redlands, CA92373, http://www.esri.com.

Geodatabase Designer v2 – Developed by Richie Carmichael, Environmental Systems Research Institute (ESRI) Inc., 380 New York St., Redlands, CA 92373, http://arcscripts.esri.com.

Microsoft Office Visio Professional 2003, Microsoft Corporation, http://www.microsoft.com/.

Pathways-Phoenix: A Web Application for Decision Support using OGC Standards

By Boyan Brodaric, Joost Van Ulden, Ryan Grant, Otto Krauth,
Murray Journeay, Sonia Talwar, and Shannon Denny

Geological Survey of Canada
234B-615 Booth St., Ottawa
Canada
Telephone: (613) 992-3562
e-mail: brodaric@nrcan.gc.ca

ABSTRACT

In this paper we describe a web application, called Pathways-Phoenix (PX), that facilitates collaboration between scientists, planners, and the public. PX leverages open geospatial standards and technologies for information delivery, and enables users to deliberate about the information in custom designed projects. To demonstrate its re-usability and effectiveness, PX is being tested in Canada and South America by a variety of geoscience agencies.

INTRODUCTION

An increasingly important task faced by geological surveys is the use of geoscience information in non-traditional societal applications such as land-use planning. This non-traditional use often requires the transformation of geoscience information into value-added products that demonstrate the socio-economic value of the underlying information. The creation, evaluation and use of such products often involves multi-disciplinary teams whose members are located in different places, but who share common tasks, plans, information sources, and various software tools. The emergence of geospatial data standards and related technologies makes it possible in theory for team members to share many of these things over the web, but in practice there is a lack of available web applications to facilitate such sharing and to promote collaboration within a team. The Pathways-Phoenix (PX) application is designed to overcome these shortcomings. It is a web application that enables information sharing and collaboration within project teams. It can be used by scientists to share scientific results and to confer about them; it can also be used by decision-makers to communicate amongst themselves and scientists, and by the public to provide feedback on implications of scientific results and on proposals by policy-makers. As such, PX represents a component of a decision-support toolbox—it is the web interface to information products and collaborative project activity.

THE NEED FOR PX

Geological agencies are increasingly being asked to demonstrate the societal relevance of their work. Many agencies are pursuing a two-pronged strategy for accomplishing this: (1) by making their information and knowledge more widely accessible, and (2) by providing value-added products more tailored to support decision-making on societal problems.

A growing trend in information access is the move away from human-driven data access, where humans go to a web site to access data from proprietary systems, and instead toward web service access, where computers access data by communicating with web servers directly using standard protocols. This trend is reinforced in the geospatial community by the emergence of standards and technologies for accessing geospatial data through such services, and by the commitment of many significant data providers to adopt these standards. The trend signals a need to develop mechanisms for managing networks of data made available via web services. An important component of such strategies are web applications, or 'clients', that serve as human interfaces to the data networks. PX is a client intended to address the problem of managing geospatial and other data distributed over the web. It achieves this by providing an out-of-the-box solution that makes transparent many of the technologic hurdles that must be overcome, such as the installation and configuration of web mapping services and associated technologies.

Once information is available via standard protocols and access is enabled via effective clients such as PX, it becomes important to facilitate transformation of traditional geoscience data, such as geological maps, into derived products that are directly useful to decision-makers, such as risk maps (e.g. of natural hazards, groundwater contamination, etc.). Add to this the fact that the systems involved in the derivation of such value-added products are also migrating to web service interfaces, then two clear needs arise: (1) the need to manage workflows consist-

ing of web service chains that process data into derived products, and (2) the need to develop web applications that enable scientists, decision-makers and the public not only to catalog such services and products, but also to deliberate about their fitness for use in specific societal problems. PX is designed to respond to both needs: it is developing mechanisms for managing human workflows, but not for specifying and executing computational workflows, and it provides a web-based collaboration environment that facilitates deliberation about information sources.

DESCRIPTION OF PX

PX is the result of significant collaboration by initiatives from several agencies, including:

- The GSC Pathways Project (http://sdki.nrcan.gc.ca/path/index_e.php),
- The Multi-Andean Project: a consortium of six South American geological survey organizations and Canada (http://www.pma-map.com/en/gac),
- The Sustainable Development Research Institute, University of British Columbia (http://www.sdri.ubc.ca/), and
- The ResEau Project of Environment Canada (http://map.ns.ec.gc.ca/reseau/en/).

PX can trace its technologic roots, and ideological inspiration, to the Georgia Basin Digital Library (Talwar, et al., 2003) and the CordLink Digital Library (Brodaric, et al., 1999), which are its predecessors. Both of these earlier efforts explore web-based management and use of geoscience information, increasingly to augment various decision-making processes.

PX Features

Standards-driven: access and manipulation of all geospatial data is via OGC web service protocols (http://www.opengis.org/). PX has mechanisms for registering, searching and retrieving geospatial data using such web services, as well as mechanisms for automatically creating web services for geospatial data stored locally by PX. PX currently supports the WMS (Web Mapping Service, for browsing images) and WFS (Web Feature Service, for retrieving and updating vector data) web service standards; support of the WCS (Web Coverage Service, for retrieving raster data) standard is planned. Visualization of data is supported via the SLD (Styled Layer Descriptor) and WMC (Web Map Context) standards, which enable users to select data subsets and assign symbolization to them. PX also manages metadata by using the Z39.50 protocol for information access, and the Canadian Geoscience Knowledge Network metadata content standards, which are a subset of the U.S. Federal Geographic Data Committee specification.

Project-driven: all data and related activities are organized in projects (called workspaces). Projects are managed by administrators, who act as gate-keepers: they decide who has access to the project and they set security levels, thereby determining what actions can be taken by individuals. Projects contain a specific subset of all resources available to PX: e.g., users can register a subset of all available data with a project, manage tasks, post news items, conduct surveys, etc.

User-driven: PX is designed to be configurable and easy to install. PX enables users to design custom user interfaces, such that the look and functionality of PX can differ from one installation to the next, and indeed from one project to the next. Such customization can be accomplished in its basic form without additional programming. Once custom styles are in place, they can be selected from a pull-down menu, resulting in a different screen appearance and potentially different functionality; users can also toggle between the English, French and Spanish languages at any time. Installation is also being designed to insulate users from the complexity of managing web services. The overall notion behind PX is re-usability: it is designed to be a plug-and-play web mapping and collaboration environment that facilitates decision support with geoscience information.

PX Functions

Information management functions: PX enables three types of information to be registered, deleted, somewhat updated and queried within specific folders in a project. These information types include: (1) geospatial layers (such as WMS, WFS, ESRI Shape and raster layers, e.g., a layer containing polygons classified as ‘geological units’), (2) combinations of layers called maps, and (3) other documents (such as text documents, spread sheets, databases, etc.). Information can be located remotely, on some distant server, or it can be uploaded to a PX server. Once maps are created from geospatial layers they can be viewed from other systems that support OGC access protocols, such as various GIS systems.

Collaboration functions: PX enables different-time, different-place collaboration. This assumes that users are not co-located (i.e., different-place), and that they do not attempt to create or edit the same resource simultaneously (i.e., different-time). PX enables users to add, edit, delete, and organize a variety of collaboration resources in folders within a project, including: announcements, contacts (e.g., project members), a discussion forum, an email archive, a list of events, a catalog of web links, polls (questions with multiple-choice answers), surveys (questions with free text answers), tasks, weblogs, news feeds, and publication mechanism that allow informa-

tion to be broadcast outside the project. The intent of the collaboration functions is to enhance communication and interaction amongst project participants.

Project administration functions: these functions enable projects to be created and managed, including (1) configuring a project by changing its visual appearance or by selecting which information and collaboration functions will be made available to project members; (2) adding and deleting project members, and editing their profiles including security levels for accessing and publishing information; and (3) managing project content, including the recycle bin, which enables recovery of deleted items, and project statistics, which summarize project activity.

These three groups of functions are typically made visually available as selections in the user interface, i.e., note in Figure 1 that the column on the left contains links to the information, collaboration, and administration functions.

PX Architecture

PX is designed using a three-tiered system architecture, including tiers for: (1) content, (2) business functions, and (3) presentation. The PX user interface components, including customized user content in the form of graphics and text, as well as project contents (e.g., maps) that are displayed and managed by that interface, are all stored in a SQL-Server relational database. Access to the PX elements in the database is enabled by software middleware consisting of functions written in the Microsoft ASP.NET environment. These functions are then used by style sheets in the presentation tier, which expose the functionality to the user via the customizable user interface. Installation of PX on a server requires the SQL-Server relational database to be pre-loaded, and the open-source MapServer software is then semi-automatically loaded during installation. Once a PX server is in place, projects can be easily established. PX is a very thin

Figure 1. PX's map viewer: note the information, collaboration and administration functions on the left (www.pathways.geosemantica.net).

client—users only require a web browser with the Flash plug-in to be able to work with PX projects.

PX USE

At the time of this paper, PX is being used mainly by two GSC projects: (1) in the Pathways Project, PX is deployed to increase use of geoscience information in land-use planning, and (2) in the South American-based Multi-Andean Project it is geared toward the management and sharing of natural hazards information amongst six geological survey organizations.

Pathways Project

The Pathways project is responding to the perception that regional planners are not often using geoscience information in the development of growth strategies and plans, or are not using it as effectively as they might. To overcome this situation, the Pathways has adopted a risk-based approach in which consecutive modeling steps ultimately lead to the generation of alternative land-use scenarios, with attached socio-economic risk levels, that can be considered by planners. PX is being tested in two ways under this approach: (1) as a means of enabling collaboration amongst a team of water scientists jointly developing hydrogeological models in the Okanagan Basin, and (2) as a means of distributing geospatial information, including model results, to decision-makers, and encouraging web-based deliberation over those results. Figure 2 shows the PX home page for this project, which can be accessed via pathways.geosemantica.net.

Figure 2. PX home page of the Groundwater Assessment in the Okanagan Basin project.

Multi-Andean Project: South American Geoscience Information Network

The Multi-Andean Project aims to decrease the negative impact of natural hazards on communities in six countries (Argentina, Bolivia, Colombia, Chile, Ecuador, Peru, Venezuela) in the South American Andes region. An important component of the work involves the development of an information system, called GeoSemantica, that enables the various agencies to deliver, manage, and share information over the web, including information from joint project work. GeoSemantica is composed of PX installations, called nodes, that are customized for use within individual countries. At this time there exist nodes in Chile, Colombia and Canada, with more to follow soon. These nodes are being heavily used, so much so that they are being adopted in non-geoscience domains by other government agencies and universities. Figure 3 shows the home page for the Canadian node, which can be accessed via http://can.geosemantica.net/—note the customized graphical user interface and its contrast to the user interface adopted by the Canadian Pathways Project, which shown in Figure 2.

CONCLUSIONS

Pathways-Phoenix (PX) is a web application designed to enable (1) management of geospatial information provided via open geospatial standards and technologies, and (2) deliberation over such information between scientists, policy-makers, and the public to aid in decision support. It is designed to be easily installed on a web server, and can be used in any web browser that has installed the Flash plug-in. PX development is on-going, with v1.0 expected to be released over the web to the public, at no cost, in March 2006. It is currently being tested in Canada and South America in various geoscience projects that require collaboration amongst individuals from various disciplines and jurisdictions. Additional capabilities to be added to PX in the near future include:

- editing non-spatial attributes of geospatial features,
- adding geospatial points to maps, and editing their attributes,
- enhanced querying of geospatial features,
- managing human workflows within projects, and
- managing complex documents.

REFERENCES

Talwar, S., Journeay, M., Harrap, R., Brodaric, B., van Ulden, J., Grant, R., and Denny, S., 2003, The Georgia Basin Digital Library Project: a framework for a place-based exploration of sustainability issues: Geomatica, v.57, no.2, p.155-164.

Brodaric, B., Journeay, M., and Talwar, S., 1999, Using the Proposed U.S. National Digital Geologic Map Data Model as the Basis for a Web-Based Geoscience Library Prototype, in D.R. Soller, ed., Digital Mapping Techniques '99—Workshop Proceedings: U.S. Geological Survey Open File Report 99-386, p.83-91, available at http://pubs.usgs.gov/of/1999/of99-386/brodaric1.html.

Figure 3. Home page of the Canadian GeoSemantica node.

Capturing and Vectorizing Black Lines from Greenline Mylars

By Thomas G. Whitfield, P.G.

Pennsylvania Bureau of Topographic and Geologic Survey
3240 Schoolhouse Road
Middletown, PA 17057
Telephone: (717) 702-2017
Fax: (717) 702-2065
e-mail: twhitfield@state.pa.us

ABSTRACT

Land development pressures in glaciated northeastern Pennsylvania and the Poconos have resulted in a great demand for information about the surficial deposits of the area. Surficial deposit mapping of this area has been an on going STATEMAP project for many years (STATEMAP is a component of the USGS National Cooperative Geologic Mapping Program). Two or three 7.5-minute USGS quadrangles are usually mapped each year. Until recently, finished (but not finalized) map projects consisted of a text report and one or more clear or greenline mylar quadrangle maps. A finished map project is one in which the author has completed his or her fieldwork, maps, and documentation, and has had a minimum level of review. A finalized map project is one that has had a more formal review and has met all the standards necessary for formal publication. Surficial geology contact lines, isochores, bedrock outcrop ledges, etc. were drafted directly onto mylar maps or on mylar overlays. Other features were hand drafted or rub-on transferred to the mylar sheets.

The initial intent was to release these maps as formal publications at a later date, but given the demand for the data, they were released in the open file series. Each open-file report consisted of large, at-scale photocopies of the mylar maps and various overlay combinations, in widely varying quality, and a copy of the report.

When GIS and digital map data began to be widely used in the 1990's, users began to request these maps in a digital format, preferably as a georeferenced GIS file. Early attempts to convert the mylar maps to digital were problematic. Many of the greenline mylars had black ink contact lines drafted directly on them. Scanning these maps and separating the drafted line from the background was very difficult. Most of the digitizing had to be done by hand.

New and improved scanning techniques and software solved the problem of capturing lines drafted directly on a greenline mylar. Drafted lines are captured at the scanning station and saved as a separate binary image. A binary image is a raster image with just two values. Each pixel is either a one (1) or a zero (0). Improved auto-vectorization software (ESRI's ArcScan 9.x) that also allowed interactive image editing was also a great step forward. ArcScan reduced the digitization process by several days.

This particular open-file series of maps is now completely digital. When new maps are released, they include many different georeferenced and attributed data layers and data-sets, instead of one or two large photocopies of the originals. Also included is a PDF file of the finished map for those who wish to print their own copy.

GETTING DRAFTED LINES OFF A GREENLINE MYLAR

Heads-up digitizing of a scanned image is generally a straightforward process, but can consume many hours. Automated or semi-automated digitizing speeds the process up considerably. For successful tracing, however, most automated digitizing programs require a binary or black and white image. The line tracer will follow pixels with ones or zeros, but not number ranges associated with color designations. Producing a usable binary image from a greenline mylar can be a difficult task.

The key to getting a good scan of a mylar is good contrast. Because a mylar is translucent, the scanner will often pick up the color of the hold-down bar behind the mylar as it is scanned (Figure 1). If the hold-down is white, then there usually won't be a problem. But, more often the paper hold-down is pitted, scratched, and discolored, and therefore it does not make a good contrasting background for images on the mylar. Creating a sheath out of a folded piece of clear acetate, then putting a scrap piece of white plotter paper inside the sheath, behind the mylar, makes an excellent mylar scanning "packet" (Figure 2). The mylar then has a solid white background for contrast, and the sheath keeps everything together. This type of sheath is also good for scanning worn, tattered, or delicate maps. It protects the maps from friction associated with the hold down bar and traction friction from the scanner rollers (Figure 3), and it lets you carefully piece a tattered map back together, keeping loose map pieces in place while being scanned.

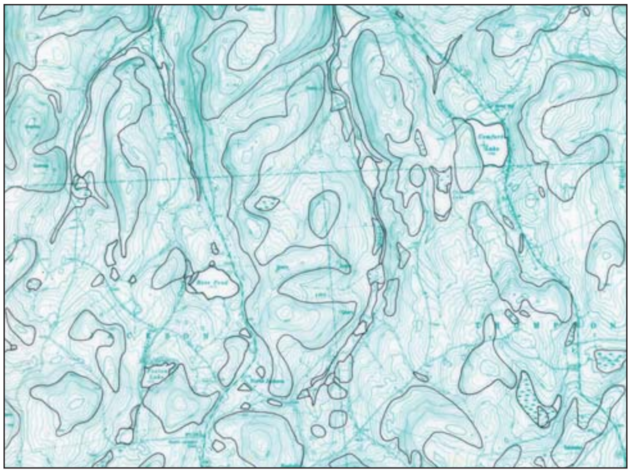


Figure 1. Surficial geology contact lines drafted directly onto a greenline mylar.

Figure 2. Folded clear acetate sheath for scanning. A sheet of plotter paper is placed behind a greenline or overlay mylar to provide a white background for contrast. This set-up is also good for protecting fragile maps from friction involved in the scanning process.

Practicing good scanner hygiene is equally important and will prevent problems in the future. The scanned media should contain no tape, glue, or staples. Tape and glue residue can transfer to the scanner glass, causing streaks and lines to appear in the scanned image, and also can be transferred to other originals that are scanned. Staples can permanently scratch the surface of the glass. Scanner glass is optical quality glass and is often softer than normal plate glass and can be quite costly to replace. Also, the scanner cameras have a precise focal field and they focus on the upper surface of the scanner glass. Any glue, debris, or scratches on the glass can adversely affect the quality of the scanned image. Always check the scanner glass for foreign matter, and clean regularly. Also ensure

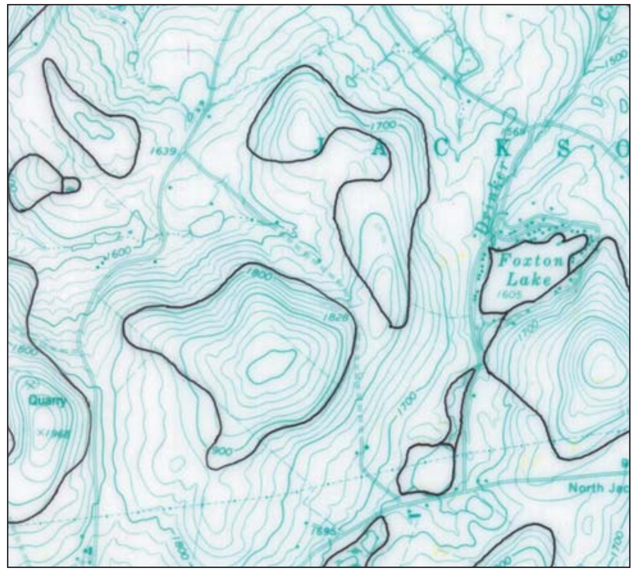


Figure 3. A closer look of the surficial geology contact lines drafted directly on a greenline mylar. The variable line quality is there, but not easily seen.

the mylar or other original map is free of dirt, eraser dust, etc. A horsehair drafting brush is great for dusting off originals.

Drawing a line on mylar with a drafting pen is not difficult, but getting consistent ink line quality can be. Because mylar does not absorb the ink as paper does, the drafted lines can vary in thickness and density. Lines can be thick and dark in some places, and thin and light in others. These variations in the drafted lines can make it difficult to capture them in a consistent manner. The lines in Figure 3 appear to be consistent. Their variations, however, will not become apparent until the mylar is scanned.

One other detail we nearly overlooked in this project was the preservation of control points or tic marks on each map. All the line work on a greenline mylar is of a 7.5-minute topographic map is, of course, colored green, including the tic marks (Figure 4). Because the objective of scanning the greenline mylars was to make the green lines disappear, we had to either draft the tic marks onto the mylars in black, or use black rub-on transfers to ensure we retained control points on the scanned images. Usually the other mylar overlays had the tic marks already on them. If they did not, we used a light table to manually add the control points.

During the scanning operation, a threshold setting determines the sensitivity of the scanner. The threshold sets the values that the scanner uses for dividing tonal ranges into black and white output. Setting the threshold high enough to drop out the greenline background and noise in one area may cause fainter black lines to be dropped out in another area (Figure 5). Setting it too low will increase noise (speckling) and will pick up unwanted background

Figure 4. Location control points (tic marks) on the greenline base map are colored green. In order not to lose the control points when the green lines are deleted from the mylar base during scanning, they had to be redrawn in black. Note the location of the green tic mark inside the circle, northwest of the "L" in Comfort Lake. This tic mark had to be changed to black before scanning.

lines (Figure 6). Finding the right setting often involves a lot of trial and error.

Many of the newer scanning interfaces have an automatic thresholding feature. During the scanning process, the scanner will analyze small sections of the object map and determine the optimal threshold setting for that section within a variability range set at the interface. By independently varying the threshold for each section of the object map, noisy areas are cleaned up and light or faded object lines are more reliably detected. Although this process was designed for maps such as blue-line ozalids where print quality can vary widely across the map and for older maps that tend to degrade to a yellowish color, it worked very well for us in dropping the greenline background from our mylars while preserving the black contact lines (Figure 7).

We used a Vidar Titan II scanner (http://www.vidar.com/wideformat/). It is a color scanner capable of scanning maps up to 40-inches wide and (assumed) unlimited length. It has a dual roller feed, three cameras, and an optical resolution of 400 dpi. The dpi can be increased in the software, but anything above the optical resolution of the cameras is done through software interpolation. The scanning software we used is Vidar TruInfo v1.4.6, which was supplied with the scanner. The scanner and software were purchased in 2000 and there have not been any updates to the software since then.

File size and image formats can become significant issues during scanning. We found that the easiest and safest image format to use is TIFF. It is a very common image

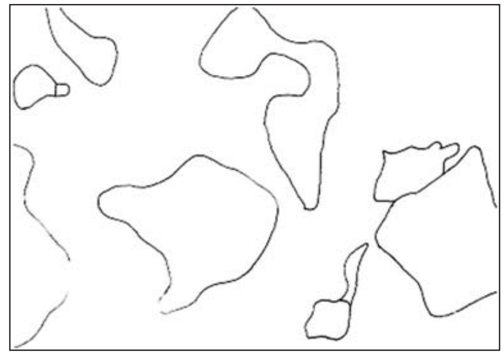


Figure 5. Setting the threshold too high causes the background and noise to drop out, but often fails to retain fainter parts of the contact lines.

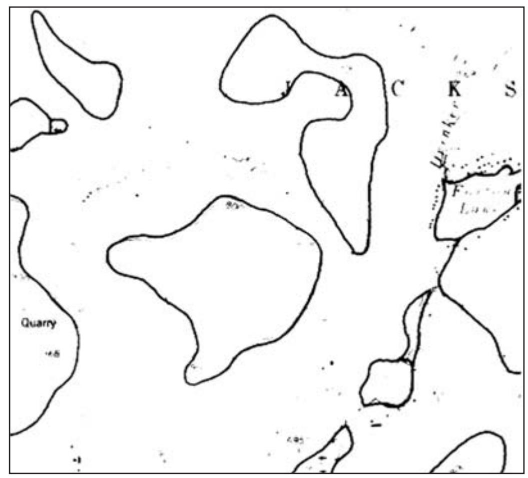


Figure 6. Setting the threshold too low picks up unwanted background lines from the greenline, and "noise" along with the geologic contact lines.

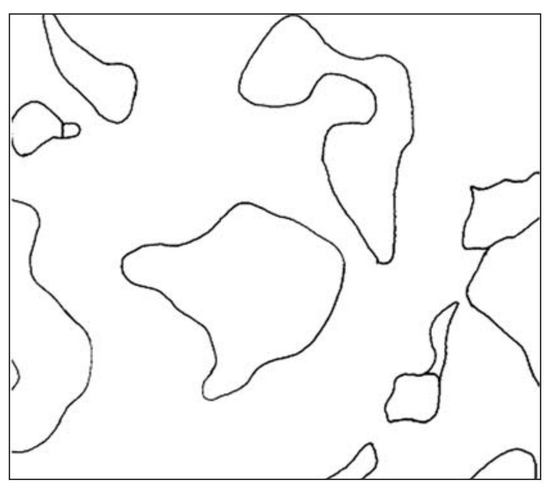


Figure 7. The automatic thresholding option effectively dropped out the greenline background and noise while preserving fainter drafted contact lines.

format, and is easily read by most GIS and image software packages. File size varies by the image format chosen, but more significantly by the resolution (dpi, or dots per inch) chosen for the scanner. "Dpi" is somewhat of a misnomer when applied to images. Dpi is more appropriate when working with inkjet plotters to specify how much ink per linear inch the plotter will "dot" on the paper. Dpi when applied to images designates how many pixels a scanner will record in an inch of measure. The higher the dpi, the more (and subsequently smaller) pixels the scanner will fabricate in an inch in both the x and y directions. As the pixels get smaller, the resolution and defined detail of the image increases, as well as file size, sometimes exponentially. Image dpi settings of 300 or 400 are usually more than adequate for most uses. Higher dpi settings result in large, unwieldy files that are difficult to handle and have little or no noticeable gain in clarity. In our case, we worked primarily with 300 and sometimes 400 dpi images. The 400 dpi setting was used for mylars where the contact lines were very close together. The increase in detail kept close lines from melding together. Also, as noted above, the cameras of our scanner have a resolution maximum of 400 dpi; therefore, dpi settings above 400 can only be achieved by software re-interpolation.

VECTORIZATION

When we completed the mylar scanning process, we moved the TIFF image files over to a workstation for further processing. Although we could have vectorized the scanned images, we found it more efficient to first georeference the images to the appropriate map projection. It saves time later, and allows us to compare it to other georeferenced data layers during the vectorization process. We used the georeferencing module in ArcGIS 9.x, with a pre-defined 2.5-minute point grid coupled with a 7.5-minute quadrangle boundary line grid. Each image was brought in to ArcGIS and geo-referenced to the 16 control points (2.5-minute tic marks) on each scan (Figure 8).

We used the ArcScan extension module of ArcMAP 9.0 and 9.1 for line vectorization. ArcScan was an optional module in ArcMAP 8.x through 9.0, but has been made a part of the core functions in version 9.1. ArcScan is also a module available in Arc/Info Workstation 7.x and higher, but the only similarity is that the Workstation version does trace lines. The ArcMAP version of ArcScan is a vast improvement over the Workstation version.

ArcScan draws vector lines based on how it traces contiguous raster pixels. In order for ArcScan to trace

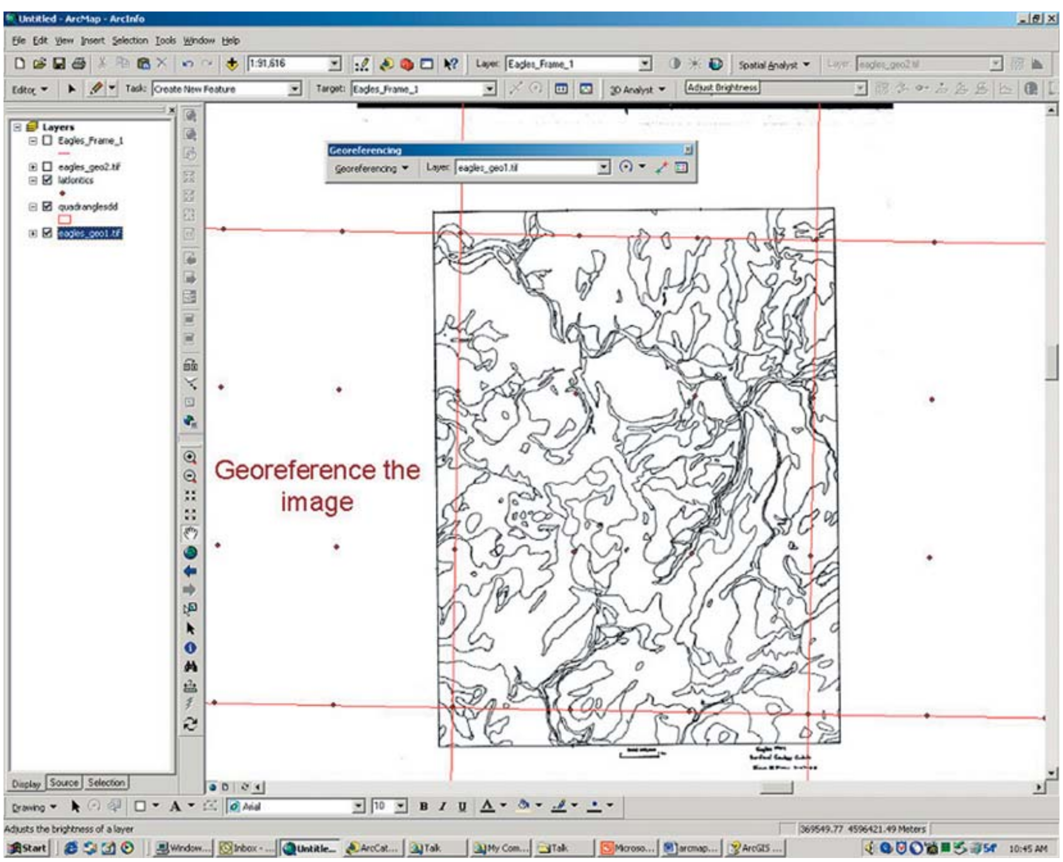


Figure 8. Using ArcGIS to georeference the scanned image to a 2.5-minute point grid and 7.5-minute quad boundary data set. The projection is set to match the projection of the original map.

contiguous pixels, they must be the same value, and surrounding pixels must have a contrasting value. As a result, ArcScan will only work on binary images where values are either 1s or 0s (zeros). It cannot distinguish single pixel values (or near the same color values) from 256-color raster color, grayscale, or 3-band color (RGB or CMYK) images.

Once the georeferenced image has been brought into ArcMAP, it is best to change the classification designation (found under the symbology tab) to unique values. It does not matter what colors the user designates to represent the 1s and 0s on screen, as long as there are only two values on the scanned image.

Raster line intersections have always been one of the hardest things for ArcScan (or any vectorization software) to interpret. "T" intersections would commonly have a deep "V" in them where the tracer would move to the geographical center of the pixel cluster in the middle of the "T" before continuing down the pixel line. ArcScan now offers new intersection solutions under the vectorization settings: geometrical, median, and none. The "geometrical" option tries to preserve angles and straight lines; in other words, it tries to keep "T" intersections as "T" intersections. "Median" is designed to work for non-rectilinear angles; this is presumably for use in depicting natural resources where right angles are rarely observed. "None" is designed for nonintersecting features like contours, etc. Although one would assume that the median option is best for our use, we found the geometrical option gave far better results with less clean up needed.

ArcScan also has problems tracing lines intersecting at low angles. Often there are pixel in-fills between the lines as they approach the actual line intersection (Figure 9). The tracer interprets the line intersection to be somewhat short of its actual location and at a larger angle than intended. This problem can be addressed by the interactive raster editing capabilities of ArcScan.

The interactive raster editing module and the tracing preview option, used in conjunction with each other, are by far the biggest time savers of the ArcScan extension. The interactive raster editor allows the user to edit the raster image on the fly. Pixels can be erased or filled individually, in blocks, by "painting" (Figure 10), or by a number of different options. The preview option shows the user how ArcScan intends to vectorize the pixel lines as they are shown on the screen. The vectorization preview can be set to refresh after each raster edit. If the user erases a number of pixels at once, after the mouse key is released, the preview will refresh to show how the vectorizing will change. The user can then tweak individual pixels, if necessary, to obtain the best results. Raster editing not only gives the most optimal vectorization results, thereby reducing clean up, but also produces in a very clean raster image (Figure 11).

Another of the raster editing tools that is actually fun to use is the "magic eraser". The magic eraser interactively erases connected pixels. It will erase a feature by touching it or by drawing a box around it. This is quite useful if, for example, a name happens to appear on the scan. Touch it or surround it, and the name disappears. The magic eraser, however, will not erase a pixel string if it passes through the magic eraser bounding box. This is quite useful if there are a number of random dots (noise) appearing on both sides of a contact line. To erase the noise, simply surround the noise with a magic eraser bounding box, making sure the contact line passes through the bounding box, and the noise within the box is erased leaving the contact line intact (Figures 12 and 13).

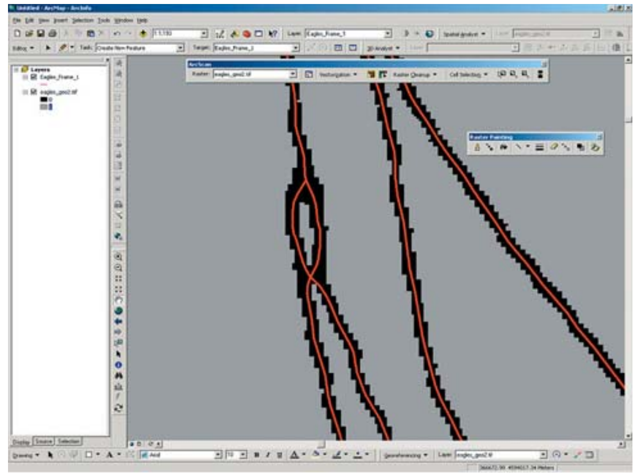


Figure 9. ArcScan module showing a preview of how the tracer will vectorize this area of the scan. The contact lines were very close, so when scanned, the pixellated lines merge into one. The vectorization tracer will try to cross the vector lines because the pixel lines are not separated.

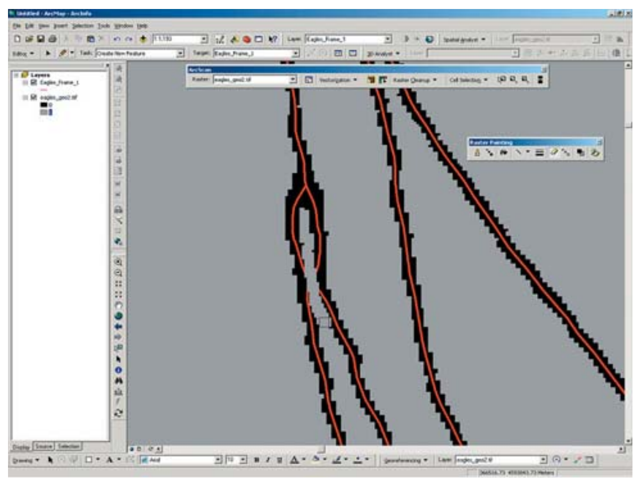


Figure 10. Interactively erasing the pixels between the lines.

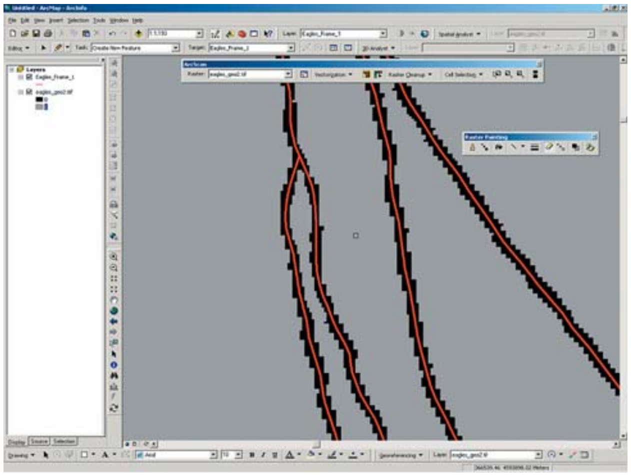


Figure 11. The tracer will now vectorize the area correctly.

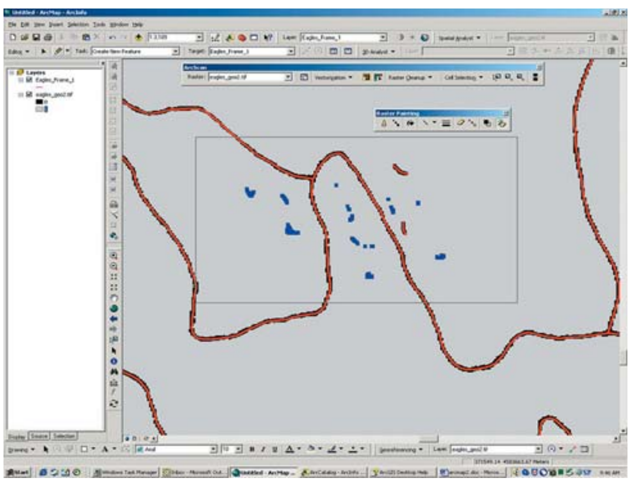


Figure 12. Random noise (dots) near a contact pixel line to be vectorized, surrounded by a "magic eraser" bounding box.

Once the raster editing was done, we used the ArcScan batch vectorization feature to vectorize the entire scan (Figure 14). We could have used the interactive tracing tool to digitize the images; it operates in a manner similar to the interactive tracing tool in Arc/Info Workstation's ArcScan and requires tweaking of the detection and direction settings to actually get it to run smoothly. The interactive tracer will follow a pixel line until it encounters an intersection or a cluster of pixels with an unclear exit. It then waits until the user decides which way the tracer should go. This interaction continues until the scan, or parts thereof, are vectorized. We found the batch vectorization and subsequent minor clean up to be much less time-consuming than interactive vectorization.

The vectorization results were very good, but some final clean up was necessary. It was easier to do clean up at this stage than after the data set had been converted to polygons (e.g., only one line is being edited at a time, so there is no danger of creating sliver polygons; also, discontinuous lines are more easily edited). Checking the topology for each scanned map layer or theme is very important. We did not want dangling lines or disconnected lines (undershoots) present before creating polygons. We also did not want lines that self-intersect or overlap themselves or other lines. Bypassing this step can lead to hours of corrections later. Points and other features can be hand-digitized into their own data files. Line and point placements should be checked for accuracy against the scanned images and corrected where necessary.

Once we were satisfied with the positioning of these features, and the data sets were free of errors, we then converted the appropriate line files to polygon data sets (Figure 15). Converting line data to polygons in ArcGIS can be done in two different ways with the same results. In ArcMap, there is a "construct polygons from line features" button on the topology toolbar. In ArcCatalog, the construct polygons from line features option is found by

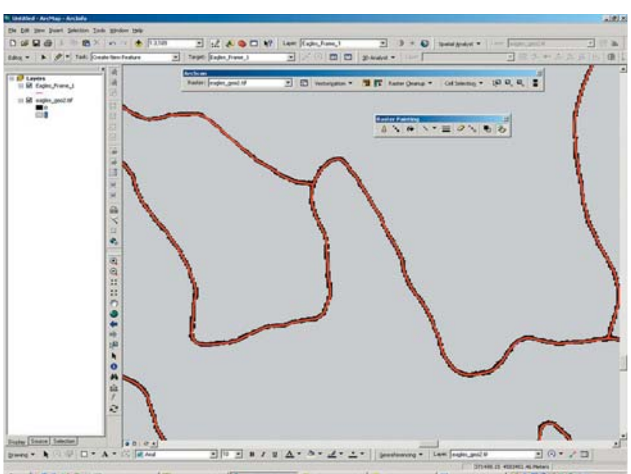


Figure 13. Random noise shown in Figure 12 removed, leaving the contact pixel line intact.

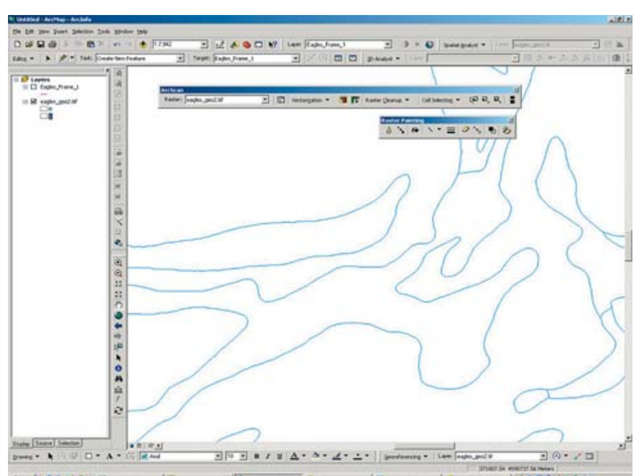


Figure 14. A newly-vectorized data set.

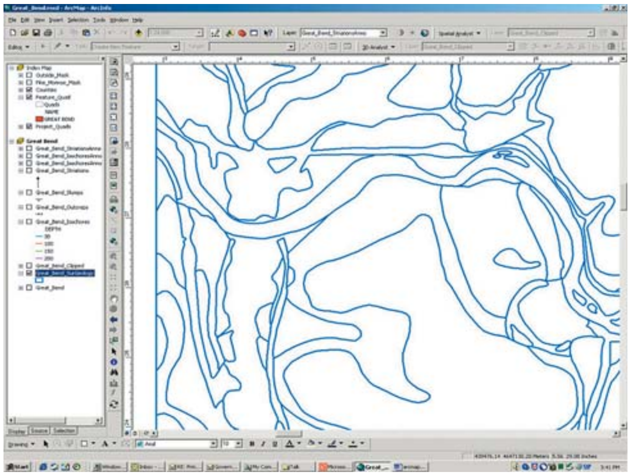


Figure 15. Polygon data set created from a line data set.

right clicking the "new" tab. A new polygon feature class is created without destruction of the original lines data set.

Although it may seem redundant, it is a good idea to create and maintain a topology rule set on the newly created polygon data set. Sliver polygons, polygon gaps, and overlaps, etc., are rare just after creating the polygon data set, but an inadvertent, undetected move of a polygon during attribution or other editing process can create problems later.

Assigning attributes to the various layers of the map project was a straightforward process. We created appropriate fields in each data set's attribute table and populated them accordingly (Figure 16). Extensive use of look up tables for many of the textural attributes saved a significant amount of time when assigning attributes. For example: we attributed the surficial geology layer, polygon by polygon, with just the geologic symbol (e.g., Qa, Qat, Qwic, etc.). We then used the join feature to link to a standard database containing a more detailed narrative field keyed to each geologic symbol. The resulting layer was then exported to a shapefile or geodatabase layer, making the joined narrative fields a permanent part of the data layer.

COMPLETION AND RELEASE

When we were satisfied with the vectorization and

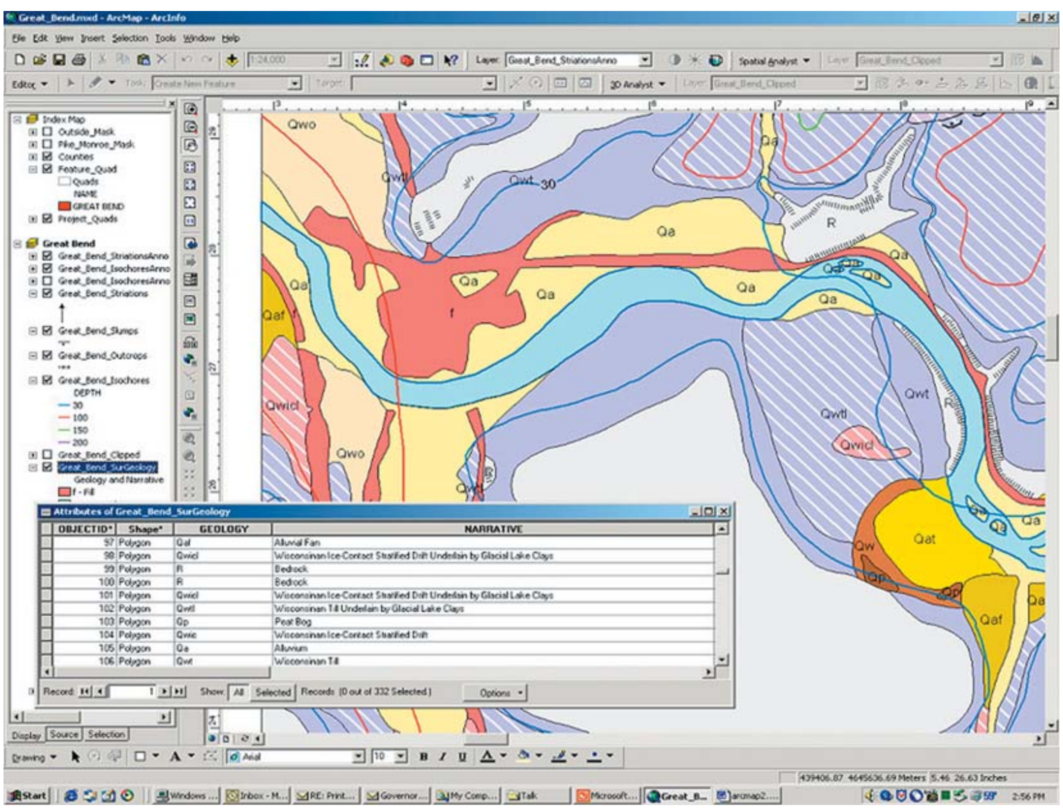


Figure 16. An attributed, symbolized, and completed polygon data set.

attributing of the various data layers, we made check plots of the maps and submitted them back to the authors for their review. Because this is such a large project, a generic map template was constructed in which any of the quads can be placed, with minimal editing and adjustment, and printed. In some cases, the authors made changes or clarifications to the data, which was edited accordingly. The completed maps and digital files were then submitted for internal review and approval.

We made the decision to release these maps as part of our Open File Series of publications; these publications have undergone a level of review, but have not been subjected to rigorous formal publication reviews. The purpose of these open-filing the maps is to quickly get them to our customers. Caveats apply to the data until they have undergone a more rigorous review and are formally published.

Data sets released to the public are in several digital formats including ArcGIS Geodatabases and shapefiles. For those using the digital data in ArcGIS, we include the ArcMap MXD (ArcMap document) file, and a PMF (ArcPublisher-created) file for use with ArcReader, a limited version of ArcGIS and free download from ESRI. A PDF of the map document is included for those not using GIS, or those who just want to print the map.

CONCLUSION

In northeastern Pennsylvania, we have more than 30 USGS 7.5-minute quadrangles with the surficial geology already mapped in analog form. Each quad has from 3 to 5 mylar overlays along with a greenline quad base. The greenline mylar base usually has one of the data layers, surficial geology, drafted directly on the mylar. Digitally lifting the surficial geology contacts from the greenline mylar was a challenge. By adapting scanning and rendering techniques designed for other purposes, we are able to successfully digitize the surficial geology data layer with minimal effort and in a timely manner.

The overall goal of this project was to get highly sought-after information to the general public quickly, and in a digital format. Although the user must be aware that the data has not been through the formal review process and is subject to change, it is still the best data available right now.

Geologic Quadrangle Mapping at the Illinois State Geological Survey

By Jane Domier

Illinois State Geological Survey
615 E Peabody Drive
Champaign, IL 61820-6964
Telephone: 217-244-2513
Fax: 217-244-7004
e-mail: domier@isgs.uiuc.edu
Web site: http://www.isgs.uiuc.edu/
Powerpoint presentation: http://ngmdb.usgs.gov/Info/dmt/docs/domier05.ppt

INTRODUCTION

Since 1996, a primary goal of the Illinois State Geological Survey (ISGS) has been to digitally map the geology of the state from the land surface to the bedrock surface, using USGS 7.5-minute (scale 1:24,000) topographic quadrangles as the fundamental mapping area. Principal map themes include surficial geology, bedrock geology, bedrock topography, and drift thickness. Additional themes are included depending upon the resources and needs of each area.

Initiated in 1993, the STATEMAP program is funded in part by federal appropriations to the U.S. Geological Survey and is a component of the National Cooperative Geologic Mapping Program. Over the past ten years, ISGS has compiled 65 quadrangle-based geologic maps for STATEMAP, but more than one-half of them remain unpublished, due to lack of sufficient staff and inefficiencies in map production.

This paper reviews the procedures implemented during the past three years at ISGS for improving the quality and efficiency of geologic quadrangle map production. An overview of large-scale USGS digital base map products currently in use at ISGS is presented first, followed by a discussion of new cartographic standards being developed for geologic mapping. Finally, I present a summary of how software use and production work flow modifications have impacted the efficiency of the ISGS map production process.

BASE MAPS FOR 7.5-MINUTE GEOLOGICAL QUADRANGLE MAPS

The selection of an appropriate base map has been an important element in our geologic mapping program. Two categories of large-scale digital spatial data are currently used at ISGS for map production: a) USGS Digital Line Graph (DLG), and b) USGS Raster Feature Separates (RFS).

USGS Digital Line Graphs

ISGS has used large-scale USGS DLG data for approximately 15 years. DLG data are comprised of vector data digitized from 7.5-minute topographic quadrangle maps, with updates conducted using National Aerial Photography Program (NAPP) or NAPP-like aerial photography. The resulting vector data produce high quality line work that is used as the base for geologic quadrangle maps. The various primary map features are contained in separate, topologically structured layers, thus allowing for selection or omission of features to be shown with the geologic map information. DLGs must first be converted to a data format compatible with GIS software, and the feature types within each layer must be coded and symbolized to recreate the appearance of the published topographic quadrangle map. Most features in DLG files do not have place name attributes, so usually the black feature text is scanned from a USGS mylar feature separate to produce a raster text layer for the geologic quadrangle map. DLG files require significant processing (approximately one week) with even an experienced GIS analyst before they can be utilized as the primary base for a geologic quadrangle map.

The availability of large-scale DLG data for Illinois is limited, and a majority of the quadrangles have no DLGs at all. In the past, updating of some of the more critical DLG feature layers, e.g. hypsography, hydrography, transportation and PLSS, could be produced at the USGS for approximately $7,500 for each quadrangle. Because the DLGs do not include cartographic text, a copy of the the mylar feature separate for the cartographic text has also been purchased at a cost of $60.00 for each quadrangle.

As of August 2003, copies of the mylar feature separates were withdrawn for purchase from the USGS. In addition, creation of new 7.5-minute DLG data at the USGS has ceased, but existing DLG data will continue to be available for purchase through EarthExplorer (http://edcsns17.cr.usgs.gov/EarthExplorer/). As of July 2005, individual 7.5-minute DLG data layers will only be accessible through a new USGS Tiled Data Delivery System (TDDS) that will be a component of the Seamless Data Distribution System (http://seamless.usgs.gov/).

USGS Raster Feature Separates

ISGS continues to use DLG data for those 7.5-minute quadrangles that are currently available, but an alternative map base must be used for the quadrangles for which no DLG data exist. When a USGS quadrangle map is revised, typically the original mylar feature separates are scanned at 1,000 dpi and updated using raster revision methods. The resulting Raster Feature Separates (RFS) data files can be used to create press negatives for printing, or can be combined into a single Raster Color Composite (RCC) file. These RCC data files are also used for the creation of a Digital Raster Graphic according to the USGS revised DRG product standard implemented in October, 2001 (http://topomaps.usgs.gov/drg/drg_standard_change.html).

RFS data files can also be produced directly from the mylar feature separates without any revision. In this situation, the RFS and RCC files will be identical to the most current published edition of the 7.5-minute quadrangle map. The cost of RFS and RCC data files for a single quadrangle may range between $700-$1,200, depending upon the number of original map feature separates. The resulting RFS and RCC data are delivered as both 1,000 dpi and 500 dpi resolutions in GeoTIFF format, and can readily be used in GIS software systems. USGS standards for RFS data are available at http://rockyweb.cr.usgs.gov/nmpstds/rfsstds.html. Currently, RFS products are produced using a full-repay funding agreement with the USGS, which means that the user must pay the entire production cost, and this situation is not likely to change in the near future. The USGS RFS product provides a base map that is high resolution, affordable, and easy to use (Figure 1).

MAP PRODUCTION PROCESS AT ISGS

Early attempts at ISGS in the digital production of geologic quadrangle maps involved experimenting with a number of GIS and graphics software programs. Cartographic standards and map layouts were inconsistent. Originally, compilation of the geology began using GIS and was completed in a graphics program, where changes

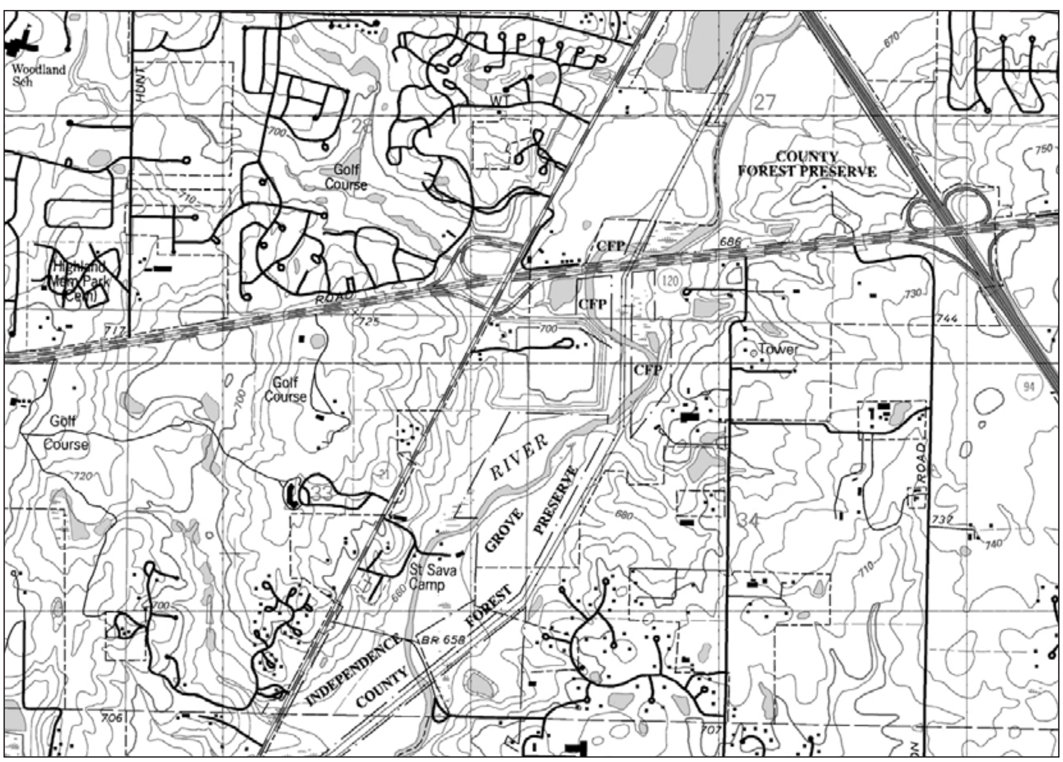


Figure 1. Raster Color Composite (RCC), with green, pink, and gray features layers removed, from Libertyville, IL 7.5-minute quadrangle.

were made without updating the GIS data files. During the past few years there have been significant changes in the map production process at ISGS that have helped to increase efficiency and consistency of the geologic quadrangle maps.

The software programs used at ISGS for creating geologic quadrangle maps are ESRI ArcGIS, Adobe Illustrator, InDesign and Photoshop, and Avenza MaPublisher. The geological units, data points and other geologic features are compiled and symbolized in ArcMap. The completed geology is exported from ArcMap to an Illustrator file, where final adjustments to color and label placement are completed. When necessary changes are made to the geologic features of the map, they are first created in the GIS data file and subsequently re-exported as an updated Illustrator data file. DLG base maps are also coded and symbolized in ArcMap and then exported to an Illustrator file. RFS base map files are converted to bitmap mode files in Photoshop before being placed into Illustrator file format via MaPublisher. All other components such as legends, stratigraphic columns, cross sections and other figures are accomplished using Illustrator. Once the individual Illustrator components are completed as .eps files, they are placed in Adobe InDesign for final layout of the map sheets (Figure 2).

A number of organizational efforts have improved the efficiency of map production at the ISGS. All of the digital data for each geologic quadrangle map are maintained in one location and are organized and named in a consistent manner, thus facilitating the map production procedures. GIS data files are retained as an ArcGIS geodatabase, within which all map editing is completed. A geodatabase template is under development at ISGS that will increase efficiency in map production and also facilitate the movement of mapped data into a statewide geology geodatabase to be available in the future.

In order to maintain a focused effort in the production of each geologic quadrangle map, all of the map compilation is conducted at ISGS by only one or two GIS/graphics personnel, in addition to several GIS/graphics student interns under the direction of the staff cartographer. The cartographer closely monitors all maps currently in production, the staff who are working on each quadrangle map, the stage of review, and what is required to move it on through to publication. The review progress of each geologic quadrangle map is available on-line to all ISGS staff.

Because there are new student interns joining the map production each academic year, there has been a need for detailed documentation for all mapping procedures. Internal documents are now available to interns, new GIS/graphics staff, or to geologists with instructions for each of the procedures involved in creating and producing a geologic quadrangle map. These internal documents, along with sample components and map layouts, will be publicly available on-line by December, 2005, at the ISGS

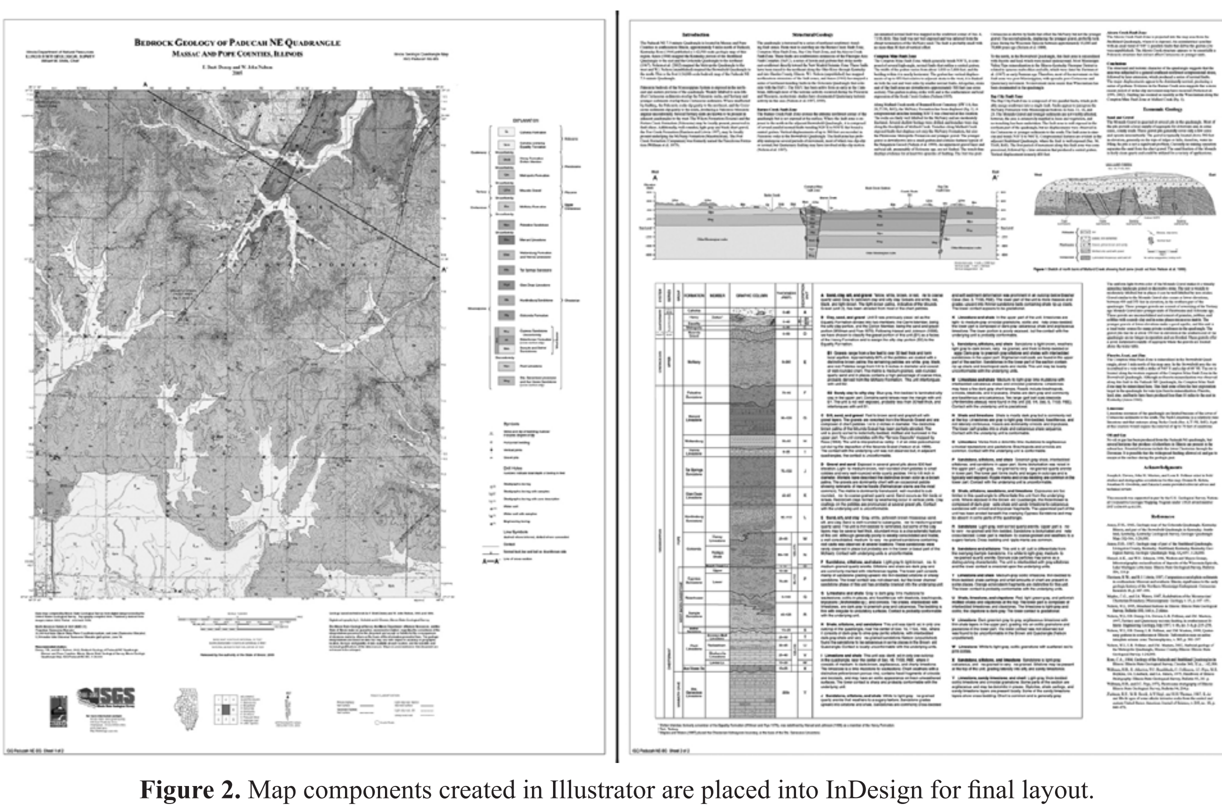


Figure 2. Map components created in Illustrator are placed into InDesign for final layout.

website (http://www.isgs.uiuc.edu/). Such documentation has greatly reduced the amount of time spent training staff in mapping procedures and ensured a more standardized format. This on-line documentation will also be a resource for university geology departments who are participants in the USGS EDMAP program and would like to make use of the mapping standards developed at the ISGS.

A number of standards have been created at ISGS for use in the geologic quadrangle mapping program, and include the following:

- Consistency in production of base map materials
- Consistency in compilation procedures
- Consistency in use of data and geologic unit codes
- Use of ArcGIS geodatabases; digitizing and error validating with topology and topology rule sets
- Use of domains and attribute pull-down menus for quick and consistent coding of geology features.
- Color standards based on the FGDC Digital Cartographic Standard for Geologic Map Symbolization (http://pubs.usgs.gov/of/1999/of99-430/)
- Symbol style sets for geologic features based on FGDC standards
- Symbol style sets for DLG base maps
- Standards for cross sections, columns and legends
- Map sheet layout templates.

PUBLICATIONS AND AVAILABILITY

Geologic quadrangle maps are published in the ISGS Illinois Geologic Quadrangle (IGQ) map series. As of July 2005, a total of 50 IGQ maps had been published since the series was established in 1999. Geologic quadrangle maps that are produced for the STATEMAP program are created in one year, and usually these maps have not gone through IGQ review when they are delivered to USGS as a STATEMAP product. In part to facilitate STATEMAP production, the ISGS established the Illinois Preliminary Geologic Map (IPGM) series in 2004, and a total of 29 IPGM maps have been produced as of July, 2005. IPGM map products are created using the same compilation, cartographic and graphic standards as the IGQ map series, but IPGM maps differ in that they have not completed the full review and revisions necessary for IGQ publication. Many of these IPGM maps will eventually be published as part of the IGQ map series, and in the meantime they are publicly available on-line (see Related Information below).

RELATED INFORMATION

Domier, J. E., 2003, Retiring of the USGS map separates and the emergence of the USGS raster feature separates: presentation at Illinois Mapping Advisory Committee meeting, November 28, 2003, Springfield, Illinois, http://ngmdb.usgs.gov/Info/dmt/docs/domier03.html.

Domier, J. E., 2005, Geologic Quadrangle Mapping at the Illinois State Geological Survey: presentation at Digital Mapping Techniques Workshop, April 26, 2005, Baton Rouge, Louisiana, http://ngmdb.usgs.gov/Info/dmt/docs/domier05.ppt.

Illinois Geological Quadrangle (IGQ) Digital Data Products, Illinois Natural Resources Geospatial Data Clearinghouse, http://www.isgs.uiuc.edu/nsdihome/ISGSindex.html.

ISGS On-Line Geologic Quadrangle Maps (.pdf files), http://www.isgs.uiuc.edu/isgshome/online_maps_data.htm.

ISGS Cartographic Standards for ISGS Geologic Quadrangle Mapping, http://www.isgs.uiuc.edu/ (available fall 2005).

USGS Digital Line Graphs, http://edc.usgs.gov/products/map/dlg.html.

USGS Standards for Raster Feature Separates, http://rockyweb.cr.usgs.gov/nmpstds/rfsstds.html.

Using XML for Legends and Map Surround

By Victor Dohar

Natural Resources Canada
601 Booth Street
Ottawa, Ontario, Canada K1A 0E8
Telephone: (613) 943-2693
Fax: (613) 952-7308
e-mail: vdohar@NRCan.gc.ca

WHAT IS XML?

XML is an acronym for E**x**tensible **M**arkup **L**anguage. Basically, it is a readable text file used to store information in a structured manner. Just as HTML (Hypertext Markup Language) was designed to display data on web pages, XML was designed to store data. It is important to note however, that an XML document by itself does not do anything. It cannot be executed, or perform any function. It is simply a means of storing information and passing it from application to application. Thus, it is widely accepted as a means to allow for the exchange of data between incompatible systems.

The structure and syntax rules of an XML document are fairly straightforward. The information conveyed in an XML document must be enclosed between standard markups, or more commonly known as tags or nodes. The result is a start and end tag with a value in between, forming an element. The start tag can also include element attributes, which are used to describe the value between the tags. The use of tags is important as they allow a computer application (or human) to quickly locate a piece of information, much like a directory structure on a hard disk. Unlike HTML where tags are predefined, XML tags are defined and named by the user or the application that creates the XML document. The syntax rules are not very complicated. Listed below are a few to help you understand the basic rules of an XML document:

1. All XML documents must contain a declaration and one unique root element,
2. All elements must have matching start and end tags,
3. Tag names are case sensitive,
4. All elements must be properly nested,
5. Element attribute values must always be double quoted.

An XML document is considered to be well-formed when none of these syntax rules are broken.

The following is a sample XML document, displaying one root element <Paper> containing three additional elements with some information about this paper. For legibility reasons in this paper, values between the tags are displayed in bold, and nested tags are indented.

```
<Paper>
   <Title>Using XML for Legends and Map Surround</Title>
   <Author>Vic Dohar</Author>
   <Organization>Natural Resources Canada</Organization>
</Paper>
```

The following is a similar XML document with more information:

```
<Conference>
   <Name>DMT ‘05</Name>
   <Papers>
     <Paper>
       <Title> Geologic quadrangle mapping at the ISGS</Title>
       <Author>
         <Surname>Domier</Surname>
         <GivenName>Jane</GivenName>
```

```
      </Author>
      <Organization>Illinois State Geological Survey</Organization>
    </Paper>
    <Paper>
      <Title>Using XML for Legends and Map Surround</Title>
      <Author>
        <Surname>Dohar</Surname>
        <GivenName>Vic</GivenName>
      </Author>
      <Organization>Natural Resources Canada</Organization>
    </Paper>
  </Papers>
</Conference>
```

The two examples above contain the same type of information, yet some information is stored differently. This variance in structure is driven and controlled by an XML Schema. An XML schema is used to define the structure or elements that exist in an XML document. They are the legal building blocks of an XML document as defined by the originator. Schemas define each element, the data type for each element, each element’s attributes, the number of occurrences of an element, whether or not an element is optional or mandatory, its child elements, and the order of elements, just to list a few. XML schemas are also written as an XML document, but are saved with the .xsd file extension, thus they are at times referred to as XSD documents. At the top of an XML document, a reference is usually made to a schema in order to validate the content and structure of the XML document.

The diagram in Figure 1 is a graphic representation of a schema for the above XML document, produced using the software XMLSpy by Altova (http://www.altova.com). This software allows schemas to be created graphically, much like UML (Unified Modeling Language) diagrams. The diagram basically states (from left to right) that the root element is called Conference, and it must contain elements called Name and Papers. Name contains a text string representing the name of the conference, and Papers must contain any number of Paper elements. Each Paper element must contain a Title, an Author, and an Organization element. Finally, each Author element must contain a GivenName and Surname element, along with an optional MiddleInitial element.

XML Resources

The above should provide a basic level of understanding when discussing the use of XML for map surround and legend creation. There are many resources available for you to get a better understanding of XML. Two that I use often when creating applications utilizing XML are W3 Schools (http://www.w3schools.com/xml/default.asp) and the Microsoft Development Network (http://msdn.microsoft.com/xml/). In addition to learning XML, you will also need software to manage, view, and edit XML documents in human-friendly form. Some are free like Peter’s XML Editor (http://www.iol.ie/~pxe/) with limited capabilities, whereas others such as Altova’s XMLSpy charge a fee and have many bells and whistles.

APPLYING XML FOR MAKING MAPS

Using XML for Map Surround Elements

The Publication Process and Integration (PPI) is an electronic web-based system to manage each Geological Survey of Canada (GSC) publication through its vari-

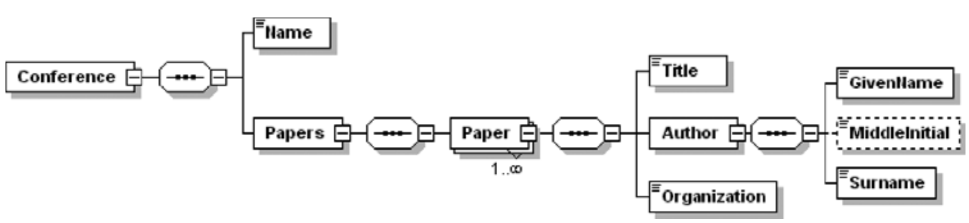


Figure 1. A graphic representation of an XML schema based on the sample XML document produced using XMLSpy software (Altova, Inc.). It clearly displays the relationships between the elements, the order of elements, and the element type. Each element can be dragged and edited in order to create schema variations.

ous stages. The system replaces with web-based forms the many paper submission forms that were required of authors in order to publish reports, open files, bulletins, and maps. The information entered in these web forms is stored in an Oracle database, where it can be extracted to an XML document. Some of the information that is entered is metadata which can be used for generating various map surround elements such as title block and recommended citation.

The following sample XML document generated from Oracle is then used in an ArcMap VBA (Visual Basic for Applications) application to display in ArcMap the title block shown below (see display in Figure 2).

```
<PublicationInformation>
   <Authors>
     <Author>
       <Surname>Smith</Surname>
       <Initial>L</Initial>
     </Author>
   </Authors>
   <Language>english</Language>
   <Bilingual>no</Bilingual>
   <Publication>
     <Series>A-series map</Series>
     <Number>2059</Number>
     <Title>Sandilands</Title>
   </Publication>
   <Map>
     <Feature>surficial geology</Feature>
     <Coverage>
       <District></District>
       <Province>Manitoba</Province>
     </Coverage>
   <ScaleDenominator>100000</ScaleDenominator>
   </Map>
</PublicationInformation>
```

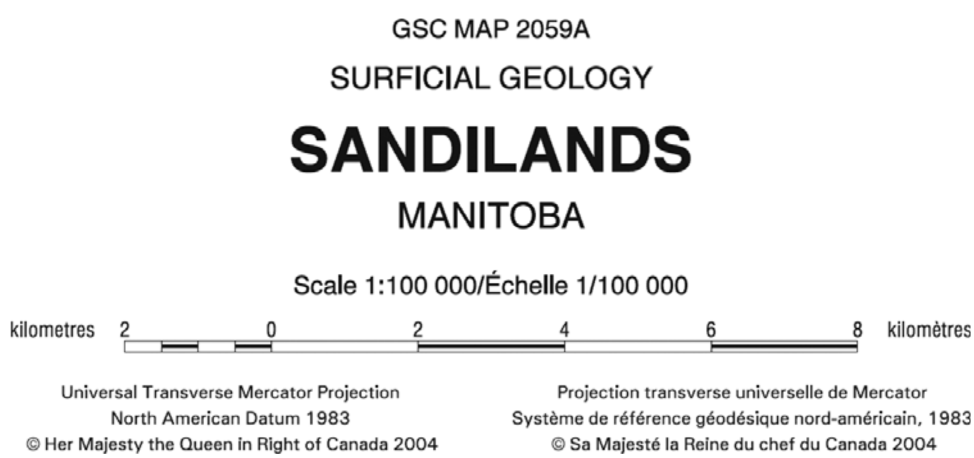


Figure 2. Image of title block from a geological map. The content was extracted from an Oracle database as an XML document. A VBA script in ArcMap generates the title block along with a second XML document that stores the GSC design specifications.

In addition to the above XML document containing the information for the title block, another XML document stored on a central server is used for storing the GSC Design Specifications or the rendering of these elements in ArcMap. This XML document is used to store the properties of these elements; such as font name, font size, colour, justification, indentation, and line spacing. Should a change in design be required, only the values in this XML document need to be updated, without the need to modify the VBA script.

Shown below is an excerpt from the GSC Design Specifications XML document for the map title element of the title block. The same XML schema exists for other surround elements.

```
<GSCDesignSpecifications>
   <TitleBlock>
     <MapTitle>
       <Font>
         <Name>Arial</Name>
         <Style>Regular</Style>
       </Font>
       <Size units="points">24</Size>
       <Colour>
         <Cyan>0</Cyan>
         <Magenta>0</Magenta>
         <Yellow>0</Yellow>
         <Black>100</Black>
       </Colour>
       <LeadingFactor>1.25</LeadingFactor>
       <HorizontalAlignment>HaCenter</HorizontalAlignment>
       <VerticalAlignment>VaBaseline</VerticalAlignment>
       <LineSpacings>
         <LineSpacing>
           <FromElement>default</FromElement>
           <Distance units="points">32</Distance>
         </LineSpacing>
       </LineSpacings>
       <LineLimit units="picas">36</LineLimit>
       <Indent units="picas">0</Indent>
     </MapTitle>
   </TitleBlock>
</GSCDesignSpecifications>
```

The use of these XML documents and the VBA application in ArcMap provides an efficient means of adding this information to maps thereby ensuring quality and consistency in all the maps published at the GSC. The key benefits are that this approach reduces errors and omissions by reducing the need for user intervention, and provides consistent rendering of the information based on established design specifications.

Using XML for Geological Legends

A similar approach utilizing XML documents is used for rendering geological legends in ArcMap. In most instances, the text of a geological legend is initially created by the author/geologist as a Microsoft Word document. By utilizing the styles and formatting capabilities of paragraphs in Microsoft Word, custom formatting styles are created and applied to each paragraph. The custom formatting styles reflect the content of a geological legend (i.e., geological unit description) as well as resembling the geological legend XML schema.

VBA scripting and a toolbar in Microsoft Word allow the user with the click of a mouse to apply the desired custom formatting style to each paragraph. Paragraphs are then formatted visually according to the settings of each style; however it is only meant as a visual aid and has no bearing on the final appearance of the legend in ArcMap (see Figure 3). The important aspect is that each paragraph is formatted correctly. Based on the formatting style applied to each paragraph, a VBA script in Microsoft

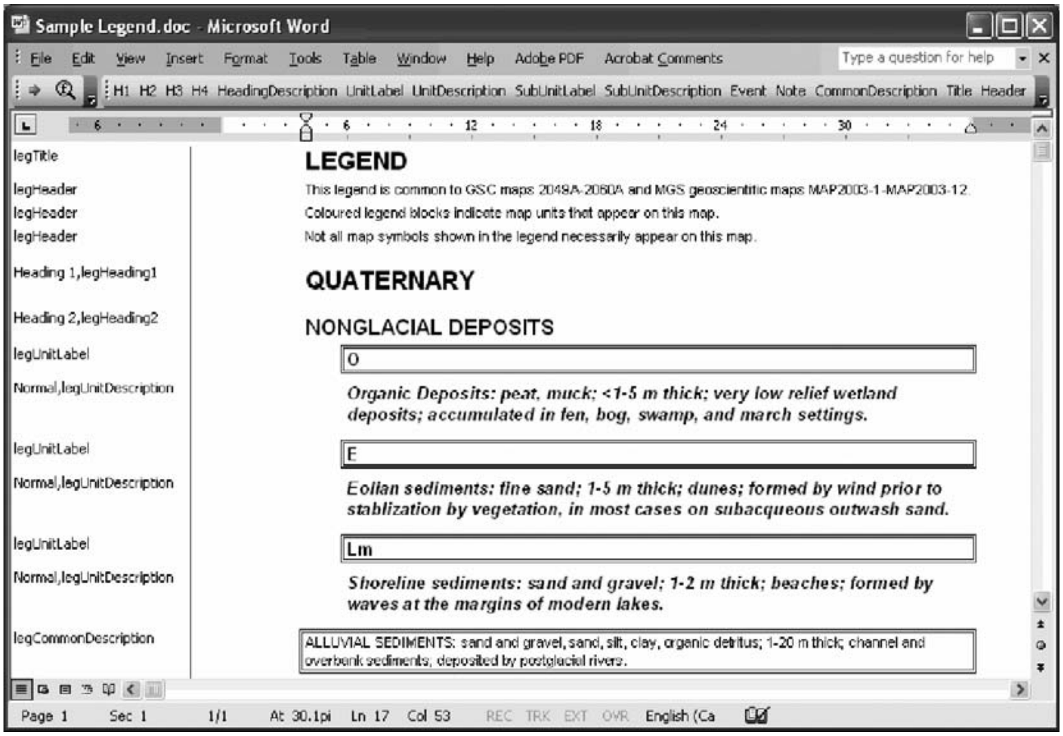


Figure 3. Screenshot of Microsoft Word document, displaying a sample geological legend (shown on right side) that has its paragraphs formatted to custom styles (shown on left side).Also shown is the toolbar for applying a custom formatting style to each paragraph. Based on the formatting, the content of each paragraph is written to an XML document accordingly.

Word transfers the content in each paragraph to an XML document, placing the content within the corresponding element tags (see XML document below that has been translated from the Word document in Figure 3). The XML document in turn is validated against the legend content schema XSD document (see Figure 4) before being processed in ArcMap.

```
<LegendContent>
  <LegendTitle>
    <Title legID="1">LEGEND</Title>
  <Header legID="2">This legend is common to GSC maps 2049A - 2060A, and MGS
        geoscientific maps MAP2003-1 - MAP2003-12.</Header>
  <Header legID="3">Coloured legend blocks indicate map units that appear on
        this map.</Header>
  <Header legID="4">Not all map symbols shown in the legend necessarily ap-
        pear on this map.</Header>
  </LegendTitle>
  <UnitLegend>
    <Heading>
      <HeadingLabel legID="5" level="1">QUATERNARY</HeadingLabel>
    </Heading>
  </UnitLegend>
  <UnitLegend>
    <Heading>
      <HeadingLabel legID="6" level="2">NONGLACIAL DEPOSITS</HeadingLabel>
    </Heading>
  </UnitLegend>
```

```
    <UnitLegend>
      <Units>
        <Unit boxID="1">
          <UnitLabel legID="7">O</UnitLabel>
        <UnitDescription legID="8">Organic deposits: peat, muck; &lt;1-5 m
            thick; very low relief wetland deposits; accumulated in fen, bog,
            swamp, and marsh settings.</UnitDescription>
        </Unit>
      </Units>
    </UnitLegend>
    <UnitLegend>
      <Units>
        <Unit boxID="2">
          <UnitLabel legID="9">E</UnitLabel>
          <UnitDescription legID="10">Eolian sediments: fine sand; 1-5 m thick;
            dunes; formed by wind prior to stabilization by vegetation, in most
            cases on subaqueous outwash sand.</UnitDescription>
        </Unit>
      </Units>
    </UnitLegend>
    <UnitLegend>
      <Units>
        <Unit boxID="3">
          <UnitLabel legID="11">Lm</UnitLabel>
        <UnitDescription legID="12">Shoreline sediments: sand and gravel; 1-2
            m thick; beaches; formed by waves at the margins of modern lakes.</
            UnitDescription>
        </Unit>
      </Units>
    </UnitLegend>
    <UnitLegend>
    <CommonDescription legID="13">ALLUVIAL SEDIMENTS: sand and gravel, sand,
            silt, clay, organic detritus; 1-20 m thick; channel and overbank
            sediments; deposited by postglacial rivers.</CommonDescription>
    </UnitLegend>
</LegendContent>
```

In ArcMap, a VBA script is used to generate the geological legend (see Figure 5) using three XML documents. The content of the legend is extracted from the XML document generated from Microsoft Word (described above). The rendering or design specifications of the legend (i.e., fonts, colours, legend box sizes, line spacing) is obtained from the GSC Design Specifications XML document noted above. A third XML document is used to control the layout of the legend on the paper. This is used primarily for the legend's location on the paper, number of columns, and aligning geological units chronologically in multiple columns. In addition, when the VBA script generates the legend, the symbology used for each of the geological units in ArcMap is transferred to the legend.

It is important to note that the legend created by this method is not dynamically linked to the ArcMap table of contents (TOC). If any edits are required to the legend, either to the content in the Word document, or the symbology of a geological unit in ArcMap, the simplest task is to delete the current legend from ArcMap and regenerate the legend with the updated XML documents and ArcMap symbology. This method utilizing three XML documents ensures a consistent level of quality and output from ArcMap.

Next Steps

The next steps in using XML documents for geological legend generation is to complete and fine tune the VBA scripting in ArcMap. After doing so, the XML schema for the legend will be expanded to include geological and mineral symbols that also occur on maps. Since the content of the legend exists in an established XML schema, other applications can be developed, such as a

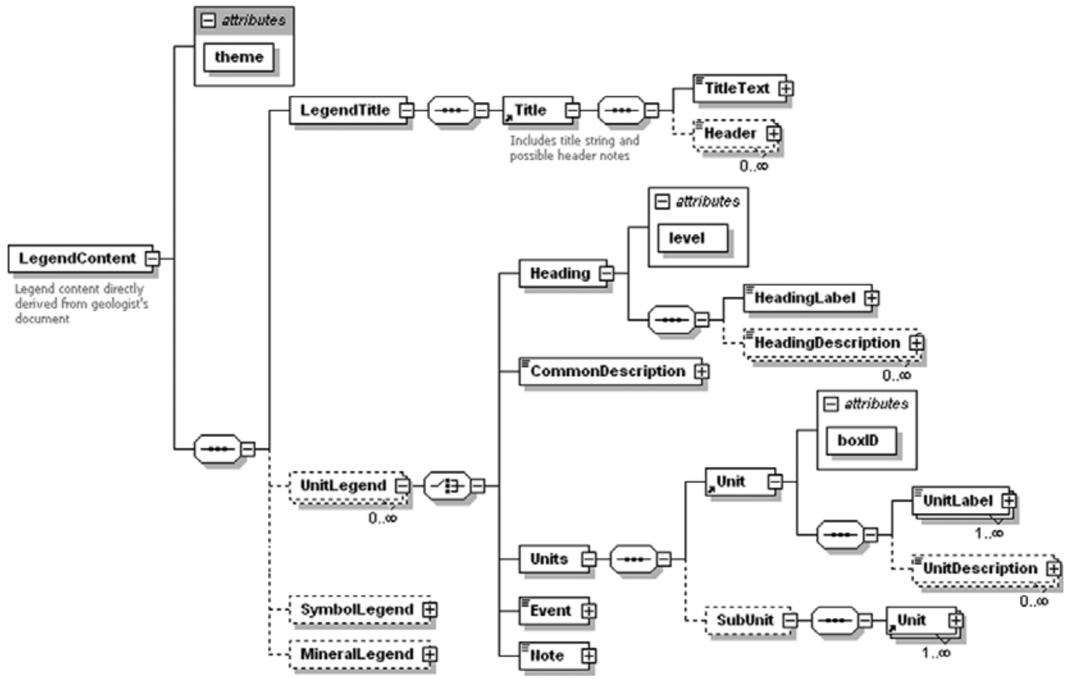


Figure 4. XML schema representing the structure of XML documents for legend content schema, generated using XMLSpy. XML documents that are generated from Microsoft Word are validated against this schema before being processed in ArcMap. This schema diagram states which elements are required (boxes with solid outline), those that are optional (boxes with dashed outline), the number of occurrences of each element (0…∞), and the lineage between elements (symbols between elements indicating either a choice, or a sequence).

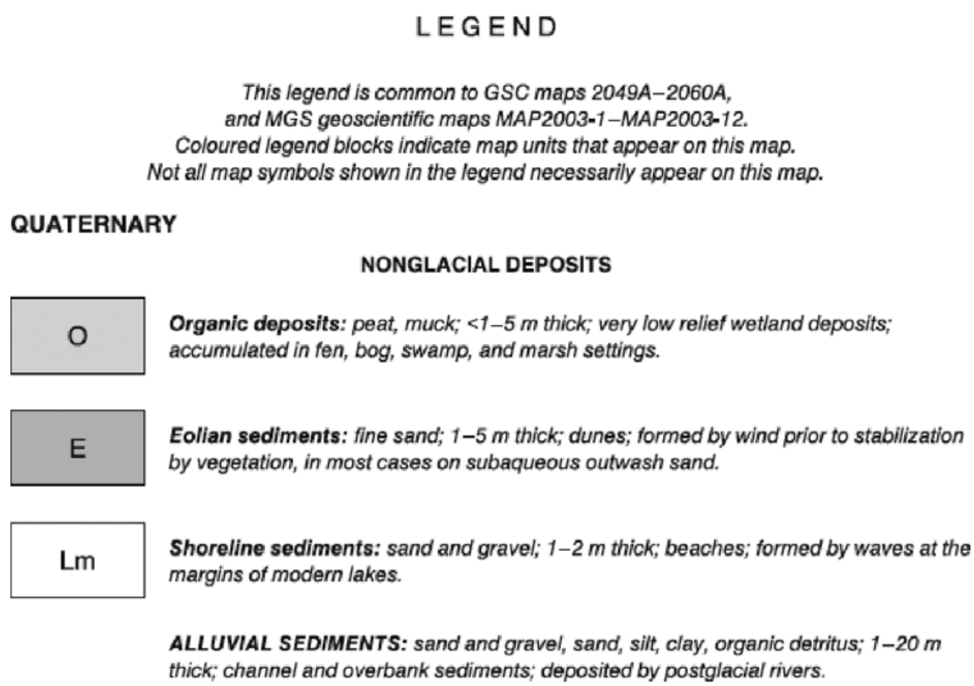


Figure 5. Top portion of a legend for a published surficial geological map. (NOTE: The legend in this image was not produced in ArcMap, as the VBA script is still in beta testing status. The goal is to achieve results similar to current production methods.)

customized query tool either for ArcMap or web mapping. By having data stored in a structured manner and widely accessible, the possibilities are limitless.

REFERENCES

ArcMap and ArcGIS, ESRI Inc., http://www.esri.com.

Microsoft Development Network, XML Development Center, http://msdn.microsoft.com/xml/.

NADM Data Interchange Technical Team, 2003, XML Encoding of the North American Data Model, in D.R. Soller, ed., Digital Mapping Techniques ’03—Workshop Proceedings: U.S. Geological Survey Open-File Report 03-471, p. 215-221, available at http://pubs.usgs.gov/of/2003/of03-471/boisvert/index.html.

Peter’s XML Editor, Peter Reynolds, http://www.iol.ie/~pxe/.

W3 Schools, XML Reference, http://www.w3schools.com/xml/default.asp.

XMLSpy, Altova, Inc., http://www.altova.com.

MAPublisher 6.1—Presentation Exercise

By David Andrec

Avenza Systems Inc
124 Merton Street
Toronto, Ontario M4S 2Z2
Telephone: (416) 487-5116
e-mail: david@avenza.com

The following is a summary of the Avenza MAPublisher presentation given at the Digital Mapping Techniques 2005 conference in Baton Rouge, LA, on April 26, 2005. As many of those present were either already users of MAPublisher or were familiar with it in one way or another, it was decided that a short presentation outlining some of the newest features in MAPublisher 6.1, in the form of a small forum-specific exercise would serve everyone well. It is assumed that users of this Exercise are familiar with Adobe Illustrator. All data files used in these examples can be found at ftp://ftp.avenza.com/pub/misc/dmt05_files.zip.

DATA SOURCE: Geologic data mapped by the National Wetlands Research Center for the Louisiana area (http://sabdata.cr.usgs.gov/sabnet_pub/pub_sab_app.aspx?prodid=1403), in ArcInfo (v.7.0.4) export format. Map projection was Datum NAD27, Projection UTM, Feature Polygon, Units Meters, Resolution 15 Meters.

PART 1: IMPORTING THE DATA

1. Create a new Adobe Illustrator document.
   Custom > Landscape format
   17W x 15H
   Reset Rulers

How to Import Data in a Similar Projection/Co-ordinate System:

2. (SIMPLE IMPORT) Import the "gelogicpoly.shp" shapefile from the "projected" data folder and make the line 50% blk, 0.25pts wide.

Note: auto recognition of projection/co-ordinate system (Nad 27 Zone 15 North, Meter). Data will import with current fill and stroke properties as currently set in Illustrator.

3. (SIMPLE IMPORT) Import the "geologypoint.shp" shapefile from the "projected" data folder using the "Add To:" Destination Map View and choosing the "gelogicpoly" Map View.

Note: this will import the 2nd file into the existing Map View and register it to the existing data, which is desirable as both files are in the same projection. The point data will import as the default of a small solid black symbol.

How to Rescale and Reposition all Data Located Under a Map View:

4. (EDIT MAP VIEW) Open "gelogicpoly" Map View and change the name of the View to " Louisiana" and enter a new scale value of 1: 1,300,000 choosing the LL center justification from the "LL Corner" control widget.

Note: The Map View is renamed and the map data is rescaled and moved to the left centre of the art board.

How to Import Multiple Data Files in a Different Coordinate System and Reproject to Match That of Existing Data:

5. (ADVANCED IMPORT) Select both the "rivers.shp" & "rrline_arc.shp" shapefiles from folders under the "unprojected" data folder, choosing the "Use existing" option for the Destination Map View. Specify the "Louisiana" Map View as the choice. Make the rivers blue 0.25pts wide and the rrline brown, 0.75 pts wide.

Note: this will import both the unprojected "river" and "rrline" shapefiles and reproject them to match the projection and locational settings used in the "Louisiana" Map View. The data is imported and reprojected into the chosen view.

How to Import Data in a Different Coordinate System and Reproject to Match That of Existing Data By Dragging the Data From One View to Another:

6. (ADVANCED IMPORT) Import the"rdline.shp" shapefile from the "roads" folder within the "unprojected" data folder using the "New based on" option for the Destination Map View. Make rdline black, 1.5 pts wide.

Note: New Map View is created (latitude/longitude_ Degree) and the states layer is contained in it.

7. (OPEN MAP VIEW - "rdline") Open the view to display that the view is in LatLong format.

8. (REPROJECT DATA – drag & drop) Select the "rdline" data layer in the Map Views palette and drag and drop it into the "Louisiana" view. Delete the now empty "rdline" view.

Note: data is reprojected to (Nad 27 Zone 15 North, Meter) from (latitude/longitude_Degree) and is added to the view.

How to Import Data and Easily Create an Inset Map Based on an Existing Views Projection:

9. (SIMPLE IMPORT) Import the"states.shp" shapefile from the "inset – usa" map folder found within the "unprojected" data folder, using the "New based on" option for the Destination Map View. Data is fit to page based on new View. Double click on the "States" View to open it. Enable the "same as" option at the bottom of the dialog and choose "Louisiana".

Note: Data is reprojected to same (NAD 27 Zone 15 North, Meter).

Rename the view to "Inset" and enter a scale value of 1: 40,000,000, choosing the top right justification from the "LL Corner" control widget.

Note: User may move inset around by moving the green data extents box in the Map View dialog.

PART 2: STYLIZING DATA

10. (OPEN FILE) Open "Louisiana.ai" file (file contains all imported data above and has been slightly color modified).

11. (OPEN STYLE FILE) Open the "samplestyles.ai" file located in the "USGS KEYNOTE\Data\styles and symbol samples" folder.

12. (LOAD STYLE FILE) Load the "samplestyles.ai" file located in the "USGS KEYNOTE\Data\styles and symbol samples" folder into the "Louisiana.ai" file.

Note: Drag these user created styles into the current default style list so that they can be used in stylesheets.

13. (CREATE STYLESHEET AREA) In the Map Style Sheet Palette create a new stylesheet named "geologic" of type "Area" and drag the "geologicalpoly_area" layer under this stylesheet.

Note: Use the "Geologic Code" attribute column to assign values in the stylesheets. The attribute column is automatically remembered when assigning. Use the category names listed on the right side of the map and sequentially assign the attribute to the accompanying style until complete, then click "Apply".

14. (CREATE STYLESHEET POINT) First load the "samplesymbols" symbol set into the AI document. In the Map Style Sheet Palette create a new stylesheet named "symbols" of type "Point" and drag the "geologypoint_ point" layer under this stylesheet. Now assign the symbols from ECF-EJ-L. Click Apply and Ok.

Note: Use the "Geology_ CO" attribute column to assign values in the point stylesheets. The symbols being assigned are those in the current symbol library and they can be scaled.

PART 3: MANIPULATING POINT DATA

How to Rotate Point Data Based on Attribute Values as Well as Joining Point Data Based on Attribute Values:

15. (ROTATE SYMBOLOGY- Based on attribute column) Make the "geologypoint" layer active in the Illustrator doc and lock all other layers. Select all on layer. Open Edit Map Columns and set the following options:

Layer->"geologypoint"
Expression Column->Rotate_By
Result Type->Properties
Result Property->Rotation
Click Ok

Note: The symbols rotate according to attribute value.

16. (IMPORT POINT – GPS WAYPOINTS) Continue using the "Louisiana.ai" file. Use SIMPLE IMPORT to import the "gpspath.txt" file located in the following location USGS KEYNOTE\Data\Louisiana\Unprojected\gps waypoint.

Import this file as a new Map View.
Assign "SOURCE PROJECTION" as "WGS84 Lat-Long" click "Ok"
Now drag this layer into the " LOUISIANA " Map View (data gets reprojected)
Now while points selected rescale to 35% of original size (Transform Each)
Zoom in to show point locations
Click Apply and Ok

Note: Use the "Geology_CO" attribute column to assign value to the point stylesheets. The symbols being assigned are those in the current symbol library and they can be scaled.

17. (JOIN POINTS – GPS WAYPOINTS) Select all the points on the "gpspath_point" layer and then open the Join Points dialog (Filter->Map Lines->Join Points) and make the following settings:

Input Layer->"gpspathpoint"
Output Layer->New layer
Group By->Group
Sort By->Sequence
Click Ok (points are joined)
Use Object->Path->Simplify to experiment with smoothing the line.

18. (OPEN FINISHED FILE) Open the "Louisianafinal" file and view finished map sample.

A New Day Has Come! The New Marine Map Standards for Natural Resources Canada-Geoscience for Oceans Management (GOM)

By Gary M. Grant, Phil O'Regan, Sheila Hynes,
John Shaw, Brian J. Todd, and Phil Moir

Geological Survey of Canada (Atlantic), Natural Resources Canada
1 Challenger Drive, Room M330
Dartmouth, NS, Canada B2Y 4A2
Telephone: (902) 426-2773
e-mail: gagrant@nrcan.gc.ca

Earth Sciences Sector Contribution No. 2005148

INTRODUCTION

Surficial geology mapping at the Geological Survey of Canada (Atlantic) or GSCA, has gone through several revolutions since inception in the early 1970s. Then, extensive mapping was based on echo sounder data, complemented by data from low-resolution sub-bottom profiling systems and sidescan sonar systems. The advent of high-resolution sub-bottom profilers (e.g., the Huntec DTS, or Deep-Towed System) augmented the suite of tools available to marine geologists.

The standard outcomes of the work completed between the 1970's and the mid 1990's were surficial geology maps, accompanied by reports. These maps commonly allocated areas of the sea floor to one of a number of formations (e.g., Scotian Shelf Drift Formation). Workers uncomfortable with the formation concept used genetic schema which had similar outcomes. For example, a regional equivalent of the Scotian Shelf Drift Formation was mapped as ice-contact sediment.

The advent of multibeam bathymetry mapping in the mid-1990s provided geologists with images of the sea floor that had unprecedented clarity. From the beginning, multibeam bathymetric data have been portrayed in different formats at various scales, perhaps reflecting the absence of a systematic offshore mapping program. The advent of mapping of larger areas, allied with the possible advent of systematic mapping involving collaboration between GSCA and the Department of Fisheries and Oceans (DFO) (similar to the collaboration involving echo sounder data in the 1970s), revealed a pressing need for a more systematic approach to the content and appearance of multibeam imagery, and in particular, standardization of the types of maps produced (Shaw and Todd, 2003).

MULTIBEAM BATHYMETRIC SURVEY

Multibeam bathymetric survey techniques used in the St. George's Bay area (Figure 1) provide a rapid means of determining the morphology and nature of the sediments on the seafloor (Figure 2). A multi-element transducer provides many (30-150) individual soundings of the water depth and echo strength for each ping. Automatic seafloor tracking programs determine depths and echo strengths for each transducer element, correct for transducer motion, and calculate a geographic co-ordinate for each individual sounding.

A wide swath (up to 7 times the water depth) can be surveyed in a single pass through an area. Survey lines are spaced to provide overlapping coverage of the seafloor. The data are used to generate high-resolution images that contain information about the morphology of the seabed.

BATHYMETRY MAP

The sun-illuminated DEM of bathymetry (Figure 3) is based on a GEOTIFF exported from the GSCA GRASS software. Many analysts prefer grey-scale images for detection of morphologic features on the sea floor, arguing that colour can mask variability. Commonly, hot-to-cold colour ramps are used (also referred to as "rainbow" colour ramps). In this scheme, blue is used to represent the deep areas, green the intermediate depths, and red the shoals. Some problems arise concerning the range of depth values represented by individual colour bands. Since the inception of multibeam processing at GSCA, the rainbow scheme has been used to represent depth. When a colour bar is applied in the GRASS software, the rainbow scheme is automatically used.

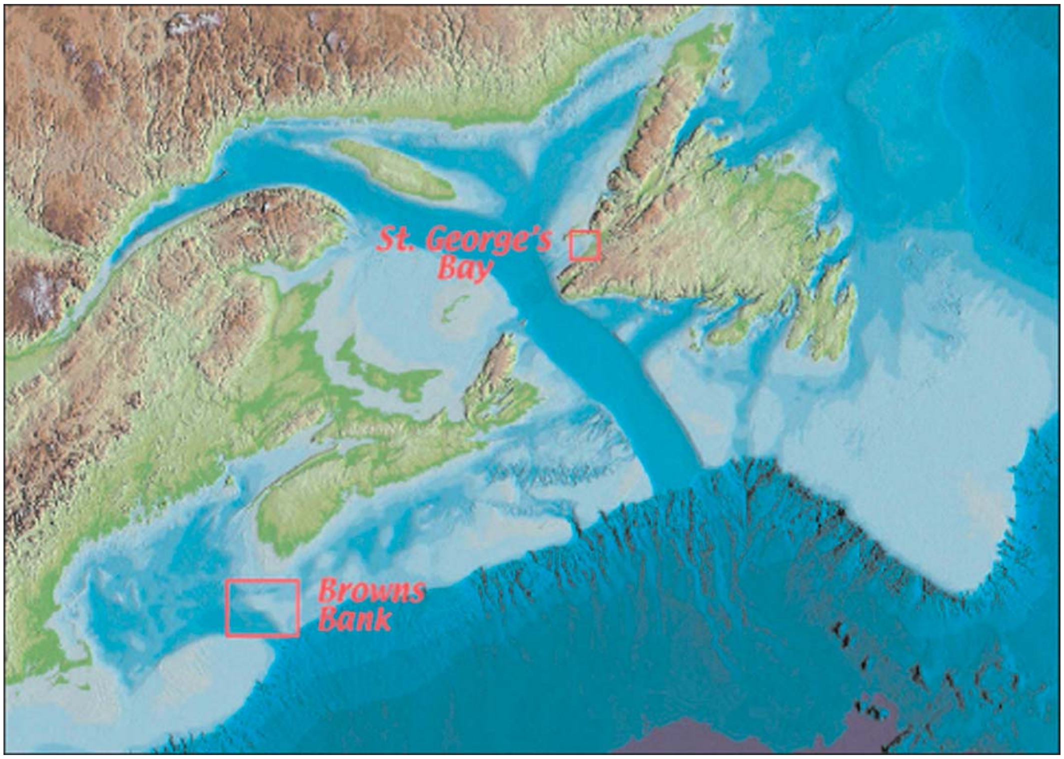


Figure 1. Location map showing two sample map areas, St George's Bay and Browns Bank (from Shaw and Courtney, 2004).

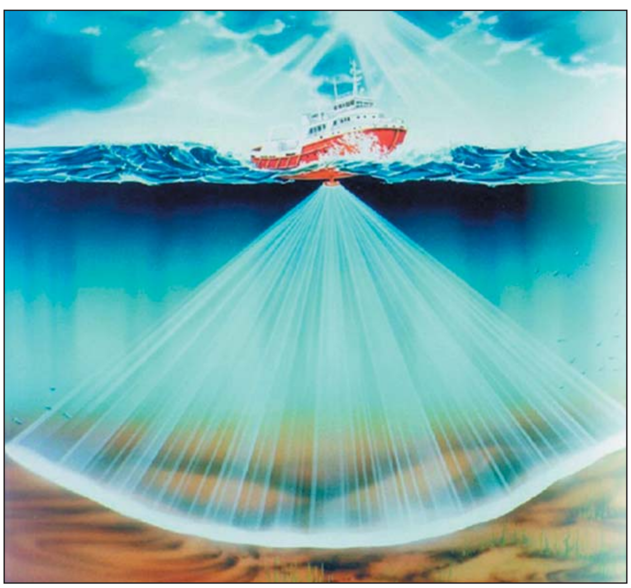


Figure 2. Multibeam bathymetric survey techniques provide a rapid means of determining the morphology and nature of the sediments on the seafloor. A wide swath (up to 7 times the water depth) can be surveyed in a single pass through an area (used with permission of Kongsberg Simrad).

In many multibeam bathymetry map areas, much of the bathymetry variation is in a narrow band. For example, on the Scotian Shelf banks, the predominant depth may be 100-110 m while intervening basins and troughs may be 200-300 m deep. However, if colours are selected automatically, all depths on the bank might be allocated the same shade of red, and this obscures important morphologic variability. It would be preferable to have the subtle depth variations on the bank highlighted by allowing greater colour variation over the 100-110 m interval, for example from red to green. Colour distribution optimization can be achieved by first determining the hypsometric frequency of the data and then tuning the colour bar accordingly.

BACKSCATTER MAP

GSC has traditionally represented backscatter on a grey scale. The usual practice has been to portray areas of high backscatter (commonly gravel and rock) as dark tones, and to portray areas of low backscatter (sand and mud) as light tones (Figure 4). This approach parallels the customary depiction of the sea floor on sidescan sonograms: areas of high reflectivity (gravel, rock, shipwrecks etc.) have dark returns, whereas muddy and sandy sea floors with low reflectivity have light returns. The inverse approach is used by University of New Brunswick Ocean Mapping Group, who commonly depict mud as dark-

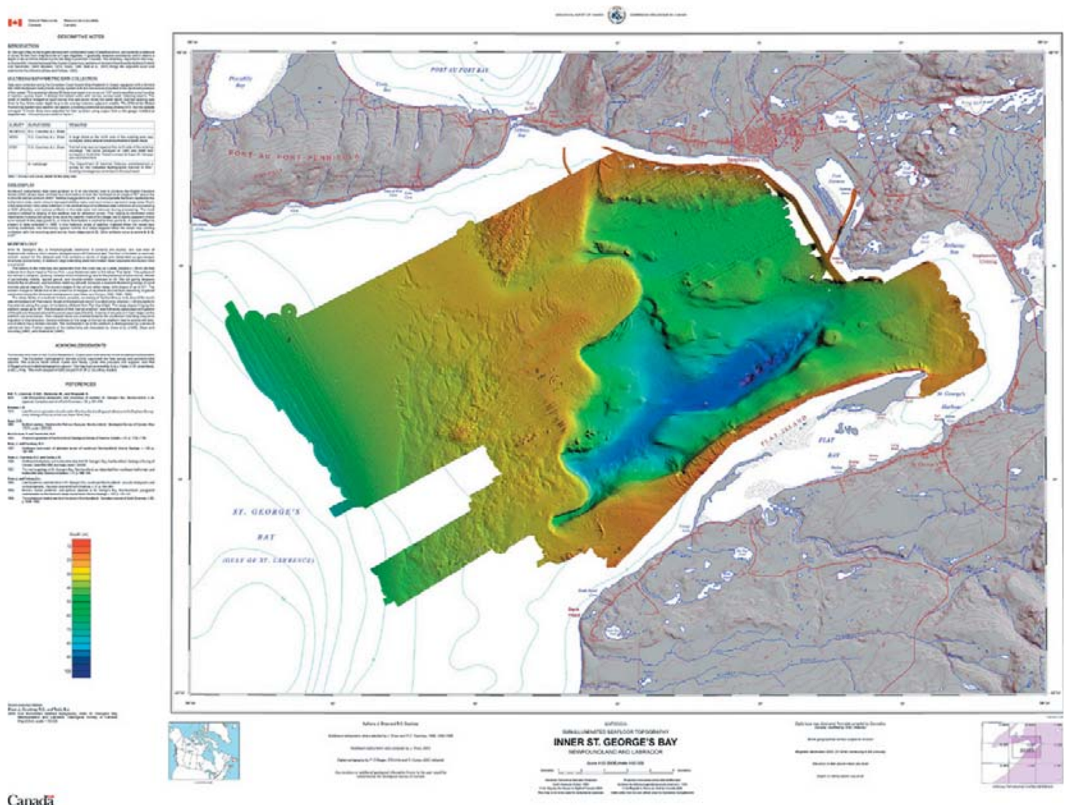


Figure 3. Bathymetry map—the sun-illuminated DEM of bathymetry is based on a GEOTIFF exported from the GSCA GRASS software. Depth is represented by colour. Additional information includes isobaths (contours of depth), a representation of land areas, and isobaths from Canadian Hydrographic Service nautical charts. In the printed version of the Proceedings, these figures are in black and white; however, the web version shows these figures in full colour.

toned. There are no advantages of one method over the other: it is mostly a matter of personal preference, in which those accustomed to examining sidescan sonograms prefer the first option, However, backscatter is also very effectively represented using colour (Figure 4).

SURFICIAL GEOLOGY MAP

This map (Figure 5) is the vehicle for an interpretation of the imagery. We take the shaded relief imagery and explain what is seen in terms of geological processes. In other words, it is the opportunity to turn information into knowledge, with several approaches possible:

- Pragmatic approach: Delineate, describe, and interpret features observed on multibeam imagery in a pragmatic fashion.

- Surficial materials approach: Show the distribution of materials at the sea floor as remotely observed with cameras, ROVs, sidescan sonar systems, etc., or collected by grab samplers and corers.

- Formations approach: Segregate the image into formations (lithologically similar stratigraphic units). Previous mapping on eastern Canadian shelves has resulted in formational frameworks that have regional validity.

- Genetic approach: Apportion elements of a stratigraphic framework to the imagery, that is, divide the image into areas of surficial sediment with common genesis. Mostly this results in classification according to age: e.g. glacial sediments formed in contact with grounded ice, proximal and distal glaciomarine sediments, and postglacial sediments. This approach is used in the example shown in Figure 5.

BENTHIC HABITAT MAP

Based on the sea floor sediment maps and statistical analyses of benthos (Figure 6), habitats and corresponding associations of benthic animals are mapped. In the case of the Browns Bank habitat map, each of the habitats is

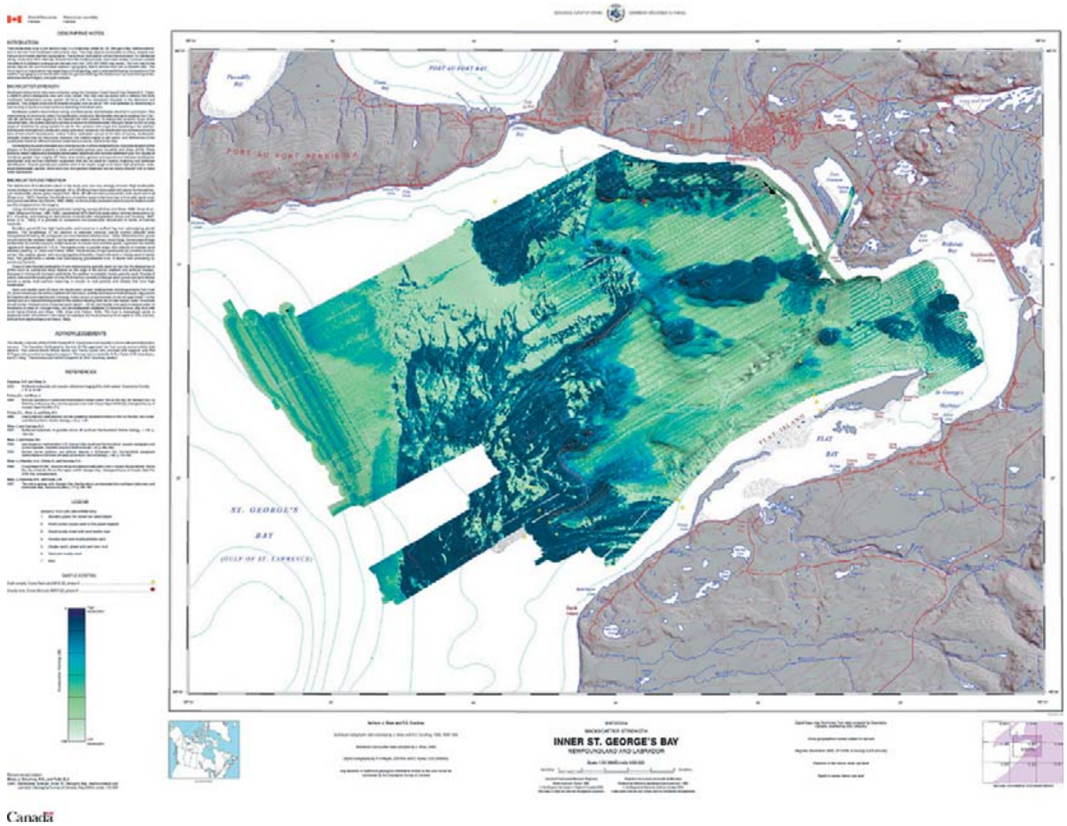


Figure 4. Backscatter map—backscatter is very effectively represented using colour. Here an indigo to pale green colour scheme is applied, so that indigo represents high backscatter and pale grey represents low backscatter. There are also some advantages to draping the coloured backscatter over the sun-illuminated topography.

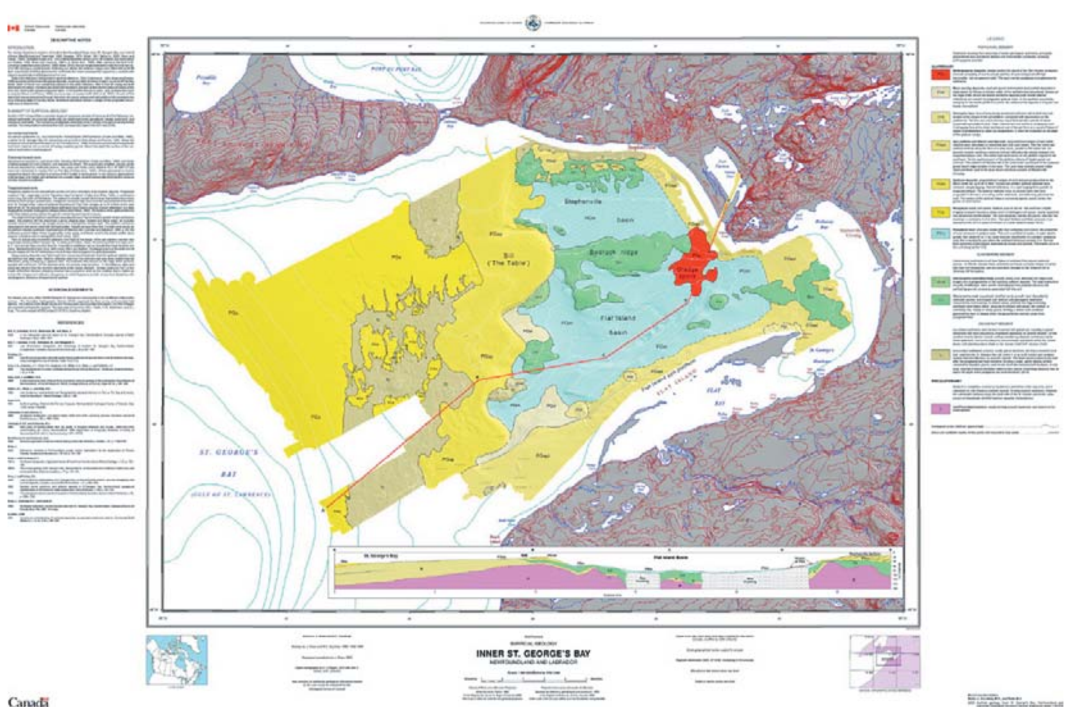


Figure 5. Surficial Geology map—this map shows how the shaded relief imagery is used and explains what is interpreted in terms of geological processes.

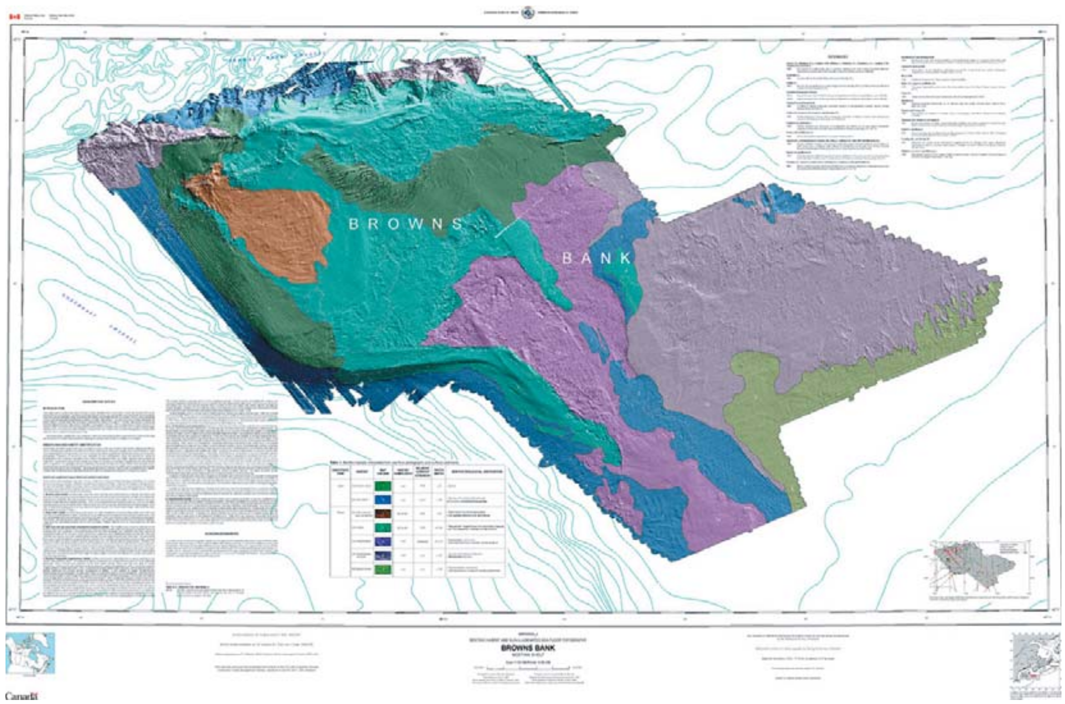


Figure 6. Benthic habitat map—based on the sea floor sediment maps and statistical analyses of benthos, habitats and corresponding associations of benthic animals are mapped.

distinguished on the basis of substrate, habitat complexity, relative current strength and water depth. Different locations may require utilization of other information to distinguish habitats. The spatial allocation of samples, and abundance and commonness of species are used as additional guidelines for identification of habitat zonation.

MARINE MAPPING SYMBOL SET

The aim of this symbol set (Figure 7) is to promote symbol design and usage which will help to eliminate confusion and inconsistency on published maps, as well as promote a more quantitative definition of interpreted information.

A large volume of marine geological data has accumulated worldwide over the last number of decades and has stimulated the publication of large numbers of interpreted marine geological maps. While the raw data for these interpretations may derive from numerous sources—government surveys, the offshore oil and gas industry, the offshore mining industry—it has generally been the task of government agencies to compile and synthesize the information into map form. A characteristic of marine geological maps, particularly those representing aspects of the surficial geology, is the diversity not only of the symbols used to represent features or zonations, but also in the levels of interpretation implied by those symbols. For example, where one worker may use an "S" to denote only the occurrence of scouring on the seabed, another may mark the location with an "X" varying in size according to the typical scour width, and qualified by an adjacent number indicating scour density at that location of depth of scour. Reconciliation of these approaches is not only frustrated by the difference in representative symbols, but also in the amount of information conveyed (Fader and Peters, 1988).

ON-LINE ACCESS TO THE MAP PRODUCTS

The new policy of Geological Survey of Canada is to provide free, on-line access to all new map products (Figure 8). When completed these maps will become part of the national Geoscience Data Repository (GDR), a collection of geoscience databases managed by a series of standardized Information Services (http://gdr.nrcan.gc.ca/). The aim of the GDR is to standardize corporate geoscience databases and make them interoperable in order to increase the discovery, access, and use of the information that the GSC has been mandated to collect and maintain. The maps will be available on-line as printable image files (PDF, MrSid), and digital GIS layers accessible through Open Geospatial Consortium (OGC) standards such as Web Map Service or WMS.

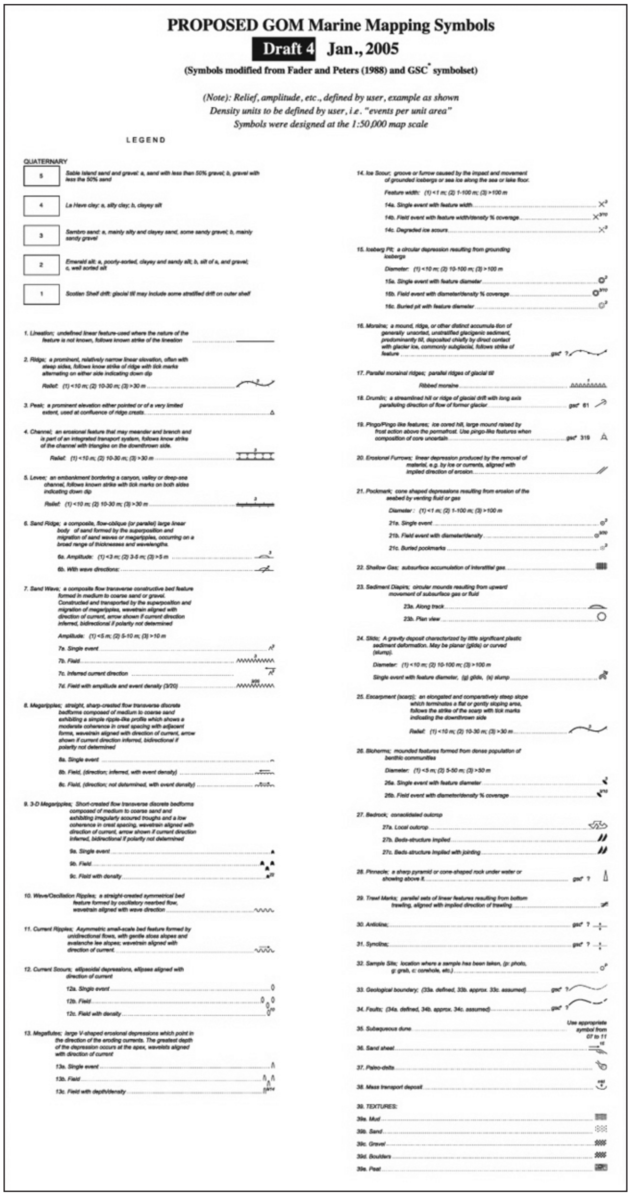


Figure7. Marine mapping symbol set—the goal of this symbol set is to promote use of symbol design, which will help to eliminate confusion and inconsistency on published maps, as well as promote a more quantitative definition of interpreted information.

TECHNICAL NOTES ON THE PRODUCTION OF THE FOUR A-SERIES MAPS: BATHYMETRY MAP, BACKSCATTER MAP, SURFICIAL GEOLOGY MAP AND BENTHIC HABITAT MAP

GSC ‘A-series’ maps are maps that are subject to critical review by at least one scientific authority. They are multicoloured maps printed on demand on colour ink jet plotters or on quality stock when large pres-runs of several hundred copies are needed. ‘A-series’ maps are also available in digital format on CD-ROMs and even, in some cases, released on the Internet. Instructions on database standards and procedures for geological map production are available on the Internet at http://www.nrcan.gc.ca/ess/carto/specifications_e.html.

Bathymetry Map (Figure 3): The sun-illuminated DEM of bathymetry is based on a GEOTIFF exported from the GSC (Atlantic) GRASS (GIS) software. Depth is represented by colour. The DEM’s Geographic projection, datum, etc are made to match the base map at the time of importing (via “Add Data”) in ArcMap.

Backscatter Map (Figure 4): High backscatter (high reflectivity-gravel rock, etc.) is represented by indigo blue and low backscatter (low reflectivity-mud, etc.) by the pale grey/green. The backscatter values are draped over the greyscale multibeam bathymetry.

Surficial Geology Map (Figure 5): The interpretation of surficial geology is shown as coloured polygons that are draped over a greyscale map of the multibeam bathymetry. The geological data polygon boundaries can be as acquired by heads-up digitizing, in ArcMap on a separate layer with the backscatter duotone underneath as the guide to accurately digitizing the features. A second method can also be used; this is making a hard copy print of the backscatter and overlaying it with a sheet of translucent mylar (plastic). With the mylar taped securely to the backscatter duotone print, the outline for the geological features can be traced with a black ink technical pen or a fine tip marker. This overlay is then scanned, put through ArcScan to vectorize the polygon boundaries and with the latitude and longtitude control points is now added (rubbersheeted) as data (in a feature class) to ArcMap.

Benthic Habitat Map (Figure 6): Benthic habitat interpretation polygons were draped over a greyscale multibeam bathymetry of the mapped area and were compiled in the same way as in Figure 5 and then added to ArcMap.

For the land area a greyscale DEM was used as an underlay to the culture, drainage, etc. The greyscale DEM was derived from processing the elevations of the topographic contours layer. To allow the overlay, culture, drainage, etc., data to be viewed clearly, a transparency value is applied to the greyscale DEM; in most cases a

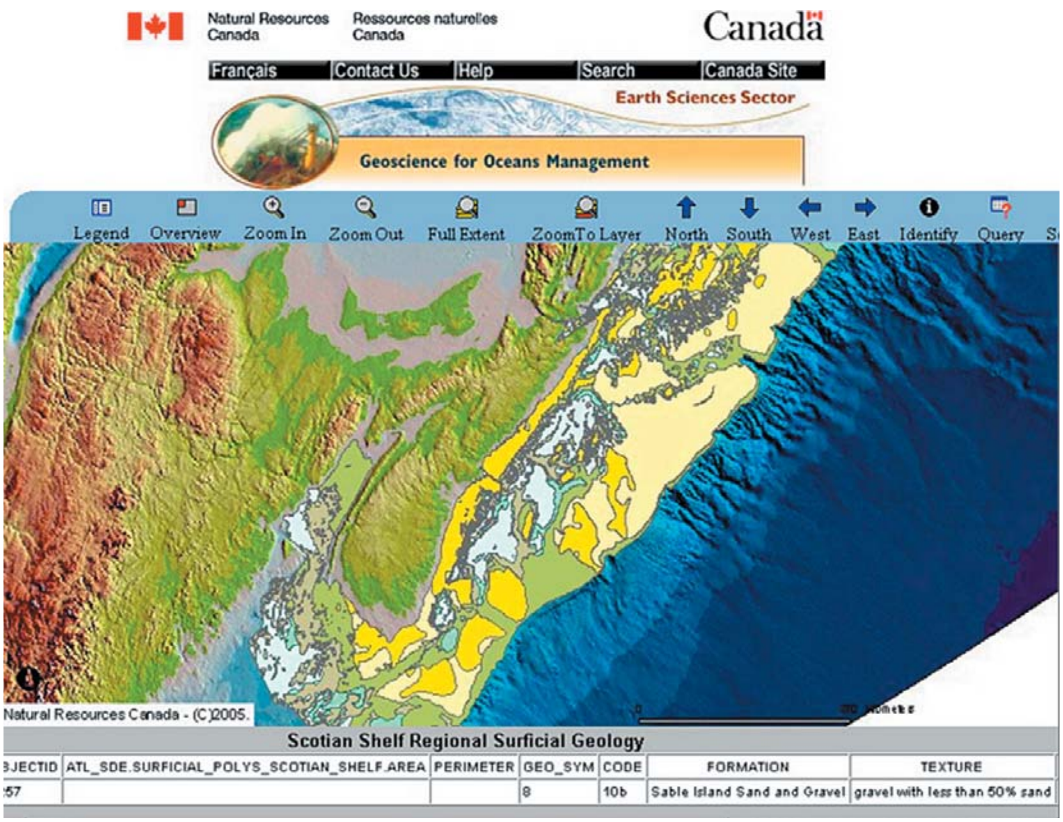


Figure 8. Online access to the map products—the policy of Geological Survey of Canada is to provide free, online access to all new map products. When completed, these maps will become part of the national Geoscience Data Repository (GDR), a collection of geoscience databases managed by a series of standardized Information Services (http://gdr.nrcan.gc.ca/).

value of 50% is adequate.

The map border, at present, is made in the GSC's Geological Mapping System "GEMS" routine, (via "Create Border"), in ArcInfo. The border is produced as an ArcInfo Coverage, complete with coordinate text, ticks, neatline, etc. This border coverage is added to the map folder in ArcCatalogue and then the border arc file and polygon file are saved into an ArcMap feature class/feature dataset, the same is done for the neatline and coordinate annotations.

Once the border is made to the required map extents (map area), then the base data is added, this data can be acquired from the National Topographic DataBase (NTDB) topo maps data library, Geomatics Canada, Ottawa www.cits.rncan.gc.ca. These data layers, topographic contours, drainage, culture and annotations can be added as an arc, node or annotation in a feature class. If the personal geodatabase for the map is set up correctly, then the feature class is inside a feature dataset that is inside the personal geodatabase (i.e. personal geodatabase—feature dataset—feature class, this is how these three items are seen in ArcCatalogue and ArcMap Table Of Contents). The personal geodatabase should be the first thing set up when a new map is started.

These file structure/management issues are usually addressed in ArcCatalogue and then added (via "Add Data") to ArcMap. There is a third ArcGIS tool used—ArcToolbox, an ArcGIS module that is used for file conversion, analysis and data management; the functions include Analysis (surface-cut and fill), Conversion (Import to Geodatabase), Data Management (Projections-Spatial adjustment or rubbersheeting).

The Location map (a small map of Canada and its offshore) for A-Series maps is now a layer (linked through a database at GSC (Ottawa) via the internet) that can be added to ArcMap (via "Add Data"). The location dot, (the study area), of the map is linked to the neatline of the map being drawn.

At this time, the National Topographic System (NTS) index map is not linked to the GSC database (Ottawa). It is made in the "GEMS" routine, (via "Create NTS Index Map") and output as a .eps file from ArcInfo, imported to CorelDraw and placed in the ArcMap map surrounds as a linked graphic.

The surround text, (descriptive notes, references, title blocks, etc.) are created in CoreDraw and imported to ArcMap (via "Insert-Object") as graphic elements that are linked files to CoreDraw. This ArcMap/CorelDraw linking allows edits to be done in CorelDraw and will automatically update the ArcMap element. Other graphic software can be used such as Macromedia Freehand and Illustrator. All the placements of cartographic elements and data in the surrounds and on the map area use the GSC's cartography specs. For A-series maps, these specifications are found at http://ess.nrcan.gc.ca/pubs/carto/specifications_e.php.

REFERENCES

Shaw, John and Todd B.J., 2003, Development of mapping standards for marine bathymetry, backscatter, geology and habitat maps: Geological Survey of Canada internal report, p. 1-22.

Fader, B.J.G., and Peters, J., 1988, Symbols for mapping of marine surficial geological features. Federal Geographic Data Committee (FGDC), digital cartographic standard for geologic map symbolization: Geological Survey of Canada internal report, p.1-16.

Shaw, J., and Courtney, R.C., 2004, A digital elevation model of Atlantic Canada: Geological Survey of Canada Open File Report 4634, Compact Disk.

Geological Data and Collections in Peril: Case Example in Georgia

By Mark D. Cocker

Georgia Geologic Survey (former)
Watershed Protection Branch (currently)
Georgia Environmental Protection Division
19 Martin Luther King, Jr., Dr., SW
Atlanta, Georgia 30338
Telephone: (404) 657-6131
e-mail: mark_cocker@mail.dnr.state.ga.us

INTRODUCTION

A report by the National Research Council Committee (NRC) on the Preservation of Geoscience Data and Collections investigated types of geoscience data and collections, their estimated volume, and factors that threaten loss or degradation of these data and collections (Musser, 2003). Types of data and collections included drill core, cuttings, thin sections, washed residues, well logs, fossils, minerals, rocks, surface geophysical surveys, scout tickets, and chemical analyses. The emphasis of the National Research Council Committee appeared to be on the vast amount of data collected by hydrocarbon and mineral exploration companies in the western U.S., as well as data collected and stored by relatively well-funded, mainly western, state geological surveys and the USGS. The Committee's recommendations focused on preservation of the data and collections and value-added functions such as documentation and outreach (Musser, 2003). Rocks, whether they are drill core, cuttings, fossils or hand specimens, represent the most voluminous (and heavy) portion of the data and collections. Rocks are, in some ways, the cheapest to preserve, and in other ways, probably the most costly to preserve and document.

A significant amount of other geoscience data that wasn't addressed in the NRC study includes field notes, maps, photographs, and publications, as well as data that are difficult to quantify (e.g. institutional memory). Some state geological surveys may depend more heavily on these types of data, which require smaller expenditures to acquire or maintain than, for example, drill core. State geological surveys are more likely to have more focused geological data and collections pertinent to their respective state, and should have the greatest interest in preservation and documentation of those data, as well as promotion of these through outreach programs. However, when a state geological survey suddenly ceases to exist, will there be any stewardship of the data and collections?

STATUS OF THE GEORGIA GEOLOGIC SURVEY

At the end of 2004, after serving the State of Georgia for 115 years, the Georgia Geologic Survey (GGS) was abruptly terminated. The State Geologist retired in August of 2005. A small handful of geologists continue to work on geologic problems and mapping within the Regulatory Support Program, Watershed Protection Branch, Environmental Protection Division, Georgia Department of Natural Resources. No organization or individual has been charged with ownership of the GGS's geoscience data or collections. The threat of non-stewardship and perhaps permanent loss of a significant amount of scientific and historical data and collections pertaining to the State of Georgia is real. Permanent loss of institutional memory is highly probable.

Although the collection of geoscience data had increased significantly in Georgia during the past 30 years, the GGS's management exhibited little interest and committed inadequate funding, time and personnel to the organization, documentation, preservation and storage of these geoscience data and collections. Prior to termination of the GGS, guidelines had not been established to organize, document, manage, preserve and store these data and collections, as well as new data and additions to the collections that continue to accumulate. Maintenance and updating of digital data and media have not been addressed. A few geology programs continue, e.g., the National Cooperative Geologic Mapping Program's STATEMAP component, and these accumulate new data, such as field observations, photographs, maps, and core logs, and new collections, such as drill core.

TYPES OF GEOLOGICAL DATA AND COLLECTIONS

GGS's geological data and collections are basically the same as most governmental geological surveys. These data and collections include written field observations in notebooks and maps, well logs, petrographic and XRD analyses, geochemical analyses, geophysical well logs, photographic records (film and digital), drill core, cuttings, washed residues, minerals, rocks and fossils. Mapping projects conducted under the STATEMAP program offer a good example of the breadth of these collections. These projects have allowed the field geologist to observe the geology and geological relations that describe outcrops, roadcuts, and mine exposures and to record these observations by a variety of methods. During the course of mapping 7.5' quadrangles, 300 to 500 sites per quadrangle may be examined and recorded. During the GGS mapping program, these sites are continuously added to a GIS outcrop database, which now contains more than 6000 sites. The database currently includes the quad number, the site number, an interpretation of the geological unit (i.e., the formation symbol), as well as a coded shade-set number for map plotting purposes. (Note: at the beginning of this GGS mapping program, publication of digital geological database examples was rather limited, and the year-to-year continuity of the program was not established, so only the minimum amount of geological data was entered into the database. Future tasks may include coding of site descriptions and perhaps linking of digital photos to each outcrop). With this database, outcrops can be quickly plotted on a topographic base map either as a hard copy or on-screen. Contacts for an interpretive geologic map are then digitized on-screen in relation to the outcrops. Outcrops are assigned a slightly darker color shade and can be plotted relative to the interpretive geology. Relative, size, shape and distribution of outcrops are also apparent on the outcrop coverage. The addition or linking of digital photographs can document lithology, sedimentary structures, alteration, mineralization, structural deformation, and geological hazards.

Another aspect of the STATEMAP program involves shallow core drilling. This drilling is invaluable in areas where outcrops are poor to non-existent. Because of equipment limitations, hole depth is limited to 50 feet. Sites are selected mainly for the opportunity of locating geologic contacts, and core is logged principally for lithology and contacts. In addition to the written core logs, digital photographs of the core are taken with a 2.5-foot scale marked in inches. These images are clipped, and the core is reconstructed into 10-foot lengths by digitally pasting the images end-to-end.

A part of the product produced by the GGS for the STATEMAP program consists of a geological report that includes descriptions of the formations, structure, mineralization, aquifers, and geologic hazards. Selected annotated photographs of outcrops document observations in the report. Descriptions of new core are also included. Because the file size of a digitally reconstructed core hole is on the order of several tens of megabytes, current hardware and software cannot handle these files in a text document, and so they are not included in the published STATEMAP product.

Older data, maps and manuscripts exist only in hard copy paper or mylar formats. More recently, data, maps and manuscripts were compiled or created digitally and stored on a variety of evolving formats and media. A significant amount of data exists only in the form of hard copy publications. Some of the newer publications are available only on CD ROMs. Without a management plan and support, will those publications on CDs be readable in 10 or 20 years?

The GGS lacked a management plan to develop consistent data-recording methodologies and store and preserve that data. Over a period of many years, diverse types of geologic data were collected and recorded by numerous staff members with different education and experience levels, employing a variety of evolving techniques, tools and media. In addition, geologists were not required to provide copies of the data to the technical files. This resulted in the actual physical loss of unpublished data or misplacement of data files when staff members left the GGS.

During the past 17 years, digital technology advanced slowly within the GGS. In 1988, one personal computer was available to a staff of approximately 40. Computers were gradually acquired over the next seven years so that eventually the entire staff had access to a personal computer. Data storage was problematic, with inadequate hard drives and a policy that limited the number of available diskettes. With no linkage to a common server, file sharing was difficult. Reusing diskettes was a management policy as diskettes were "expensive" and long term data storage was a foreign concept. Even as technology advanced and file sizes grew rapidly, only one CD writer was made available to the entire staff of the GGS. Files from a PC were transferred to a server and then to another PC where the CD writer was installed. This procedure remains as computers and other related hardware have not been updated since 1999. As with the hard copy data files, the GGS did not develop a strategy for planning how data were to be stored, backed-up, or archived. Software acquisition and software training were neglected by GGS management, with few staff members advancing beyond basic word processing and spreadsheet computer literacy.

Migration of data to newer formats is vital, as technology continues to advance and older technology is no longer supported. A change in software approximately 5 years ago resulted in many data files becoming inaccessible or corrupted. Over the course of 10 to 12 years, numerous document files were created with one particular

word processing software, with data files and accompanying graphs prepared with that company's spreadsheet software. Manual entering of large amounts of data into spreadsheets represented a considerable investment of time. A change in software vendors by the State of Georgia resulted in the removal of the previous software, installation of another company's software, and the resultant loss or corruption of a significant amount of digital data.

ARCHIVING DATA FILES

The archiving of data files and collections is a critical function of a geoscience organization. A collection of data files should be easily searchable and accessible. Most of the GGS's data files are referred to as the technical files; these are housed mainly in standard file drawers and flat files (for maps). Despite a recent, multi-year attempt to develop a digital catalog of the technical files, the most effective search technique remains the manual method. The present digital catalog database is an alphabetical file listing and is not searchable by key words, topics, authors, dates, or subject areas. This digital catalog was developed by people with no technical background, and no input from the geologic staff was considered. Recently, a compilation of drill hole data for a selected depth interval in a selected multi-county area required a month-long manual search of file drawers to find and retrieve logs from five different locations. The existence or location of some drill hole logs remains unknown.

PUBLICATIONS

GGS publications should also be regarded as data sources, as these publications contain data unavailable anywhere else. Ideally, data would be archived in a data repository. Depending on the author or the reviewers, some or all of the collected data may be included in the publication. Some publications, e.g. maps, may be compilations of new and older published and unpublished data from a variety of sources. As these publications are data sources, they should also be documented, preserved and made available for access by other geologists and customers. Other manuscripts and maps were at various stages of completion when the GGS was terminated and continue to be published. Still other geologic projects, including STATEMAP mapping, will continue to produce more data and publications.

Publication, sales, and preservation of the publication inventories, require an agency to commit funds, sales staff and space. Documentation is especially important to the customer, in order to search for, and find, what they need. Traditionally, a geological survey's publications are documented in a catalog. The GGS's annual catalog of publications consists of a simple sequential (mainly chronological) listing by type of publication, i.e. bulletin, open-file report, hydrologic atlas, etc., and by title, author and date. The publications are not arranged by subject matter, e.g. economic geology, or other logical method to quickly find a publication of interest. An annotated bibliography could provide more pertinent information regarding the publications. More recently the GGS produced an on-line catalog, but it is just a digital version of the hard-copy catalog without a key word search. Even the most recent catalog of GGS publications is far from complete, e.g., it does not indicate the existence of 26 geologic maps and 8 open-file reports completed during seven years of GGS participation in STATEMAP mapping. Customers searching the catalog would not be aware of these publications, and the staff servicing the customers would probably also not be aware of these publications. The GGS has not funded either new publications or reprints of older publications that have been sold out. As publications begin to be sold out, at what point will termination of the GGS affect accessibility and availability of their publications and data?

STORAGE OF COLLECTIONS

Drill core and cuttings, petrologic, mineral and fossil collections belonging to the GGS have been stored in a non-climate controlled warehouse in Atlanta, GA. Many core and cuttings boxes are up to 40 to 60 years old and have suffered the effects of high temperatures, high humidity, dust from nearby industrial activities, and neglect. Because of poor lighting, security issues, access, air-conditioning and heating, and the lack of other basic facilities, the warehouse has never served as a research facility. Project files (data), mylar originals of published maps (required for reprints), office files, rare and historic USGS Professional Papers and Bulletins, etc., bound professional periodicals, and excess older GGS publications also have been semi-permanently stored at the warehouse. Deterioration of materials and data over time has been inevitable.

Prior to the termination of the GGS, an unknown quantity of drill core and cuttings, rock, mineral and fossil collections, maps, project files, equipment, and GGS publications were discarded, as a result of a lack of interest and understanding by decision-makers regarding the present and future value of the data and collections.

INSTITUTIONAL MEMORY

The institutional memory of a geoscience organization consists of: the undocumented experiences, observations and interpretations that are accumulated by an organization's personnel mainly during the course of their field and laboratory work, conversations with colleagues both within and outside their organization, knowledge gained at professional meetings, and reading or knowledge of pertinent published literature and unpublished

or “gray” literature. Institutional memory also includes other types of information and knowledge such as road or property access, new roadcut or other excavation exposures outside of one’s current study area, and professional contacts outside the agency (e.g., consultants and industry geologists, who may have little or no publication record). A discussion of what constitutes institutional memory is open-ended, but essentially it is undocumented knowledge and expertise that, in order for the organization to survive and flourish, can (and must) be passed on to other personnel.

During the past 25 years at the GGS, an unknown and immeasurable amount of institutional memory was permanently lost, as experienced geological personnel were reassigned, retired, or moved on to new employment. Nearly all of the GGS geologists who were reassigned or acquired employment with other Georgia state agencies have retired or are within a few years of retirement. In this author’s experience, the institutional memory of former staff generally fades rapidly with time. Currently, the two remaining GGS geologists have about 10 years to retirement age with no new or potential opportunities for new geological staff to pass on this institutional memory.

RECOMMENDATIONS FOR PRESERVATION AND INCREASED AVAILABILITY OF GEOSCIENCE DATA AND COLLECTIONS IN GEORGIA

These recommendations may be specific to Georgia because of the current circumstances, but may serve as a guide if other state geological surveys risk termination:

- Restore a GGS that has a legislative mandate to collect, preserve, document, and disseminate geoscience data and collections.
- Adequately fund and staff a geological survey to be able to achieve that mandate.
- Construct a climate-controlled core warehouse facility with a permanent, geologically-trained staff.
- Provide training for new and current employees and encourage retention of employees or emeritus employees to preserve and transfer institutional memory.
- Scan older publications, maps, and data files and make them available in a digital format.
- Update older digital files to current standards and maintain newer digital files with newer digital formats and technological advances.
- Develop a digital on-line catalog of GGS publications, geoscience files, and archival maps, etc., such as that developed by the North Carolina Geological Survey (available at http://www.geology.enr.state.nc.us/).

REFERENCES

Musser, L.R., 2003, Preservation of Geoscience Data and collections, in Soller, D.R., ed., Digital Mapping Techniques ’03—Workshop Proceedings: U.S. Geological Survey Open File Report 03-471, p. 195-196, available at http://pubs.usgs.gov/of/2003/of03-471/musser/index.html.

National Research Council, Committee on the Preservation of Geoscience Data and Collections, 2002, Geoscience Data and Collections—National Resources in Peril: Washington, D.C., National Academy Press, 107 p., available at http://www.nap.edu/catalog/10348.html.

Creation of a New Statewide Bedrock Geologic Map for Missouri

By Tom Hertel, Edith Starbuck, Patrick Mulvany, Mark Middendorf, Hal Baker (retired), David Erickson, and Karen Loveland

Missouri Department of Natural Resources
Geological Survey and Resource Assessment Division
P.O. Box 250
Rolla, MO 65402
Telephone: (573) 368-2100
FAX: (573) 368-2111
e-mail: edie.starbuck@dnr.mo.gov

INTRODUCTION

The previous edition of the 1:500,000-scale *Geologic Map of Missouri* was published in 1979. Since that time, the bedrock geology for approximately 40% of the state has been revised, updated, or mapped in greater detail. Beginning in 1997, ArcView (versions 3 through 3.2) has been used to compose 1:24,000-scale geologic maps. Some 1:100,000-scale maps had been compiled from these, but no statewide digital compilation of existing geologic maps had been attempted. The occasion of the 150th anniversary of the creation of the first Missouri Geological Survey seemed an appropriate time to make a new edition of the 1:500,000-scale map of the bedrock geology of Missouri, and to create an easily updateable digital version.

Development of the Missouri Environmental Geology Atlas (MEGA) was also proposed to be completed for the sesquicentennial. MEGA is a collection of statewide Geographic Information System (GIS) data layers that were produced as a reference resource to aid in environmental evaluations. The shapefile for the 2003 edition of the *Geologic Map of Missouri* was incorporated into the MEGA project.

ASSEMBLING THE DATA

Sixty-five 1:24,000-scale bedrock geologic maps were already available in a digital format in the form of shapefiles. The 1979 version of the *Geologic Map of Missouri* was also available in a digital format. Approximately 22,000 square miles (400 7.5' areas) of geologic contacts needed to be digitized in order to incorporate all new mapping into the sesquicentennial map. These maps were at various scales ranging from 1:24,000 to 1:250,000. Several university students were hired to scan and georeference paper maps, perform heads-up digitizing with ArcView 3.2, and attribute the maps. Old geologic maps that had been hand-drawn on paper or mylar were scanned at 200-300 dots per inch using a 36-inch scanner. Maps were georeferenced, generally with respect to topographic quadrangle boundaries, using ESRI Image Analysis software. In order to be compatible with available USGS digital raster graphics (DRGs) of topographic maps, the location data were stored with the NAD 27 datum.

Most of the existing 1:24,000-scale digital maps had been created solely to place in an ArcView layout to produce a hard copy map of an individual 7.5'quadrangle, with no intent to provide a digital product. Many individual maps were composed of multiple ArcView themes that had been created with one theme per bedrock unit in order to ease editing. These themes had to be unioned and properly attributed.

COMPILATION

The statewide geologic map was compiled using ArcView 3.2. Compilation involved deleting some areas from the 1979 statewide map and then appending the new mapping into the shapefile. Overlapping polygons with the same attributes were unioned together. Mismatches along quadrangle boundaries and abrupt changes in detail were obvious at this stage. Also, irregularities along some of the quadrangle boundaries, such as gaps or holes, and long, thin polygons (which we referred to as slivers) caused resolution problems. Many of these were removed using the Dissolve Adjacent Polygons tool from Jenness Enterprises (http://www.jennessent.com/index.html). Some of the largest of the slivers were found by calculating polygon areas to find small polygons, and having a geologist decide which polygons were slivers and which were legitimate polygons. Other spatial problems were corrected using tools that are within ArcView's Compiled Theme Tools. These include:

Shape Tools
- Detect Intersecting Polygons
- Dissect Intersecting Polygons
- Detect Hole

Miscellaneous Tools
- Simplify Shape by Removing Vertices

At this point, the map needed to be edited by staff geologists who corrected boundary mismatches, and inaccurate attributes, and smoothed the transitions between areas mapped at varying scales. However, the data layer was too big to be edited on any of the geologists' computers. The size of the shapefile needed to be reduced. This was accomplished in four ways.

1. The number of vertices was reduced using a vertex weeding routine.
2. Large, complicated polygons were split into 2 or 3 pieces
3. The theme was "cleaned" in both ArcView 3.2 and ArcInfo 8. (ArcView 3.2 "CLEANS" were made for speed and convienence, to gain a better broad understanding of the effect of splitting on the process. ARCINFO 8 allowed more precise control of "CLEAN" parameters.)
4. The state map was split into 3 pieces.

The fourth step also allowed more than one geologist to edit the map at the same time. The three pieces of the map were later put back together using the Geoprocessing Wizard extension. The final editing process determined the necessary map units and any exaggeration of features for viewing at 1:500,000-scale.

CREATION OF THE CROSS SECTION

ArcView 3D Analyst was used to create a topographic profile from a statewide compilation of USGS DEMs. The location of the cross section was selected to pass through an interesting and representative section of the geology of the state emphasizing some of the major structural features. The locations of deep, logged wells were also considered when selecting the cross section location. These wells would provide the formation tops for the subsurface part of the cross section. The cross section was split into two parts in order for it to fit on the map layout. The state capitol, Jefferson City, was a convenient and significant location to divide the cross section.

The line of cross section was placed on the geologic map in ArcView 3.2. A theme was created that reflected the surface geology along the line of cross section by clipping a narrow polygon strip from the bedrock geology shapefile along that line. This clipped theme retained the formation attributes.

The profile was then scaled to the desired vertical exaggeration and placed on a scaled grid. The strip of surface geology was repositioned and projected onto the profile and the geology was transposed to the profile. Formation tops were added to the wells by posting attributes from a modified well log database. The geologist then drew the cross section as a polygon theme using these data points.

THE LAYOUT

The layout was created in ArcView 8 however; to facilitate labeling, point themes with the label attributes were made in ArcView 3.2. These "labeling themes" were created for roads and rivers, as well as for the bedrock. For these themes, reference points were created for each of the lines and polygons that were to be labeled. These were selected based on size and density of features. These themes were brought into ArcView 8 and the reference points were labeled. In this way the position of labels could be more easily controlled while still allowing the use of autolabeling.

Text was created using Microsoft Word software and was copied and pasted into text boxes on the layout. Multiple "data views" were created for the primary map, the cross section, and the inset maps. Images of the logos for the Department of Natural Resources and the Geological Survey Sesquicentennial were inserted as pictures.

Problems encountered during the layout phase included proximity of objects on the layout. When a graphic or text item was placed too close to the edge of the layout or too close to another item, these graphic and text items did not legibly print. Color selection was a more time consuming process than expected since colors that contrasted well on the computer screen did not always contrast well in print. Many variations in color combinations were selected and printed before the final selection was made.

To allow for easier reference to more detailed mapping, 7.5' quadrangle outlines and names were printed on the map. To improve readability, the township and range lines, present on the 1979 version, were not included. The earthquake epicenters in southeast Missouri from 1973 to 1999, and a map of major tectonic features were added as insets.

Updates and changes to the map explanation column were extensive. For simplicity and accuracy it was decided not to include a graphic representation of the lithology on the geologic column, as had been done on the 1979 version of the map. Series names, particularly in the Carboniferous had changed considerably since 1979.

REVIEW AND RESULTS

The map was printed and reviewed by the geologists involved in producing the map. The map was also posted so geologists throughout the survey could comment on

the map. Several errors were found and corrected. Final changes to the colors were made at this time.

The layout was printed to a postscript file from ArcMap 8 and sent to Adobe Distiller 5.0 with custom settings for the HP3500 printer. Problems encountered at this stage included obtaining the desired fonts and paper sizes. Correcting for this was primarily a matter of trial and error. This required all the fonts in ArcMap to be downloadable. A major obstacle was that the postscript driver did not recognize the landscape format and ArcMap will not create a postscript file if the Windows driver is used. When landscape is attempted, the map will rotate but the right side will be cut off. To work around this situation, the map was given a false paper size, but this required extensive trial and error. After a large amount of experimentation with settings and numerous test prints, the file was sent to the printer for production.

The shapefile that had been created for the 2003 *Geologic Map of Missouri* is included in the Missouri Environmental Geology Atlas, or MEGA. MEGA is a collection of statewide Geographic Information System (GIS) data layers. The variation in scale of the various source maps for the statewide map is noted in the metadata. The bedrock geology shapefile has been updated subsequent to the publication of the 2003 *Geologic Map of Missouri*. In the year since its creation, approximately 200 MEGA CDs and 300 paper geologic maps have been distributed. This is the approximate number that was anticipated.

The National Park Service Geology-GIS Geodatabase Data Model: A Story of Migration

By Stephanie A. O'Meara[1], Heather S. Stanton[1], James R. Chappell[1], Greg Mack[2], Anne R. Poole[3] and Georgia Hybels[4]

[1]Colorado State University/National Park Service Cooperator
1201 Oak Ridge Drive, Suite 200
Fort Collins, CO 80525
Telephone: (970) 225-3584
e-mail: Stephanie_O'Meara@partner.nps.gov; Heather_Stanton@partner.nps.gov; Jim_Chappell@partner.nps.gov

[2]National Park Service – Pacific West Region, Seattle Office
e-mail: Greg_Mack@nps.gov

[3]United States Forest Service – Chippewa National Forest
e-mail: apoole@fs.fed.us

[4]University of Denver/National Park Service Cooperator
e-mail: Georgia_Hybels@partner.nps.gov

INTRODUCTION

Geologic maps are an integral component of the physical science inventories stipulated by the National Park Service (NPS) in its Natural Resources Inventory and Monitoring (I&M) Guideline. The NPS has identified Geographic Information Systems (GIS) and digital cartographic products as fundamental resource management tools. There are few geologists employed at parks, thus these tools are particularly important to the NPS to aid resource managers in using geologic data for park management decisions (O'Meara et. al., 2003).

The NPS Geologic Resources Division (GRD) is currently developing a Geologic Resources Evaluation (GRE) that includes a geologic bibliography, a summary report of each park's geologic resources, and the development of a geology-GIS data model for implementation in the production of digital geologic-GIS data for each park. Colorado State University is a partner in the development and production of these products.

The present NPS Geology-GIS Data Model (O'Meara et. al., 2005a) for park digital geologic-GIS data is based upon Environment Systems Research Institute (ESRI) ArcInfo coverage and shapefile vector file formats and provides a robust method for storing geologic data. Recently, ESRI developed the geodatabase format for storing spatial data within a relational database management system (RDBMS). A geodatabase stores data in point, line or polygon data layers (feature classes) that can be grouped into a feature dataset, a logical groupings of vector feature classes that share the same spatial extent. The geodatabase has the added strength of allowing attribute validation rules, relationship classes, and topological rules that maintain data integrity within and between feature classes.

In order to take advantage of this new, enhanced functionality, migration from the coverage/shapefile-based geology-GIS data model to a geodatabase geology-GIS data model is underway. Using geologic data from Glacier National Park, Montana (GLAC), our poster presented at this meeting outlines 1) the present NPS Geology-GIS coverage/shapefile data model, and its benefits and drawbacks, 2) the ESRI geodatabase architecture and key components, and 3) the implementation of the NPS Geology-GIS Data Model within the geodatabase architecture, and its benefits and drawbacks. The Glacier National Park, Montana (GLAC) digital geologic-GIS map was produced from existing published USGS paper maps (Whipple, 1992 and Carrara, 1990). Of note, Figures 1 through 6 and 10 in this paper were produced using Microsoft Office Visio objects created from Geodatabase Diagrammer, an ArcScript created by ESRI.

THE NPS GEOLOGY-GIS COVERAGE/ SHAPEFILE DATA MODEL

The present data model defines how geologic features are captured, grouped and attributed in coverage and shapefile formats. In addition, relationships with ancillary source map and geologic unit information tables are also established in the data model. Acceptable data model data layer attribution values, or domain lists, are stored in separate table files.

Features Supported by NPS Coverage/ Shapefile Data Model

In the existing coverage/shapefile data model (O’Meara et. al., 2005a), geologic features are currently divided into 30 coverages and 38 shapefiles. The discordance in the corresponding number of files between coverages and shapefiles is the result of the capability of the coverage format to store multiple geometries within one coverage file. Shapefiles do not support multiple geometries and therefore multiple files are needed to represent some features. Table 1 lists coverage/shapefile data model data layers by geologic feature and spatial type (i.e. polygon, line and point).

Coverage/Shapefile Feature Attribute Tables

Coverage and shapefile feature attribute tables consist of descriptive attribute fields that contain information about geologic features in the data model. Attribute field parameters include field name, data type, field definition, and field width parameters. Figure 1 presents the attribute table for the Geologic Units (GLG) coverage/shapefile data layer.

Coverage/Shapefile Data Model Benefits

Several benefits exist for a coverage/shapefile-based data model. These include:

- Features are stored in discrete data files (E00 and SHP) with inherent topology.
- Relationships between data layers and/or tables can be established using joins and relates.
- Works well with ArcInfo, ArcView and related modules.

Coverage/Shapefile Data Model Drawbacks

Several drawbacks also exist for the coverage/ shapefile-based data model. These include:

- No efficient way to implement within a Relational Database Management System (RDBMS).
- Coverage format easy to corrupt as a coverage spatial component and attribute table are stored in separate folders. This also limits the portability of a coverage as export files are required to efficiently transfer a coverage.
- Relationships between data layers and/or related ancillary tables not stored with data files. These must be re-established for each project/map document.
- Difficult to maintain topological relationships across multiple data layers.
- Large datasets are difficult to manage.
- File size limitations.

THE NPS GEOLOGY-GIS GEODATABASE DATA MODEL

A geodatabase stores spatial and non-spatial data, including attributes, in tables, feature classes, and feature datasets. In addition, the geodatabase stores attribute validation rules, relationship classes, and topological rules for ensuring data integrity.

Two types of geodatabases exist: personal and multi-user (enterprise). Personal geodatabase support is implemented in ArcGIS using the Microsoft Jet Database Engine and is suitable for project-level GIS. Enterprise databases are deployed using ArcSDE and require a DBMS such as IBM DB2, Informix, Oracle, or Microsoft SQL Server. The NPS geodatabase data model presented here was constructed in a personal geodatabase.

Similar to the coverage/shapefile data model, the proposed NPS Geology-GIS Geodatabase Data Model (O’Meara et. al., 2005b) includes a list of feature classes, and feature attribute tables. More importantly the geodatabase data model also includes: 1) a geodatabase relational schema, 2) domains, 3) subtypes, 4) topological rules, and 5) relationship classes.

Feature Class List

Geologic features are currently divided into 44 geodatabase feature classes. Table 2 lists these feature classes by geologic feature type and spatial type (i.e. polygon, line and point).

Feature Datasets

In the NPS Geology-GIS Geodatabase Data Model, feature classes representing each of the geologic data layers on a single map are grouped into a feature dataset. All feature classes that participate in the feature dataset share the same spatial reference (i.e., projection and datum). Feature datasets store spatial data and relationships, but do not store tables. Tables are stored within the geodatabase, but outside the feature dataset. Feature datasets are

Table 1. List of coverage/shapefile data model data layers by geologic feature and spatial type (i.e., polygon, line, and point). Each data layer is assigned a three- or four-letter abbreviation (e.g., GLG for Geologic Units, GLGA for Geologic Contacts) that is indicated in the table. Data layers highlighted in gray are present in the Glacier National Park (GLAC) digital geologic-GIS map.

Geologic Features	Coverage Abbreviation (Spatial Type)	Shapefile Abbreviation (Spatial Type)
Geologic Units and Contacts	GLG (polygon and line)	GLG (polygon) and GLGA (line)
Linear Geologic Units	GLN (line)	GLN (line)
Point Geologic Units	GPT (point)	GPT (point)
Faults	FLT (line)	FLT (line)
Folds	FLD (line)	FLD (line)
Attitude Observation Localities	ATD (point)	ATD (point)
Age-Date Localities	DAT (point)	DAT (point)
Volcanic Line Features	VLN (line)	VLN (line)
Volcanic Point Features	VPT (point)	VPT (point)
Linear Dike Units	DKE (line)	DKE (line)
Area Dike Swarms and Contacts	DKS (polygon and line)	DKS (polygon) and DKSA (line)
Mine Point Features	MIN (point)	MIN (point)
Cross Section Lines	SEC (line)	SEC (line)
Area Volcanic Units and Contacts	ASH (polygon and line)	ASH (polygon) and ASHA (line)
Metamorphic/Alteration Boundaries	MET (line)	MET
Linear Glacial Features	MOR (line)	MOR (line)
Structure Contour Lines and Other Lines	LN# (line)	LN# (line)
Joints	JLN (line)	JLN (line)
Sensitive Geologic Point Features	SPF (point)	SPF (point)
Unique Geologic Point Features	UPF (point)	UPF (point)
Surficial Units and Contacts	SUR (polygon and line)	SUR (polygon) and SURA (line)
Measured Unit Thickness Localities	MUT (point)	MUT (point)
Mine Area Features	MAF (polygon and line)	MAF (polygon) and MAFA (line)
Seismic Data Localities	SMC (point)	SMC (point)
Sample Localities	SAM (point)	SAM (point)
Area Deformation Zones	DEF (polygon and line)	DEF (polygon) and DEFA (line)
Geologic Hazard Line Features	HZL (line)	HZL (line)
Geologic Hazard Point Features	HZP (point)	HZP (point)
Glacial Area Features	AGF (polygon and line)	AGF (polygon) and AGFA (line)
Geologic Hazard Area Features	HZA (polygon and line)	HZA (polygon) and HZAA (line)

necessary to the creation of topological rules, which will be discussed later. Figure 2 presents the geodatabase data model schematic for the Digital Geologic Map of Glacier National Park, Montana (GLAC), which includes the map's feature dataset and related ancillary tables.

Feature Attribute Tables

Feature attribute tables in the geodatabase consist of descriptive attribute fields that contain information about the features within a feature class. Attribute field parameters include field name, data type, whether or not to allow null values, field definition, domain type (coded value or range), and field width (precision, scale and length). Figure 3 presents the attribute table for the Geologic Units (GLG) feature class.

Attribute Domains

Attribute domains define acceptable values for fields in attribute tables that are contained in the geodatabase. Coded value domains are used in the NPS Geology-GIS Geodatabase Data Model to define acceptable values for various feature class type fields and their subtypes (see Subtypes section) including: Attitude Type (subtypes – planar measurements, linear measurements, etc.), geologic contact type, fault type, and fold type. Range domains are used to define ranges of acceptable values for attribute fields such as

Multiple-Geometry feature class

Geologic Units (glacglg)

Field Name	Data Type	Field Definition	Input Width	Output Width	Dec. Places
AREA	Binary	area of feature	8	18	5
PERIMETER	Binary	perimeter of feature	8	18	5
GLACGLG#	Binary	unique internal feature ID #	4	5	-
GLACGLG-ID	Binary	unique internal ID#	4	5	-
GLG_IDX	Integer	unqiue ID#	6	6	-
GLG_SYM	Character	unit age-lithology symbol	12	12	-
USGS_SYM	Character	source unit age-lithology symbol	12	12	-
GLG_AGE_NO	Number	unit age-sort number (young to old)	8	8	4
GMAP_ID	Integer	source map ID # (in MAP table)	6	6	-
HELP_ID	Character	unit (page) variable for Help File	12	12	-

Figure 1. Geologic Units (GLG) coverage/shapefile data layer attribute table. Attribute table fields highlighted in white store information about geologic unit polygons (areas) such as: geologic feature identification number (GLG_IDX); geologic unit symbol (GLG_SYM); source map geologic unit symbol (USGS_SYM); an age number sorting units from youngest to oldest (GLG_AGE_NO); a source map ID number (GMAP_ID); and a variable used to 'link' a geologic unit to a map help file containing geologic unit descriptions (HELP_ID). Standard ArcInfo polygon attribute fields are highlighted in medium gray. These are created automatically, and include area (AREA), perimeter (PERIMETER), a unique internal feature ID (GLACGLG#), and a unique internal ID (GLACGLG-ID).

Table 2. List of geodatabase feature classes by geologic feature and spatial type (i.e., polygon, line, and point). Each feature class is assigned a three- or four-letter abbreviation (e.g., FLT for faults, GLN for Linear Geologic Units) that is indicated in the table. Feature classes highlighted in medium gray are present in the Glacier National Park (GLAC) digital geologic-GIS map.

Geologic Features	Feature Class Abbreviation	Spatial Type
Geologic Units	GLG	polygon
Geologic Contacts	GLGA	line
Linear Geologic Units	GLN	line
Point Geologic Units	GPT	point
Surficial Units	SUR	polygon
Surficial Contacts	SURA	line
Volcanic Ash Units	ASH	polygon
Volcanic Ash Contacts	ASHA	line
Linear Dike Units	DKE	line
Dike Swarm Units	DKS	polygon
Dike Swarm Contacts	DKSA	line
Deformation Zones	DEF	polygon
Deformation Zone Boundaries	DEFA	line
Faults	FLT	line
Folds	FLD	line
Linear Joints	JLN	line
Attitude Observation Localities	ATD	point

Geologic Sample Localities	GSL	point
Cross Section Lines	SEC	line
Structure Contour Lines and Other Value Lines	CN#	line
Observation, and Observed Extent and Trend Lines	LIN	line
Volcanic Linear Features	VLF	line
Volcanic Point Features	VPF	point
Geologic Linear Features	GLF	line
Geologic Point Features	GPF	point
Glacial Area Features	GAF	polygon
Glacial Area Feature Boundaries	GAFA	line
Glacial Linear Features	GFL	line
Glacial Point Features	GFP	point
Mine Area Features	MAF	polygon
Mine Area Feature Boundaries	MAFA	line
Mine Linear Features	MLF	line
Mine Point Features	MIN	point
Geologic Hazard Area Features	HZA	polygon
Geologic Hazard Area Feature Boundaries	HZAA	line
Geologic Hazard Linear Features	HZL	line
Geologic Hazard Point Features	HZP	point
Alteration and Metamorphic Areas	AMA	polygon
Alteration and Metamorphic Area Boundaries	AMAA	line
Alteration and Metamorphic Linear Features	AML	line
Alteration and Metamorphic Point Features	AMP	point
Geologic Measurements Localities	GML	point
Seismic Localities	SMC	point
Geologic Observation Localities	GOL	point
Map Symbology	SYM	point

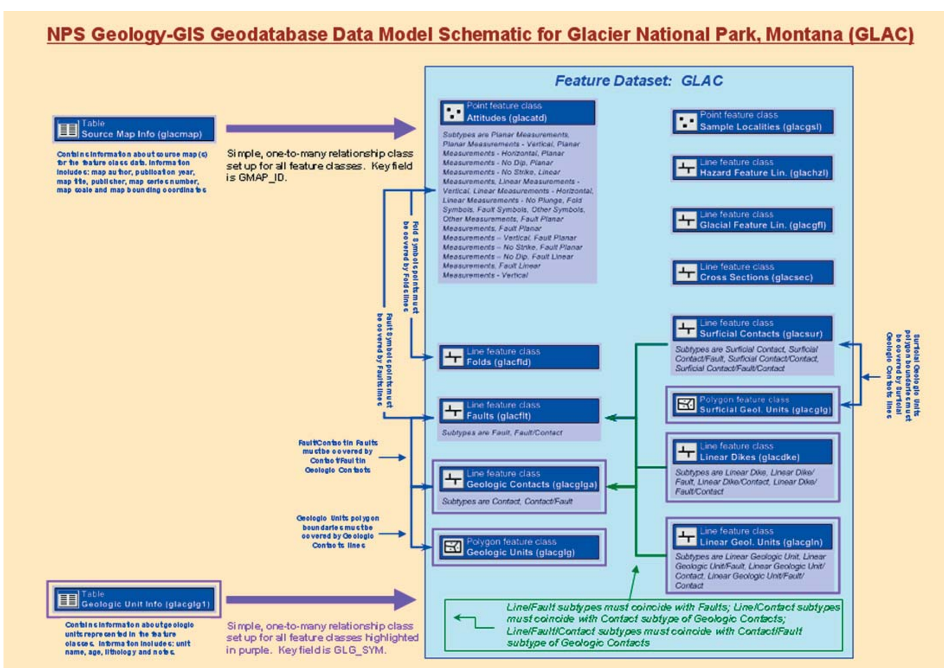


Figure 2. Geodatabase data model schematic for the Digital Geologic Map of Glacier National Park, Montana (GLAC) showing map feature dataset, feature classes, ancillary map tables, and implemented relationship classes.

Simple feature class

Geologic Units (glacglg)

Geometry *Polygon*
Contains M values *No*
Contains Z values *No*

Field Name	Data Type	Allow nulls	Field Definition	Domain	Pre	Scale	Length
OBJECTID	Object ID	-	feature object ID	-	-	-	-
SHAPE	Geometry	Yes	feature geometry	-	-	-	-
GLG_ID	Long integer	No	feature ID #	-	0	-	-
GLG_SYM	String	No	unit age-lithology symbol	-	-	-	12
SRC_SYM	String	No	source unit age-lithology symbol	-	-	-	12
NOTE	String	No	text notes and remarks	-	-	-	254
GMAP_ID	Long integer	No	source map ID # (in MAP table)	-	0	-	-
HELP_ID	String	No	unit (page) variable for Help File	-	-	-	12
SHAPE_Length	Double	Yes	feature perimeter	-	0	0	-
SHAPE_Area	Double	Yes	feature area	-	0	0	-

Figure 3. Geologic Units (GLG) feature class attribute table. Attribute table fields highlighted in white store information about geologic unit polygons (areas) such as: geologic unit symbol (GLG_SYM); source map geologic unit symbol (SRC_SYM); a notes and remarks field (NOTE); a source map ID number (GMAP_ID); and a variable (HELP_ID) used to 'link' a geologic unit to a map help file containing geologic unit descriptions. Attribute fields highlighted in medium gray are created automatically in a geodatabase, and include a unique feature ID (OBJECTID), as well as geometry type (SHAPE), length (SHAPE_Length), and area (SHAPE_Area).

strike, dip, and rotation in the Attitude Observation Localities (ATD) feature class. By placing limits and definitions on acceptable values, domains help to ensure consistency when attributing the features. Figure 4 presents the coded value domain used to define acceptable values for horizontal planar measurements (planar horizontal values subtype) in the Attitude Observation Localities (ATD) feature class. Figure 5 presents a range domain defining acceptable values for attitude strike/trend values for numerous subtypes in the Attitude Observation Localities (ATD) feature class.

Subtypes

Geodatabase subtypes are used to subdivide data in feature classes into groups that share the same attribute or topological validation rules and/or default values. Subtypes are defined by integer values stored in a field in a feature class's attribute table.

The purpose of the contact and contact/fault subtypes in the Geologic Contacts (GLGA) feature class (Figure 6) is to enforce topological rules and attribute validation. For example, the "fault/contact" subtype in the Faults (FLT) feature class must spatially coincide with the "contact/fault" subtype in the Geologic Contacts (GLGA) feature class (see Topology section). Subtypes in the Geologic Contacts (GLGA) feature class share the same domain for positional accuracy (CNT_POS).

Nineteen subtypes were created for topological rule enforcement and to control attribution for attitude type, attitude strike/trend, and attitude dip/plunge in the Attitude Observation Locality (ATD) feature class. For instance, the planar measurements subtype consists of planar measurements that have an azimuth (strike) that as a range from 0 to 359 degrees and are inclined at an angle from 0 to 89 degrees. Both measurement values, strike and dip, are restricted by range domains. Another subtype, planar vertical measurements, shares the same strike domain, from 0 to 359 degrees, however, the dip value is restricted by a coded domain to 90 degrees. Figure 7 presents an example of the Attitude Observation Localities (ATD) feature class attribute table.

Topology

In a geodatabase, topological rules govern spatial relationships within and between different feature classes. In the NPS Geology-GIS Geodatabase Data Model, topological rules are used to ensure that: faults exactly coincide with geologic contacts where a fault is also a contact; geologic contacts coincide with the boundaries of geologic units; and fold symbols and fault symbols lie along folds and faults, respectively. Topological rules also stipulate that no gaps or overlaps and no self-intersections do not occur in the various polygon and line feature classes in the geodatabase. Figure 8 illustrates an example of a topology rule error caused by incorrect attribution. Figure 9 presents topology rules for the Glacier National Park, Montana digital geologic-GIS map.

Coded value domain
Attitudes–Planar Horizontal Values
Description
Field type *Short integer*
Split policy *Default value*
Merge policy *Default value*

Code	Description
4	horizontal beds
25	horizontal foliation
112	horizontal schistosity
125	horizontal cleavage
133	horizontal inclusion

Figure 4. Coded value domain used to define acceptable values for the planar horizontal measurements subtype in the Attitude Observation Localities (ATD) feature class. All acceptable coded domain values and their description are listed for the subtype.

Range domain
Attitudes–Strike Values
Description
Field type *Short integer*
Split policy *Default value*
Merge policy *Default value*

Minimum value	Maximum value
0	359

Figure 5. Range domain defining acceptable values for attitude strike for numerous subtypes in the Attitude Observation Localities (ATD) feature class. Acceptable strike/ trend values or azimuths are between 0 and 359 degrees, and are defined in a range domain by a minimum value (0) and maximum value (359).

Simple feature class
Geologic Contacts (glacglga)
Geometry *Polyline*
Contains M values *No*
Contains Z values *No*

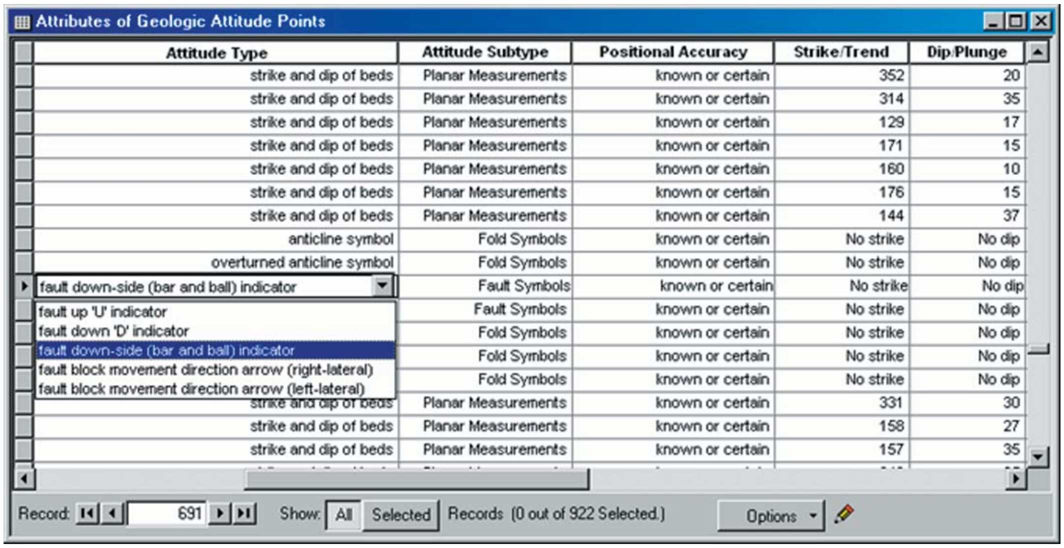
Field Name	Data Type	Allow nulls	Field Definition	Domain	Pre	Scale	Length
OBJECTID	Object ID	-	feature object ID	-	-	-	-
SHAPE	Geometry	Yes	feature geometry	-	-	-	-
GLGA_ID	Long integer	No	feature ID #	-	0	-	-
GLGA_SUB	Short integer	No	contact/faulted contact	Coded	0	-	-
POS	Short integer	No	contact position/concealment	Coded	0	-	-
NOTES	String	No	text notes and remarks	-	-	-	254
GMAP_ID	Long integer	No	source map ID # (in MAP table)	-	0	-	-
SHAPE_Length	Double	Yes	feature length	-	0	0	-

Subtypes of Geologic Contacts (glacglga)

List of defined subtype definitions and domains in this class

Subtype Code	Subtype Description	Field name	Subtype Definition	Domains
0	Contact	GLGA_SUB	geologic contact	CNT_POS
1	Faulted Contact	GLGA_SUB	faulted geologic contact	CNT_POS

Figure 6. Geologic Contacts (GLGA) feature class subtypes. Note that the subtype field (GLGA_SUB) is defined as a short integer field. In the attribute table, the subtype description text, in this case "contact" and "faulted contact", appears in the actual attribute table, and not the actual coded value, 0 for contact and 1 for faulted contact. Subtype descriptions appear by default when viewing attribute data and aid in attribution during object creation.



Attitude Type	Attitude Subtype	Positional Accuracy	Strike/Trend	Dip/Plunge
strike and dip of beds	Planar Measurements	known or certain	352	20
strike and dip of beds	Planar Measurements	known or certain	314	35
strike and dip of beds	Planar Measurements	known or certain	129	17
strike and dip of beds	Planar Measurements	known or certain	171	15
strike and dip of beds	Planar Measurements	known or certain	160	10
strike and dip of beds	Planar Measurements	known or certain	176	15
strike and dip of beds	Planar Measurements	known or certain	144	37
anticline symbol	Fold Symbols	known or certain	No strike	No dip
overturned anticline symbol	Fold Symbols	known or certain	No strike	No dip
fault down-side (bar and ball) indicator	Fault Symbols	known or certain	No strike	No dip
	Fault Symbols	known or certain	No strike	No dip
	Fold Symbols	known or certain	No strike	No dip
	Fold Symbols	known or certain	No strike	No dip
	Fold Symbols	known or certain	No strike	No dip
strike and dip of beds	Planar Measurements	known or certain	331	30
strike and dip of beds	Planar Measurements	known or certain	158	27
strike and dip of beds	Planar Measurements	known or certain	157	35

Figure 7. Attitude Observation Localities (ATD) feature class attribute table showing attribute fields with coded domains for attitude observation type, attitude subtype, and positional accuracy fields, and ranges domains for strike/trend and dip/plunge fields. The pull-down list displays attitude type values (ATD_TYPE field) dependent on the subtype field (ATD_SUB), in this case attitude type values that are fault symbols.

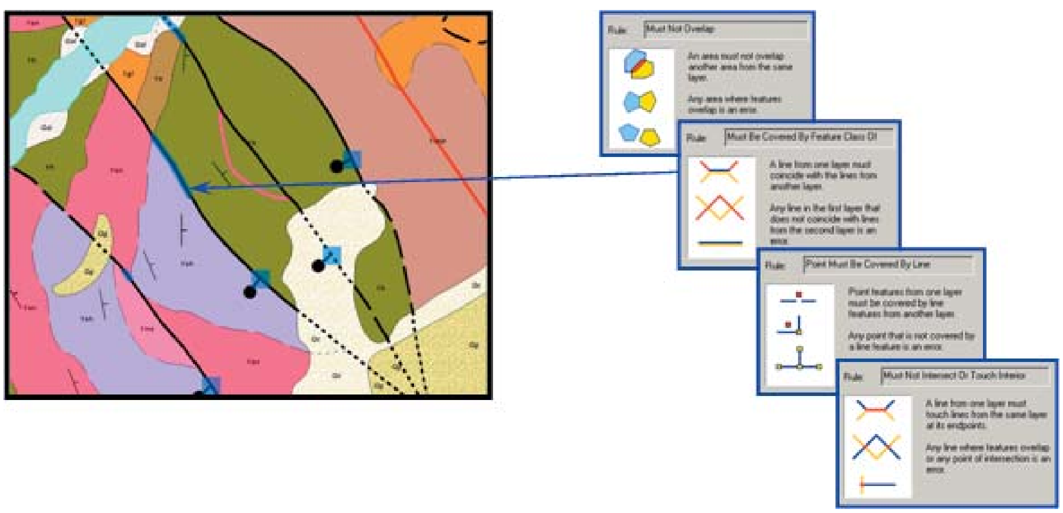


Figure 8. Snapshot of the Glacier National Park, Montana (GLAC) digital geologic-GIS map in ArcMap showing, in gray, a line segment where the rule "Must Be Covered By Feature Class Of" has been broken. Here, a fault/contact" in the Faults (FLT) is covered by a "contact" in the Geologic Contacts (GLGA). To correct the topological error, the contact should be attributed in the GLGA feature class as a "contact/fault".

Rules between feature classes that are specific to the geodatabase

Rules that apply only within a single feature class; mimic coverage rules

Rules between feature classes that mimic coverage rules

	GLGA	FLT	SURA	DKE, GLN	FLD, GFL, HZL, SEC	GLG	SUR
GLGA	H	1-A-1, 0-B-0	-	-	-	C	-
FLT	1-A-1	G	-	-	-	-	-
SURA	0-B, 2-A-0, 3-A-1	0-B, 1-A-0	H	-	-	-	C
DKE, GLN	0-B, 2-A-0, 3-A-1	0-B, 1-A-0	-	G	-	-	-
FLD, GFL, HZL, SEC	-	-	-	-	G	-	-
GLG	D	-	-	-	-	F	-
SUR	-	-	D	-	-	-	F

A Must Be Covered By Feature Class Of
B Must Not Overlap With
C Must Be Covered By Boundary Of
D Boundary Must Be Covered By
E General Polygon Rules: Must Not Overlap
F General Polygon Rules 2: Must Not Overlap, Must Not Have Gaps
G General Line Rules: Must Not Intersect Or Touch Interior, Must Be Single Part, Must Not Self-Intersect
H General Line Rules 2: General Line Rules Plus Must Not Have Dangles

Figure 9. Topological rules for the Glacier National Park, Montana (GLAC) digital geologic-GIS map are represented with the source or origin feature classes in rows and the destination feature classes in columns. Letters indicate the type of topology rule, as presented right of the principal table. A number before a letter (a topology rule) indicates the subtype in the origin feature class to which the topology rule is applicable. A number following a number (origin feature class) and letter (a topology rule) indicates the subtype of the destination feature class to which the topology rule is applicable. For instance, the Faults (FLT) feature class (source) is related to the Geologic Contacts (GLGA) feature class (destination) using the rule 1-A-1, where the first 1 is the "fault/contact" subtype (see line subtypes below figure) of the Faults (FLT) feature class, the A represents the topological rule (see list right of table), and the second 1 is the "contact/fault" subtype of the Geologic Contacts (GLGA) feature class. In other words, a "fault/contact" in FLT must coincide with a "contact/fault" in GLGA. If there is no number to represent a subtype for the source and/or destination feature class, the rule applies to the entire source and/or destination feature class. Note that one or more rules may apply to a given source and destination feature class.

Relationship Classes

Relationship classes store information about how geodatabase objects such as tables and feature classes are interrelated. In the NPS Geology-GIS Geodatabase Data Model, they are used to relate the table of Geologic Unit Information (GLG1) to the Geologic Units (GLG) feature class, as well as to other feature classes containing geologic unit information. They are also used to relate the Source Map Information (MAP) table to all of the feature classes in the geodatabase. See Figure 2 for an example of implemented relationship classes for the Glacier National Park, Montana (GLAC) digital geologic-GIS map.

Relationships are implemented with simple, one-to-many relationship classes, where records in the tables exist independently from feature class objects, as opposed to composite relationship classes where the records in the source table control the deletion/addition of destination feature class objects. For instance, if a composite relationship existed between the Geologic Unit Information (GLG1) table and the Geologic Units (GLG) feature class, with the GLG1 table defined as the source, and the GLG feature class defined as the destination, if a record in GLG1 (source) table was deleted, all related records (in this case actual features) in the GLG (destination) feature class would automatically be deleted as well. Figure 10 displays a relationship between the Geologic Units (GLG) feature class and the Geologic Unit Information (GLG1) table.

Geodatabase Data Model Benefits

Several benefits exist for a geodatabase-based data model. These include:

- All geographic data is centrally stored and managed in one database.
- The availability of subtypes, domains, relationship classes, and topological rules help maintain database integrity and reduce database maintenance by making data entry and editing more efficient and accurate.
- Previous data formats (i.e., coverage and shapefile) can be created via data export from a geodatabase.
- Geodatabase specifications or schema can be replicated and reused for production purposes using Microsoft Visio, Extensible Mark-up Language (XML) and/or Computer-Aided Software Engineering (CASE) tools.
- Geodatabase annotation can be linked to respective features. When annotated features are altered, the

Simple feature class

Geologic Units (glacglg)

Geometry *Polygon*
Contains M values *No*
Contains Z values *No*

Field Name	Data Type	Allow nulls	Field Definition	Domain	Pre	Scale	Length
OBJECTID	Object ID	-	feature object ID	-	-	-	-
SHAPE	Geometry	Yes	feature geometry	-	-	-	-
GLG_ID	Long integer	No	feature ID #	-	0	-	-
GLG_SYM	String	No	unit age-lithology symbol	-	-	-	12
SRC_SYM	String	No	source unit age-lithology symbol	-	-	-	12
NOTE	String	No	text notes and remarks	-	-	-	254
GMAP_ID	Long integer	No	source map ID # (in MAP table)	-	0	-	-
HELP_ID	String	No	unit (page) variable for Help File	-	-	-	12
SHAPE_Length	Double	Yes	feature perimeter	-	0	0	-
SHAPE_Area	Double	Yes	feature area	-	0	0	-

Relationship class
Unit Info to Geologic Units
One to many

Table

Geologic Unit Information (glacglg1)

Field Name	Data Type	Allow nulls	Field Definition	Domain	Pre	Scale	Length
OBJECTID	Object ID	-	feature object ID	-	-	-	-
GUNIT_ID	Long integer	No	unit ID #	-	-	-	-
GLG_SYM	String	No	unit age-lithology symbol	-	-	-	12
GLG_NAME	String	No	unit name	-	-	-	100
GLG_FM	String	Yes	formation or equivalent grouping	-	-	-	100
GLG_MBR	String	Yes	member or equivalent grouping	-	-	-	100
G_REL_AGE	String	No	relative age of unit	-	-	-	6
GLG_AGE_NO	Float	No	unit age-sort number (young to old)	-	6	3	-
G_AGE_TXT	String	No	unit geologic time text	-	-	-	50
G_AGE_PER	String	No	unit geologic time period	-	-	-	100
G_MJ_LITH	String	No	unit major lithology type	Coded	-	-	3
G_LITH_ID	Long integer	No	unit lithology	Coded	-	-	-
G_LITH_TXT	String	No	lithology text of unit	-	-	-	254
G_NOTE_TXT	String	Yes	notes and remarks of unit	-	-	-	254
GMAP_SRC	String	No	list of source map GMAP_IDs	-	-	-	100

The relationship class helps in data quality assurance by displaying an error during validation when values in the Geologic Unit Symbol (GLG_SYM) field do not match what is in the Geologic Unit Information table.

Figure 10. A relationship class relating the Geologic Unit Information (GLG1) table to the Geologic Units (GLG) feature class. The field relating the Geologic Unit Information (GLG1) table to the Geologic Units (GLG) feature class is the geologic unit symbol (GLG_SYM) field. Through the relationship, data from all other fields in the Geologic Unit Information (GLG1) table can be accessed, preventing duplication and data redundancy throughout the database.

feature-linked annotation also is altered.

- Provides more intuitive data objects instead of generic points, lines, and polygons.
- Geodatabases can accommodate very large datasets without the need for tiling or spatial partitions.
- More portable than coverages and shapefiles as data layers and tables can be stored in one geodatabase file, and not as multiple files.
- Enterprise geodatabases allow for multiple users to access data at the same time via versioning, and they can leverage additional functionality from additional connected robust databases.

Geodatabase Data Model Drawbacks

A few drawbacks also exist for the geodatabase-based data model. These include:

- Significant learning curve when migrating from coverage or shapefile format to a geodatabase for-

mat and data model.

- Requires duplication of data where polygon boundaries overlap and where bounding lines carry attribution.
- Geodatabases have not yet been universally adopted by GIS users, requiring those who use them to import from, and export to, other data types (shapefiles and coverages, for example).
- At its current version, ArcGIS still has many functionality problems (bugs).

IMPLEMENTATION USING XML SCHEMA

Because geologic features present on a geologic map frequently vary, a flexible approach to data model implementation using XML files for each feature class or two interdependent feature classes (i.e., Geologic Contacts (GLG) and Geologic Units (GLGA)) has been adopted. For example, not all geologic maps have faults. Having the functionality to implement a feature class to store faults only if these features are present decreases time spent creating the data layer if needed; including specifying attribute field parameters, domains, and participating relationship classes, and eliminates the need to delete components of a geodatabase that only pertain to a faults data layer should these not be desired in the final digital map.

Methods for implementing topological rules and relationship classes are currently under development. XML files are in the format accepted by Geodatabase Designer version 2, an ArcScript created by ESRI. Figure 11 shows a screen capture of Geodatabase Designer in ArcCatalog.

CONCLUSIONS

The current NPS Geology-GIS Coverage/Shapefile Data Model provides a robust method for storing geologic map data in a GIS. ESRI's new geodatabase model offers features and functionality that enhance the quality, portability, and scalability of digital geologic map data.

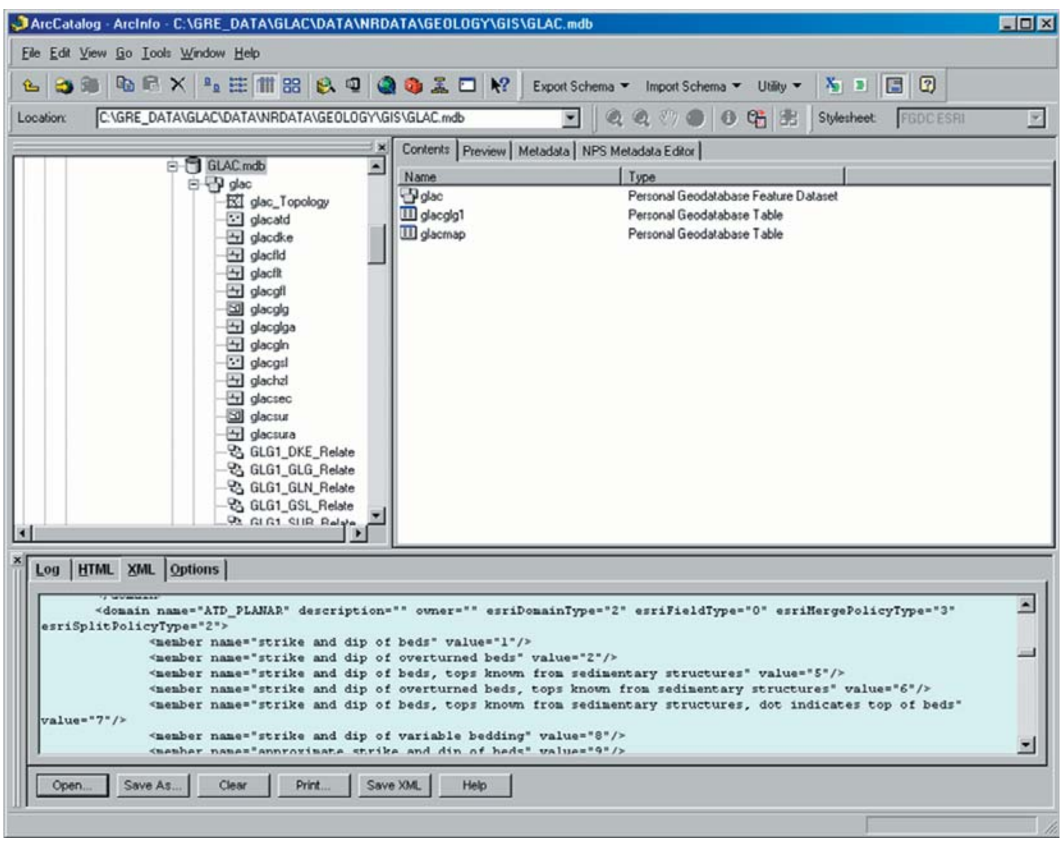


Figure 11. Screen capture of the Geodatabase Designer in ArcCatalog. The Geodatabase Designer is used to implement geodatabase feature class schema stored in XML files. A feature class XML schema file includes name and alias of the feature class(es), as well as field parameters, subtypes, and associated domains. XML schema for the Attitude Observation Localities (ATD) feature class is visible in the lower window of the figure. The schema is implemented in the Glacier National Park, Montana (GLAC) digital geologic-GIS map.

The decision to migrate to a geodatabase-based data model was influenced by the potential value that these new features and functionality could bring to the Geologic Resources Evaluation's (GRE) digital geologic map data program. The new NPS Geology-GIS Geodatabase Data Model incorporates the functionality of a geodatabase and enhances attribution and data integrity through the use of domains, subtypes, topology, and relationship classes. Current data formats (i.e., coverage and shapefile) also can be supported through export functionality included with ArcGIS.

FUTURE PLANS

The GRE program has identified several ideas to further develop, implement, and integrate a geodatabase-based data model into the production of digital geologic-GIS maps. These include:

- Further develop procedures for creating and presenting digital geologic-GIS map data in a geodatabase.
- Continue to refine database design, including information stored in ancillary tables (e.g., geologic unit information and source map).
- Develop improved methodology for storing map symbology (i.e., fault and fold symbols).
- Continue to refine current Geodatabase Designer XML file implementation of data model, including incorporation of topological rules and relationship classes.
- Further investigate other methods of data model implementation, including XML functionality included with ArcGIS 9.x and CASE tools.
- Reproduce existing coverage-based procedures for Quality Control and feature class digitizing using ArcObjects programming. Create new schema implementation, data loading, and export routines using ArcObjects.
- Develop methodology for efficiently producing FGDC metadata for feature datasets and object classes (i.e., feature classes and tables) within a geodatabase.

REFERENCES

O'Meara, Stephanie, Joe Gregson, Anne Poole, Greg Mack, Heather Stanton and Jim Chappell, 2005a, National Park Service Geologic Resources Evaluation Geology-GIS Coverage/Shapefile Data Model, http://science.nature.nps.gov/im/inventory/geology/GeologyGISDataModel.htm.

O'Meara, Stephanie A., Heather I. Stanton, James Chappell, Greg Mack, Anne R. Poole and Georgia Hybels, 2005b, National Park Service Geologic Resources Evaluation Geology-GIS Geodatabase Data Model (Draft v. 1.2), http://science.nature.nps.gov/im/inventory/geology/GeologyGISDataModel.htm.

O'Meara, Stephanie A., Trista L. Thornberry, Anne R. Poole, and Victor deWolfe, 2003, The National Park Service Geologic Resources Inventory "From Paper to Digital: Exploring a Geologic-GIS Map", in D.R. Soller, ed., Digital Mapping Techniques '03—Workshop Proceedings: U.S. Geological Survey Open-File Report 03-471, p. 247-254, available at http://pubs.usgs.gov/of/2003/of03 471/omeara/index.html.

Environmental Systems Research Institute (ESRI) Inc, 2002, Building a Geodatabase: ESRI Inc., 380 New York St., Redlands, CA 92373, http://www.esri.com.

Whipple, James W., 1992, Geologic Map of Glacier National Park, Montana: U.S. Geological Survey Miscellaneous Investigations Series Map I-1508-F, scale 1:100,000.

Carrara, Paul E., 1990, Surficial Geologic Map of Glacier National Park, Montana: U.S. Geological Survey Miscellaneous Investigations Series Map I-1508-D, scale 1:100,000.

SOFTWARE REFERENCES

ArcGIS 8.3 and 9.0 (ArcCatalog and ArcMap) - Environmental Systems Research Institute (ESRI) Inc., 380 New York St., Redlands, CA92373, http://www.esri.com.

Geodatabase Diagrammer – Developed by Michael Zeiler, Environmental Systems Research Institute (ESRI) Inc., 380 New York St., Redlands, CA 92373, http://arcscripts.esri.com.

Geodatabase Designer v2 – Developed by Richie Carmichael, Environmental Systems Research Institute (ESRI) Inc., 380 New York St., Redlands, CA 92373, http://arcscripts.esri.com.

Microsoft Office Visio Professional 2003, Microsoft Corporation, http://www.microsoft.com.

The New Mexico Bureau of Geology and Mineral Resources' Preliminary Map Series: Expediting Dissemination of STATEMAP Geologic Map Data to the Public

By David J. McCraw, J. Michael Timmons, and Adam S. Read

New Mexico Bureau of Geology and Mineral Resources
801 Leroy Place
Socorro, NM 87801
Telephone: (505) 835-5594
Fax: (505) 835-6333
e-mail: djmc@nmt.edu

ABSTRACT

Each year, new geologic map data produced under contract to the USGS National Cooperative Geologic Mapping Program's STATEMAP program are in high demand in New Mexico. Scientists, decision makers, engineers, and others request or need these data for many reasons, as soon as possible. Given that cartographic production, reviews, and the application of map standards set by the New Mexico Bureau of Geology and Mineral Resources are lengthy processes, the Bureau has created a new Preliminary Map Series. These draft map data are now available for download as PDF files from the Bureau's website in early June of each year, immediately after completion of the mapping and delivery of the STATEMAP contractual requirement to the USGS. These preliminary maps are either cleaned and rectified scans of authors' original greenline mylars, or original digital (vector) linework data that are placed into a draft layout template that includes rudimentary ancillary data (cross sections, map unit descriptions, correlation charts, photos, etc.). The draft map is then uploaded to the Bureau website, for public access. After digitization, attributing, and labeling of the map and cross sections, this online preliminary map is replaced by a full-color, but still preliminary, version of the map; after peer review, that map is superceded by the final version published in the Bureau's formal Geologic Map Series.

PURPOSE OF THE PRELIMINARY MAP SERIES

In New Mexico, geologic maps are in high demand by the government (scientists, decision makers), private industry (engineers, etc.), academia, and by the general public. Geologic maps combine descriptive information (such as materials and structures) and interpretations (e.g., geologic processes) into a conceptual framework that relates all of the geologic elements through time. It is a powerful tool that describes the geologic environment and permits us to predict how natural systems are likely to behave in the future. In addition, they provide immediate economic benefits, adding up to many millions of dollars saved.

Studies have shown that geologic information is important to government and private industry for a variety of environmental and economic applications, with the following being the most common applications:

1. Exploration and development (ground water, industrial minerals, metallic minerals, oil and gas, coal),
2. Environmental consulting (pollution prevention, site cleanup, industrial issues),
3. Hazard prevention and protection (landslides, earthquakes, soil stability, mine subsidence, sinkholes, volcanic eruptions, floods),
4. Engineering applications (buildings and foundations, roads, pipelines, dams, utilities, railroads),
5. City planning (zoning decisions, landscape planning, building codes),
6. Regional planning (regional water plans, waste disposal, industrial permits, planning transportation corridors), and
7. Property valuation (land acquisition, property tax assessment, cost-benefit analysis).

Specifically in New Mexico, new geological quadrangle maps produced by the New Mexico Bureau of Geology and Mineral Resources' STATEMAP program are being used to support a great variety of environmental and hydrologic work throughout the state, especially along the Rio Grande where a majority of the state's population lives.

The Bureau's STATEMAP mapping priorities are set annually by a 35-member STATEMAP Advisory Com-

mittee (SMAC),, composed of hydrologists, geologists, and planners from state, local, federal, tribal, and private agencies and entities. The quadrangles to be mapped are selected based on their potential to provide essential earth science data to planners, engineers, geologists, and hydrologists. Our mapping program is especially important to New Mexico because, of the approximately 2000 7.5-minute quads in the state, less than 25% have been mapped at the standard scale of 1:24,000. The most critical unmapped areas are along the population centers of the Rio Grande corridor. Therefore, it is not surprising that much interest is generated each spring when the Bureau finishes a new set of 24K geologic quadrangles (usually around 10-12 per year).

Unfortunately, formal cartographic production of these maps cannot keep pace with actual geologic mapping. In addition to the time consuming processes of digital data capture and layout, each map must go through both scientific and editorial review, and must meet Bureau map standards, before they are finally released as "official" Bureau geologic maps in its Geologic Map (GM) publication series. This often adds an additional year or more to each quad beyond the point where a) the map is delivered to the USGS as per contractual requirement, and b) the geologist(s) deem the mapping sufficiently stable (complete) for production to commence. For many people who want the map data, this lengthy process is unacceptable. Hence, the Bureau developed its "Preliminary Map Series," to expedite map data dissemination.

Now, each May, prior to contractual delivery to the USGS, each map is put into a DRAFT preliminary layout template, which contains ancillary data such as cross sections, correlation charts, report, photos, well data, etc. (Figure 1a). This is usually accomplished by scanning, cleaning, and rectifying each new map's greenline mylar (which shows the geology and the base map), and inserting these data into an ArcGIS preliminary map template (see Figure 1b for details of the map shown in Figure 1a). Occasionally, actual digital linework vector data are provided by the compiler, and these are used. As hard copy delivery is made to the USGS at the end of each May, a PDF file of each draft preliminary map is uploaded to the Bureau's STATEMAP webpage, for public access. Thus, map data are immediately available. After the map linework has been digitized and map polygons have been attributed and labeled in ArcGIS, the map proceeds to the second phase of the Preliminary Map Series, as the rectified scan of the greenline is replaced with the ArcMap file and uploaded to the public site to replace the draft map (Figure 2a and 2b; note that whereas these figures are in grayscale, the available maps are in 24 bit CMYK color). The map then is processed through final layout and review. Upon completion, the Preliminary Map is superseded and replaced by a map formally published in the Bureau's Geologic Map (GM) Series.

PROCEDURE

The greenline, a digital green ink base map provided by the Bureau to the mappers both digitally and on mylar.The procedure for generating a digital greenline was described at the DMT '02 meeting, by Read, et al. (2002). The geology is inked by the compiler, and then is scanned on a Colortrac 5480 scanner to produce a 400-dpi RGB (24 bit) file. In Adobe Photoshop, the brightness and contrast are adjusted and the image is converted to the USGS standard DRG color palette (13 colors), thus greatly reducing file size and isolating the geologic linework into 1 or 2 color bins. The magic wand tool is then used to remove artifacts of the scanning process. The file then is rectified in ArcMap, and the colors are isolated into 2 bins: linework in black and base material in light green. The map then is ready for export to Adobe Acrobat, as well as for digitizing in ArcMap.

In Adobe Acrobat, the cleaned and rectified map image is combined with ancillary data files. Unit descriptions are taken directly from the report (a Microsoft Word document) and placed next to the map; the result is saved as a PDF file. Graphics (cross sections, correlation chart, location maps, etc.) are delivered from the geologists either as Adobe Illustrator files (which can go directly to the PDF file) or as drawings, etc., which are scanned and saved as PDF files in Adobe Photoshop. Photos are handled in a similar fashion. These materials then are assembled onto a 36"x 50" map sheet in a draft layout, and the "reduce file size" command is run, resulting in a PDF file of about 3-4 MB (e.g., Figure 1a). Finally, the map is printed, submitted to the USGS as the STATEMAP contract deliverable, and uploaded to the Bureau website for public access.

Next, the map linework is digitized and attributed, the map polygons are built and attributed, and map features are labeled in ArcGIS, the "second phase" preliminary map is generated. In this new version of the map, the greenline scan is replaced by the ArcMap project, and a new PDF file is created (e.g., Figure 2a). Overall quality and legibility are thereby greatly enhanced, even though final cartographic production, final map layout to Bureau standards, and scientific/editorial reviews remain pending.

SUMMARY

To expedite dissemination of new geologic mapping funded by the USGS National Cooperative Geologic Mapping Program's STATEMAP program in New Mexico, the New Mexico Bureau of Geology and Mineral Resources has developed the Preliminary Map Series. Draft geologic map data are scanned, cleaned, and rectified into a draft or preliminary map layout and are uploaded to the web when they are delivered to the USGS to fulfill the contractual obligation to the STATEMAP program. As the map pro-

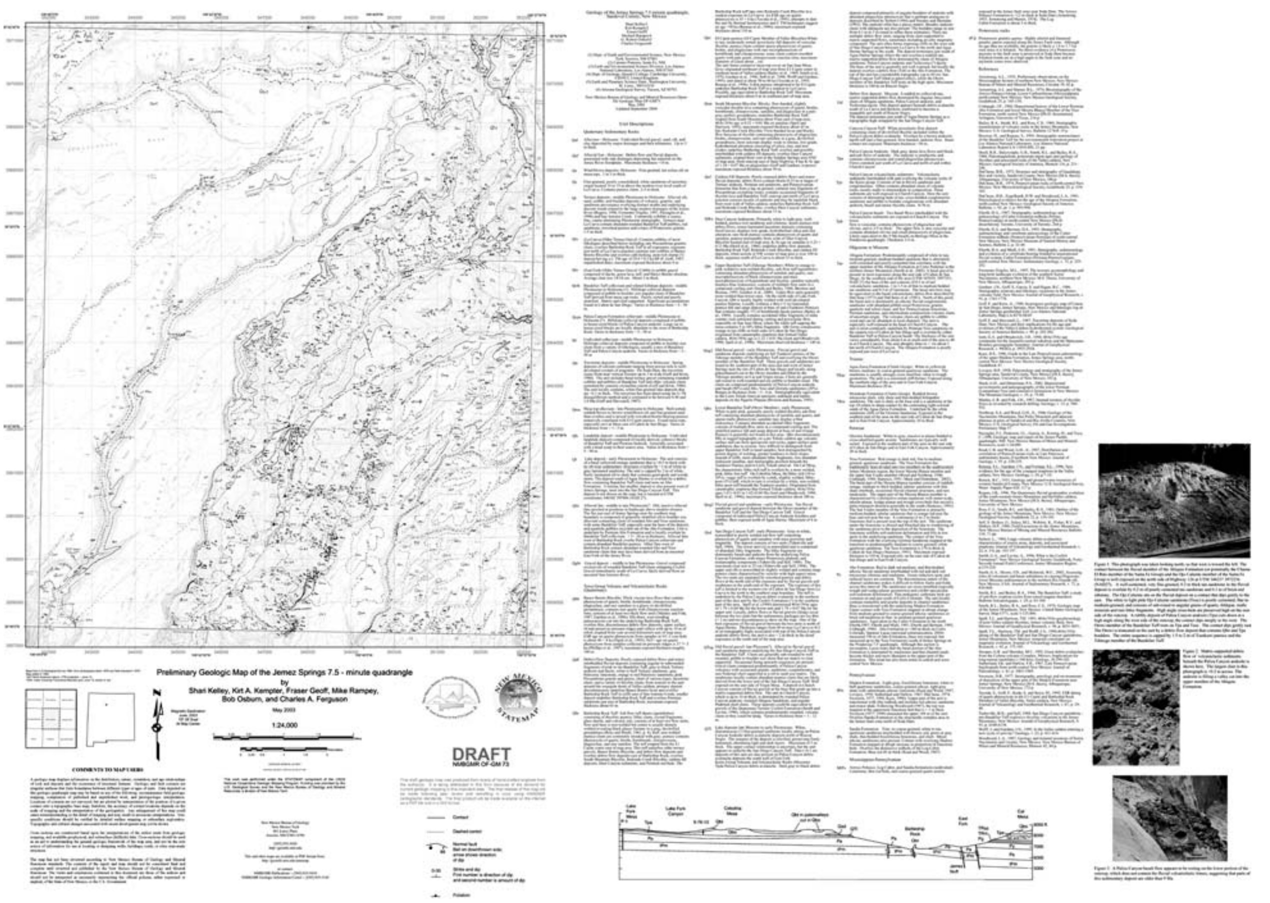


Figure 1a. Draft version of the geologic map of the Jemez Springs quadrangle, published in the Bureau's Preliminary Map Series (map scale 1:24,000; paper size 36"x 50"). Scanned geologic greenlines, or digital vector linework data placed upon a greenline raster base, are cleaned and rectified, and placed in the draft map layout, with Microsoft Word text (map unit descriptions, etc.) and original hand-drafted ancillary data (cross sections, correlation charts, etc.) and photos. This map is available in PDF format.

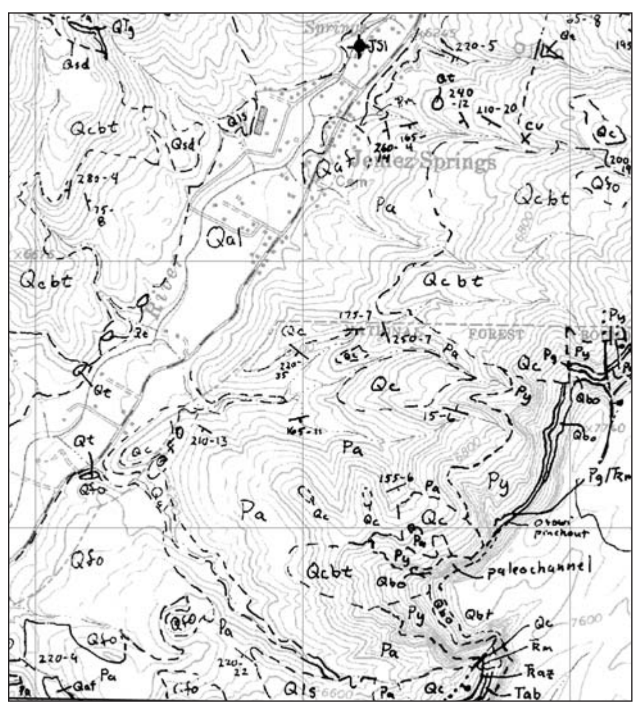


Figure 1b. Detail of geologic greenline from Figure 1a, showing the area around the village of Jemez Springs, NM (Note: on the published map, geologic linework is depicted in black and base is depicted in green).

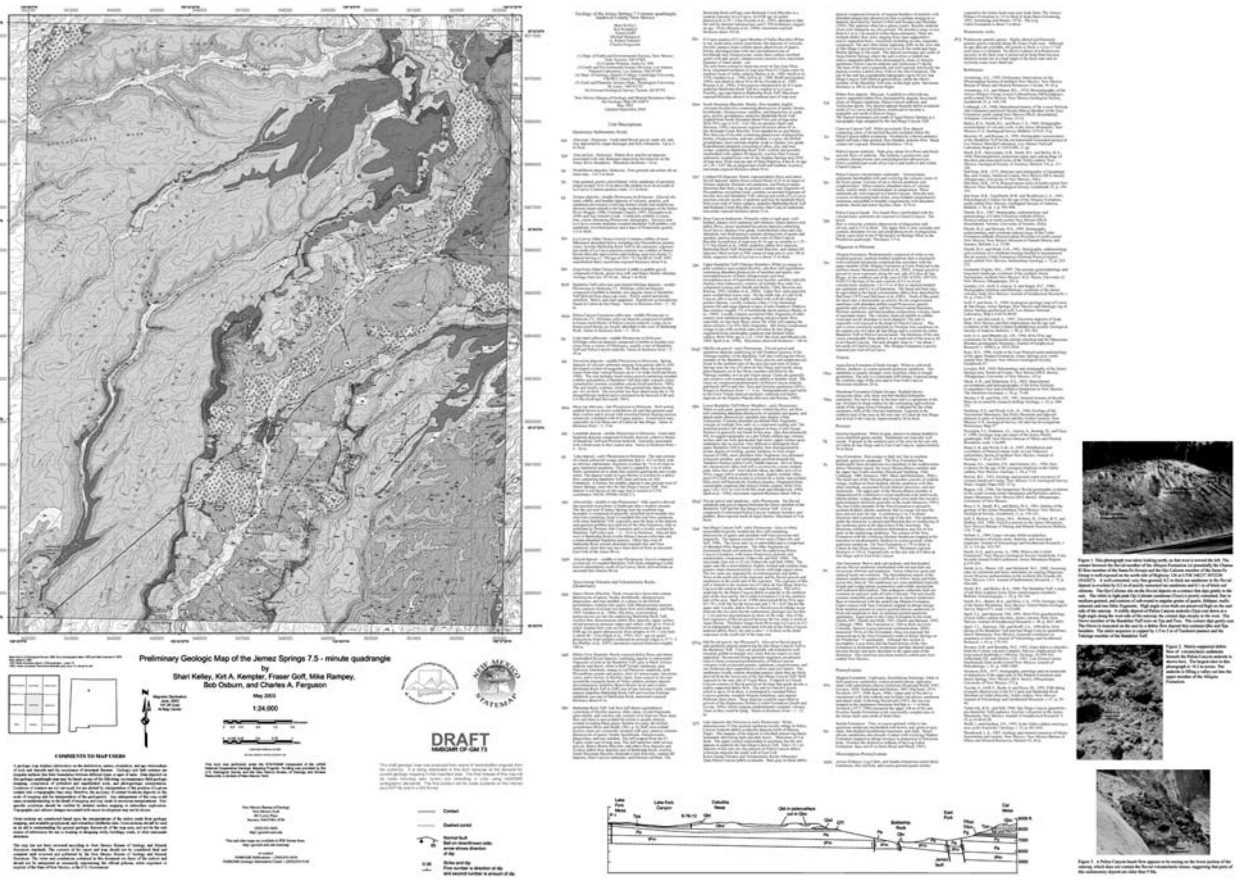


Figure 2a. The "second phase" version of the geologic map of the Jemez Springs quadrangle, which replaces in the Bureau's Preliminary Map Series the draft version shown in Figure 1a. Here, the original greenline has been replaced by the ArcGIS digitized map (and displayed in color, although map units may not meet Bureau color standards).

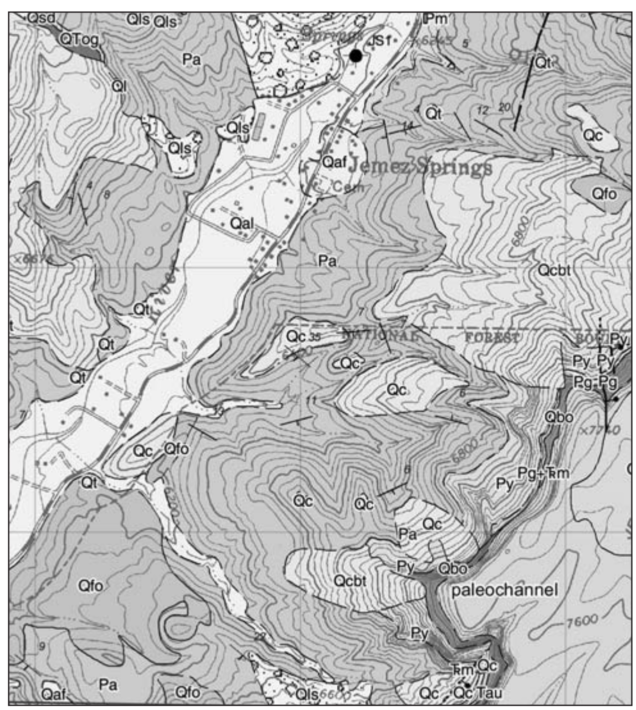


Figure 2b. Detail of geologic map from Figure 2a, showing the area around the village of Jemez Springs, NM (for comparison, this is approximately the same area as depicted in Figure 1b).

gresses through digital data capture, scientific peer review, editing, and final layout commensurate with Bureau map standards, what is available to the public via the PDF file on the Bureau website is updated. Hence from "Day One" after contractual deadlines are met for each quadrangle, the public has access to the most current geologic map data via the Preliminary Map Series. Upon completion of all phases of map and editorial production, each quadrangle within the Preliminary Map Series is replaced and superseded by the final map in the Bureau's GM series.

REFERENCE

Read, A.S., McCraw, D.J., and Rawling, G., 2002, New tricks for old dogs: A digital technique for producing mylar base maps for geologic mapping and compilation, in D.R. Soller, ed., Digital Mapping Techniques '02—Workshop Proceedings: U.S. Geological Survey Open-file Report 02-370, p. 111-117, available at http://pubs.usgs.gov/of/2002/of02-370/read.html.

Potential Issues with the Use of LIDAR for Geologic Mapping in Louisiana

By Richard P. McCulloh

Louisiana Geological Survey
Louisiana State University
3079 Energy, Coast and Environment Building
Baton Rouge, LA 70803
Telephone: (225) 578-5327
FAX: (225) 578-3662
e-mail: mccullo@lsu.edu

ABSTRACT

The detail afforded by LIDAR imagery gives striking views of surface features previously known much more poorly or not at all. For some previously known features such as the traces of active faults in the south Louisiana coastal plain, LIDAR permits mapping with much greater accuracy and detail than previously was possible. Nevertheless, experience with newly available LIDAR imagery in Louisiana suggests that not all of the aspects resolved by the improved detail are meaningful in the context of mapping surface geology. For example, a physiographic character dominated by an abundance of fine-scale surface relief may obscure gross-scale characteristics diagnostic of the units composing the mapping schema of a project, i.e., the units that a given mapping exercise is intended to delineate. Abundant detail may also create suggestions of subdivisions of primary map units that are not borne out by observation. Even where the detail resolves surface aspects considered geologically meaningful, fitting LIDAR-based interpretations of them to a suitable base map in some areas remains a practical problem for the near-term future, because topography compiled by non-photogrammetric means on some older base maps cannot be matched to such detail. When custom production of user-designed large-scale topographic base maps from the same LIDAR DEMs becomes cost-effective and routine, use of LIDAR imagery in geologic-mapping applications likely also will become commonplace; until then, such use probably will remain selective, though LIDAR imagery should continue to be a medium of discovery, and a tool for revision and amplification of map interpretations within the constraints of existing map editions.

INTRODUCTION

Between late 1999 and early 2000, a program was begun to generate LIDAR data coverage for the state of Louisiana, the funding source being the Louisiana Federal Emergency Management Agency and the U.S. Army Corps of Engineers, St. Louis District. At this time, coverage is primarily restricted to south Louisiana, with limited portions of north Louisiana (Rapides, Ouachita, and Lincoln parishes) having received coverage in recent years. This paper summarizes impressions regarding the use of LIDAR imagery in geologic-mapping applications during the first several years of its availability in Louisiana, by an author having no previous exposure to it. As such, the views expressed herein reflect a limited baseline of experience. All the LIDAR digital elevation models (DEMs) figured herein were generated by the program mentioned above and obtained as downloads from the Atlas website (http://atlas.lsu.edu/), and all the views of DEMs in these figures were created with Global Mapper DEM viewer.

Existing 7.5-minute topographic quadrangles in south Louisiana use a 5-ft (1.5-m) contour interval, whereas the LIDAR datasets generated by the above program exceed this vertical precision by greater than an order of magnitude, with an effective vertical resolution of approximately 0.03 m (0.1 ft) following correction for vegetation[1] (the instrumental vertical resolution is reported to be 0.003 m (0.01 ft) at a horizontal grid spacing of 5 m). This increased detail (Figure 1) enables the imaging of features previously perceived only dimly or not at all. Some features previously known appear stunningly and unambiguously imaged with LIDAR data, especially active surface faults in the south Louisiana coastal plain, the traces of which appear with unprecedented clarity. One issue preventing the immediate wholesale use of LIDAR imagery for large-scale geologic-mapping applications, however, is that in many areas the conflicts between the existing large-scale USGS base map and the LIDAR-derived feature renderings are nontrivial and intractable, and

[1]Effective vertical resolution assessed by 3001, Inc. (New Orleans), which flew the LIDAR.

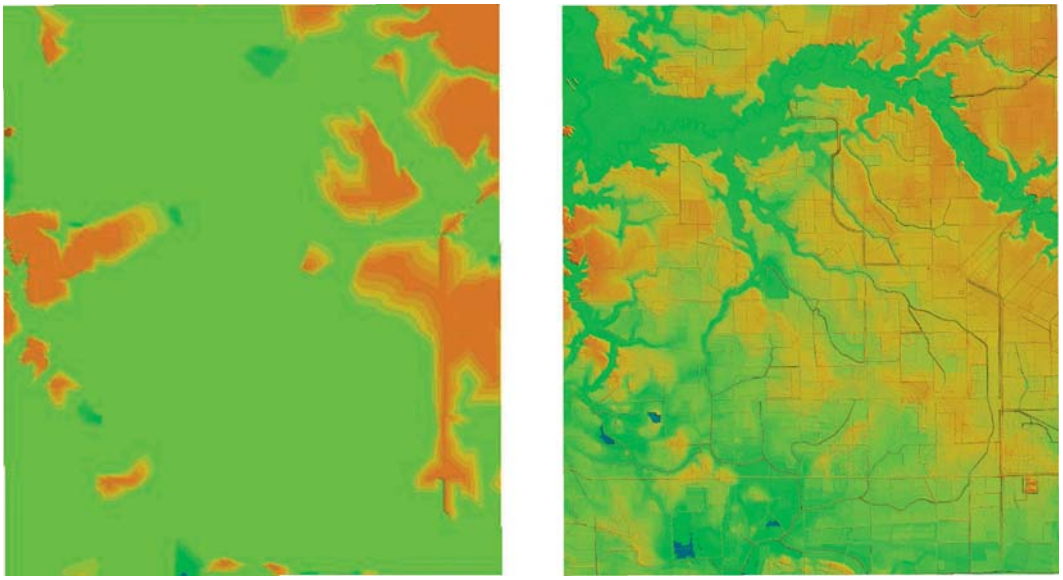


Figure 1. Example of improved detail afforded by LIDAR DEMs (right) over standard 7.5-minute quadrangle DEMs (left) in a generally low-relief area (Gueydan quadrangle, southwestern Louisiana). Colored shaded-relief views of digital elevation models in this and subsequent figures were downloaded from the Atlas website (http://atlas.lsu.edu/) and viewed with Global Mapper DEM viewer. LIDAR data source: Louisiana Federal Emergency Management Agency, and U.S. Army Corps of Engineers, St. Louis District.

as a result standard large-scale topographic base maps on which new LIDAR-based interpretations may be properly georeferenced are not available for these areas. This is especially the case with quadrangle sheets for which topography was compiled by nonphotogrammetric means.

ASSORTED ISSUES, SELECTED EXAMPLES

1. Artifacts (artificial differences in perception of the same DEM in this case) may result from (1) choice of a particular shader option, because of varying degrees of differentiation of the vertical color scale by different shaders; (2) change of the elevation scale range and elevation color contrast associated with an area of interest, after mosaicking an image of the area into a composite image of a larger area (with shader options that employ a uniform color spectrum over the entire vertical scale applicable to an image); and (3) the random association of particular elevation values with particular hues as a result of these choices. (This is somewhat analogous to the differences one can obtain in contour maps of the same landscape by selecting slightly different contour intervals, or by selecting the exact same contour interval but using different datums that are not whole-number multiples of it.) Such accidental associations and results may affect user perception in nontrivial ways (Figure 2).

One way to decrease the influence of such artifacts on viewer perception generally is to create several DEM views, each with a different shader option. Effect (2) listed above can be eliminated by viewing the DEM of both the single image and the mosaicked composite image with at least one shader that assigns unique hues to particular elevation values in a series that repeats itself rather than covering the entire elevation range for the image area with a single color spectrum.

2. Preparation of an interpretive map using any new technique is an iterative process that begins with a period of orientation involving the calibration of sense perceptions to cues considered meaningful or found to be so by the mapper. Following this initial adaptation the mapper conducts ongoing tests of these cues while employing them as recognition criteria to organize the growing body of perceptions into a conceptual framework. Greater experience leads to more effective focus on what is judged to be essential.

In Louisiana, the advent of LIDAR coverage has implications both for the recognition of map units against a field of abundant surface detail largely undiagnostic of them, and for the potential suggestion of phantom subunits by such detail. Fine surface detail may mask the gross geomorphic signature of the units of surface geology that the mapper considers it meaningful to distinguish,

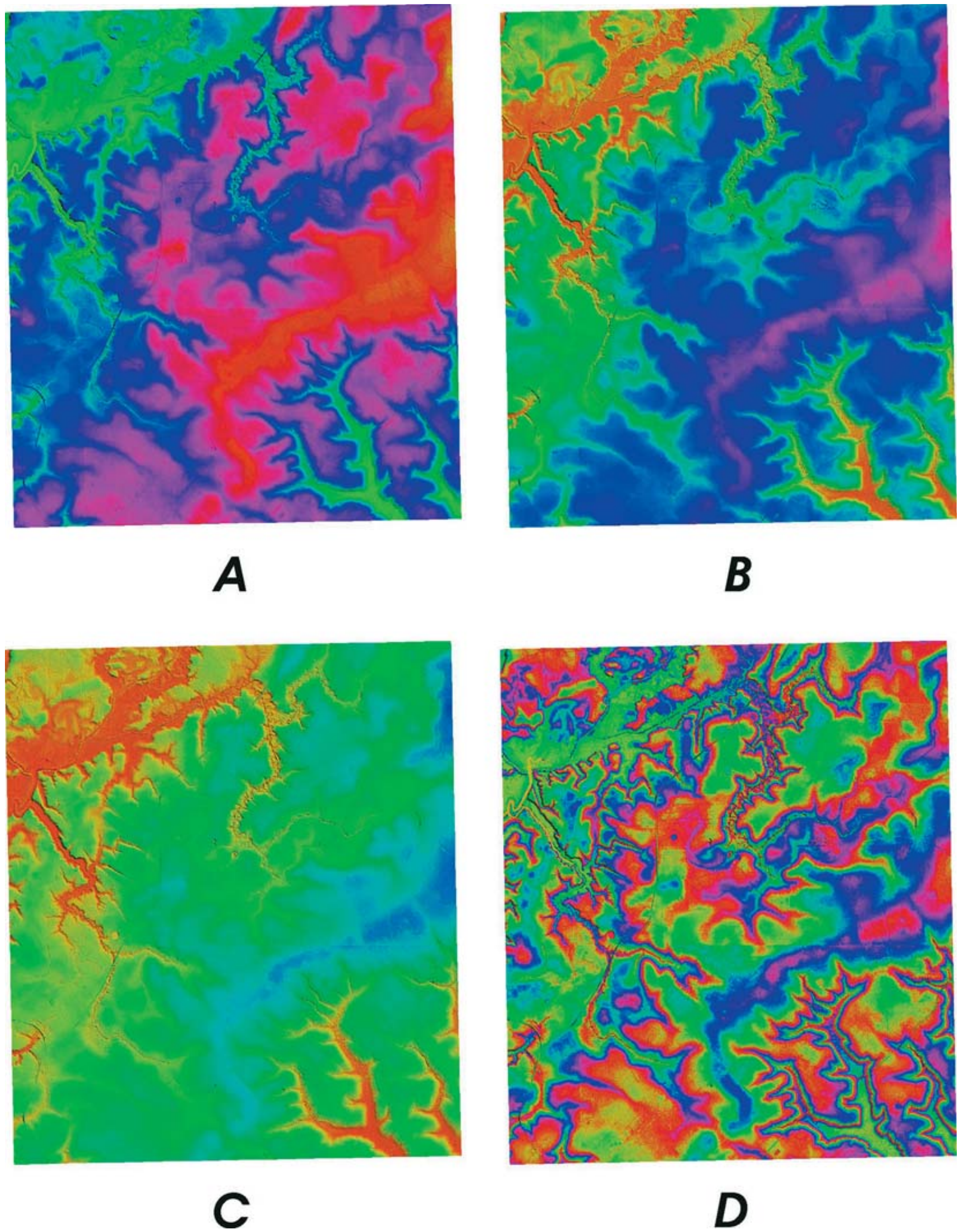


Figure 2. Artificial differences in appearance created using different shader options for the same LIDAR DEM of an area: ***A*** and ***B*** represent different starting colors, ***C*** and ***D*** represent different color ranges (northeast quarter of French Settlement quadrangle, southeastern Louisiana).

e.g., a dramatic increase in dissection and relief across a drainage divide associated with proximity to a river or major stream may obscure a contact between two mappable bedrock geologic units in the same area. In south Louisiana the opposite problem more likely occurs, in that land surface relief on the coastal plain may, if rendered in fine detail, generate what appear to be potential contacts within the outcrop belts of the units being mapped (Figure 3). Upon examination, however, these putative contacts typically do not correlate readily with recognizable subdivisions of the primary map units; as a result, careful evaluation of every such suggestion by means of other criteria is essential.

3. The land surface of Louisiana shows a perceptible, albeit intermittent, rectilinearity of its drainage courses, which is suggestive of control by systematic fractures (McCulloh, 1995 and 2003). Older topographic maps compiled nonphotogrammetrically at scales of 1:62,500 and larger may miss this character completely where it occurs at smaller dimensions in smaller areas. An example is shown from the Fred 7.5-minute quadrangle (Figure 4): the headward area of one drainage system shows a strikingly rectilinear pattern on the LIDAR DEM, whereas the non-photogrammetrically compiled topographic base gives little or no suggestion of this. This aspect (as the distribution of interpreted alluvium) cannot be rendered on a map of surface geology fitted to the existing base, and no alternative base is likely to become available for some time unless custom production of user-designed base maps becomes cost-effective and routine.

4. Rapid change of surface detail in an area also generates conflict with existing base maps, making them obsolete regardless of how they were compiled. In Louisiana, this is especially characteristic of areas experiencing changes in the position of the land-water interface, such as rapidly eroding and accreting banks within flood plains and along estuaries of the major rivers, and shorelines subject to high rates of coastal erosion and land loss. In such areas, images of the DEM generated from recently flown LIDAR data will be in conflict with existing base maps regarding areas corresponding to land and open water (Figure 5). Such change may be largely unrelated to contacts between mapped geologic units; if the surface-geology layer incorporates substantial LIDAR-based interpretations, however, fitting it to a suitable 7.5-minute base map will then require either custom creation of a new base from the same LIDAR DEM or revision of the existing base involving substantial adjustment in places.

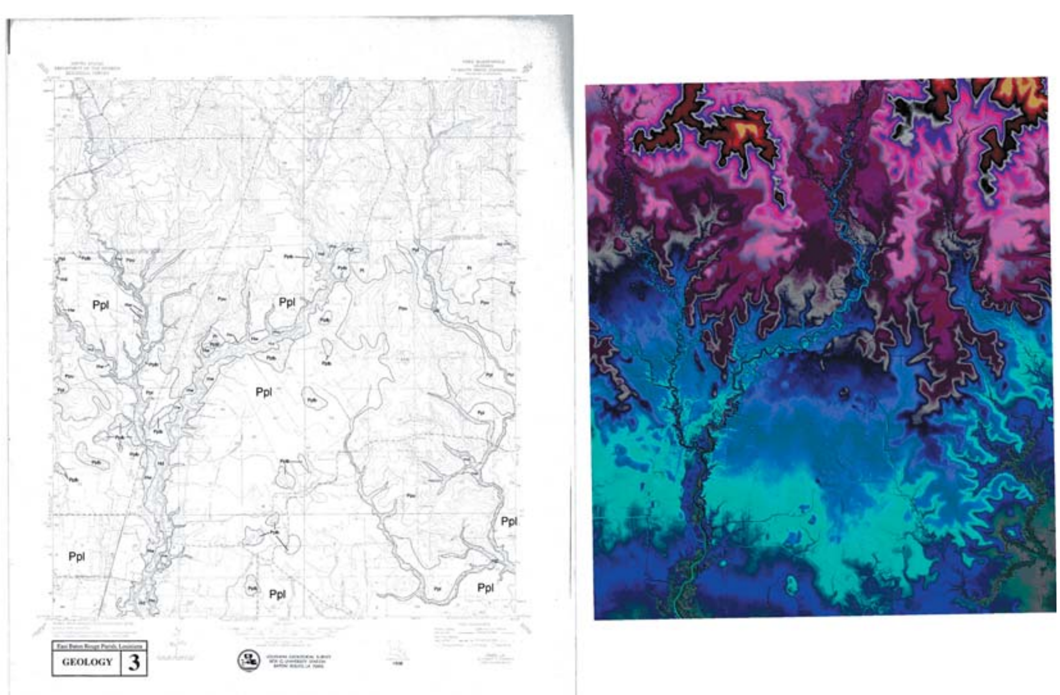


Figure 3. Putative subunits of a geologic map unit suggested by the detail on a LIDAR DEM (Fred quadrangle, north-central south Louisiana). In this case the large Ppl polygon in the southern portion of the quadrangle (geologic map—Plate 3 of Autin and McCulloh 1991, left) belongs to a unit the surface of which is tilted to form a ramp, and in this view of the LIDAR DEM (right) the unit appears potentially subdivisible because of the particular suite of shader attributes chosen. (Geology is mapped only in that portion of the quadrangle lying in East Baton Rouge Parish.)

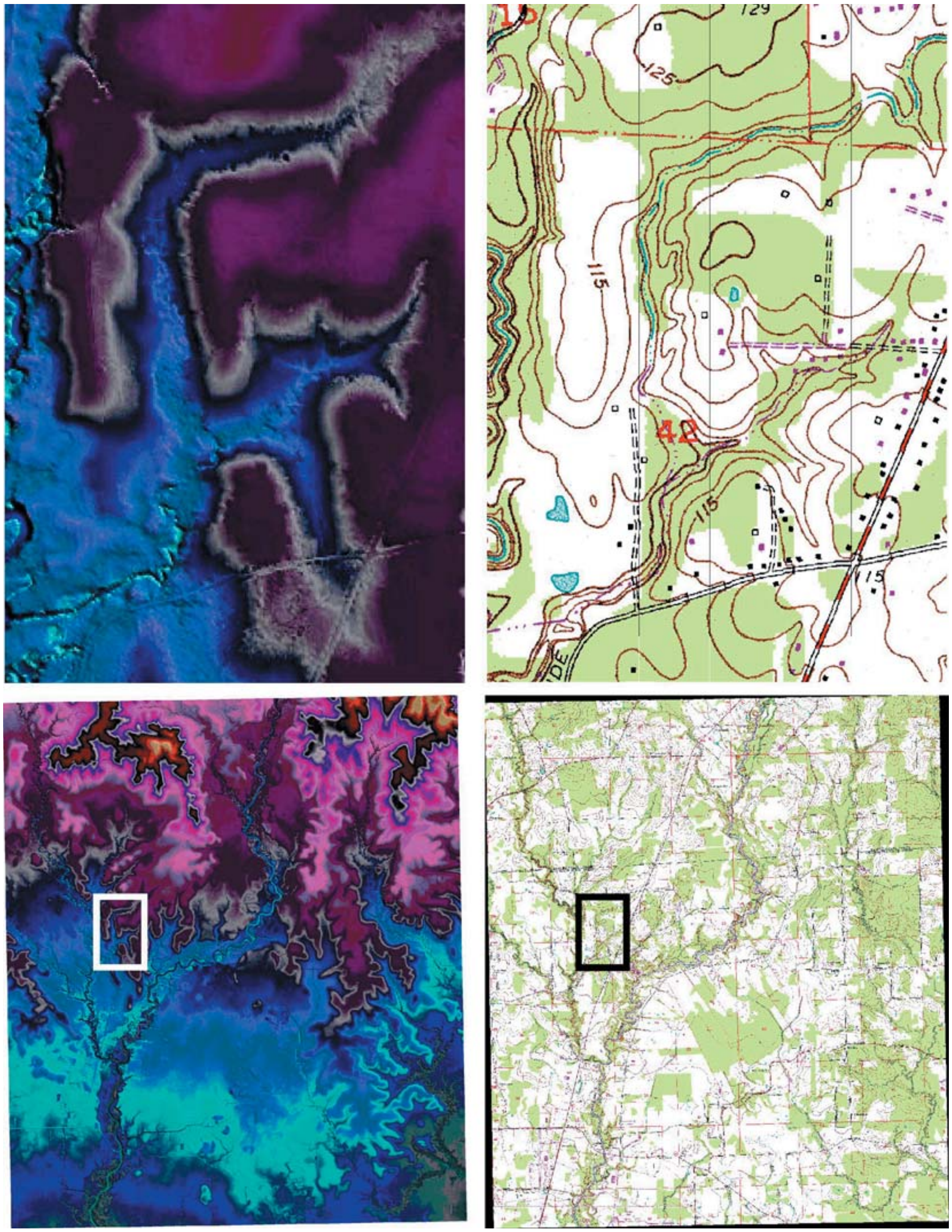


Figure 4. Rectilinear aspect of the headward portion of a drainage shown by LIDAR (left) is much less obvious on the 7.5-minute topographic quadrangle (right), which was non-photogrammetrically compiled (Fred quadrangle, north-central south Louisiana: LIDAR quarter-quadrangle mosaic, lower left, and enlargement of inset area, upper left; USGS 7.5-minute topographic quadrangle, lower right, and enlargement of same inset area, upper right).

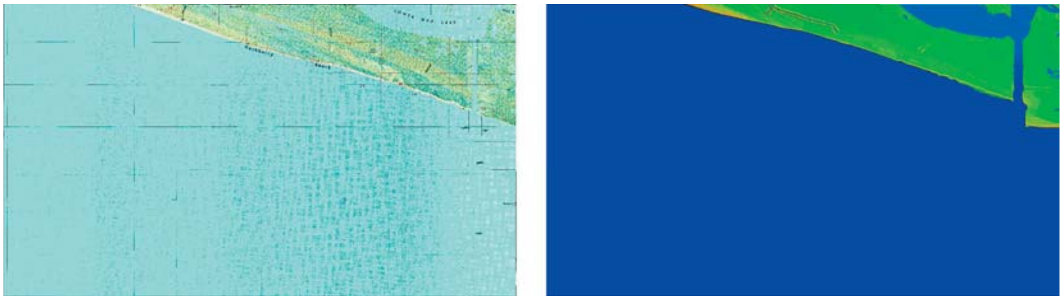


Figure 5. LIDAR DEM (right) depicting shoreline substantially changed from that depicted on the 7.5-minute topographic quadrangle (DRG, left—Hackberry Beach quadrangle, southwestern Louisiana).

The author's experience in Louisiana thus far suggests that for the present, LIDAR imagery is (1) a medium of discovery, revealing hitherto unknown surface features with topographic expression but of such slight relief that elevation contours on standard topographic maps previously had failed to resolve them adequately; and (2) a tool for the modification and/or refinement of previous interpretations (e.g., of active faults displacing surface Quaternary strata) within the constraints of existing editions of topographic and geologic maps. Until it becomes feasible to rapidly prepare custom topographic base maps on suitable media using elevation contours generated from LIDAR DEMs, depiction on published geologic maps of certain LIDAR-based interpretations (those of an intricate nature and/or requiring fitting to the drainage courses of the hydrographic layer) is likely to remain limited.

SUMMARY

Although LIDAR-based imagery provides views of the ground surface with unprecedented detail, only some of this detail reflects aspects likely to prove meaningful in the context of most surface-geologic mapping projects. In places, such surface detail may in practice obscure the comparatively gross-scale characteristics considered to have some diagnostic value for the units being mapped. Even where detail provided by LIDAR imagery is found useful in the context of mapping, fitting of more intricate and/or complicated LIDAR-based interpretations of surface geology to suitable large-scale base maps will remain a basic problem for the near-term future, until it becomes feasible for users to rapidly generate new topographic base maps based on LIDAR DEMs. For now, LIDAR imagery is a medium enabling discovery of previously unknown surface features, and a tool for updating map interpretations within the constraints imposed by existing map editions.

ACKNOWLEDGMENTS

Paul V. Heinrich (Louisiana Geological Survey) informed the author of the example of shoreline change shown in Figure 5. David R. Soller (USGS) reviewed and edited the manuscript.

REFERENCES

Autin, W.J., and McCulloh, R.P., 1991, Geologic and derivative engineering geology hazard maps of East Baton Rouge Parish, Louisiana: Louisiana Geological Survey Open-File Series no. 91–01, prepared for East Baton Rouge Parish Department of Public Works under project no. 90-MS-CP-0024, 31 plates plus index and explanation.

McCulloh, R.P.,1995, Fisk's NE- and NW-striking "faults" and the potential for joint control of stream courses in Louisiana: Gulf Coast Association of Geological Societies Transactions, v. 45, p. 415–422.

McCulloh, R.P., 2003, The stream net as an indicator of cryptic systematic fracturing in Louisiana: Southeastern Geology, v. 42, no. 1, p. 1–17.

Improving Geologic Data through Aerial Photography in the Ashepoo-Combahee-Edisto River Basin

By Erin E. Hudson and Jennifer L. Krauser

South Carolina Geological Survey
5 Geology Road
Columbia, SC 29212
Telephone: (803) 896-7714
Fax: (803) 896-7695
e-mail: hudsone@dnr.sc.gov, krauserj@dnr.sc.gov

INTRODUCTION

During a 5-year period, the South Carolina Geological Survey mapped 23 7.5-minute quadrangles in the Ashepoo-Combahee-Edisto River ("ACE") Basin and Hilton Head areas in South Carolina. This mapping was funded by the U.S. Geological Survey STATEMAP program. The ACE Basin is a rapidly changing coastal ecosystem at the convergence of these three rivers.

Doar and Hudson (2004) stated, "Geology has a major part to play in coastal zone resource and hazard management. Understanding the physical environment of the coastal zone, in particular its geomorphological evolution, is essential for determining a strategy for land use and development that is both sustainable and sensitive to the rich wildlife of the ACE Basin. Geologic information also provides a long-term perspective on coastal zone issues. This perspective is important because it is largely unaffected by tidal or seasonal variability. Mapping provides information on tectonic processes that influence uplift or downwarp in the coastal zone. Such uplift or downwarp determines whether large estuaries will develop along any particular part of the coastline. Defining tectonic processes affecting the coastal zone also is important in understanding relative sea-level change. Sea levels have been rising since glacial ice began to retreat in the northern hemisphere 18,000 years ago. The location and age of depositional units within the coastal zone provide information on how the coastline developed. Integrating such information with offshore data will show how quickly the coastline has changed or can change during periods of rapid sea-level rise."

Accurate base maps provide a basic tool for delineating geomorphic features throughout an area. While developing digital maps, the Geological Survey realized that the USGS topographic quadrangles, originally used by field geologists to delineate geologic contacts, were outdated. In some cases they were more than 40 years old. The dynamic nature of this coastal environment and the availability of current aerial imagery led to the conclusion that a much better geologic data set for the area could be created.

PROCEDURE

Step 1: Digitization

The geologist's field maps are drawn on a 7.5-minute topographic base. Finished maps are scanned at 200-300 dpi, depending on the complexity of the map, and saved as a TIFF image. The TIFF is then opened in R2V (raster to vector software), and control points are selected. Most often the four corners of a 7.5-minute quadrangle are suitable control points. The hydrography DLG is imported into the project, the geologic contact lines are digitized, and these two are edited together. A line shapefile of the combination is exported. The bounding edge of the quadrangle is not included in the editing procedure in R2V, because the quadrangle border can inadvertently be pulled or snapped away from its correct position. The quadrangle boundary is merged with the geologic-contact line file in ArcMap, using the geoprocessing tool. The result is a new line shapefile containing all base and geologic lines.

Step 2: Creating a Geodatabase

In ArcCatalog, a personal geodatabase is created for the mapping area. A feature dataset is created by importing the digitized polyline shapefiles as feature classes. Topology rules are assigned in order to validate the topology of the new line-feature classes. Examples of line rules include: must be single part, must not intersect, and must not have dangles. A first-pass validation of these topology rules is performed in order to edit all the geology and hydrography lines that intersect the quadrangle boundary, and to snap them accordingly.

Step 3: Edit, Validate, and Code Lines

In ArcMap, the entire quadrangle is edited to ensure that all of the topology rules for lines are followed. After the lines are validated, they are coded for type and position. In lower Coastal Plain maps, the lines are coded as geologic contacts or water boundaries. In the Piedmont and upper Coastal Plain of South Carolina, the line coding process is more involved. In most map projects, once the lines are coded, a new empty-polygon feature class is brought in and features are constructed with the topology toolbar, which builds the polygon file.

At this point, the accuracy and precision of geologic contacts is evaluated. Because the coastal environment is constantly changing and the topographic base maps are typically 20 to 30 years old, there are significant changes to the base map that affect contact-line accuracy. These changes can be natural or manmade. Natural changes are common on the active beach front through the addition and subtraction of deposits. Manmade changes are brought about by the intense development this area has experienced over the last 20 years. According to a recent U.S. Census Bureau report, Beaufort County had an estimated population growth rate of 11.2 percent for the years 2000-2004. The State average for the same period was about 4.3 percent.

Step 4: Editing the Geology

During the ACE Basin geologic mapping project, available aerial photography was used to improve the accuracy of the geologic information delineated by the geologist on field maps. Interpretation of aerial photos was used to more precisely delimit geological features in the study area. Photos were loaded into the ArcMap project, and edits were made with the editing, snapping, and topology tools. The precision and smoothness of lines were corrected up to a scale of 1:5000. Items most often corrected included: moved earth or artificial fill; water bodies; islands; and salt-marsh features. This area has a large amount of new development, such as golf courses, residential areas, and roadways that has altered the geologic landscape. Even the most recent air photos (2-3 years old) can be out of date for some areas.

Edge matching adjacent quadrangles was also a priority. Usually the cartographer completed the editing process. When the geologic contacts were not clear a marker was placed, and the geologist would carefully review and describe what edits needed to be made in order to ensure edge matching. The examples shown as Figures 1, 2, and 3 are common types of edits that were made.

Step 5: Final Validation and Coding Polygons

After corrections are made to each quadrangle, a final validation is performed. At this point, an empty polygon feature class is brought in, all lines in the quadrangle are selected, and polygon features are constructed with the topology toolbar. Then the polygons are coded and attributes are assigned. In the ACE Basin and Hilton Head areas there are 40 geologic units across 7 terraces.

Step 6: Creating Cross Sections

Portraying accurate cross sections digitally has been a challenge for the digital-mapping group. Often, the final cross sections do not accurately match the map units or ground distances. Additional hours of corrections, or even construction of a new cross section, were often the only answer. Accordingly, some aspects of the process have been adjusted. Most importantly, the construction of a digital cross section is delayed until the digital map layers are completely edited. Using the following procedure has greatly reduced errors.

1. Intersect the cross-section line with the final geologic-map polygon layer by using ArcMap’s Toolbox. Choose to keep attributes from polygon layer and symbolize the new line with the value field for map units.
2. Import DLG hypsography (contour) data into the project and overlay the new cross section line.
3. Label the contour elevations and plot the results with the horizontal scale at which the cross section will be drawn. The geologist then uses the plot as a guide under the graph paper when drawing the cross section.
4. Digitize, build, and code the resulting cross section. It was found that distance errors and missed map units were greatly reduced when cross sections were produced by this method.

Step 7: Final Cartography

Final cartography is accomplished by using Adobe Illustrator software with a map production plug-in (MAPublisher 6). The plug-in imports geographically referenced data into a graphics environment. An import tool, along with style sheets and symbol sets, allows map products to be easily updated. Templates are used for most of the non-map features of the layout, thus saving more time.

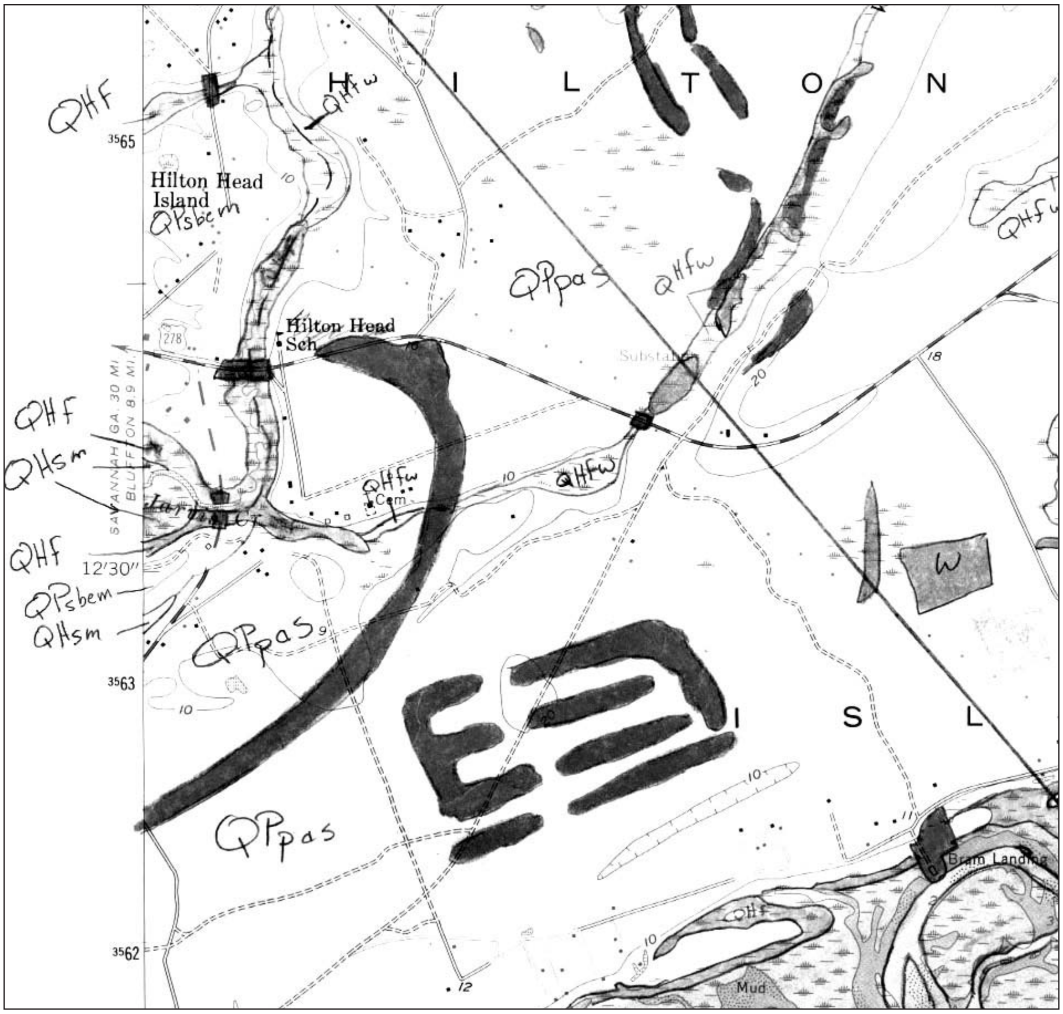


Figure 1. The geologist's field map, drafted onto a USGS topographic quadrangle base. This map is then scanned and digitized.

FUTURE PLANS

The South Carolina Geological Survey, upon completion of this digital compilation and metadata, intends to release the geology layers through the South Carolina Department of Natural Resources, GIS Clearinghouse. Geographic data in the Clearinghouse are organized in 7.5-minute quadrangles. This 23-quadrangle compilation represents the first set of digital data layers made available by the South Carolina Geological Survey.

ACKNOWLEDGMENTS

W. R. Doar, III (SCGS geologist) mapped the ACE Basin and Hilton Head areas, assisted by drillers Joe Koch and Gary Taylor. Malynn Drescher assisted with the poster design.

SOFTWARE CITED

R2V – Able Software, http://www.ablesw.com
ArcGIS – ESRI, http://www.esri.com/
Adobe Illustrator – Adobe Systems, http://www.adobe.com/
MaPublisher – Avenza, http://www.avenza.com/

REFERENCES

Doar, W.R., III, and Hudson, E.E., 2004, Geology of the ACE Basin, Beaufort, Charleston, and Colleton Counties, South Carolina: South Carolina Geological Survey Open-File Poster 150, 1:62500, accessed at http://www.dnr.state.sc.us/geology/publications.htm#ofp.

Figure 2. Geologic contact lines digitized from the field map, before corrections were made.

Figure 3. Geologic contact lines after corrections were made by using available aerial photography. In this example, dramatic changes were made to a large area of moved earth created by a road interchange on Hilton Head Island. In addition, locations of water bodies and golf-course fairways (moved earth) were significantly altered from the original field map. All of the digitizing and aerial photography interpretation was done by a cartographer and later reviewed and accepted by the project geologist.

Creating and Managing Digital Geologic Cross Sections within ArcGIS

By Evan Thoms

U.S. Geological Survey
4200 University Ave.
Anchorage, AK 99508
Telephone: (907) 786-74096
Fax: (907) 786-7401
e-mail: ethoms@usgs.gov

INTRODUCTION

While geologic maps are now almost always created within a geographic information system (GIS), geologic cross sections still are commonly portrayed and archived as simple un-attributed graphics in either paper form or digital files. Yet, there are several advantages to creating cross sections within the same GIS as the map. Features can be richly attributed and symbolized with the same symbology as in the map, the scale of the cross section can be more closely controlled than in a graphics program, and, once created, it is possible to project the cross section back into the coordinate system of the map with elevation values so that it could be used for 3D analysis or viewing. Furthermore, with the introduction of the Environmental Systems Research Institute (ESRI) geodatabase format, it is possible to store the map and the cross section within the same data container, which simplifies distribution and archiving.

In this paper, I describe a Visual Basic extension to ESRI's ArcGIS program that takes relevant information from a geologic map—the digital elevation model (DEM), geologic polygon layer, structural point measurements, and borehole information (Figure 1)—and builds a template from which the user can begin to digitize a geologic cross section. A surface profile is interpolated from the DEM, apparent dip is calculated for structural measurements, boreholes are shown as stick columns, and the amount of vertical exaggeration can be controlled. The output data are stored within either ESRI shapefiles or ESRI geodatabase feature classes, depending on the format of the input layers.

THE CONCEPT

A key concept employed within this extension is that of "linear referencing", a concept likely to be unfamiliar to most GIS users who have a geology backgrounde, as it was primarily designed to manage transportation systems. There are two intrinsic components to linear referencing. The first is a "route", that is, a line feature that has been "measured" so that every vertex stores not only the X and Y attributes of its position but also a value, M, related to its distance from the starting point of the line. The second component is a table of point or line "events" that occur along that route. In a transportation-based GIS, a table of point events might represent the locations of bus stops along a bus route, and a table of line events might represent sections of a road network that are under construction. When the table of events is referenced to the route, a new feature class is created without requiring the user to know or calculate the XY coordinates of the new features.

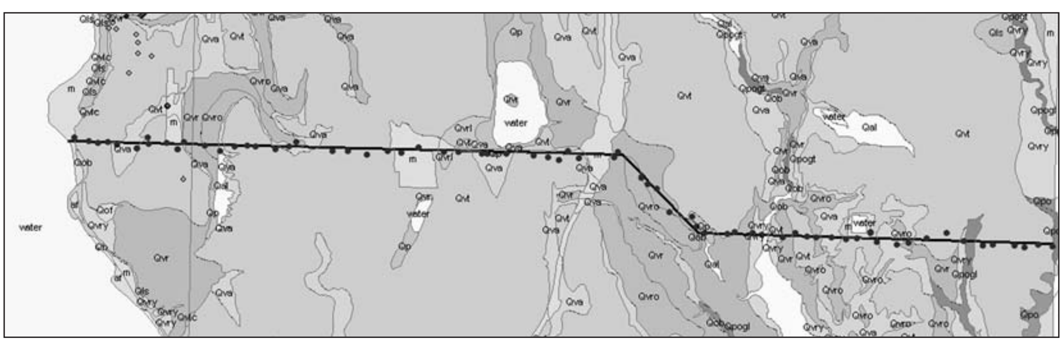


Figure 1. Cross section line in map view, above a geology polygon layer with boreholes.

Linear referencing is used in the cross section extension, in three ways. First, the line of cross section on the map is assigned elevation values based on the underlying DEM and then converted into a route so that every vertex stores an X, Y, Z, and M attribute (Figure 2). Next, the extension builds a new line by setting the X value equal to the M and setting the Y value equal to the Z at every vertex (Figure 3). This step "flips" the cross section line on to its side so that it is represented in the vertical plane. The resulting surface profile is drawn so that the X coordinate at the starting point equals 0 and the X coordinate at the ending point equals the length of the line in the units of the map view. It is also a route retaining the M values from the previous step against which events can be referenced.

In this second step, linear referencing is used to build tables of line and point events that occur along the cross section route in map view. Line events represent the locations of the intersection between the line of cross section and the boundaries of polygons in an underlying polygon layer, which the geologist interprets as geologic contacts. Point events represent the nearest projected locations of points, either locations of structural measurements or boreholes, along the cross section route within a specified buffer distance. The resulting line events table carries all of the fields from the original polygon layer with the addition of three: a route key field, a from-M field, and a to-M field (Figure 4). The route key field carries a value equal to the user-chosen name of the cross section route so that the linear referencing engine can place the events along the correct route. The from-M and to-M fields carry values that specify the beginning and ending distances along the route of each event. Once the table has been generated from the map view, it is referenced to the surface profile route (Figure 3) in the cross section. The resulting feature class consists of a collection of contiguous lines, each of which represents the intersection of a geologic unit with the topographic surface (Figure 5). Lines can be symbolized by attribute, or with a short perpendicular tick at the end (or beginning) of each line to mark the unit contacts. Point events tables, similarly, retain all of the fields from the original layers as well as an additional route key and M field. Structural measurements are also run through a routine that converts the true dip into an apparent dip.

The final use of linear referencing is to build and attribute borehole "stick" columns in the cross section view. This process begins with building a line for every borehole, based on the elevation of the borehole collar and the final depth of the hole. Each line representing a borehole is then measured and turned into a route where some unique value, either a numerical id or text name

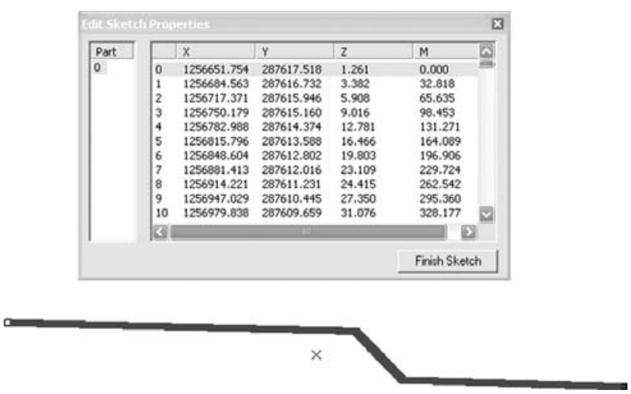


Figure 2. The cross section line after Z values have been interpolated from the DEM and after its conversion to a route. Note the first, or starting, vertex has been selected.

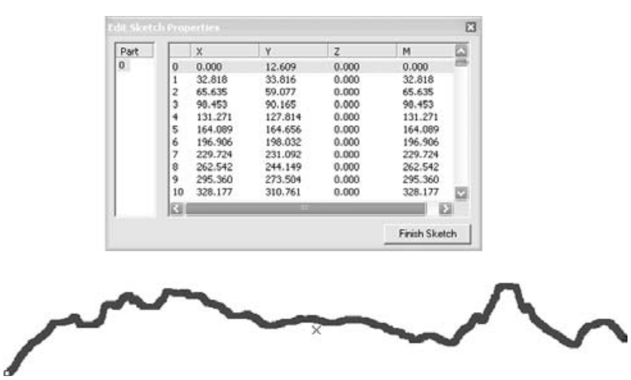


Figure 3. The new surface profile route (with 10 times vertical exaggeration) made by setting X equal to M, Y equal to Z (x10), and keeping the M value from the step before.

OID	RKEY	FROM_M	TO_M	Unit
18	A	0	27.6430660695	water
38	A	27.6430660695	259.165140961	m
39	A	259.165101999	1586.5049399	Qvt
17	A	1586.50496003	2583.17176377	Qva
40	A	2583.17179997	4077.5247778	Qvt
16	A	4077.52481401	4183.18903187	Qva
19	A	4183.18903187	4818.33637439	Qvr
32	A	4818.33637439	5370.2783969	Qvro
11	A	5370.2783969	5565.34711963	Qva
41	A	5565.34712477	7311.92061377	Qvt
8	A	7311.92057525	7472.94093269	Qva
42	A	7472.94083659	7563.41336139	Qvt
9	A	7563.41333755	7954.44685395	Qva
20	A	7954.44685395	8296.11017021	Qvr
10	A	8296.11017021	8928.19080091	Qva
43	A	8928.19080231	11626.4539526	Qvt
13	A	11626.4539037	14172.7463307	m
44	A	14172.746794	14267.951052	Qvt
14	A	14267.9511236	14311.5574283	m
15	A	14311.5574283	19373.306524	m

Figure 4. A portion of the line events table, which is made by locating the geologic unit polygons along the cross section route.

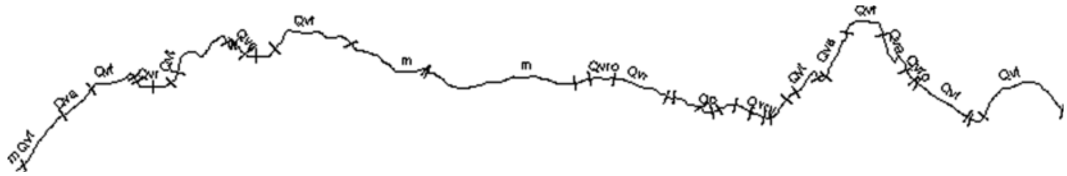


Figure 5. The resulting feature class of lines representing the contacts between the geologic unit polygons.

field, is stored as the route key. At this point, a non-spatial table of borehole intervals or layers is required. This type of table is often already part of a borehole database; existing in a one-to-many relationship with another table or feature dataset of borehole locations. The intervals table is used in the same way that the extension-generated events tables (Figure 4) are used. It must contain a route key field of unique values (usually a borehole id or name field) and two fields that describe the top and bottom depth of each interval. Once the intervals events table is referenced against the borehole routes, each borehole column consists of a collection of contiguous lines representing the intervals. The intervals can be symbolized and labeled according to any attribute in the table, to aid in interpretation of the subsurface geology.

USING THE EXTENSION

The Geologic Cross Section Tool extension is available from me (see email address in title) as a dynamic link library (DLL) file. It has been compiled for ArcGIS 9. Once the DLL has been registered with the host computer, the tool can be added to any ArcMap toolbar through the Customize interface. It appears as a button with a simple "faulted strata" icon.

To use the tool, first add appropriate layers to an ArcMap data frame. These are: a polygon layer of geology units, a line layer containing one or more lines of cross section, a digital elevation model or similar raster with elevation values, an optional layer showing structural measurement points, an optional layer showing borehole locations, and a related table of borehole intervals.

The following are some requirements for the formats of the layers:

1. Before being added to the map, all layers should be projected to the same coordinate system.
2. All X, Y, and Z values among all layers must be in the same units. In particular, check that all elevation and depth values in the boreholes layer and the borehole intervals table agree with the DEM.
3. The cross section lines attribute table must have a name field with which to uniquely name each cross section. Avoid using a single quote character in the name, that is, don't use AA' or BB'.
4. The structural measurements attribute table must have strike, or equivalent, and dip, or equivalent, fields. Convert any quadrant strike values (e.g., N30E) into azimuth values between 0 and 360.
5. The borehole locations attribute table must have a field that carries the name of each borehole and a field that carries the total depth of the hole. It is best if there is also a field carrying the elevation of the top of the hole, but the user has the option of letting the extension calculate an elevation value from the DEM.
6. The table of borehole intervals must have a field that carries the name or id of the borehole that each interval is associated with (values must exactly match the name or id values in the borehole locations attribute table), a field that carries the depths to the top of each interval, and a similar field of depths to the bottom of each interval.
7. It may be necessary to make partial copies of the borehole locations layer and the related borehole intervals table that only include the boreholes lying very near, if not completely within, the desired buffered width of the cross section line. The extension tends to commit fatal errors when trying to process large tables.

Once the map and tables are ready, the user selects a line of cross section (the selection does not have to occur within an editing session) and clicks on the tool button. A two-page form is displayed, on which the user selects the appropriate values from a series of picklists or enters information (Figure 6). The user must supply a text string, which will be used as the prefix to the feature class names that will be created. Lastly, the user can choose to have the new feature classes added to a new data frame in the map document. The new feature classes are created in the same format and placed in the same workspace (a geodatabase for geodatabase feature classes or a directory for shapefile feature classes) as the layers on which they are based.

Here are suggestions for managing and editing the cross section layers in the cross section view:

1. Feature classes generated by the extension are assigned an "unknown" coordinate system, thus,

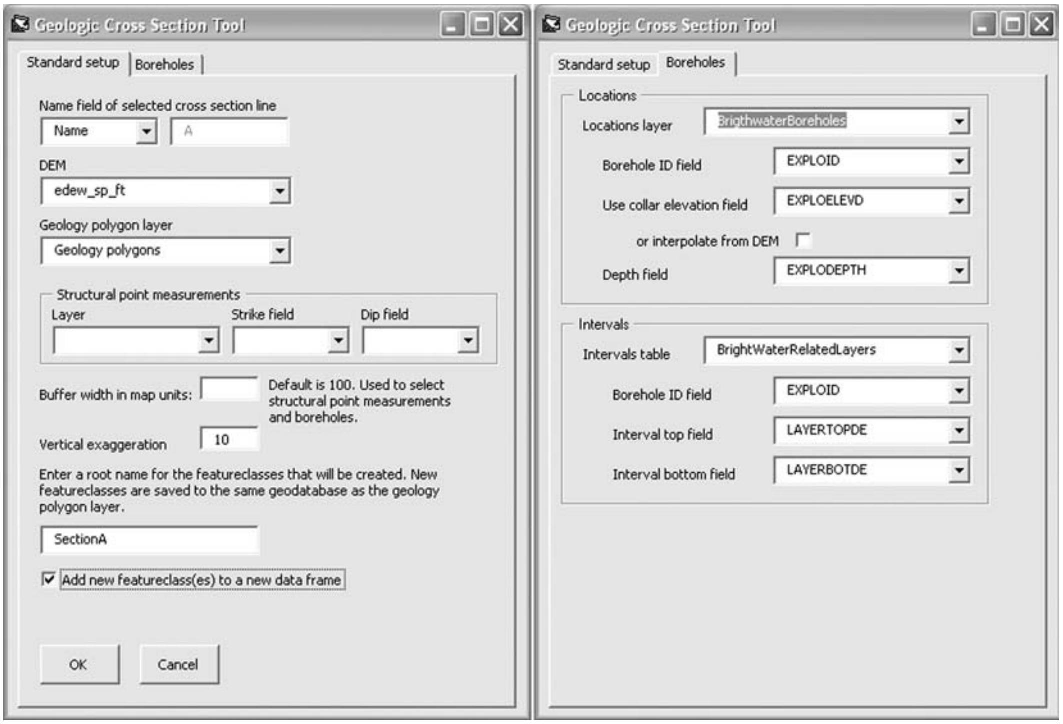


Figure 6. Pages 1 and 2 of the Geologic Cross Section Tool form.

the user cannot explicitly set a scale until the 'Map units' are set through the data frame properties dialog. Set these to be the same as in the map view (usually feet or meters).

2. Before creating geology polygons, create a feature class by importing the attribute table schema from the geology polygons feature class being used in the map view, and import the symbology.
3. Establish topology between the contacts and geology polygons layers (with either a map topology or a geodatabase topology relationship class) and use the relationship to build the polygons from the contacts.
4. For viewing in layout view, create a custom data frame grid with an origin of 0,0 and with intervals appropriate to the length and vertical exaggeration of the cross section. Unfortunately, you must always divide the Y value by the factor of vertical exaggeration to obtain the correct elevation. I know of no way to force ArcMap to make that calculation for grid labels.
5. Ultimately, you will probably want to store the cross section layers in the same geodatabase as the map layers. In that case, store them in a feature dataset with an unknown coordinate system.

CONCLUSION

I wrote this extension with two primary goals in mind. First, I wanted to be able to build geologic cross sections in ArcGIS using borehole data without having to purchase an expensive plug-in or a separate software package. In this respect, I have mostly succeeded (Figure 7). The greatest inflexibility is that data in the output layers are not dynamically linked to the input layers or selections, that is, many output files are often generated before an acceptable set is decided upon. As a result, the

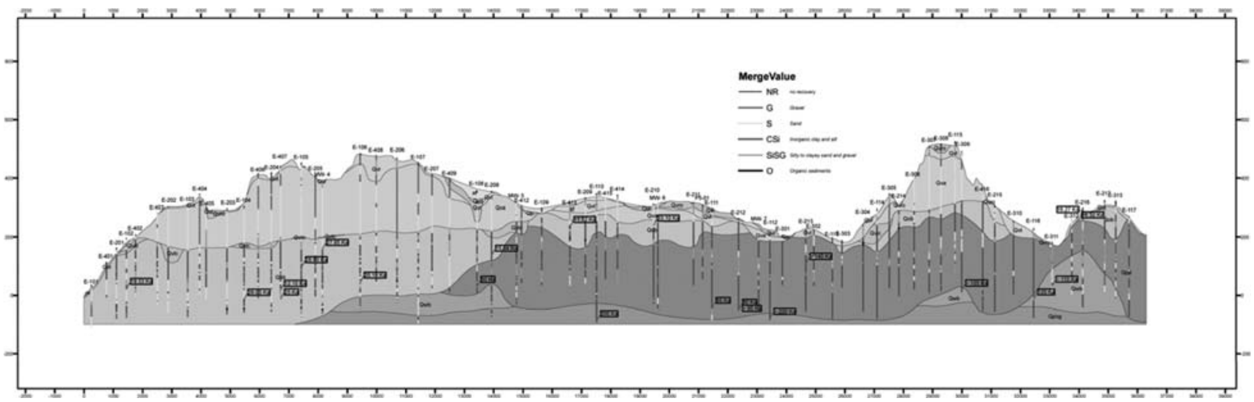


Figure 7. Nearly final version of interpreted geologic cross section showing subsurface geology and boreholes (10x VE). Borehole intervals have been symbolized by classes defined by groups of unified soil classification system codes.

user must exercise good file management. But the primary advantages are that the symbology of the layers in the map view can be used for the cross section layers and, at least if the geodatabase format is being used, the cross section layers can accompany the map view layers in the same data container.

My second goal was to devise a method for taking interpreted cross section data from the cross section view and converting them into 3D features that would be viewable in a 3D viewing program, such as ArcScene or Earthvision. The extension does not yet do this, but I describe the presumed method here in the event that another developer will wish to implement it. Again, linear referencing will be the key. The process would be to take the X coordinate of every vertex or point in the cross section view and treat that value as an M value which can be referenced to the cross section route in the map view. The X and Y at that measure of M could then be used to create a new vertex or point where the Z value is equal to the Y coordinate from the cross section view. At that point, the process will have come full circle and the new features will have been elevated from simple graphics to attributed spatial features that can be viewed and analyzed alongside other subsurface data.

3D Modelling Techniques for Geological and Environmental Visualisation and Analysis

By Mauro De Donatis, Giuliano Gallerini, and Sara Susini

LINEE – Laboratory of Information-technology for Earth and Environmental Sciences
Universita' degli Studi di Urbino "Carlo Bo"
Campus Scientifico
Loc. Crocicchia
61029 Urbino – Italy
Telephone and FAX: int.+ 39. 0722.304295
e-mail: dedonatis@uniurb.it

INTRODUCTION

Three dimensional (3D) geological models represent an advanced synthesis of different kinds of data (geological and geophysical, surface and subsurface data). This is useful to obtain a geological vision more realistic than a 2D map, and to perform complex analyses and computations on the geological framework.

The accumulation of geological information in digital form has introduced the possibility of applying geographical information system technology to the field geology. To achieve the benefits in information management and in data analysis and interpretation, however, it is necessary to develop spatial models that are specifically designed for working in three dimensions. An overview of examples for various geological settings and scale were elaborated with different software and the analysis technique described below.

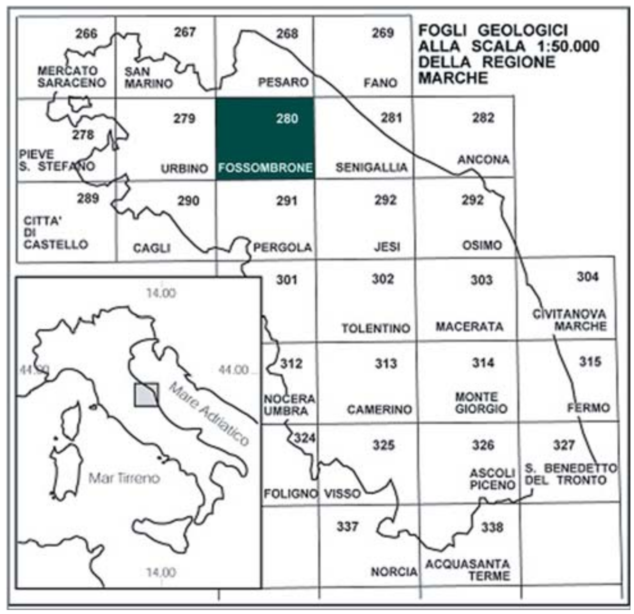


Figure 1. Location Map of Italian Geological Survey—Sheet 280-Fossombrone.

3D GEOLOGICAL MAP FOR THE ITALIAN GEOLOGICAL SURVEY

The goal of the national project on 3D Geological Cartography of Italy, carried out by LINEE (Laboratory of Information-technology on Earth and Environment) and the Italian Geological Survey with collaboration of ENI and Midland Valley Exploration, is to study, represent and better understand the relationship among geological features (stratigraphy, structures, geomorphology, etc.) by using new tools for 3D geological modelling (*Slatt et al., 1996; De Donatis, 2001*).

The selected area is contained within Sheet 280 – Fossombrone (SGN-Italian Geological Survey , scale of 1:50,000 – Figure 1) which was surveyed by the SGN with the collaboration of the Universities of Urbino and Siena, for the new Geological Map of Italy Project (CARG CARtografia Geologica). The in-depth knowledge of the regional and local geology combined with the availability of subsurface (well and seismic by ENI-Agip) data makes the area ideal for defining a suitable methodology for creating a 3D geological model for other areas of Italy (De Donatis et al. 2002).

The field geological map (Figure 2) and the CARGtype database (Figure 3) were integrated by geophysical data (seismic profiles calibrated with well logs) in order to build a series of parallel cross-sections. For the topographic surface, a DEM was developed from the new detailed Technical Map of Regione Marche, at the scale of 1:10,000. The later was used as the base upon which the geological boundaries were draped, and the geologic units were colored, to show the geology in 2.5 dimensions (Figure 4).

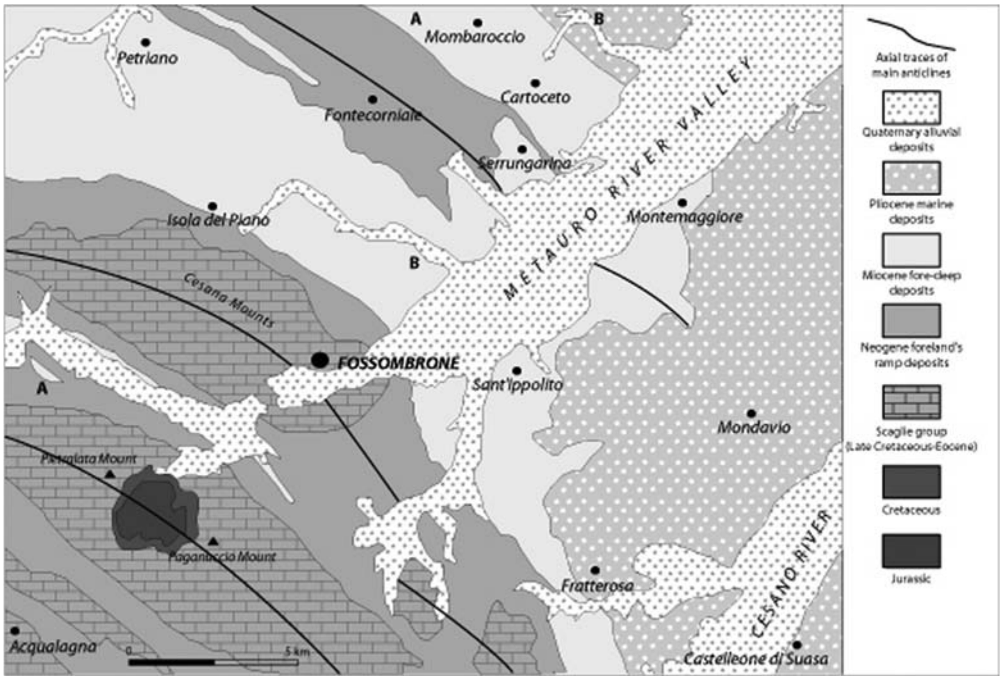


Figure 2. Geological sketch map of Sheet 280.

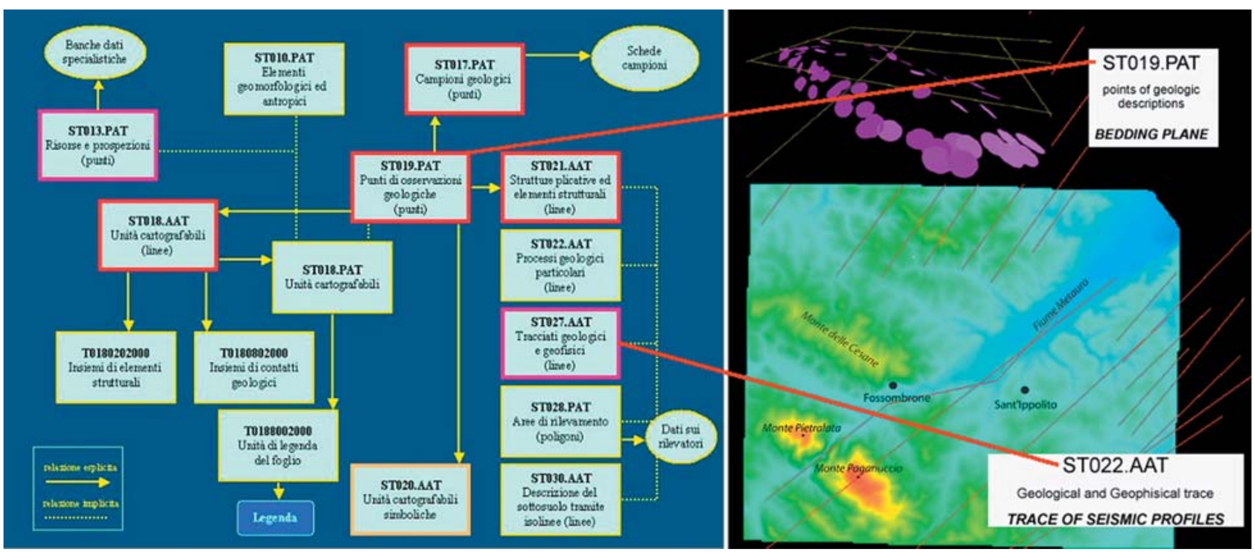


Figure 3. The CARG (*CARtografia Geologica*: Geological Cartography) database.

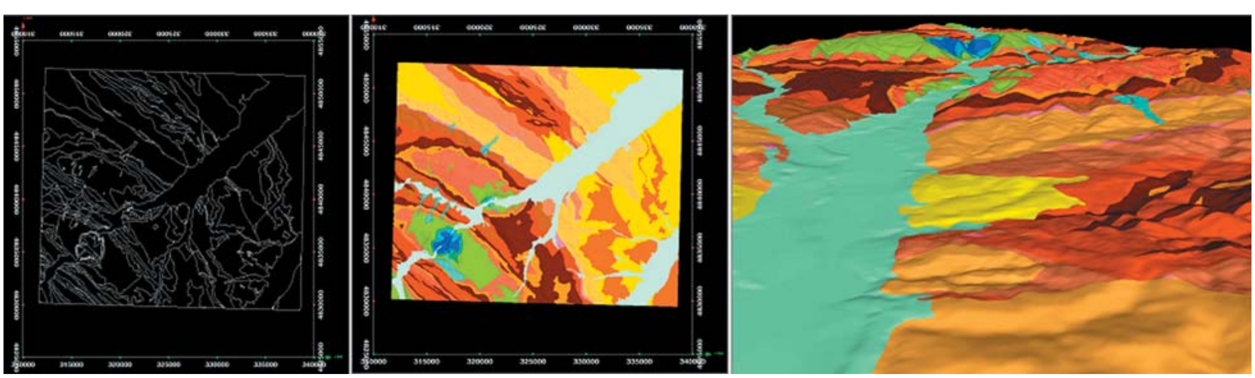


Figure 4. Combining geology and the DEM, to show geology in 2.5 dimensions.

The interpreted seismic profiles (Figure 5a) were time-depth converted and, integrating field data, a number of georeferenced geological cross-sections were built using 2DMove (Figure 5b) and imported into the 3DMove environment (Figure 5c).

By correlating lines between cross-section, the 3D surface of each geological horizon was developed. These were edited many times, in order to ensure a coherent, integrated data set. The result of this work is the geological model shown in Figure 6. Figure 7 shows details of the top of the Furlo Gouge structure, and its 3D form (this feature is located in the western part of the map area).

3D GIS: ALLUVIAL DEPOSITS OF THE FOGLIA RIVER, MAPPED FOR THE MUNICIPALITY OF PESARO (MARCHE-ITALY)

The aim of this study was to build a 3D geological model using subsurface data, using more then 700 logs. Every log was digitised and interpreted. A geological database was built in a GIS environment. After we imported the log data into Voxel Analyst software, we could also compute data by the means of three dimensional geostatistical methods.

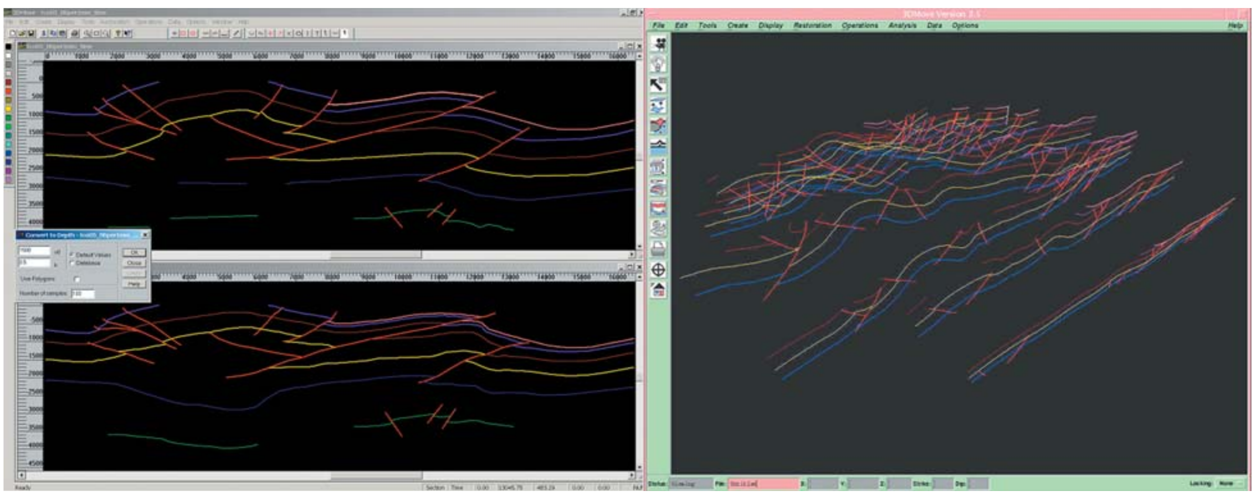


Figure 5. Construction of geological cross-sections (left windows) and importing into 3D environment (right window).

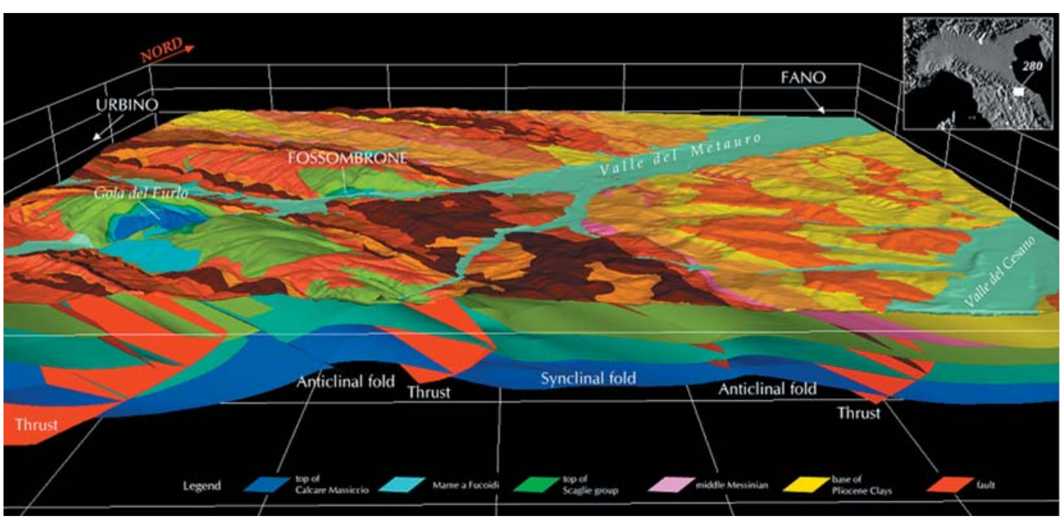


Figure 6. The 3D geological model of Sheet 280-Fossombrone in a perspective view from the south.

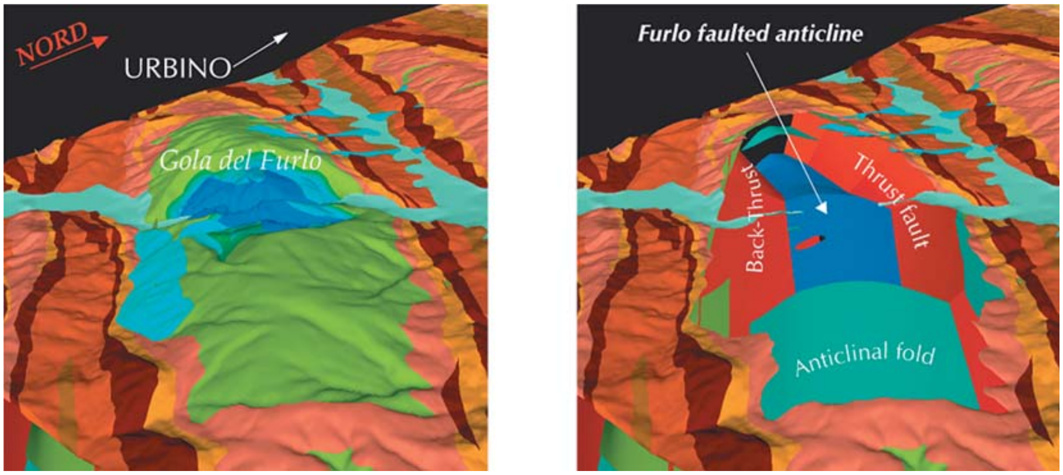


Figure 7. Detail of the structure at the Furlo Gouge, at land surface (left diagram) and at depth (right diagram), in the western part of the map area, near Urbino.

We accomplished the following:

- built a 3D geological model of a sector of alluvial plain of the Foglia river around the Pesaro municipality (Figures 8a and b),
- compared the model to older, two-dimensional models, and
- built a 3D geo-environmental model of a hydrocarbon-contaminated site, and a volumetric evaluation of the pollution plume (Figure 8c).

THE 3D HYDROGEOLOGICAL MODEL OF A SECTOR OF THE PO PLAIN, MAPPED FOR THE REGIONAL GEOLOGICAL SURVEY OF LOMBARDIA REGION (ITALY)

The 3D geological modelling has also been applied for hydrogeological characterization to an area in Lombardia, around the Po River Plain. Data, coming from a previous subsurface map publication by Regione Lombardia in collaboration with ENI, were georeferenced and all the contour lines were assigned an altitude (z) value. We imported the data into two different software (3DMove by Midland Valley Exploration Ltd and Voxel Analyst by Intergraph) in order to develop horizon surfaces (Figure 9a and b) and the volumes between the surfaces. This model allowed us to better assess volumes and to build flux models by attributing parameters such porosity, permeability, etc.

THE GEODYNAMICS OF SOUTH SCOTIA RIDGE, MAPPED FOR THE ITALIAN ANCTARTIC RESEARCH PROGRAM.

The main purpose of this three-dimensional model was to better understand the geodynamics of the South Scotia Ridge (SSR), a submerged structural high representing the eastern continuation of the Antarctic Peninsula at sea, and one of the major transcurrent plate boundaries on Earth. The SSR runs approximately E-W for about 500 km, separating the oceanic Scotia Plate

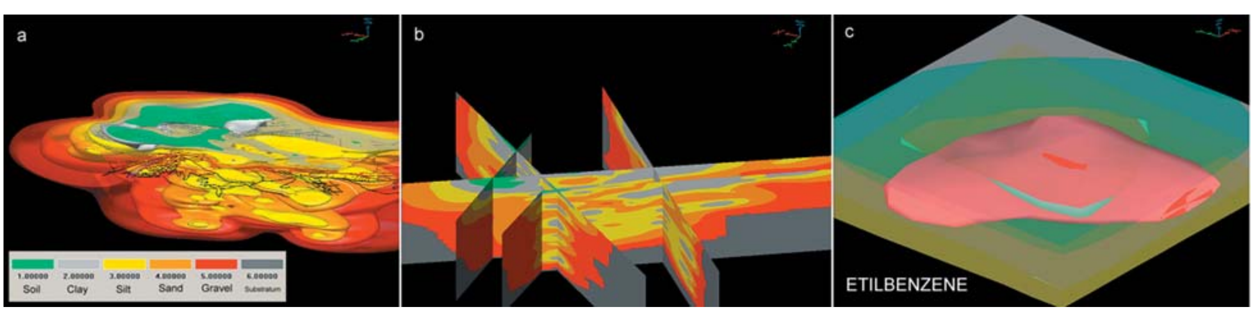


Figure 8. 3D geostatistical model of the lithological bodies with different grain size (see legend in 8a) of the Foglia river valley (8a); some cross sections (8b); and a hydrocarbon pollution plume (8c) in the porous media (sand and gravel deposits).

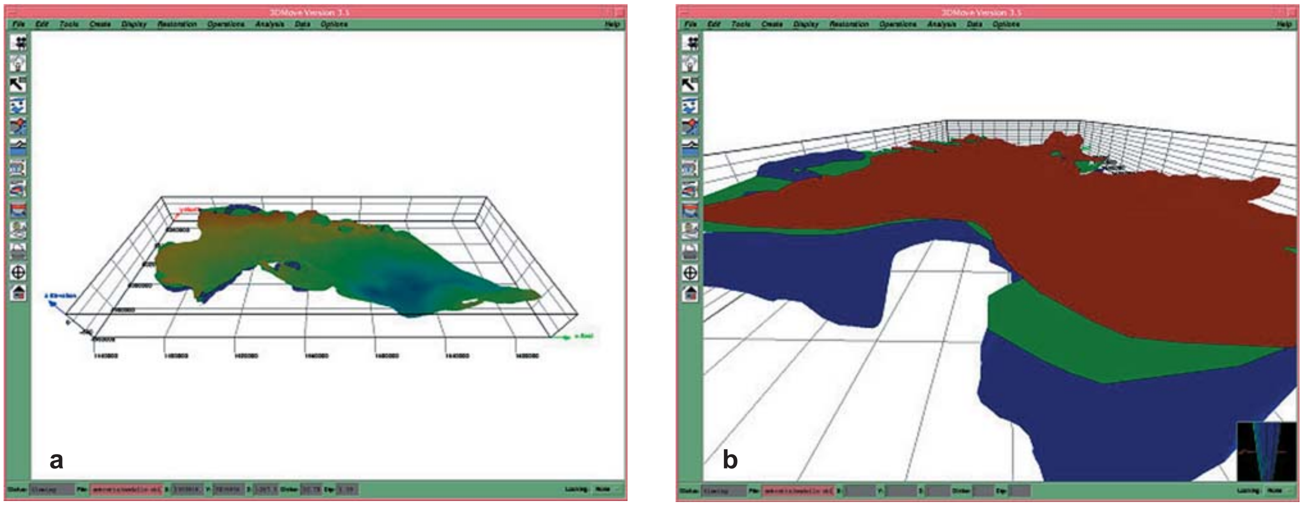


Figure 9. 3D geological model showing the horizons of sequence-tops bounding the three principal aquifer groups in the Lombard sector of the Po Plain. Total view of the model from South (9a), and view from East (9b).

from the Antarctica Plates, and mainly composed of fragments of continental crust. The seismic lines that we used (Italian campaigns IT91 and IT95 onboard the R/V OGS-Explora) were collected mainly orthogonal to the geological structure.

The interpreted seismic lines were digitised and acquired in a format useful for processing with 2D-Move and 3D-Move software packages. After georeferencing and converting from travel-time to depth, geologic horizons recognised from the network of seismic lines were correlated in three dimensions with boreholes, to create the 3D model (Figure 10).

FINAL REMARKS

The ability to collect, visualize and analyse spatially-accurate surface and subsurface datasets in three-dimensions gives us new opportunities to capture and interpret three-dimensional geological architectures. This method brings new research opportunities allowing:

- New representations and new ideas on the concept of Geological maps,
- Spatial analysis of geological architectures and processes at a range of scales, and
- Evaluation of uncertainty in geological data and interpretations.

REFERENCES

De Donatis, M., 2001, Three-dimensional visualization of the Neogene structures of an external sector of the Northern Apennines, Italy: AAPG Bulletin, v. 85, no. 3, p. 419-431.

De Donatis, M., Jones, S., Pantaloni, M., Bonora, M., Borraccini, F., Gallerini G., and D'Ambrogi, C., 2002, A National Project on Three Dimensional Geology of Italy: 3D model of Monti della Cesana from Sheet 280 – Fossombrone: Episodes, March 2002, p. 29-32.

Slatt, R.M., Thomasson, M.R., Romig, P.R., Pasternack, E.S., Boulanger, A., Anderson, R.N., and Nelson, H.R., 1996, Visualisation Technology for the Oil and Gas Industry: Today and Tomorrow: AAPG Bulletin, v. 80, no. 4, p. 453-459.

See also http://www.uniurb.it/ISDA/Linee/linee.html.

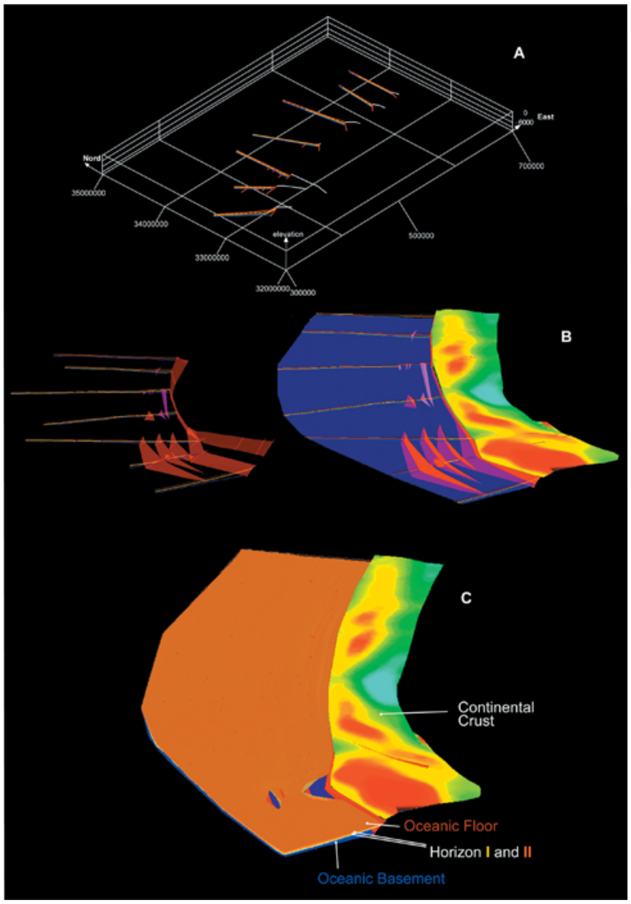


Figure 10. Construction phases of 3D model of South Scotia Ridge. Geological sections geo-referenced in the 3D virtual space (10a). Creation of surfaces representing geological and structural elements (10b). Complete 3D model showing oceanic basement, faults, horizons I and II, Oceanic floor and DEM of the continental crust (10c).

New Geologic Mapping and Geologic Database for the Urbanized Puget Lowland, Western Washington State, USA

By Derek B. Booth, Kathy Goetz Troost, Scott A. Shimel,
Michael A. O'Neal, and Aaron P. Wisher

Pacific Northwest Center for Geologic Mapping Studies
Department of Earth and Space Sciences
University of Washington, Box 351310
Seattle, WA 98195
Telephone: (206) 543-7923
e-mail: dbooth@u.washington.edu

INTRODUCTION

The cities of the Puget Lowland region (Figure 1) have been built atop a complex sequence of deposits with widely varying engineering strengths and an irregular bedrock surface at depth. They lie in one of the most seismically active regions of North America, with moderate earthquakes virtually assured during the lifetime of any structure, most recently the Nisqually earthquake of 2001. Many contain steep hillslopes that are marginally stable in wet weather; because of shallow water tables, underlying sandy deposits are particularly susceptible to liquefaction during strong ground shaking. As the center of both population and economic activity of the Pacific Northwest, geologic events of even moderate intensity can and do result in substantial human and economic losses. Seattle was recognized by Federal Emergency Management Agency (FEMA) in 2000 as the city with the seventh-highest annualized earthquake loss in the United States, and the highest outside of California. At the state level, Washington has the second highest risk (2nd only to California) of suffering economic loss due to earthquakes.

Geoscientists and engineers recognize that the Quaternary deposits of the Puget Lowland are primary determinants of the magnitude and location of strong ground shaking. Knowledge of the geometry and variability of these deposits—the *geologic framework*—is critical to the support of ongoing seismic evaluations across this region, which will ultimately determine the necessary measures, and the cost, of adequate preparation and hazard mitigation. Such a framework comprises a detailed representation of the sequence, chronology, structural history, distribution, lateral lithologic variability, and geotechnical properties (such as strength and permeability) of geological materials.

The Pacific Northwest Center for Geologic Mapping Studies (http://geomapnw.ess.washington.edu) is a collaborative effort to develop new data and greater under-

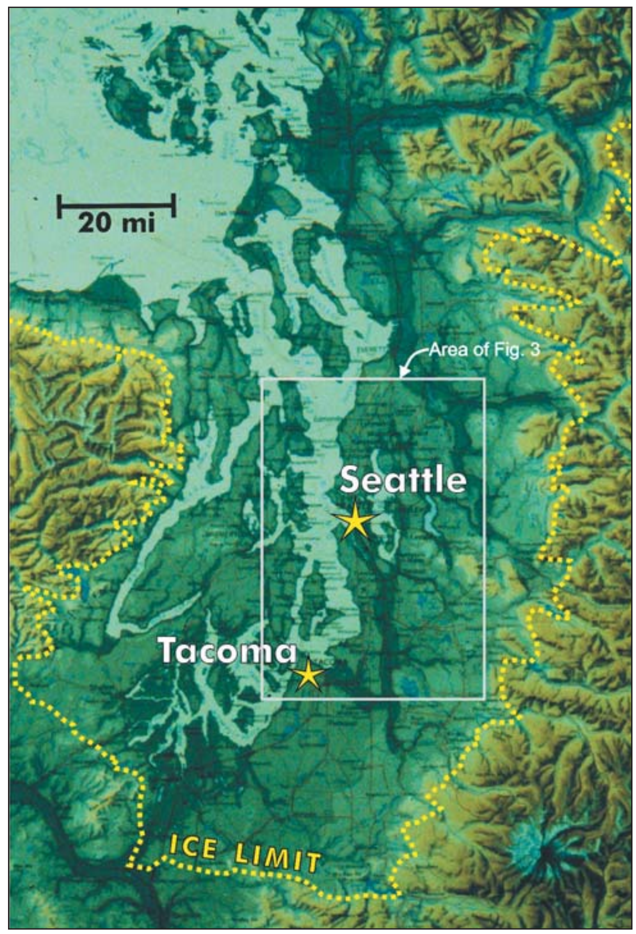


Figure 1. Location map of the Puget Lowland region, showing the southern extent of the Puget lobe of the Cordilleran ice sheet about 16,000 calendar years ago (dashed line; Booth and others, 2004).

standing of the geology of the central Puget Lowland. The project was initiated in 1998 through collaboration with the U.S. Geological Survey, the University of Washington, and the City of Seattle, to provide state-of-the-art geologic data to support geologic hazard mitigation in the City. Since that beginning, its scope has broadened to include other geographic areas and a broadened range of research topics. The project goals are to acquire existing geologic data and create new geologic information; to conduct geologic research and produce new geologic maps; and to support the wide variety of additional research, hazard assessments, and land-use applications of other scientists, organizations, and agencies throughout the region.

Our efforts to improve the regional understanding of western Washington’s geologic framework consist of several interrelated elements:

- *Scientific studies of the regional geologic framework,* including determinations of the age and identification of geologic materials to help understand the history of crustal deformation and develop standardized nomenclature for all geologists working in the central Puget Lowland;
- *A subsurface database of existing geologic data,* built to include new geographic areas and accept new data fields as the needs arise;
- *Geologic maps across the central Puget Lowland,* replacing preliminary documents that are locally almost 50 years old and establishing a new standard of consistency and geologic mapping for the region;
- *Public access to geologic data* via web-based interfaces for both subsurface geologic data and geologic maps; and
- *Outreach to varied audiences,* particularly the technical and planning community, and research scientists.

SUBSURFACE GEOLOGIC DATABASE FOR THE GREATER SEATTLE AREA

Geologic investigations in urban areas, regardless of location, all face the same quandary—the value and potential applicability of the data are high, but the same human infrastructure that makes these data so valuable also obscures the very source of that information. Fortunately, that infrastructure also creates some of the most valuable geologic data to be found in urban areas, namely subsurface explorations. Although abundant, most of these exploration data are widely scattered and poorly organized in building and utility departments, transportation agencies, and private consulting firms. To be able to take full advantage of these data, we have developed and are continuing to populate a GIS-based relational database to efficiently store, manipulate, and display the vast amount of subsurface geologic data available for the Seattle area. Geologic data from tens of thousands of field explorations, exposures, and excavations have been entered into the database and are now accessible and available to a much wider audience than ever anticipated.

Partnerships have been formed with a number of local public agencies (such as building departments, public utilities, port authorities, transportation agencies, and natural resource departments) both to acquire the raw data from geologic and geotechnical studies and to return the populated database and GIS interface to those agencies and the public. As a result of continued partnerships over the past seven years, we have developed and streamlined processes for identifying and acquiring geologic data from a variety of sources, with our data largely obtained from public-agency, reports, permit files, and other records.

A basic three-level structure was adopted for the database to provide a common framework for all data and to allow for future expansion (Figure 2). Information about the *document* (i.e. the physical report for a property, a structure, or other type of project) that contains the geologic data and its spatial coverage are stored at the first level (in the GEOTECH_DOC table in Figure 2). Within that document, the attributes and location of subsurface *explorations*, of which there may be just one or many, and which may range from shallow test pits to deep water wells, are stored at the second level (EXPLORATION table). For each exploration point, all the related subsurface *layers* described in each exploration log are stored at the third level. Any additional layer-based information, either comments made on the original logs or subsequent geologic interpretations of the individual layers themselves, are stored at this level as well. The structure of the database and the fields were designed to accommodate geologic data from a variety of sources and formats, to create a common interface for entering and displaying data, and to support current and future scientific and engineering studies.

Data are entered through customized GIS and database interfaces. Spatial data, namely the area covered by a document and the data points representing the explorations, are entered through a GIS interface along with their associated attributes; the nonspatial data (i.e. the subsurface geologic layer data associated with a specific exploration data point, together with any comment or interpretation) are entered through customized database forms.

Guidelines have been developed to ensure that the data are entered in a uniform and consistent manner. These guidelines provide normalization of data collected from boring logs, test pits, and other exploration types that were prepared by many different consultants and agencies under a variety of classification systems and protocols. Geologic layer-entry guidelines were developed to facilitate translation from the logs to the database. Similar guidelines exist for document and exploration point entry. The guidelines define the fields, give default values, and

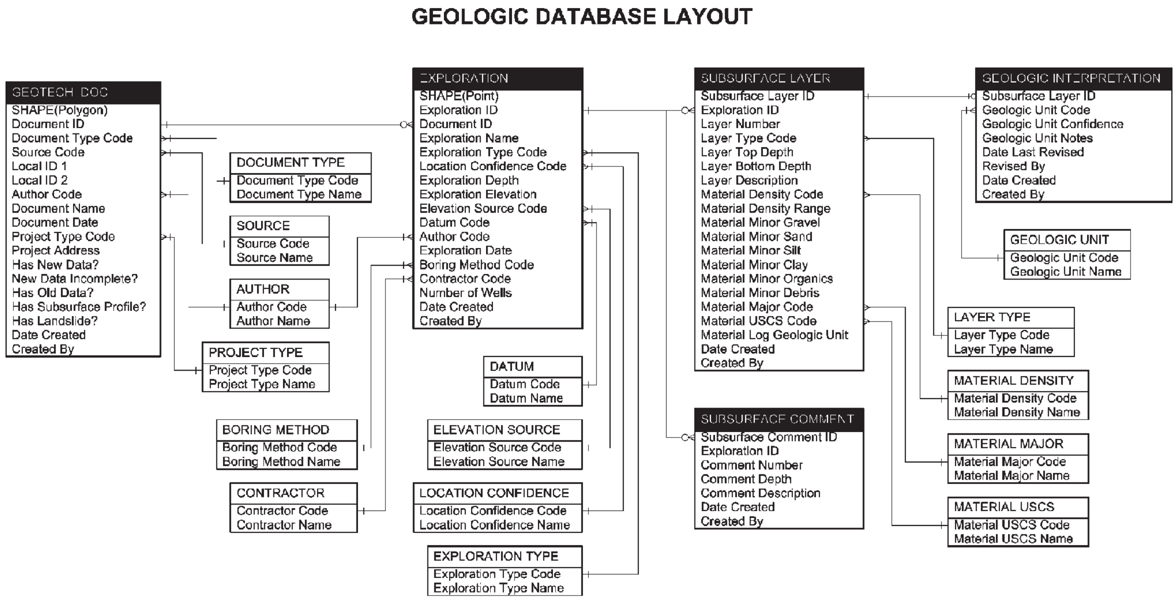


Figure 2. Three-level database structure, showing the data fields and their relationships for the spatial data (GEOTECH_DOC and EXPLORATION) and the nonspatial data (SUBSURFACE LAYER, SUBSURFACE COMMENT, and GEOLOGIC INTERPRETATION).

describe what to do if data are missing from the log.

The database contains "raw" data, in particular the verbatim transcription of the original on-site geologist's or well-driller's description of each layer. This information is then parsed manually into fields for density, major and minor materials, and the presence of organics and debris in order to facilitate future database queries. Fields are also available for geologic interpretation, the metadata on original source documents, and anticipated accuracy of point locations.

Since 1998, we have populated the main database tables with a significant amount of data:

TOTAL STUDY AREA (as of 1/06)	
Geotechnical Documents	14,251
Exploration Points	70,355
Subsurface Layers	291,101

Because there are no fixed limits on the prospective area of database coverage, we cannot estimate an ultimate magnitude of data acquisition. Mainly by increasing the geographic area, 1300-2800 documents per year have been added to the database. Within the city of Seattle, where we have been working steadily since the project's inception, we have an ongoing program to add new data as it is received by the City; nearly 200 new documents were added from there in 2005. The geographic areas covered by subsurface information are illustrated in Figure 3.

When we began the project, data were entered into the database through customized ArcView and Microsoft Access interfaces to take advantage of readily available

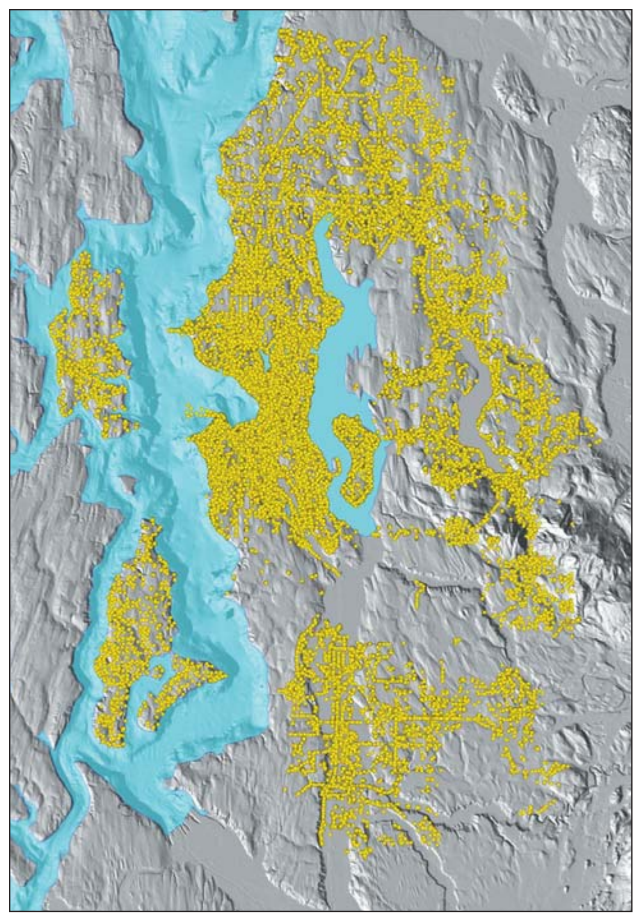


Figure 3. Database coverage currently available, as indicated by the distribution of exploration points (light-colored circles).

software and to simplify interactions between multiple (mainly municipal) users, nearly all of whom had access to these software tools but not to anything more sophisticated. The spatial data (document areas and exploration data points) were entered through an ArcView interface along with their associated attributes. Once the spatial data were recorded, nonspatial data (layer data and associated comments) were entered through customized Microsoft Access forms. The two phases of the data-entry process corresponded to the two main components of the database: the spatial data, stored in ArcView shapefile format, and the nonspatial data, stored in a Microsoft Access file. This approach was chosen to take advantage of the relational database capabilities of Microsoft Access while keeping the spatial data in a common format.

Increasing volumes of data, the desire to accommodate multiple simultaneous users, and concerns for fail-safe back-up led us to our present system, whereby the database and corresponding GIS are stored in ESRI's geodatabase format employing ArcSDE with an Oracle database backend. ArcSDE was chosen for its ability to accommodate a multiuser editing environment for spatial data using multiversioning, and for its ability to efficiently store and deliver geospatial datasets. Access to the data stored in the geodatabase is available through a number of application program interfaces (API's) so that customized applications and services can be developed on a variety of computer platforms. Full access to the data is also available to native Oracle objects such as views, functions, and stored procedures, making it possible to programmatically query and analyze the data efficiently. The previous customized tools for entering, analyzing, and viewing data were converted for use within ArcMap by using Visual Basic and object model component technology. Our municipal partners, however, have generally required conversion of data to ESRI shapefile and dBASE dbf file formats to maintain compatibility with their ArcView systems. The database and corresponding GIS are currently stored on a Linux server and are accessed by several Windows workstations through a gigabit network.

SURFICIAL GEOLOGIC MAPS

One of the primary direct applications of the subsurface geologic database has been to support the preparation of new geologic maps. To date, the area where we first began our compilation (the City of Seattle) has been completely remapped at 1:12,000 scale; a preliminary compilation is available (Troost, and others, 2005a), with its four constituent quadrangles in various stages of USGS technical review and publication (Booth and others, 2005; Troost and others, in review a, b; Booth and others, in review a). These maps represent a dramatic increase in both the detail and quality of geologic information for the city relative to the only previously available map (Waldron and others,1962, scale 1:31,680; see Figure 4).

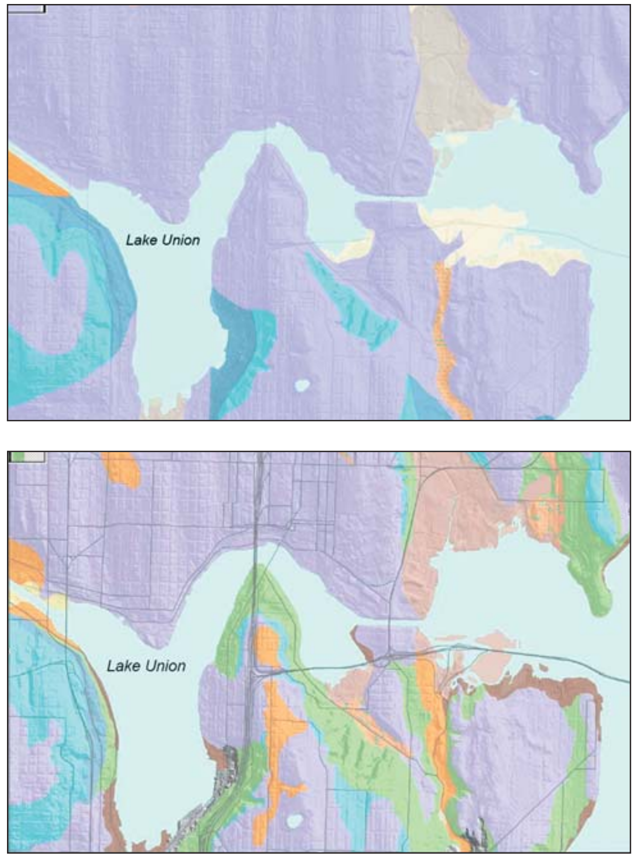


Figure 4. Comparison of old and new geologic maps of Seattle (differences only evident in online color version; printed version available only in grayscale). A portion of the geologic map of Seattle from Waldron and others (1962; top), and from Troost and others (2005a; bottom).

In areas where both local-agency concerns and regional geologic questions have warranted intensive study, and where funding was provided, this database has been applied to the development of new geologic maps. These include the westward and eastward extension of the Seattle fault (Haugerud, 2005; Booth and others, in review b; Troost and others, in prep.) and planned expansion areas of the regional wastewater-treatment system, particularly just north and east of Seattle.

REGIONAL STRATIGRAPHY AND CHRONOLOGY

In addition to the focused acquisition of data and development of very large-scale geologic maps, we are developing a chronological and lithologic context for the complex sequence of glacial and nonglacial deposits in the central Puget Lowland, one that can be used to evaluate the distribution, correlation, and deformation of individual geologic units across the region. As a result of the mapping and stratigraphic and chronologic work being done for our

geologic maps, we have established a regional stratigraphic nomenclature and updated timescale. Fundamental errors of stratigraphic (mis-) assignment in the southern Puget Lowland have been recognized over the last two decades, reflecting profound differences between stratigraphic sections exposed in the southeastern (Crandell, 1963) and northern Puget Lowland (e.g., Easterbrook and others, 1981; Blunt and others, 1987). Regional mapping and chronologic efforts (e.g., Hagstrum and others, 2002; Mahan and others, 2003; Figure 5) are now beginning to reconciling these differences (see also Booth and others, 2004). Through collaboration with USGS scientists, for example, we have shown that the stratigraphic units identified at type sections on Whidbey Island (Easterbrook, 1986), 40 km north of Seattle, can be identified more than 70 km south in the Tacoma area using absolute age control (Troost and others, in press), and we have identified deposits from mid-Pleistocene climatic stages previously undocumented anywhere in the Puget Lowland.

PUBLIC, AGENCY, AND CONSULTANT ACCESS TO DATA

The manner of data distribution outside of our immediate research group has been guided by the individual users. For those public agencies that have provided us with sources of data and, commonly, funding as well, we have been delivering quarterly (static) updates of the database, generally as ESRI shapefiles of the documents and exploration points and dBASE dbf files for subsurface layers and comments. The agencies, in turn, load these data onto their intranets, to be available to staff (Figure 6).

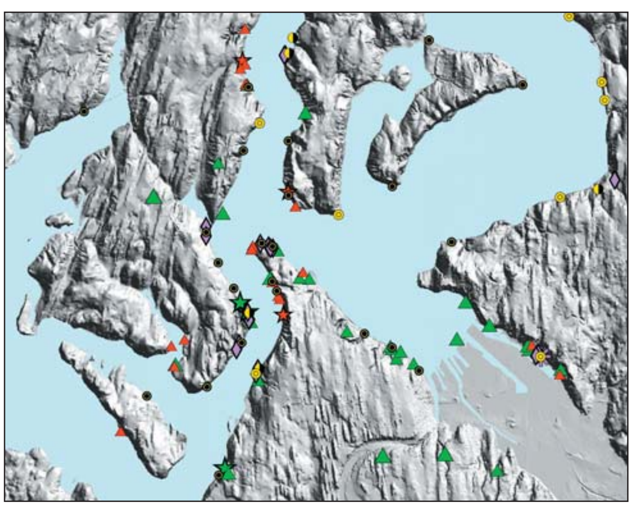


Figure 5. Map of analytic samples of Quaternary sediments collected, dated, and/or compiled by the project. Key: circles = paleomagnetic samples; diamonds = IRSL age samples; triangles and stars = ^{14}C age samples; snowflake = fission-track age sample.

Actual use of the data, however, is almost certainly quite variable. In the City of Seattle, for example, where our interaction and funding spans seven years, engineering and building departments use the database regularly and we maintain a systematic program of adding new information and delivering it to the City. For some of the smaller cities, however, usage by staff is probably less common; in addition, many of these smaller jurisdictions were only contacted by us during a single interval of data collection, and so the one-time digital compilation of geologic explorations will drift inexorably more and more out-of-date. We have not yet solved the logistical and financial problem of maintaining a truly "current" data set in each of the areas once visited for data acquisition.

We also provide a point of public access to our data, in part to satisfy our funders' goal of public data access, and in part to provide a broader service to the geotechnical and engineering community without making undue demands on our time. Access is through the Center website, http://geomapnw.ess.washington.edu (Figure 7a); the most heavily used links are those for downloading of publications and geologic maps (Figure 7b) and for individual queries of the geologic database (Figure 7c and d), for which we upload a static update on a roughly quarterly basis. Typical rates of access for the first half of 2005 have been about 700 unique visits per week, with 75 downloads/week of reports and maps and about 300 queries/week of individual exploration logs. At the continuing request of colleagues in the consulting community, we are in the process of scanning all of our borehole data and posting those scanned images on the web as pdf files. Currently almost two-thirds of our files are scanned and available.

EDUCATION AND TECHNICAL OUTREACH

We have actively participated in and led seminars, field trips, professional short courses, and workshops, to educate the scientific and nontechnical community about the baseline geologic setting of the Seattle and Tacoma areas. This acknowledges a critical emphasis in urban-area geology, namely bridging the gap between research and consulting geology. This is an ongoing effort with steadily increasing attention and influence. It also requires a significant expenditure of time, but one that we feel is critical to the long-term viability and value of our work.

To further support this outreach, a technical advisory group was established early in our first year to enhance communication between this project and the end users of the products, especially consultants and agency representatives. The group's membership, several dozen in number, emphasizes senior members of the region's geologic, geotechnical, hydrogeologic, and engineering consulting firms, and also includes representatives from state, city,

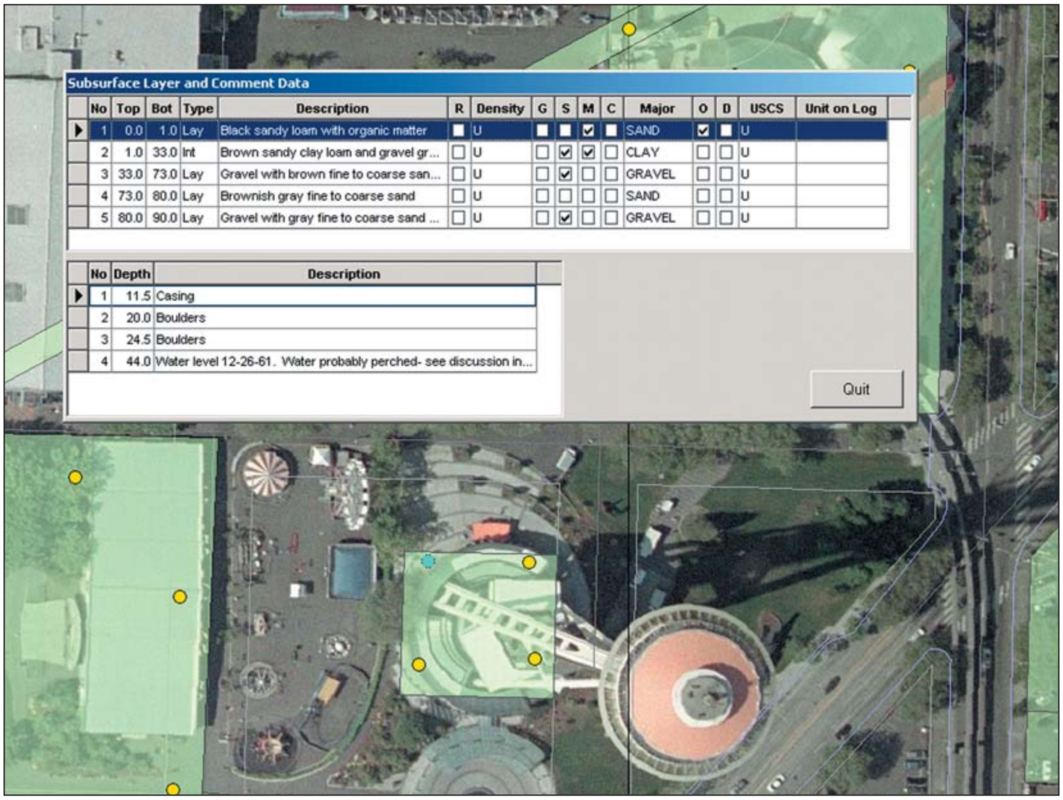


Figure 6. Example of an ArcView data query screen. The base aerial photograph is of the Space Needle; document areas are shaded (green in color online version). Of the four explorations originally drilled for the Space Needle foundation, that in the upper left-hand corner of the Needle footprint (turquoise highlight in online version) has been selected; the pop-up window shows the description of the five geologic layers in the exploration log and the dominant and secondary grain sizes as parsed from the layer description (upper table), and any comments (lower table).

and local agencies who are both the major users and the major contributors of data.

Our partnerships have permitted the digital archiving of some of the very best data—closely spaced, deep, linear transects of continuously sampled borings—provided by large capital projects. Together with new field mapping and the many additional sites of prior study by both public agencies and private individuals, these data are now starting to provide excellent opportunities to learn about the region’s geology. They are also forming the basis for the new, detailed, large-scale geologic maps of the region’s urban and urbanizing areas that are now being prepared and published.

FUTURE PLANS AND ISSUES

Although the project in its current form has demonstrated the value of detailed data compilation within the framework of regional scientific investigations, the full range of this approach to geospatial data has been explored only modestly. We recognize several additional areas in which this work could expand to the greater benefit of current and future users:

1. Creating a data model for incorporating other types of spatial information, emphasizing widely available base data that is available not only across all of the Puget Lowland but also nationwide.
2. Expanding the existing geologic data compilation, both spatially and thematically, to achieve spatially contiguous coverage over our region of interest and to incorporate geospatial data types not part of our current data model into a relational database structure.
3. Integrating these disparate data types into a single access interface.
4. Expanding how users, both members of the project team and the broader public, can view, query, and analyze the data for scientific, engineering, and educational applications, emphasizing web-accessed map-based interfaces.
5. Developing a systematized approach to data

Figure 7. Screenshots of the types of data access available from our website. (A), Center home page (http://geomapnw.ess.washington.edu). (B), view of index screen for downloading geologic maps. Queries for those maps available only in draft form are served in .pdf format from this site directly; queries for those maps that are already published by the USGS are redirected to the corresponding USGS page. (C) view of part of central Seattle in the ArcIMS window used to view and select explorations in map view. Zooming in to a local area (D) allows selection of an individual point (highlighted in white circle), which opens windows for the point's layer information and for the metadata on the data point and the source document.

delivery and outreach to known and potential users.
6. Creating new geologic products, particularly subsurface visualizations and 3-D representations of surfaces and stratigraphic layers.

Although these future plans would expand the value of detailed geologic information, the *current* costs of the present effort are already quite substantial: for example, a detailed, digital, USGS-published 7.5' geologic quadrangle map based on new field work and a subsurface database has averaged $250,000 at 1:24,000-scale and about twice that amount at 1:12,000 scale (i.e. across the City of Seattle). Derivative maps are not nearly as expensive, but they too add an incremental expense. In an urban area such as Seattle, the cost of detailed geologic mapping and a subsurface database is more palatable when expressed as a function of population density, with rates of about $1.75 to $2.00 per person (Troost and others, 2005b). Ultimately, however, the value of detailed mapping and geologic data must be quantified wherever we try to initiate or continue support for them. The question we therefore face is whether these new geologic products are worth their cost; and even if they are, can we find funding agencies with the foresight to recognize that value and to bear the expense?

REFERENCES

Blunt, D.J., Easterbrook, D.J., and Rutter, N.W., 1987, Chronology of Pleistocene sediments in the Puget Lowland, Washington, in Schuster, J.E., editor, Selected papers on the geology of Washington: Washington Division of Geology and Earth Resources Bulletin 77, p. 321-353.

Booth, D.B., Troost, K.G., and Shimel, S.A., in review a, Geologic map of the Seattle NE quadrangle: U.S. Geological Survey Scientific Investigation Map, scale 1:12,000.

Booth, D.B., Troost, K.G., and Walsh, T.J., in review b, Geologic map of the Issaquah quadrangle: U.S. Geological Survey Scientific Investigation Map, scale 1:24,000.

Booth, D.B., Troost, K.G., and Shimel, S.A., 2005, Geologic map of northwestern Seattle (part of the Seattle North 7.5' x 15' quadrangle), King County, Washington: U.S. Geological Survey Scientific Investigations Map 2903, scale 1:12,000, available at http://pubs.usgs.gov/sim/2005/2903/.

Booth, D.B., Troost, K.G., Clague, J.J., and Waitt, R.B., 2004, The Cordilleran ice sheet: Chapter 2, in Gillespie, A., Porter, S.C., and Atwater, B., eds., The Quaternary Period in the United States: International Union for Quaternary Research, Elsevier Press, p. 17-43.

Crandell, D.R., 1963, Surficial geology and geomorphology of the Lake Tapps quadrangle, Washington: U.S. Geological Survey Professional Paper 388-A, 84 p., 2 plates.

Easterbrook, D.J., 1986, Stratigraphy and chronology of Quaternary deposits of the Puget Lowland and Olympic Mountains of Washington and the Cascade mountains of Washington and Oregon: Quaternary Science Reviews, v. 5, p. 145-159.

Easterbrook, D.J., Briggs, N.D., Westgate, J.A., and Gorton, M.P., 1981, Age of the Salmon Springs glaciation in Washington: Geology, v. 9, no. 2, p. 87-93.

Hagstrum, J.T., Booth, D.B., Troost, K.G., and Blakely, R.J., 2002, Magnetostratigraphy, paleomagnetic correlation, and deformation of Pleistocene deposits in the south-central Puget Lowland, Washington: Journal of Geophysical Research–Solid Earth 107 (B4): art. no. 2079, Apr. 10, 2002.

Haugerud, R., 2005, Preliminary geologic map of Bainbridge Island, Washington: U.S. Geological Survey Open-File Report 2005-1387, scale 1:24,000, available at http://pubs.usgs.gov/of/2005/1387/.

Mahan, S.A., Troost, K.G., and Booth, D.B., 2003, Dating sediments older than 100 ka in the Seattle-Tacoma urban corridor: A test for infrared stimulated luminescence and thermoluminescence on fine grain deposits: Geological Society of America, Abstracts with Programs, Annual Meeting, v. 35.

Troost, K.G., Booth, D.B., and Wells, R.E., in press, Geologic map of the Gig Harbor 7.5-minute quadrangle, Washington: U.S. Geological Survey Scientific Investigation Map, scale 1:24,000.

Troost, K.G., Booth, D.B., and Wisher, A.P., in preparation, Geologic map of the Mercer Island quadrangle: U.S. Geological Survey Scientific Investigation Map, scale 1:24,000.

Troost, K.G., Booth, D.B., Shimel, S.A., Blakely, R.J., and Wells, R.E., in review a, Geologic map of the Seattle SW quadrangle: U.S. Geological Survey Scientific Investigation Map, scale 1:12,000.

Troost, K.G., Booth, D.B., and Shimel, S.A., in review b, Geologic map of the Seattle SE quadrangle: U.S. Geological Survey Scientific Investigation Map, scale 1:12,000.

Troost, K.G., Booth, D.B., Wisher, A.P., and Shimel, S.A., 2005a, The geologic map of Seattle—a progress report: U.S. Geological Survey Open-File Report 2005-1252, scale 1:24,000, available at http://pubs.usgs.gov/of/2005/1252/.

Troost, K.G., Booth, D.B., Shimel, S.A., Wisher, A.P., and O'Neal, M., 2005b, Detailed geologic mapping—is it worth the cost? Applications of a geodatabase of the Seattle, Washington area: Geological Society of America, Abstracts with Programs, Annual Meeting, v. 37.

Waldron, H.H., Liesch, B.A., Mullineaux, D.R., and Crandell, D.R., 1962, Preliminary geologic map of Seattle and vicinity, Washington: U.S. Geological Survey Miscellaneous Investigations Map I-354, scale 1:31,680.

APPENDIX A

List of Workshop Attendees

[Grouped by affiliation]

Adobe Systems, Inc.
Mike Bennett

Arkansas Geological Commission
Jerry W. Clark
William D Hanson

Avenza Systems, Inc.
David Andrec
Douglas Smith

Arizona Geological Survey
Stephen M Richard

Colorado Geological Survey
Jason Wilson

Colorado State University/National Park Service Cooperator
James Chappell
Stephanie O'Meara
Heather Stanton

ESRI, Inc.
Brig Bowles
Veronica Schindler

Florida Geological Survey
David Anderson

Geological Survey of Canada
Guy Buller
Eric Boisvert
Peter Davenport
Parm Dhesi
Gary Grant
Roger Macleod
Stephen P. Williams

Idaho Geological Survey
Loudon Stanford
Jane Freed

Illinois State Geological Survey
Sheena Beaverson
Jane Domier
Marie-France Dufour
Rob Krumm
Deette Lund

Intrepid Geophysics
Philip McInerney

Kansas Geological Survey
William E. Harrison

Kentucky Geological Survey
Gerald Weisenfluh

Louisiana Geological Survey
Roger Barnaby
Reed Bourgeois
Clayton Breland
Douglas Carlson
Weiwen Feng
Bill Good
Brian Harder
Paul Heinrich
Chacko John
Jeanne Johnson
John Johnston
Bobby Jones
Rick McCulloh
Byron Miller
Riley Milner
Patrick O'Niell
Robert Paulsell
Hampton Peele
Lisa Pond
John Snead
Anne R. Tircuit
Thomas Van Biersel

Louisiana State University
Sidney Egnew

MD Atlantic Technologies
Ray D. Dupre
Kevin S. Lim

Minnesota Geological Survey
Harvey Thorleifson

Missouri Geological Survey and Resource Assessment
Edith Starbuck

National Park Service
Gregory Mack
Anne R. Poole

Natural Resources Canada
Ruth Boiuin
Benoit Chagnon
Nathalie Cote
Victor Dohar
Dave Everett
Terry Houlahan
Andrew Moore

Nevada Bureau Of Mines and Geology
Elizabeth Crouse

New Hampshire Geological Survey
Frederick Chormann, Jr.

New Mexico Bureau of Geology and Mineral Resources
David J. McCraw

Noranda Falconbridge
Pierre St. Antoine

North Dakota Geological Survey
Lorraine Manz

Ohio Geological Survey
Thomas Berg
James McDonald

Ontario Geoscience Survey
Zoran Madon

Oregon Dept. of Geology and Mineral Industries
Paul Staub

Pennsylvania Geological Survey
Thomas G Whitfield

Portland State University
David Percy

South Carolina Geological Survey
Erin Hudson

U.S. Fish & Wildlife Service
Douglas Vandegraft

U.S. Forest Service
Andrew Rorick

U.S. Geological Survey
Karynna Calderon
E. Allen Crider, Jr.
Alex Donatich
Joseph East
James Flocks
Christopher Garrity
Charles Groat
Linda Masonic
David R. Soller
Nancy Stamm
Evan Thoms
Ronald R. Wahl

University of Alabama
Douglas Behm

University of Arizona—ESPRI
Jonathan Craigue
Harry McGregor

University of Denver/National Park Service Cooperator
Georgia Hybels

University of Massachusetts
Chris Condit

University of Tennessee—Knoxville
Robert D. Hatcher, Jr.

Universita' Degli Studi Di Urbino, Italy
Mauro De Donatis

Utah Geological Survey
Kent Brown

Virginia Division of Mineral Resources
Elizabeth M. Campbell

West Virginia Geological Survey
Jane S. McColloch
Gayle H. McColloch, Jr.

Washington Division of Geology and Earth Resources
Charles Caruthers

Wisconsin Geological Survey
Michael L. Czechanski
Peter Schoephoester

Wyoming State Geological Survey
Allory Deiss
Phyllis Ranz