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

Digital Mapping Techniques '04—Workshop Proceedings

Edited by David R. Soller

May 16-19, 2004
Portland, Oregon

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

Hosted by the
Oregon Department of Geology and Mineral Industries

U.S. GEOLOGICAL SURVEY OPEN-FILE REPORT 2004-1451
2004

CONTENTS

Introduction
By David R. Soller (U.S. Geological Survey) 1

Oral Presentations

Completion of the First Phase of the Kentucky Digital Geologic Mapping Program
By Warren H. Anderson, Thomas N. Sparks, and Gerald A. Weisenfluh (Kentucky Geological Survey) 5

Progress Towards an Agency-wide Geologic Map Database at Alaska Division of Geological & Geophysical Surveys
By Larry Freeman and Fred Sturmann (Alaska Division of Geological and Geophysical Surveys) 9

The National Geologic Map Database Project: Overview and Progress
By David R. Soller (U.S. Geological Survey), Thomas M. Berg (Ohio Geological Survey), and Nancy R. Stamm (U.S. Geological Survey) 15

Migrating From ArcInfo Workstation to ArcGIS
By Victor Dohar (Natural Resources Canada) 33

Investigating the San Simeon Earthquake Using ArcPad and GPS
By Lewis I. Rosenberg (San Luis Obispo County Planning and Building Department) 41

GanFeld: Geological Field Data Capture
By Guy H.D.P. Buller (Geological Survey of Canada) 49

Geologic Data Assistant (GDA): An ArcPad Extension for Geologic Mapping
By Ralph A. Haugerud and Evan E. Thoms (U.S. Geological Survey) 55

Creating 3D Models of Lithologic/Soil Zones using 3D Grids
By Skip Pack (Dynamic Graphics, Inc.) 65

Science Language, Parsing and Querying: the Surficial Side of Things
By Andrée M. Bolduc, Andrew Moore, Kaz Shimamura, Matt Pyne, and Ron DiLabio (Geological Survey of Canada) 75

Report on Progress to Develop a North American Science-Language Standard for Digital Geologic-Map Databases
By North American Geologic Map Data Model Steering Committee Science Language Technical Team 85

GML Encoding of NADM C1
By Eric Boisvert (Geological Survey of Canada), Bruce R. Johnson (U.S. Geological Survey), Simon Cox (CSIRO, Australia), and Boyan Brodaric (Geological Survey of Canada) 95

Geologic Web Services, Planning and Design
By Sheena K. Beaverson (Illinois State Geological Survey) 105

Implementing NADM C1 for the National Geologic Map Database
By Steven M. Richard (Arizona Geological Survey and U.S. Geological Survey), Jon Craigue (University of Arizona/ESPRI), and David R. Soller (U.S. Geological Survey) 111

Photogrammetry Methods at the Utah Geological Survey: From Field Mapping to Published Map
By Kent D. Brown (Utah Geological Survey) 145

Digital Cartographic Production Techniques Using Airborne Interferometric Synthetic Aperture Radar (IFSAR): North Slope, Alaska
By Christopher P. Garrity (U.S. Geological Survey) 153

The Use of Topology in Geologic Maps
By Ron Wahl (U.S. Geological Survey) 159

Poster Presentations

Geology, Stewart River Area (Parts of 115 N/1,2,7,8 and 115-O/2-12), Yukon Territory
By S. P. Gordey, J.J. Ryan, and R.B. Cocking (Geological Survey of Canada) . . . 165

Surface Terrain of Indiana—a Digital Elevation Model
By Robin Rupp (Indiana Geological Survey) 171

Cartography at the Montana Bureau of Mines and Geology
By Susan Smith (Montana Bureau of Mines and Geology) 177

Conversion of Surficial Geologic Maps to Digital Format in the Seacoast Region of New Hampshire
By Derek S. Bennett, Frederick H. Chormann, Jr., Carl Koteff, and David R. Wunsch (New Hampshire Geological Survey) 181

Brownfield and Contaminated Site Remediation: Historic Fill Series
By Michael W. Girard (New Jersey Geological Survey) 189

Exploring Shaded Relief: The Shaded Drift-Thickness Map of Ohio
By Donovan M. Powers (Ohio Division of Geological Survey) 191

Conversion of a Digital Raster Graphics Image from Color to Black & White
By Clark A. Niewendorp (Oregon Department of Geology and Mineral Industries) 195

Oregon Statewide Geologic Map Data: A Pilot Project Where Digital Techniques Changed the Geologic Map Compilation Process and Product
By Mark L. Ferns, Ronald P. Geitgey, Margaret D. Jenks, Lina Ma, Ian P. Madin, Vicki S. McConnell, and Paul E. Staub (Oregon Department of Geology & Mineral Industries) 197

Going From Greenline Mylar to Digital Geologic Map
By Thomas G. Whitfield (Pennsylvania Bureau of Topographic and Geologic Survey) 203

Geoenvironmental Map of the Christmas Point Quadrangle, Texas Gulf Coast
By Thomas A. Tremblay, William A. White, Rachel L. Waldinger, and Edward W. Collins (Texas Bureau of Economic Geology) 207

Process of Converting Geologic Maps from Coverages to Geodatabases
By Karen Wheeler, Philip Dinterman, Russell Evarts, and Ray Wells (U.S. Geological Survey) 209

Appendix A. List of Workshop Attendees 219

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 '04 (DMT'04) workshop was attended by about 100 technical experts from 40 agencies, universities, and private companies, including representatives from 22 state geological surveys (see Appendix A). This workshop was similar in nature to the previous seven 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), and in Millersville, Pennsylvania (Soller, 2003). This year's meeting was hosted by the Oregon Department of Geology and Mineral Industries, from May 16-19, 2004, on the Portland State University campus in Portland, Oregon. As in the previous meetings, the objective was to foster informal discussion and exchange of technical information. This objective was well met, as attendees continued to share and exchange knowledge and information, and to renew friendships and collegial work begun at past DMT workshops.

All the DMT workshops have 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, Berg, and Stamm, 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 National database—and for the State and Federal geological surveys—to provide more high-quality digital maps to the public.

At the 2004 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; 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 Oregon Department of Geology and Mineral Industries (DOGAMI) and their Director and State Geologist, Vicki McConnell, for hosting this meeting. In the tradition of past DMT meetings, the attendees were given a very informative, productive, and enjoyable experience. I especially thank Paul Staub (DOGAMI), who coordinated the events; Paul provided excellent support for the attendees (including large amounts of good, strong coffee) and offered a fine range of technical and social activities (for example, an evening at one of Portland's brewpubs). Thanks also to Kate Halstead for managing the registration and logistics, James Roddey for managing the meeting's website, and Lu Clark, Mark Neuhaus, and Clark Niewendorp for providing support for the meeting logistics. I also enthusiastically thank Portland State University for providing an excellent venue and support for our meeting, and in particular, Dr. David Percy (Department of Geology). In addition to providing for our many audio-visual and related needs, David was an active participant in the meeting, and I hope he enjoyed our company as much as we enjoyed his.

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, Kentucky 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 help 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) and Jane Freed (Idaho Geological Survey) for moderating the discussion sessions; they kept us focused and greatly improved the productivity of these sessions. 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 22 oral presentations. Nearly all 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. "Rapid implementation of the Geodatabase and the Geology Data Model" by Alon Yaari, ESRI,
2. "3-D zone modeling from uncorrelated well/boring data" by Skip Pack, Dynamic Graphics, Inc. (see paper in these Proceedings),
3. "VIEWLOG: a tool for borehole data analysis and constrained geologic modeling" by Dirk Kassenaar, Viewlog Systems, and
4. "Current trends and future developments in HP Large Format printing technology" by Randy Heilbrunn, Hewlett-Packard.

POSTERS AND COMPUTER DEMOS

Nearly 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.

ESRI GEODATABASE SESSION

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. Attendees therefore requested a special session to discuss how to do this. ESRI graciously agreed to lead this session; I thank Alon Yaari, who led this session, and Joe Breman, who contributed to its content. This session provided hands-on instruction in the process of converting Coverages and Shapefiles to Geodatabase format, using selected geologic maps. Instruction also included Geodatabase design concepts. This topic will receive further emphasis at DMT'05.

DISCUSSION SESSIONS

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.

Session 2 was held on the final day of the meeting, and produced many good ideas and recommendations that will be discussed by the Data Capture Working Group and DMT'04 attendees via the DMTListserve. Jane Freed (Idaho Geological Survey) and Dave Soller (USGS) moderated this session. The most significant outcome was the recommendation that the NGMDB build a web Clearinghouse to provide links to cartographic standards, geologic map layout templates in various formats (e.g., Adobe Illustrator), and general information about the design of geologic maps. This website is intended to support digital cartographers and GISers in State and Federal agencies, as well as students who are making their first map (e.g., in the EDMAP program). This website is available in prototype form, at http://ngmdb.usgs.gov/info/cartores/; all interested persons are encouraged to comment on the site, provide guidance, and contribute links to cartographic resources.

THE NEXT DMT WORKSHOP

The ninth annual DMT meeting will be held in late April 2005, in Baton Rouge, Louisiana. Please consult the Web site (http://ngmdb.usgs.gov/info/standards/datacapt/)

for updated information. While planning for that event, the Data Capture Working Group will carefully consider the recommendations offered by DMT'04 attendees.

REFERENCES

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/.

Completion of the First Phase of the Kentucky Digital Geologic Mapping Program

By Warren H. Anderson, Thomas N. Sparks, and Gerald A. Weisenfluh
Kentucky Geological Survey

Room 228 Mining and Minerals Resources Building
University of Kentucky
Lexington, Kentucky 40506
Telephone: (859)-257-5500
Fax: (859)-257-1147
e-mail: wanderson@uky.edu

ABSTRACT

The Kentucky Geological Survey (KGS) has completed the vectorization of the 707 geologic quadrangle maps (7.5 x 7.5-minute) that cover the state. Under the National Cooperative Geologic Mapping Program and other State funding awards from 1996-2004, this milestone was reached in May 2004. It provides Kentucky with complete digital geologic map coverage at a scale of 1:24,000. At the completion of each year's contractual phase of data capture and attribution, the project deliverables consisted of (1) a hard-copy map product (1:100,000-scale) and (2) the digital geologic data (both as individual quadrangles and as compiled geologic data layers, which represent spatially referenced features including faults, formation contacts, coal beds, fossil, and quarry locations. In 2005, these maps will be compiled into 32 individual 1:100,000-scale geologic maps by edge-matching each boundary (Figure 1). In the process of this compilation, the 1:24,000-scale geologic map data are generated for separate release as Digitally Vectorized Geologic Quadrangles (DVGQ's). Some of these DVGQ's have already been released to the public, the remainder are undergoing final review, editing, and file processing. In addition to the new DVGQ product, seven of the compiled 30 x 60 minute geologic quadrangle maps (1:100,000-scale) have been published in a new Geologic Map series at KGS.

During the past decade, the KGS Digital Mapping Program employed over 50 staff and students, who converted geologic quadrangle maps into digital format at a cost of about $3.8 million. The geologic quadrangle maps that were digitized were created during the Kentucky Geological Survey–U. S. Geological Survey mapping program from 1960 to 1978, which mapped the entire state at 1:24,000-scale, at a cost of $20.9 million. That original mapping program employed over 250 geologists. The dedicated work of many people in both programs made the mapping of Kentucky possible, and provides the most detailed digital geologic maps of any large area in the United States.

CHALLENGES AND LESSONS LEARNED

Challenges during the first phase of the Digital Geologic Mapping Program included: (1) maintaining sustained funding levels, (2) discriminating between disparate mapping philosophies used during the original mapping, (3) resolving complex stratigraphic problems, and (4) establishing standards for GIS platforms, vectorization, attribution, and database issues.

Many of the original geologic mappers had different mapping philosophies, and we categorized them as "lumpers or splitters" of stratigraphic units. Some of the more common issues dealt with were stratigraphic name changes, incorrect correlation of formation contacts, and lumping of minor members or splitting of formations. Normal facies, lithologic, and geologic changes in rocks and in their nomenclature also had a great impact on stratigraphic correlation.

KGS instituted a stratigraphic committee to resolve many of the stratigraphic inconsistencies, but many correlation issues could not be addressed until digitizing began. This caused many problems in edge matching adjacent quadrangles, and is unavoidable when dealing with stratigraphy in sedimentary rocks. Facies changes, pinchouts, unconformities, fluvial and marine channels, and transgressive sequences were also a major influence on the lithostratigraphic representation of rocks as portrayed on geologic quadrangle maps. Because different mappers often worked in adjacent quadrangles at different times, many of the quadrangles did not correlate properly. Each of these problems made it challenging to create the 30 x 60 minute quadrangles, but their resolution also revealed new geologic features or interpretations which had not been previously mapped.

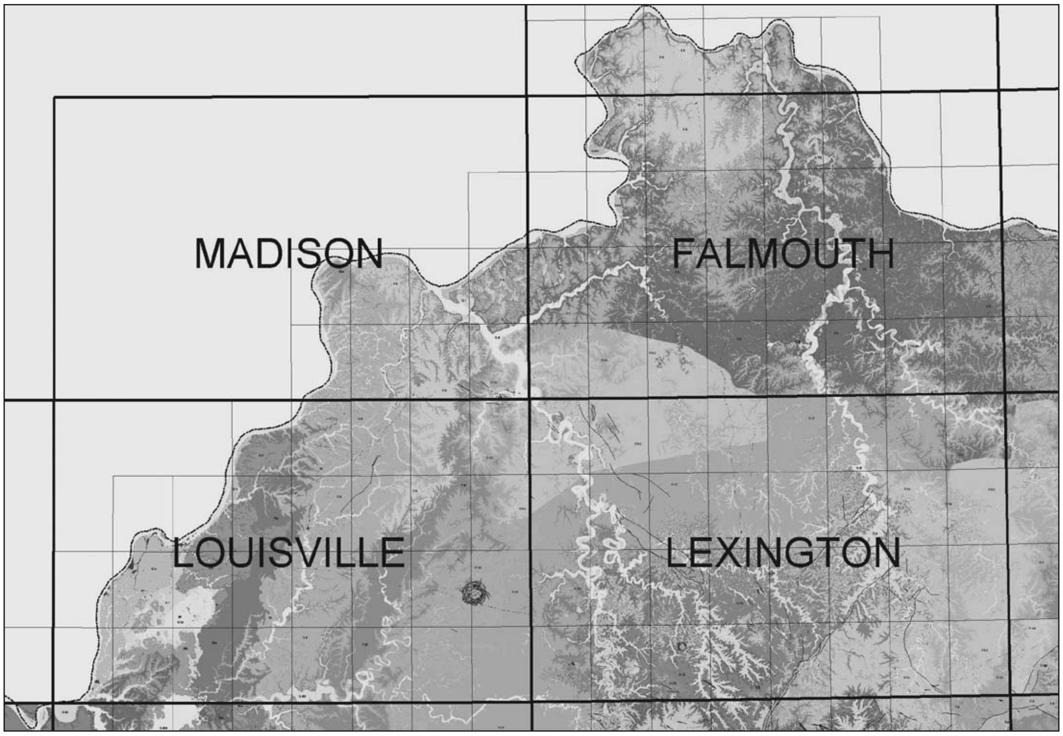


Figure 1. A part of central Kentucky showing geology of four 1:100,000-scale published maps, each covering an area of 30 x 60 minutes. These maps were compiled from digital 1:24,000-scale geologic quadrangle data (up to 32 of these 1:24,000-scale quadrangles occur in each map; their outlines are shown by the thin black lines).

Early in the program, KGS created several committees to address the issues of computer systems, GIS platforms, database organization, vectorization and attribution (Anderson and others, 1997). The committee's decisions would affect the capture, classification, content, and storage of geologic map data. ArcInfo and ArcView are robust GIS software and were used to capture the data. KGS was a founding participant in the North American Data Model (NADM) Steering Committee, and continues to contribute to that body through the Data Model Design Team (Weisenfluh, 2001). The Kentucky data have been used as a prototype database for developing the NADM (Soller and others, 2002).

Because more than 50 employees were digitizing quadrangles, some issues arose regarding the standard procedure for digitizing. Initially, we used an automatic vectorization procedure (Anderson and others, 1997). Many of our staff did not like this method because of the time required for post-digitizing cleanup and because the resultant vector lines were very jagged; this occurred where lines were closely spaced, causing the autovectorizing routine to "jump" from one line to a line nearby. Therefore, many of our staff preferred to use the manual method for digitizing to resolve those problems. Another important issue was how many vertices to use on rounded contact lines, since the more vertices used, the longer the time required to complete a quadrangle. KGS created an in-house manual describing the procedure for digitizing geologic quadrangles in order to create a standard digital product. With refinement of our work process (sometimes through trial and error), modification of our techniques, and the maturity of the staff both in technical ability and responsibility, we created three or more teams and were able to complete our data capture ahead of our original schedule.

THE KENTUCKY SPATIAL DATABASE

KGS has a long history of providing access to geologic information via the Internet and has sought to distribute its tabular and spatial databases to the public (Anderson and others, 1999; Curl, 2000). As the databases grew, the complexity of serving the data in an efficient manner became more important. During the initial stages of digital mapping, KGS began to examine the issues of serving the spatial data via the Internet. KGS formed the Geospatial Analysis Section to specifically address the problems of managing and serving the large amount of geologic map data being generated.

KGS decided to construct and maintain its own spatial databases, composed of a large geologic map database with numerous related, specialized databases (Anderson and others, 1999; Weisenfluh and others, 2002). ArcSDE was chosen as the software platform to manage these data. Data layers are being prepared in a single-coordinate system and datum (NAD 83, decimal degrees or Kentucky Single Zone) to simplify data integration. The specialized KGS databases (oil, water, coal, sample locations, and minerals) are maintained in a relational database, but have been spatially enabled by adding locations to feature classes in SDE (Weisenfluh and others, 2002).

Demand for the data

Within weeks of the first release of our digital data, numerous Federal and State agencies and University researchers actively sought it. One University purchased the entire DVGQ set before it was even completed. Other Federal and State agencies and private industry have sought this digital data for environmental, mapping, planning, land use, engineering and exploration purposes.

DERIVATIVE MAPS

The 7.5-minute DVGQ dataset is the basis of, and fundamental map for, many derivative products that KGS is and will be developing in the future. Currently, we are producing county geologic maps, planning and land use county maps, and hydrologic atlases for each county. A prototype of a karst potential index map (Crawford and Currens, 2004) has been developed and demonstrated by using lithologic characteristics of each geologic map unit in a GIS environment. Geologic hazards maps and mineral resource are being considered for future products as derivative maps.

REFERENCES

Anderson, W.H., Morris, L.G., and Sparks, T.N., 1997, Semi-automated data capture for vectorizing geologic quadrangle maps in Kentucky, *in* Soller, D.R., ed., 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, p. 9-13, accessed at http://pubs.usgs.gov/of/1997/of97-269/anderson.html.

Anderson, W.H., Sparks, T.N., Patton, J., and Yang, X.Y., and Sergeant, R.E., 1999, Integration of digital geologic databases and a map database in Kentucky, *in* Soller, D.R., ed., Digital Mapping Techniques '99—Workshop Proceedings: U.S. Geological Survey Open File Report 99-386, p. 123-125, accessed at http://pubs.usgs.gov/of/1999/of99-386/anderson1.html.

Crawford, M.M., and Currens, J.C., 2004, Digital Geologic data and a new karst potential index map: a work in progress (abstract): Geological Society of America, Abstracts with Programs, Vol. 36, No. 2, p. 137.

Curl, D.C., 2000, Integrated geospatial data: Irvine 30 x 60 minute quadrangle, east-central Kentucky, *in* Soller, D.R., ed., Digital Mapping Techniques '00—Workshop Proceedings: U.S. Geological Survey Open-File Report 00-325, p. 175-176, accessed at http://pubs.usgs.gov/of/2000/of00-325/curl.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, 40 p., accessed at http://pubs.usgs.gov/of/2002/of02-202/.

Weisenfluh, G.A., 2001, Map unit descriptions and the North American Data Model, in Soller, D.R., ed., Digital Mapping Techniques '01—Workshop Proceedings: U.S. Geological Survey Open-File Report 01-223, p. 79-86, accessed at http://pubs.usgs.gov/of/2001/of01-223/weisenfluh.html.

Weisenfluh, G.A., Eaton, N.K., Nelson, K., 2002, Distributed spatial databases-The MIDCARB Carbon Sequestration Project, *in* Soller, D.R., ed., Digital Mapping Techniques '02—Workshop Proceedings: U.S. Geological Survey Open-File Report 02-370, p. 13-19, accessed at http://pubs.usgs.gov/of/2002/of02-370/weisenfluh.html.

Progress Towards an Agency-Wide Geologic Map Database at Alaska Division of Geological & Geophysical Surveys

By Larry Freeman and Fred Sturmann

Alaska Division of Geological & Geophysical Surveys
3354 College Road
Fairbanks, AK 99709-3707
Telephone: (907) 451-5027
Fax: (907) 451-5050
e-mail: Larry_Freeman@dnr.state.ak.us
Fred_Sturmann@dnr.state.ak.us

INTRODUCTION

The Alaska Division of Geological & Geophysical Surveys (DGGS) has been producing geologic maps using a Geographic Information System (GIS) since 1983 (Davidson, 1998). To take advantage of changing GIS technologies and to be able to provide its customers with quality digital geologic data, DGGS reviewed the agency geologic map data structure and geologic mapping process. This review resulted in a new geologic map data model to be used in an agency-wide database system that will support DGGS geologic map production and geologic map data distribution into the future. DGGS geologists are now incorporating the basic components of that model into the process of compiling and producing new geologic maps utilizing the ESRI personal geodatabase framework.

DGGS intends to integrate the geologic map database into a comprehensive, centralized division-wide database that includes bibliographic information about DGGS publications, metadata about DGGS GIS datasets, geochemical analysis data, field locality information, and other data (Freeman, 2001; Freeman and others, 2002). DGGS recently moved its publications database and scanned document index (Davidson, and others, 2002) into the centralized database and began serving the publications through dynamic web pages, at http://www.dggs.dnr.state.ak.us/pubs/pubs.jsp. DGGS currently is developing a set of web-based forms that will incorporate DGGS metadata for project GIS data into the database (Browne and others, 2003) and will facilitate serving DGGS data through its publications web pages. DGGS also is developing a web-based search engine for DGGS geochemical data that will provide tabular views of the geochemical analyses and documentary data, including links to the original publications. Integrating the geologic map features into its division-wide database will allow DGGS to provide fully documented geologic map data to both web and local GIS clients in the same way that we will soon provide geochemical data.

LEGACY DGGS GEOLOGIC DATA

Geologic maps contain information about geologic features and earth materials that can be used in spatial analysis to facilitate earth resource assessment, development, and planning. Providing consistent, reliable geologic map data for decision makers in industry, government, and the general public is an essential function of a state geological survey. For DGGS to provide consistent data, it must have a consistent organization or data structure.

Legacy Data Structure

Older, legacy DGGS map data is structured in various ways. Most of the older legacy data is managed in a hierarchical GIS file directory system organized by project, map product, data type, and theme (Figure 1). Each thematic data set (e.g., "bedrock") represents a different geometry, topologic grouping, or geologic feature classification. Attributes of each theme consist of codes that pertain to cartographic elements such as annotation or symbolic representation. To perform a geologic analysis, a data user needs to execute scripts to plot the information and needs a published map legend to determine the geologic context of the features. The existing directory and file structures, codes, and scripts have evolved and changed over time, from project to project, resulting in inconsistencies. Additionally, documentation of the legacy data is limited. As a result, to get full use of the DGGS geologic map data, users must depend on the specific knowledge and institutional memory of the people that created the data.

DGGS legacy datasets can easily be converted into ESRI geodatabase format. However, direct conversion of the data does not take full advantage of the relational structure inherent in the ESRI geodatabase because the legacy datasets contain limited attribute context and documentation. Significant work is required to convert the cartographic codes contained in the legacy data into geologic

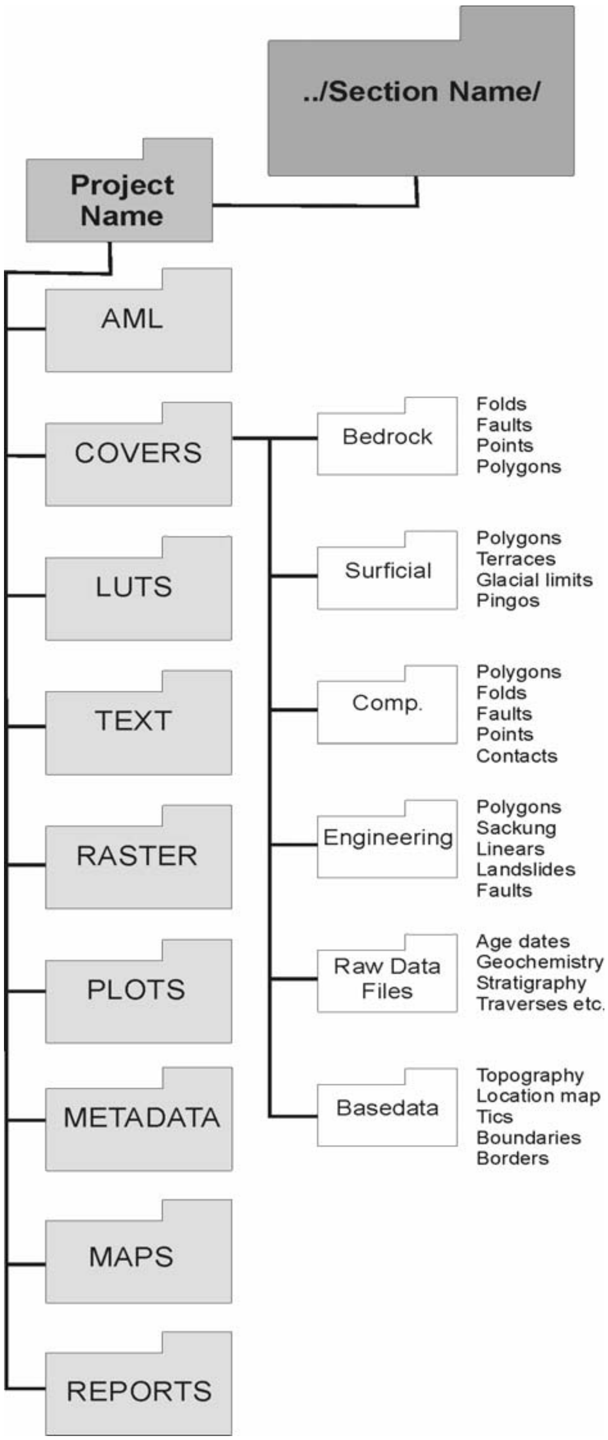


Figure 1. Hierarchical data structure for DGGS legacy maps, using the ESRI coverage data model. DGGS legacy map data is stored as coverages organized on a disk in subdirectories by Section (DGGS is divided into four sections, see http://www.dggs.dnr.state.ak.us/sections.html), Project, and Theme. For instance the bedrock fold features for any given map are stored as a coverage in the ../Covers/Bedrock/ subdirectory. The ../Comp./ subdirectory contains the data resulting from merging the bedrock geology data with the surficial geology data.

data. Therefore, DGGS suspended data conversion until it could adopt a consistent data structure that could be used for creation and integration of new geologic maps as well as for conversion of legacy data.

DGGS GEOLOGIC MAP DATA MODEL

To update the agency GIS data structure and integrate the legacy geologic map data with other geologic data, DGGS staff in 2003 reviewed both the concepts recorded in geologic maps and the agency geologic map-making process. In this review a cross section of DGGS geologists examined the definitions of geologic maps and their components to determine the information that is essential to geologic maps. The review considered the existing DGGS data structure and mapping processes, the data structure of other agencies, and geologic map data models and lexicon concepts being developed by the North American Geologic Map Data Model Design Team (see http://nadm-geo.org).

DGGS Geologic Map Definition

The DGGS discussions started by adapting the geologic map definition in Jackson (1997) to DGGS geologic maps:

> A geologic map is a two-dimensional graphic representation of selected geologic features on a part of the earth as observed and interpreted by the authors. Composition, physical characteristics, and relationships of earth materials (rocks and unconsolidated surface materials) in the area covered by the map are portrayed by graphical juxtaposition, symbols, and labels; supplemented by explanatory material presented with the map.

In current practice, DGGS geologic maps portray bedrock and surface geologic features in particular areas of Alaska, typically coinciding with a USGS 15-minute series quadrangle map, and often at a nominal scale of 1:24,000 or 1:63,360. Generally these maps are distributed as a series of four maps with different themes including a comprehensive geologic map, a surficial geologic map, a bedrock geologic map, and an engineering geology map.

The geologic map components shown in Figure 2 represent the physical parts of the map as defined by the DGGS working group. Components are shade-coded to indicate whether the component is "essential"; without the essential components, a geologic map or data do not contain usable information.

The group also looked at the science and institutional processes involved in evolving empirical data (field observations, sample analyses, and geophysical measurements) to geologic map data and finally to a published product. Understanding these processes helps to better define the

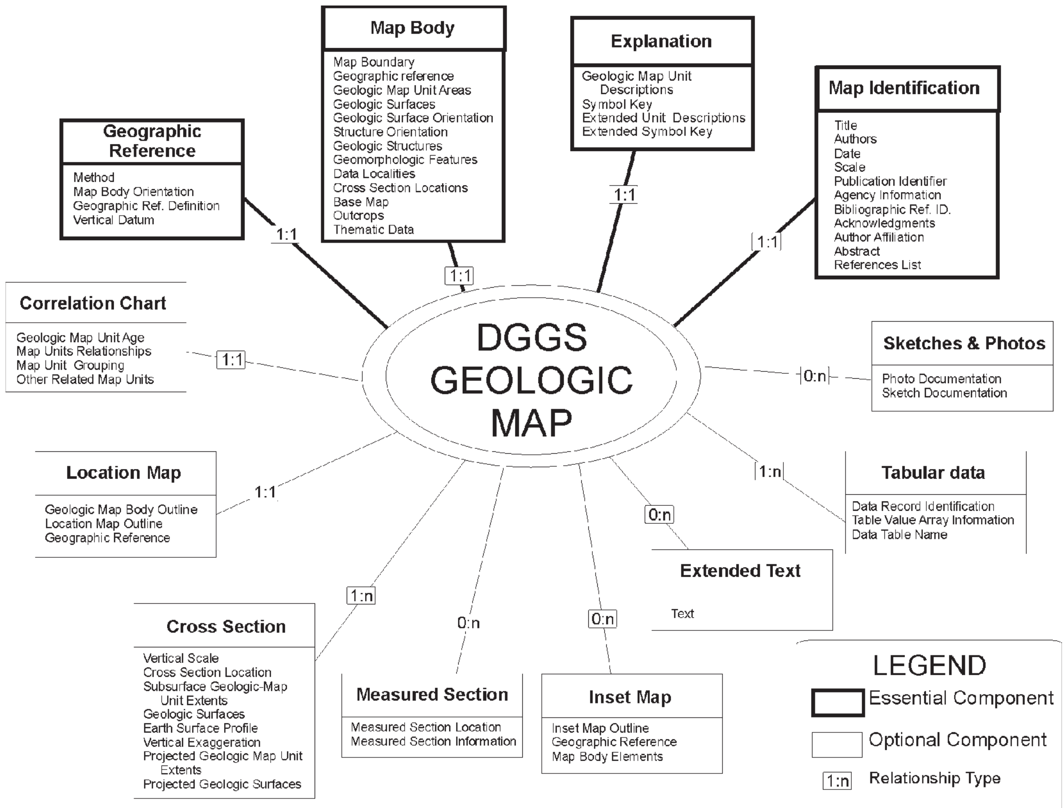


Figure 2. DGGS geologic map components as defined by DGGS geologists. Each component box contains a list of the elements of the component.

concepts and science logic that should be built into an object-relational database.

DGGS Geologic Map Conceptual Data Model

The map components identified by the DGGS working group are depicted as classes in a Unified Modeling Language (UML) Class diagram (Fowler and Scott, 2000). A "class" in a UML diagram (e.g., "data localities" in Figure 3) represents a group of objects that share characteristics and behavior. Within each class a box contains the attributes that characterize the class, and another box contains the operations that need to be carried out on the class or its attributes. Associated classes are tied together with connecting lines.

The data model is illustrated by a series of data centric, conceptual-level diagrams (Figures 3 and 4) that indicate the cartographic and scientific concepts contained in DGGS geologic maps and their essential components. For example, the "bedrock materials areas" class is a component of the geology class (Figure 4), which is in turn a component of the "map body" class (Figure 3), which is an essential component of a DGGS geologic map (Figure 2). The diagrams depict only the generalized, interpreted, and classified data that are contained in a completed geologic map. Operations in the class diagram are only those required to ensure data integrity and enforce standard vocabulary.

The geologic map body component of a DGGS Geologic Map (Figure 3) is the "two-dimensional graphic representation of selected geologic features on a part of the earth" and is the essential core of the geologic map. Map identification, geographic reference, and explanation are already integrated into our geologic database to at least some degree. Within the geologic map body the most essential subcomponent is the geology class (Figure 4).

The geology class contains information about the earth materials and features depicted on a geologic map. The components of the geology class include earth materials (both bedrock and surficial) classification, which have an area (polygon) geometry; structural measurements, which have a point geometry; and geologic structures, which may be planar but are represented in the GIS with a line geometry. Earth material areas are grossly classified into surficial material areas and bedrock material areas, to reflect the agency mapping process and map distribution. Contacts comprise a class that defines the geometric boundaries of the earth material areas. The connecting line between contact and geologic structures (Figure 4) shows that geologic structures sometimes form the contacts between different earth materials. Structural measurements are subdivided into subclasses that have measured

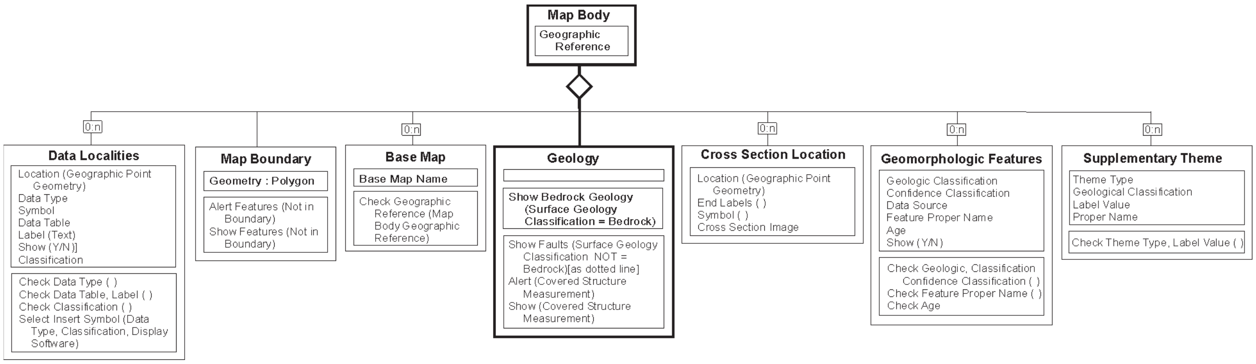


Figure 3. Geologic map body part of the DGGS data model, expressed as a UML class diagram. Each box represents a class. Within each class there are two sub-boxes; the upper sub-box is a list of attributes of the box, the second is a list of required actions.

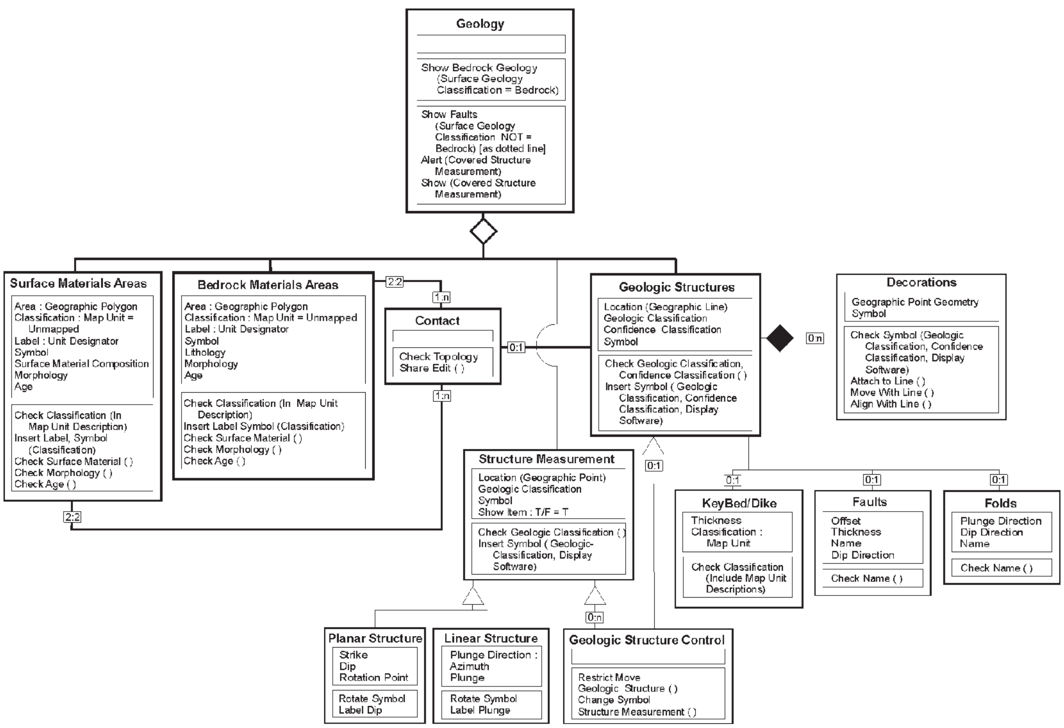


Figure 4. Subcomponents of the geology class of DGGS data model, expressed as a UML class diagram. Each box represents a class, within each class there are two sub-boxes; the upper sub-box is a list of attributes of the box, the second is a list of required actions. This diagram shows the geometric, logical, and content rules developed by DGGS staff for implementation into a geologic map database.

orientations and depict either planar or linear structures. Geologic structures have sub-classes including key beds, faults, and folds each with special characteristics.

Science and cartographic rules that apply to individual classes are listed in Figures 3 and 4 as actions, these rules apply only to the class, such as the rule: "geologic material classification must be listed in the description of map units" is represented as a "Check Classification" action. The connecting lines between classes depict science and cartographic rules between two classes. For example, a rule "lines representing contacts must be present on the boundary between two geologic material classification areas" is represented by the lines connecting the Contact class and the Bedrock and Surface materials areas classes.

Neither the process of generating a geologic map from empirical data nor the relationships between geologic map data and empirical data are represented in the model. For example, the trace of a fault on a geologic map is in part interpreted from geophysical data, spatial and time relationships between two adjoining geologic materials, and part from the author's intuition; the DGGS model only captures that information in the Geologic Classification and Confidence Classification of the fault.

IMPLEMENTATION

The ESRI personal geodatabase model was used for DGGS mapping projects completed during 2004 (Athey and Craw, 2004). The geodatabase model was chosen because it supports geographically referenced polygon, line, and point geometries as feature classes. Additionally, spatial relationship rules between and within geometric classes can be enforced using topology classes, and logical rules between classes can be enforced using joins and relationship classes. Finally, the geodatabase can be used in a single user environment or in a multiple user environment (e.g., in ArcSDE, or Spatial Database Engine), allowing us the flexibility to integrate the geologic map database into our comprehensive agency-wide database and to take the same data structure and user interface out to remote field locations.

In the implementation of the DGGS geologic map data model in the geodatabase framework, each of the Geology subcomponent classes (Figure 4) was created as a feature class. Feature classes related by theme and geometric association are grouped into feature datasets (Figure 5). For example, the geodatabase used for the 2004 mapping projects (Figure 5) contains a "bedrock" feature dataset which includes feature classes "bedpolys" for bedrock materials areas, "bedcont" for contacts and faults, and "livfold_polyline" for folds. Geometric rules between contacts and polygons were enforced within the "bedrock" feature dataset using a topology class. This initial implementation did not include the use of relationship classes or domains to control the scientific language. DGGS plans to make a more robust use of topologic rules, relationship classes, and domains to ensure data integrity in future mapping projects.

Once the geologic map data model is fully developed and tested in the personal geodatabase, we will implement the model in an agency-wide spatial database system. The map data will be fully integrated with publications index information, dataset index information, field locality, analytical data, lexicon control, and project records contained in an agency-wide geologic map database (Freeman and others, 2002). This data integration will allow DGGS staff geologists to access geologic map features and analytical data from a single data repository while working at their networked desktop computers. When this geologic data integration is complete, DGGS staff geologists will be able to conduct spatial analysis and create and edit new geologic map data in a shared, multi-user environment.

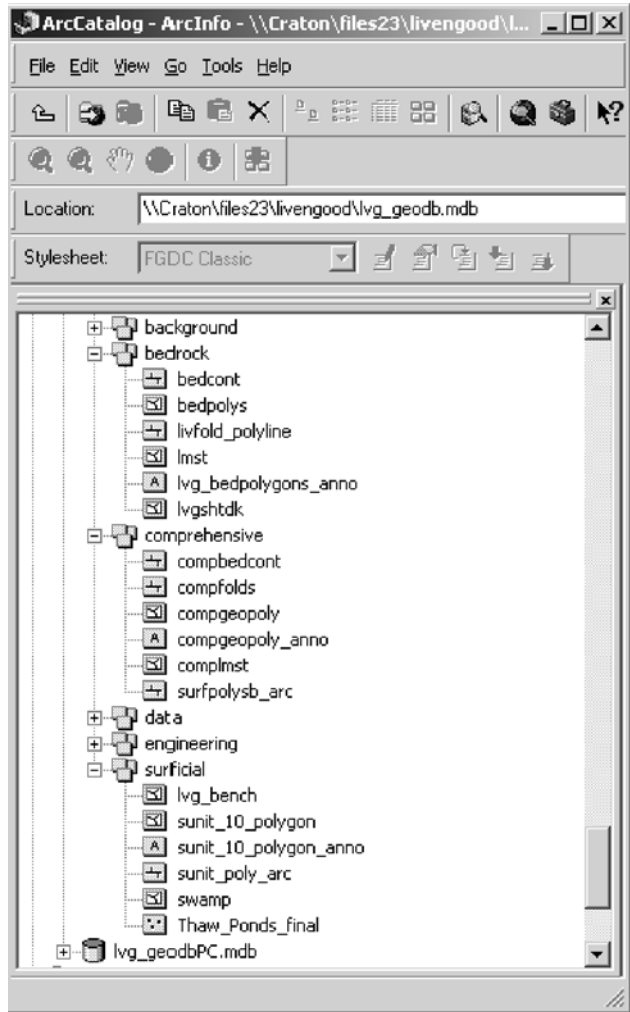


Figure 5. Screen shot of an ArcCatalog navigation tree of a geologic map database used for the 2004 DGGS geologic map products.

REFERENCES

Athey, J.E., and Craw, P.A., 2004, Geologic maps of the Livengood SW C-3 and SE C-4 quadrangles, Tolovana mining district, Alaska: Alaska Division of Geological & Geophysical Surveys, Preliminary Investigative Report 2004-03, 24 p.

Athey, J.E., Werdon, M.B., Newberry, R.J., Szumigala, D.J., Craw, P.A., and Hicks, S.A., 2004, Geologic map of the Livengood SW C-3 and SE C-4 quadrangles, Tolovana mining district, Alaska: Alaska Division of Geological & Geophysical Surveys, Preliminary Investigative Report 2004-03a, 1 sheet, 1:50,000 scale.

Browne, C.L., Freeman, L.K., and Graham, G.R.C., 2003, The Alaska Division of Geological & Geophysical Survey's Metadata Policy Development and Implementation, *in* Soller, D.R., ed., Digital Mapping Techniques '03—Workshop Proceedings: U.S. Geological Survey, Open-File Report 03–471, p. 201–208, accessed at http://pubs.usgs.gov/of/2003/of03-471/browne/index.html.

Davidson, Gail, 1998, Can we get there from here? Experiences of the Alaska Division of Geological and Geophysical Surveys, *in* Soller, D.R., ed., Digital Mapping Techniques ‘98—Workshop Proceedings: U.S. Geological Survey, Open-File Report 98-487, p. 13–15, accessed at http://pubs.usgs.gov/of/of98-487/davidson.html.

Davidson, Gail, Staft, Lauren, and Daley, E.E., 2002, The Alaska DGGS Scanning Project: Conception, Execution, and Reality, *in* Soller, D.R., ed., Digital Mapping Techniques ‘02—Workshop Proceedings: U.S. Geological Survey, Open-File Report 02-370, p. 105–109, accessed at http://pubs.usgs.gov/of/2002/of02-370/davidson.html.

Fowler, Martin, and Scott, Kendall, 2000, UML distilled: a brief guide to the standard object modeling language, second edition: Boston, Addison Wesley Longman, Inc., 186 p.

Freeman, Larry, 2001, A Case Study in Database Design: The Alaska Geologic Database, *in* Soller, D.R., ed., Digital Mapping Techniques ’02—Workshop Proceedings: U.S. Geological Survey, Open-File Report 01-223, p. 31-34, accessed at http://pubs.usgs.gov/of/2001/of01-223/freeman.html.

Freeman, Larry, Engle, Kathryn, and Browne, Carrie, 2002, Alaska Division of Geological & Geophysical Surveys Geologic Database Development—Logical Model, *in* Soller, D.R., ed., Digital Mapping Techniques ’02—Workshop Proceedings: U.S. Geological Survey, Open-File Report 02-370, p. 157–160, accessed at http://pubs.usgs.gov/of/2002/of02-370/freeman.html.

Jackson, J.A., ed., 1997, *Glossary of Geology*, 4th edition: Alexandria, Virginia, American Geological Institute, 769 p.

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 65,000 map products published by about 270 organizations including the U.S. Geological Survey (USGS), 45 state geological surveys, universities, and scientific societies and organizations,
- an extension of the Map Catalog, the Geologic Map Image Library, has evolved into a useful collection of about 1000 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 45,000 times by 17,000 users each month. 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 eighth annual Digital Mapping Techniques workshop, bringing together about 100 technical experts from 40 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.

INTRODUCTION

This project provides an unusual if not unique opportunity to foster better relations and technical collaboration

[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, 2003b) in order to provide a comprehensive.

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 advise and contribute 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 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

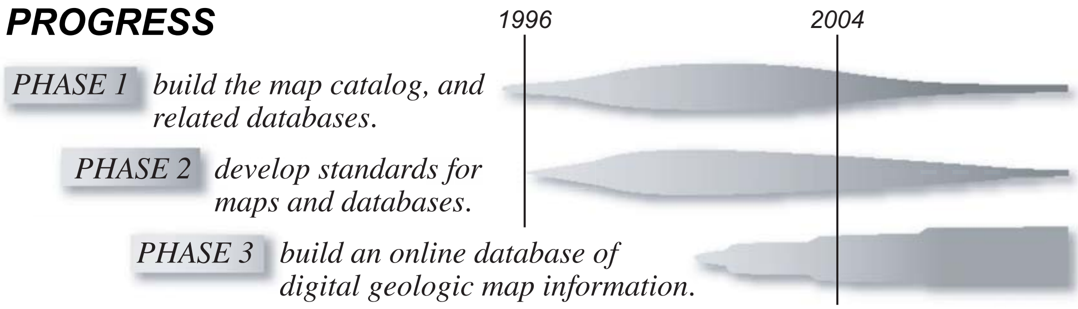


Figure 1. Diagram showing the three NGMDB Phases, and progress toward our goals (for example, documenting in the Geoscience Map Catalog all maps and related products for the United States and its territories and possessions).

‘04 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-2003. 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 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 65,000 products from about 270 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, 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 3000 published, downloadable products of the USGS and the state geological surveys. These links are established only to stable Web pages that provide the official

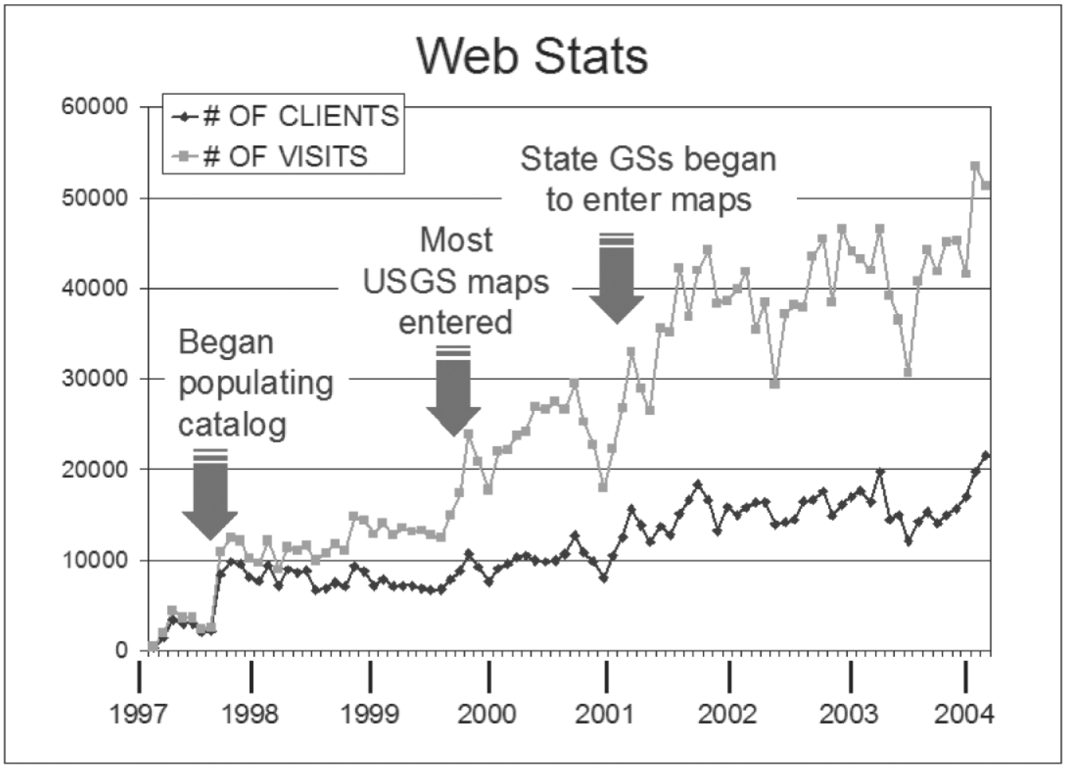


Figure 2. Web usage for the Geoscience Map Catalog, GEOLEX, Image Library, and Mapping in Progress Databases. This diagram shows that the number of people (actually, the number of unique IP addresses or computers) using the NGMDB has gradually increased as these resource databases become more widely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

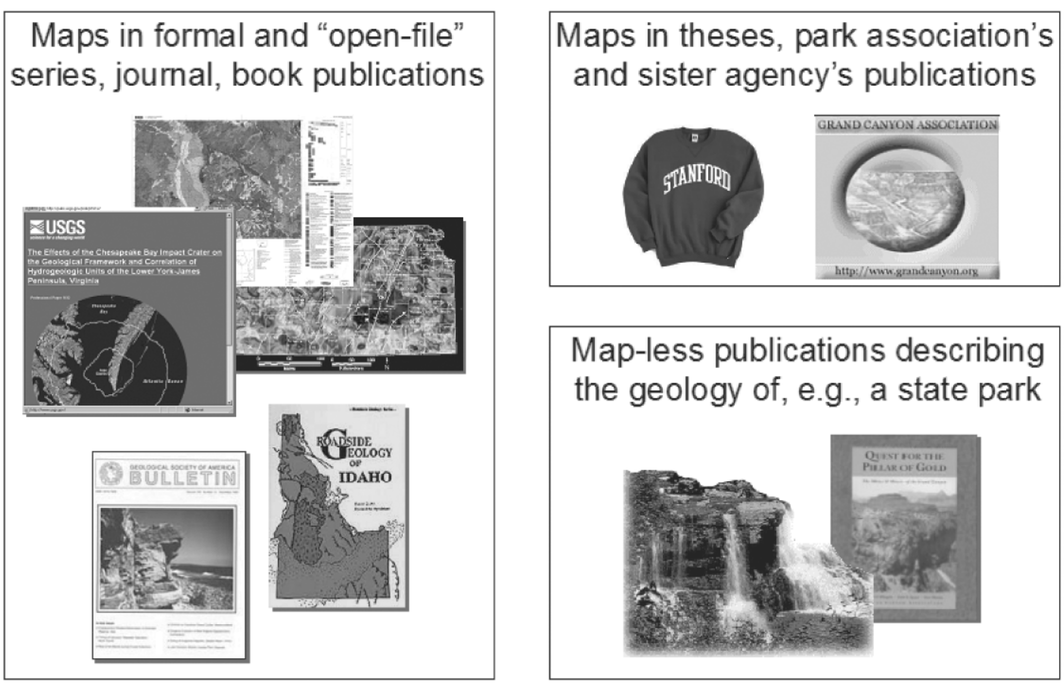


Figure 3. Bibliographic records in the Geoscience Map Catalog are drawn from a diverse group of more than 320 publishers.

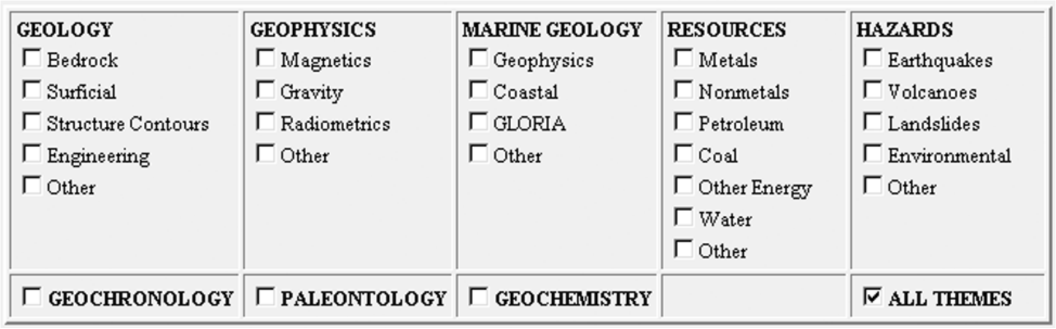


Figure 4. A portion of the Geoscience Map Catalog search page, showing the types of products included.

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 Geologic Map Image Library (Soller and Berg, 2003a). 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 become more fully integrated with the Geoscience Map Catalog because: 1) the Image Library's database is based on a subset of the 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

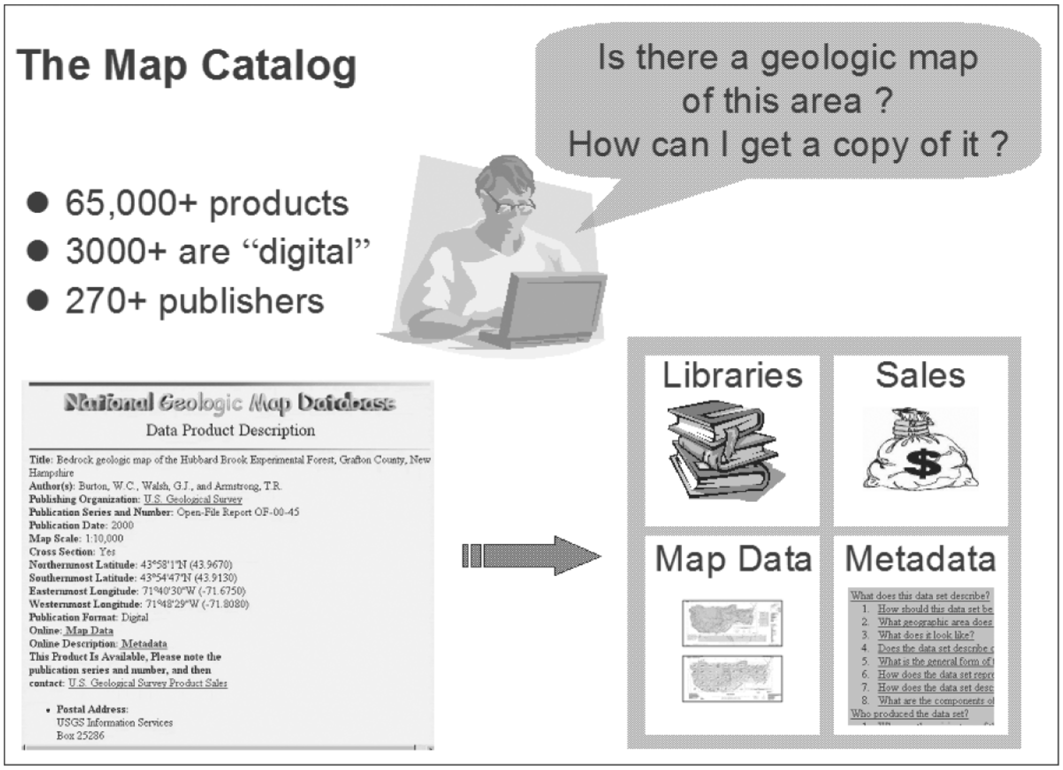


Figure 5. Interested in knowing something about the geology of an area (such as the land beneath their house), the user queries the Geoscience Map Catalog, which returns a hit list of possibly useful maps and related products. The user selects one of these and, from the Product Description Page (shown 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.

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. In 2005 we will be re-engineering and repopulating this database.

Paleontology Database “I want to know if there is any fossil data from this area…”

The NGMDB project has designed, and is planning to 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 have been delayed somewhat, in order to assess what can be accomplished with funding and personnel resources more modest than earlier anticipated. We now envision a system that: 1) includes new data provided by the NGMDB, and 2) archives and serves modestly-sized 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 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

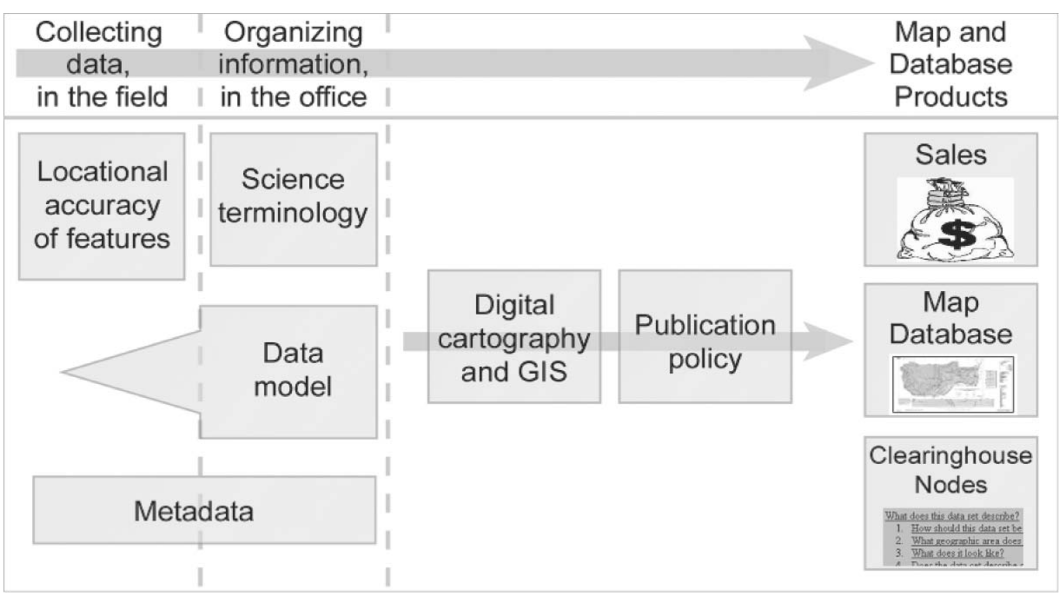


Figure 6. Diagram showing how the standards and guidelines under development by the NGMDB and related groups relate to the process of creating and publishing a map and database.

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,
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), and
4. development of a map database and standards Clearinghouse (http://ngmdb.usgs.gov/intdb/) that is endorsed by the IUGS CGI and the International Association for Mathematical Geology (http://www.iamg.org/).

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 is being prepared for submittal to FGDC, for final discussion and adoption as a Federal standard. This process is expected to conclude in 2005. 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 eight 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 Portland, Oregon, and hosted by the Oregon Department of Geology and Mineral Industries, was attended by about 100 representatives of 40 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 had 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,

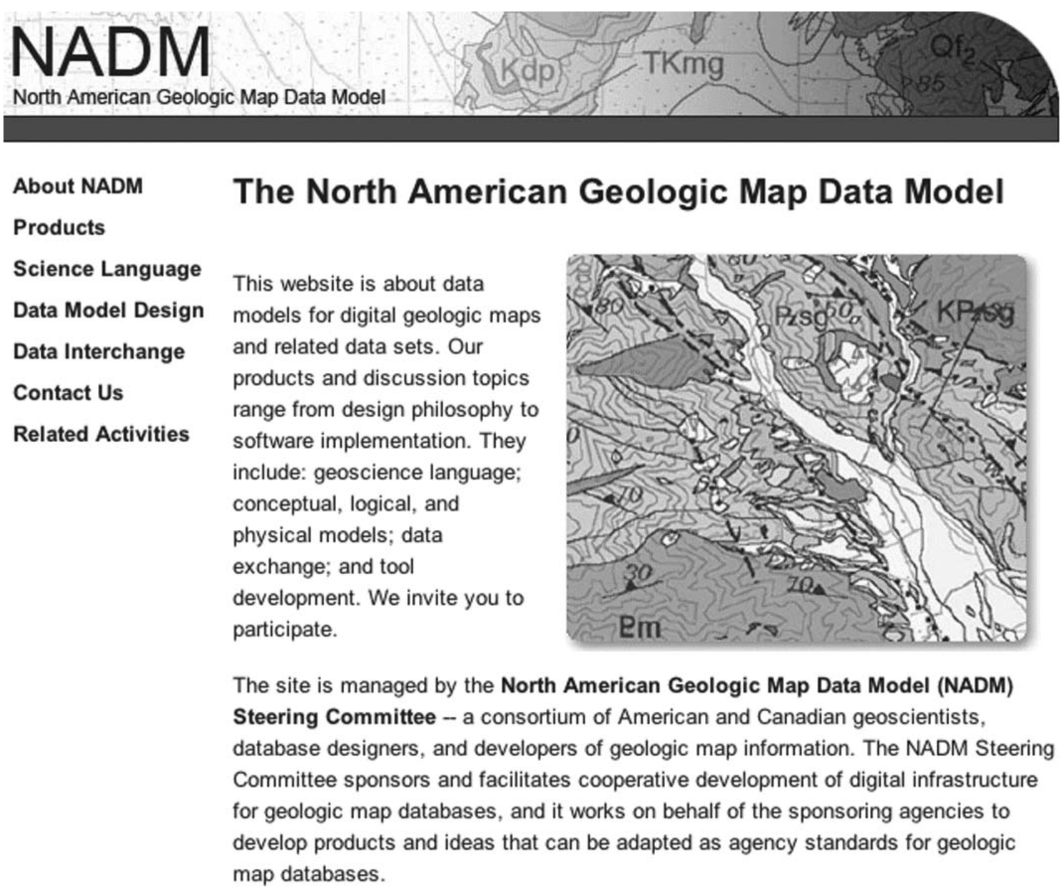


Figure 7. Website of the North American Geologic Map Data Model Steering Committee.

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 a significant accomplishment—a conceptual data model known as NADM C1 (NADM, 2004); this model was published simultaneously by the USGS and GSC, and is available at http://pubs.usgs.gov/of/2004/1334/. 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 standard science language that accompanies it. Information about NADMSC activities is provided in two papers in this volume: 1) the development of a GML-based interchange format (see Boisvert and others); and 2) the development of standard science language to describe the lithology of earth materials (see NADM SLTT).

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 has begun 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, using existing and new mapping in disparate field areas (for example, central Arizona, northern Virginia, Kentucky, and southern California). 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 SLTT, in these Proceedings).

The data-entry tool is being designed as a standalone application that will connect to a relational database that implements the NGMDB design (see Richard and others, 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. We anticipate that this tool will be available in 2005 for our cooperators to evaluate and use.

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:

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 sys-

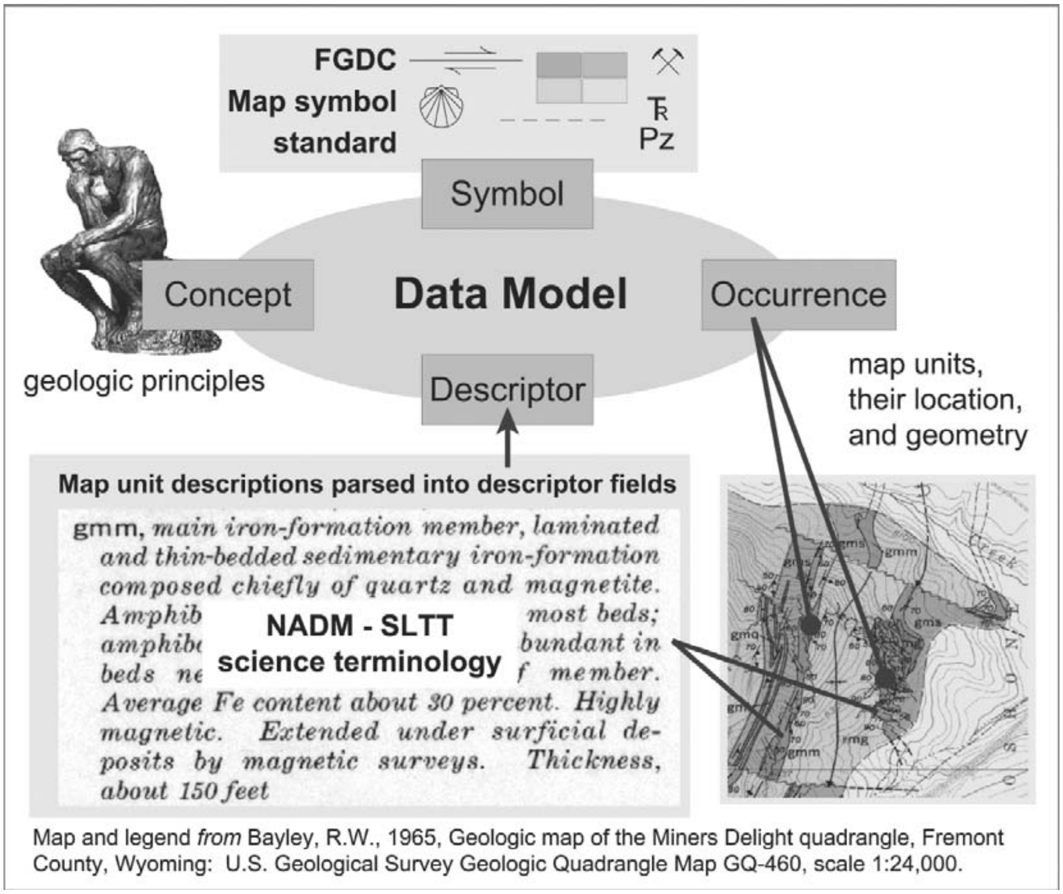


Figure 8. Simplified representation of the data model and its application to a typical, 2-D geologic map. The presence of a geologic unit on the map, referred to in the data model as an “occurrence” of that map unit, is described by: 1) its bounding contacts and faults, whose coordinates are stored as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

tem 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, 2003). 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 will be used to print the map, in order to engage GSA in a plan to develop a database for the map. When compilation and review of the map has been completed, and the map printed in early 2005, we will propose a database design and begin to populate the digital files made available from cartographic production of the map. This work will be conducted in collaboration with GSA and interested national geological surveys.

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).

ACKNOWLEDGMENTS

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 manage-

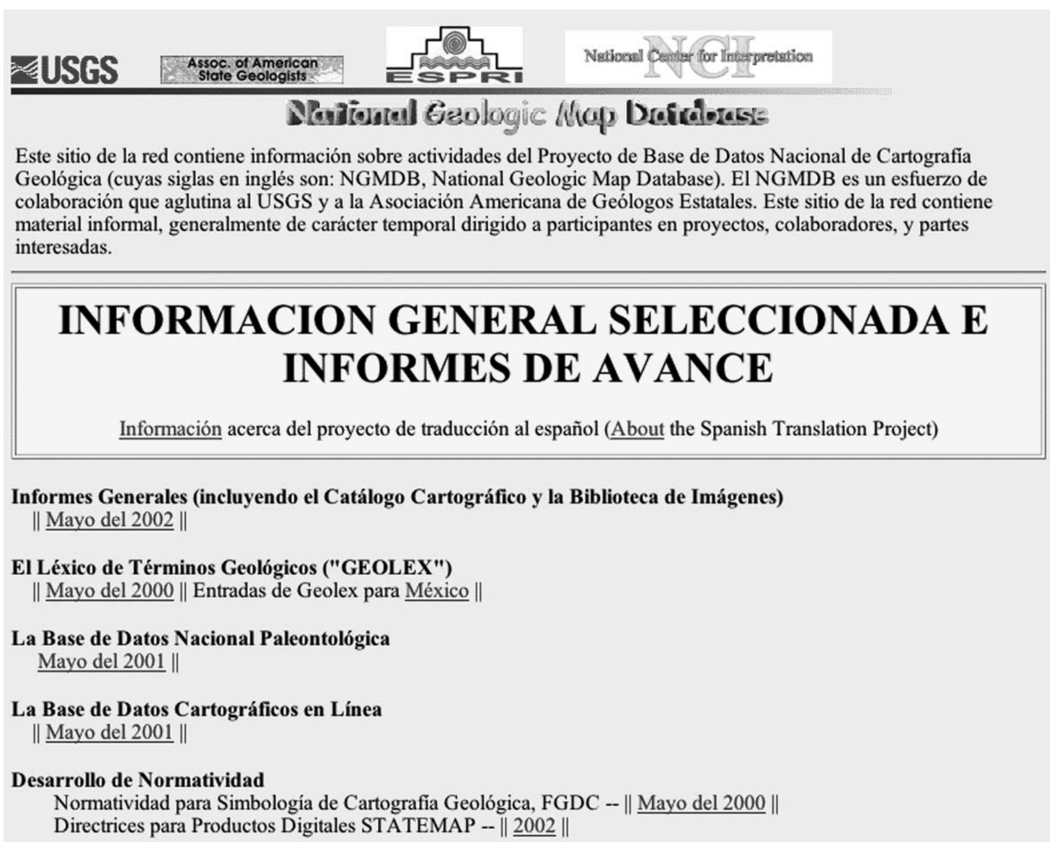


Figure 9. The NGMDB Spanish language website, containing translations of selected technical reports and geologic names of Mexico as found in GEOLEX.

ment), Chuck Mayfield and Nancy Blair (USGS Library; Map Catalog content), Bruce Wardlaw (Paleontology Database), Robert Wardwell (USGS, Vancouver, WA; Image Library) and Kevin Laurent, Jeremy Skog, and Ben Carter (USGS, Reston, VA; Image Library), Steve Richard (Arizona Geological Survey, Tucson, AZ; data model and science terminology), Jonathan Matti (USGS, Tucson, AZ; data model and science terminology), and Jordan Hastings (USGS, Santa Barbara, CA; data model).

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

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, 2004, 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/.

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.

Soller, D.R., and Berg, T.M., 2003a, 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 Berg, T.M., 2003b, 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 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)

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)

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 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

AASG/USGS Metadata Working Group:
Peter Schweitzer (U.S. Geological Survey and 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 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 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)

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)
Jonathan Matti (U.S. Geological Survey and Chair of the Science Language Technical Team)
Scott McColloch (West Virginia Geological and Economic Survey)
Steve Richard (Arizona Geological Survey)
Peter Schweitzer (U.S. 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.

APPENDIX B

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

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., 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.

Migrating from ArcInfo Workstation to ArcGIS

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

INTRODUCTION

During the past ten years, geological maps at the Geological Survey of Canada (GSC) have been produced using ESRI's ArcInfo Workstation software along with a custom application called GEMS (GEological Mapping System – available at http://www.nrcan.gc.ca/ess/carto/toolbox_e.html). Supporting this application in delivering high-quality publications, Cartographic Digital Standards and Cartographic Design Specifications are used to manage the spatial data and surround elements that appear on the map, respectively.

With the recent introduction of ArcGIS and the benefits that this new technology offers, considerable re-tooling of the process will be required to achieve the product quality that we achieved with ArcInfo Workstation. This will cover all aspects of production, from data management to developing and implementing new applications. From a cartographic perspective, this document will explain the migration strategies for developing a geodatabase design to manage the data, as well as some ideas for producing maps with ArcMap. The content of this document is based on the ESRI product ArcGIS 8.3. General knowledge of ArcGIS, particularly geodatabase elements and design, will be beneficial.

MIGRATING TO ARCGIS

Goals

One of our main goals is to be able to produce geological maps with ArcMap as effectively and efficiently as with ArcInfo Workstation. The majority of this effort has involved designing a geodatabase schema to store and manage the geological features that constitute a map. The geodatabase also provides the means for feature-validation through the use of domains for standardizing feature names and user-defined topologies. Domains can be used to restrain the user in assigning values to an attribute field from an established list of coded values or a numeric range. Topologies will ensure users adhere to the established spatial relationships between features.

Another goal is to keep the design of the geodatabase simple and easy to understand. This will be achieved by eliminating unnecessary data fields, using intuitive and obvious field names, implementing the use of domains, and using tables and defined relationships where necessary.

Migration Strategy

The gradual migration from ArcInfo to ArcGIS can be characterized as occurring in three stages. These are outlined in ESRI's help documentation, and are summarized below:

1. The first stage is to use ArcGIS for data use and visualization, not data management. Workstation would still be used to edit and maintain coverage data, but ArcGIS's ArcMap would be used for producing the geological map.
2. The second stage is the most complicated. This involves moving the data management from coverage-based files used in ArcInfo Workstation to the geodatabase model of ArcGIS. Designing an effective geodatabase schema will require prototype implementations to determine the most appropriate design.
3. The final stage revolves around the geo-processing tools that will be introduced in ArcGIS 9.0. These tools will be used to create applications to upgrade data from Workstation to ArcGIS.

In our migration strategy, we decided to merge steps 1 and 2, delivering both at the same time. The primary reason for this approach is to eliminate the transition phase that could affect our service delivery of map products. Using two software suites to produce maps and manage data could potentially create bottlenecks, and data management and versioning issues that could affect the timely delivery of map products (both paper and digital). In addition, based on past experiences, it will be easier to train everyone at the same time to make the transition.

DESIGNING A GEODATABASE

A geodatabase contains the geographic or spatial features and their attributes, which can be used to produce a geological map. Geographic features of a homogenous theme and spatial type (e.g., contacts, which are lines) are stored in feature classes. Feature classes that are physically related to one another (e.g., contacts, faults, and geological polygon units) are best to be grouped in a feature dataset. This grouping permits the application of user-defined topological rules that express the relationships between each of the feature classes. The feature name, definitions, attribute fields, and properties of these feature classes determine the schema or design used in the geodatabase. There are three basic methods in designing a geodatabase schema:

1. The first is using ArcCatalog's menus and wizards. This is probably the simplest method of creating a geodatabase schema. The geodatabase, feature datasets, feature classes, relationship classes, and tables can all be created using these wizards. Domains can also be entered as properties of the geodatabase. All the menus are simple to use and navigate, yet some general understanding is required when entering all of the geodatabase information. After becoming familiar with designing geodatabases, the procedures become very repetitive and somewhat rudimentary. Transferring or sharing the geodatabase schema and its properties with other geodatabases, however, is not possible. Even the smallest of changes will require extensive re-entry.
2. The second method uses ArcCatalog's CASE tool to import a schema that is created by another software, such as Visio. Although creating schemas using UML (Unified Modeling Language) is the preferred choice for developers, access to this type of software was unfortunately not possible for us, therefore this option was not investigated. One thing to point out, however (which is also the case for the first method), is that existing geodatabase schemas cannot be exported from ArcCatalog and used in CASE tools. Any changes must be performed using the CASE software and then the schema must be imported into ArcCatalog. The CASE tool is expensive and is not available for ArcView licenses.
3. The third method uses Geodatabase Designer to define the geodatabase schema. The main benefit of this tool is that it can take an existing geodatabase and export its schema to XML (extensible mark-up language). Changes to the XML (the geodatabase schema) can be performed in an XML editor and then imported either into the same geodatabase or to a new one. Geodatabase Designer is fully integrated with ArcCatalog, and free to download from ESRI's Support Center website (http://support.esri.com/index.cfm?fa=downloads.gateway). At that site, search for "geodatabase designer"; however, it must be noted that it is an unsupported application.

Designing a Geodatabase for Map Production

Our map products are either bedrock or surficial maps, and so we decided to tailor the design of geodatabases to this type of information. Therefore, separate geodatabases will be created for bedrock and surficial geological datasets. This may seem redundant, but there are areas where both thematic features have been mapped, either on different maps or on one map. Storing the data in separate geodatabases will provide a better means for managing the data as well as referencing it when making maps with ArcMap. Some of the criteria we will use for creating a geodatabase schema are as follows:

- create separate geodatabase schemas for bedrock geology data, surficial geology data, and cartographic elements (non-geological features like borders, index maps)
- do not store data in a geodatabase if it already exists in another database and is accessible when making a map (e.g., geochronology database)
- avoid duplication and redundancy in feature attribute values by retaining only the fields that are necessary, and use tables where appropriate to expand attribute meaning (e.g., a single field can be used to define each geological polygon, which in turn relates to another table containing information of each unit)
- use domains and topology to validate features
- use sub-types along with domains to group/categorize simple features
- use relationship classes to define relates between tables.

Managing Geological Polygons and Legend Information

At this time, much of the effort at GSC has involved implementing a surficial geological geodatabase; however, there will be many aspects applicable to a bedrock geological geodatabase. In the design of a surficial geodatabase schema, the first feature we examined was the geological polygons. The surficial map sample from ArcMap (Figure 1) displays one of the methods authors from the GSC use to classify geological polygons. Many of the surficial geological polygons are labeled on the map with one or two units, where each unit is listed and described separately in the legend. These compound units (e.g. "Cv/Tm") can consist of veneers (in this case, "Cv")

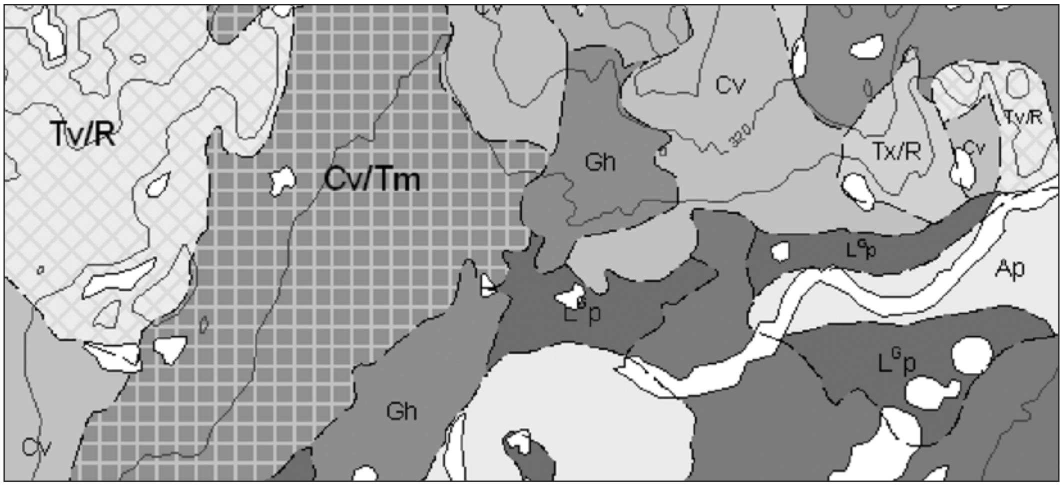
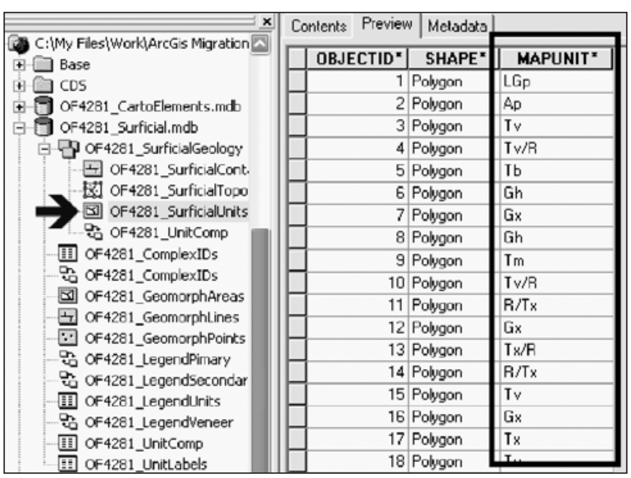


Figure 1. Sample black & white image of a colour surficial map from ArcMap.

overlying a dominant unit ("Tm") or as a proportional mix between two geological units (Tx/R). In most cases, veneers include the lowercase character "v", and for compound units, a symbol is used between the two geological units to express the type of relationship and ratio of each unit. These methods are intended as a visual aid for the map-reader, but a different approach is required to better manage this type of information and structure in a geodatabase.

The values assigned to the geological polygons, including the relationship information for these compound units, is stored in the attribute field MAPUNIT, in the polygon feature class table (Figure 2). The values do not necessarily have to match the polygon label on the map.



OBJECTID*	SHAPE*	MAPUNIT*
1	Polygon	LGp
2	Polygon	Ap
3	Polygon	Tv
4	Polygon	Tv/R
5	Polygon	Tb
6	Polygon	Gh
7	Polygon	Gx
8	Polygon	Gh
9	Polygon	Tm
10	Polygon	Tv/R
11	Polygon	R/Tx
12	Polygon	Gx
13	Polygon	Tx/R
14	Polygon	R/Tx
15	Polygon	Tv
16	Polygon	Gx
17	Polygon	Tx
18	Polygon	[illegible]

Figure 2. Snap shot from ArcCatalog. The left-hand side displays the catalogue tree, with the arrow pointing to the currently selected polygon feature class "OF4281_SurficialUnits". The attributes of these features are displayed on the right-hand side. The bold outline contains the attribute values of each feature in the MAPUNIT field.

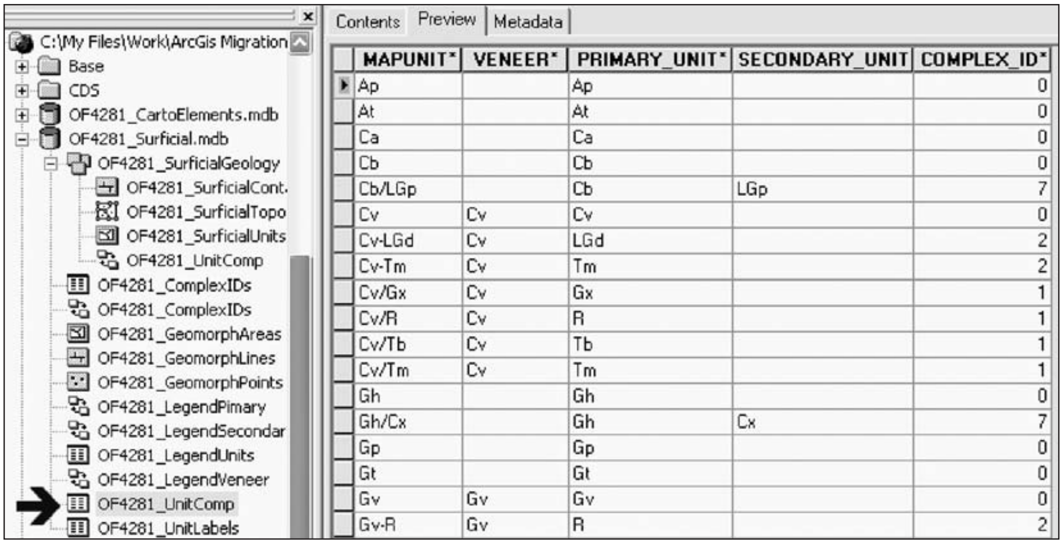
In fact, they can be any value, string or numeric, as the value relates to another table that describes the composition of each unique polygon value as well as the label to be placed on the map.

The table UnitComp (Figure 3) is used to describe the composition of each unique polygon value. In this table, a field also named MAPUNIT contains all the unique polygon values from the above polygon feature class. Three additional fields are used to define the composition. The VENEER field contains geological units that are classified as veneers. The PRIMARY_UNIT and SECONDARY_UNIT fields contain geological units that are classified as the primary and secondary units respectively. The values in these three fields relate directly to the geological units that appear in the legend.

A fourth field named COMPLEX_ID is used to identify the relationship or proportional mix between each of the geological units in a composition as defined by the author. The integer values in this field relate to the ComplexIDs table (Figure 4), where each type of relationship or complex ID is explained. On the map, symbols such as · (middle dot), - (dash), / (slash), and // (double slash) are used as visual aids to express these relationships (e.g., the middle dot between two units is used to express a 50/50 proportional mix between a primary and secondary unit).

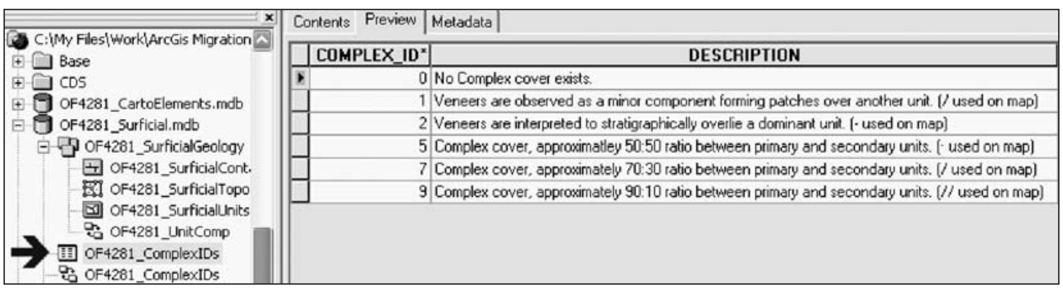
Another table, called LegendUnits (Figure 5) contains all the geological units that appear on the legend and provide the definitions for all surficial geologic units listed in the UnitComp table. In the LegendUnits table, each geological unit contains a description, as well as heading information (not shown in figure) that would appear on the legend (e.g., Glacial Till).

The final table associated with geological polygons is used for dynamically labeling the polygons in ArcMap. This table, called UnitLabels (Figure 6), also contains all unique polygon values for the field MAPUNIT found in the above polygon feature class. The field MAPLABEL



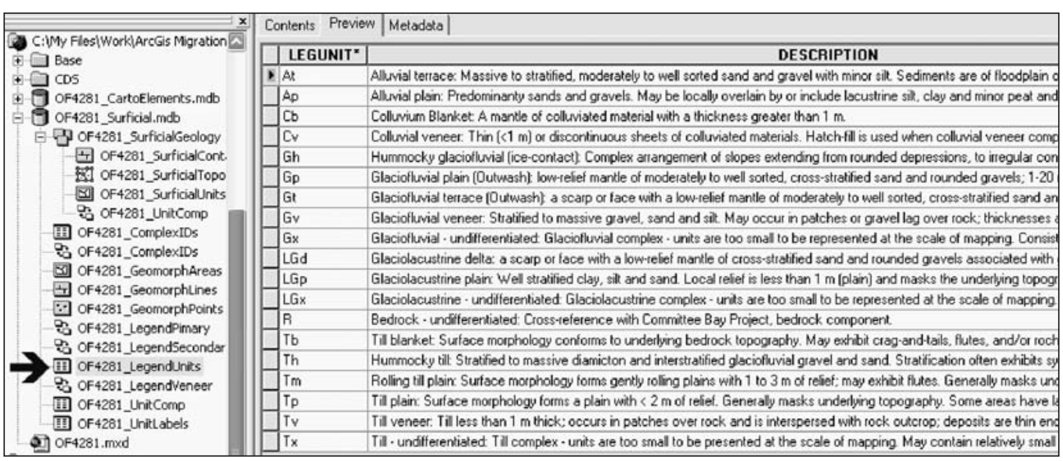
MAPUNIT*	VENEER*	PRIMARY_UNIT*	SECONDARY_UNIT	COMPLEX_ID*
Ap		Ap		0
At		At		0
Ca		Ca		0
Cb		Cb		0
Cb/LGp		Cb	LGp	7
Cv	Cv	Cv		0
Cv-LGd	Cv	LGd		2
Cv-Tm	Cv	Tm		2
Cv/Gx	Cv	Gx		1
Cv/R	Cv	R		1
Cv/Tb	Cv	Tb		1
Cv/Tm	Cv	Tm		1
Gh		Gh		0
Gh/Cx		Gh	Cx	7
Gp		Gp		0
Gt		Gt		0
Gv	Gv	Gv		0
Gv-R	Gv	R		2

Figure 3. Snap shot from ArcCatalog. The arrow is pointing to the currently selected table "OF4281_UnitComp", which is displayed to the right. This table contains all unique polygon values in the field MAPUNIT. Subsequent fields further define the composition of each unique polygon value and the nature of the composition.



COMPLEX_ID*	DESCRIPTION
0	No Complex cover exists.
1	Veneers are observed as a minor component forming patches over another unit. (/ used on map)
2	Veneers are interpreted to stratigraphically overlie a dominant unit. (- used on map)
5	Complex cover, approximatley 50:50 ratio between primary and secondary units. (· used on map)
7	Complex cover, approximately 70:30 ratio between primary and secondary units. (/ used on map)
9	Complex cover, approximately 90:10 ratio between primary and secondary units. (// used on map)

Figure 4. Snap shot from ArcCatalog. The arrow is pointing to the currently selected table "OF4281_ComplexIDs", which is displayed to the right. This table lists the different types of relationships for geological units that participate in a composition.



LEGUNIT*	DESCRIPTION
At	Alluvial terrace: Massive to stratified, moderately to well sorted sand and gravel with minor silt. Sediments are of floodplain o
Ap	Alluvial plain: Predominanty sands and gravels. May be locally overlain by or include lacustrine silt, clay and minor peat and
Cb	Colluvium Blanket: A mantle of colluviated material with a thickness greater than 1 m.
Cv	Colluvial veneer: Thin (<1 m) or discontinuous sheets of colluviated materials. Hatch-fill is used when colluvial veneer comp
Gh	Hummocky glaciofluvial (ice-contact): Complex arrangement of slopes extending from rounded depressions, to irregular con
Gp	Glaciofluvial plain (Outwash): low-relief mantle of moderately to well sorted, cross-stratified sand and rounded gravels; 1-20
Gt	Glaciofluvial terrace (Outwash): a scarp or face with a low-relief mantle of moderately to well sorted, cross-stratified sand an
Gv	Glaciofluvial veneer: Stratified to massive gravel, sand and silt. May occur in patches or gravel lag over rock; thicknesses
Gx	Glaciofluvial - undifferentiated: Glaciofluvial complex - units are too small to be represented at the scale of mapping. Consist
LGd	Glaciolacustrine delta: a scarp or face with a low-relief mantle of cross-stratified sand and rounded gravels associated with
LGp	Glaciolacustrine plain: Well stratified clay, silt and sand. Local relief is less than 1 m (plain) and masks the underlying topogr
LGx	Glaciolacustrine - undifferentiated: Glaciolacustrine complex - units are too small to be represented at the scale of mapping.
R	Bedrock - undifferentiated: Cross-reference with Committee Bay Project, bedrock component.
Tb	Till blanket: Surface morphology conforms to underlying bedrock topography. May exhibit crag-and-tails, flutes, and/or roch
Th	Hummocky till: Stratified to massive diamicton and interstratified glaciofluvial gravel and sand. Stratification often exhibits sy
Tm	Rolling till plain: Surface morphology forms gently rolling plains with 1 to 3 m of relief; may exhibit flutes. Generally masks und
Tp	Till plain: Surface morphology forms a plain with < 2 m of relief. Generally masks underlying topography. Some areas have la
Tv	Till veneer: Till less than 1 m thick; occurs in patches over rock and is interspersed with rock outcrop; deposits are thin end
Tx	Till - undifferentiated: Till complex - units are too small to be presented at the scale of mapping. May contain relatively small

Figure 5. Snap shot from ArcCatalog. The arrow is pointing to the currently selected table "OF4281_LegendUnits", which is displayed to the right. This table contains each geological unit and its description as it appears in the legend on the map.

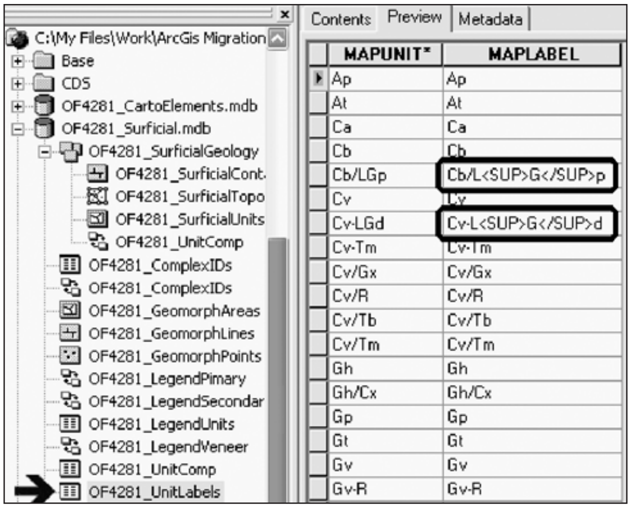


Figure 6. Snap shot from ArcCatalog. The arrow is pointing to the currently selected table "OF4281_UnitLabels", which is displayed to the right. This table is used in ArcMap for dynamically labeling polygons. The field MAPLABEL is used as the label text field for each unique polygon. The bold outline displays two examples of using text-formatting tags, in this case to achieve the superscript character.

contains the text string that will be used by ArcMap to label the polygons. Using this kind of table ensures consistency and allows the use of text formatting tags to show unit labels with superscript, subscript, bold, and coloured type (see text formatting tags in ArcMap Help for more information). In the example map in Figure 1, the geological unit $L^{G}P$ contains a superscript character (the superscript G for this unit was requested by the author for this particular map). The field MAPUNIT is defined as a text field and cannot contain any superscript or special characters, therefore the value "LGP" is used for these types of polygons. The field MAPLABEL, also defined as a text field, contains the superscript formatting tags to achieve the superscript character when labeling (circled in Figure 6).

In order to use this table in ArcMap for labeling, an ArcMap relate must be specified between the layer (polygons) and the table. A relationship class (discussed below) is not necessary, as this table is only used for labeling purposes, and not used for querying the polygon and legend information. Furthermore, when labeling polygons in ArcMap, classifications can be created based on the area of the polygons being labeled to achieve different point sizes for the labeled text strings. For example, small polygons (area ≤ 10,000 m^2), medium polygons (area > 10,000 m^2 and ≤ 100,000 m^2), and large polygons (area > 100,000 m^2) can be labeled with 8, 10, and 12 pt. type respectively. As with all GIS software, the labels are arbitrarily placed, and therefore further cartographic refinement is required. The labels can be converted to annotation and their placement adjusted to provide a more legible map. At this time, we are unsure if this table and/or the use of relates in ArcMap can be used in advanced labeling engines like the MAPLEX extension.

For data analysis and map production, these many tables must be related by creating relationship classes. A relationship class defines which fields are related between tables, as well as the nature of the relate. A total of five relates exist, each defined as a separate relationship class. Figure 7 shows the table and field names, and their relationships. You can see the relationship classes in the previous figures, preceded by the icon (two small boxes with arrows pointing to each other in the catalogue tree). These tables and relationship classes exist in the geodatabase, but they could also exist in the feature dataset containing the SurficialGeology polygon feature class

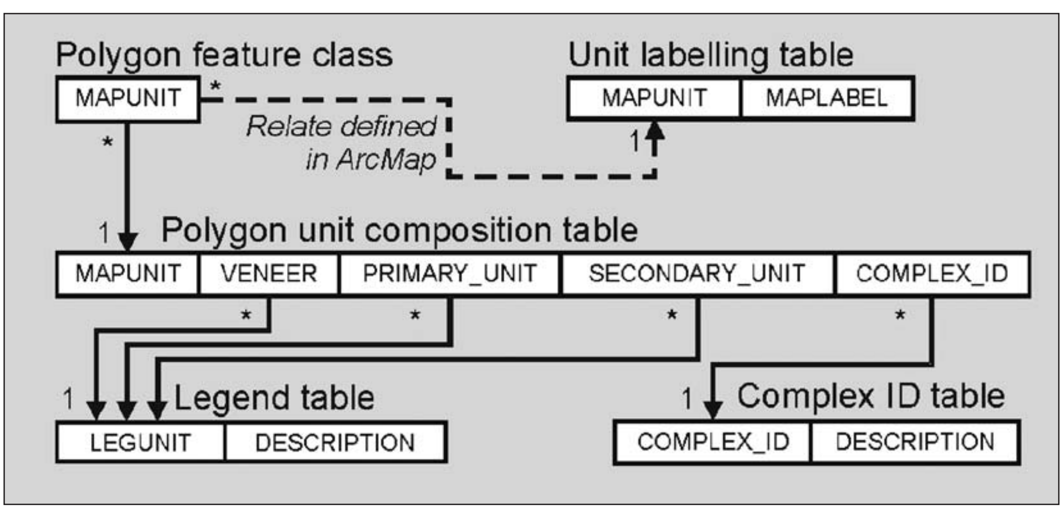


Figure 7. Diagram showing the relationship between the fields in the tables. Each line is defined as a relationship class in the geodatabase. Cardinality is expressed as 1 (one) and * (many).

to which they are related. Storing them at either location is acceptable, but it would have been better in these examples, from a file management perspective, to store the tables and relationship classes in the SurficialGeology feature dataset.

Creating relationship classes will allow end-users to query information from ArcMap (at this point it is unsure if this can also be achieved with ArcIMS, ArcReader, or ArcExplorer). By selecting a polygon, the end-user can display the attributes from all of the tables by navigating through the established relates (Figure 8).

Managing Geological Line, Point, and Other Area Features

Our current plan for managing all other surficial geological line, point, and area features on a geological map is to store them in separate, stand-alone feature classes. In the surficial geodatabase schema discussed above, these would be the feature classes GeomorphLines, GeomorphPoints, and GeomorphAreas. These feature classes would also incorporate the use of sub-types and domains to assist in categorizing and validating the features.

Domains are a property of the geodatabase, and as such can be applied to any feature class. In fact, the domains created for these features will be generic to all geodatabases. Domains are used to constrain values for a field, and can include acceptable coded values or a numeric range. In most cases, coded value domains are used to specify each type of geological feature, and numeric range domains are applied where numeric fields are used to store attribute information, such as the strike (rotation) and dip of a geological symbol. Domains are used in all custom fields in a feature class; thus validating a feature class will ensure that the attribute values of all features are acceptable and valid. In the following example (Figure 9), the Glacial domain has been created as a coded value domain that lists valid values and their descriptions.

Sub-types provide a means for categorizing or grouping features within a feature class by assigning unique values to a long-integer defined field. In the example below (Figure 10), the polygon feature class has a sub-type field named CATEGORY, with two valid values: 1 (to store glacial features), and 2 (to store glaciolacustrine features). Sub-types allow for different domains to be applied to each sub-type classification. For example, for the glacial and glaciolacustrine feature sub-types, the field FEATURE will only accept coded values from the Glacial and Glaciolacustrine domains respectively. This is highly useful, because each sub-type classification can have its own unique attribute values, which are specified by the domain. Furthermore, these domains can be applied to more than one feature class. Depending on the compilation scale, features can be stored either as points, lines, and/or areas in the respective feature classes (GeomorphPoints, GeomorphLines, GeomorphAreas). For example, drumlins are usually stored as point or line features. In both feature classes they would exist in the sub-type Glacial, and the Glacial domain would apply to the field FEATURE. Therefore, these features must be coded as “drumlin” based on the coded value from the Glacial domain, regardless of the type of feature.

User-defined Topologies

In addition to feature validation, another benefit of geodatabases is the ability to create custom topologies

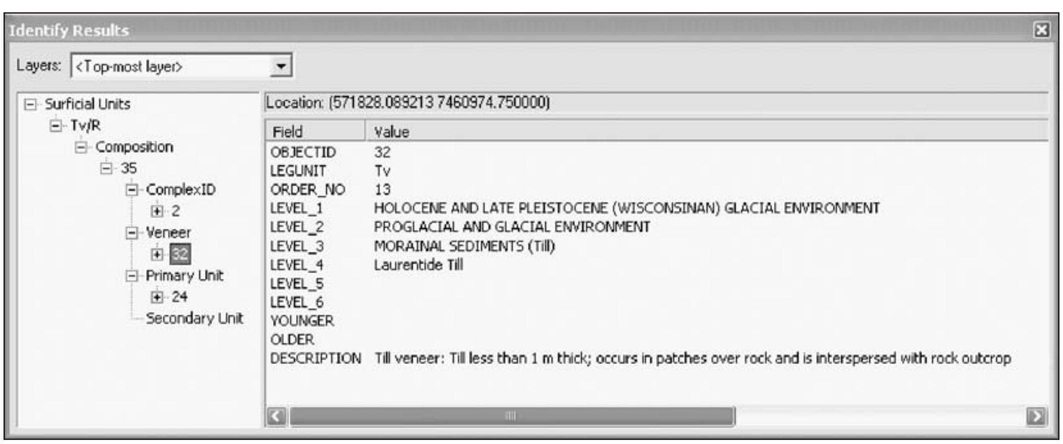


Figure 8. ArcMap snap shot displaying results of an identify procedure on a selected feature. The left-hand side displays the tables in a tree format as one migrates from the selected feature, through the established relates to other tables (just as confusing as it appears). The right-hand side displays the attributes of the geological unit “Tv” from the legend table. For this selected feature, the geological unit “Tv” also participates in a composition as a veneer with the map unit “Tv/R”.

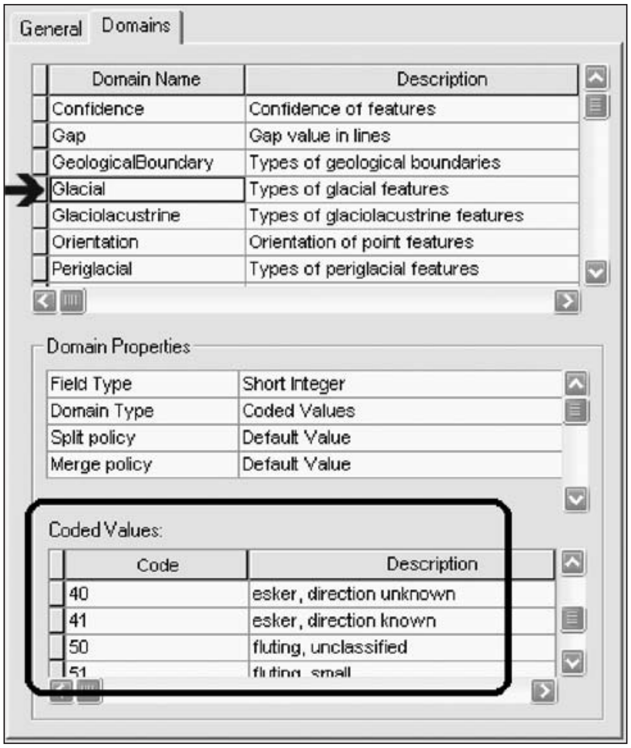


Figure 9. Snap shot from ArcCatalog displaying the domains as a property of the geodatabase. The arrow is pointing to the "Glacial" domain with valid coded values and descriptions shown below in the bold outline.

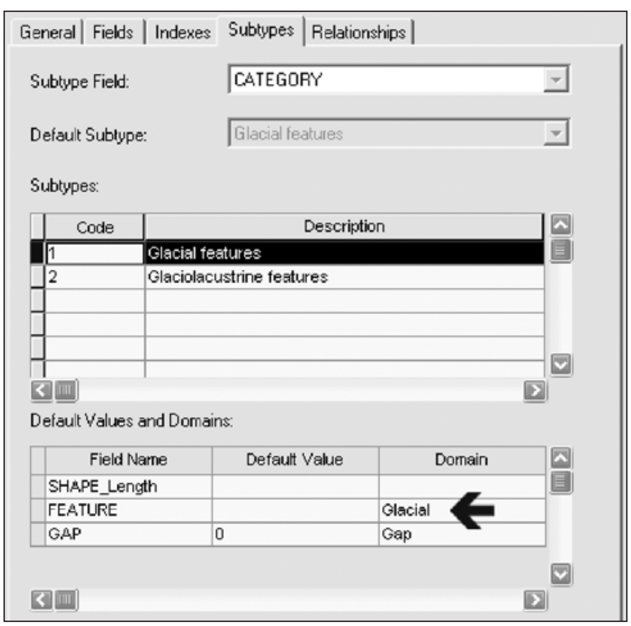


Figure 10. Snap shot from ArcCatalog displaying the subtypes as a property of a feature class. Highlighted is the "Glacial features" subtype, in which case values for the field FEATURE must adhere to the "Glacial" coded value domain.

based on feature classes that exist in a feature dataset. Custom topology involves specifying which feature classes will be participating, the relationship rules between the features, and the cluster tolerance (similar to the fuzzy tolerance from ArcInfo Workstation). For our application, user-defined topologies apply only to the geological polygons datasets, where the SurficialContacts and SurficialUnits feature classes participate. The two rules that are implemented are:

1. Polygons from the feature class SurficialUnits must not overlap,
2. The lines (contacts) from the feature class SurficialContacts must overlie the boundary of the polygons from the feature class SurficialUnits.

The second rule simply states that the contact lines must coincide with the outline of the polygons. The rule could have been expressed the other way around: the outlines of polygons must overlie the contact lines, but this would have caused excessive topological dirty areas (errors) where contact lines do not exist (i.e., at the edges of bodies of water which close some polygons).

MAKING MAPS WITH ARCMAP

Style Files and Symbols

Unfortunately, customized symbolsets from ArcInfo Workstation cannot be used directly or imported into ArcMap. This has required a significant investment of time, to rebuild the libraries for approximately 400 line and 700 point symbols. The majority of the time has been spent creating TrueType fonts containing the point symbols, using third-party software (Font Lab and Fontographer). Our current set of symbols/style files for ArcMap are available free of charge from our "Migration to ArcMap" website (http://www2.nrcan-rncan.gc.ca/ess/carto/arcgis_e.asp).

Using XML for Map Surrounds and Legends

During the past ten years of producing geological maps using ArcInfo Workstation, several customized commands have been developed to assist cartographers in placing surround information (title blocks, legends, logos, credits, descriptive notes, recommended citation, figures, etc…). Unfortunately, these commands are all written in AML (Arc Macro Language), and as such cannot be used to position these elements, or to aid in the layout of a map using ArcMap. In order to provide the same level of functionality, scripts will have to be written using VBA (Visual Basic for Applications) and ArcObjects. In addition, the information associated to each of the surround elements will be stored in XML. The source of this information will either be obtained from existing metadata or inputted by the cartographer. Storing this information in XML will not only assist in the placement of these surround elements, but the XML file can be used by third parties to extract the information about the map they require. This effort is currently being developed, and is not yet incorporated in the production of maps with ArcMap.

Another aspect of the map surround information is the geological legend. These legends, created using ArcInfo Workstation, also rely on AML programming. A similar approach will also be taken in producing legends using ArcMap, utilizing XML to store content derived from the legend tables. Much of this work is just beginning.

ACKNOWLEDGMENTS

I wish to acknowledge the efforts of Roger MacLeod and Parm Dhesi, in migrating the ArcInfo symbolsets to style files; Norah Brown for her expertise in ArcObjects programming and XML; and Sheila Hynes, Barb Szlavko, and Terry Houlahan for their support and expertise in geodatabase, SDE, and Oracle administration.

REFERENCES

ESRI Inc., 1999-2002, ArcGIS Desktop Help 8.3: ESRI Inc.

ESRI Inc., 2003, Geodatabase Designer Help: ESRI Inc. (developed by the Applications Prototype Lab).

ESRI Inc., 2003, Geodatabase Case Studies: ESRI Inc., Geodatabase Case Studies, Chapter 9 Topography and the Base Map (draft publication).

Fontlab software is available from Fontlab Ltd., Inc., accessed at http://www.fontlab.com.

Fontographer software is available from Macromedia, Inc., accessed at http://www.macromedia.com/software/fontographer.

Investigating the San Simeon Earthquake using ArcPad and GPS

By Lewis I. Rosenberg

San Luis Obispo County Planning and Building Department
County Government Center
San Luis Obispo, CA 93408
Telephone: (805) 781-4577
Fax: (805) 788-2413
e-mail: Lrosenberg@co.slo.ca.us

INTRODUCTION

The December 22, 2003, Magnitude 6.5 (M6.5) San Simeon, CA, earthquake provided a unique opportunity to use mobile digital mapping techniques to rapidly and effectively investigate the effects of the earthquake. This was accomplished by using a combination of mobile Geographic Information System (GIS) software (specifically, ESRI ArcPad) and global positioning system (GPS) technology on a handheld computer, or Personal Digital Assistant (PDA). This approach improved upon traditional techniques of disaster assessment and geologic field mapping, and resulted in a detailed, accurate database of earthquake effects and damage. This database has proven useful for a variety of needs including scientific investigations, building department permitting, and presenting information to decision-makers in the public sector. Thus, the combination of ArcPad and GPS will likely become a widely used tool for damage assessment in future earthquakes.

THE NEED FOR MOBILE GIS IN ASSESSING THE SAN SIMEON EARTHQUAKE

On December 22, 2003, a M6.5 earthquake struck the central coast of California, with the epicenter near the community of San Simeon (Figure 1). The San Simeon earthquake produced a large amount of damage, including collapse of unreinforced masonry buildings, ground cracking, widespread landslides, liquefaction, and changes to the flow from springs and water wells. The affected area extended in a southeastern direction from the epicenter, with significant damage as far as 80 km away. The estimated damage was approximately $239 million (Kircher & Associates, 2004).

Because the earthquake happened only a few days before Christmas, many people were on vacation. This resulted in a shortage of personnel available to assess earthquake damage. The State of California Governor's Office of Emergency Services elected not to set up a data clearinghouse for the San Simeon earthquake. Therefore, there was a great need to acquire in a timely manner information about the effects of the earthquake, for use by county officials and the geologic community. Faced with this task, as the San Luis Obispo County Planning and Building Department (SLOCPBD) County Geologist, I used mobile GIS/GPS to collect data about earthquake damage and effects. The combination of ArcPad and GPS was ideal for this task and enabled me to share data with other local, state, and federal agencies through an informal, Web-based county data clearinghouse, which I set up a few days after the earthquake.

EQUIPMENT NEEDED FOR EARTHQUAKE ASSESSMENT

For a mobile GIS that can be used for assessing earthquake damage, the basic components are fairly simple: a PDA, a GPS, and a copy of ArcPad. Edmundo (2002) presented a comprehensive review of the various types of PDA and GPS hardware available at that time and the reader is referred to that article for more detailed information. The following section outlines the typical items used for earthquake damage assessment.

Hardware

The basic requirement is a PDA with the Microsoft Pocket PC operating system. Most business-type PDAs have sufficient computing power and memory to operate ArcPad. Also, sub-meter precision and ruggedized hand-held computers are becoming more common. Screen visibility under a variety of lighting conditions is an important issue with any computer. Optional features that prove useful are memory card slots and wireless capabilities, such as wireless Local Area Network (WLAN, also known as "Wi-Fi"). Depending on the type of system, the GPS can connect to the PDA via a storage card slot, or as

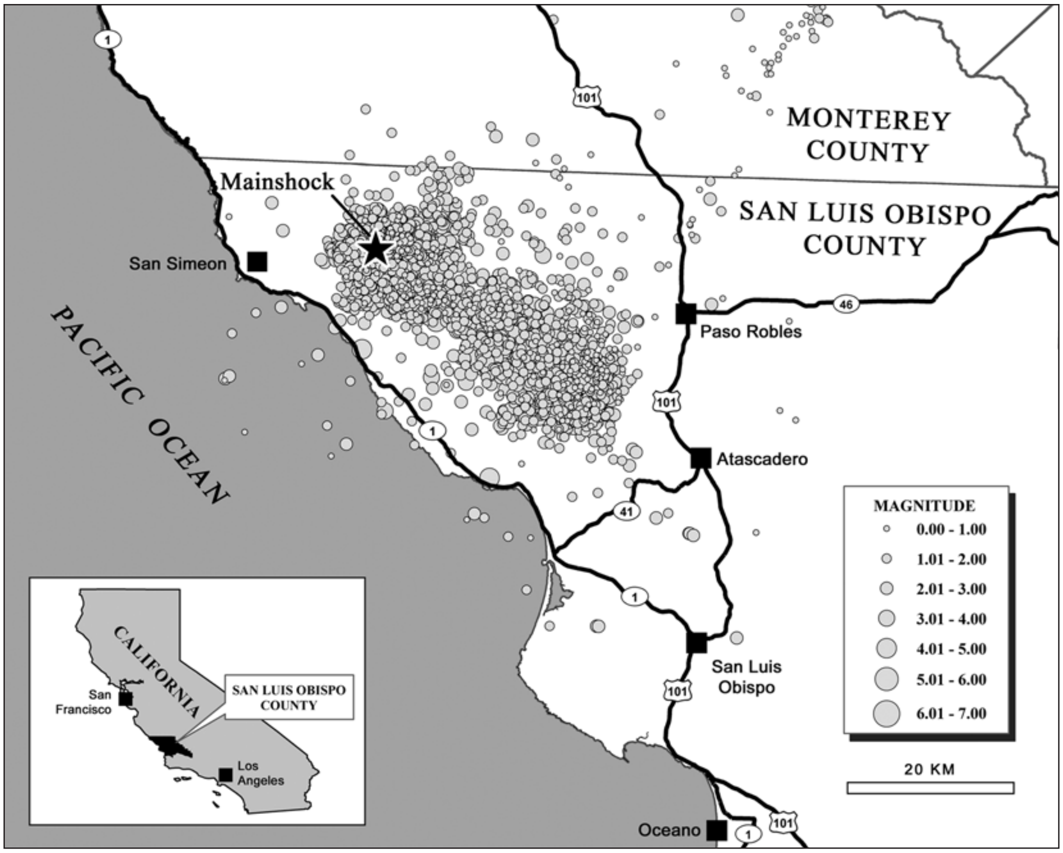


Figure 1. Location map of San Luis Obispo County, showing San Simeon earthquake epicenter (labeled "Mainshock") and aftershocks (indicated by gray circles).

a separate unit connected by cable. Storage cards facilitate carrying large amounts of data such as topographic maps. In the San Simeon example, a 512MB storage card contained the assessor parcel database, several scanned large-format geologic maps, a digitized streets database, and vectorized geologic maps for all of San Luis Obispo County, an area of nearly 8,600 km^2.

Alternatively, a tablet or laptop computer running Windows 2000 or XP software and ESRI ArcView connected to a GPS could be used for collecting data. The advantage of the tablet or laptop computer is the enhanced capability of the ArcView software, while the disadvantages are reduced portability, ease of damage, and increased cost.

Although there are many different GPS units available, I prefer the Garmin 12XL because of its relatively low cost (about $300 in 2004), ruggedness, availability, long battery life (24 hours) and AA battery power supply. The observed horizontal accuracy of this unit is about 5-10 m, which is acceptable for reconnaissance-level hazard mapping. Units such as the Trimble GeoXM have greater horizontal accuracy (2-5 m), but are more costly (about $2,800 in 2004).

Software

The minimum software requirement is a copy of ArcPad, which is the mobile, reduced-capability version of desktop GIS programs such as ArcView or ArcInfo. ArcPad allows the user to collect new data and to have the capability of displaying the data along with georeferenced photographs, maps, and vector files such as parcels and roads. For more advanced users, ArcPad Application Builder is helpful in creating and customizing ArcPad forms. However, ArcPad Application Builder is not included with ArcPad and it has a steep learning curve, so there are significant cost and training issues associated with it.

By means of automated data analysis from seismograph stations, California Integrated Seismic Network (CISN) posts on their ShakeMap website (http://earthquake.usgs.gov/shakemap) various ground motion values; this is done shortly after the earthquake occurs. These ground motions are available for download as either contoured polygons in ESRI Shapefile format or as raw grid (x,y,z) text files. These files are useful for initial estimates of where the strongest shaking and greatest damage are located, which is critical for emergency first-responders. At this time, only California, the Pacific Northwest, and the Salt Lake City, Utah areas have ShakeMap websites.

TECHNIQUES FOR MOBILE EARTHQUAKE DAMAGE ASSESSMENT

The techniques described in this section are intended for earthquake damage assessment. However, with minor modification, many of these techniques could be used for other types of natural disaster assessment such as landslides, flooding, or tornadoes.

Air Reconnaissance

Although not all mappers will have the opportunity to use helicopters or airplanes for earthquake damage assessment, it is one of the best ways to quickly cover a large area. By adding ArcPad and GPS, air reconnaissance becomes even more valuable. Here, I present a case study on how air reconnaissance combined with ArcPad and GPS were used in the San Simeon earthquake, along with lessons learned and issues that need to be resolved.

Following the earthquake, geologists from the USGS (John Tinsley and Kevin Schmidt) and the SLOCPBD (Lewis Rosenberg) flew by helicopter to the earthquake epicenter. We navigated to the epicenter using a Garmin 12XL GPS connected to a Hewlett-Packard Jornada 568 PDA running ArcPad version 6.03, supplemented by two Garmin GPS III GPS units. The epicenter location was downloaded from the USGS ShakeMap website (California Integrated Seismic Network, 2004) into ArcPad, and the GPS was used to guide the helicopter (Figure 2). We circled in order to get a feel for the level of detail we could and could not observe at a given elevation. Then, flying as low as the pilot liked, we circled about the epicenter at radial distances increasing at 3-km increments, at roughly 150 m above the terrain. Our goal was to search for fault rupture, landslides, and building damage. ArcPad allows the user to store the flight path, which can be superimposed on simulated terrain by means of a digital elevation model (Figure 3). This mode of presentation was useful to show other scientists and non-technical audiences the area covered in the initial air reconnaissance.

After circling the epicenter out to about 6–8 km radii, we then flew southeast along the mapped trace of the San Simeon fault. In places, this fault is relatively easy to follow from the air, as it is a contact between distinctive rocks (Monterey Formation shale and Franciscan Complex mélange) with a break in slope and change in lithology denoting the different units. In other areas, the fault is concealed beneath marine terrace deposits. ArcPad and GPS were extremely useful in these areas where the fault was concealed, owing to the ability to navigate in real time using a scanned, georeferenced geologic map (Figure 4).

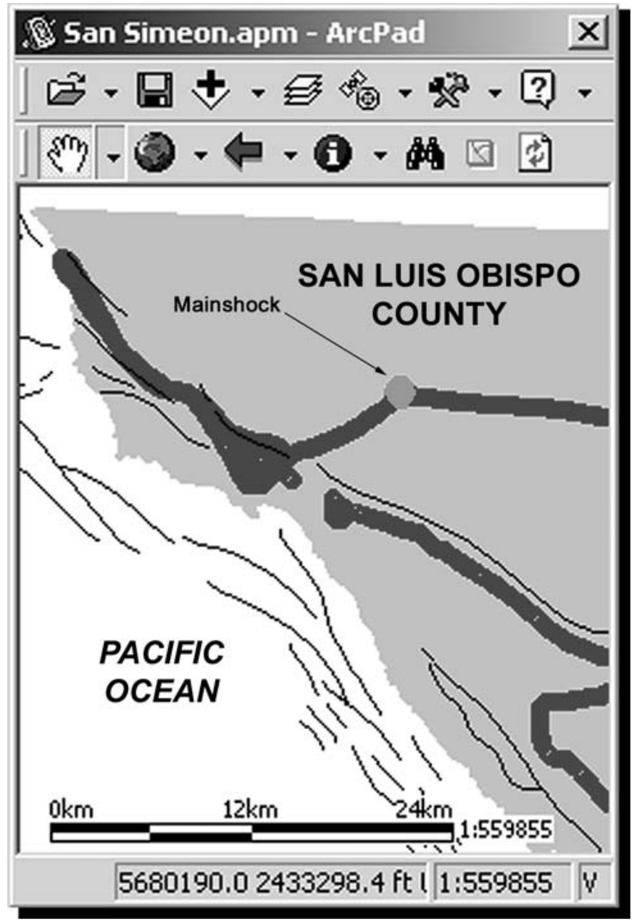


Figure 2. Screenshot of helicopter flight path (thick black lines), faults (thin black lines), and mainshock epicenter.

One of the lessons learned in studying California earthquakes is that seismically triggered rockfalls and landslides correlate with areas of high peak ground velocity (McCrink and Wilson, 2004). By loading into ArcPad the Shapefiles of peak ground velocity obtained from the ShakeMap website, we were able to explore the areas most likely to have landslides (Figure 5). This approach worked well—we noted numerous rockfalls and sheared and uprooted oak trees in areas of high peak ground velocity (30-40 cm/s) compared with other areas. These features suggested that the actual ground shaking was stronger than indicated by the initial automated ground motion reports.

We also used ArcPad and digital cameras to record damage, by creating point features linked to photographs of some of the more heavily damaged homes and roads in the foothills southwest of Paso Robles. This enabled us to find damage in a densely vegetated area with few roads, a task that would have been much less successful by ground-based methods. However, one problem was the issue of accuracy and precision owing to taking measurements in a moving aircraft. The airspeed of the helicop-

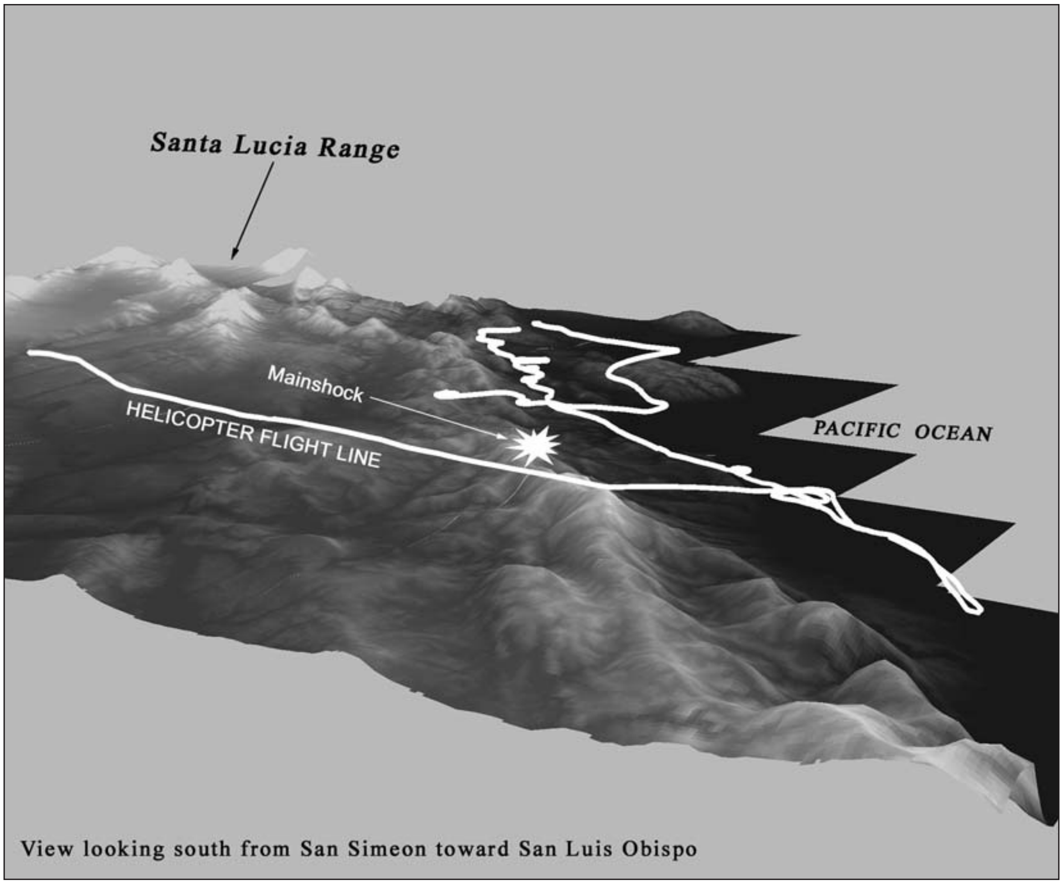


Figure 3. Oblique view of digital elevation model showing helicopter flight path on return to Paso Robles and earthquake mainshock.

ter was as much as 130–145 km per hour, with slower airspeeds where we were circling to evaluate observations made at higher rates of airspeed. We tried to account for this by each of us taking independent GPS readings for each point recorded. Although this approach offered the prospect of increasing the precision, it also complicates data reduction.

We anticipate that by using the photos we took, in relation to the trackline data, we can sort out and eventually report the best locations for observed features. We do not have a GPS location for each photo, as the camera shutter cycles much more quickly than does the waypoint acquisition function in the Garmin unit, but hopefully we acquired enough control to fit the photos into the terrain with some reliability. Regardless, the precision of GPS locations relative to observed features in the terrain will vary, because during the time we logged a given waypoint our helicopter's elevation would vary. In addition, the azimuth of the helicopter changes with each photograph taken or waypoint reading. It would seem to be helpful to have the helicopter's azimuth logged as a function of time, to help restrict the camera's possible orientations while locating features photographed from the air. This could be accomplished by incorporating the bearing ("BRG") field that corresponds to compass direction and is calculated by ArcPad.

As a general comparison with ground accuracy, we observed from the helicopter a prominent headscarp to a deep-seated landslide located on a hill near Paso Robles. The landowner had independently reported the fissure to me, and I subsequently located the headscarp on the ground using the Garmin GPS. This comparison showed that our helicopter GPS location was within 150 m of ground location. It is encouraging that our airborne-determined GPS locations may be close enough to be considered useful.

Ground Reconnaissance

Ground reconnaissance is the most commonly used technique for earthquake damage assessment. One of the useful features about using ArcPad and GPS is that sites can easily be located where street addressing is poor or non-existent. This is accomplished by having street map files such as U.S. Census "TIGER" files, or even better, having digital assessor parcel maps (Figure 6). For homeowners who do not work at home, this capability is essential; with digital parcel maps and cell phones, we were able to instantly contact landowners

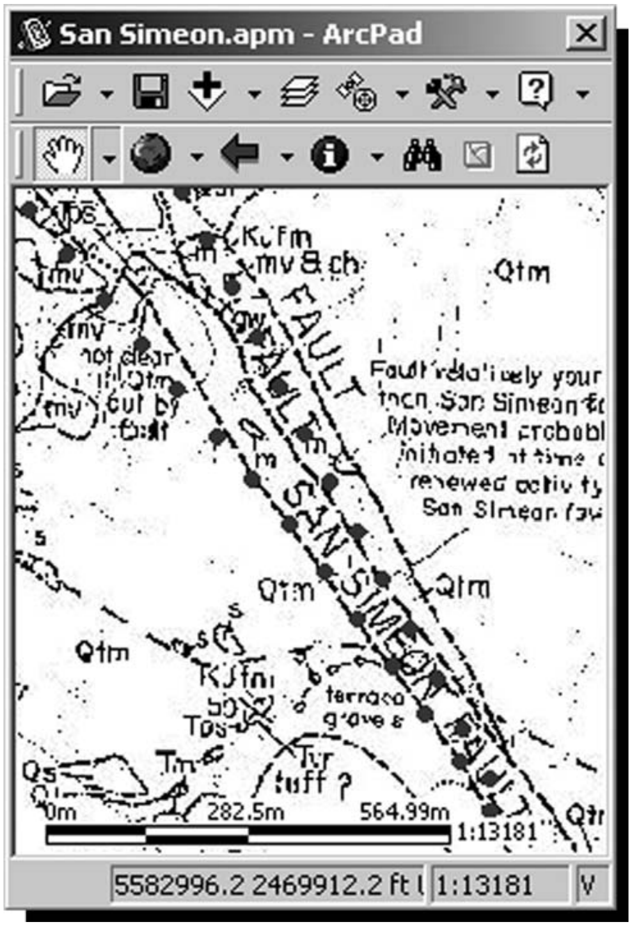


Figure 4. Screenshot of helicopter flight path (dotted lines) and scanned geologic map (Hall, 1974) showing San Simeon fault.

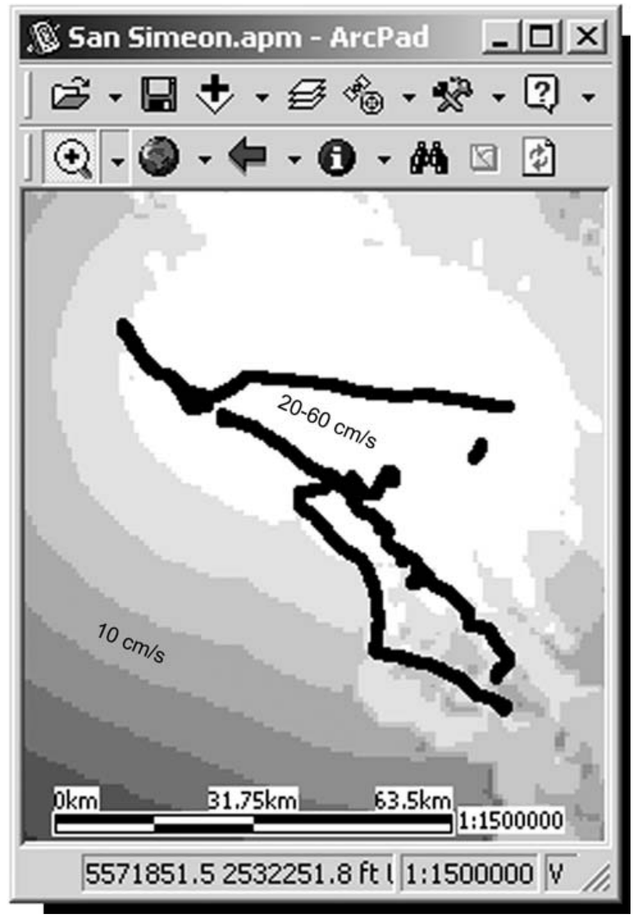


Figure 5. Screenshot of helicopter flight path and ShakeMap peak ground velocities (lighter color indicates relatively higher velocity).

for permission to trespass on their land.

Another purpose of ground reconnaissance is to map ground cracks. Although the typical horizontal precision of a recreational-grade GPS such as the Garmin 12XL is only 5 m, it is sufficient for reconnaissance-level mapping in rural areas, especially if a high-resolution aerial photograph is available for ground truthing (Figure 7).

Public Relations

Although public relations are rarely presented as a topic in disaster assessment papers, it is a key issue in dealing with the public, especially following earthquakes. Landowners were much more receptive to allowing geologists to map their land once they were given a color aerial photograph of their property. Mobile GIS/GPS makes this a simple task if a color printer is available.

SOME IMPORTANT ISSUES

Although ArcPad and GPS are extremely useful for earthquake damage assessment, there are significant advantages and disadvantages associated with this approach. These are summarized below.

Advantages

- It is fast and allows one person to collect large amounts of relatively accurate data.
- It allows users to share data both in the field and in the office.
- Data combined with appropriate imagery makes it easy to visualize the distribution and degree of damage.
- Common file format (Shapefile) facilitates simple exchange with desktop GIS software.

Disadvantages

- The cost (in 2004 dollars) for a recreational-grade GPS, a business-type PDA, and a single ArcPad license is about $1,500. Sub-meter precision and

Figure 6. Screenshot of county parcel map and liquefaction sites (black dots) located with GPS.

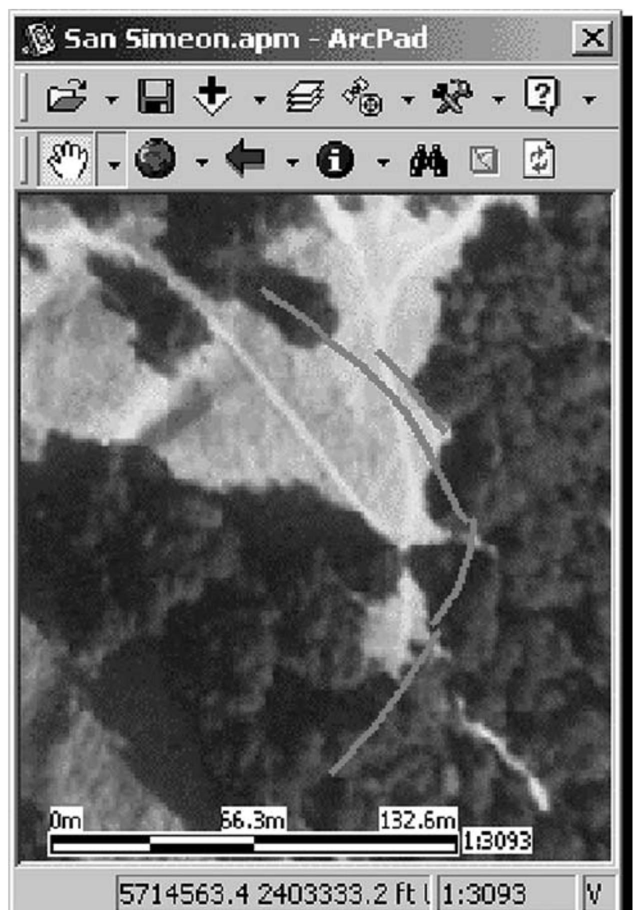


Figure 7. Screenshot of aerial photograph and landslide scarps (shown by the gray lines) located with GPS.

ruggedized hand-held computers cost upwards of $2,500. A desktop GIS software package of ArcView would add approximately $1,500 at a minimum; the costs would be more if add-on extensions such as Spatial Analyst or 3D Analyst were included. For agencies with limited budgets, these could be significant expenditures, especially if more than one ArcPad/GPS unit was required.

- Not all agencies have GIS, so sharing data can be difficult. Alternatively, the database part of the Shapefile (the .dbf file) can be exported and read by commonly available programs such as Microsoft Excel. The downside is that the spatial component is lost. However, this did not seem to be an obstacle to sharing data with local building departments, who were mainly interested in "red-tagged" (structurally unsafe) building locations.

CONCLUSIONS

Our method of using ArcPad and GPS allowed us to acquire data effectively and quickly. It is a significant improvement to the technique of "pen and paper" mapping used in previous earthquakes.

The widely used Shapefile format enables exchange of data inside and outside of the agency collecting the data.

For maximum effectiveness in hazards mapping, the equipment must be available in advance, and an event response plan must be in place. Although it is possible to learn basic proficiency with ArcPad in a short time, training individuals before the earthquake event allows them to focus on damage assessment during a time when every minute counts.

There are significant cost and training issues for acquiring the hardware and software. However, ArcPad and GPS can be used for routine, non-emergency applications such as building code enforcement, so the equipment is not for emergencies only.

ACKNOWLEDGMENTS

Because the San Simeon earthquake was an unplanned event, many of the techniques described in this paper were developed in the days following the earthquake. Thanks go to my co-workers J.C. Tinsley, III and K.M. Schmidt (USGS); T.P. McCrink, R.C. Loyd, and K.L. Knudsen (California Geological Survey), and M.J. Johnsson (California Coastal Commission) for taking the time to spend long days in the air and field during the holiday season—their help and insight were invaluable. Parts of Tinsley's and Schmidt's helicopter report were used in this paper. John Kelly (SLOCPBD Supervising Mapping Systems Specialist) is commended for his foresight in placing San Luis Obispo County's GIS data on the Internet so that other workers could independently get the base map files they needed to perform their work.

REFERENCES

California Integrated Seismic Network, 2004, ShakeMap Home Page, accessed at http://earthquake.usgs.gov/shakemap/nc/shake/index.html.

Edmundo, G.P., 2002, Digital geologic field mapping using ArcPad, *in* Soller, D.R., ed., Digital mapping techniques '02—workshop proceedings: U.S. Geological Survey Open-File Report 02-370, p. 129-134, accessed at http://pubs.usgs.gov/of/2002/of02-370/edmondo.html.

Hall, C.A., 1974, Geologic map of the Cambria region, San Luis Obispo County, California: U.S. Geological Survey Miscellaneous Field Studies Map MF–599, 2 sheets, scale 1:24,000.

Kircher & Associates, Consulting Engineers, 2004, M6.5 San Simeon earthquake—December 22, 2003, comparison of HAZUS99-SR2 estimates and observed damage and losses: unpublished report dated March 29, 2004, 5 p, accessed at http://www.hazus.org/BAHUG/BAHUG%20Documents/KIRCHER_San_Simeon_Analysis.pdf.

McCrink, Timothy, and Wilson, Rick, 2004, Summary of Seismic Hazards Mapping Program post earthquake field activities following the magnitude 6.5 San Simeon earthquake of December 22, 2003: Part 1—Summary of reconnaisance field activities and observations: California Geological Survey, Seismic Hazard Mapping Program, Technical Document 2004-1, 44 p.

VENDOR CONTACT INFORMATION

ESRI ArcInfo, ArcPad, ArcPad Application Builder, ArcView—Environmental Systems Research Institute (ESRI), Inc., 380 New York St., Redlands, CA 92373-8100 USA, (909) 793-2853, accessed at http://www.esri.com.

Garmin 12XL—Garmin International Inc. 1200 East 151st St., Olathe, KS 66062 USA, (913) 397-8200, accessed at http://www.garmin.com.

HP Jornada 568—Hewlett-Packard Development Co., L.P., 3000 Hanover St., Palo Alto, CA 94304-1185 USA, (650) 857-1501, accessed at http//www.hp.com.

Microsoft Pocket PC, Windows 2000, Windows XP—Microsoft Corp., One Microsoft Way, Redmond, WA 98052-6399 USA, (425) 882-8080, accessed at http://www.microsoft.com.

Trimble GeoXM—Trimble Navigation Limited, 749 North Mary Ave., Sunnyvale, CA 94085 USA, (408) 481-8000, accessed at http://www.trimble.com.

GanFeld: Geological Field Data Capture

By Guy H.D.P. Buller

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

OVERVIEW

GanFeld:
gan – Old English meaning Open
feld – Old English meaning field

The collection of geological information has been an ongoing process in Canada even before the inception of the Geological Survey of Canada (GSC) by William Logan 162 years ago. Much of the geological information collected was based on the need to discover new mineral resources for a new country and to further understand geological phenomena. Geoscientists go out to the field, collect samples to examine more closely and model a geological process that often leads to a report on their findings. The difference between then and now is the speed at which things happen as well as the precision of the location, the accuracy of the results, and how information is disseminated to the world.

The importance of geological information in the past 40 years has expanded beyond the focus of pure research or mineral resource discovery. With this change in scope there is the need to share accurate, up to date information between different groups and disciplines, whose demands continue to grow. Geological surveys around the world acknowledge that they house a wealth of information that can be critical to the enhancement of business opportunities, the environment, and the citizens of their respective nations. Many of these organisations also recognize that the information that is easily available often is not current, or failing that, is difficult to discover, and is often in formats that are archaic to modern technology. These hurdles limit easy access and are not sympathetic to the demands of today's rapid and potentially critical decision-making processes. In response to these issues, a concerted effort has taken place to find solutions that make the products and complete data of geoscientists' work more accessible and of an assured high quality for today's needs.

To meet the above challenges and to streamline efforts, organisations are looking at the collection, processing and analysis, and final dissemination of information from a business perspective. The different steps that are involved in making a geological report have been examined as critical parts of a whole process, rather than being end results in themselves. This 'holistic' view of the information process has shown where and how there can be immediate improvements in data quality and a reduction in the time it takes to publish final results.

A significant improvement in this information processing can be achieved by reducing our reliance on paper formats at various phases of geological research; this finding has been well documented by the British Geological Survey. By evolving from paper field notes to electronic field data-capture, we reduce the likelihood that transposition or scientific errors will enter the data collection process. Collecting information in this way also increases the ability to search and manipulate the field data. By using such electronic systems, quality assurance and quality control (QA/QC) begin at the inception of data collection.

FIELD DATA CAPTURE

Many different computer applications are being used to capture field data electronically, and several groups both nationally and internationally (e.g., http://www.bgs.ac.uk/dfdc/home.html) are working hard to make data collection applications fit their specific set of business rules. In almost all cases, field applications have been developed 'in house' by manipulating existing software applications to fit into the requirements that geoscientists demand for data capture. Some of these applications are full geological mapping systems that "capture" nearly all the geologic information for a single point of interest, whereas other applications limit their information to XY coordinates with a few electronic notes and then compliment this electronic data with hand written notes in field journals. In the case of two full scope well-known field applications, *FieldLog* (Brodaric, 2004) and *GeoMapper* (Brimhall and Vanegas, 2001), a laptop computer is used to gather and hold information electronically; these two systems have had broad acceptance by researchers and have been very successful in their many varied field deployments. Yet, the use of a laptop computer often means that field data gathering systems still rely on the use of paper forms and the manual transfer of this data to an electronic type format. Alternately, if true site-to-site specific information is to be captured using a lap top computer, a vehicle of some description (either a truck or a 4-wheeler) is required to transport the computer to each location due to the weight and size of laptop computers. This can also mean that special vehicle mounts are required to hold the computer in place on board the vehicle.

Handheld Computers

With the advances of technology, computers have become smaller and personal digital assistants (PDAs) have become more powerful, more rugged, and more conducive to being used while on traverse. An initial trial in Canada of electronic data capture using PDAs was researched and developed by Gilbert, Parlee and Scott (2001) and further extended by Celine Gilbert and Edward Little (GSC, oral communication) using the Palm handheld devices that leverages the power of Microsoft Access through software known as Pendragon Forms. The continued success of these data gathering devices in arctic terrains have prompted their use in other areas of Canada and have proved themselves as reasonably inexpensive mobile field data gathering systems. The idea of using a truly relational data structure for capturing data has been extended further with the goal of loading information directly into a data base structure that mimics the data structure found at the corporate level (Buller, 2002). This single database structure concept would greatly facilitate the transfer of information to a corporate data holding and reduce the data manipulation needed to facilitate the transfer.

While many of these systems work well for an individual project or survey, extending their functionality to other projects or groups often means substantial redevelopment of the application. In real terms, this means that there has been no improvement in the level of data access or sharing, because data created with the various different applications is not easily exchanged; in reality, we have only altered the format, from paper to various, less accessible electronic formats. Furthermore, the information captured or the terminology being used for a project may be specific to a single researcher and thus only may have a life span equal to the length of time that the geologist is employed at the organization. This lack of interoperability is well known amongst researchers and organisations, and much work has been done to 'translate' information into different formats to enhance communications between these different groups (Brodaric, 2004). These data translation activities have met with limited success and are recognised as large consumers of time and resources.

Searching for Solutions

The Canadian federal government's *Government On Line* initiative (http://www.gol-ged.gc.ca/index_e.asp) intends to have most government information available on line by 2005. This initiative has been an incentive to put geological data into electronic formats and to increase the accessibility of this information via the Internet. To achieve these ends, the overall business of geological information collection and distribution needed to be examined.

At the Geological Survey of Canada (GSC) Terrain Sciences Division in Ottawa, there has been a concerted effort over the past few years to streamline the workflow of sample processing. As a result, the GSC has developed a laboratory information system (LIMS). The LIMS has been effective in improving the quality of the analytical results, and assists in expediting quality information transfer to the final publication process (R. Laframboise, oral communication, 2002). As the LIMS was being finalized, steps were being taken to integrate geochemical data across three divisions of the GSC into a common geochemical data structure. These developments were seen as part of the foundation that would facilitate a geochemical mapping web presence using ESRI's MapObjects, thus making large gains towards the government's on-line initiative.

The MapObjects application requires ESRI Shapefiles to deliver maps to clients via the web. In the case of many raw data gathering systems information is captured in formats that are not visual and therefore are not spatially referenced. This difference in data formats is similar to the translation problem mentioned earlier, in that there is a need to manipulate raw information to develop Shapefiles for use in GIS systems. This activity of altering data formats may not be problematic but it splits the raw information into a spatial file and a data file. This division of information is counter productive and it is felt that capturing map data directly in the field on a station-to-station basis would be an effective way of streamlining the information process while at the same time capturing vital geological information. The challenge to capture this sort of map data as well as other information was met with the development and release of ESRI's ArcPad handheld mapping application (http://www.esri.com/software/arcpad/index.html).

ArcPad

The ArcPad application has essentially put a GIS in the geologist's pocket, giving them the ability to plot a variety of map information (polygon, line and point) and to directly capture point information from a GPS receiver. By setting up the Shapefile data table (a DBF file) to capture a wide range of other information in addition to spatial data, it is possible to have the best of both worlds, a digital data capture system as well as a visual display of the map data while out in the field. A further advantage with having a map interface is that other map information, such as gravimetric data or geological feature sets (polygon data, outcrop delineation sets, etc.) can be accessed using the same device, and viewed in the field with newly captured map data. This combination of spatially-related raw data and maps effectively means that researchers at the end of a field season have a preliminary map that is available for publishing, as well as an easily searchable data set that is geographically referenced. ArcPad has allowed us to reach a main goal, which is to better help the geologist in their work.

Though there is some resistance toward the new way of collecting data it must be kept in mind that, at one time, the use of paper forms in the field was considered as an inconvenience to the geologist but are now often seen as an indispensable aid to the systematic capture of information in the field. Creating electronic forms allows the geologist to retrieve, share, and examine data more easily, and in turn allows the geologist to think about geology rather than be concerned with the input of raw information into computer systems. As an additional bonus, any functional coding to customize ArcPad is done using VBScript, meaning that web developers who build active server pages (ASPs) can transfer their expertise directly to the development of ArcPad applications.

Field To Curatorial Project

Over the past year there have been far-reaching changes at the GSC. These changes have instituted a project-based system, and the *Field to Curatorial Project* (FTC) is one of these projects. This project's main goal is to track a physical sample from initial collection, through processing and analysis, to the archive. The desire to have this broad spectrum of information accessible at all times was seen by management as an important goal. Under these guidelines the FTC project has three distinct modules (field, laboratory, and archival). The need for continued refinement and development of a robust field data-gathering application became clear, as it is the first step to making field data more interoperable between groups and the project modules.

To meet the goals of the FTC project it was necessary to look at the work that geologists do and to consider this work from a business model perspective. This modelling was extended to the field module and at the earliest stages of the project, it was recognised that although geologists do similar activities they do not always use the same language for describing these activities or common things in the field. Therefore, one of the first steps for the field module was to try to standardise some of the common terms and expressions that are used in three different phases of geological field research. These strictly field related phases of geological research are an arbitrary breakdown and have no set time period and are recognised as pre-field research, fieldwork, and post-field data manipulation or research.

During the first phase there can be a considerable amount of paper research for an area that may include interdisciplinary discussion. Often, common terms are used in several different ways and are mostly non-scientific words that are used in daily speech. These words are often applied to items or actions specific to the discipline or to the researcher during the various phases of geological research activities. Between researchers the 'translation' of words to gain meaning is not a problem as it is an organic process that humans use in everyday conversation, but because there is no such intrinsic translation between computers, this level of ambiguity causes havoc when dealing with databases. In order to facilitate the input of data into relational database systems, terms that have specific definitions must be agreed upon by a variety of users.

These words have been developed by consulting different ISO publications, considering other developments within the GSC as well as soliciting information from a number of researchers. The original set of words were then reviewed by a number of people inside and outside the project, and have become the starting point of a lexicon to be used for information collection standards. Table 1 shows an example of some of the developed lexicon.

The set of common definitions in turn determines the minimum information required for any geological project, and gives more consistent information sets between different field projects. Through the use of a questionnaire, geologists were surveyed to gather information about the various aspects of fieldwork carried out in different parts of the GSC. These questionnaires were followed up with discussions and meetings as a way of clarifying the modeled work process and determining clear definitions for terms that an individual researcher uses. It must be kept in mind that the common word set is not a static entity, as new words and definitions will be added over time. This type of development, which builds a common word system, is similar to other efforts (e.g., ISO Standards) in attempting to standardise an activity that is carried out by many individuals in a branch of research. By following such a standard system, an individual's specific words can be matched to the common word set and, subsequently, field project-specific data can easily be transferred and stored in a relational database along with other interdepartment survey data. This common word set accomplishes two goals,

- the facilitation of data input to a common model, and
- a reduction of data ambiguity both now and in the future.

Over time, the accumulated information will become more of an asset to the pre-field research phase and will allow researchers to more easily share post-field information. Furthermore, general field information, such as observations about vegetation or morphological descriptions outside of the research scope, can also be extended to web applications or web services to further promote geological works and to demonstrate to managers and general public, in a timely manner, of the work that is being completed by a research group. This up-to-date information becomes more important when dealing with regions that have surface-access-rights issues or areas of a high sensitivity (i.e. environmentally protected arrears) and gives a clear indication to the general public as to the extent of the work being done by a party.

Part of the challenge faced by the FTC project was the introduction of a new business paradigm by management that has focussed on stronger accountability of expenses and more intra-department collaboration of activities. Thus, in preparation for this change in business process and before any of the software development began, measures were taken by the FTC project Information Management (IM) coordinator (Richard Laframboise) to extensively plan the project's time line. Over the past year the use of business requirements analysis has been our main focus in an effort to both document the development of the project and also give direction and communication to the various groups within the GSC that are distributed throughout the country.

This business-centric effort has placed a large emphasis on the planning stages of this project and has used the Zachmann framework model and business requirements analysis as demonstrated by Hay (2003). These planning and analysis activities have been invaluable in understanding the scope of the application and the roles of the different individuals who are involved in the FTC project. In the past, this type of planning activity has had limited use because many previous projects have had impacts that were limited to very small groups or individual researchers. As there is a desire by management to expand the extent of web accessible information and to have more accountability, a project of this kind needs to extend its contacts to as many groups as possible in order for the long term planning phase to be successful. The use of the planning tools have been invaluable for focusing the

Table 1. A working example of the lexicon that has been developed at the Geological Survey of Canada's Field To Curatorial Project, for the collection and storage of field information. (Note: Information in ***bold italics*** indicates an edit to the lexicon that has yet to be considered as accepted).

Word	Definition	Source
Sample	1) Portion of material selected from a larger quantity of material. 2) The raw material collected in the field and shipped back to the lab.	1) ISO 11074-2 2) Geochemistry Database concept
Specimen	Specifically selected unit/portion of a material taken from a dynamic system and assumed to be representative of the parent material at the time it is taken. NOTE 1: A specimen may be considered as a special type of sample, taken primarily in time rather than in space. NOTE 2: The term "specimen" has been used both as a representative unit and as a non-representative unit of a population, usually in clinical, biological and mineralogical collections.	ISO 11074-2
Activity	***An action carried out by a field party at a specific station that in some way gathers information about that specific field station. Examples: observations (including no activity), picture, drawing, sampling.***	***Guy Buller***
Sampling	Process of drawing or constituting a sample (ISO 3534-1:1993). NOTE: For the purpose of an investigation, "sampling" also relates to *in situ* testing carried out in the field without removal of material.	ISO 11074-2

development plan for the present fiscal year and helping the individual developers recognise how they fit within the project itself.

CONCLUSION

Computerized mapping is finally being widely adopted for field use. Continued advancements in technology will make data collection systems commonplace and will be an even greater asset to the geologist in the future. As the costs of running a field camp increase, no longer do we have the luxury of letting field data collections languish in obscurity by having data in a multitude of formats that are not interoperable or easily accessible. The data collected by the geologist at an individual station has importance, as do observations and 'thoughts'. These thoughts, at the time of data capture, can give clarity to a geological model when contrary models are introduced. Furthermore, field data that today may seem unimportant may in the future become extremely useful.

Any data gathering system that is developed, regardless of operating platform, must have interoperability of data as a main goal. This means that some of the focus of any development has to be centered on a data storage structure that allows researchers to use any data capture tool available. This data storage needs to be able to extend the availability of data to researchers and also allow for single queries to access multiple, seemingly disparate data sets.

The planning stages for the Field to Curatorial project have been most helpful in understanding the scope of the project. It is hoped that by following such a stringent planning stage, others who intend to develop a similar system can learn from the process that is being documented. The planning stage is critical to determining the actual needs of the business and the process to meet those needs.

To simply capture field information for the specific use of a single researcher limits the sharing of data and ultimately does not advance the scientific process. The information must be accessible by others, now and in the future, in order to serve the public and the science.

REFERENCES

Brodaric, Boyan, 2004, The design of GSC FieldLog: ontology-based software for computer aided geological field mapping: Computers & Geosciences, v. 30, issue 1, p. 5-20, accessed at http://www.sciencedirect.com/science/article/B6V7D-4B1SFSX-1/2/e76e0967fde762328b7080f865c7fbc8.

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

Buller, Guy, 2002, Ganfield: data integrity from field to final product, *in* Capturing Digital Data in the Field Workshop, British Geological Survey, Nottingham, England 25 & 26 April 2002, accessed at http://www.bgs.ac.uk/dfdc/gbuller.html.

Gilbert, C., Parlee, K., and Scott, D.J., 2001, A Palm–based digital field-data capture system: Geological Survey of Canada, Current Research 2001-D23, 10 p.

Hay, D.C., 2003, Requirements Analysis: From Business Views to Architecture: Prentice Hall PTR, Upper Saddle River, New Jersey, 496 p.

HARDWARE AND SOFTWARE CITED

Operating System

PalmSource, Inc., 1240 Crossman Ave., Sunnyvale, CA 94089, (408) 400-3000, Fax: (408) 400-1500, accessed at http://www.palm.com/us/.

Hardware

palmOne, Inc. Corporate Headquarters, 400 N. McCarthy Blvd., Milpitas, CA 95035, (408) 503-7000, Fax: (408) 503-2750, accessed at http://www.palm.com/us/.

Hewlett-Packard Company, 3000 Hanover St., Palo Alto, CA 94304-1185, (650) 857-1501, Fax: (650) 857-5518, accessed at http://welcome.hp.com/country/hk/en/welcome.html.

Software

ESRI, 380 New York St., Redlands, CA 92373-8100, (909) 793-2853

Pendragon Software Corporation, 1580 S. Milwaukee Ave., Suite 515, Libertyville, IL 60048, (847) 816-9660, Fax: (847) 816-9710, Sales Email: info@pendragonsoftware.com, Support Email: support@pendragonsoftware.com, accessed at http://www.pendragon-software.com/index.html.

Geologic Data Assistant (GDA): An ArcPad Extension for Geologic Mapping

By Ralph A. Haugerud[1] and Evan E. Thoms[2]

[1]U.S. Geological Survey
University of Washington
Dept. Earth and Space Sciences
Box 351310
Seattle, Washington 98195
Telephone: (206) 713-7453
Fax: (206) 553-8350
e-mail: rhaugerud@usgs.gov

[2]U.S. Geological Survey
Alaska Science Center
4200 University Drive
Anchorage, AK 99508
Telephone: (907) 786-7409
Fax: (907) 786-7401
e-mail: ethoms@usgs.gov

INTRODUCTION

For two centuries, geologists have traveled with pencil, notebook, and base map and returned with records—field data—that support the production of geologic maps. At earlier DMT workshops there were reports on experiments using a computer to collect such field data, including a special session at DMT'01 (Soller, 2001). Reasons for exploring digital solutions have ranged from more rapid production of paper maps (Brodaric, 1997; Williams, 1997; Walsh, 1999), more efficient field work (Pavlis and Little, 2001), and more efficient creation of digital map data (Williams, 1997; Brimhall and Vanegas, 2001; Brimhall and others, 2002; Edmondo, 2002; van de Poll and Parsons, 2003), to better visualization of spatial relations (see the GeoPad effort at http://geopad.org).

We have followed these experiments with interest, looking for a digital field solution that increases the efficiency of our regional-scale (1:12,000 to 1:100,000) geologic mapping. In particular, we desire a field data collection system that:

- Transparently incorporates location that is determined by GPS,
- Allows use of digital base materials (assessor's parcel maps, lidar topography, high-resolution orthophotos) that are commonly not available in paper form,
- Facilitates the collection of feature-level metadata (observer, time of observation, spatial accuracy and precision, etc.),
- Encourages standardization of observations, and
- Records data in structures and formats that can be translated by machine into our project-level geologic map databases.

We wish to accomplish these objectives while maintaining the portability, richness, flexibility, and graphical nature of the traditional notebook and field map.

Some digital solutions have depended on the processing power, operating system, and storage of a laptop (or tablet) computer. These are a problem because of the consequent bulk, weight, limited battery life, and excessive cost of ruggedized systems. Other solutions have used handheld or palmtop computers (personal digital assistants or PDAs) with software that does not support on-the-fly plotting of new observations. These are unacceptable to us because spatial reasoning from new observations is an important part of the mapping process.

ESRI's ArcPad™ software, running on an off-the-shelf PDA with a daylight-readable color display, has the potential to bypass these limitations (e.g., Edmondo, 2002). At version 6, ArcPad incorporated an easy interface to commonly-available GPS units. ArcPad can now support many of the needs of geologic mapping, but to do so requires the development of a customized application.

Such customization is costly: $1,500 to purchase ArcPad Application Builder, weeks to months of time to learn VBScript and the ArcPad object model, and days to weeks of coding and experimentation. The difficulty of customization seems to be reflected by the fact that in early July 2004, a Web search for *"ArcPad application" geologic mapping* yielded only a handful of pages, none of which offered a system that could be used for general geologic mapping without further modification.

Geologic Data Assistant (GDA) is our effort to fill this void. It is an ArcPad application that works well enough for Haugerud to use it routinely instead of a field map and notebook. We are publishing the GDA code (Thoms and Haugerud, in preparation) for the benefit of geologists who may find it a useful tool, and to encourage others to improve upon our efforts.

In the course of developing GDA we found ourselves examining our geologic mapping procedures more closely. We made compromises between the richly-structured geologic database of our dreams and the capabilities of ArcPad. We learned by trial and error that certain design choices were critical. This paper describes GDA and what we learned while developing it.

HOW GEOLOGIC MAPPING WORKS

Many person-years of experience with regional mapping suggest the following essential elements of a field-based geologic mapping process:

1. Before going into the field, one:

- Assembles base maps, and
- Compiles the results of previous geologic investigations.

2. While in the field, each geologist in a party develops a set of field data. This consists of several elements:

- Field notebook, organized by station number, with stations numbered consecutively. Notes for each station commonly are formatted in some fashion, but may contain extensive free-form text,
- Field map, that includes locations for all stations as well as graphical representations of many of the elements (particularly structure data, map units) recorded in the field notebook,
- Short segments of contacts, faults, and unit polygons that are sketched on the map, and
- Samples, photos, sketches, and other items. There may be zero to many of these elements at each station.

3. The mapping party maintains a compilation map that commonly has two elements:

- Station (and sample) inventory, and
- Compilation geologic map that adheres to geologic map rules: it shows units, contacts, and faults; it shows selected structure data; it has topology (i.e., certain topological constraints among lines and polygons are enforced), and it is inked and colored.

When one geologist is working alone, this process is still maintained.

The best geologic mapping is an exercise in graphical reasoning, using the topology and geometry of previous observations to guide the remainder of a traverse. We find that as we mature as mappers, we do more of our work on our field map and record less information in our field book. A digital field map is at least as important as a digital field notebook.

Note that field data are not simply a preliminary version of the completed geologic map. First,—at least at the beginning of a mapping project—field data serve to document an exploratory phase: What units are mappable? What lithologic and structural details are key to distinguishing them? At this stage of investigation one's data structure must allow for evolving lithologic and stratigraphic vocabularies. A paper notebook and pencil are quite effective for this. Only later does one shift to the focused observations that make database design easy, and the controlled vocabularies that make database implementation efficient. Second, field data commonly contain numerous elements that are not preserved on the final map: the name of a landowner; whether a dog is friendly; the weather; who (in a multi-worker effort) made an observation; how, and how accurately, a position was determined; and, extensive lithologic detail and stratigraphic and structural speculation. Third, most workers find that their field maps are not geologic maps because not all contacts and geologic units are identified. With more complete fieldwork, a field map may closely resemble a geologic map but, commonly, full interpretation of an area depends on laboratory analyses, additional data from a colleague, or simply further thought.

ABOUT GDA

GDA is an extension to ArcPad, which is a simplified GIS for Microsoft Windows CE / PocketPC that also runs on Windows 95/98/NT/2000/XP. ArcPad is published by Environmental Systems Research Institute, Inc. (ESRI) and has moderately good display capabilities, good projection capabilities, and an excellent GPS interface. ArcPad cannot build, or even enforce, topology (e.g., find line intersections, interconnect line segments to form polygons) or handle sophisticated database operations (e.g., natively support relationships between tables).

ArcPad supports the limited display of raster images and shapefile-format vector files. Raster images are shown as opaque layers and cell values cannot be queried, thus they are most useful as backgrounds to vector data. Shapefile features are symbolized simply: polygons must be empty or filled with a solid opaque color, lines cannot be dashed, and basic geometric shapes or TrueType characters are used for points. Individual point symbols can also be rotated based on the value of a feature attribute. Shapefile features can be created, deleted, or moved. Editing of the geometry within individual features (e.g., moving part of a line) is difficult or impossible. Attributes of shapefile features can be edited through forms where values are entered either by hand (using an on-screen keyboard or through handwriting recognition software) or by choosing values from picklists that allow standardized attributes. Multiple shapefiles can be displayed as superimposed layers.

GDA (Figure 1, Table 1) extends ArcPad with custom XML (St. Laurent, 2001) and VBScript code. It adds 4 predefined layers in shapefile format (Stations, Structures, Geolines, Mylar), a handful of external DBF tables (Geologists, StrucTypes, Units, Photos, Samples), a toolbar to readily add new features and manipulate these layers and tables (Table 2), forms for creating and attributing features, and code to enforce one-to-many relations between the Stations shapefile and external subtables (Samples and Photos, see Table 1). Optionally, a free-form text file of extended notes can be created for any station.

Most GDA picklists are stored within external DBF files and are related to a particular shapefile's data entry form through VB script. GDA incorporates code so that some of these lists are self-updating. For example, at the start of the day the choices for Lithology of a certain unit might be `gravel` and `sand`. If one enters, as free text, `gravelly sand` for Lithology, the next time that particular unit is mapped the picklist for Lithology will have three values, `gravel`, `sand`, and `gravelly sand`. Each instance of GDA supports independent picklists—though as a mapping project evolves, it may be desirable for project participants to standardize picklists through compilation and discussion.

GDA can support multiple ArcPad mapping projects on a single PDA. Each project is a separate directory in the file system and has its own map projection, data layers, base layers, and picklists.

Because ArcPad performance declines with large, editable layers, and to add discipline to the process of uploading data from the PDA to the PC where map compilation occurs, GDA provides an archive feature. When invoked (typically daily), current data-layer shapefiles are moved to a new subdirectory; DBF files that define picklist vocabularies are copied to the same subdirectory; and empty, data-layer shapefiles are installed in the working directory.

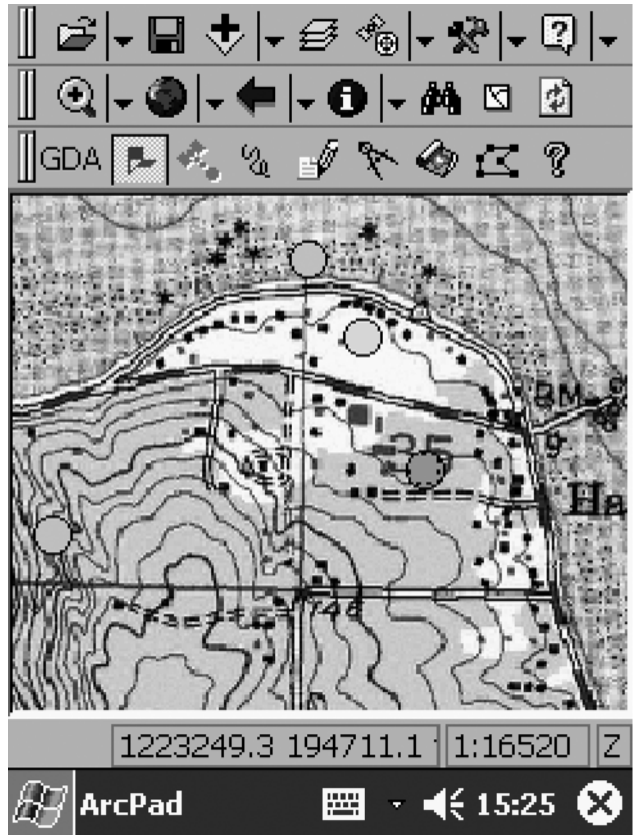


Figure 1. A GDA screen view. The upper two toolbars are standard ArcPad, whereas the lower toolbar is created by GDA. The map shows a digital raster graphic (DRG) of a standard 1:24,000-scale quadrangle map as backdrop, and dots, colored by map unit, at previously established station locations.

Hardware and software needed

To use GDA in the field, one needs a Windows CE / PocketPC-based PDA with daylight-readable color screen (e.g., a Hewlett-Packard (HP) IPAQ). Some form of non-volatile memory for field data, coupled with a large amount of memory for images, is highly desirable. We add to the PDA an expansion pack with an extra battery and support for a Compact Flash memory card.

A GPS unit that communicates with the PDA is useful. For mapping at scales of 1:12,000 and smaller we use credit card-sized recreational GPS receivers (circa 10m accuracy, with averaging) that have no display and communicate via Bluetooth (short-range high-frequency radio; http://www.bluetooth.com) with the PDA. Because the PDA often is carried in a vest pocket, a GPS unit inside the PDA would be shadowed by the user's body and thus is less desirable.

An IPAQ with expansion pack provides 12-14 hours of battery life. The Bluetooth-compatible GPS units

Table 1. Files associated with the GDA ArcPad extension.

ArcPad SYSTEM FILES

Name	Description
ArcPad.apx	XML configuration file that controls the appearance of ArcPad.
ArcPad.vbs	VB script routines that run on map-level and GPS related events.
ArcPadPrefs.apx	XML file that contains user preferences and paths to default directories.
ArcPad.apm	ArcPad map file.

GDA TOOLBAR FILES

Name	Description
GDA.apa	XML file that controls the appearance of the GDA toolbar.
GDA.vbs	VB script routines that manage the toolbar.

SHAPEFILES

Name	Feature Type	Form File	Code File	Related Picklists
Stations.shp	point	stations.apl[1]	stations.vbs[2]	geologists.dbf, units.dbf, grainsizes.dbf, samptype.dbf
Structure.shp	point	structure.apl	structure.vbs	stype.dbf
Geolines.shp	line	geolines.apl	geolines.vbs	
Mylar.shp	line			

[1]ArcPad layer definition file, in XML
[2]VB script file of routines that manage the layer's data entry form

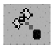
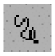
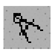
SUBTABLES

Name	Related to:	Relationship Type
Samples.dbf	Stations shapefile	one station, many samples
Photos.dbf	Stations shapefile	one station, many photos

PICKLISTS

Name	Bound to:
Geologists.dbf	Concatenated with date/time string to create a unique station id
Units.dbf	'Mapunit' field in Stations.shp
Grainsizes.dbf	'Min-/MaxGrnSz' fields in Stations.shp
StrucType.dbf	'StrucType' field in Structure.shp
SampType.dbf	'Purpose' field in Samples.dbf

Table 2. Tools on the GDA toolbar.

Name		Description
	AddStation	User adds a point where the station is located, and the Stations data entry form appears.
	StationAtGPS	Station is located by GPS, and the Stations data entry form appears.
	AddLine	User adds a line to Geolines.shp, and the data entry form appears.
	DigMylar	User adds scratch lines, annotation, sketches, etc. to Mylar.shp.
	OffsetStation	User provides a bearing and distance from current station to the next, and the Stations data entry form appears.
	Archive	Current data shapefiles and .dbf files are copied into an archive directory, and the map is re-opened with empty shapefiles.
	ToggleDrawToolbar	Toggles the ArcPad Draw toolbar on and off to either quickly expose shape-editing tool or maximize screenspace.

have internal batteries that run for 6-7 hours before they need recharging. With ruggedized case, total cost for this hardware is about $1,000. Long traverses might require an external battery or a second GPS unit. Bulk and weight of this configuration are minimal (Figure 2).

ArcPad version 6.0.3 costs $495 for a single-use license. An evaluation copy with full functionality—but that must be restarted after 20 minutes—is available for free download from the ESRI web site. GDA is freely available on the web (Thoms and Haugerud, in preparation).

A DBF editor is handy for creating initial picklists and for modifying picklists outside of an ArcPad mapping session. A DBF editor that runs on the PDA would be ideal, but we haven't found one. We use Microsoft Excel on a Windows PC.

You can use GDA in the field with this hardware and software alone. To create custom base maps, to specify symbology, to compile multiple GDA data sets, and to create a compilation database with full topology—the digital equivalent of a compilation map—from GDA notes, most users will find it necessary to have a full-fledged PC—either laptop or desktop—with ESRI's ArcGIS available at camp or back in the office.

Uploading and compilation

We upload GDA data to the PC by copying each completed archive directory. We commonly do this by removing the Compact Flash card that hosts the GDA project directories from the PDA and installing it in the PC, then copying the archive subdirectories to the PC's hard drive. When appropriate, we incorporate digital photographs and other non-GDA digital data elements into the archive directory.

How we import GDA data into the compilation database on the PC depends on the GIS that is used for compilation. In ArcInfo, we run a script that converts the shapefiles in the archive into coverages. In ArcGIS, the shapefile features are appended to geodatabase feature classes in either ArcMap or ArcCatalog.

On the PC we then symbolize stations with spots of color that correspond to their geologic map unit. With map units thus color-coded at stations, with field-observed contacts as guides, and perhaps with other resources such as topography and aerial photography at hand, the geologist then digitizes more extensive linework (contacts and faults) and ultimately closes these lines to create unit polygons in the compilation database.

Customizing GDA

Some customization (symbolization, map projection, picklists) of GDA is easy, and necessary for GDA to be useful. Specification of symbolization, definition of map projection, and optimization of base materials are best accomplished in ArcGIS, then exported using the ArcPad Tools for ArcGIS extensions that ESRI supplies with ArcPad. Most picklists are easily modified and extended by editing the associated DBF files. Some picklists are defined in the .apl (layer-definition) file for a particular layer (Table 1). These files can be modified with any text editor (e.g., WordPad). Thoms and Haugerud (in preparation) provide instructions. Other customization (e.g., modifications to database structure and forms) requires a skilled programmer, probably working with ArcPad Application Builder.

Figure 2. GDA hardware. Top: Silva compass for scale. Middle (from left to right): ruggedized case, Hewlett-Packard IPAQ PDA with CF-card expansion pack, Socket Bluetooth GPS receiver. Bottom: the top of a scepter improvised from plastic pipe fittings that drops into a rucksack and carries the GPS receiver above shoulder level.

SOME DESIGN CHOICES

The process of geologic mapping, the opportunities and constraints associated with explicit use of a structured database for storing geographic knowledge, and the capabilities of ArcPad conflict at certain points. While designing GDA we had to navigate these conflicts and make some compromises. We focus here on five design choices that are key to making GDA work.

Field notes, not a geologic map constructed in the field

When we began working with ArcPad, we assumed that because we were taking a GIS into the field we could treat fieldwork as a direct extension of building the compilation database in ArcInfo or ArcGIS (Black and Walker, 2001; Brimhall and Vanegas, 2001; Brimhall and others, 2002; van de Poll and Parsons, 2003; Geopad, http://geopad.org). This didn’t work for us. First, ArcPad doesn’t have the GIS functionality needed to build a completed units-contacts-faults layer for a geologic map. (Note that these cited efforts are laptop-based.) Second, we found that we soon had two versions of the same map database, one on the PDA and one on the PC, and no straightforward protocol for incorporating one into the other. Data remained on the PDA in volatile main memory–one day the battery on the PDA failed, and several days’ work was lost.

Thus, we learned that we needed to distinguish clearly between field data collected via PDA and the compilation database. Rather, field data stand alone as one or more separate databases that underlie the compilation database. This is consistent with traditional mapping practice. As Pavlis and Little (2001) suggest, “a standard map compilation step … may be a preferred method to insure map accuracy.”

While our decision to not construct a geologic map in the field primarily reflects our innate conservatism and the limitations of ArcPad, it is also a step towards clarifying the distinction between observation (field data) and inference (completed geologic map database)—a distinction that geologists don’t always make. (Note, however, that as in most other sciences, most of our so-called observations are actually sophisticated abstractions or inferences.)

A partly-normalized database: Locations in one layer (almost)

A database that is "normalized"–stores each elemental fact (observation, inference, etc.) once and only once–is easier to update and less prone to accumulate errors than an "unnormalized" one. For example, if locations are stored both with the list of samples collected and with the list of map-unit observations, and a location is later corrected, it must be corrected in both the sample list and the map-unit observation list for the database to remain error-free. This can be difficult to achieve. Ideally, locations would be recorded once in a single file (i.e., shapefile) and then referenced as needed. For example, a single location might have multiple lithologies, map units, and samples; thus lithology, map unit, and sample information would be stored in related tables.

ArcPad does not enforce the relations between tables that would be required in such a database. Furthermore, ArcPad cannot symbolize observations (such as map unit) that are not stored directly in a shapefile. Thus we compromise, and store location, map unit, some outcrop information, lithology, and Station ID in the Stations (location) layer. We rationalize this practice by noting that a geologic mapper should observe and record geologic map unit at every possible occasion. Furthermore, if stations have infinitesimal extent there can be only one lithology and one map unit at each station. Multiple map units and lithologies correspond to multiple, closely spaced stations—which while not practical on an analog map, are easily accommodated in a digital map database.

Data about samples and photographs are recorded in separate tables without location information. Station ID is used as a key to relate sample and photo data to the Station layer. GDA includes code to explicitly enforce one-to-many relations between records in the Station layer and the records in these tables. We have not provided tables for other occasionally-repeated observations, such as property ownership and access.

GDA supports freeform text notes by opening, as requested, a text file named with the Station ID. Multiple notes can be appended to such files; associated with each note are the date and time at which it was initiated. Similarly, a sketch file can be created for any station. All related files are stored in the same folder as the associated shapefiles.

Because it is important to symbolize structural observations on the fly, and because ArcPad cannot symbolize observations that are not part of a shapefile, we repeat locations that correspond to structural observations in a second, Structure, shapefile. There are no locations in the Structure shapefile that are not also in the Stations shapefile, thus the Structure shapefile can later (within a more powerful GIS) be reduced to a location-less table related, by station ID, to the Stations shapefile. Our data structure for observations at a point is depicted in Figure 3.

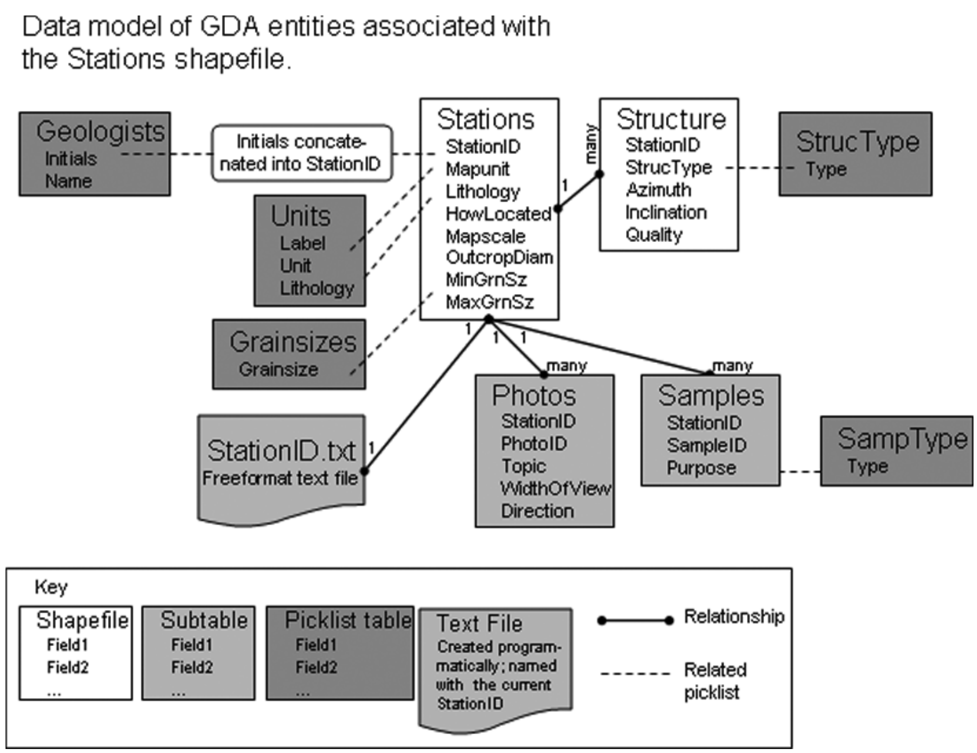


Figure 3. Data model of GDA entities associated with the Stations shapefile.

Pick-lists: Project-specific vocabularies

The appropriate data structure to be used for description of geologic features and the vocabulary used to populate this data structure are controversial. There are great benefits to be gained from a universal data structure populated from a closed (controlled) vocabulary. On the other hand, there is no agreement on what this data structure and vocabulary should be: after years of discussion to resolve such inconsistencies, the geologic community does not agree on definitions for the common terms mylonite and Holocene. Precise geologic communication still requires explicit definition of terms, usually by incorporation of external definitions (e.g., 'This report uses the GSA time scale and Streckeisen's nomenclature for igneous rocks.'). The "standard terminology" problem is most acute for lithology and map-unit attributes.

The geologic community has expended much effort on standardized map-unit vocabularies (e.g., USGS, undated; British Geological Survey, undated) and there is a strong argument to be made for incorporating appropriate subsets of these vocabularies into a digital geologic mapping tool. On the other hand, much geologic mapping is essentially stratigraphic research aimed at modifying and extending these vocabularies. It would be unfortunate to hobble this research with closed, or difficult-to-extend, vocabularies.

Working with pencil, notebook, and field sheet, we record extensive outcrop-level descriptions of lithologies and map units mostly during the early stages of a project when we are learning or creating the local stratigraphic vocabulary. We discover which lithologic attributes distinguish map units. For mapping in one area, we learn that shale-sand ratio, syndepositional deformation, and sheen (corresponding to the grain size and degree of segregation of microscopic metamorphic phyllosilicate grains) are important discriminators. In another area we find it useful to concentrate on observing the presence or absence of visible titanite and the extent of post-crystallization deformation of biotite. In another project, we note whether sandstones are cross-bedded and zeolitic or plane-bedded and altered to low greenschist facies. Project-specific shorthand vocabularies for lithology and map unit eventually emerge.

Explicit recording of such attributes in a form-driven database would require the development, in the early stages of each mapping project, of a project-specific Station-layer DBF that included the relevant attributes and new forms and picklists to populate the DBF. Alternately, one could create a Station-layer DBF, forms, and picklists that encompass the universe of all possible attributes, but the vast majority of these attributes would remain unspecified at most stations and the system would be impossibly unwieldy. A universal Station-layer DBF with lithologic attributes restricted to manageable dimensions (e.g., a few dozen rock types, dominant grain size, and color) would make it impossible to document important distinctions between map units in many areas.

We found it useful to mimic our paper-based protocol in GDA. Extensive descriptions are recorded in free-format text files. Database fields are used for a few simple shorthand attributes, including map unit and lithology, which are populated from picklists with open, project-specific vocabularies. For simplicity and flexibility we sacrifice the benefits of discipline-wide controlled vocabularies.

Multiple projects, multiple directories. Archive to subdirectories

A GDA project can be defined as data with a particular extent, map projection, symbolization rules, and set of vocabularies. We have adopted the convention that each GDA project resides in its own working directory. Each project directory contains an ArcPad map definition file that defines the current layers and extent of the project; current GDA shapefiles; an optional GPS tracklog shapefile that records traverse routes, generated by ArcPad; DBF files for samples, photos, and picklists; and optional per-station files of extended notes and sketches. Also present are subdirectories for base data, earlier geologic mapping which is used as a backdrop for the current project, and blank GDA template layers. Data for a project are archived by moving data shapefiles, the tracklog shapefile, sample and photo DBFs, and copies of the current picklist DBFs to a new archive subdirectory on the PDA. These archive subdirectories are later copied to the PC for incorporation into the project compilation database. This convention allows us to maintain multiple projects on a single PDA, each with its own map projection (e.g., State Plane for one project, UTM for another) and its own lithologic and stratigraphic vocabularies.

Unique, independently-calculated Station IDs

Initially we used record number in the Station layer as the Station ID. It quickly became apparent that merging multiple record sets (from multiple workers, or multiple sets of field data from one worker) would require complicated and error-prone updating of the Station IDs in the Stations layer and in all the related tables in order to keep Station IDs unique. It is much simpler if, *a priori,* all Station IDs are unique, at least within the universe of a single mapping project.

We found it difficult, within ArcPad, to maintain a continuously incrementing station number across multiple projects and multiple archive events. Therefore, GDA uses Station IDs of the form IIYYMMDDHHNN, where II are the initials of the geologist making the observation, YY are the last two digits of the year, MM is the month of the year (01-12), DD is the day of the month (01-31), HH is the (24-hour time) hour (00-24), and NN is the minute (00-60). Advantages of this convention are:

- Station IDs can be calculated independently and simultaneously by several different workers and remain unique. The last (or next) Station ID need not be carried from one ArcPad session to another,
- Station IDs encode the time and date at which the location was established, and
- Station IDs can be sorted alphanumerically and their chronologic sequence recovered.

Multiple stations created within a single minute, such as offsets generated by a laser rangefinder, could be distinguished by appending additional characters (seconds, tenths of seconds). An alphanumeric sort would still recover station sequence. GDA uses a field length for Station ID that is longer than presently needed, to allow for this expansion.

SUMMARY

GDA uses the capabilities of ArcPad to turn a lightweight, low-cost PDA into an effective field tool for collecting geologic mapping data in digital form. Implementation of the GDA interface required comprises between desirable database structures, the ideal user interface, and the capabilities of ArcPad. In making these compromises we were in large part guided by previous experience with pencil-and-paper mapping.

We expect that many geologists will find GDA sufficiently complete and flexible to be useful in its present form. We hope that others will build upon our effort and experience to create even better mapping tools.

ACKNOWLEDGMENTS

We thank Ray Wells for continued support of our efforts and encouragement to collaborate. Jordan Hastings and David Soller provided very useful reviews of this manuscript.

REFERENCES CITED

Black, R., and Walker, J.D., 2001, Development and use of a laptop-based geological mapping system: experiences at the University of Kansas, *in* Soller, D.R., editor, Digital Mapping Techniques ’01—Workshop Proceedings: U.S. Geological Survey Open-File Report 01-223, p. 127-131, accessed at http://pubs.usgs.gov/of/of01-223/black.html.

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* Soller, D.R., editor, Digital Mapping Techniques ’01—Workshop Proceedings: U.S. Geological Survey Open-File Report 01-223, p. 103-114, accessed at http://pubs.usgs.gov/of/of01-223/brimhall.html.

Brimhall, G.H., Vanegas, A., and Lerch, D., 2002, GeoMapper Program for Paperless Field Mapping With Seamless Map Production in ESRI ArcMap and GeoLogger for Drill-Hole Data Capture: Applications in Geology, Astronomy, Environmental Remediation, and Raised-Relief Models, *in* Soller, D.R., editor, Digital Mapping Techniques ’02—Workshop Proceedings: U.S. Geological Survey Open-File Report 02-370, p. 141-152, accessed at http://pubs.usgs.gov/of/2002/of02-370/brimhall.html.

British Geologic Survey, undated, The British Geological Survey Lexicon of Named Rock Units, http://www.bgs.ac.uk/lexicon/lexicon.html, accessed 9 February 2005.

Brodaric, B., 1997, Field data capture and manipulation using GSC Fieldlog v3.0, *in* Soller, D.R., editor, Digital Mapping Techniques ’97—Workshop Proceedings: U.S. Geological Survey Open-File Report 97-269, p. 77-81, accessed at http://pubs.usgs.gov/of/1997/0f97-269/brodaric.html.

Edmondo, G.P., 2002, Digital geologic field mapping using ArcPad, *in* Soller, D.R., editor, Digital Mapping Techniques ’02—Workshop Proceedings: U.S. Geological Survey Open-File Report 02-370, p. 129-134, accessed at http://pubs.usgs.gov/of/2002/of02-370/edmondo.html.

Pavlis, T.L., and Little, J., 2001, Using handheld personal computers as field data collection tools: some lessons learned in the school of hard knocks in the Wingate Wash Project and related projects using Fieldlog/Fieldworker software exported into ArcInfo, *in* Soller, D.R., editor, Digital Mapping Techniques ’01—Workshop Proceedings: U.S. Geological Survey Open-File Report 01-223, p. 115-121, accessed at http://pubs.usgs.gov/of/of01-223/pavlis.html.

Soller, D.R., 2001, editor, 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.

St. Laurent, S., 2001, XML: A Primer: M&T Books, New York, 560 p.

Thoms, E.E., and Haugerud, R.A., in preparation, GDA (Geologist’s Digital Assistant), an ArcPad extension for geologic mapping: Code, prerequisites, and operating instructions: U.S. Geological Survey Open-File Report.

USGS, undated, National Geologic Map Database Geologic Names Lexicon “GEOLEX”, http://ngmdb.usgs.giv/Geolex/geolex_home.html, accessed 9 February 2005.

van de Poll, H.W., and Parsons, C., 2003, From CARIS GEMM 4 to Carta for Geology software: Our responses to geology education challenges in the age of digital mapping, *in* Soller, D.R., editor, Digital Mapping Techniques ’03—Workshop Proceedings: U.S. Geological Survey Open-File Report 03-471, p. 47-55, accessed at http://pubs.usgs.gov/of/2003/of03-471/vandepoll/index.html.

Walsh, G.J., Reddy, J.E., and Armstrong, T.R., 1999, Geologic mapping and collection of geologic structure data with a GPS receiver and a personal digital assistance (PDA) computer, *in* Soller, D.R., editor, Digital Mapping Techniques ’99—Workshop Proceedings: U.S. Geological Survey Open-File Report 99-386, p. 127-131, http://pubs.usgs.gov/of/of99-386/walsh.html.

Williams, V., 1997, Using the GSMCAD program with GPS for data collection in the field and as a quick and efficient way of creating Arc/Info geologic map coverages, in Soller, D.R., editor, Digital Mapping Techniques ’97—Workshop Proceedings: U.S. Geological Survey Open-File Report 97-269, p. 83-87, accessed at http://pubs.usgs.gov/of/of97-269/williams.html.

Creating 3D Models of Lithologic/Soil Zones using 3D Grids

By Skip Pack

Dynamic Graphics, Inc.
1015 Atlantic Avenue
Alameda, CA 94501
Telephone: (510) 522-0700, ext. 3118
Fax: (510) 522-5670
e-mail: skip@dgi.com

THE PROBLEM AND THE GOAL

Aside from describing and recording surficial geology, mapping of the subsurface is, for geologists, one of the most frequently performed tasks. Traditionally, surface contacts are integrated with data gathered in vertical borings, usually drillers' logs, and correlations are manually interpreted on sections between the borings. Contour maps can then be made, manually, or by computer from the surface and correlated borehole data. In this paper, I will describe a work path using the same data, but bypassing the manual correlation step, to create a 3D model of zones (volumes) representing soil or lithology categories or types, or even larger scale geologic zonation. Maps of zone tops, if needed, can then be calculated from the 3D model with rigorous interzone consistency. This involves a pastiche of known techniques, and offers real advantages in some cases, and none in others.

The process using traditional manual correlation works very well when each zone can effectively be defined with a top and bottom surface, that is to say there is little or no lobing or interfingering of zones, and when zones extend far enough to be encountered in several borings. If the geologic environment is relatively chaotic over short distances relative to the spacing of borings, traditional correlation becomes difficult or impossible. Glacial and fluvial environments commonly produce deposits that are difficult to correlate over even short distances.

In our method, we begin by creating an indicator variable for each possible soil or lithology type at each data location, whose value simply indicates whether each type of material is present or absent. Using those indicators as input to 3D grid calculation results in a grid which estimates the likelihood of occurrence of each soil/lithology type at each node of a 3D grid. Because each grid was calculated independently from those representing other soil/lithology types, there will be locations away from the sample points where more than one type is shown to probably exist. This requires a reconciliation process, which I will describe, that determines which soil/lithology type is most likely at each location. Finally, I will describe a method for labeling each separate occurrence of a given type, rather than grouping them. For example, we can create a number of zones, each representing a distinct, separate sand lens, rather than lumping them all together as a single discontinuous sand zone.

Several organizations have found this method useful when manual correlation is difficult or impossible to achieve. I also believe that this method may prove useful as a precursor to manual correlation. Once a model has been created using this method, model-derived cross sections can be generated on traverses using the boring locations as vertices. Use of these cross sections as background reference for the manual correlations may offer time savings over the purely manual process. The geologist starting with the model-derived sections would need only to focus on correcting the geologically implausible aspects of the automated method output.

I have used EarthVision (EV), a geologic modeling software package from Dynamic Graphics, Inc., my employer, to implement this process. I will, in this discussion, focus on the conceptual process, which could be implemented using tools and software packages from other sources.

A SIMPLE TEST CASE

This technique requires a data set of lines, each having numerous points, where each point indicates the local soil or lithology (i.e., each line is a continuous, vertical sampling through the soil, sediment, and/or rock in the study area). Vertical borings are the most common source of such data, and can be conventional wells, test borings, or data gathered by direct push technologies such as cone penetrometers. In many cases, the data points are derived by interpolation or expansion of the actual information available for the borehole. For example, a driller's log indicating the top of each zone implicitly states that more or less the same material exists until a change is logged that indicates a new material. A script or spreadsheet would then be used to fill in the intervening interval with the value from the last log entry uphole. The following is an excerpt from a cone penetrometer data set. The first three columns contain the X, Y, and Z coordinates of the data point and the fourth column contains an integer code for the soil or lithology type encountered. The header line that begins with # Description matches the integers 1, 2, and 3 with the soil category names. The Z field (or column) of data is expressed in units increasing upwards both above and below a 0 datum, which, in this case, is mean sea level. Figure 1 provides a perspective view of this data.

```
# Type: property scattered data
# Version: 7
# Description: 1=clay, 2=silt, 3=sand (soilcat)
# Format: free
# Field: 1 x
# Field: 2 y
# Field: 3 z
# Field: 4 soilcat
# Projection: Local Rectangular
# Units: unknown
# End:
14191624.16    1270905.512    -7.42    1
14191624.16    1270905.512    -6.77    1
14191624.16    1270905.512    -0.86    1
14191624.16    1270905.512     1.11    1
14191641.82    1270917.503    -6.11    1
.
.
.
14191624.16    1270905.512    -8.73    2
14191624.16    1270905.512    -8.08    2
14191624.16    1270905.512    -6.11    2
.
.
.
14191624.16    1270905.512    -9.39    3
14191624.16    1270905.512    -4.8     3
14191624.16    1270905.512    -4.14    3
```

The next step is to create one new column per soil/lithology type. In this case, there are three types, so we create an indicator value for each type and record it in the type-specific column. Using a spreadsheet, a script, or a suitable program, we tested the integer value in the fourth column of the original file and set the appropriate indicator column to “1” where that type occurred at that location and “0” where it did not. The following is an excerpt of that file after the indicators were added:

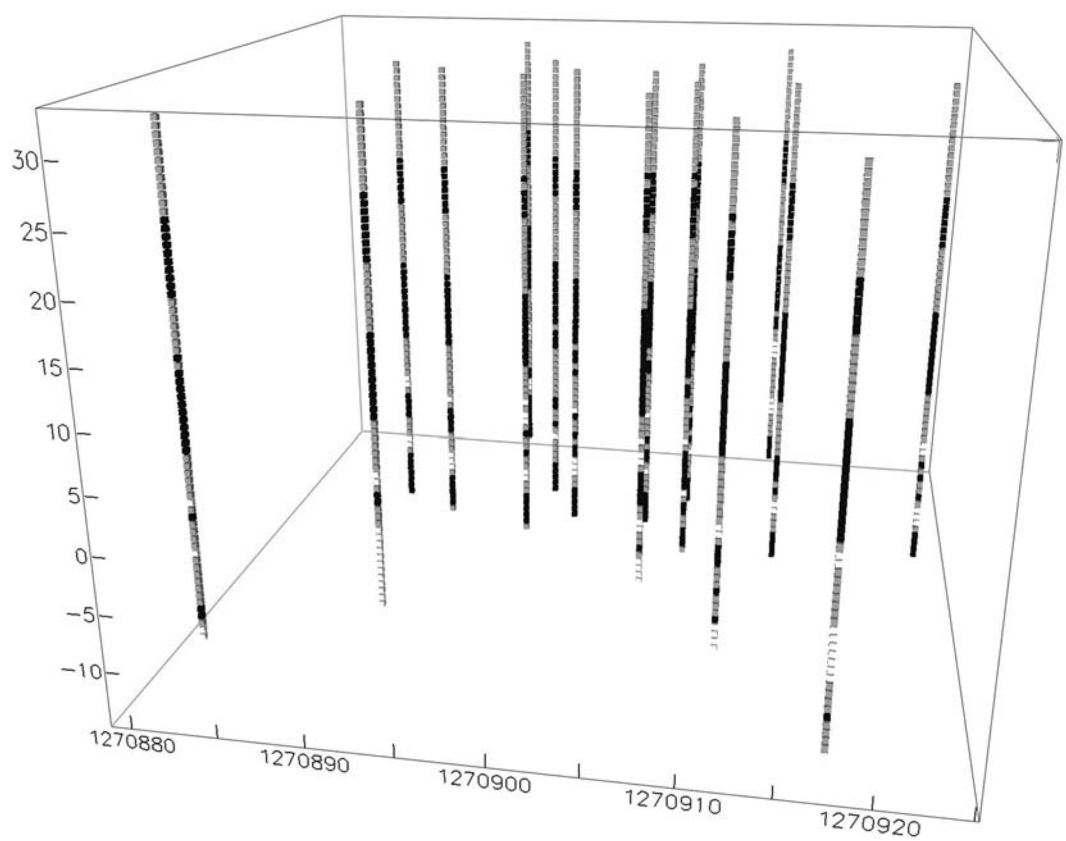


Figure 1. Input data file indicating soil types interpreted from cone penetrometer tip pressure and sleeve friction readings processed through a lookup table to determine soil/lithology type. Since cone penetrometer readings are almost continuously sampled, the data points (small cubes) in the illustration are measured points.

```
# Type: property scattered data
# Version: 7
# Description: Created with formula processor (skip, 30
Mar 2003)
# Format: free
# Field: 1 x
# Field: 2 y
# Field: 3 z
# Field: 4 soilcat
# Field: 5 I_one
# Field: 6 I_two
# Field: 7 I_three
# Projection: Local Rectangular
# Units: unknown
# End:
14191624.16    1270905.512    -7.42    1    1    0    0
14191624.16    1270905.512    -6.77    1    1    0    0
14191624.16    1270905.512    -0.86    1    1    0    0
.
.
.
14191624.16    1270905.512    -8.73    2    0    1    0
14191624.16    1270905.512    -8.08    2    0    1    0
14191624.16    1270905.512    -6.11    2    0    1    0
.
.
.
14191624.16    1270905.512    -9.39    3    0    0    1
14191624.16    1270905.512    -4.8     3    0    0    1
14191624.16    1270905.512    -4.14    3    0    0    1
```

Using this augmented file, I calculate one 3D grid per category using the indicator data for that category as input. With most 'representational' or deterministic gridding methods (which models natural surfaces or volumes as opposed to analytical gridding like trend surface analysis), the resulting grid node values will have values from around 0.0 to around 1.0; but, unlike the input values (0.0 or 1.0), they will vary continuously where the grid nodes are interpolated or extrapolated. This transformation from the discrete input values to the continuously varying grid node values creates a probability-like value that expresses the likelihood of the soil/lithology type existing at each node location.

If you choose kriging for this gridding step, you are following a very standard path that has long been used on discrete data. This is not to be confused with indicator kriging using thresholds on continuous data. In this case I used EarthVision's 3D minimum tension gridding, though I have also used simple nearest neighbor gridding, where the gridder sets each node to the value of the nearest input data point. Initial tests suggest that the results are a little less sensitive to differences in gridding techniques than single grids calculated from continuous numeric input variables, but not enough cases have been run to state that with any conviction. Kriging would be desirable when variogram analysis or prior knowledge of the area indicates anisotropic variation in the zonal orientation. Without some indication or prior knowledge regarding the existence and nature of the anisotropy, deterministic gridding methods will yield a defensible result more quickly, and, potentially, with fewer gridding artifacts.

The resulting grid, using a higher order gridder like a minimum tension algorithm, will have node values ranging from somewhat below 0 to somewhat above 1. A linear weighted average gridding algorithm likely would have values ranging from slightly above 0 to slightly below 1, unless a node happens to be coincident with a data point. In that case, 0 or 1 would be assigned to the node. In any case, the resulting grid contains values that I would call pseudo-probabilities. While I am not sure that there is any statistical rigor in their generation, they serve well when used as probabilities. When the values exceed 1 or fall below 0, their qualification as probabilities is dubious by definition, but this seems not to matter in practical terms since the comparison of the 'probabilities' of several soil/lithology types at any one node is only in question when two or more types have similar 'probabilities'. In this case, the 'probabilities are likely to be well inside the 0-to-1 range. Thus, where a soi/lithology type shows a value greater than 0.5, it is assumed to exist at that location. Figures 2 and 3 show the volume for the Type 1 (Clay) grid above 0.5 represented by blocky cells and by smooth contour surfaces. These are two representations of the same grid, with the difference in display resulting from 3D viewer options.

Reconciliation

Because these grids of individual soil/lithology type probabilities are created independently, the sum of the probabilities at each grid node location does not equal 1.0, except for those locations where a sample point is coincident with a grid node. For this reason, it is almost certain that there will be node locations where more than one type is indicated to be likely to exist. Thus, the next step is a reconciliation process to select only one type as present at each grid node location.

I have simply compared the grid node values of each of the soil/lithology type grids at each location and selected as present the type with the highest pseudo-probability. I can then create a single 3D grid containing an integer value showing which type is present, in the same way that integers were used in the unmodified input file.

This method of reconciliation has been used on a number of projects to date (approximately 6), and has served well. At most node locations, the choice of which soil or lithology type is present is quite clear, that is to say, the pseudo-probability of one type is distinctly higher that those of the other categories. However, some node locations may have pseudo-probability values for one or more types that have very similar values. At those loca-

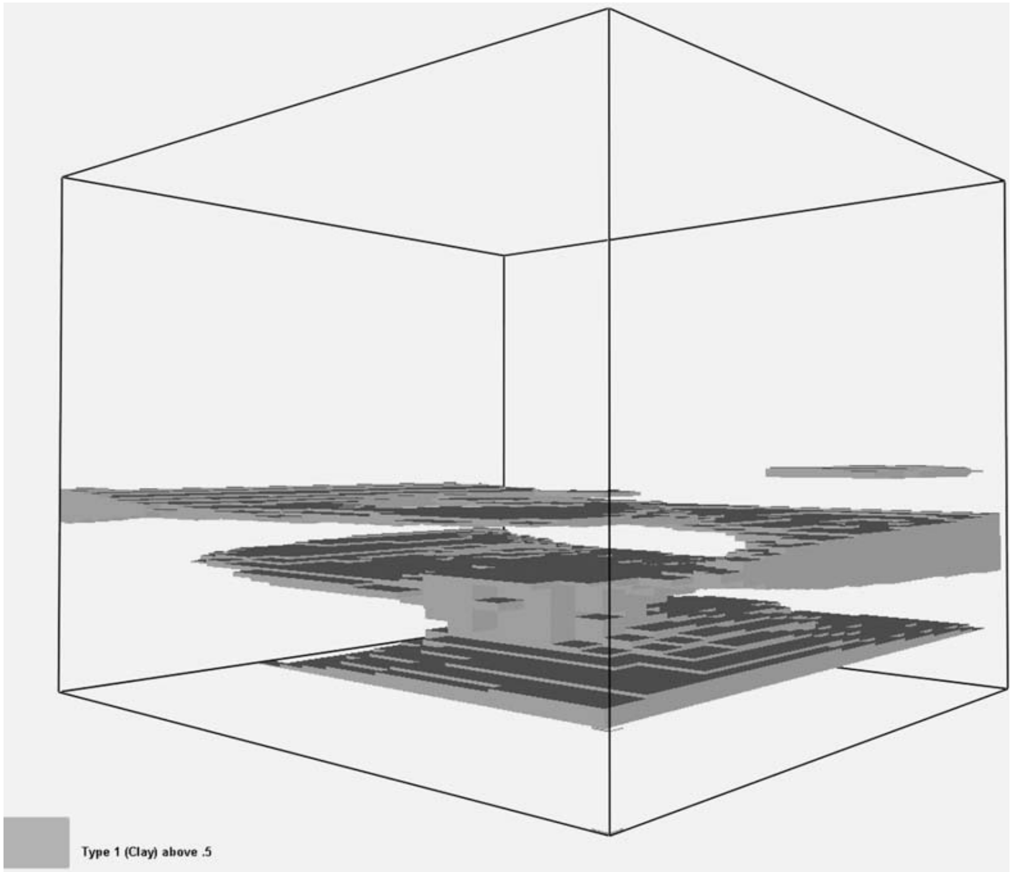


Figure 2. Type 1 (Clay) 3D grid where pseudo-probability is > 0.5 (cells shown as voxels).

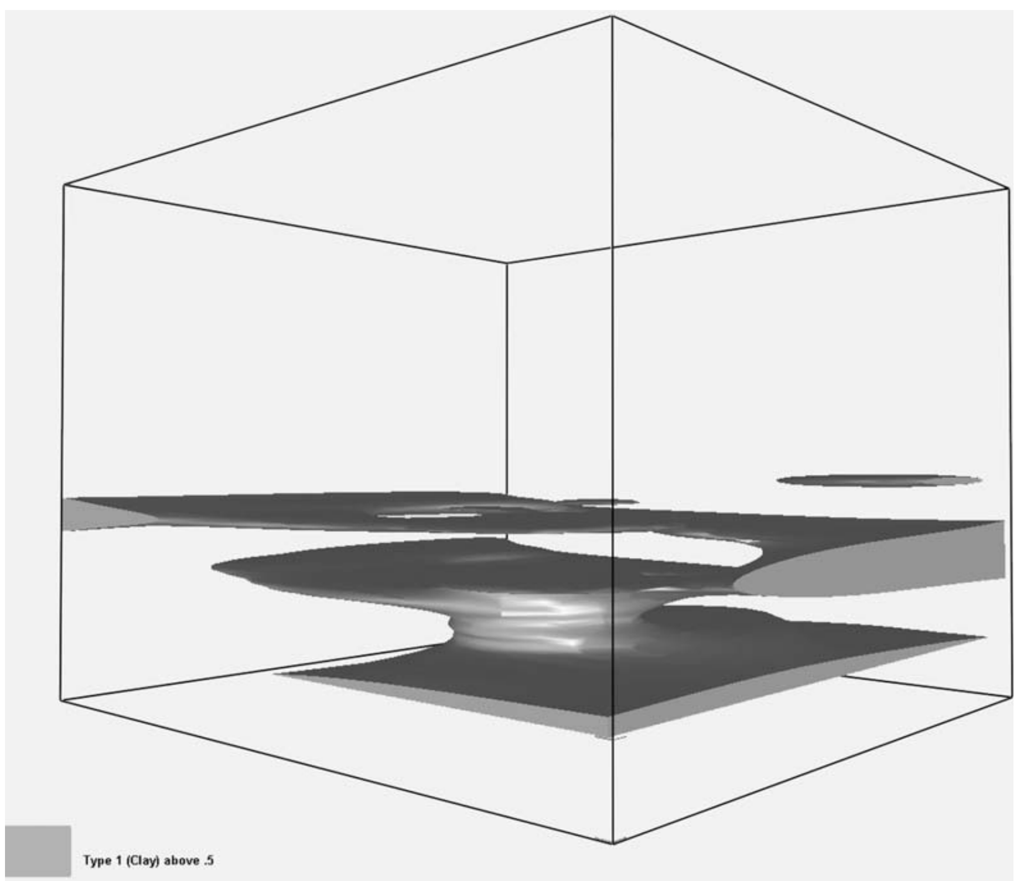


Figure 3. Type 1 (Clay) 3D grid where pseudo-probability is > 0.5 (3D oblique contouring).

tions, selecting the highest numerical is somewhat crude and questionable. Use of secondary information derived from geophysics or stochastically inferred tendencies could greatly improve the reconciliation process.

Model Building

I have, so far, generated and used two outputs from the reconciliation process. The first is the combined integer grid described above. The second is a set of individual soil/lithology type grids indicating that the type is present or absent at each grid node. Thus, with three soil/lithology types, we have three reconciled grids containing indicator codes, identical in concept to the indicator codes in the modified input file where the indicators were put into three additional fields. In this case each node of the grid indicates present or absent. Earthvision allows us to create volumes with oblique boundaries that contain the soil/lithology types from 3D grids. These volumes can then contain properties such as porosity or hydraulic conductivity (permeability) that can, in turn, vary continuously in three dimensions within each zone, but discontinuously from zone to zone. Figure 4 shows a reconciled 3D grid of combined zones cut away to display the internal variation. The 'blocky' nature of the grid is obvious. Figure 5 shows a model with the clay zone projected into a cutaway. The smoother representation using oblique triangles is based on the same grids shown in Figure 4.

You will notice the multiple occurrences of the clay zone, which may or may not be connected outside of the model volume. This leads to the next topic, differentiation between potentially separate occurrences of similar or identical soil/lithology types, that is to say, spatially separated volumes where the same type was assigned during the input data interpretation.

Clustering (3D volumetric classification)

A final, and very interesting, step in this technique is processing the combined integer grid to detect and label each volumetrically separate occurrence of each soil/lithology type. Graham Brew, of Dynamic Graphics, has developed a script that uses the 3D grid containing integer soil/lithology codes output from the reconciliation process. The output of this script is also a 3D grid of integer values, but these integers are compound labels that show both the soil/lithology type and a unique integer value for each spatially separate volume where that type

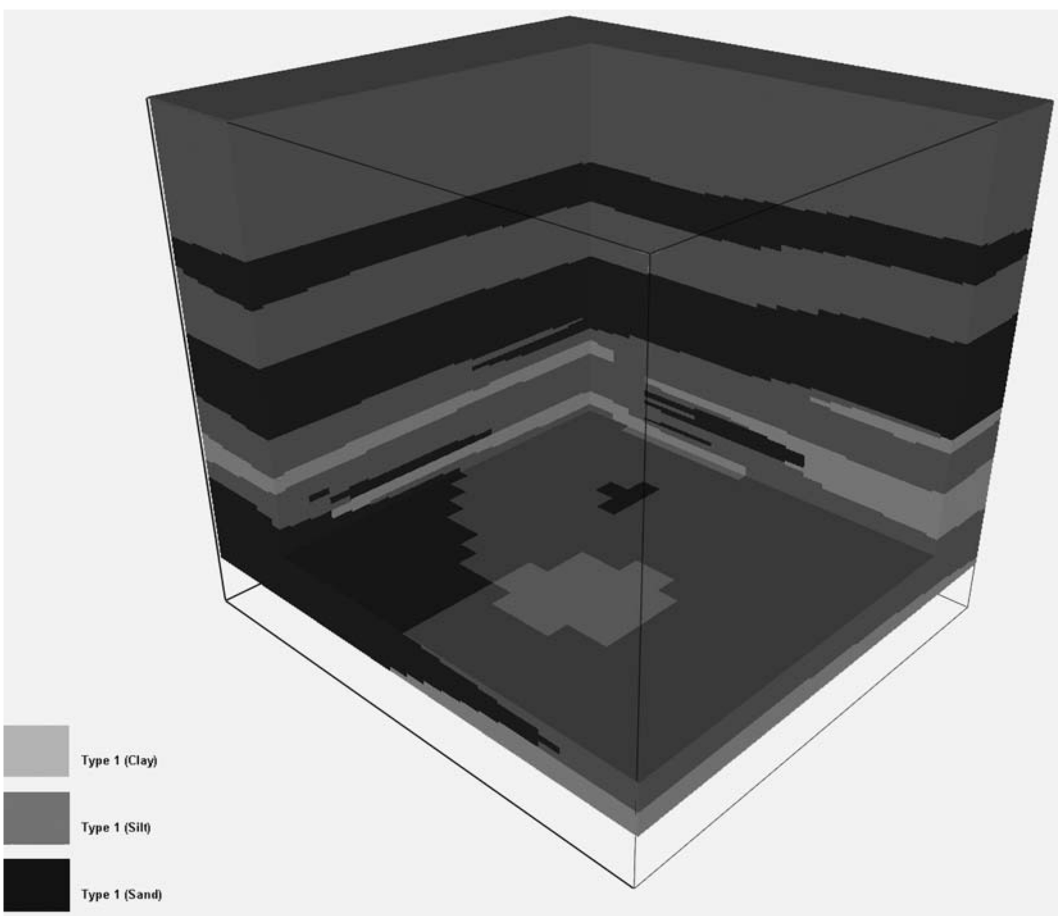


Figure 4. Combined, reconciled 3D grid of all soil/lithology types.

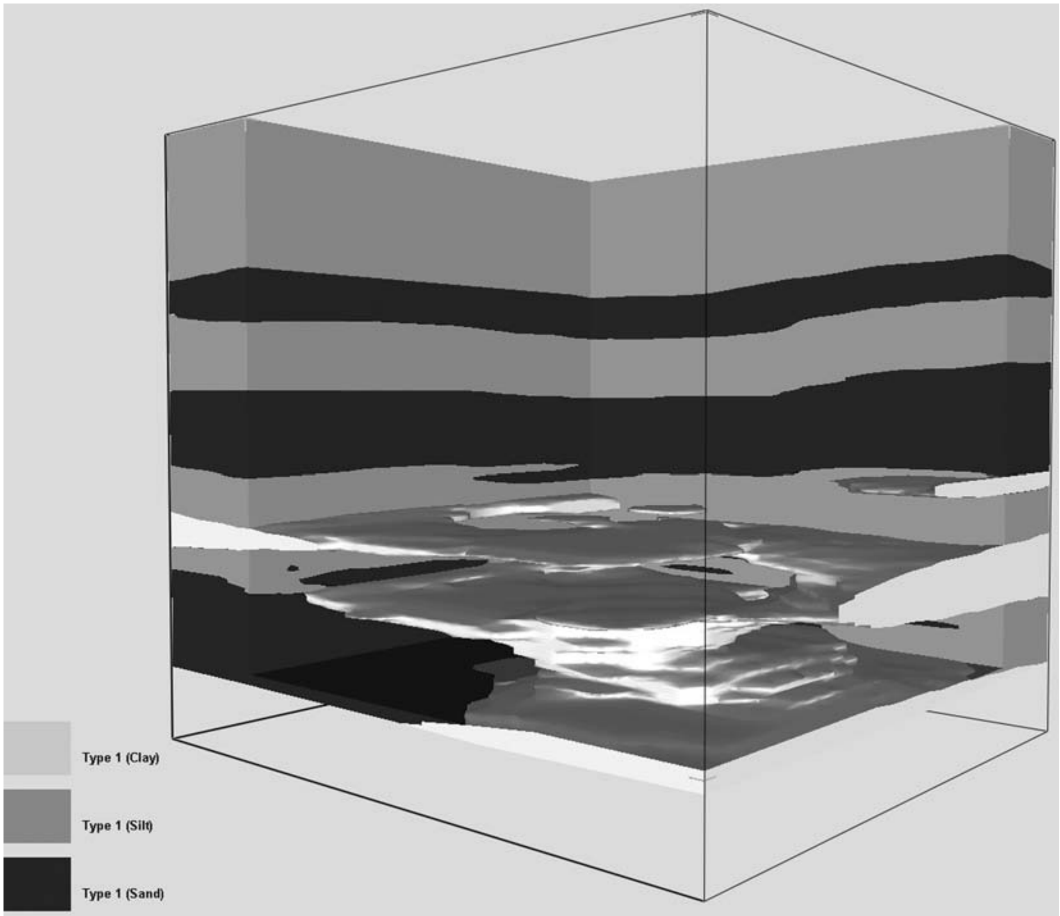


Figure 5. 3D volumetric model with clay zone displayed in cutaway.

is present. These separate volumes are sometimes referred to as 'geobodies'.

In essence, this process involves starting at one node location, determining what type is present at that node, then looking at each adjacent node and determining if the same type is present. The script repeats this process, looking vertically and horizontally until all adjacent nodes of the same type have been detected. All of the adjacent nodes are given a unique label and written to the output grid. The script then moves on to the next node in the input grid, which has not yet been labeled, and conducts the adjacency process again to build the next geobody.

In the sample case used here, the 3D grid output from the reconciliation step contained values of 1, 2, or 3 indicating clay, silt and sand, respectively. The output of the clustering script output values of 1 and 2 for each of the two clay bodies, 18, 19, and 20 for each of the three silt bodies, and 44, 45, 46, and 47 for each of the four bodies of sand. These numbers are arbitrary and have no significance beyond designating a separate integer per body while still indicating the soil/lithology type. Using this grid to generate a 3D model with oblique zone boundaries, we can see each of the sand zones projected into the cutaway model in Figures 6, 7, and 8.

There are a number of issues to resolve in the clustering process, which I will not discuss in detail. They are generally simple, and can be addressed in the script used for the process. One example is the degree of spatial adjacency of grid nodes required for assignment to the same geobody. Generally, nodes of an identical soil/lithology type, which are above, below, or directly beside, are considered to be in the same body. It is less simple when the adjacency is diagonal laterally, vertically, or both. The user needs to determine if this diagonal adjacency is sufficient to provide a geobody connection.

Similarly, it is useful to avoid classifying nodes into geobodies where the 'zone' created would be too thin or too small in volume. A frequent complication of the clustering/classification process is creation of a "noisy" volume with a large number of small, separate bodies. As always, you must select the resolution (scale or granularity?) of your classification based on the uses to which the result will be applied. The overall goal in visualization or further analysis should determine the level of generalization you select in the clustering process.

One natural result of applying rules for geobody generation, such as those discussed above using adjacency, thickness, and volume, is the generation of some number

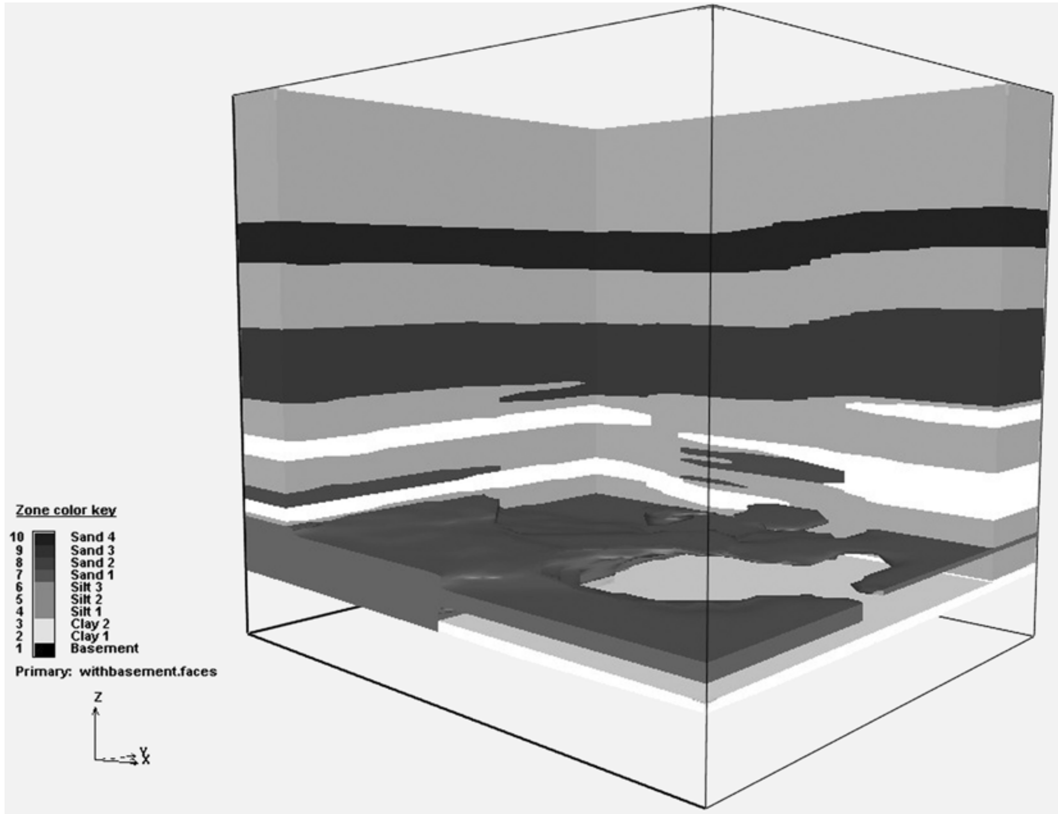


Figure 6. Sand Zone 1 projected into cutaway model.

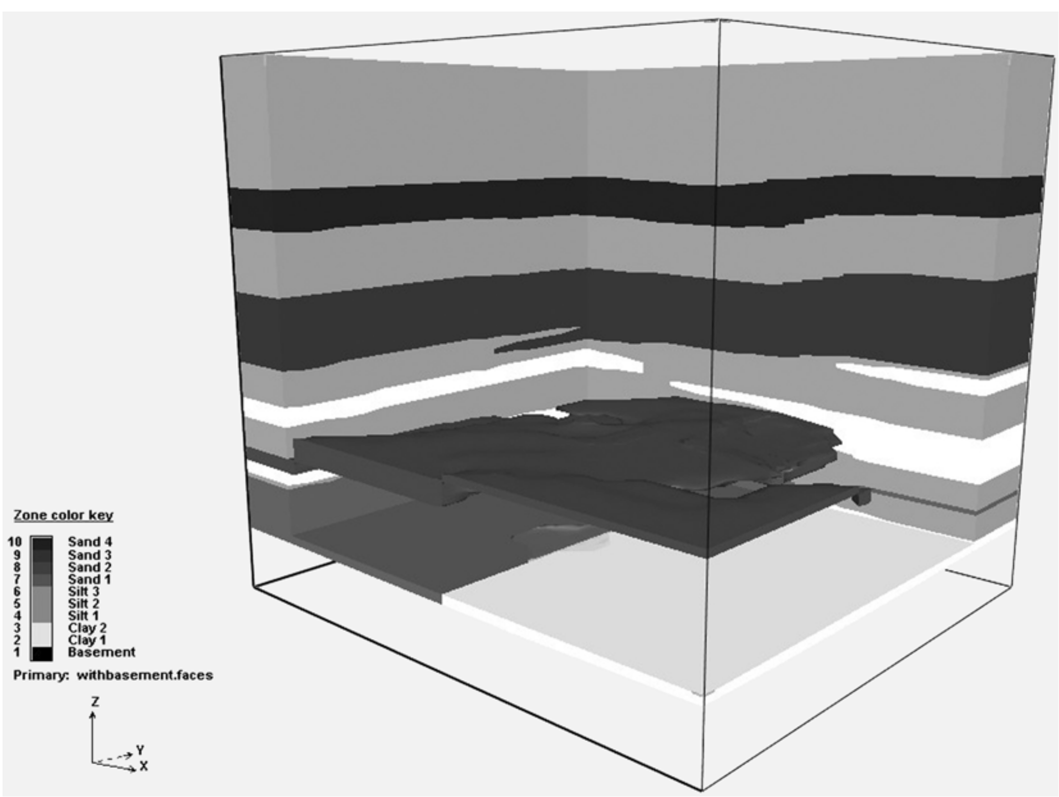


Figure 7. Sand Zone 2 projected into cutaway model.

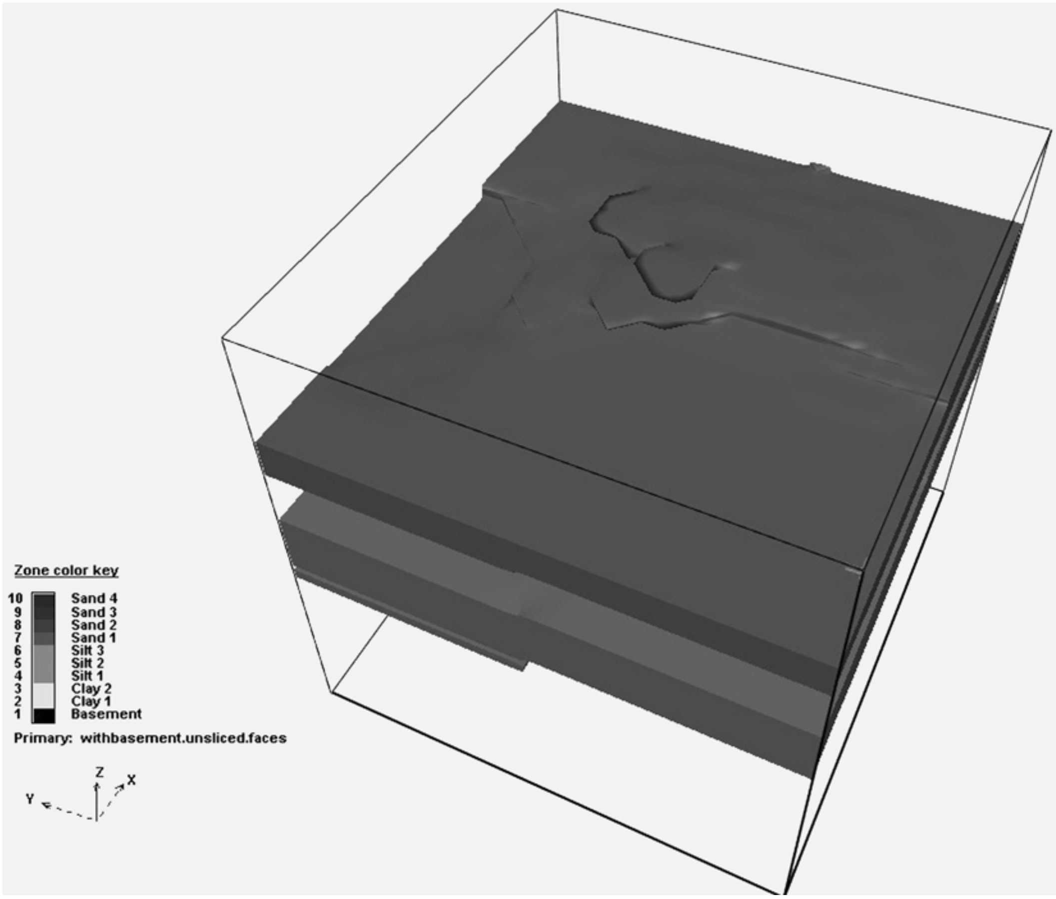


Figure 8. Sand Zones 3 and 4 shown in isolation.

of unclassified (formerly classified before the clustering) cells. Development of rules that can be used to select the most appropriate adjacent geobody type into which each unclassified cell should be merged is one of the current topics under consideration.

While the clustering process presently used in this larger technique is relatively simple and deterministic, it has worked quite well on several cases. This is a natural area for more sophisticated stochastic methods, which could allow for inclusion of anisotropic/directional continuity determination.

CONCLUSION

This general technique for generating 3D models of zones from data that indicates soil/lithology types has worked well in a number of cases. The data set and model used in this paper cover a very small area in a glacial environment. The largest number of soil/lithology types that we have used in a model with this technique is 24, and the largest range of a model thus far is 42 km x 40 km x 2 km. In the example used in this paper, another model was generated using traditional correlation prior to model generation. The traditionally correlated model had somewhat less detail as a consequence of each zone being modeled with only a top and bottom surface (no intricate lobing). In the wide-area model mentioned above, traditional methods were used to develop a Paleozoic basement surface and a top of Tertiary volcanics surface, with the 3D method described here used for five shallow, mostly unconsolidated, units in between.

Early in this paper I mentioned the potential use of this method as a timesaving precursor to traditional correlation. To date, no one who has used this method has gone back and performed the 3D model supported manual correlation/edit, though I believe this is desirable for quality control. The modest number of projects performed thus far suggest that this 3D automated technique can perhaps reduce time needed to complete the model by about 75 percent compared to traditional methods. In the other two cases mentioned above, the 24-zone case and the wide-area model, traditional methods would have been very difficult or impossible within a reasonable timeframe.

There are many promising avenues for improvement and extension of the workflow outlined above. Both the reconciliation and clustering/classification steps seem to be natural candidates for application of stochastic techniques such as those included in the transitional probability methods developed by Weissmann, Carle, and Fogg

(1999). Additional value should be available through use of geophysical data to help differentiate between soil/ lithology types when the basic method does not indicate a clear choice. Graham Brew has made improvements to the clustering methods recently, and I have begun to study the reconciliation process to gauge which node locations are candidates for use of secondary data such as gamma ray, resistivity, and statistically inferred criteria.

REFERENCE

Weissmann, G.S., Carle, S. F., and Fogg, G. E., 1999, Three-dimensional hydrofacies modeling based on soil surveys and transition probability geostatistics: Water Resources Research, v. 35, no. 6, p. 1761-1770.

Science Language, Parsing and Querying: The Surficial Side of Things*

By Andrée M. Bolduc[1], Andrew Moore[2], Kaz Shimamura[3], Matt Pyne[3], and Ron DiLabio[3]

[1]Geological Survey of Canada
GSC-Québec
490 de la Couronne
Québec, Canada G1K 9A9
Telephone: (418) 654-2635
Fax: (418) 654-2615

[2]Natural Resources Canada, ESS Info
[3]Geological Survey of Canada

e-mail: {abolduc, amoore, kshimamu, mapyne, rdilabio}@nrcan.gc.ca

ABSTRACT

Communicating geological concepts amongst geologists as well as from geologists to the general public requires that a well established and documented science language be available. This task is tackled by various groups that have an interest in making available geological data to a broad community. This paper focuses on science language needed to adequately represent geological concepts portrayed on surficial geology maps in Canada. Words necessary to describe glacial and ice-contact deposits, environments, landforms, etc., have been extracted from map legends. They have been organized hierarchically under high level concepts such as earth materials, physical characteristics, and genesis. The science language will be an evolving tool used by geologists and the general public. Therefore, it needs to have the flexibility to accommodate modifications under the supervision of an expert committee to insure that such modifications are made according to the accepted philosophy.

INTRODUCTION

Words are a powerful tool for communication, if the definition of the word is agreed upon by all users. Too many times do we hear: “in this context, we use the word (*insert your choice*) to mean (*a definition of the word*).” With the advent of digital mapping, it has become more and more important to develop common terminologies in order to facilitate the querying of maps by knowledgeable and lay users of geological information. The Science Language Technical Teams of the North America Data Model Steering Committee (http://nadm-geo.org) have undertaken this very task. We will refer to reports by the Science Language Technical Team on Sedimentary Materials (North American Geologic Map Data Model Science Language Technical Team, 2003 and their paper in this volume) as needed, using the acronym “SLTTS_1.0”.

Traditionally, surficial geology has been regarded as a completely separate geological topic, especially in Canada where it is largely equivalent to Quaternary glacial geology. This is not rigidly enforced, and one could rightly argue that recently erupted lavas are surficial materials. There remain problematic issues (pyroclastic materials are a good example) that will not be resolved here. Rather, we will treat this subject from a Canadian surficial geological map point of view, recognizing that it is incomplete. The main objective of this paper is to present the science language that is being used with the parsing and querying tools currently under development at the Geological Survey of Canada (GSC).

Recently, the point has been made that surficial geology describes unconsolidated earth material of sedimentary origin. Glacial geology is thus a subset of sedimentary geology where the sedimentation agent is ice, rather than water or wind, although both (water and wind) are involved in forming the materials and landscape of the ice contact environment. This clarification comes at a time when the GSC is developing web-based tools (http://cgkn.net/) to help a variety of geological data users to sort through the information available on maps (Moore and

*GSC contribution 2004057

others, 2003). The challenge remains: providing a coherent science language adapted to surficial geology. Specialized terminology needed to adequately describe glacial deposits, especially landforms, had not yet been developed under SLTTS_1.0. We hope to contribute to science language development in North America by providing such terminology and a coherent classification scheme. The reader should keep in mind that the rest of this paper is focused on the ultimate GSC goal, parsing map legends (by the geologist) and querying the data (by any user).

THE SCIENCE LANGUAGE

The goal of the GSC is to develop a science language to be used in the parsing and querying of surficial geology maps, supported by a glossary of all terms used. The science language needs to adhere to a coherent classification scheme. It also should be intuitively simple enough for non-expert users (novice geologist-users as well as the general public), and complete enough for experienced users. For bedrock geology, the problem was tackled using NADM-C1 (2004) as a starting point (Davenport and others, 2002; Struik and others, 2002).

The group working on the science language for surficial geology adopted a "*legend to science language*" approach, similar to the one described in Thorleifson and Pyne (2003). As a starting point, we identified the science words most commonly used in surficial geology map legends, rather than using an already existing classification. The rationale behind this approach was to quickly provide users with a working science language that would cover most of the terminology used on maps. Following two pan-Canadian workshops, the results of this exercise were organised into a flat table, the heading of each column representing first order concepts such as depositional environment or sedimentary material. Quickly, it became clear that this simple schema was not satisfying if it was to be incorporated into SLTTS_1.0.

The next step was to look at the results of the legend analysis exercise and to compare it with SLTTS_1.0. A large proportion of the words readily found a niche in SLTTS_1.0. Words for earth materials, sedimentary structures, and, to a certain extent, depositional environments already existed. We put together a series of tables, each representing a high level concept (e.g., earth materials (Figure 1), physical characteristics (Figure 2), genesis (Figure 3)). Each of the tables then proposes a hierarchical classification of terminology into lower level concepts until the words resulting from the map legend analysis exercise were reached. In most cases, words from the legends were introduced on the 4th level or lower.

Earth Material

Earth material ("*a naturally occurring substance formed in or on the Earth by physical, chemical, or biogenic processes that produce solid particles or crystals of mineral and (or) rock*"; SLTTS_1.0 (NADM, unpublished document) is the starting point for all map unit descriptions. As shown in Figure 1, following SLTTS_1.0 philosophy, this first level concept is divided at level two according to the primary genetic process, into igneous, composite-genesis (includes metamorphic rocks), and sedimentary earth material. Sedimentary earth materials are further subdivided into consolidated sedimentary materials and unconsolidated sedimentary materials (level 3). Terminology for Quaternary glacial geology resides underneath the latter. Any older consolidated glacial deposits such as tillites would reside under consolidated sedimentary earth material. There is no need to introduce additional terminology for earth material, since sand is sand whether deposited by water, wind or ice. Therefore, for surficial geology science language needs, we extract from SLTTS_1.0 the subset of terms needed to adequately describe map units on surficial geology maps.

However, a "miscellaneous" category was required to accommodate polygons that are identified as "bedrock" without further description, or ice, in areas where icefields and/or glaciers are large enough to be mapped. In the case of bedrock, we would suggest that the polygon, whenever possible, be identified at least to its 2nd level (sedimentary, igneous, or composite genesis) or 3rd level (general rock name) concept, and where possible, a link to the appropriate bedrock geology map (if web-available) should be included with the description of the polygon.

Because map units are rarely composed of a unique earth material, we suggest the use of a table of relative abundance (as proposed in NADM-C1, http://nadm-geo.org/) in conjunction with earth materials. The combination of the two tables would yield the information necessary to answer a data user query such as "*Show me, in this region, places where the surficial deposits are mostly sand*".

Physical Characteristics

SLTTS_1.0 has proposed language to describe map unit characteristics. As shown in Figure 2, we have chosen to group under the first level concept of "physical characteristics" second level concepts such as outcrop characteristics, map-unit thickness, sedimentary fabric, and sedimentary structures. Descriptions of map units are usually qualitative, but there may be some quantitative information where there exists, for example, grain size analysis or others. Therefore, we offer the parser the choice between using qualitative terminology or quantitative terminology at the same level.

Sedimentary structures are an important aspect of all sedimentary materials, but perhaps because surficial geology is partly done via air-photo interpretation,

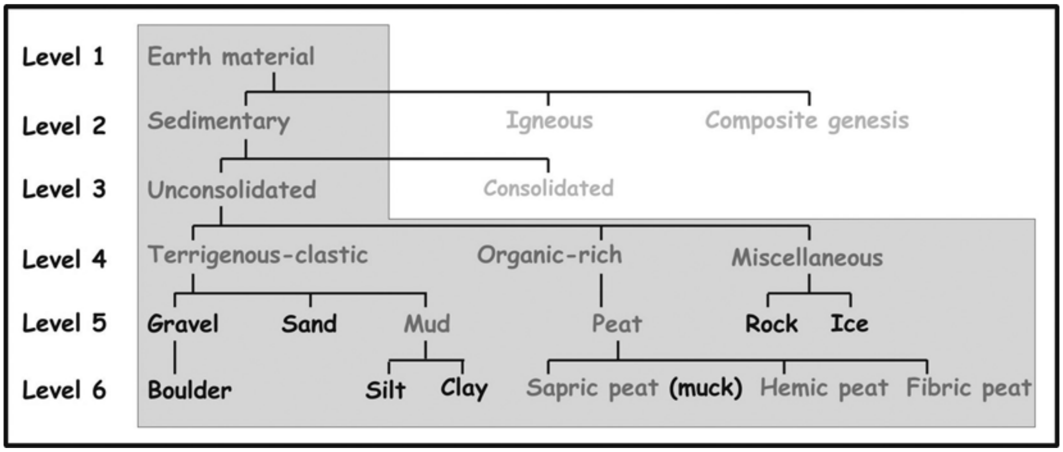


Figure 1. Classification of earth materials for surficial geology. The subset of words in the greyed area represents surficial glacial materials. Words in bold were either already in SLTTS_1.0 and in map legends, or they were added by the legend analysis exercise.

descriptions of sedimentary structures are rarely included, or briefly mentioned, in surficial geology map legends. When this information is provided in the map legend, it can be captured using SLTTS_1.0 proposed hierarchy, following the path: sedimentary structures → primary → inorganic → syngenetic (or penecontemporaneous) and depositional (or erosional) structures (Figure 2B). Again, because this paper provides science language derived from the analysis of map legends, Figure 2 only shows a very restricted number of words, but all the terminology proposed in SLTTS_1.0 is available as needed.

There are a number of features that may be grouped under secondary sedimentary structures, like those produced by the growth of ground ice. Because the development of such features is normally of a scale large enough to allow mapping of individual features, we suggest that these may be better classified as landforms. Micro-landforms such as striation could also be argued to be secondary sedimentary structures imposed on bedrock surfaces. We prefer, however, to include these in glacial landforms, where they more intuitively belong.

Genesis

Considerable attention in SLTTS_1.0 has been given to the high level concept of genesis. Depositional processes, environments, and products are three interrelated second level concepts; in Figure 3 the hierarchy of glacial processes and environments are shown. We would like to introduce a fourth concept, landforms, in order to address the needs of surficial geologists (Figure 4). The rationale is as follows.

We believe that depositional products (the map unit itself) and the landform (of which there may be one or more within a map unit) are both the direct results of a depositional process taking place in an environment. In the most simple case, the landform is a flat plain that does not contribute very much to the description of the map unit or the interpretation of the map. In the most complex case, a map unit would host many landforms that result from interrelated processes, such as the association of recessional moraines with kames and kettles (ridges, mounds and depressions to speak in non-genetic terms).

In devising the hierarchy for depositional processes and environments related to a glacial setting, we were inspired by the comprehensive work produced by the Commission on Genesis and Lithology of Glacial Quaternary Deposits of the International Union for Quaternary Research [INQUA] (in Goldthwait and Matsch, 1988), specifically those papers by Dreimanis (1988), Goldthwait and others (1988), and Lundqvist (1988). We subdivide ice flow processes and environments into “glacial” (*requiring the direct action of glacial ice*; Lundqvist, 1988) and “ice-contact” rather than “glacigenic” (*processes, depositional environments, deposits or landforms that require glacial ice, but not necessarily directly related to it*; Lundqvist, 1988). We prefer the term “ice-contact” because it is in common usage amongst glacial mappers. Under “ice-contact”, we find concepts such as glaciofluvial, glaciolacustrine and glaciomarine. We also introduce the concept of “periglacial” processes and environments for those regions where extreme cold conditions result in unique depositional products and landforms.

Clearly, the dividing line between glacio (fluvial-lacustrine-marine) and strictly fluvial-lacustrine-marine terminology requires some thought. It has been argued amongst members of the SLTTS working group that processes and environments outside of a glacier already have terminology and descriptions. Fluvial processes are the same whether the water comes from direct precipitation or from the melting of glacier ice. Similarly, lacustrine and marine processes/environments exist whether in con-

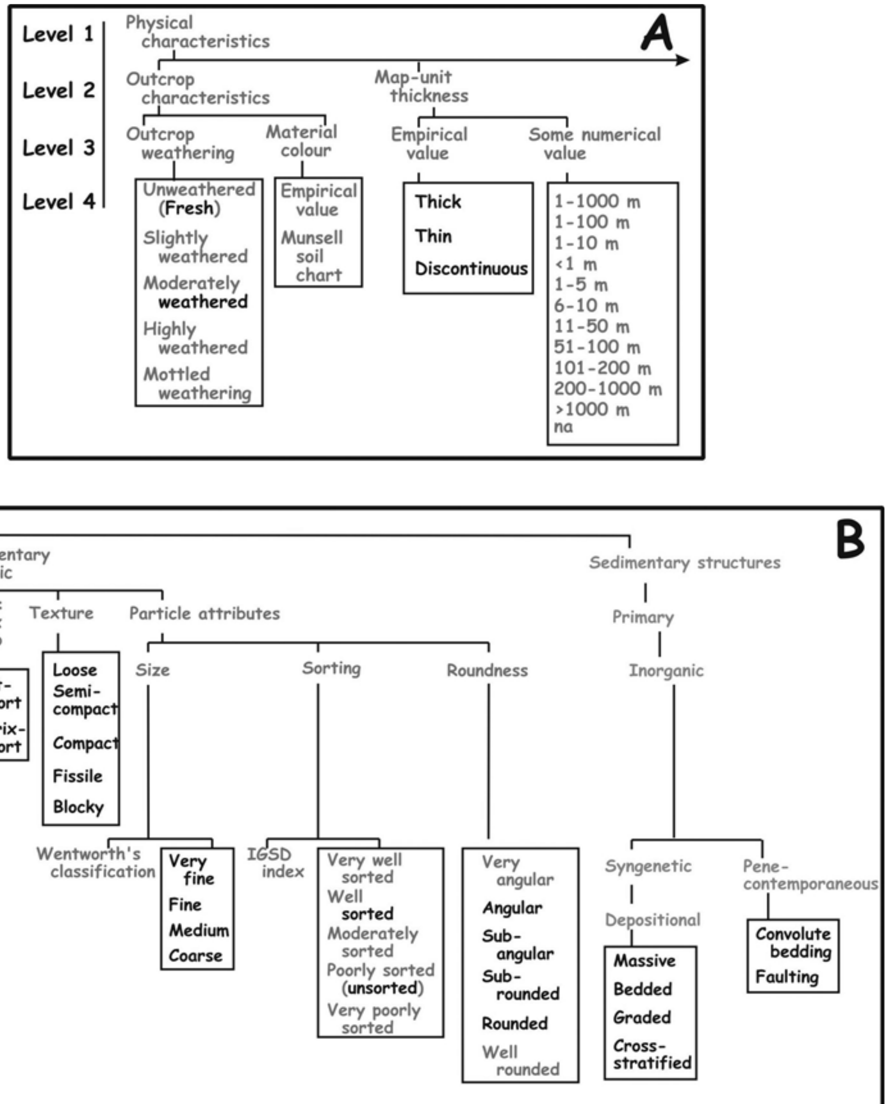


Figure 2. Classification of physical characteristics for surficial geology: A) Outcrop characteristics and map-unit thickness; B) Sedimentary fabric and sedimentary structures. Words in bold were either already in SLTTS_1.0 and in map legends, or they were added by the legend analysis exercise. It is possible to provide some choices for “material colour” such as grey(ish), yellow(ish), red(dish), etc. if the Munsell chart is not used. The distinction between map-unit thickness values “thick” and “thin” will need to be decided by the CGKN stewardship committee. Although “thin” is usually associated with “discontinuous”, i.e. with numerous bedrock outcrops found in the deposit, it may actually be continuous and range from less than 1m to 5 m depending on the map maker.

tact with a glacier or not. Therefore, these should not be duplicated under “ice flow processes”, or “glacial-related settings”. Perhaps a multihierarchical classification, such as proposed in Struik and others (2002) would better serve this complex concept of genesis.

We believe there is a logical reason to address glaciofluvial, glaciolacustrine and glaciomarine processes/environments separately, because there are some unique features associated with them. For example, landforms such as eskers and kames are not formed without the presence of glacial ice. Similarly, deposits such as varves and glaciomarine diamictons will not be formed in non-glacial lacustrine/marine environments. We thus use in Figure 4 some specialized terminology related to those transitional environments, where the “glacio-” component is critical in the development of deposits and landforms. Because there is a continuous evolution between the glacially-influenced and the non-glacially influenced, the distinction is seldom evident in surficial geology map legends. This paper will not address the non-glacial fluvial, lacustrine and marine processes/environments that are present on surficial geology maps,

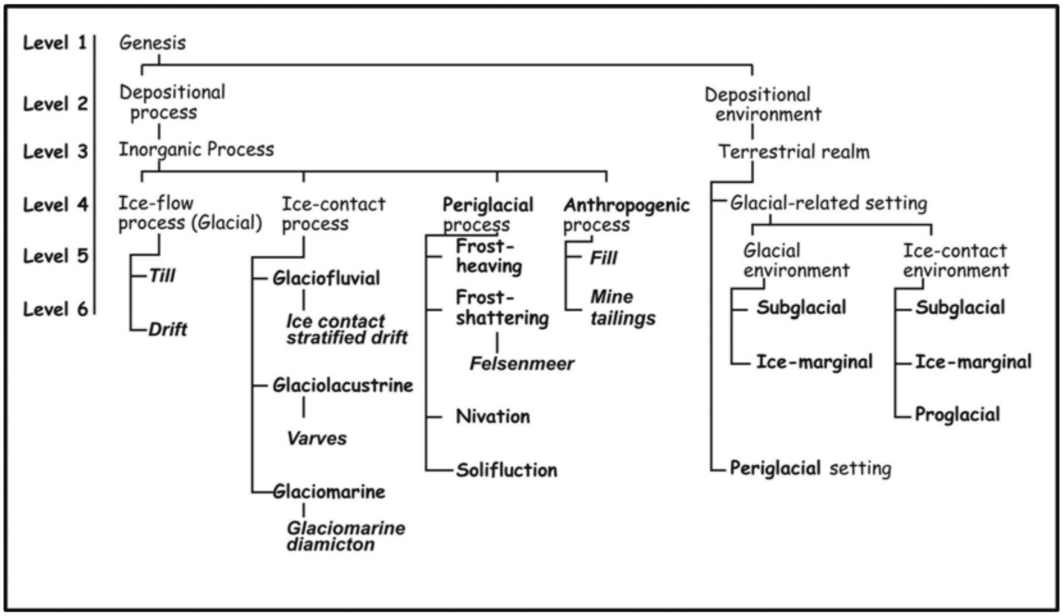


Figure 3. The high-level concept of "genesis" and its constituents, processes and environments. Words in bold-italics represent depositional products, and words in bold were either already in SLTTS_1.0 and in map legends, or they were added by the legend analysis exercise.

as we recognize that the terminology and hierarchy provided by SLTTS_1.0 satisfies our needs.

Two more concepts need to be introduced to adequately capture the information on surficial geology map legends: periglacial and anthropogenic. We use Lundqvist's (1988) definition of the term periglacial. It encompasses processes, depositional environments, deposits or landforms "*caused by cold climate and thus often occurring outside a glacier but not requiring its presence*".

From the Glossary of Geology (Jackson, 1997), periglacial is more narrowly defined as "*said of the processes, conditions, areas, climates, and topographic features at the immediate margins of former and existing glaciers and ice sheets, and influenced by the cold temperature of the ice*". This definition is extended to include "*an environment in which frost action is an important factor, or of phenomena induced by a periglacial climate beyond the periphery of the ice*" (Jackson, 1997). This extension of the term "periglacial" might include features that do not require perennially frozen ground, but that do require extreme cold temperatures.

We believe that the heading "cryogenic" (processes, depositional environments, deposits or landforms that are related to extreme cold conditions including but not restricted to perennially frozen grounds) should be the higher-level concept under which periglacial (requiring perennially frozen grounds) and extreme cold climate would reside. Whether this distinction is needed is open to discussion. This is why Figures 3 and 4 do not reflect this thinking. Similarly, how we choose the cut-off percentage for the volume of ground ice required before a process, depositional environment, deposit, or landform is classified under periglacial needs to be fine-tuned. We offer language to at least cover the terminology found in map legends. We will need to address this issue with expert references such as those available at the Frozen Ground Data center web site (http://nsidc.org/fgdc/glossary/description.html), and specifically Everdingen (1998).

Finally, a concept that also requires some thought is the action of man on its environment. Anthropogenic processes result in such deposits as fill or mine tailings, which are important mappable units, especially where they conceal the underlying geology.

CONCLUSIONS

Science language development is not a minor task. Those involved in this effort each have a different view of the hierarchy and organisation of the terminology depending on their professional background. A carefully considered philosophy brings together ideas and concepts into a largely accepted consensus from which the science language is built. This is by no means a static issue. Users will need to work with the proposed science language and offer suggestions to improve its usefulness. Discussions will result in modifications, minor and perhaps major, which will lead to a new release of the science language. In order to facilitate this process, the GSC has established a stewardship policy (http://cgkn.net/2002/working/surficial_e.htm;

	GLACIAL PROCESS		**ICE-CONTACT PROCESS**		**PERIGLACIAL PROCESS**	
	Subglacial	**Ice-marginal**	**Subglacial**	**Ice-marginal**	**Perennially frozen**	**Extreme cold temperatures**
Plain	Cover moraine, veneer			Outwash plain		
	Cover moraine, blanket			Kame terrace		
	Cover moraine, undulating			Apron		
Ridge (stream-lined drift)	Flute		Flute			
	Drumlinoid		Drumlinoid			
	Drumlin		Drumlin			
	Crag-and-tail		Streamlined drift			
	Till ramp					
Ridge (stream-lined bedrock)	Rock drumlin (whaleback)					
	Roche moutonnée					
	Rat-tails					
Ridges (drift, not stream-lined)	Ribbed (Rogen) moraine	Terminal moraine	Crevasse filling		Ice-wedge polygons	
	DeGeer moraines	Frontal moraine	Esker			
	Thrust moraine	Recessional moraine Lateral moraine Interlobate moraine				
Mound	Hummocky moraine		Kame	Outwash fan	Pingo	Solifluction lobe
				Subwash	Palsa	Sorted circles
Depression	Cirque			Channel		
	Glacial furrow			Kettle		
	U-shaped trough					
	Fjord					
	Glacial-lake basin					
	Striation					

Figure 4. Glacial, ice-contact, and periglacial landforms. All the words found in this table were added by the map legend analysis exercise.

see Appendix for the description of the CGKN working group on surficial materials).

The progress report that we present here shows the situation at the GSC regarding surficial geology. Future versions of this science language will most likely include more specialized terminology, and perhaps change the classification scheme somewhat in order to better comply to SLTTS_1.0. We welcome discussion and comments on this proposition. We also believe that we need to publicly discuss our thought processes and science language so that potential users can critically comment on it. We think the approach we have taken will satisfy users’ needs and allow the geologists to find the terminology they want to use. It will take some time before we are all comfortable with the science language and its use, but it is a powerful tool for earth scientists and for decision-making authorities, planners, and the general public.

ACKNOWLEDGMENTS

This work is being carried out by the authors under, and financially supported by, project GK4350 "Integrated Information System for Bedrock, Surficial, Geochronological, Stratigraphic and Paleontological Data" of the program "Consolidating Canada's Geoscience Knowledge" of the Earth Science Sector, Natural Resources Canada. We wish to acknowledge the numerous discussions we have had with our colleagues, as well as Dr. Alain Plouffe for critically reviewing this manuscript. The first author has nagged repeatedly some specific individuals, Éric Boisvert of GSC-Québec, Peter Davenport of GSC-Calgary, and Jonathan Matti of the USGS. Thank you all!

REFERENCES

Davenport, P., Boisvert, É., Quat, M., Okulitch, A., Brodaric, B., Colman-Sadd, S., Noland, L., Struick, B., MacIntyre, D., Tzeng, P., Scott, D., Gilbert, C., Abbott, G., Bassan, A., Journeay, M., Francis, J., Houlahan, T., 2002, A Scalable, Digital Map Database of Bedrock Geology for Canada: A Progress Report, *in* Soller, D.R., ed., Digital Mapping Techniques '02—Workshop Proceedings, U.S. Geological Survey Open-file Report 02-370, p. 54-73, accessed at http://pubs.usgs.gov/of/2002/of02-370/.

Dreimanis, A., 1988, Tills: their genetic terminology and classification, *in* Goldthwait, R.P., and Matsch, C.L., eds., Genetic classification of glacigenic deposits (Final report of the Commission on Genesis and Lithology of Glacial Quaternary Deposits of the International Union for Quaternary Research [INQUA]): Rotterdam, A.A. Balkema, p. 17-83.

Everdingen, Robert van, ed., 1998 revised January 2002, Multi-language glossary of permafrost and related ground-ice terms: Boulder, CO, National Snow and Ice Data Center/World Data Center for Glaciology, accessed at http://nsidc.org/fgdc/glossary/description.html.

Goldthwait, R.P., and Matsch, C.L., eds., 1988, Genetic classification of glacigenic deposits (Final report of the Commission on Genesis and Lithology of Glacial Quaternary Deposits of the International Union for Quaternary Research [INQUA]): Rotterdam, A.A. Balkema, 294 p.

Goldthwait, R.P., Matsch, C.L., and Dreimanis, A., 1988, Preface, *in* Goldthwait, R.P., and Matsch, C.L., eds., Genetic classification of glacigenic deposits (Final report of the Commission on Genesis and Lithology of Glacial Quaternary Deposits of the International Union for Quaternary Research [INQUA]): Rotterdam, A.A. Balkema, p. vii-ix.

Jackson, J.A., 1997, Glossary of geology, 4th ed.: Alexandria, Virginia, American Geological Institute, 769 p.

Lundqvist, J., 1988, Glacigenetic processes, deposits, and landforms, *in* Goldthwait, R.P., and Matsch, C.L., eds., Genetic classification of glacigenic deposits (Final report of the Commission on Genesis and Lithology of Glacial Quaternary Deposits of the International Union for Quaternary Research [INQUA]): Rotterdam, A.A. Balkema, p. 3-16.

Moore, A., Dilabio, R., Thorleifson, H., Matile, G., Pyne, M., Murray, R. and Bolduc, A., 2003, Developing and deploying a surficial Geology Science Language: Geological Society of America Abstracts with Programs, No 6, September 2003, p. 278, accessed at http://gsa.confex.com/gsa/2003AM/finalprogram/abstract_66536.htm.

North American Geologic Map Data Model Science Language Technical Team, 2003, Science Language for Geologic-Map Databases in North America: A Progress Report, *in* Soller, D.R., ed., Digital Mapping Techniques '03—Workshop Proceedings: U.S. Geological Survey Open-file Report 03-471, p.109-138, accessed at http://pubs.usgs.gov/of/2003/of03-471/matti/index.html.

North American Geologic Map Data Model Steering Committee, 2004, NADM Conceptual Model 1.0, A Conceptual Model for Geologic Map Information: preliminary website release under the auspices of the Geological Survey of Canada, the U.S. Geological Survey, and the Association of American Geologists, accessed at http://nadm-geo.org/.

North American Geologic Map Data Model Steering Committee, unpublished document, Sedimentary materials: science language for their classification, description, and interpretation in digital geologic-map databases, Version 1.0.

Struik, L.C., Quat, M.B., Davenport, P.H., Okulitch, A.V., 2002, A preliminary scheme for multihierarchical rock classification for use with thematic computer-based query systems: Geological Survey of Canada Current Research, 9 p., accessed at http://geoscan.ess.nrcan.gc.ca/cgi-bin/starfinder/20479/geoscan_f.txt.

Thorleifson, L. H. and Pyne, D. M. 2003, Conversion of lithological data in the Manitoba water well database (GWDrill) to a mappable format, *in* Soller, D.R., ed., Digital Mapping Techniques '03—Workshop Proceedings: U.S. Geological Survey Open-file Report 03-471, p. 139-156, accessed at http://pubs.usgs.gov/of/2003/of03-471/thorleifson/index.html.

APPENDIX

Roles and responsibilities of the Canadian Geoscience Knowledge Network’s Working Group on Surficial Geology.

MANDATE

To provide, through the Internet, coherent visualizations of the nation’s surficial geology from the map collections of Canada’s geological survey agencies. This will be done from distributed databases maintained by individual agencies that are based on a common data model, including science language, which makes the information from mapping published at different scales and to different standards both interoperable and scalable.

PRINCIPLES

- Each geological map database will remain under the control of the source agency, which would be responsible for maintaining and updating it.
- The Subgroup will cooperate to ensure a stable and robust science language and glossary.
- The Subgroup will communicate regularly with geologists and information managers to ensure that this model addresses concerns and requirements.
- The data model will support both national «standards» for a common CGKN portal, and allow flexibility so that individual agencies can provide information to their own «standards» directly from their web portals.
- The data model should respect the well-established geological principles of geologic time, superposition, and correlation and conventions such as the North American Stratigraphic Code.
- The data model will accommodate both the original geological information from the source maps, and common science language and high-level classifications that would follow international standards and protocols where available.
- Source information, including original citations, must be explicitly linked to all information in the databases.

MEMBERSHIP

All Territorial, Provincial and Federal agencies who are actively involved in the mapping of surficial geology will be represented on the subgroup to guide the development of the data model, its science language, and to populate the geological map databases. Members must be able to effectively communicate the decisions and concerns of the Subgroup to their own organizations and to represent fairly to the Subgroup the views of their organizations. Each Agency will:

- Nominate a member for a period of one year and this individual will:
  a) Have a background in the Earth Sciences
  b) Have published at least one surficial geology map for the agency they represent
  c) Participate in technical teams when needed
  d) Be available to co-chair the subgroup if elected for this role

- Provide their representative with sufficient time and resources to participate on htis subgroup
- Recognize the contribution of the representative in their annual performance appraisal so that participation in this subgroup will not be detrimental to career progression
- Make available, where possible, additional expertise to technical teams struck by this subgroup

CHARGE AND ROLES

1. Overall guidance. The Subgroup specifies the scope of activities for development and implementation of the data model. The Subgroup provides authoritative statements of the model’s purpose, its intended use and users, and its relationships with other specifications such as the Canadian Geospatial Data Infrastructure (CGDI), the Open GIS Consortium (OGC), the USGS National Geologic Map Database (NGMDB), and the North American Data Model (NADM) Steering Committee.
2. Coordination of technical teams. The Subgroup identifies functional goals for the data model in sufficient detail that a technical team working towards each goal can accomplish clearly specified tasks within one year. For each

functional goal, the Subgroup will identify the technical need, state the immediate goals of the team, identify people who can work on the team, and facilitate, evaluate, and disseminate the work of the team. Examples of technical teams may include: conceptual data model design; scientific terminology; software tool development; and policy evolution.

3. Publicity and organizational liaison. The Subgroup works to publicize the model and its supporting products in appropriate conventions, meetings, workshops, and publications. The Subgroup provides information on its progress to related groups such as the Bedrock Subgroup (and others through the Data Infrastructure Working Group meetings), the NADM Steering Committee, and the NADM Surficial Science Language Technical Team.
4. Communication and Support. The Steering Committee facilitates public discussion and individual guidance regarding both broad issues and technical details of concern in the data model. These discussions and guidance are supported by technical information exchange at the CGKN web site. Members will actively solicit comments and guidance from the technical experts whose interests they represent to the Committee, and will respond regularly to that constituency.

STEWARDSHIP

Two co-leaders, one representing Federal agencies and the other the Provincial and Territorial agencies will lead the Subgroup. GSC management and the Committee of Provincial Geologists will appoint the co-leaders, respectively. These leaders will act as the Stewards for this data model and will:

- Lead Subgroup meetings
- Work with Subgroup members to develop meeting agendas
- Compile minutes for each Subgroup meeting
- Oversee the work of contracts let by the Subgroup
- Coordinate technical team activities
- Serve as the official point of contact between the Subgroup and the public

FUNDING

Each agency will be responsible for maintaining its own database, including the provision of the necessary computer systems to support it and personnel to populate it. Funding will be sought from national programs such as GeoConnections for support in developing common software and documentation, and to support travel to regular subgroup meetings (two per year). Travel support will also be required so that subgroup members can actively participate in international activities related to data model development, especially science language.

PROJECTS

The subgroup will establish annual and longer-term (3 year) plans to estimate funding requirements from member agencies and from national programs.

COORDINATION

The subgroup will meet via conference call regularly (6-10 times per year), and face-to-face twice annually. The subgroup will submit its annual plan to the CGKN secretariat for approval, and will prepare a written report to the secretariat annually. At least one member of the Data Infrastructure Coordination subgroup will be a member of the surficial subgroup.

TECHNICAL TEAMS

Technical teams are formed by the Subgroup to develop: detailed functional specifications of the data model; software tools and test-data meeting the specifications; and standard scientific terminology. The work of the technical teams will be defined specifically enough that the tasks can be accomplished within one year. This requirement is meant to allow the membership of technical teams to vary from year to year as needed by the teams and by the organizations that employ the team members. Technical team members are specialists in scientific or technological disciplines, and usually the work of the team falls within the member’s professional responsibilities.

Report on Progress to develop a North American Science-Language Standard for Digital Geologic-Map Databases

North American Geologic Map Data Model Steering Committee
Science Language Technical Team
http://nadm-geo.org/sltt/

INTRODUCTION

With the increasingly widespread production and use of digital geologic-map databases it has become clear that, to more effectively serve their constituencies, geoscience agencies need to develop several vital pieces of digital infrastructure:

1. A standard conceptual model for describing geologic phenomena, and for manipulating related data in a database environment,
2. Standardized science language that allows geologic materials and geologic structures to be described, classified, and interpreted,
3. Software tools for entering data into a standardized database and for retrieving the information according to user's needs, and
4. Methodologies and techniques for exchanging data sets having different structures and formats.

A single uniform language to classify and describe earth materials and their genesis is especially needed because users of geoscience information apply names, terms, and icons to communicate information about geologic objects and concepts. To the extent possible in a world where words are used diversely and inconsistently, standardized terminology is useful to facilitate information exchange among these users.

To address development of the infrastructure noted above, public-sector geologic-mapping entities in the United States and Canada formed a partnership called the North American Geologic Map Data Model Steering Committee (NADMSC, http://nadm-geo.org). NADMSC is sponsored by cooperative agreements between the U.S. Geological Survey (USGS) and the Association of American State Geologists (AASG), and between USGS and the Geological Survey of Canada (GSC). In the United States, NADMSC is linked to the database and standards development activities of the USGS National Geologic Map Database (http://ngmdb.usgs.gov/); in Canada, NADMSC is linked to database-development activities under the auspices of the Canadian Geoscience Knowledge Network.

The NADMSC chartered technical teams to develop resources and prototype standards for geologic map databases. These include: 1) the Data Model Design Team, which recently published the design for a conceptual data model (NADM DMDT, 2004); 2) the Data Interchange Technical Team, which provides in these Proceedings a report of progress (Boisvert and others); and 3) the Science Language Technical Team (SLTT), whose work is the subject of this report.

Between April 2000 and November 2004, the SLTT developed a prototype science language for the naming and describing of earth materials in geologic map databases produced by public-sector entities in North America. When the SLTT began its work, the intention was to produce a draft standard that could be evaluated, revised, and adopted by agencies and geologists working in North America. By the end of this process it became clear that, although this goal might be ultimately attainable, over the short term the SLTT's resources and lack of administrative authority prevent it from facilitating and executing the ambitious scope it originally envisioned.

Some committee members have proposed that the SLTT documents be published in a peer-reviewed venue. This is a logical suggestion, considering that the science-language reports are a comprehensive resource. However, this would require that the documents undergo an extensive review, and the NADMSC neither had the resources to conduct such a process nor does it have the formal mandate or permanent mechanism for archiving its documents. Therefore, in December, 2004, the SLTT posted the prototype science language as a set of "working documents" (see http://nadm-geo-org). Geologists and agencies are encouraged to evaluate and use the documents, to modify them as necessary for their purposes, and to offer recommendations for their modification.

In lieu of formal publication, the working documents are included in this open-file report in order to allow them to be permanently archived. The body of this report is an abbreviated summary of the SLTT's results. The Team's administrative procedures and the general nature of the science language classification were documented in a progress report at DMT'03 (NADM SLTT, 2003), and so are not repeated here. The appendices include the working documents and the executive summary from which this report was adapted. Because of their significant length, the appendices are available only in the web version of this open-file report.

The SLTT

The SLTT was formed in 2000, to identify and/or develop science language that allows information about geologic materials and geologic structures to be described in a standard way, and to promote wider use and more efficient exchange of geologic information. SLTT members were identified in the following ways:

1. Most participants from the U.S. Geological Survey were identified by Regional Geologic Executives from the USGS Western, Central, and Eastern Regions. This group includes representatives of the geologic-map editorial standards units of the regional publications groups. Additionally, some USGS scientists were appointed by Coordinators of USGS line-item science programs,
2. Scientists from the Geological Survey of Canada were identified by Canadian members of the NADMSC,
3. Scientists from State geological surveys were identified by the Digital Geologic Mapping Committee of the Association of American State Geologists (AASG),
4. Scientists from the U.S. Forest Service, National Park Service, U.S. Bureau of Land Management, and Natural Resources Conservation Service were selected by the committee chair, and
5. Scientists from academic institutions were selected by SLTT subcommittee co-chairs.

The assembled group (Table 1) represents a cross section of public-sector geologic map producers and users in the United States and Canada.

Table 1. NADMSC Science Language Technical Team committee members (Jonathan C. Matti, Chair)

Participant	Affiliation	SLTT Role
Lee Allison	Kansas Geological Survey	General scientific review
Brian Berdusco	Ontario Geological Survey	General scientific review
Richard C. Berg	Illinois State Geological Survey	Sedimentary subgroup
Thomas Berg	Ohio Geological Survey	General scientific review
Sam Boggs, Jr.	University of Oregon	Sedimentary subgroup
Eric Boisvert	Geological Survey of Canada	Sedimentary subgroup
Andrée Bolduc	Geological Survey of Canada	Sedimentary subgroup (co-chair)
Mark W. Bultman	U.S. Geological Survey	Sedimentary subgroup
William F. Cannon	U.S. Geological Survey	Metamorphic subgroup
Robert L. Christiansen	U.S. Geological Survey	Volcanic subgroup (co-chair)
Jane Ciener	U.S. Geological Survey	Geologic-map editorial standards
Stephen P. Colman-Sadd	Geological Survey of Newfoundland and Labrador	Metamorphic subgroup
Peter Davenport	Geological Survey of Canada	General scientific review
Ron DiLabio	Geological Survey of Canada	Sedimentary subgroup (co-chair)
Lucy E. Edwards	U.S. Geological Survey	Sedimentary subgroup
Robert Fakundiny	New York State Geological Survey	General scientific review
Kathleen Farrell	North Carolina Geological Survey	Sedimentary subgroup
Claudia Faunt	U.S. Geological Survey	Volcanic and sedimentary subgroups
Mimi R. Garstang	Missouri Department of Natural Resources	Sedimentary subgroup
Joe Gregson	National Park Service	General scientific review
Ardith K. Hansel	Illinois State Geological Survey	Sedimentary subgroup
Thomas D. Hoisch	Northern Arizona University	Metamorphic subgroup
J. Wright Horton, Jr.	U.S. Geological Survey	Metamorphic subgroup (co-chair)
David W. Houseknecht	U.S. Geological Survey	Sedimentary subgroup
Bruce R. Johnson	U.S. Geological Survey	Volcanic and metamorphic subgroups
Robert Jordan	Delaware Geological Survey	General scientific review
Ronald Kistler	U.S. Geological Survey	Plutonic subgroup (co-chair)
Alison Klingbyle	Geological Survey of Canada	Geologic-map editorial standards
Dennis R. Kolata	Illinois Geological Survey	Sedimentary subgroup
Elizabeth D. Koozmin	U.S. Geological Survey	Geologic-map editorial standards
Hannan LaGarry	Natural Resources Conservation Service	Sedimentary subgroup
Diane E. Lane	U.S. Geological Survey	Geologic-map editorial standards
Victoria E. Langenheim	U.S. Geological Survey	Plutonic and Sedimentary subgroups

Reed Lewis	Idaho Geological Survey	Plutonic and Volcanic subgroup
Stephen D. Ludington	U.S. Geological Survey	Volcanic subgroup (co-chair)
Jonathan C. Matti	U.S. Geological Survey	Sedimentary subgroup (co-chair)
James McDonald	Ohio Geological Survey	Sedimentary subgroup
David M. Miller	U.S. Geological Survey	Sedimentary subgroup (co-chair)
Andy Moore	Geological Survey of Canada	Sedimentary subgroup
Douglas M. Morton	U.S. Geological Survey	Plutonic subgroup
Patrick Mulvany	Missouri Department of Natural Resources	General scientific review
Carolyn Olson	Natural Resources Conservation Service	Sedimentary subgroup (co-chair)
Anne Poole	National Park Service	Plutonic and sedimentary subgroups
Stephen M. Richard	Arizona Geological Survey	Metamorphic subgroup (co-chair)
Andrew H. Rorick	U.S. Forest Service	Sedimentary subgroup
William Shilts	Illinois State Geological Survey	General scientific review
David R. Soller	U.S. Geological Survey	Sedimentary subgroup (co-chair)
Roy Sonenshein	U.S. Geological Survey	Sedimentary subgroup
William Steinkampf	U.S. Geological Survey	Volcanic and sedimentary subgroups
Douglas Stoeser	U.S. Geological Survey	Plutonic subgroup
Lambertus C. Struik	Geological Survey of Canada	General scientific review
John F. Sutter	U.S. Geological Survey	General scientific review
Harvey Thorsteinson	Minnesota State Geological Survey	Sedimentary subgroup
Robert J. Tracy	Virginia Polytechnic Institute and State University	Metamorphic subgroup
David Wagner	California Geological Survey	Volcanic subgroup
Richard Waitt	U.S. Geological Survey	Sedimentary subgroup
Peter D. Warwick	U.S. Geological Survey	Sedimentary subgroup
Richard Watson	U.S. Bureau of Land Management	General scientific review
Gerald A. Weisenfluh	Kentucky Geological Survey	Sedimentary subgroup (co-chair)
Carl Wentworth	U.S. Geological Survey	Sedimentary subgroup
Michael L. Williams	University of Massachusetts	Metamorphic subgroup
Ric Wilson	U.S. Geological Survey	Volcanic and plutonic subgroups
Robert P. Wintsch	University of Indiana	Metamorphic subgroup
Michael L. Zientek	U.S. Geological Survey	Plutonic and metamorphic subgroups

Rationale for standard science language

Standardized science language is needed to increase the usability and comparability of information contained in geologic map databases. A map user might conclude that terms occurring in map-unit explanations and in database fields have identical meanings from map to map and from region to region. This certainly is true for some specialized terms, and especially for more generalized terms. However, for some terms used in geologic maps, subtle to significant differences in geologic meaning can occur from map to map. This happens for various reasons:

1. The field description and interpretation of earth materials and geologic structures is as much an art as a science, and is predicated on the experience, training, intuition, skill, and persistence of the geologic-map maker. Moreover, each field area presents unique challenges to the geologic-mapping process (outcrop quality, climatic setting, accessibility, etc.). These realities open the door to differences in science language usage from map to map.
2. The meaning of some terms changes subtly to significantly from generation to generation as academic traditions change, and as new analytical techniques and geologic perspective influence and modify research results and teaching curriculums. New and different science language commonly emerges from these activities.
3. Some geologic terms once in vogue may completely disappear from the geologic lexicon as they are replaced with terms that are more accurate or precise or that better reflect current usage.
4. Some geologic terms take on meanings and applications specific to a particular geologic terrain or region; beyond that region, these terms may have a slightly different meaning, or may not even be used.
5. In a climate of open and competitive academic research, scientists constantly are experimenting with new, more creative, and more effective terminology to communicate information about earth materials that have complex combinations of composition, structure, fabric, and genesis.

For these reasons, the vocabulary (science language) of both historic and current geologic maps can vary—in some instances enough to create uncertainty on the part of the map user as to whether earth materials and geologic structures in one map are similar to or different from those in another. To minimize this problem, standardized science language that classifies and describes earth materials and their genesis is helpful, especially to facilitate information exchange.

Purpose and Intended Use for the SLTT Prototype Standard

The SLTT prototype standard provides a logical, consistent, hierarchical framework for naming and classifying earth materials, and for describing their physical characteristics and genesis—based on the way geologic maps are made by the field geologist or assembled by a science compiler. It is intended for use by persons and agencies submitting digital geologic-map data into public-domain databases that are managed by various State/Provincial and Federal agencies. Intended users include:

- geologists who collect original data in the field while making a geologic map,
- geologists who compile geologic-map data from legacy sources and must interpret and translate these data for representation in the compilation, and
- information-users who query public-domain geologic-map databases for information appropriate to their interests and applications.

It has been the SLTT's intention to break down common terms for earth materials into their fundamental science concepts. This is based on our belief that it is not so much what an object or concept is called, but what the name means in terms of the science concepts it represents. The SLTT documents provide specific defined names for earth-material objects and concepts, with the hope that they will be familiar and palatable to the average geologic-map maker and map user. However, we understand that each map producer and map user will have their own favorite names, and that humans are reluctant to abandon terms and meanings with which they are comfortable. With that recognition, we believe SLTT will have served its purpose if it provides a yardstick against which terms can be compared and translated—the true meaning of a "standard".

Related science-language efforts

SLTT deliberations benefited from previous and ongoing science-language efforts being conducted by other entities.

British Geological Survey

In 1999 the British Geological Survey (BGS) issued four reports that presented science language for earth materials from a geologic-mapping point of view:

- science language for igneous materials (Gillespie and Styles, 1999)
- science language for metamorphic materials (Robertson, 1999)
- science language for sedimentary materials (Hallsworth and Knox, 1999)
- science language for surficial and man-made materials (McMillan and Powell, 1999).

The SLTT adopted major elements of the BGS approach, but found that in order to accommodate North American geologic-mapping traditions and approaches we had to develop slightly modified terminology and taxonomic hierarchies.

International Union of Geological Sciences (IUGS)

SLTT activities benefitted from a series of IUGS subcommissions chartered to develop uniform classifications of earth materials:

- *Igneous materials:* A long-standing IUGS Subcommission on the classification of plutonic and volcanic igneous rocks (http://www.minpet.uni-freiburg.de/IUGS-CSP.html) has led to a widely accepted standard (IUGS, 1973; MacDonald, 1974; Streckeisen, 1974, 1976, 1978, 1979; Schmid, 1981; Heiken and Wohletz, 1985; Foley and others, 1987; Le Bas and others, 1986; Le Maitre and others, 1989; Le Bas and Streckeisen, 1991; Le Maitre and others, 2002).
- *Metamorphic materials:* An IUGS Subcommission on the classification of metamorphic rocks (http://www.bgs.ac.uk/SCMR/scmr_products.html) is underway, and is stimulating wide-ranging discussion of terminology for the naming, description, and genesis of metamorphic rocks.
- *Sedimentary materials:* An IUGS Subcommission on the classification of sedimentary materials (http://www.iugs.org/iugs/science/sci-cgsg.htm) is in the initial phases of its activities.

Science language for glacial sedimentary materials

The International Union for Quaternary Research [INQUA] in the 1970's sponsored a Commission on Genesis and Lithology of Glacial Quaternary Deposits

(Commission C-2). The results of Commission C-2 were published in Goldthwait and Matsch (1988; see Commission summaries in Goldthwait and others, 1988, p. vii-ix, and Dreimanis, 1988, p. 19-25). The SLTT used this document to develop science language for sedimentary materials of glacial origin.

Geological Survey of Canada science language

Concurrent with SLTT activities, the Geological Survey of Canada (GSC) is developing science language for use by GSC projects producing digital geologic-map databases. Through a series of projects, GSC has investigated approaches to developing geological map databases, including prototype data models and user interfaces. Bedrock and surficial geological maps have to date been addressed separately. As part of data modeling, based on variants of NADM, several approaches have been tested to enable interoperability among maps that use varied, usually undefined and sometimes inconsistent science language, particularly for the earth-material constituents of map units.

Two main approaches have been tried, both relying on map context and geological experience as guides to the authors' meaning. For surficial geological maps, the uncontrolled and variable terminology is reinterpreted within a controlled set of defined terms (a translation, in effect). For bedrock maps, earth material names are "reverse-engineered" into the properties (genetic process, composition, texture, etc.) implied by each name (single word or phrase), using sets of keywords for these properties (Davenport and others, 2002). In both approaches, a hierarchical organization of terms is applied wherever possible to allow for categorization at variable levels of precision in accordance with the information available, and to enable efficient querying of the databases.

For bedrock maps, Struik and others (2002) followed a different approach, recognizing that earth material names are multi-dimensional and can be organized in a variety of hierarchies depending on the choice of criteria (genetic process, composition, texture, etc.). The earth material names that Struik and others (2002) considered were uncontrolled terms gleaned from several published geological maps, but were neither exhaustive nor representative of the entire collection of published maps for Canada. This approach has been extended to collect earth material names in a master list as additional maps are brought into the database, and associate controlled keywords for earth material properties to each unique term (single word or phrase) through a data model that supports multiple ontologies. This enables map units to be searched or grouped by one or several of these keywords. User interfaces have been written to streamline the analysis of map unit descriptions, extraction of earth material types, and the assignment of keywords.

Federal Geographic Data Committee (FGDC) science language

Within the United States, an important science-language activity is occurring under auspices of the Federal Geographic Data Committee (FGDC) Geologic Data Subcommittee (http://ncgmp.usgs.gov/fgdc_gds/). The FGDC has developed a draft cartographic standard for polygon, line, and point symbols that depict geologic features on geologic maps and digital displays. Although primarily concerned with cartographic technical specifications, the FGDC cartographic standard contains science-language concepts that should be integrated with the science-language in these SLTT documents.

THE SLTT WORKING DOCUMENTS

Working documents versus a "standard"

As originally envisioned by the NADMSC and by the SLTT charter (see http://nadm-geo.org), our intent was to develop formal science-language standards for evaluation and use by the North American geologic-mapping community. Based on the charter and early discussions among SLTT members, it seemed logical to pursue the following strategy:

- develop formal science-language standards for the major classes of earth materials (metamorphic, plutonic, sedimentary, and volcanic). Do this by creating a set of SLTT subgroups, one for each earth material class,
- submit the standards for peer review and for simultaneous release as official publications of the U.S. Geological Survey and the Geological Survey of Canada,
- upon publication of the formal standard, obtain peer review and feedback from the North American geologic-mapping community,
- use this feedback to revise and refine the standard through a stewardship process maintained by the NADMSC and its SLTT, and
- on an as-needed basis, archive and distribute subsequent versions of the science-language standard.

This strategy proved unsupportable for the following reasons:

- Differences in philosophy among the various SLTT participants led to science-language approaches that differ from subgroup to subgroup, with the result that the SLTT documents do not have commonality of purpose, content, and scope,
- Participation from a broad cross-section of U.S. and Canadian agencies proved elusive, and the SLTT chair became concerned that the SLTT

documents would not be perceived as a truly North American science-language standard,
- SLTT subgroup leaders concluded that technical peer-review prior to USGS and GSC publication would lead to significant editorial revision and response by the SLTT subgroups—each of which already was overwhelmed by the weight of its SLTT responsibilities. Moreover, the SLTT documents would have been completely out of context for the average peer reviewer not already involved in the science-language process or its philosophical and operational complexities; hence, the agency peer-review process would have been lengthy and difficult to execute and would have been of uncertain benefit,
- The SLTT charter did not anticipate or identify science-language stewardship as a mandated function, nor did it include mechanisms for responding to community feedback or for preparing and releasing revised versions of the science-language documents,
- The NADM SLTT process, although sanctioned generically by various memoranda of understanding between the USGS, GSC, and AASG (but not the Canadian Provincial geological surveys), has no formal mechanism for communicating science-language issues and results to their respective agencies and downward to their geologic-mapping projects (for evaluation and testing). Until such mechanisms are defined and tested, it is premature to consider standardization, stewardship, and versioning, and
- In short, the SLTT process does not have the mandate or the personnel to execute a formal science-language process on behalf of the various North American federal, state, and provincial agencies that conduct geologic mapping.

For these reasons, the NADMSC accepted the SLTT chair's recommendation that formal publication of the science-language document be reconsidered. NADMSC agreed that the best approach was to post the various SLTT reports on the NADM website, and to present them as a work in progress (i.e., as "working documents"). The inclusion of the working documents as appendices to this report serves to fulfill a principal NADMSC objective—to publish and archive these documents as a permanent record of the SLTT's endeavor.

This strategy allows the SLTT to conclude its responsibilities, and to present the North American geologic-mapping community with a range of science-language approaches and issues for their evaluation and discussion—pursuant to any next steps in the science-language process that are determined necessary by the NADMSC or by any geological survey.

A philosophical issue

Early in the SLTT process, tensions developed between two very different science-language goals and strategies:

1. Classifying the terminology of geologic maps so that each term commonly used in map legends and map-marginal explanations can be found in science-language classification schema. This objective focuses on *legacy* geologic-map information and on science language that enables the compilation of such information, without having to determine how the author of the map used the geologic terminology. By this rationale, science-language deliberations should determine how to organize *existing* earth-material names, based on the premise that the names are the principal basis for conveying science content.
2. Creating science-language schema that allow the map author or map compiler to represent what actually is known about the earth materials portrayed on a geologic map. This objective focuses on the geologic-mapping process itself—that is, on the way geologists use terms to express what they see in outcrops and in hand specimens, how they make mapping decisions in the field, how they organize and present their map data to express confidence in their observations and interpretations, and how the scientific content of current and future geologic-map databases can be improved and clarified. By this rationale, science-language deliberations should provide the map maker or map compiler with (1) very specific names that can be used where field data warrant or where legacy map terminology is clear, or (2) higher-level general names that can be used where field data are ambiguous or where the use of legacy map terminology is not clear. This rationale is driven by the premise that the scientific content, not just the names, is what geologic-map users are looking for.

These two objectives are equally legitimate. However, they reflect different philosophies and lead to different science-language strategies. Tensions between them were not resolved during the course of SLTT deliberations and, as a consequence, significant differences in scope, content, purpose, and philosophy exist among the various SLTT reports. This did not invalidate the SLTT effort, but it does illustrate the complexity and challenges of developing a standard science-language. Moreover, it should be a valuable lesson for agencies that conduct geologic mapping and that intend to develop local, regional, and national geologic-map databases that have uniform science content.

The composite-genesis and sedimentary subgroups concluded that their principal objective was to examine the science concepts embedded in geologic-map terminology, and to develop classification schema organized around that conceptual content. This philosophical approach forced a re-examination of how traditional map terms are used, and in some instances led these subgroups either not to adopt as controlled terms some familiar earth-material names, or to position these names in classification hierarchies in a different place than where some workers might expect to find them. For future geologic-mapping activities and their resulting databases, this probably will not create any long-term problems—provided future geologic mappers understand and agree with SLTT approaches. For legacy geologic-map information, the approach adopted by the composite-genesis and sedimentary groups might require some decision making on the part of the information compiler: (1) for a legacy term whose original meaning was not clear, the map compiler might have to use a higher-level, more generalized SLTT term instead, or (2) where a legacy term is understood to have a different meaning than the SLTT rendering of the same term, the map compiler may have to use a different SLTT term for the same concept.

General classification

The classification adopted by the SLTT follows this high-level architecture for earth-material name (see also Appendix A):

- Earth Material
  - Igneous earth material
    - Volcanic rock
      - lithologic class based on composition
      - lithologic class based on texture
      - lithologic class based on emplacement characteristics
    - Hypabyssal rock (BGS classification, Gillespie and Styles, 1999)
    - Plutonic rock (BGS classification, Gillespie and Styles, 1999)
  - Sedimentary earth material (unconsolidated, consolidated)
    - Sedimentary material, unclassified
    - Terrigenous-clastic sedimentary material
    - Carbonate sedimentary material
    - Organic-rich sedimentary material
    - Non-clastic siliceous sedimentary material
    - Noncarbonate-salt sedimentary material
    - Phosphate-rich sedimentary material
    - Iron-rich sedimentary material
  - Composite-genesis earth material
    - Cataclastic rock
    - Impact-metamorphic material
    - Metamorphic rock (traditional sense) (including hydrothermally-altered rock)
      - granoblastic rock
      - foliated metamorphic rock

These high-level categories fundamentally are genetic: they reflect how earth material was formed (genetic process, crustal depth, etc.). This raises the irony that, while deeper levels of the earth-material classification hierarchy are based on what the mapping geologist can see in the outcrop (empirical factors such as composition, structure, and texture), upper-level categories are based on interpretations about how the material was formed. Once this choice is made, an earth material is classified in more detail based on textural or compositional criteria—criteria that actually can be satisfied on the basis of empirical observation.

The use of standardized science language in digital geologic-map databases is a new frontier that is likely to evolve with time and experience. With this in mind, we developed classifications of earth materials that we believe reflect not only how mapping geologists view them but also how such materials might be queried and analyzed in geologic-map databases. No single classification of earth materials will please all workers. However, the schemes we propose hopefully will be clearly understandable, internally consistent, and usable by both data-producer and data-user.

Detailed Classification

Volcanic SLTT

The volcanic SLTT document (see Appendix D) provides a concise look at the science language of unconsolidated and consolidated volcanogenic earth materials. The goal of the volcanic subgroup was:

"...to develop standardized nomenclature for use in digital geologic map databases, specifically to describe lithologies in volcanic rock units. Although this nomenclature takes the form of a hierarchy of terms, it is important to note that this is not the same as a formal rock-naming system....

We consider it critical to remember that the purpose of our hierarchical subdivision of terms is to describe the *lithologic characteristics* of *geologic map units*. [Our hierarchical subdivision] is to be used to logically retrieve or select those map units that contain a specified set of lithologic characteristics. Thus, it must be flexible enough to accommodate the extremely varied and unsystematic way in which map units are described and defined by various authors. This report groups lithologic features necessary to adequately characterize **volcanic materials** in the map units of a geologic map database into three fundamental classes based on **composition**, **texture**, and **emplacement characteristics**.

No one of these classes is primary, and any or all may be used to select the lithologies of map units. The subdivision of any one of the fundamental classes consists of a list of words, arranged in a hierarchy that can be used to select lithologies. The words that describe these subdivisions are not given formal definitions here, but brief descriptions are given in the appendices. Many of the words have multiple, sometimes conflicting definitions and have been used differently over the years by different map authors. We have attempted to make the hierarchy sufficiently comprehensive, especially at the higher levels, to allow adequate lithologic characterization and to accommodate the vast majority of lithologic descriptions on existing geologic map legends."

The volcanic SLTT subgroup focused on how to bring the variable and inconsistent usage of legacy geologic maps into a modern database. To accomplish this, they characterize volcanic materials using three fundamental classes: composition, texture, and emplacement characteristics. Their report provides informal characterizations of volcanogenic materials in terms of these three aspects, but does not provide formal material descriptions, deferring instead to other sources (such as Le Maitre and others, 2002). The report does not provide a comprehensive listing of petrologic descriptors, as the subgroup felt it was beyond their mandate.

Plutonic SLTT

Owing to conflicting agency science-project obligations, members of the plutonic SLTT subgroup were unable to conclude their deliberations and were unable to develop plutonic science-language standards for use by geologic-mapping projects in North America. In the interim, the NADMSC recommends that the British Geological Survey report on plutonic science language (Gillespie and Styles, 1999) be used for North American geologic-map databases.

Sedimentary SLTT

The sedimentary subgroup produced a comprehensive analysis of the attributes for sedimentary earth materials, both consolidated and unconsolidated (see Appendix C), that includes the following components:

- attempts to identify from a database point of view the essential science concepts that underlie sedimentary terminology,
- science language for the various lithologic classes of sedimentary earth material,
- science language for the physical properties of sedimentary earth materials, including outcrop characteristics, consolidation state, sedimentary structures, sedimentary texture and fabric, particle composition, and material strength,
- science language for upper-surface attributes of sedimentary earth materials, including depositional and erosional landform features and surface-modification features (e.g., surface smoothing, surface dissection, surface armoring, particle weathering, pedogenic modification, cryogenic modification, and microrelief),
- science language for the genesis of sedimentary earth materials, including particle origin, depositional process, depositional place, geomorphic configuration, ambient conditions, and tectonic setting and basin type, and
- science language for human-affected landscapes, including made ground and worked ground.

Composite-genesis SLTT

Science language for metamorphic rocks and for other earth materials that form through modification of pre-existing earth material owing to the effects of temperature, pressure, and deformation, is discussed in the SLTT report on composite-genesis materials (see Appendix B). The domain of this classification system includes metamorphic rocks as commonly understood, as well as impact metamorphic rocks, hydrothermally altered rocks, mylonite-series rocks, and cataclastic rocks. These composite-genesis rocks are classified according to descriptive properties that are interpreted to reflect processes that made the rock composite.

The Composite-genesis subgroup members discussed whether or not to include within the composite-genesis domain earth material such as pedogenic soil that forms at the earth's surface through low temperature-pressure processes that modify pre-existing sediment and rock. No consensus was reached on this subject, hence pedogenic materials are not currently included in any of the SLTT science-language documents, except as a modifier to describe the upper surface of sedimentary earth materials.

Preliminary results of the SLTT process

The SLTT process was an experiment with mixed outcomes:

- We produced documents that can be evaluated for their contribution to the science content and increased uniformity of North American geologic-map databases.
- However, committee deliberations revealed significant differences in how various individuals, agencies, and scientific programs view geologic-map

databases and how they should be constructed to further their science missions.

The NADMSC believes the SLTT documents will be of significant value to the North American geologic-mapping community: hopefully, the effort will stimulate discussion about how the content of geologic-map databases is used, how it is accessed, and how it can be structured and represented through the use of standard science language. Such discussions could lead to future work that will build on SLTT accomplishments.

Finally, and because the SLTT process was conducted to support agency needs for standardized map databases, we offer the following recommendations to high-level science managers in agencies that execute geologic mapping:

1. understand and appreciate the fundamental importance and intellectual complexity of a geologic-map data-model standard and its scientific content,
2. require your agencies to develop such a standard, or to adapt and build on the SLTT standard,
3. encourage your scientific workforce to participate fully and legitimately in standards development, and to implement the standards once they are developed, and
4. mandate and empower a single entity within your agency to take the lead on standards development on behalf of all other producers and users of geologic-map information within your agency.

If these four requirements are not advocated and facilitated, then science-language standards will be neither robust nor comprehensive, and most likely they will not be viewed seriously by a workforce that may (or may not) be asked to adopt them.

CREDITS

The SLTT chair (Jon Matti) prepared the summary document (Appendix A) in coordination with the SLTT subgroup leaders (Table 2), each of whom contributed to the SLTT subgroup narratives in this report. Dave Soller assisted with preparation of this report by adapting it from the summary document.

REFERENCES CITED

Davenport, P., and others, 2002, A Scalable, Digital Map Database of Bedrock Geology for Canada: A Progress Report, *in* Soller, D.R., ed., Digital Mapping Techniques ‘02—Workshop Proceedings: U.S. Geological Survey Open-File Report 02-370, p. 47-66, accessed at http://pubs.usgs.gov/of/2002/of02-370/davenport.html.

Dreimanis, A., 1988, Tills: their genetic terminology and classification, in Goldthwait, R.P., and Matsch, C.L., eds., Genetic classification of glacigenic deposits (Final report of the Commission on Genesis and Lithology of Glacial Quaternary Deposits of the International Union for Quaternary Research [INQUA]): Rotterdam, A.A. Balkema, p. 17-83.

Foley, S.F., Venturelli, G., Green, D.H., and Toscanni, L., 1987, The ultra potassic rocks: characteristics, classification and constraints for petrographic models: Earth Science Reviews, v. 24, p. 81–134.

Gillespie, M.R., and Styles, M.T., 1999, Classification of igneous rocks: Volume 1 of the BGS rock classification scheme: British Geological Survey Research Report Number RR 99-06, 52 p.

Goldthwait, R.P., and Matsch, C.L., eds., 1988, Genetic classification of glacigenic deposits (Final report of the Commission on Genesis and Lithology of Glacial Quaternary Deposits of the International Union for Quaternary Research [INQUA]): Rotterdam, A.A. Balkema, 294 p.

Goldthwait, R.P., Matsch, and Dreimanis, A., 1988, Preface, in Goldthwait, R.P., and Matsch, C.L., eds., Genetic classification of glacigenic deposits (Final report of the Commission on Genesis and Lithology of Glacial Quaternary Deposits of the International Union for Quaternary Research [INQUA]): Rotterdam, A.A. Balkema, p. vii-ix.

Hallsworth, C.R., and Knox, R.W.O’B., 1999, Classification of sediments and sedimentary rocks: Volume 3 of the BGS rock classification scheme: British Geological Survey Research Report Number RR 99-03, 44 p.

Heiken, G., and Wohletz, K.H., 1985, Volcanic ash: Berkeley, University of California Press.

Table 2. SLTT Subgroup leaders who contributed to this report.

Robert L. Christiansen	U.S. Geological Survey	Volcanic subgroup
Andrée Bolduc	Geological Survey of Canada	Sedimentary subgroup
Ron DiLabio	Geological Survey of Canada	Sedimentary subgroup
J. Wright Horton, Jr.	U.S. Geological Survey	Metamorphic subgroup
Stephen D. Ludington	U.S. Geological Survey	Volcanic subgroup
Jonathan C. Matti	U.S. Geological Survey	Sedimentary subgroup
David M. Miller	U.S. Geological Survey	Sedimentary subgroup
Carolyn G. Olson	Natural Resources Conservation Service	Sedimentary subgroup
Stephen M. Richard	Arizona Geological Survey	Metamorphic subgroup
David R. Soller	U.S. Geological Survey	Sedimentary subgroup
Gerald A Weisenfluh	Kentucky Geological Survey	Sedimentary subgroup

IUGS (International Union of Geological Sciences Subcommission on the Systematics of Igneous Rocks), 1973, Plutonic rocks: Geotimes, v. 18, no. 10, p. 26-30.

Le Bas, M.J., and Streckeisen, A., 1991, The IUGS systematics of igneous rocks: Journal of the Geological Society of London, v. 148, p. 825–833.

Le Bas, M.J., Le Maitre, R.W., Streckeisen, A., and Zanettin, B., 1986, A chemical classification of volcanic rocks based on the total alkali-silica diagram: Journal of Petrology, v. 27, p. 745–750.

Le Maitre, R.W., and 11 others (editors), 1989, A classification of igneous rocks and glossary of terms: Recommendations of the International Union of Geological Sciences Subcommission on the Systematics of Igneous Rocks: Oxford, Blackwell Scientific Publications.

Le Maitre, R.W. (editor), Streckeisen, A., Zanettin, B., Le Bas, M.J., Bonin, B., Bateman, P., Bellieni, G., Dudek, A., Efremova, S., Keller, J., Lameyre, J., Sabine, P.A., Schmid, R., Sørensen, H., and Woolley, A.R., 2002, Igneous rocks: A classification and glossary of terms: Recommendations of the International Union of Geological Sciences Subcommission on the Systematics of Igneous Rocks: Cambridge, Cambridge University Press, 236 p.

MacDonald, R., 1974, Nomenclature and petrochemistry of the peralkaline oversaturated extrusive rocks: Bulletin Volcanologique, v. 38, p. 498–516.

McMillan, A.A., and Powell, J.H., 1999, Classification of artificial (man-made) ground and natural superficial deposits: applications to geological maps and datasets in the UK: v. 4 of the BGS rock classification scheme: British Geological Survey Research Report Number RR 99-04, 65 p.

NADM DMDT (North American Geologic Map Data Model Steering Committee Data Model Design Team, 2004, NADM Conceptual Model 1.0—A Conceptual Model for Geologic Map Information: U.S. Geological Survey Open-File Report 2004-1334, 58 p., accessed at http://pubs.usgs.gov/of/2004/1334/.

NADM SLTT (North American Geologic Map Data Model Steering Committee Science Language Technical Team), 2003, Science Language for Geologic-Map Databases in North America: A Progress Report, *in* Soller, D.R., ed., Digital Mapping Techniques '03—Workshop Proceedings: U.S. Geological Survey Open-File Report 03–471, p. 109-138, accessed at http://pubs.usgs.gov/of/2003/of03-471/matti/.

Robertson, S., 1999, Classification of metamorphic rocks: Volume 2 of the BGS rock classification scheme: British Geological Survey Research Report Number RR 99-02, 24 p.

Schmid, R., 1981, Descriptive nomenclature and classification of pyroclastic deposits and fragments: Recommendations of the IUGS Subcommission on the Systematics of Igneous Rocks: Geology, v. 9, p. 41–43.

Streckeisen, A., 1974, Classification and nomenclature of plutonic rocks: Recommendations of the IUGS Subcommission on the Systematics of Igneous Rocks: Geologische Rundschau Internationale Zeitschrift für Geologie: Stuttgart, v. 63, p. 773–785.

Streckeisen, A., 1976, To each plutonic rock its proper name: Earth Science Reviews, v. 12, p. 1–33.

Streckeisen, A., 1978, IUGS Subcommission on the Systematics of Igneous Rocks: Classification and nomenclature of volcanic rocks, lamprophyres, carbonatites and melilitic rocks. Recommendations and suggestions: Neues Jahrbuch für Mineralogie, v. 134, p. 1–14.

Streckeisen, A., 1979, Classification and nomenclature of volcanic rocks, lamprophyres, carbonatites and melilitic rocks: recommendations and suggestions of the IUGS Subcommission on the Systematics of Igneous Rocks: Geology, v. 7, p. 331–335.

Struik, L.C., Quat, M.B., Davenport, P.H., and Okulitch, A.V., 2002, Multi-hierarchical rock classification for use with thematic computer-based query systems: a preliminary scheme: Geological Survey of Canada, Current Research 2002D, 9 p.

APPENDICES

(NOTE: these Appendices are only available in the Web version of the document, at http://pubs.usgs.gov/of/2004/1451/.)

Appendix A. Philosophical and operational guidelines for developing a North American science-language standard for digital geologic-map databases.

Appendix B. Classification of metamorphic and other composite-genesis rocks, including hydrothermally altered, impact-metamorphic, mylonitic, and cataclastic rocks.

Appendix C. Sedimentary materials: Science language for their classification, description, and interpretation in digital geologic-map databases.

Appendix D. Volcanic materials: Science language for their naming and characterization in digital geologic-map databases.

GML Encoding of NADM C1

By Eric Boisvert[1*], Bruce R. Johnson[2*], Simon Cox[3], and Boyan Brodaric[4]

[1]Geological Survey of Canada
490 de La Couronne
Québec, G1K 9A9 Canada
e-mail: eboisver@nrcan.gc.ca

[2]U.S. Geological Survey, 954 National Center, Reston, VA 20192, U.S.A.
[3]CSIRO, ARRC, PO Box 1130, Bentley, WA 6102, Australia
[4]Geological Survey of Canada, 615 Booth Street, Ottawa, K1A 0E9, Canada

*Members of NADM DITT (Digital Interchange Technical Team)

ABSTRACT

The North American Geologic Map Data Model Steering Committee's (NADM) Digital Interchange Technical Team (see http://nadm-geo.org) is tasked to create an interchange format compliant with the North American Geologic Map Data Model conceptual model, known as "NADM C1". XML was unanimously selected as the technology of choice, and after initial attempts, it was realised that leveraging existing work on GML (Geographic Markup Language; OpenGis, 2004) would improve the interchange format. GML is a library that provides essential GIS features that can be reused in any geospatial application, such as NADM. GML provides reusable objects and design patterns. The NADM conceptual model has been analysed to create a GML application; this paper describes that process and provides examples of NADM GML encoding for interchange of geoscience data.

INTRODUCTION

In the latest report on this technical team (Digital Interchange Technical Team, 2003), we discussed the challenges of encoding the NADM conceptual model in an XML document. The process involved converting classes from the UML diagram into meaningful XML tags. We pointed out that the principal difficulty involved creation of a consistent logical schema (in UML), such that patterns in the schema could be mapped in a regular manner onto XML document structures, easing the transition from modeling to encoding. The solution to this problem came to us in the fall of 2003 at an international meeting in Edinburgh (Laxton & Brodaric, 2003), where it was decided that GML could be used to constrain and direct the XML encoding process, and thus that GML will be used as the encoding standard for sharing geological datasets among the participants (see http://ncgmp.usgs.gov/intdb/dmic/dmic-rep1.html). GML is itself an XML encoding of ISO standards designed to represent geographical features. Of importance to this discussion is the notion that GML provides a much needed design pattern which focuses encoding choices to a manageable subset of the very large set of choices possible in XML.

WHAT IS GML?

GML is the abbreviation for *Geographic Markup Language*. It is an XML encoding of an ISO *feature model* standard that describes classes required to represent common geospatial features (see Open GIS Consortium, 2004, for complete normative references). In the standard, there are descriptions for spatial objects (points, lines, polygons …), projections, dictionaries, topology, time, etc. The state of a geographic feature is described by properties, where each property is formed by a name, a type, and a value. GML is a framework, or a library, of reusable building blocks for any group that needs to create a "*GML application*". GML by itself is not intended to be used directly, because it is domain neutral; it only describes geography, and there are only a small number of concrete feature types. To turn GML into a useful application, it must be expanded to describe the content in a specific domain, such as biology, forestry or geology. XMML is one example of such an application which extends GML to address specific issues of mining and mineral exploration (https://www.seegrid.csiro.au/twiki/bin/view/Xmml/WebHome).

WHY GML?

Several benefits can be obtained by adapting GML rather than starting with a new XML encoding. First, GML is an international standard developed by the Open GIS Consortium (OGC) that is currently being revised for publication as ISO 19136, being derived from a related set of international standards (the ISO 19100 series). It provides a formal development framework and comes with a large group of practitioners to help guide and support development. It also ensures that encodings developed in other domains can be reused by sharing common GML constructs. Some examples of reuse will be discussed later.

GML is also at the root of other standard technologies and protocols fostered by OGC. One important technology is WFS (Web Feature Service), which provides a mechanism for interaction with a geospatial database using GML. Software developed according to this standard can be reused for any GML encoding. By connecting to the GML community, we benefit from software vendors who support WFS in their server and client products. One possible scenario is to integrate information from various sources for decision making. Figure 1 shows a typical architecture of servers (on the left) translating their content into GML and serving the content through WFS to a GML-enabled client. Because the schemas are GML-based, the application can handle the common parts (the geography) without any prior knowledge of the domain.

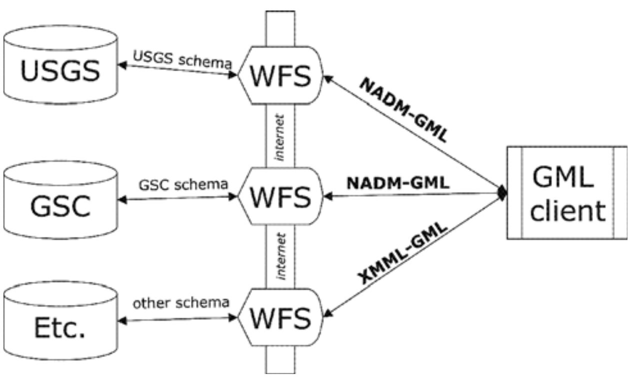


Figure 1. Example of a distributed system sharing geological information using GML and WFS (Web Feature Service).

GML also provides a 'Design pattern' that is one of the most significant additions to our design process. Although the literature does not 'sell' GML as such, the rules and constraints set by GML designers are a great help to any team working on XML schema design. GML in effect provides guidelines for consistent schema design. Without such rules and constraints XML (and UML) is almost too flexible, allowing many alternative ways to encode the same content, and requiring potentially different tools to be developed in order to read and manipulate the content. This implicit and understated aspect of GML proved to be of significant importance in our design process.

XML VS GML

It is important to realise that GML is XML. GML components are defined using XML schemas. An XML schema is a W3C standard that defines the structure of an XML document (see http://www.w3.org/XML/Schema). We extend GML by using an XML schema that allows us to re-use tools from our first attempt to encode the NADM conceptual model in pure XML. GML provides a conceptual model based on the ISO feature model. A GML application must reuse core GML features defined in the conceptual model. Figure 2 shows a UML representation of a very simple example of a *Street* feature that we might want to model. The street inherits from the abstract GML feature and reuses standard GML spatial geometries and property structures for the new street features. By inheriting the Street from a GML Feature, we turned our street into a formal GML feature (the diagram reads 'A Street is a kind of GML feature').

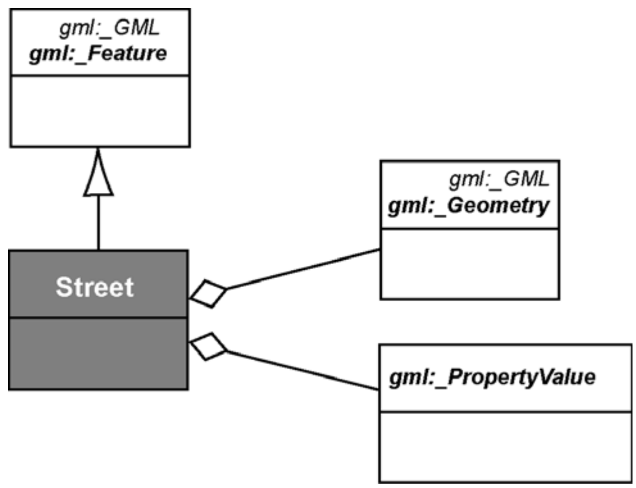


Figure 2. UML schema representing how GML components are reused in a specific application. In UML, a line with a triangle can be read as 'is a kind of' and the lines with diamonds are read as 'has'. This diagram is read as follows: 'A **street** is a kind of GML feature that has geometry and property'.

However, GML also introduces design patterns, conformant with the General Feature Model defined in ISO 19109:

- Classes are associated to other classes (or to simple data types such as strings or integers) using GML properties (i.e., the Class-property model). GML properties must have meaningful names and defined types. This rule forces relations to be qualified, documenting why or how classes are related.
- All class names must start with a capital letter (`'Street'`, and not `'street'`) and all property names must start with a lower case letter. More precisely, GML uses the 'CamelCase' structure where names can be formed from several words using capitalisation as a separator, for

example **'MainStreet'** as a class name and **'coveringMaterial'** as a property name.
- All classes must have IDs. In a complex document, the same feature can be referenced several times in the same document, or from outside the document. To ensure that every piece of information can be successfully reused, it must be uniquely identified. GML provides this functionality with the **gml:id** attribute.
- XML element attributes should be avoided. GML uses attributes in two very specific cases: 1) to assign ID and references to other features in the document or in another document; and 2) for limited metadata, in particular indicating the reference system for simple values (e.g. spatial coordinates, quantities, codes). This means that all first order information should be modelled as XML elements.

This snippet of GML code is an example of these encoding rules applied to the model depicted in Figure 3 (adapted from Galdos System Inc., 2003):

```
<app:RoadSegment gml:id="RS1"[1]>
   <app:name>Handbury Road North</
      app:name>
   <gml:centerLineOf[2]>
      <gml:LineString gml:id="L1">
         <gml:posList srsName="#
              localCRS2a">
         1,2 2,3 3,4 4,0
         </gml:posList>
      </gml:LineString>
   </gml:centerlineOf>
   <app:numberOfLanes>2</app:
      numberOfLanes>
   <app:surfaceType>Asphalt</app:
      surfaceType>
   </app:RoadSegment>
```

[Note the typical
```
   class (UpperCase "RoadSegment")
      property (lowercase "center
         LineOf")
         class (UpperCase "Line
              String")
```
structure].
[1]Required ID.
[2]Since RoadSegment inherits from Feature, it can have 0 or more geometries. One type of GML geometry is **centerLineOf**.

"app:" and **"gml:"** are namespace prefixes. The namespace mechanism provided by XML prevents 'name clashing' when several schemas are brought together in the same document. In this case, the designer chose to create a **<app:name>** tag to hold the name, but a **<gml:name>** also exists in GML. To differentiate between those two elements, a namespace mechanism is used.

Another way to encode the same piece of information (Galdos System Inc., 2003) is:

```
<app:RoadSegment gml:id="RS1">
   <app:name>Handbury Road North</
      app:name>
   <gml:centerLineOf href="#L1" />
      This points to a LineString
   <app:numberOfLanes>2</app:
      numberOfLanes>
   <app:surfaceType>Asphalt</app:
      surfaceType>
   <app:width uom="m">7.0</app:
      width>
</app:RoadSegment>…
…
…
<gml:LineString gml:id=»L1»>   This is
   the LineString, located further in document
   <gml:posList srsName="#localCR
      S2a">1,2 2,3 3,4 4,0</gml:
      posList>
</gml:LineString>
```

where "localCRS2a" is the value of an ID attribute on the definition of the coordinate-reference system.

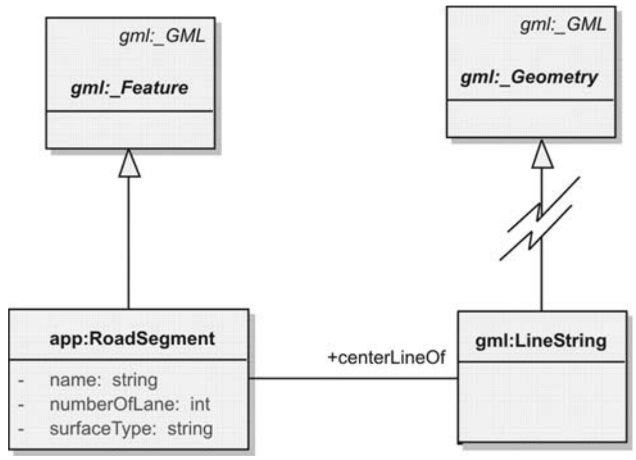


Figure 3. UML representation of a GML **RoadSegment**. RoadSegment inherits from the abstract **gml:_Feature** , so a **RoadSegment** is a GML Feature. The **RoadSegment** declares a series of simple properties (**numberOfLane**, etc..) and a more complex property named '**centerLineOf**' that relates to a complex geometric feature from gml called **LineString** (which is a special kind of **Curve**) that derives (through a series of intermediate geometry types) from an abstract **_Geometry** class.

ENCODING OF NADM C1

The NADM Data Model Design Team (DMDT) recently released version 1 of its conceptual model (C1), which is available at http://nadm-geo.org. The model is described in a document containing UML diagrams, descriptions of classes, and accompanying text. The conceptual model is 'implementation neutral', in that it does not describe how it can be implemented in any specific technology, be it XML or in a relational database. Conversion of the conceptual model to an XML implementation first required adaptation of the conceptual model into a logical model expressed in a "GML-friendly" profile of UML, in preparation for formal encoding. This required: a) converting NADM C1 into a feature model, and b) replacing some specific UML structures with structures that are directly compatible with GML/XML. The main objective was to retain the original meaning of NADM C1 and to only reinterpret it for conversion to the GML application.

The Class-property model used in GML encoding is represented in a UML logical model (Figure 4) where:

- GML Classes (including Feature types) correspond to UML Classes defined in NADM C1, in a derivation hierarchy; all NADM C1 classes derive ultimately from abstract **`_Feature`** and abstract ***`_GML`***. Note that the underline *(_)* preceding the name is a syntaxic convention to represent abstract classes.
- GML properties are modelled as UML attributes and associations (i.e., the lines between the classes), where the GML property name is the same as the UML attribute name or the role name associated with the target class from the original NADM C1 model.
- Thus, all associations must carry role names at the end corresponding to the child element. See association between **`_EarthMaterial`** and **`Fabric`** on Figure 6.
- Association classes from the original NADM C1 are replaced in GML by intermediate classes, with additional role names as required and fixed cardinalities (as demonstrated later on Figure 6).

Several difficulties in GML encoding were encountered, for example:

- We discovered possible improvements to NADM C1, such as modifying the relations between the **`ParticleSize`** and **`ParticleShape`** classes and the **`ParticleGeometry`** class. We then struggled to choose between the improved model and strict adherence to the published (NADM C1) model.
- Some UML constructs could not be ported to GML; for example, multiple inheritance (see **`Fossil`** class in NADM C1), although some techniques can be used in GML/XML to simulate multiple inheritance.

The resulting GML code or implementation, although depicted as a UML model is not a representation of the original conceptual model, but rather is a model of the GML application in which model elements are derived from the GML classes and some relation types are interpreted as XML structures, e.g., aggregation implies nesting tags into another tag in the XML document. This is a very important distinction; the GML-friendly UML diagram is not an amendment or change to the official NADM C1 conceptual model. It is a tool that is used to bridge between the conceptual model and the GML application.

Shown on Figure 5 are structural changes to the hierarchy of NADM C1. In GML, all classes must inherit from core GML classes (**`_Object`**, **`_GML`**, **`_Feature`**).

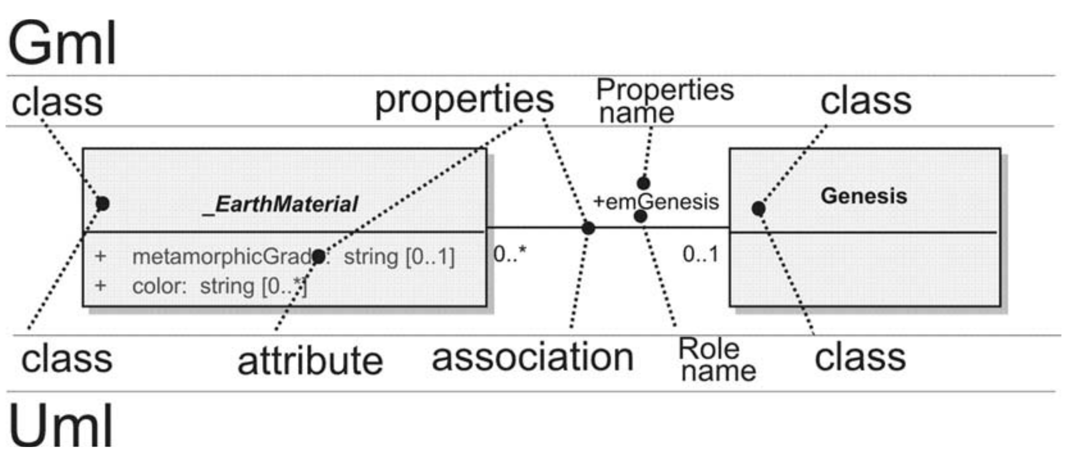


Figure 4. Representing GML using UML. The original UML diagram (top) is converted in a GML-friendly representation (bottom). Changes are described in the text.

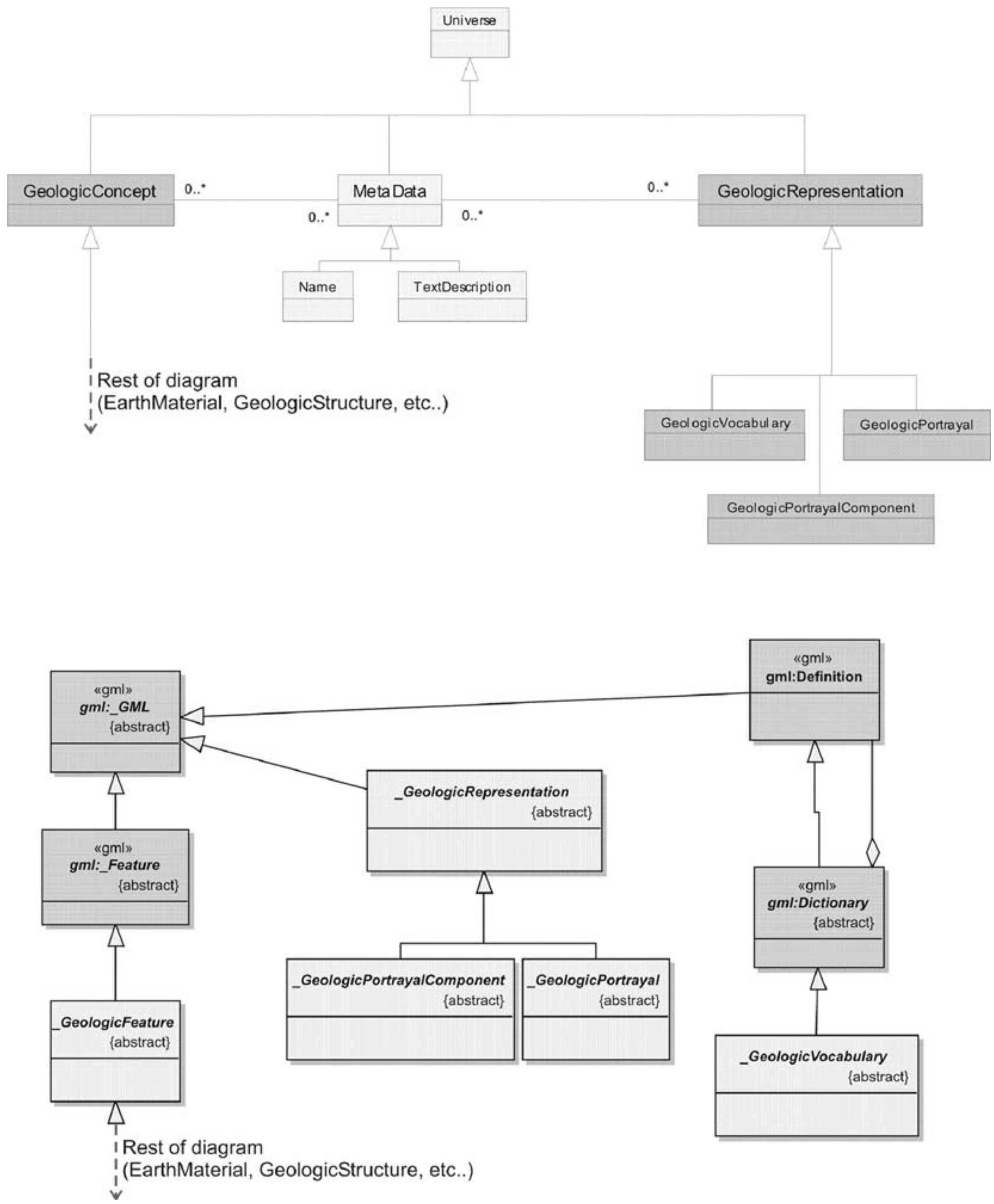


Figure 5. Adaptation of the top level of the model, essentially redirecting top level classes from NADM C1 to GML. Top part is original NADM C1 while bottom part is the GML interpretation. The original root of NADM C1 is a concept called `Universe`, from which derives `GeologicConcept`, `MetaData` and `GeologicRepresentation`. GML offers an alternative set of root concepts from which GML application must derive. For instance, `GeologicConcept` has been renamed `_GeologicFeature` to match GML syntax and now derives from GML's `_Feature`, which provides `_MetaData` functionalities. `_GeologicRepresentation` now derives from `_GML`, which is a high level abstract object. `_GeologicVocabulary` has been moved under GML `_Dictionary`, since this GML class offers the functionalities `GeologicVocabulary` intended to offer.

Therefore, top level classes of NADM C1 have been remodelled as descendents of GML features. Although multiple inheritance is allowed in UML, it is almost always suggested to avoid it in any kind of implementation, hence we did not keep the original top classes from NADM C1, but renamed the original **`GeologicConcept`** to **`_GeologicFeature`**. GML already provides the naming and documenting mechanism (**`_MetaData`**) so they don't need to be duplicated in NADM C1 encoding. The **`GeologicVocabulary`** from NADM C1 takes advantage of the GML Dictionary that provides the same core functionality.

Figure 6 shows some typical examples of changes that were made:

1. **`EMConstituent`** has been converted from an association class into a bridge class—a class that acts as a connector between two classes (**`EMConstituent`** was linked to the association between **`EarthMaterial`** and **`CompoundMaterial`**). Therefore, **`EMConstituent`** will behave as a container in GML that will wrap the **`_EarthMaterial`** class (see GML example, Figure 7).
2. **`EMCRelation`** has been remodelled as a GML property, because it essentially links two classes, **`EMConstituent`** and **`EMCRole`**. We also redesigned the way relations were described in the model. **`EMCRelation`** is a kind of **`_GeologicRelation`** (see NADM C1 documentation) and shows the parent class to emphasise that **`EMConstituent`** can have any kind of relation.
3. **`Proportion`** and **`EMCRole`** have been wrapped into properties of **`EMConstituent`**. This representation has the same meaning as the change made to **`EMCRelation`** (#2), but this representation is preferred when we feel that the property will

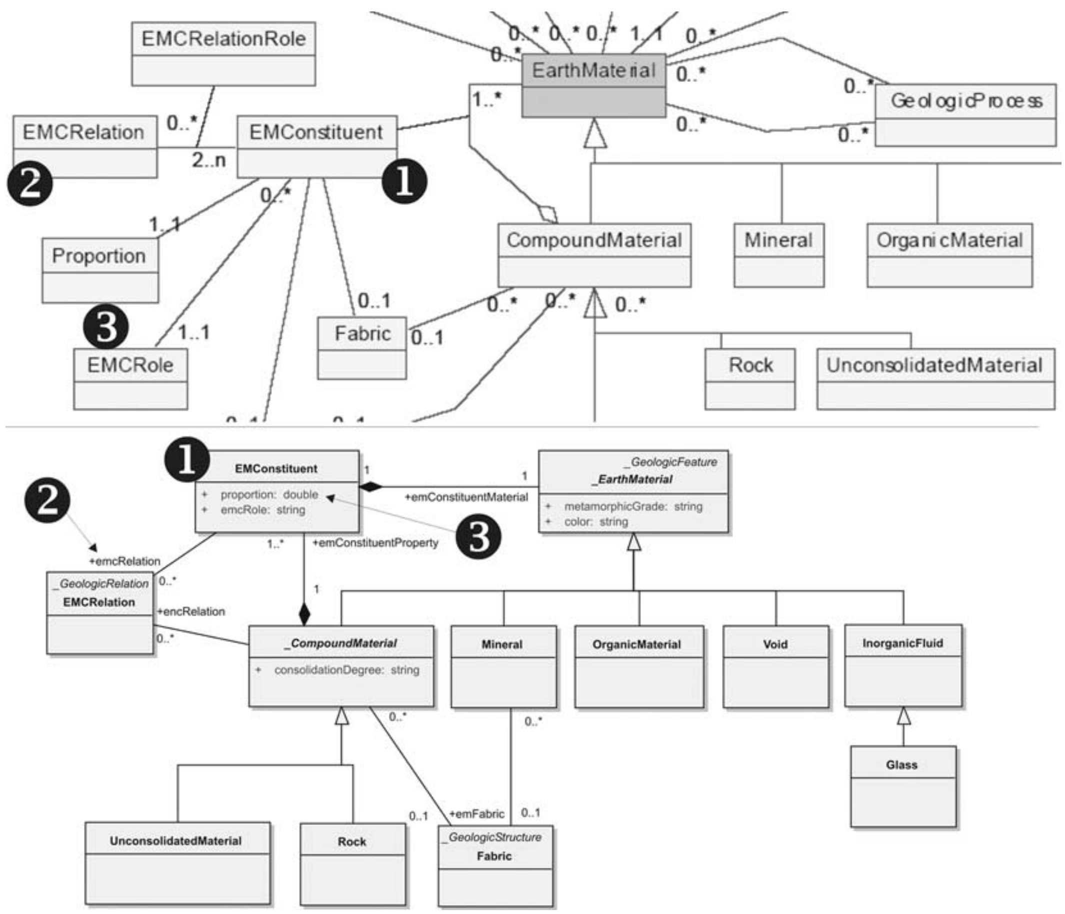


Figure 6. Reinterpretation of the conceptual model to conform to GML (and XML) constraints. The top part of the figure is from the original NADM C1 document while the bottom part shows adaptations to create a GML-friendly model. Principally, NADM association classes are characterized as GML bridge classes (see label #1), some NADM classes are characterized as GML complex properties that are depicted as UML associations (see label #2), and some NADM classes are characterized in GML as simple properties that are depicted as UML attributes (see label #3).

contain simple values (such as text or number), as opposed to complex values. We expect these properties in GML to be modelled as simple XML entities, ie, tags that cannot contain any other tags.

In some instances, the adaptation was not trivial. For example, **`GeologicAge`** might be replaced by the more detailed encoding designed by the XMML group. The issue at this point is: do we keep the simple structure that NADM C1 proposes (follow the guiding principle of being consistent with NADM C1) or do we simply reuse XMML, which essentially is the same thing but uses a different vocabulary and is more detailed (so, we would not reinvent the wheel).

EXAMPLE OF A NADM GML DOCUMENT

This example sums up the discussion about GML encoding. The portion of the document (the full document is located at https://www.seegrid.csiro.au/twiki/pub/CGIModel/EarthMaterial/earthMaterial.xml) should be compared to our model depicted on Figure 7 (numbers correspond on document, comments and figure).

```
<?xml version=”1.0“ encoding=”UTF-8“?>
<NADM xmlns=”http://geology.usgs.gov/dm/NADM/v1.0“ xmlns:xsi=”http:
//www.w3.org/2001/XMLSchema-instance“ xsi:schemaLocation=”http://
geology.usgs.gov/dm/NADM/v1.0 NADM-0_1.xsd“ xmlns:xlink=”http://www.w3.org/
1999/xlink“>
  <!-- Current max id number = 27 -->
  <featureMember>
    <Rock id=”nadm-1“>  1
      <!-- Rock connected to a genetic event, w/constituent minerals -->
      <name>Joe’s Granite</name>
      <description>Intrusive rock</description>
      <color>light gray</color>
      <consolidationDegree>consolidated</consolidationDegree>
      <emGenesis xlink:href=”earthMaterial.xml#nadm-19“/>
      <emFabric>  2
        <!-- fabric of Rock -->
        <Fabric id=”nadm-26“>
          <name>Fabric description</name>
          <pervasiveness/>
        </Fabric>
      </emFabric>
      <emConstituentProperty>  3
        <EMConstituent id=”nadm-10“>
          <!-- mineral w/ it’s own fabric -->
          <proportion>35</proportion>
          <emcRole>Mineral</emcRole>
          <emConstituentMaterial>
            <Mineral>
              <name>Quartz</name>
              <description>Silica</description>
              <color>white</color>
              <emFabric>
                <Fabric>
                  <name>Aligned C-axes</name>
                  <pervasiveness>pervasive</pervasiveness>
                </Fabric>
              </emFabric>
            </Mineral>
          </emConstituentMaterial>
        </EMConstituent>
      </emConstituentProperty>
```

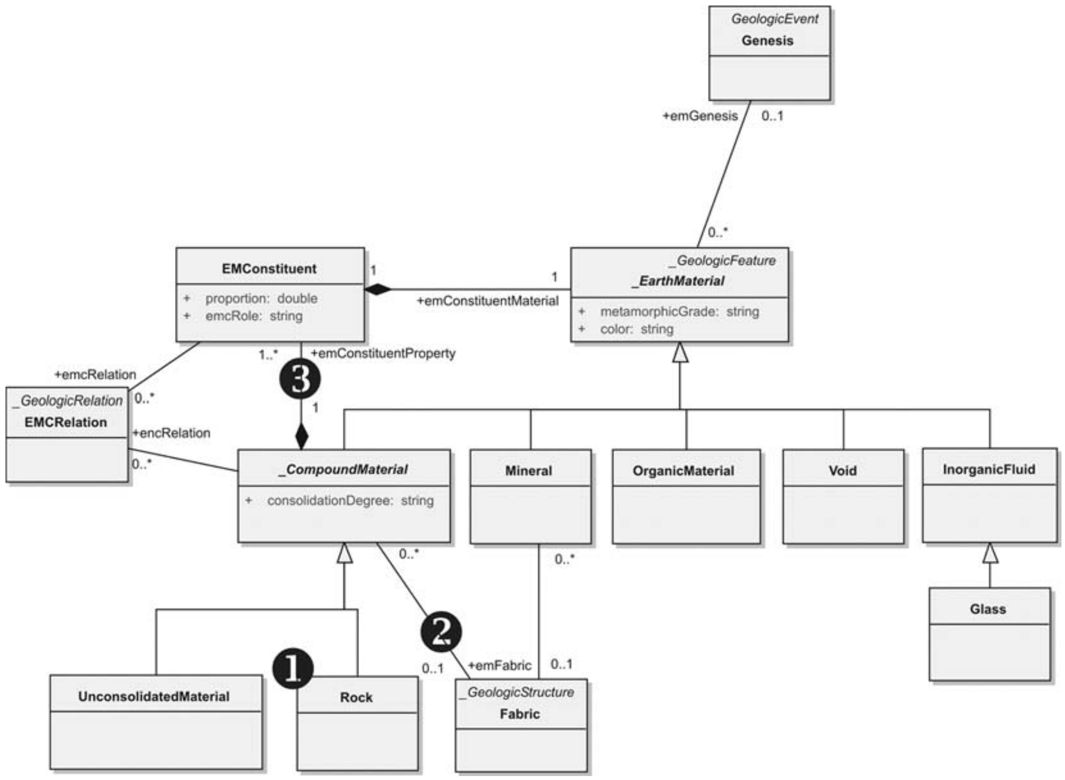


Figure 7. Example of a GML-compliant XML document with accompanying UML representation. The numbers on the figure refer to the example document and the notes in the text.

The main points noted in this example are:

1. A **`Rock`** class inherits from **`_CompoundMaterial`**, which itself inherits from **`_EarthMaterial.`** In other words, a **`Rock`** is a **`_CompoundMaterial,`** which is an **`_EarthMaterial`**. So, a **`Rock`** inherits the **`color`** and **`consolidationDegree`** properties (and also **`metamorphicGrade`**, but this property is optional). **`color`** is a simple property, shown as a UML attribute instead of an association. A rock also has a **`Genesis`**; in this example, we chose to point to a description in another document (earthMaterial.xml). **`name`** and **`description`** are inherited from far above the hierarchy shown in Figure 7, they are properties of the top-most class of the model.
2. A **`_CompoundMaterial`** can have a **`Fabric`**; note on the UML diagram that **`emFabric`** is the name of the association. Note the *CamelCase* structure; you can distinguish properties from classes quite easily by looking at how the name is capitalised.
3. A **`Rock`** is made of constituents (because it's a **`_CompoundMaterial`**); **`proportion`** and **`EMCRole`** are inherited from the **`EMConstituent`** class. The constituent material in this example is a Mineral, but other kind of **`_EarthMaterial`** can be substituted (**`Rock`** or **`Glass`** for instance).

INTERNATIONAL ACTIVITIES

This activity is a contribution to a larger international project hosted by the IUGS Commission for the Management and Application of Geoscience Information (CGI) (see http://www.bgs.ac.uk/cgi_web/tech_collaboration/tech_collab.html). As stated there, "*The overall objective of the Working Group is to develop international standards for the structure of geological information (i.e. data model standards) to enable interoperability among several national geological survey agencies.*" (see http://www.bgs.ac.uk/cgi_web/tech_collaboration/data_model/data_model.html). The working group progress is documented at https://www.seegrid.csiro.au/twiki/bin/view/CGIModel/WebHome.

CONCLUSION

Moving from generic XML encoding to GML has been beneficial for several reasons: a) it allowed encoding of the model to follow broadly accepted standards (ISO, OGC); b) it provided a much needed design pattern that resolved many consistency issues; and c) it opened NADM to other standards that are based on GML, such as WFS, and to other tool developers working with GML. As a by-product, encoding of the model was also an excellent review process and, for better or worse, generated a series of revision requests and comments to the NADM Data Model Design Team.

ACKNOWLEDGMENTS

Peter Davenport, Andrée M. Bolduc and Kathleen Lauzière (Geological Survey of Canada) have kindly reviewed this manuscript. Dave Soller's (U.S. Geological Survey) comments greatly improved this paper. Eric Boisvert has been financially supported by *the Consolidation of Canadian Geocience Knowledge* and *Groundwater* programs of the Earth Sciences Sector of Natural Resources Canada. Simon Cox's contributions were supported by the Minerals and Energy Research Institute of Western Australia and the sponsors of project M340, and by the Open GIS Consortium. Boyan Brodaric was supported by *Sustainable Development through Knowledge Integration* and *Consolidation of Canadian Geocience Knowledge* programs of Earth Science Sector of Natural Resources Canada.

REFERENCES

Digital Interchange Technical Team, 2003, XML Encoding of the North American Data Model, *in* Soller, D.R., ed., Digital Mapping Techniques '03—Workshop Proceedings: U.S. Geological Survey Open-File Report 03–471, p. 215-221, accessed at http://pubs.usgs.gov/of/2003/of03-471/boisvert/index.html.

Galdos System Inc., 2003, Developing and Managing GML Application Schemas, http://www.geoconnections.org/developersCorner/devCorner_devNetwork/components/GML_bpv1.3_E.pdf.

Laxton J., and Brodaric, B., 2003, Data Model Collaboration Working Group Report, Edinburgh, Nov. 2003, accessed at http://ncgmp.usgs.gov/intdb/dmic/dmic.html.

North American Geologic-map Data Model Steering Committee, 2003, NADM Conceptual Model 1.0. A Conceptual Model for Geologic Map Information: preliminary website release under the auspices of the Geological Survey of Canada, the U.S. Geological Survey, and the Association of American State Geologists, accessed at http://nadm-geo.org, now published and available at http://pubs.usgs.gov/of/2004/1334/.

OpenGIS Consortium, 2004, OpenGIS Geography Markup Language (GML) Implementation specification version 3.1, Cox et. al., ed.: OGC Document 03-105r1, accessed at http://portal.opengis.org/files/?artifact_id=4700.

Geologic Web Services, Planning and Design

By Sheena K. Beaverson

Illinois State Geological Survey
615 East Peabody Drive
Champaign, IL 61820
Telephone: (217) 244-9306
Fax: (217) 333-2830
e-mail: beavrsn@isgs.uiuc.edu

INTRODUCTION

Illinois State Geological Survey (ISGS) staff members have been developing interactive maps since June, 2001. These web resources provide access to scientific data for a diverse array of audiences. This paper will present the planning procedures developed, and technical lessons learned, from the successful completion and public release of multiple interactive maps.

Map Services and Interactive Maps from the ISGS

The Open Geospatial Consortium (OGC; http://www.opengeospatial.org/) Web Map Service Specification (WMS 1.3) states that a web map service (WMS) "produces maps of spatially referenced data dynamically from geographic information" and that "WMS-produced maps are generally rendered in a pictorial format such as PNG, GIF, or JPG" (Beaujardiere, 2004). In this paper, the term 'map service' represents the active service created in ArcIMS Administrator, referencing the .mxd map document developed with ArcMap software. The term 'interactive map' collectively represents a map service and the accompanying collection of HyperText Markup Language (HTML) .html files and javascript .js files that allow a web visitor to query and customize pictorial data display.

The ISGS currently publishes map services with the use of ESRI ArcIMS 4.0.1 software (www.esri.com). This software is further supported by ESRI ArcMapServer4.0.1 software, to enable the utilization of .mxd map documents generated by ESRI ArcMap 8.3 Geographic Information Systems (GIS) software.

The ISGS has released five interactive maps and two map services at multiple web sites. These are available at:

- The Illinois Natural Resources Geospatial Data Clearinghouse (http://www.isgs.uiuc.edu/nsdihome). Interactive maps are publicly available for the following data collections:
  - Illinois Digital Orthophoto quarter Quadrangle (DOQ) files from National Aerial Photography Program Three flights. Map initially released in June, 2001. Revised version released July, 2004.
  - Illinois Historical Aerial Photographs, 1936 through 1941. Map initially released in September, 2003. Revised version released October, 2004.
- The ISGS web site (http://www.isgs.uiuc.edu/). Interactive maps are publicly available for the following data collections:
  - Illinois Oil and Gas Resources, with access to more than 180,000 oil and gas wells. Map initially released in February, 2004. Revised version released November, 2004.
  - Illinois Natural Gamma-Ray Logs; released in June, 2004.
  - Illinois Water and Related Wells, with access to more than 275,000 water, engineering, and stratigraphic wells. Map released in July, 2004.

Map services featuring data for Illinois have also been developed for the Midcontinent Interactive Digital Carbon Atlas and Relational dataBase (MIDCARB; http://www.midcarb.org) project, the National Interactive Digital Carbon Atlas and Relational dataBase (NATCARB; http://www.natcarb.org) project, and for a custom file-sharing application for the ISGS and the Illinois Department of Transportation.

GEOLOGIC WEB SERVICES

Project Planning

To develop an interactive map, a complex sequence of steps has been delineated. This multi-phased approach has proven to be crucial, in order to usher a project smoothly through the various stages of planning, development, and testing, to public release.

Define the Scope of the Project

The first project phase introduces team members to obstacles inherent in the development of an interactive map. Scientific staff unfamiliar with GIS software may need a demonstration of the fundamental differences between paper and interactive maps. A demonstration should be held to review scale dependent rendering of data layers, outline priorities for allocating computer screen real estate between the various map manipulation controls and data display frames, and illustrate the limited viewing area available for the central map frame.

Initial discussions will logically lead to a review of potential data layers to be featured. It is important to decide on the order of data display as early in the project planning process as possible, to save time while implementing data-specific programming customizations. A series of meetings will be needed to accomplish this task, as this is an iterative process. It is the responsibility of the GIS specialist to guide the team into making decisions about logically ordering the collection of data layers. Data layers must be ordered in a vertical sequence that conforms to the rendering capabilities of GIS software. Data layers featuring points and lines must overlay layers featuring polygons, which can in turn overlay base imagery. Reference data, including points, lines, polygons and imagery will provide familiar context for the interactive maps' feature data layers, be they points, lines, or polygons.

An all-inclusive list of potential data layers (Table 1) should be provided at the initial meeting, to enlist help in identifying inappropriate data. A list that seems overwhelming in length should inspire the team to limit reference data options, in favor of showcasing the feature data layers. Presenting an abbreviated list of reference data to the team will cause 'what about my favorite data' additions, which can lengthen the overall list by tens of data layers. This is counter-productive to the ultimate goal for the project, which is to guide web visitors to the feature data quickly and logically.

Team members should be encouraged to approach project development with the goal of creating a tightly focused application. Data heavy applications tend to run slowly, and overwhelm new users with their large number of map customization options. An application incorporating such traits will cause customers to quickly become dissatisfied with the interface, and leave the web site before achieving their research goals.

Another early goal in the planning process is to identify the primary target audience. Potential audiences can be broken down into the broad classifications of Internet, Extranet, or Intranet. Resources available on the Internet are accessible to the general public. This can include landowners, policy makers, industry consultants, K-12 educators and students, and many others. Extranet, or inter-agency, resources can be targeted to a professional audience; people familiar with geologic maps. These customers include surveyors, drillers, college instructors and students, exploration geologists, or private consultants. Intranet resources, geared toward internal agency staff, can be customized for an audience with fast, in-house network connections, uniform access to support software, and clearance to confidential data elements.

Project Design

Data Collection and Assessment

The second project phase is simple in scope, but time-consuming to complete. It involves the collection, assessment, and cartographic display of reference and feature data resources. Assessment of existing data layers may trigger modifications based on incompatible data file

Table 1. Sample of an all-inclusive list of data layers, order determined by the rendering capabilities of GIS software.

- Reference Data (lines and/or points)
  - Power Plants
  - Municipal Water Sources
  - Roads
  - Rivers
  - Counties
  - Township, Range and Section
  - Quadrangle Boundaries
- Key Data (lines and points)
  - Data index points
  - Well locations
  - Geologic surface contours
- Key Data (polygons)
  - Quaternary geology
  - Bedrock geology
  - Other geologic data
- Reference Data (polygons)
  - Land Cover
  - Municipalities
  - Water Bodies
- Reference Data (imagery)
  - Ortho photography
  - Digital Raster Graphic files
  - Digital Elevation Models
  - Satellite Imagery

formats, the need to construct data sub-sets and/or screen confidential data, and requirements for programming custom data attribute display or hyperlink capabilities. Compilation of detailed information about the cartographic display of each data layer can prove to be an invaluable time-saving exercise. Sources of previously determined cartographic standards and symbols for feature data may be found on published maps, described in journal articles, archived in digital symbol sets, or recorded in the attribute definition fields of metadata files.

Development of the Map Document

Once the data layer order has been established and any necessary data revisions have been completed, development of the map document can commence. Tasks include importing base and feature data layers, and setting cartographic display parameters for individual data layers. Minor revisions to the established order of data layers will become evident during development of the map document, with respect to emerging complexities with information display. All project team members should be given a preliminary demonstration of the map document, so the GIS specialist can compile their feedback and address their concerns.

After issues resulting from the preliminary feedback have been addressed, technical staff can apply the predetermined customizations to the map document and support files. Such customizations could include: selecting a sub-set from the complete list of data attributes archived for each data layer, substituting easily interpreted aliases for scientific codes, enabling hyperlinks to related documents or images, and creating short but intuitive labels for data attribute fields. Because of the complexity of these tasks, it is important to delay these programming customizations until after the data order has been finalized and approved

Beta Testing, Round One

The first round of formal beta testing should be open to all team members to ask them to provide input on all features of the interactive map and the array of map tools in the display. Relevant areas for review are: font selection and font size, display of data layers, cartographic distinctiveness of data attributes, appropriateness of maximum and minimum settings for scale-dependant rendering of each data layer, display characteristics between multiple data layers with respect to symbology and settings for scale dependant rendering, required map manipulation and query tools, and completeness and usability of the overall interface. This level of review should be lengthy and in-depth.

Customization of the overall web page in which the interactive map frames are embedded should be delayed until this phase has been completed. A web interface that appears to be nearly complete can inspire reviewers to limit their editorial input, and thereby leave serious design flaws undiscovered. Fixing such flaws can be extremely time-consuming if done in a later phase of project development.

Beta Testing, Round Two

The second round of beta testing should be opened to all team members, agency management, and other interested staff members. Technical staff should seek input to ensure that previously identified editorial points have been addressed, verify that no new problems have been created, and certify that no major flaws were overlooked in round one testing. It is also constructive to solicit review from other developers of interactive maps. If the map tools, data display, and navigation elements aren't intuitively designed, these technical consultants will quickly identify key problem areas. This level of review should focus primarily on cosmetic features.

Project Completion

Support Documentation

The standard array of software required for accessing an interactive map will incorporate data, database software, software for map service presentation and customization, web server software, operating system software, plus other plug-ins or standalone components. Because of this complexity, it is advised that all map access be routed through a hyperlink on a single web page. Notices about map service interruptions, support documentation, and project information can also be featured at this location.

Support documentation should be provided for a variety of project details. Information about the overall project, appropriate metadata, an outline of the map tools (Figure 1) and default navigation settings (Figure 2), and a detailed legend (Figure 3) are useful to on-line customers. Prior to public release, the support resources and the interactive map will benefit from a final review by editorial staff, a professional cartographer, and the agency webmaster.

Navigation Strategies

While Internet consumers are familiar with tools embedded in web browser software, they are unfamiliar with map manipulation tools and the ways in which interactive map components can either work in tandem with or counteract browser tools. It is advisable to utilize the default map tool icons provided by the software provider, to educate users about the basic characteristics of interactive map tools over time. An agency should also create and adhere to design standards for the layout of all inter-

Navigational Tools of the ILH2O Interactive Mapping Web Interface	
Overview Map	Turn the Overview Map displayed in the upper left corner of the map frame on or off.
Zoom In	Enlarge the map image by clicking on a spot or dragging a rectangle around the area of interest.
Zoom Out	Reduce the map image by clicking on a spot or dragging a rectangle around the area of interest.
Pan Map View	Drag the map image so that the area of interest is shifted within the view.
Previous Extent	Return the map view to the previous zoom level and location.
Zoom to Full Extent	Zoom the map view out to the the fullest extent of all data layers, in this case, the boundary of the state of Illinois.
Select Within	Graphically select multiple records by dragging a rectangle around the area of interest. Features in the rectangle will be selected and highlighted in light blue. A table of information pertaining to the active layer will appear. *The user MUST specify the "active layer" to the desired data layer county.*
Hyperlink	Graphically select a water well data record for viewing from the "Water and Related Wells" data layer. A separate file-viewing window appears to display the well record. *There is no need to specify an "active layer."*
Identify	Identify a selected feature. A table of information for the selective feature and within the active layer will appear.
Measure	Click on the map to create line segments. Distance (in miles) for each line segment and the total of all segments will be displayed.
Set Measure Units	Specify the "Map Units" and "Display Units" to be used with the "Measure" Tool. "Map Units" can be set to feet, meters, or degrees. "Display Units" can be set to feet, miles, meters, or kilometers.
Clear Selection	Unselect all previously selected features.
Print Map View	The user is prompted to submit a map title. The user types and submits the title, a new window opens that contains a formatted title, map image, and legend suitable for printing. Use the browser "Print" button to print the map.
Legend or Layers	Toggle between the layer list and the map legend. The layer list enables the user to set the active data layer and also set the display of individual data layers to be on or off. The legend is a graphical representation of the symbology of all visible map layers.

Figure 1. Detailed description of navigational tools inherent in the interactive map, available at http://www.isgs.uiuc.edu/wwdb/launchims.htm.

Instructions for Accessing Illinois Water and Related Well Data Resources

Map manipulation tools are offered on the left side of the interface. The currently selected tool will be outlined in red, as the "Select Within" tool is outlined in the screen shot shown below. Data layer visibility and activation control is located on the right side of the interface. The visitor must set desired "Active" layer combinations and then implement "Refresh Map" in order to create a custom map view.

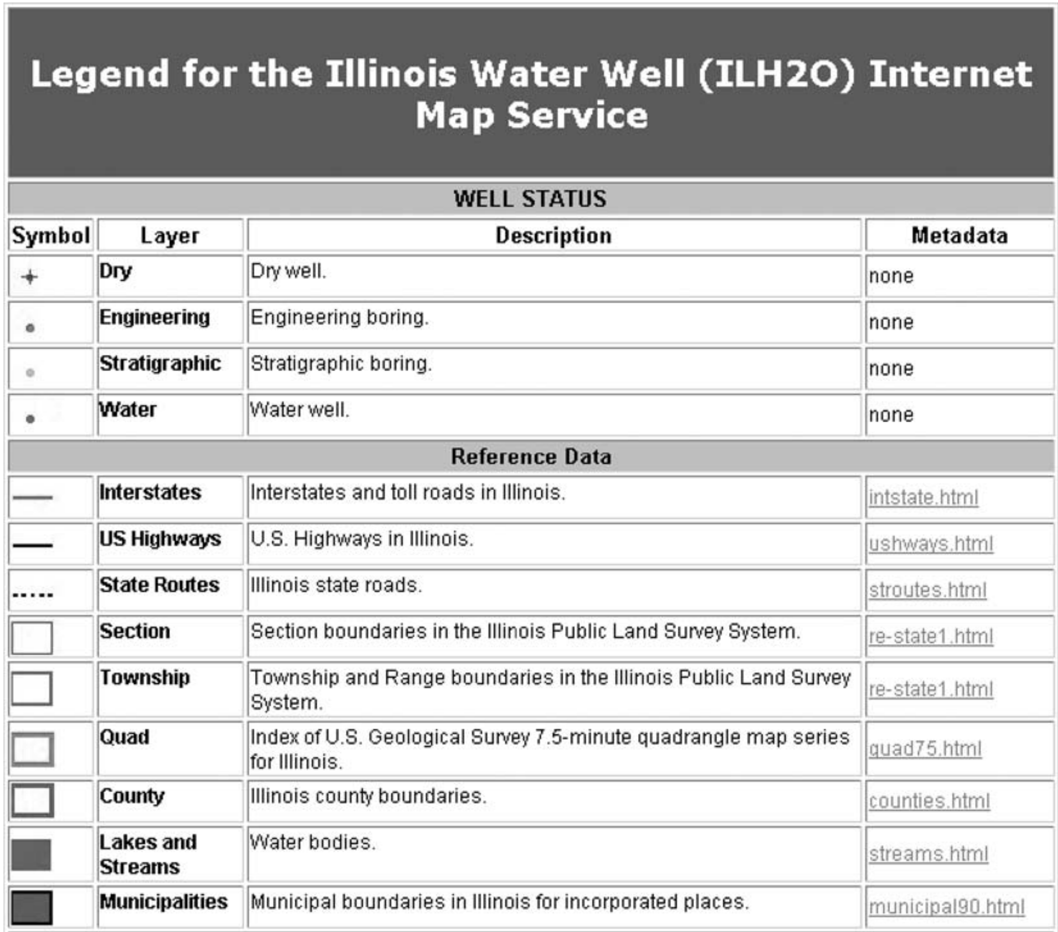
Step 1: Enlarge an appropriate area.
The "Zoom In" tool is the default tool upon launch of the Interactive Mapping Interface.

Step 2: Set to "Visible" any data layers you wish to enable.
Geologically relevant data layers are listed in order by Depth of Aquifer. By default, the "Major Sand and Gravel Aquifers" data layer is set to "Visible." Many combinations of data layers can be set to "Visible."

Step 3: Click the "Refresh Map" button.
The "Refresh Map" button is listed below all data layers and will complete alterations to the selection of "Visible" layers. Data layers "Water and Related Wells," "Interstates," "US Highways," "State Routes," "Sections," "Townships," "Counties," "Lakes," "Municipalities," and "Major Sand and Gravel Aquifers" are set to "Visible" in the screen shot shown below.

Step 4: Access Data
To use the "Select Multiple" rectangle tool, first set the "Active" layer to the desired data layer by clicking the corresponding radio button. The "Water and Related Wells" data layer is set as "Active" in the screen shot shown below.

Figure 2. Instructions for interacting with the data and customizing the map display.



Legend for the Illinois Water Well (ILH2O) Internet Map Service			
WELL STATUS			
Symbol	Layer	Description	Metadata
	Dry	Dry well.	none
	Engineering	Engineering boring.	none
	Stratigraphic	Stratigraphic boring.	none
	Water	Water well.	none
Reference Data			
	Interstates	Interstates and toll roads in Illinois.	intstate.html
	US Highways	U.S. Highways in Illinois.	ushways.html
	State Routes	Illinois state roads.	stroutes.html
	Section	Section boundaries in the Illinois Public Land Survey System.	re-state1.html
	Township	Township and Range boundaries in the Illinois Public Land Survey System.	re-state1.html
	Quad	Index of U.S. Geological Survey 7.5-minute quadrangle map series for Illinois.	quad75.html
	County	Illinois county boundaries.	counties.html
	Lakes and Streams	Water bodies.	streams.html
	Municipalities	Municipal boundaries in Illinois for incorporated places.	municipal90.html

Figure 3. Legend for all data layers included in the interactive map, including cartographic symbols, data layer descriptions, and hyperlinks to metadata files.

active maps. Adherence to such standards will, with each new project release, teach visitors where to look within the layout of each map for the controls to query data and customize data display (Nielson, 1999). Other methods to improve the usability of the map itself would be to group map tools into useful categories and provide intuitive, abbreviated descriptions of each tool function (Figure 4).

Enthusiastic adoption of a new interactive map by the target audience will require advertisement. Promote the project release at workshops and professional conferences with 'how-to' navigation demonstrations. Newly developed interactive maps are typically presented to project funding supporters, at agency seminars and at meetings of the state GIS Association.

ACKNOWLEDGMENTS

The following staff members made significant contributions to the development of geologic map services at the ISGS: Steven Gustison, Christopher Korose, Robert Krumm, Alison Lecouris, Christopher McGarry, and Daniel Nelson. Editors included Jonathan Goodwin and Robert Krumm of the ISGS, and David Soller of the United States Geological Survey. Project support has been provided by many state and federal agencies, including the Illinois State Geological Survey, the Illinois Department of Natural Resources, the Illinois Department of Transportation, the Institute of Museum and Library Services, the United States Department of Energy through the National Energy Technology Laboratory, and the United States Army Corps of Engineers.

REFERENCES

Beaujardiere, Jeff de (editor), 2004, Web Map Service: Open Geospatial Consortium, Inc., OGC-04-024 version 1.3, accessed at http://portal.opengeospatial.org/files?artifact_id=5316.

Nielson, Jakob, 1999, Do Interface Standards Stifle Design Creativity?, accessed at http://www.useit.com/alertbox/990822.html.

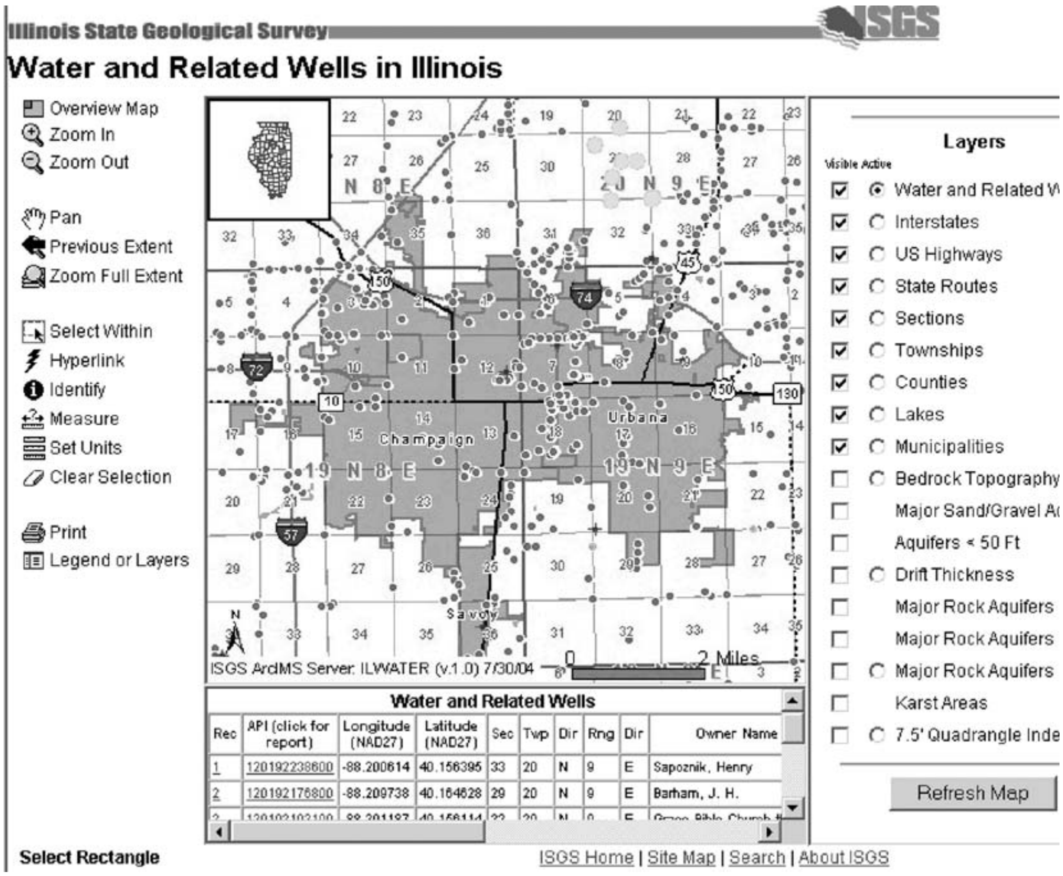


Figure 4. The "Water and Related Wells in Illinois" interactive map published by the ISGS. Map tool descriptions (Zoom In, Pan, Identify, etc.) have been included to improve the usability of the map.

Implementing NADM C1 for the National Geologic Map Database

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@usgs.gov

[3]U. S. Geological Survey
926A National Center
Reston, VA 20192
e-mail: drsoller@usgs.gov

BACKGROUND

This report is a summary of progress on implementing the NADM C1 conceptual model (NADMSC, 2004) as a production-scale prototype for a U.S. Geological Survey –Association of American State Geologists (AASG) National Geologic Map Database (NGMDB, Soller and Berg, 2003). The implementation uses standard relational database technology, and is designed to function as an ESRI geodatabase or as a standalone, non-geographic database. The implementation is designed for depth of knowledge representation and flexibility, not for simplicity or performance. The NGMDB will be a data archive for geoscience information, with provision to record alternative interpretations, evolving terminology and science paradigm, uncertainty, incomplete knowledge, and metadata pertaining to data acquisition, processing, and automation. The objective is integration of geologic data from published maps by different authors at different scales, as well as newly acquired field data.

In more concrete terms, the objective of the project is to design and implement a scalable database for storing geologic descriptions, particularly those related to geologic maps (e.g., geologic units, lithology, and geologic structures), as well as the location and geometry of mappable features. Development has been ongoing for several years. The development of an underlying conceptual model and science vocabulary has taken place in a community arena (the North American Data Model Steering Committee, http://nadm-geo.org) in order to achieve some level of standardization. Physical implementation has been guided by discussions with mappers in the U.S. National Cooperative Geologic Mapping Program. Implementation of an enterprise-scalable database requires that we address business rules for security, ownership, and authority of data contained in the database. Integration of geologic data from different authors will require the maintenance of documentation (metadata) for each data object, such that the original source of information can be determined. Because it is critical for users to get data into and out of the database, considerable effort is being spent to develop a software interface to the NGMDB prototype database

NADM MODEL

Our underlying framework for information system development is the NADM C1 conceptual model (NADMSC, 2004). This model specifies the basic kinds of geologic things of interest and how they are described. It does not specify a database implementation. Table 1 summarizes significant concepts from the NADM C1 model that are used in the NGMDB implementation.

Table 1. Major concepts in NADM C1, as used in the NGMDB implementation.

Concept	Scope and rationale
EarthMaterial	A naturally occurring substance in the Earth. EarthMaterial represents substance, and is thus independent of quantity or location. Ideally, EarthMaterials are defined strictly based on physical and chemical properties, but because of traditional geologic usage, genetic interpretations enter into the description as well. Does not include melted rock (magma or lava). Many concepts related to water or petroleum have not been modeled in this version.
GeologicEvent	An identifiable event during which one or more geologic processes act to modify geologic entities. A GeologicEvent may have a specified GeologicAge and GeologicEnvironment. An example might be a cratonic uplift event during which erosion, sedimentation, and volcanism occur.
GeologicProcess	A function, possibly complex, that acts on one geologic entity to produce another geologic entity at a later time. Process is time independent; some GeologicProcesses are observable in the present at work in the world or in the laboratory, others can only be inferred from observing the results of the process. Processes take one or more of EarthMaterial, GeologicUnit, or GeologicStructure as input and have one or more of EarthMaterial, GeologicUnit, or GeologicStructure as output.
GeologicProperty	An inherent feature used to characterize a GeologicConcept.
GeologicRelation	Any of a wide variety of relationships that can exist between two or more GeologicConcepts. For example, the GeologicRelation "intrudes" is a relationship between an intrusive igneous rock and some host rock. Includes spatial, temporal, sequence, correlation, and parent/child relations. Many of the relationships in NADM-C1.0 (particularly attribute links and parent-child links) are not explicitly modeled as kinds of GeologicRelation.
GeologicStructure	A configuration of matter in the Earth based on describable inhomogeneity, pattern, or fracture. The identity of a geologic structure is independent of the material that is the substrate for the structure. There are almost always strong dependencies between the nature of the material substrate and the kinds of structure that may be present. Geologic structures are more likely to be found in, and are more persistent in, consolidated materials than in unconsolidated materials. Properties like 'clast-supported', 'matrix-supported', and 'graded bed' that do not involve orientation are considered kinds of GeologicStructure because they depend on the configuration of parts of a rock body. Includes sedimentary structures.
GeologicVocabulary	A collection of concept definitions, each associated with a preferred name, and usually organized in some logical fashion such as in a hierarchy. The preferred name associated with a concept in a GeologicVocabulary is a proxy for the collection of property values and relationships specified in the definition. The vocabulary makes the definitions of these concept instances available to apply in other descriptions without having to reconstruct the entire description denoted by the concept definition. Examples of geologic vocabulary include a collection of standard rock types, a stratigraphic lexicon, or a geologic time scale.
Legend	An ordered collection of LegendItems. A map legend specifies a collection of symbols (including patterns and colors) displayed on a geologic map or cross-section, along with the meaning or geologic description assigned to each symbol.
LegendItem	An association of a concept or description with a symbol. Each LegendItem instance represents a single entry in a map legend that describes either a single entity or a single class of entities occurring on a geologic map or cross-section.
MapDescription	All of the descriptive information that accompanies the graphic portion of a geologic map or cross section. Includes descriptive text, symbols and their explanations, associated graphics, etc.
SpatialObject	A description of the geometry (size, shape, and location) of an occurrence. Commonly represented as points, lines, or polygons.
GeologicUnit	A geologic unit is a part of the solid Earth that is identified by its geologic characteristics, has definable, locatable boundaries, and is persistent in time. Excludes non-material, temporal units. It is a body of earth material distinguished from adjoining material on the basis of content (lithologic or fossil), inherent attributes, physical limits, geologic age, or some other property or properties [adapted from NACSN, 1983, p.22; http://www.agiweb.org/nacsn/code2.html]. Corresponds to 'stratigraphic unit' in the North American Stratigraphic Code. Commonly used properties include composition, texture, included fossils, magnetic signature, radioactivity, seismic velocity, and age. Sufficient care is required in defining the boundaries of a unit to enable others to distinguish the material body from those adjoining it [NACSN, 1983].

IMPLEMENTATION

Issues-Extensions in NGMDB Model

Because the NADM C1 model is conceptual, it does not specify the details that must be implemented in a database. The following discussion lists some of the property extensions and implementation decisions that were made in developing the NGMDB prototype.

Numerical measured quantities may be represented by specifying a typical value, for example when an analysis requires a single 'best' value for the quantity. Alternatively, minimum and maximum values may be used to specify bounds on an uncertainty envelope (either symmetric or asymmetric about the typical value), or the upper and lower bounds on a value range. Measurement units are specified from a standard set of terms. A quantity type term in the database indicates whether the value is an average with standard deviation bounds, a range value, a value with asymmetric uncertainty bounds, or a single value with no uncertainty estimate.

The NADM C1 model does not provide for schema representing basic field observation locations for newly acquired field data. The NGMDB model has an abstract ObservationLocation feature type that is the supertype for field data acquisition sites—Station for point data, and Section for observations along a track (line) and AreaOfInterest for observations pertaining to an area. These model elements are based on the XMML site and specimen model by Simon Cox (2004, https://www.seegrid.csiro.au/twiki/bin/view/Xmml/SitesAndSpecimens). An observation location may have one or more associated structure observations (measured orientations of structures), text descriptions, images (sketch, photo), or samples.

Boreholes are not accounted for by the NADM C1 model. NGMDB implements a simple model for borehole data based on XMML, in which a borehole is modeled as a subtype of Section, a kind of observation location. The borehole collar point feature represents the XYZ location of the intersection of a borehole with the Earth surface. A borehole may be reentered with new boreholes drilled as splays from an existing borehole. Thus, one or more boreholes may be associated with a single borehole collar. Each borehole may be associated with an ordered collection of borehole segments that constitute an XMML interval log. Each borehole segment may have associated structure observations, text descriptions, images, or samples.

The Morphology property of GeologicUnit (see Plate 3, NADMSC, 2004) is implemented through links to a standard list of geologic unit morphology terms, and in the case of lithostratigraphic units, to a numerical value for thickness. The GeometricDescription property of GeologicUnit in NADM C1 has been expanded for lithostratigraphic (bedded) units to describe bedding style and bedding pattern with standard terms, and bedding thickness with either a standard term or a numerical thickness value. After discussion with geologists who are mapping surficial deposits (mostly in the arid southwest of North America), the descriptive properties associated with surficial geologic units were expanded to include terms for degree of dissection, surface armoring (pavement development), soil development, clast weathering character and style, and varnish development. Additional properties are probably necessary to fully describe surficial deposits in glacial, polar, temperate, and tropical environments.

The NADM particle geometry description (see Plate 2, NADMSC, 2004) has been implemented with additional properties grouped as particle shape and particle size descriptions. Particle shape properties include roundness, form, and degree of crystal face development. Particle size description may include quantitative specification of particle diameter (mean, median, maximum, etc.), as well as terms specifying sorting, particle size (diameter), and particle size range.

The NGMDB implementation allows multiple values to be assigned for many properties, through an attribute relationship correlation table. This mechanism allows qualifiers and metadata to be associated with each attribute value assignment. Qualifiers provide information on frequency that a value is observed, confidence in value assignment, and intensity of development of an attribute. Ability to assign a frequency to some value allows the expression of negation, i.e., the fact that a particular value is not allowed. The ability to express confidence allows representation of specific, estimated or guessed property values at low confidence, and less specific generalized property values at higher confidence. Metadata associated with attribute value assignment can record the measurement procedure, or more detail on evidence for assigning a value (e.g., why the environment attribute is assigned as 'fluvial').

In a distributed information system, data from many sources may be integrated for a single analysis, and evaluation or interpretation of results will require knowledge of the provenance of individual units of data. The feature level metadata (tracking) implemented in this database allows for the recording of links to citations for published sources, the person, organization, and project responsible for original data acquisition, and the processing steps involved in automating the information.

NGMDB Implementation Framework

Based on previous experience with geologic map database design and implementation, especially the evolution of implementations based on the Johnson et al. (1998) model version 4.3 (see http://nadm-geo.org for examples listed under "Data Model Design Team"), it became apparent that a widely applicable geoscience database implementation must be adaptable to evolving data requirements. In order to provide the flexibility and expressiveness required for a widely applicable geoscience information system, we are implementing a relatively abstract logical model that allows different users to configure the data structure to include entities and properties appropriate for their requirements. The underlying goal is to code the semantics of the data into the database to the maximum extent possible, as opposed to structurally incorporating semantics into data design. This requires including vocabularies that define terminology in the data, and encoding the data schema in the data. The most extreme version of such a design results in a one-table database (see design of the Protégé knowledge tool database backend, http://protege.stanford.edu/doc/design/jdbc_backend.html). This results in maximum flexibility and minimum comprehensibility of the raw data store. The NGMDB design documented in this paper includes a relatively small number of physical 'base tables' for standard kinds of geologic descriptions, and a standardized mechanism for extending these with properties of interest for a particular application.

The implementation builds on Arizona Geological Survey (AZGS) designs documented in Richard (2003) and Richard and Orr (2001), the Canadian Cordlink 5.2 design (Brodaric et al, 1999), and previous NGMDB prototype designs (Soller et al., 2002; Brodaric and Hastings, 2001, 2002). The design has been influenced by the Ontology Web Language (OWL), analysis of the data structure used by the LegendBurster tool (http://www.georeferenceonline.com/LegendBurster/), and various models proposed as part of the International Organization for Standardization Geographic Information/Geomatics project (ISO TC211, http://www.isotc211.org/), especially the Geography Markup Language (GML, Cox et al., 2004) and Exploration and Mining Markup Language (XMML, https://www.seegrid.csiro.au/twiki/bin/view/Xmml/WebHome).

The NGMDB implementation of the NADM C1 conceptual model revolves around three logical elements—Vocabularies, Description Schema, and Data Instance (Figure 1). Vocabularies are collections of terms and text definitions, analogous to GML Dictionary and Definitions (Cox et al., 2004). A Vocabulary constitutes an enumeration of things thought to exist in a domain or of possible values for properties. A Vocabulary also may include relationships between the represented concepts (terms), in particular a 'kind of' or subsumption hierarchy where appropriate. A mature vocabulary also might include thesaurus type relationships, to allow users to map terms between different vocabularies, and search for similar or related terms within a vocabulary. The Description Schema is an explicit representation of the implemented data model that is part of a dataset; analogous to an XML schema contained in a .xsd document. The Description Schema represents the data model, including kinds of objects, their properties, relationships between objects, and rules that determine valid database conditions. Data Instances are valid descriptions based on the Description

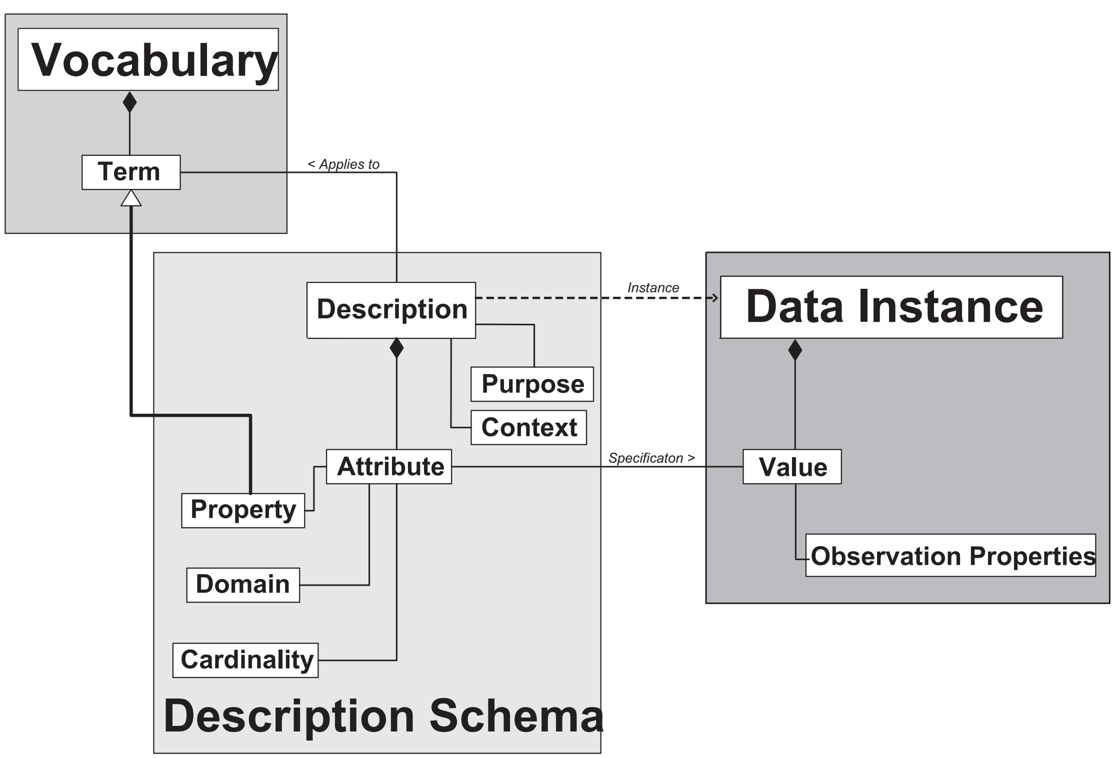


Figure 1. Logical framework for NGMDB implementation.

Schema, each of which specifies attribute values for some entity of interest.

Vocabulary Tables

Vocabulary tables contain collections of terms with associated text scope notes or definitions. If appropriate, the terms may be structured with parent-child links in the vocabulary table to define a tree hierarchy. The vocabulary defines a collection of shared concepts that may be used to classify observations, or to specify attribute values. The terms in the vocabularies are used to populate pick lists in the user interface that is under development. Each vocabulary table includes a unique identifier for each concept, the preferred name (the term) by which the concept is communicated between human users, a text notes to convey to users the meaning of the concept, and a link to a tracking record that supplies information on the source of the term and its definition.

Table 2 is a summary of the vocabulary tables implemented for the NGMDB prototype. Because the structure of each of these tables is the same, they could all be implemented in a single physical table. They have been implemented in separate tables to make them more portable – different vocabularies will be useful in different environments depending on the audience and the purpose for the geologic data. The geologic unit vocabulary will include names for geologic units that are found in the area of interest, and the terms required in the standard lithology and mineral vocabularies are determined by the kinds of Earth materials that are found. Different geologic time scales will require different vocabularies of stratigraphic eras. Richard et al. (2003) discuss some issues with integrating data using different vocabularies. The basic rule is that data integration is simplest if everyone is using the same vocabulary. Conflation of data using different vocabularies will require a thesaurus that matches terms in one vocabulary with those in the other, and may involve information loss.

The science language vocabulary is really a collection of many vocabularies that enumerate kinds of phenomena, or terminological values that may be used to quantify properties. It includes vocabularies for kinds of geologic structures, geologic process, geologic units, physical properties and geologic relationships. Examples of included property value vocabularies include Earth Material genesis terms, particle size terms, particle sorting terms, particle form terms, geologic unit rank terms, consolidation degree terms, and metamorphic grade terms.

Data Schema

The NGMDB implementation includes tables that explicitly store the data schema (Figures 1, 2, and 3). This approach is similar to that used for XML documents, which have an associated schema document (xsd file) that specifies the structure of data instances, and is similar to the schema information recorded by the internal tables (‘GDB_’ tables) in ESRI geodatabases. The schema identifies a collection of entities (ESRI ‘object classes’), a collection of properties (the property vocabulary mentioned above), associates each entity with a set of properties, and specifies a value domain and cardinality for the property in that entity. In a standard relational database implementation, each entity would be a physical table, each property associated with the entity would be a field in the table, and for properties that are specified using terms the list of terms used to populate the field might be a separate table. In the NGMDB implementation, each entity is implemented by a physical table in which each row corresponds to an instance of the entity. Attribute values are associated with entity instances through AttributeRelationship instances. Each AttributeRelationship instance (row in the AttributeRelationship table) correlates an entity instance with an attribute and a value for that attribute (Figure 2), along with observation-related metadata. Attribute values are specified by value specification instances that may be a science language term, measured quantity instance, text

Table 2. Vocabularies in the NGMDB prototype implementation.

Vocabulary Table	Content
GeologicUnit	Known geologic units within some area of interest.
StandardLithology	Terms associated with descriptions for standard kinds of rocks and unconsolidated material.
StandardMineral	Kinds of minerals that may be constituents in compound Earth materials.
EntityPropertyTypes	Kinds of properties that may be used in descriptions.
ScienceLanguage	Collection of vocabularies for kinds of things and property values. In any data repository this table will include both ‘infrastructure’ terms shared by all databases, and local terms defined for use in this repository. The infrastructure includes basic geoscience terminology, plus interdisciplinary terminology (units of measurement…), and some metadata terminology used in the information system.
StratigraphicEra	Named eras in a geologic time scale.

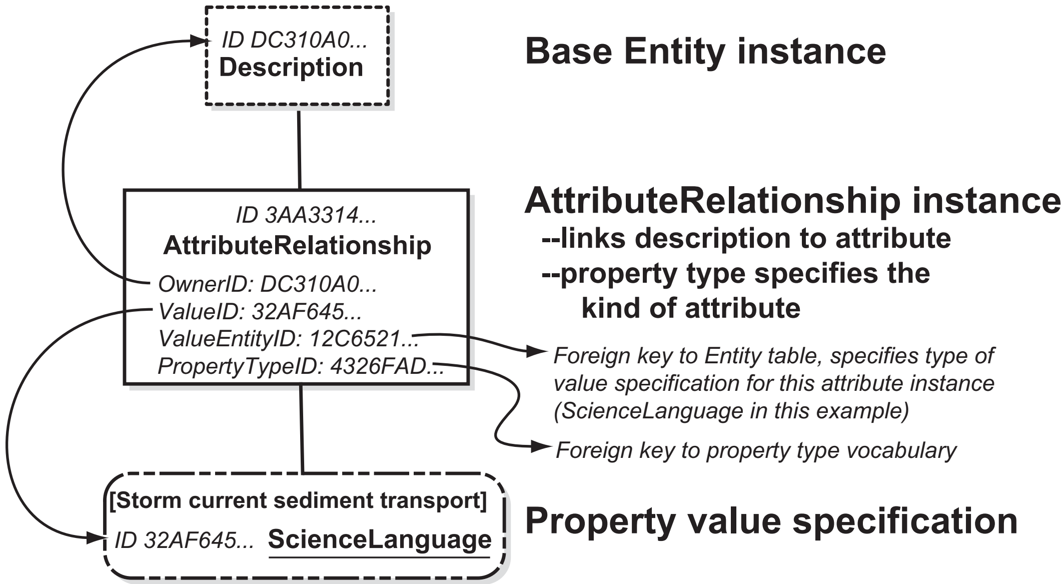


Figure 2. Diagram showing connection of logical attributes to a description through the AttributeRelationship correlation table. The AttributeRelationship table has attributes that specify the kind of attribute, and the entity type that contains the value specification. Observation-related metadata fields in AttributeRelationship are not shown.

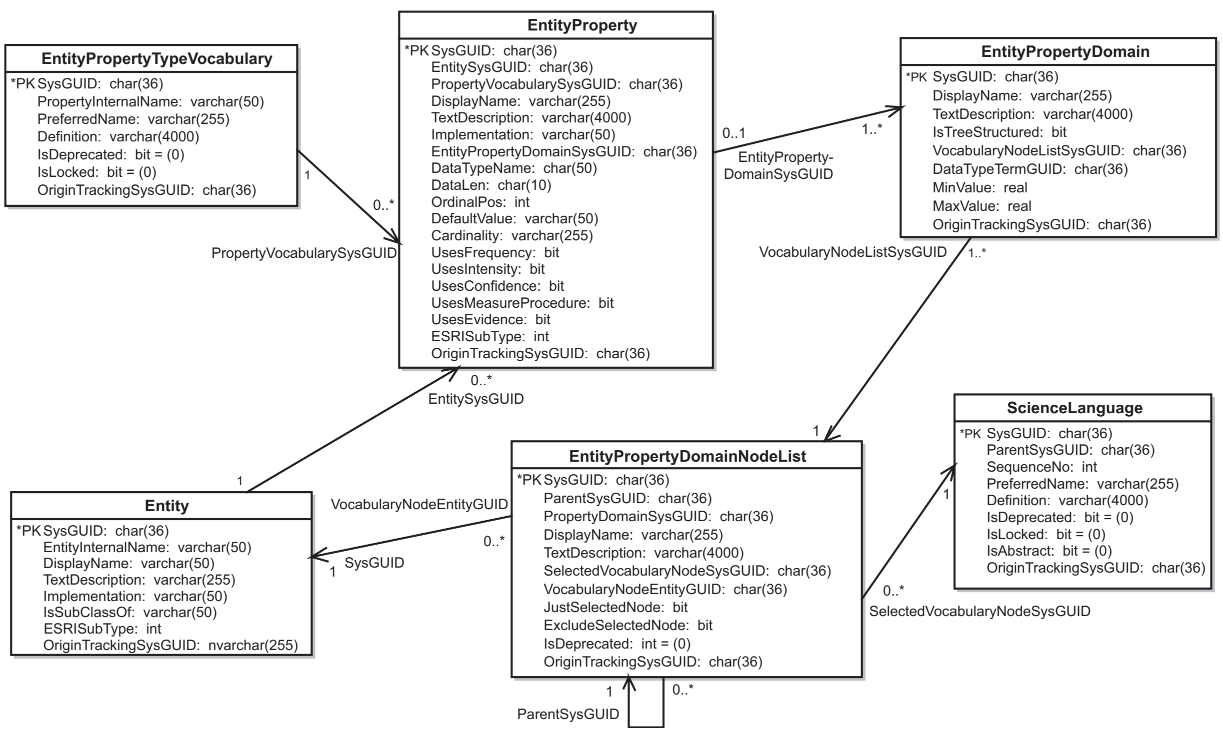


Figure 3. Data schema tables. Lines in the diagram represent foreign key relationships between tables. The field names at the end of lines adjacent to the table boxes indicate the field that is the key in the table at that end of the relationship. The EntityPropertyDomainNodeList table is used to enumerate data instances available in the pick list for a particular attribute. The SelectedVocabularyNodeSysGUID links are to any table that contains data objects that may be made available in pick lists; most commonly these will be vocabulary tables—ScienceLanguage, StandardLithology, StandardMineral. Science language link is shown here for example. The VocabularyNodeEntityGUID identifies the table that contains the referenced node.

description, or instance of another entity. The ValueEntity attribute in the AttributeRelationship table specifies the kind of value specification used.

The Data Schema tables identify the base table associated with each entity, enumerate the properties associated with each entity, indicate how each property is specified, and specify the domain associated with each property in each entity (Figure 2, Table 3). Data Schema tables all have names with the prefix ‘Entity’.

Data Instances

Data instances are implemented by a row in a base table, possibly with additional attributes associated through relationship table correlations. In the simplest case, the base table implements a single entity specified in the Data Schema. In more complex cases, multiple, related entities may be implemented in a single base table (similar to ‘object subtypes’ in ESRI geodatabase); entities that are implemented through a shared base table are referred to as logical entities. Base tables that implement logical entities include an attribute that specifies the entity type for each data instance (row) in the table. The structure of a data instance is specified by the associated entity type defined in the Data Schema.

Each Data Instance is a collection of attributes assigned to some object of interest represented in the database. The Data Instance has an associated Entity specification in the data schema tables that dictates what attributes are associated with instances of the entity, and how values for these attributes are specified. Standard, physical tables in a relational database structurally associate a data instance (row) with a collection of single-valued (0..1, depending on cardinality defined in Data Schema) attributes, each specified in a field in the table. Values may be assigned directly by numbers or strings in the field, or the field may contain a foreign key to a more complex value specification. Because the data schema is implemented directly in the table structure, entities implemented as physical tables do not require an explicit link to their entity type.

The base table for a group of logical entities (e.g. GeologicUnitDescription) contains single-valued attributes that are required by all of the logical entities implemented through that base table. The distinct logical entities implemented by a single table may have different AttributeRelationship associations, attribute domains, or cardinality constraints. Attribute values assigned through AttributeRelationship instances may have 0 to many values. Even a single-valued property may be specified using multiple AttributeRelationship instances with different observation properties. For instance a particular geologic unit may be assigned a Proterozoic age with high confidence based on stratigraphic relationship, and a middle Proterozoic age (more specific within the value range) with low confidence based on lithologic correlation. Each logical entity instance must have as one of its attributes a link to the entity definition that specifies its structure.

Base data tables are discussed in groups based on their content and use. The groups include value specification tables, GIS feature classes, section location tables, sample table, description tables, and relationship tables. The value specification tables represent observations of some individual property value specified by a numeric measurement, text, an image, or geometry. GIS feature classes represent located geographic data. Section loca-

Table 3. Data schema tables. These tables are further explained in the Appendix.

Table	Description
Entity	Vocabulary of types of data instances (entities) that may be implemented in the database. Each type specifies a collection of properties, each with a cardinality and value domain. The instances of the entity may reside in a single physical table, or be implemented as a logical structure with a base table and attribute values associated through correlation tables.
EntityProperty	Table that correlates properties defined in EntityPropertyVocabulary with entities that may specify values for the property, assign a cardinality for property values in the entity, and a domain of possible values for the property in the entity.
EntityPropertyDomain	Table that defines domains that may be used to specify property values.
EntityPropertyDomainNodeList	Table that explicitly enumerates terms in a value domain. May aggregate terms from one or more vocabulary entities into a single ‘domain’ or term pick list used to populate some property in some entity. If JustSelectedNode is false, then all children of a selected term should be included in the pick list. Exclude selected node is used to exclude an abstract term that is used in the vocabulary table as the root for some pick list.
EntityPropertyTypeVocabulary	Vocabulary of properties that may be used in descriptions. Analogous to vocabulary of classes included as subtype of GeologicProperty in NADM C1.

tion tables have to do with locations along observation tracks—boreholes, traverses, flightlines. The sample table catalogs physical specimens. Description tables are the base tables for standardized description of geologic objects, including EarthMaterialDescription, GeologicUnitDescription, GeologicAge, and GeologicStructureDescription. Relationship tables are correlation tables used to establish semantic relationships between other data instances, and include several specialized tables with different relationship properties.

Value specification tables

Value specification tables record individual property values specified by a numeric measurement, text, an image, or geometry. Table 4 lists the various value specification tables. These tables are the leaf nodes in description tree structures that specify the basic units of observation and description—numbers, text, pictures, locations. Vocabulary terms that specify property values may also be viewed as leaf nodes in description tree structures, but they are shared by many descriptions. Instances in the value specification tables are unique to some particular description or location context, and if that context object is removed from the database, the value specification become useless and should be removed as well.

GIS feature classes

Table 5 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. GeologicSurfaceTrace and GeologicUnitOutcrop are line and polygon feature classes whose locations represent observable geologic phenomena in or on the Earth. ProjectExtent is a simple polygon feature class used to define the area of interest for a project. By defining an area of interest, a spatial search can be done to locate existing data that may be of use for a project—for instance, which geologic units have been mapped in the area. AreaOfInterest, SectionLine, and Station are polygon, line, and point feature classes used to define extents associated with observations in the sense of GML Observation and Measurement (Cox et al., 2004). They represent features that are located based on where observations are made, and do not (inherently) represent the location of observable phenomena.

Section location tables

In a variety of situations, observations are located relative to a section line—for example, locations in a borehole are typically specified in length from the top of the hole. These types of locations are treated specially in the NGMDB implementation (Table 6). Although borehole, traverse, and flightline (all kinds of sections) might be considered feature classes, the actual geometry of a borehole can not reliably be represented by the 2-D geometry available in the ESRI geodatabase structure. Each type of section is related to a SectionLine feature (Table 5) that represents the projection of the 3-D section onto a map horizon (typically the Earth's surface). Coordinates of locations along a section line are not simply related, in general, to length along the projected line. Thus, each kind

Table 4. Value specification tables. These tables are further explained in the Appendix.

Table	Description
Extent	Table for specifying extents with a bounding box defined by latitude and longitude coordinates (in decimal degrees) and optional link to a spatial object. Provides mechanism for simple spatial searches in a non-GIS analysis environment.
DocumentLink	Table that contains file path information for locating auxiliary documents (especially images) associated with observations.
MeasuredQuantity	Table container for numerical specification of measured values with associated uncertainty, units, and measurement method. Type field distinguishes different semantics for DefaultValue, LowerBound, and UpperBound.
StructureObservation	Table for recording orientation measurements of geologic structures. It combines two measured quantity instances into one description record, with additional observation properties and a default symbol specification. Strike and dip orientation data are fundamental to geologic map information, and are represented in this physical table to simplify usage. The two measured quantities represent strike and dip or plunge and trend, depending on whether the orientation represents a planar or linear structure. Observation properties record classification confidence for identification of the measured structure and measurement procedure. A default symbol identifier is included to simplify quick display of the data.
TextDescription	Table for value specification using bodies of text.

Table 5. Location specification tables (ESRI geodatabase feature classes). These tables are further explained in the Appendix.

Table	Description
GeologicSurfaceTrace	Line features that represent the intersection of geologic surfaces with the map horizon.
GeologicUnitOutcrop	Polygons representing the intersection of a geologic unit with the map horizon.
ProjectExtent	Polygons that specify the area of interest for a project.
AreaOfInterest	Polygons that are associated with one or more observations.
SectionLine	Line that is the projection of a 3-D section line (e.g., borehole, flight line) into a map horizon to provide a 2-D map representation of the section. For a section line that is in the map horizon, as is typical of a measured section or traverse line, the SectionLine is the mapped trace of the section.
Station	Point location at which one or more observations are made, or samples are collected.

Table 6. Section location tables. These tables are further explained in the Appendix.

Table	Description
SectionInterval	Spatial extent located relative to section origin along the section line by a top and bottom coordinate. Ideally represents the intersection of some volume with a section.
SectionIntercept	Spatial extent, represented by a single coordinate that is located relative to section origin along the section line. Ideally represents the intersection (intercept) of a geologic surface with a section.
Borehole	Entity represents a borehole that is the result of a single drilling event. Not represented as a geodatabase feature class because the geometry is not directly represented in the currently-available, two-dimensional GIS.
FlightLine	Entity represents the course of an airborne (or waterborne) sensor.
Traverse	Entity represents the path followed by an observer on the Earth's surface.

of section includes a property that specifies the origin and metric for the coordinate reference system used to specify intervals and intercepts in that section. For instance, in a borehole the coordinate system typically is measured in linear length units downward from the ground surface or kelly bushing. In a measured stratigraphic section, the metric is thickness of strata traversed from the base of the section. A section interval is a location specification based on a start and end coordinate along a section line using the reference system defined for that section line. Section intercepts are points located by a single coordinate along a section. To convert SectionInterval and SectionIntercept locations to a true three-dimensional location, the 3-D geometry of the section must be known. For example, knowing that a sample is from 10,205 feet down in a bore hole does not fully locate the sample unless the geometry of the bore hole is known—if the hole is gently inclined, the surface projection of the sample location may fall at a significant distance from the borehole collar location.

Sample table

The sample table contains data instances that represent particular, identifiable masses of material. In this sense, they are similar to geologic units (as defined by NADM C1). The difference is that a geologic unit represents a mappable body of material—its location in the Earth is part of its identity, whereas a sample is from some location, but its identity is based on the collector's act of identifying that material by writing a number on it or putting it in a container. Many sorts of analytical data (e.g. chemical analyses, isotopic age dates) are associated with particular samples.

Description tables

Description tables (Table 7) contain data instances that are the base instances for complex descriptions of geologic units, Earth materials, structures and geologic age interpretations. Each description table listed in Table 7 includes a collection of attributes common to all of the description tables, along with attributes that are common to all descriptions of the particular type represented by that table.

Attributes common to all description tables specify the purpose and context (spatial and non-spatial) of the

Table 7. Description Tables. These tables are further explained in the Appendix.

Table	Description
GenericDescription	Physical base table that implements abstract description class as a physical table. This is a convenience for the ESRI CASE tool, so subtype integers are defined over generic descriptions only, and subtypes of other description types with other physical base tables (EarthMaterialDescription, GeologicAge(?), GeologicStructureDescription, and GeologicUnitDescription) may have their own ESRI subtype domains. Instances in this table identify descriptions of various types, identified in Geodatabase by ESRISubType, and whose attributes are defined by EntityProperty correlations for the entity (specified by DescriptionEntityGUID) associated with the description.
GeologicAge	Base table for geologic age description. Different specification details may be used through AttributeRelationship links based on the type property. Derived classes (identified by ESRISubtype / DescriptionEntityGUID) represent age specification in different ways: time instant (a number Ma before present, which may be inferred from 1 to many isotopic date measurements...), a named era (geologic time scale--e.g., Miocene), or a range specified by lower and upper bounds that may be instants, named eras, or geologic events. These different specifications are unified in this table with a best guess numerical minimum and maximum time coordinate (for analysis), and a DisplayName that summarizes the interpretation (for a data browser).
EarthMaterialDescription	Base table for compound Earth Material description. Identifies description instances with a GUID and a display name, and provides values for specifying properties common to all compound Earth Materials. Other description attributes are linked through AttributeRelationship instances. Includes all fields from GenericDescription (above).
GeologicStructureDescription	Base table for description of geologic structures. The attributes and subtypes for these descriptions are not yet fully populated. This table includes only properties common to all geologic structures. Includes all fields from GenericDescription (above).
GeologicUnitDescription	Geologic unit description object that specifies properties common to all geologic units; ESRI subtypes are used to apply rules for specific kinds of geologic units that have different combinations of properties and property value domains. Includes all fields from GenericDescription (above).

description, the described concept, and the structure of the description (see GenericDescription in Appendix Table 33). The description purpose attribute (DescriptionPurposeTermGUID) makes the intended function of a description explicit, e.g., default description, necessary property description, identifying property description, or instance description. The context specifies the domain within which the description is valid for the stated purpose. This domain may be spatial—some particular region of the Earth, or it may be human—e.g., a particular project or person, some organization, or some published authority (e.g. the Glossary of Geology…). These properties are included to solve the problem of distinguishing normative and instance descriptions, by recognizing that the distinction is always context and purpose dependent.

The link from a description to a 'described concept' (ConceptTermGUID) identifies the most specific term from an associated vocabulary that is consistent with the attributes specified in the description. For instance, an Earth material description of an indurated material of igneous origin, composed of 30% each of quartz, K-feldspar, and plagioclase would be associated the term 'Granite' from the standard lithology vocabulary. For identifying, default, or necessary property descriptions, this term will name the concept defined by the description. In this case, the vocabulary term may serve as a proxy for the attributes specified by the description. Any particular instance description need only specify attributes that are explicitly observed; other attributes may be inherited from the default description (if there is one) for the described concept. The structure of the description is defined by association with an entity definition in the Data Schema tables as described above. This entity definition may be used to validate the description before committing to the database, and for configuring the user interface during data entry or querying.

The GenericDescription table includes only the basic description property fields described above. This table serves as a base physical table for descriptions that do not have required attributes, or are deemed to not need a separate physical table. In the current implementation, the GenericDescription table is used for description of bedding fabric, genesis (geologic history), geologic event, particle shape, and particle size.

The GeologicAge table provides a mechanism for geologic age specification in as much or little detail as necessary to the user. Each GeologicAge instance represents an interpretation of one or more observations/measurements, and may be used to locate one or more other data objects (geologic units, structures, geologic events) in time. These different observations are unified in this table with a best guess numerical minimum and maximum time coordinate (for analysis), and a DisplayName that summarizes the interpretation (for a data browser). The individual observations are linked to the GeologicAge base table instance through AttributeRelationship instances, and may include time coordinates (age dates), stratigraphic eras that represent intervals (time ordinal eras in GML terminology), or individual geologic events (also from a vocabulary). Figure 4 schematically shows the data instances in various tables involved in a relatively detailed geologic age specification.

Description tables for compound Earth materials (rocks and unconsolidated materials), geologic units, and geologic structures include all the basic description attributes (purpose, context, described concept…), as well as a small set of attributes required in all descriptions of each kind. For compound Earth material, all descriptions must specify a consolidation degree, degree of crystallinity (crystalline vs. granular), grain discernibility, and representative size. Geologic unit descriptions have an age attribute. Geologic structure descriptions must specify pervasiveness, geometric aspect, and characteristic dimension. For details see the field descriptions in the Appendix.

Relationship tables

Relationship tables (Table 8) may be grouped into two types. AttributeRelationship, FractionalPartRelationship, and ObservationRelationship record information

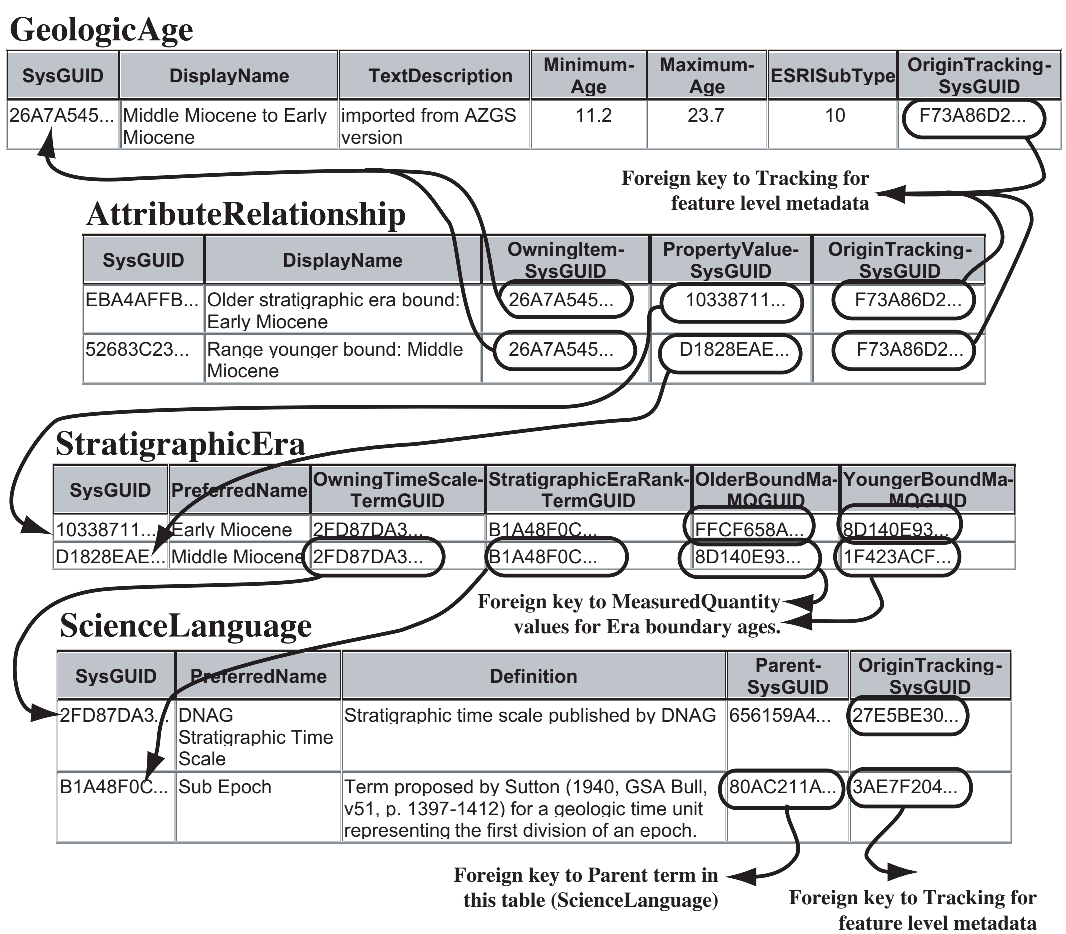


Figure 4. Example of geologic age instance showing foreign key relationships between tables involved in age specification.

Table 8. Relationship Tables. These tables are further explained in the Appendix.

Table	Description
AttributeRelationship	Represents observation/quantification/specification of the value of some property that is part of a description. Nature of value type is specified by Domain attribute of Entity-Property instance associated with the AttributeRelationship instance.
FractionalPartRelationship	Correlation that aggregates parts into a whole to represent the parts explosion (partonomy) for EarthMaterial and GeologicUnit. Includes attributes to specify proportion property as average with bounds or a range value.
ObservationRelationship	Correlation table to establish ‘science’ relationships between objects; the related objects have a lifetime independent of the observation relationship instance. A relationship type attribute specifies the semantics of each relationship instance.
SimpleRelationship	Generic correlation table, in which correlations are asserted, not observations, and have no metadata besides tracking.
MetadataRelationship	Simple relationship between metadata instances; use for associating citations with tracking records, person-organization tuples with activities, etc.

pertaining to observation and measurement of phenomena in the world, and include attributes for specifying metadata pertaining to the relationship instance. SimpleRelationship and MetadataRelationship are correlation tables that implement asserted data cardinality connections between data instances. For more information on these correlation tables, see the field descriptions in the Appendix.

The AttributeRelationship table contains data instances that link property values to a description. Property values are specified through links to a science language term, a value specification instance, or to another description instance (Figure 5). The value type and allowed values are specified in the data schema tables (see Data Schema section, above) by the Domain attribute of the EntityProperty associated with each AttributeRelationship instance. Each EntityProperty instance that may be referenced by AttributeRelationship has an associated ESRISubtype attribute integer value that is used in the ESRI geodatabase environment to specify the entity (i.e., geodatabase object class and subclass) that contains the value instance, and to specify the domain of possible values for that EntityProperty. The ESRI geodatabase domains are generated from the domain definition tables either during geodatabase setup, or dynamically with customized geodatabase behavior. Attribute relationship links are owned by a description instance in that if the description is deleted, the associated attribute links are deleted.

The FractionalPartRelationship table is used in geologic unit description and compound material description to represent compositions. The table includes a measured quantity representation (typical value, minimum, maximum, measurement method, etc.) for recording the proportion to the whole of a given part of the aggregation. Each part instance also has part type and role attributes. “Type” specifies the nature of each part. For example, in an Earth material, the mineral constituents may occur as ‘clast’, ‘fossil’ or ‘crystal’. Role specifies the relationship between one part and the aggregation as a whole, for example a mineral constituent may be of type ‘crystal’, and have a role that is either ‘phenocryst’ or ‘groundmass’. Classification of a part type is (at least conceptually) possible if the part is removed from the aggregation, whereas roles are dependent on the aggregated state of the compound material.

The ObservationRelationship table contains data instances that record observed or inferred relationships between geologic phenomena. A relationship type attribute specifies the semantics of the relationship. The ESRISubtype attribute is used in the geodatabase environment to constrain valid source and target entities for each relationship instance.

Metadata tables

Feature-level metadata is recorded principally in the Tracking table. Each tracking record has a display name, a free text description, a link to an Activity, a processing method description (similar to ‘processing steps’ in FGDC metadata), and for information derived from publications, links to one or more citations to published literature (see implementation described in Richard [2003]). Table 9 summarizes the various database tables used to implement feature-level metadata. Activity is a description that specifies one or more people involved in the work, each associated with an organization and a sponsoring project. Every data object has a link to a Tracking record that records where the data object came from (known as “origin tracking”).

Rules for the use of tracking records depend on business requirements. Given the long-term objective of a distributed and seamless database with information from a variety of sources, for both scientific and legal reasons, it seems necessary to at least be able to trace the

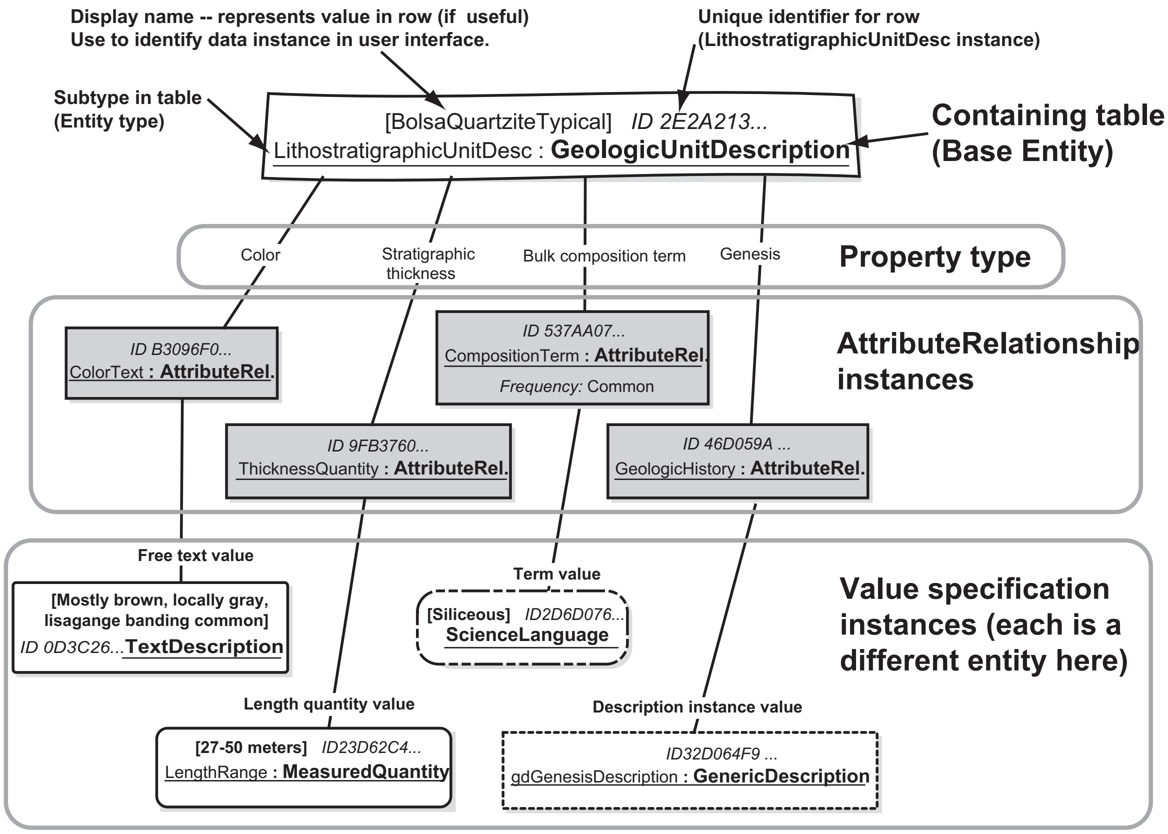


Figure 5. Association of property values to a geologic unit description using AttributeRelationship correlation links. Each box represents a row (data instance) in a table. Different shapes and line patterns indicate different tables. Identifiers for each row are 32 hexadecimal digit globally unique identifiers. These are abbreviated to 'ID nnnnn...' in the boxes. Table names are underlined. Some tables have 'subtypes' identified here by a name followed by a colon before the table name. These are subsets of rows in the table that have different value ranges defines for some fields, and may have different collections of attributes associated through AttributeRelationship links. Lines between boxes represent foreign key relationships between rows. AttributeRelationship instances (abbreviated AttributeRel.) are linked to a description through their 'owningItem' foreign key, and to a value specifier through their 'PropertyValue' foreign key. Property values may be specified by TextDescription, MeasuredQuantity, ScienceLanguage, GeologicAge or GenericDescription data instances.

Table 9. Metadata Tables.

Table	Description
Activity	Specification of involvement of a Person-Organization instance in some aspect of a Project, during some time interval.
Citation	Information for specifying a published source of information.
Organization	An administrative entity that involves one or more people, and has some physical location.
Person	Specification of an individual person.
PersonOrganization	Correlation table that records association of some person as an employee (or volunteer with some organization) during some period of time; represents institutional affiliation.
Project	Represents a planned undertaking by one or more persons, typically with funding from some organization, with a stated objective and time frame. A project can involve one or more activities.
Tracking	Specification of the intellectual source of data, and any processing history involved in automating it in the information system. Includes link to an activity (person, organization, project (as used in this database)), relevant citations, and a text description of data processing.

origin of any declarative data to the original publication or individual responsible for the scientific observation or interpretation.

PHYSICAL DATABASE

The physical database for the thematic (non-spatial) tables currently being used for NGMDB software tool development has been implemented both in SQL server 2000 on a Windows 2003 server and a stand-alone Microsoft Access 2000 database. An ESRI Geodatabase version (which includes the feature classes and internal database tables required to function as an ESRI Geodatabase) is being tested using a personal geodatabase (Microsoft Access 2000 file) generated from a Visio UML document using the ESRI CASE tools (ESRI, 2002). When records are inserted in the thematic tables (e.g., when a new map area is delineated, its extent modified, or attributes added), the SQL server version calls a user function to generate a new unique identifier (see below), whereas the personal geodatabase uses a custom class extension to generate new identifiers. Each table in the database includes a collection of standard ‘system’ fields, which are summarized in Table 10.

This design has evolved in a major step from related predecessor databases (Richard, 2003) by adopting globally unique identifiers (GUIDs) as the relational database key. These are 32 hexadecimal-digit (128 bit) numbers generated by the operating system (available on all major operating systems), guaranteed (or at least highly probable) to be globally unique (Leach and Salz, 1998). Use of these identifiers simplifies generation of unique keys for database relationships in a distributed environment. Because the information system is intended for use in a GIS environment, and because a majority of GIS systems use ESRI software, compatibility with ESRI data formats is considered essential. Standard GUIDs are binary numbers, and are incompatible with ESRI coverages used through Arc/Info version 7, version 8 Geodatabases, and ESRI shape files (which use dBase table format for thematic data). In order to maintain backward compatibility with these common data formats, GUIDs are converted to strings, formatted with hyphens according to a commonly used format (e.g. DA1AB9C6-A5D3-41DA-B3E2-66303CF231B2, hyphens after digit 8, 12, 16, and 20) to produce a 36-character string. These long strings are inefficient as keys, and in a large database would cause a performance problem, but it is anticipated that with release of ArcGIS v.9, the system will migrate to binary GUIDs for enterprise implementation, with the string GUIDs reserved for export to legacy systems and data interchange.

ENTERPRISE DATA MANAGEMENT

Because the National Geologic Map Database will be distributed in nodes maintained by various state geological surveys and the U.S. Geological Survey, some mechanism is necessary to manage data in the various nodes. The entire system is envisioned as a hierarchy of repositories (Figure 6). Each repository would be a self-contained collection of data, some of which is local to the

Table 10. Fields included in all tables.

Field	Description
SysGUID	Text, GUID (128 bit, globally unique number) converted to 32 hexadecimal digit string, with hyphens after digits 8, 12,16, and 20. Unique identifier for all data instances (rows in tables).
DisplayName	Text, identifies data instance for user in interface (renamed to PreferredName in Vocabulary tables).
TextDescription	Text, available for any comments, description, notes that user wishes to insert. Not analyzable (renamed Definition in Vocabulary tables).
OriginTrackingSysGUID	Foreign key to tracking record that records information on intellectual source of data, and data processing related to inclusion in database.
SysCreated	Date/time; this automatically-inserted value records data and time that data instance was created.
SysCreatedBy	Text, login name of user when data instance was created.
SysUpdated	Date/time; automatically inserted value records data and the time that the data instance was most recently updated.
SysUpdatedBy	Text, login name of user when data instance was most recently updated.
SysOwningRepositoryGUID	Foreign key to SysRepository table (not included in this model...) that associates each data object with its owning repository; repository designates data ownership, publication level/authority (e.g., individual, project, AZGS, NGMDB...).

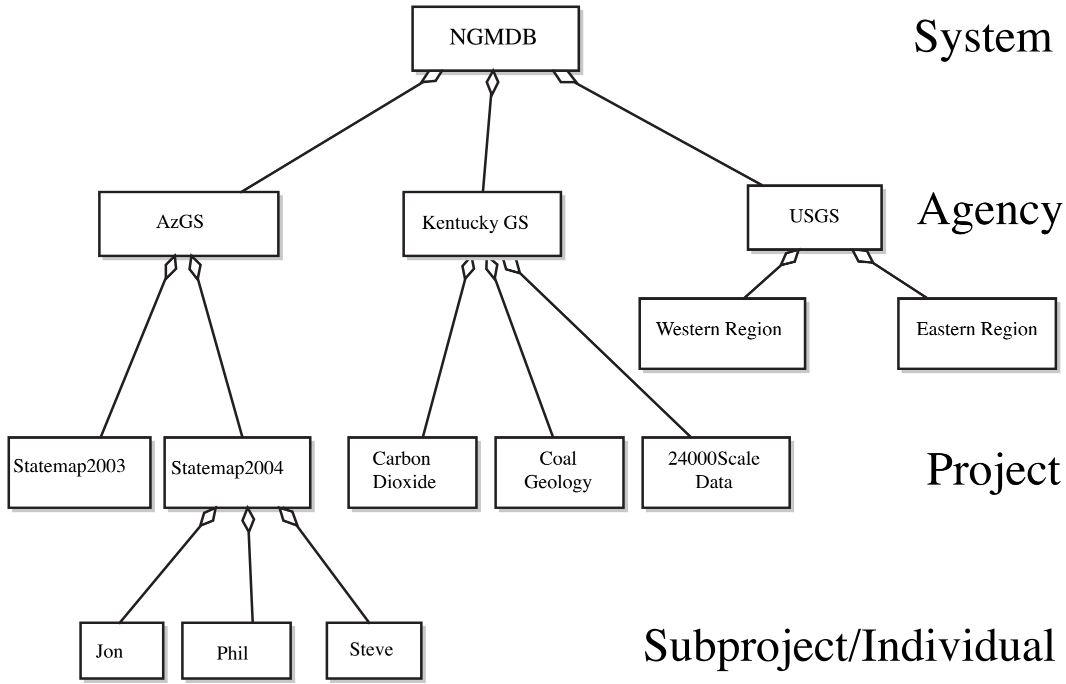


Figure 6. Schematic repository hierarchy. Higher-level hierarchies aggregate data from lower level hierarchies.

repository, and some of which is 'inherited' from parent repositories. Each repository would include the collection of tables outlined here, and possibly additional tables defined in the data schema tables for the repository.

Figure 7 is an object diagram for the proposed repository structure. A Repository is a data store composed of a TableCollection that aggregates tables (with associated domains, relationships, and constraints) defined by a standard NGMDB data schema and tables defined by local schema extensions. Each repository is contained in a physical database artifact, which is typically a file in a computer system. A repository is associated with one or more projects that use data contained in the repository. Security policies that control data entry and editing permissions are associated with projects and repositories. The repository data schema is an aggregation of schema elements from the standard NGMDB data schema (outlined in this paper), and local schema elements required for other business requirements. The tables (with associated domains, relationships, and constraints) that compose a repository realize the repository's data schema. A repository uses a science language vocabulary that is an aggregation of shared vocabulary terms from the NGMDB standard vocabularies, other standard vocabularies from the organization and project level, and locally defined vocabulary terms. The terms that are not included in the NGMDB standard vocabularies must be defined according to the process described in Richard et al. (2003).

Each repository will indicate some level of authority, and migration of data from a lower level repository to a higher-level repository will involve a publication process that includes scientific and logical review and approval. Each repository will have an owner who determines policies and procedures for inclusion of data in that repository. Access by individual users, projects, and repositories for reading, adding, and updating data are to be determined and controlled by the repository owner.

Self-contained bodies of data from one or more repositories may be 'published' as a snap shot, a read-only stand-alone dataset that may be transported into other database environments (along with the feature level tracking information for the data). Data from one repository may be linked into another repository (the owning repository is an attribute of all data instances), but if any updates are made, the repository in which the updates are made becomes the owning repository (i.e., it is responsible for the scientific content of the updated or modified information).

CONCLUSIONS

The data in an archive that is based on this design will almost certainly require some pre-processing for use in standard relational database systems with SQL-based queries. The data could equally undergo pre-processing for analysis by a description logic engine such as Racer (http://www.cse.concordia.ca/%7Ehaarslev/racer/). The implementation is essentially an implementation of a description logic (Baader et al., 2003) for science information. The description structure is assembled by links between data objects, and can be thought of as a directed

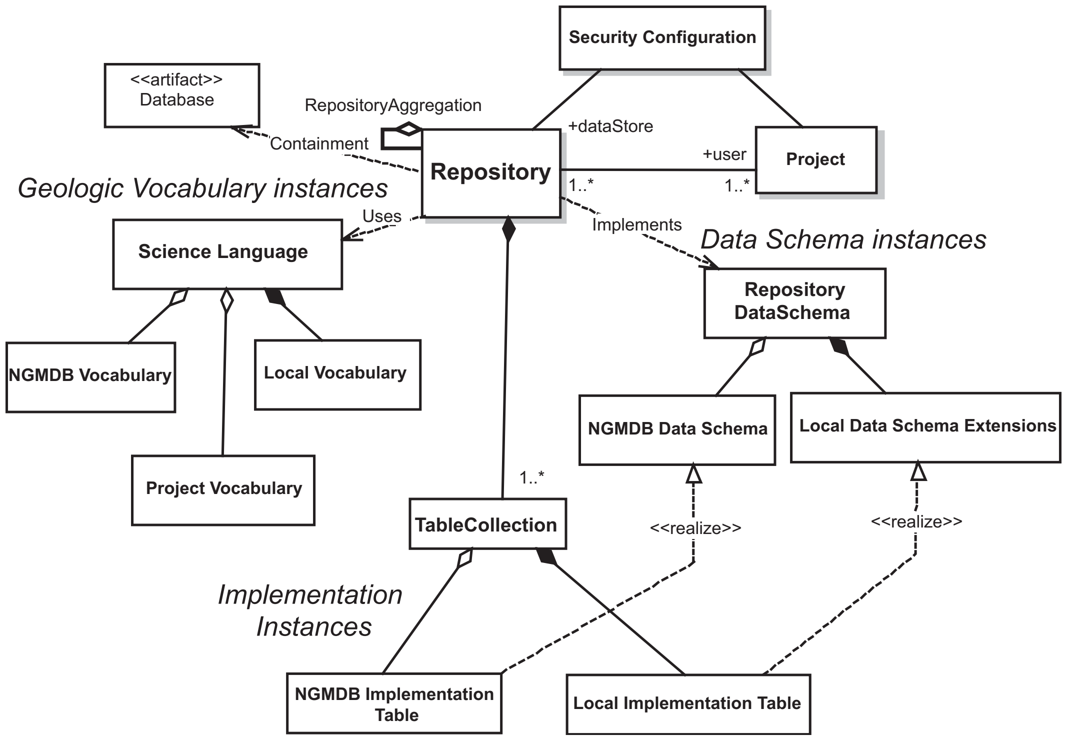


Figure 7. UML diagram for Repository. UML classes have dropshadow, Class instances and artifacts do not have drop shadow. Solid fill aggregation symbols indicate that the lifetime of the member objects is linked to the lifetime of the container object (Cascade delete).

acyclic graph, similar to an XML document (network data structure), so Xpath type search specification (Clark and DeRose, 1999) probably will also be useful. We are moving ahead with the implementation of a knowledge representation system that moves beyond the SQL-based relational database with the understanding that current technology must mature to fully reap the benefits of this approach. Given the inherently long lead time in implementing, testing, and finally populating the data in such a system, we are confident that the necessary analytical tools under development in the semantic web community will allow full utilization of the data for new and exciting applications of geoscience information.

REFERENCES

Baader, F., Calvanese, D., McGuinness, D., Nardi, D., Patel-Schneider, P., eds, 2003, The Description Logic Handbook—Theory, Implementation and Applications: Cambridge, UK, Cambridge University Press, 574 p.

Brodaric, B., Journeay, M., Talwar, S., and others, June 18, 1999, CordLink Digital Library Geologic Map Data Model Version 5.2 (Web Page), accessed June 13, 2001, at http://cordlink.gsc.nrcan.gc.ca/cordlink1/info_pages/English/dm52.pdf.

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 June 14, 2001, 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, 562 p.

Clark, James, DeRose, Steve, 1999, XML Path Language (XPath), Version 1.0, W3C Recommendation (16 November 1999), accessed November 18, 2004, at http://www.w3.org/TR/xpath.

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 November 28, 2004, at http://www.opengeospatial.org/specs/?page=recommendation.

ESRI, 2002, Building Geodatabases with CASE Tools: Redlands, CA, ESRI, 72 p., accessed November 29, 2004, at http://downloads.esri.com/support/downloads/ao_/CASE_Tools_CaseTools.pdf.

Johnson, B. R., Brodaric, B., and Raines, G. L., 1998, Digital Geologic Maps Data Model, V. 4.3 (Web Page), AASG/USGS Data Model Working Group report, accessed at http://www.nadm-geo.org//dmdt/.

Leach, P.J., and Salz, Rich, 1998, UUIDs and GUIDs, WebDAV Network Working Group Internet Draft, accessed

November 29, 2004, at http://www.webdav.org/specs/draft-leach-uuids-guids-01.txt, or at http://ftp.ics.uci.edu/pub/ietf/webdav/uuid-guid/draft-leach-uuids-guids-01.txt.

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 Report 2004-1334, 60 p., 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/.

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., Matti, Jonathan, Soller, D.R., 2003, Geoscience terminology development for the National Geologic Map Database, in Soller, D.R., ed., Digital Mapping Techniques '03—Workshop Proceedings: U.S. Geological Survey Open-File Report 03-471, p. 157-167, accessed at http://pubs.usgs.gov/of/2003/of03-471/richard1/.

Soller, D. R., and Berg, T. M., 2003, The National Geologic Map Database: 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., 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., accessed at http://pubs.usgs.gov/of/2002/of02-202/.

Struik, L.C., Quat, M.B., Davenport, P.H., and Okulich, A.V., 2002, A preliminary scheme for multihierarchical rock classification for use with thematic computer-based query systems: Geological Survey of Canada, Current Research 2002-D10, 11 p., accessed at http://www.nrcan.gc.ca/gsc/bookstore/free/cr_2002/D10.pdf.

APPENDIX

Tables 11-42, showing additional concepts, tables, and fields in the NGMDB implementation of NADM C1 (Tables 1-10 are contained in the text).

CONVENTIONS FOR FIELD NAMES IN NGMDB TABLES

- GUID—used as suffix to indicate a field contains a Globally Unique IDentifier, a 32 digit hexadecimal number calculated by the operating system based on an algorithm that produces theoretically unique values.
- Field names with the suffix SysGUID indicate fields that are foreign keys to the primary key in another entity.
- Field names with the suffix TermGUID are foreign keys to a vocabulary table, and identify a term used to specify a property value.
- FieldNames with the suffix EntityGUID are foreign keys to the Entity table, and identify an entity (physical or logical table). These are normally associated with a SysGUID foreign key that identifies a particular data object (record) in that entity (table).
- FieldNames with prefix ‘Is’ are boolean fields that may have a true or false value.

LIST OF TABLES

- Table 11. Fields in Entity table.
- Table 12. Fields in EntityProperty table.
- Table 13. Fields in EntityPropertyDomain table.
- Table 14. Fields in EntityPropertyDomainNodeList table.

Quantity Value Specification Tables (see also Table 4 in text):
- Table 15. MeasuredQuantity.
- Table 16. StructureObservation.
- Table 17. TextDescription.
- Table 18. DocumentLink.
- Table 19. Extent.

Spatial Data Tables (see also Table 5 in text):
- Table 20. NGMDBFeature.
- Table 21. ProjectExtent.
- Table 22. AreaOfInterest.
- Table 23. Station.
- Table 24. SectionLine.
- Table 25. BoreholeCollar.
- Table 26. GeologicSurfaceTrace.
- Table 27. GeologicUnitOutcrop

Section Location Tables (see also Table 6 in text):
- Table 28. Section.
- Table 29. Borehole.
- FlightLine.
- Traverse.
- Table 30. SectionInterval.
- Table 31. SectionIntercept.
- Table 32. Fields in Sample Table.

Description Tables (see also Table 7 in text):
- Table 33. GenericDescription.
- Table 34. GeologicAge.
- Table 35. EarthMaterialDescription.
- Table 36. GeologicUnitDescription.
- Table 37. GeologicStructureDescription.

Relationship Tables (see also Table 8 in text):
- Table 38. AttributeRelationship.
- Table 39. FractionalPartRelationship.
- Table 40. SimpleRelationship.
- Table 41. ObservationRelationship.
- Table 42. MetadataRelationship.

Table 11. Fields in Entity table.

Field	Description
EntityInternalName	Text string; immutable name for this entity, and should not be changed; this text string may be used to identify a physical table in software applications that use the database.
DisplayName	Text string; for identifying the entity to users in the GUI; it may be changed to suit the context.
Implementation	Text string; from a controlled vocabulary that specifies whether the description is implemented entirely as a physical table or as a base table with AttributeRelationship links for one or more properties. If the implementation is 'PhysicalTable', all properties are specified by the value in a field in the base physical table, and have cardinality 0 or 1.
IsSubClassOf	Foreign key to Entity table; if the Implementation value is 'LogicalTable', then this field is the sysGUID for the Entity that is the base Entity for the description; otherwise it is not used. The logical table includes all the properties (physical or logical) that are included with the base Entity, and one or more additional properties associated through AttributeRelationship links. Presently, the subclassing of Entity definitions is only allowed to be one level deep, that is any Entity with 'LogicalTable' implementation has a 'IsSubClassOf' link to an Entity with an Implementation value of 'PhysicalTable'.
ESRISubtype	Integer; used in ESRI geodatabase to identify different entities as 'subclasses' of the geodatabase object that is the base table for the entity.

Table 12. Fields in EntityProperty table.

Field	Description
EntitySysGUID	Foreign key to the Entity table; identifies the entity that includes the property specified by PropertyVocabulary-SysGUID.
PropertyVocabularySysGUID	Foreign key to the EntityPropertyVocabulary table; identifies the kind of property specified by the attribute.
EntityPropertyDomainSysGUID	Foreign key to EntityPropertyDomain table; identifies domain definition for this property in this entity.
DisplayName	String; name to identify this entity in the user interface.
UserInterfaceLabel	String; name to identify this property in this entity in the user interface, and will typically be a geoscientist-friendly term, which may be modified for use in different contexts.
Implementation	String; term from controlled vocabulary that specifies how the association of a property value with the entity is physically implemented; possible values are: 1) 'PhysicalField', if the property value is contained in a field in a physical table; 2) 'AttributeRelationship' if the property value is specified by a link through the AttributeRelationship table; or 3) 'PhysicalField_FK' if the property value is specified by a linked entity instance, in which case the property is a field in a physical table that contains a foreign key to the entity that contains the property value.

DataTypeName	String; term from controlled vocabulary identifying a standard data type (e.g., integer, float), using Microsoft SQL server data types.
DataLength	Integer; specifies the length of strings allowed for string or text data fields.
OrdinalPos	Integer; orders the listing of fields in the entity.
DefaultValue	String; supplies a default value to use for the property in this entity. For numeric fields, the string must be converted to a number in order to use.
Cardinality	String; specifies the number of values that may be associated with the property in this entity. For attributes implemented as 'PhysicalField' or 'PhysicalField_FK', the Cardinality is either '0..1' (optional) or '1..1' (mandatory). Attributes implemented as 'AttributeRelationship' will have cardinalities of '0..n' (optional), '1..n' (at least one required), and rarely some other value.
UsesFrequency	Boolean; if the value is true, then a value must be specified for the frequency property for each attribute value specification. Only applicable if the Implementation field contains 'AttributeRelationship.'
UsesIntensity	Boolean; if the value is true, then a value must be specified for the intensity property for each attribute value specification. Only applicable if the Implementation field contains 'AttributeRelationship.'
UsesConfidence	Boolean; if the value is true, then a value must be specified for the confidence property for each attribute value specification. Only applicable if the Implementation field contains 'AttributeRelationship.'
UsesMeasureProcedure	Boolean; if the value is true, then a value must be specified for the measurement procedure property for each attribute value specification. Only applicable if the Implementation field contains 'AttributeRelationship.'
UsesEvidence	Boolean; if the value is true, then a value must be specified for the evidence property for each attribute value specification. Only applicable if the Implementation field contains 'AttributeRelationship.'
ESRISubType	Used for integrating with geodatabase, and is the integer subtype value for AttributeRelationship instances used to specify values for this property in this entity. ESRISubType is only specified if Implementation is 'AttributeRelationship'

Table 13. Fields in EntityPropertyDomain table.

Field	**Description**
IsTreeStructured	Boolean; true if the node list is hierarchical. If true, then ParentSysGUID values in Nodes in this list define links to build tree hierarchy.
VocabularyNodeListSysGUID	Foreign Key to PropertyDomainSysGUID in EntityPropertyDomainNodeList; identifies terms to include in the vocabulary (pick list) defined by this domain.
DataTypeTermGUID	Foreign key to ScienceLanguage; specifies type of data used to specify property values for a particular entity-property combination. This value also distinguishes fields that have domains defined by the domain node list table from those that are simple foreign keys (i.e., whose domain are all the rows in the target table for the foreign key).
SimpleFKEntityGUID	Foreign key to Entity table; identifies entity if property domain is simply any instance of that entity; saves having to use the domain node list table.
MinValue	Float; if data type (specified by DataTypeTermGUID) is numeric, this value assigns the smallest valid value that may populate this field.
MaxValue	Float; if data type (specified by DataTypeTermGUID) is numeric, this value assigns the largest valid value that may populate this field
OwnerSysGUID	Foreign key to entity identified by OwnerEntityGUID; identifies owner of the domain, may be a person, project, activity, or organization. Allows context-dependent selection of appropriate domain.
OwnerEntityGUID	See above

Table 14. Fields in EntityPropertyDomainNodeList table.

Field	Description
ParentSysGUID	Foreign key to a parent node in the EntityPropertyDomainNodeList table; if IsTreeStructured is true, this field is used to define a tree hierarchy specific to the particular pick list (domain node list).
PropertyDomainSysGUID	Foreign key to EntityPropertyDomain; identifies for a particular domain, has same value for all nodes included in the domain.
SelectedVocabularyNodeSysGUID	Foreign key to entity identified by VocabularyNodeEntityGUID; identifies a data instance that is a member of the domain.
VocabularyNodeEntityGUID	See above
JustSelectedNode	Boolean; true if only the selected node is included in the domain; if false, the selected node and any child nodes in the source entity are included in the domain. If entity identified by VocabularyNodeEntityGUID does not have a ParentSysGUID field, the value is assumed to be true.
ExcludeSelectedNode	Boolean; if true then selected node is not included in the domain. Only useful for excluding particular nodes in a hierarchy identified by a parent node (for which JustSelectedNode is false).
IsDeprecated	Boolean; if true the domain value has been abandoned and is only included for backward compatibility.

Quantity Value Specification Tables (see also Table 4 in text):

Table 15. MeasuredQuantity. Table container for numerical specification of measured values with associated uncertainty, units, and measurement method. Type field distinguishes different semantics for DefaultValue, LowerBound, and UpperBound.

Field	Description
DefaultValue	Floating point number; single value that best represents the measured quantity, for use in analyses where a single value is required; determination of value is based on quantity type.
LowerBound	Floating point number; lower numerical bound for measured value, may be limit of uncertainty envelope or lower bound of value range.
UpperBound	Floating point number; upper numerical bound for measured value, may be limit of uncertainty envelope or upper bound of value range.
UnitsTermGUID	Foreign key to ScienceLanguage; identifies the unit of measurement.
ValueTypeTermGUID	Foreign key to ScienceLanguage; distinguishes quantities specified by value range, average value with symmetric uncertainty, value with asymmetric uncertainty, etc.
MeasurementMethodTermGUID	Foreign key to ScienceLanguage; specifies a measurement method; in long run may want this to be to a text description, or measurement method entity....
QuantityEntityGUID	Foreign key to entity table; specifies quantity type, e.g. length measurement, age measurement, mass measurement… Used to determine attribute domains for this value specification. Redundant with ESRI subtype, but included for consistency in data structure.
ESRISubtype	Integer; differentiates different domain subsets—e.g., age quantities, length quantities, etc.

Table 16. StructureObservation. Description table for recording orientation measurements of geologic structures. It combines two measured quantity instances into one description record, with additional observation properties and a default symbol specification.

Field	Description
StructureTypeTermGUID	Foreign key to science language, identifies the kind of structure whose orientation is described.
LocationSysGUID	Foreign key to entity specified by LocationEntityGUID; specifies the location of the measurement, typically a Station or SectionIntercept, but may be to OutcropTrace, AreaOfInterest or to SectionInterval to indicate that the measurement applies over some extended region. Associated LocationEntityGUID identifies the entity that contains the data instance identified by LocationSysGUID.
LocationEntityGUID	Foreign key to Entity table; identifies the entity that contains the data instance identified by LocationSysGUID.
Azimuth	Floating point number; default or single representative value for strike of planar feature, bearing of linear feature.
MaximumAzimuth	Floating point number; upper bound of azimuth value range or uncertainty envelope.
MinimumAzimuth	Floating point number; lower bound of azimuth value range or uncertainty envelope.
AzimuthMeasuredQuantityTypeTermGUID	Foreign key to ScienceLanguage; term distinguishes quantities specified by a value range, an average value with symmetric uncertainty, a value with asymmetric uncertainty, etc.
DipPlunge	Floating point number; default or single representative value for dip of planar feature, plunge of linear feature.
MaximumDipPlunge	Floating point number; upper bound of dip or plunge value range or uncertainty envelope.
MinimumDipPlunge	Floating point number; lower bound of dip or plunge value range or uncertainty envelope.
DipPlungeMeasuredQuantityTypeTermGUID	Foreign key to ScienceLanguage; term distinguishes quantities specified by a value range, an average value with symmetric uncertainty, a value with asymmetric uncertainty, etc.
MeasurementProcedureTermGUID	Foreign key to ScienceLanguage; term specifies the procedure for determining orientation (Brunton compass measurement on outcrop, three point determination, estimate from distance, air photo interpretation…); in long run may want this to be to a text description, or measurement method entity....
IdentificationConfidenceTermGUID	Foreign key to ScienceLanguage; term specifies the observer's confidence that the measured structure is actually is the phenomenon identified by StructureTypeTermGUID.
LabelText	String; default text to label symbol if this orientation measurement is displayed on a map.
CartoObjID	Integer; identifier for a symbol in the symbol set specified by SymbolSet used for default depiction of this spatial object.
SymbolSet	String; name of a collection of pre-defined symbols used for default depiction of this spatial object.

Table 17. TextDescription. Table for value specification using bodies of text.

Field	Description
SubjectTermGUID	Foreign key to ScienceLanguage; term classifies subject matter of text.
TextHeading	String; user-defined headings for classifying text fragments.
ContextSysGUID	Foreign key to entity specified by associated ContextEntityGUID; identifies data instance to which this description applies/belongs. Context will typically be some description instance; may also be a spatial object (e.g., station for field note TextDescription instances).
ContextEntityGUID	Foreign key to entity table; identifies entity that contains data identified by ContextSysGUID.

Table 18. DocumentLink. Table that contains file path information for locating documents associated with observations.

Field	Description
DocumentPathSpec	String; URL path for location of document file.
DocumentFileName	String; name of document file.
FileTypeTermGUID	Foreign key to ScienceLanguage; term specifies the type of document from a controlled vocabulary of document types, e.g. Tiff, Jpeg…
ContentTypeTermGUID	Foreign key to ScienceLanguage; term specifies the kind of content in the document (e.g., image, text, spreadsheet, vector graphics…).
DocumentDate	Date/Time; specifies the date the document originated or is current to.

Table 19. Extent. Table for specifying extents with a bounding box defined by latitude and longitude coordinates (in decimal degrees) and optional link to a spatial object. Provides mechanism for simple spatial searches in a non-GIS analysis environment.

Field	Description
LatMin	Floating point number; minimum latitude in decimal degrees of bounding box.
LongMin	Floating point number; minimum longitude in decimal degrees of bounding box. Use longitude west in western hemisphere.
LatMax	Floating point number; maximum latitude in decimal degrees of bounding box.
LongMax	Floating point number; maximum longitude in decimal degrees of bounding box. Use longitude west in western hemisphere.
SpatialReference	String; specification of datum and spheroid to which latitude and longitude values are referenced.
SpatialObjectSysGUID/ SpatialObjectEntityGUID	Compound foreign key to entity specified by SpatialObjectEntityGUID; specifies detailed geometry of the extent. SysGUID identifies a data instance in the entity identified by EntityGUID.

Spatial Data Tables (see also Table 5 in text):

Table 20. ***NGMDBFeature*****.** Abstract superclass for data instances that specify a geographic location. All spatial data tables (feature classes) include these fields. Location is a property that may be associated with numerous other kinds of data.

Field	Description
Label	String; default text used to label a feature in a map visualization.
DisplayName	String; for identifying the feature in text-based lists.
CartoObjID	Integer; identifier for a symbol in the symbol set specified by SymbolSet used for default depiction of this spatial object.
SymbolSet	String; name of a collection of pre-defined symbols used for default depiction of this spatial object.

Table 21. ProjectExtent. Polygons that specify the area of interest for a project. Fields are the same as NGMDBFeature.

Field	Description
ProjectSysGUID	Foreign key to project table; identifies project associated with the project extent.

Table 22. AreaOfInterest. Polygons associated with a description; may be symbolized, as in some local phenomena overprinting rock (contact aureole...), or simply represent the area to which some geologic unit description applies. Use to delineate contexts for descriptions that do not correspond to mapped outcrop polygons—for instance superimposed alteration, facies variations, phases in a pluton that are not geologic map units, area represented by a grab sample, area over which an orientation measurement applies.

Field	Description
MapHorizon	String; specifies the surface within or on the Earth that contains the area depicted by this spatial object.
DepictionScale	Integer; scale at which the spatial object was originally delineated.

Table 23. Station. Point location at which one or more observations are made or samples collected.

Field	Description
LocationDateTime	Date/Time; specifies the date and time at which the station was first occupied.
Elevation	Floating point number; surface elevation at station.
ElevationUnitsTermGUID	Foreign key to ScienceLanguage; term specifies units of measurement for elevation; default is meters.
UTME	Floating point number; UTM easting coordinate for station location.
UTMN	Floating point number; UTM northing coordinate for station location.
UTMZone	String; specifies the UTM zone and datum (e.g., xxN_NAD27 or xxS_NAD83).
PositionUncertainty	Floating point number; radius (in meters) of circle of confidence around located point; ideally interpreted to mean 'actual location is within x meters of reported coordinate with 95% confidence.'

Table 24. SectionLine. Projection of a 3-D section, along which data have been collected, into a map horizon surface, typically the Earth surface.

Field	Description
DepictionScale	Integer; scale at which the spatial object was originally delineated.
MapHorizon	String; specifies the surface within or on the Earth that contains the area depicted by this spatial object.
LocationDateTime	Date/Time; specifies the date and time at which the associated section was first defined.

Table 25. BoreholeCollar. Points that represent the location in three-dimensional space at the top of the casing (or ground surface if there is none) for a borehole.

Field	Description
Elevation	Floating point number; surface elevation at borehole collar.
ElevationUnitsTermGUID	Foreign key to ScienceLanguage; term specifies units of measurement for elevation; default is meters.
UTME	Floating point number; UTM easting coordinate for collar location.
UTMN	Floating point number; UTM northing coordinate for collar location.
UTMZone	String; specifies the UTM zone and datum (xxN_NAD27 or xxS_NAD83).
PositionUncertainty	Floating point number; radius (in meters) of circle of confidence around located point; ideally interpreted to mean ‘actual location is within x meters of reported coordinate.’

Table 26. GeologicSurfaceTrace. Represents line features that represent the intersection of geologic surfaces with the map horizon.

Field	Description
MapHorizon	String; specifies the surface within or on the Earth that is intersected by the geologic surface of interest to produce the trace located by this spatial object.
DepictionScale	Integer; scale at which the spatial object was originally delineated.
ConceptTermGUID/ ConceptEntityGUID	Foreign key to entity specified by ConceptEntityGUID; term classifies the geologic surface that crops out along this trace. EntityGUID specifies the entity that contains the instance specified by the TermGUID value.
IdentityConfidenceTermGUID	Foreign key to ScienceLanguage; term specifies the observer’s confidence that the observed phenomenon actually is the phenomenon identified by ConceptTermGUID.
ExistenceConfidenceTermGUID	Foreign key to ScienceLanguage; term specifies the observer’s confidence that mapped feature actually exists.
PositionUncertainty	Floating point number; specifies the radius of the uncertainty envelope, in meters, within which the identified feature is asserted to be located, as depicted on the base map. Includes uncertainty derived from inability to precisely locate geologist on ground (e.g., surrounded by trees), and uncertainty in ability to transfer ground location to base map point. Ideally should be read to mean ‘actual location is within x meters of the reported (mapped) location with 95% confidence’.
LocatabilityTermGUID	Foreign key to ScienceLanguage; term specifies whether a feature is directly observable, may be observable (well exposed, continuously), inferred between sparse outcrop, inferred based on physiographic expression, inferred based on other evidence (vegetation change, soil change...) or concealed by overlying material.
ESRISubType	Integer; defines subsets of geologic surface traces with different semantics.

Table 27. GeologicUnitOutcrop. Polygons representing the intersection of a geologic unit with the map horizon.

Field	Description
MapHorizon	String; specifies the surface within or on the Earth that is intersected by the geologic surface of interest to produce the trace located by this spatial object.
DepictionScale	Integer; scale at which the spatial object was originally delineated.
ConceptTermGUID/ ConceptEntityGUID	Foreign key to entity specified by ConceptEntityGUID; term classifies geologic unit that crops out within this polygon. EntityGUID specifies the entity that contains the instance specified by the TermGUID value.
IdentityConfidenceTermGUID	Foreign key to ScienceLanguage; term specifies the observer's confidence that the observed phenomenon actually is the phenomenon identified by ConceptTermGUID.

Section Location Tables (see also Table 6 in text):

Table 28. Section. Abstract class represents XMML section (https://www.seegrid.csiro.au/twiki/bin/view/Xmml/SitesAndSpecimens), which is a one-dimensional extent (path in 3-space) along which observations are made. Has concrete subtypes (physical tables) Borehole, Traverse, and FlightLine. The subtype tables include all of these fields from this abstract parent table.

Field	Description
DisplayName	String; for identifying the feature in text-based lists.
TextDescription	Free text; description or notes pertaining to the observation section.
SectionLineSysGUID	Foreign key to SectionLine feature class; identifies line representing projection of section onto a map horizon (typically the Earth's surface).
TypeTermGUID	Foreign key to ScienceLanguage; term classifies the kind of section
CoordinateReferenceSystem	String; specification of coordinate reference system for determining locations along the section line. Includes specification of the origin location (e.g., Kelly bushing, collar, south end of traverse), and the measured parameter along the section (e.g. length from origin, feet of section above starting point...).

Table 29. Borehole. A borehole instance represents a borehole interval drilled during a single drilling episode. One or more boreholes may be associated with a single collar location, either due to reentering the hole, or to branching from the same collar. The origin from which the offsets are measured will often be a collar location at the surface where the drill rig was located, but where the hole of interest splays from a primary hole, it may be an underground location. Note that in this latter case the position of the origin may again be best expressed as a 1-D coordinate relative to the coordinate reference system defined by the axis of the parent hole.

Field	Description
CollarSysGUID	Foreign key to BoreholeCollar table; identifies the location at which the borehole intersects some map horizon.
ShapeSysGUID	Foreign key to value specification table; specifies the full 3-D geometry of the section. (not currently used, and entities for 3-D geometry not yet defined…).
DrillingProcedure	String; specifies drilling procedure used.
Operator	String; name of operator responsible for borehole.
KBElevation	Floating point number; kelly bushing elevation if different from collar ground level elevation.
TotalDepth	Floating point number; total length of borehole in coordinate reference system defined for the hole.

LengthUnitsTermGUID	Foreign key to ScienceLanguage; term specifies the length units used to express total depth and kelly bushing elevation.
DrillStartDate	Date/Time; date drilling started.
DrillEndData	Date/Time; date borehole was completed.
Permit	String; state (or other local) drilling permit identification number.
APInumber	String; American Petroleum Institute identifying number for well.

FlightLine. Section line that represents the path of an aerial survey projected vertically onto a map horizon. Fields the same as Section**.**

Traverse. Line followed by a geologist on some map horizon, along which data are collected. Measured sections are implemented as a traverse. Fields the same as Section.

Table 30. SectionInterval. Interval located relative to section geometry by start and end coordinate values, for example length along a bore hole measured from the origin (collar, kelly bushing...) as defined in the associated section object. SectionInterval represents the intersection of a volume with a section geometry.

Field	Description
DisplayName	String; identifies the feature in text-based lists.
TextDescription	Free text; description or notes pertaining to the observation section.
SectionSysGUID/ SectionEntityGUID	Foreign key to section table specified by SectionEntityGUID; identifies the section in which the intercept is located.
IntervalTypeTermGUID	Foreign key to ScienceLanguage; term classifies the kind of interval.
MeasuredLength	Floating point number; derived value BottomDepth – StartDepth.
StartCoordinate	Floating point number; length in coordinate reference system for section to beginning of this segment.
EndCoordinate	Floating point number; length in coordinate reference system for section to end of this segment.
LengthUnitsTermGUID	Foreign key to ScienceLanguage; term specifies length units used. A single length units specification applies to segment length and top and bottom coordinates.
Direction	String; specifies the 3-D orientation of the interval considered as a straight line segment from start point to end point (e.g., trend/plunge); syntax not currently formalized.

Table 31. SectionIntercept. Location specified by a coordinate relative to section origin. Represents a point along a section, typically defined by the intersection of a geologic surface with the section, e.g., a formation top.

Field	Description
DisplayName	String; identifies the feature for users in text-based lists.
TextDescription	Free text; description or notes pertaining to the observation section.
SectionSysGUID/ SectionEntityGUID	Foreign key to section table identified by SectionEntityGUID; specifies the section in which the intercept is located.
LocationCoordinate	Floating point number; locates intercept with respect to section geometry based on coordinate reference system for section.
PositionUncertainty	Floating point number; ideally 95% confidence interval on each side of Location-Coordinate.

LocatabilityTermGUID	Foreign key to ScienceLanguage; term specifies whether feature is directly observable in core, may be observable (well exposed, continuously), inferred based on well logs, or concealed by gaps in data.
ConceptTermGUID	Foreign key to ScienceLanguage; term classifies the phenomenon located at the intercept.
IdentityConfidenceTermGUID	Foreign key to ScienceLanguage; term specifies the observer's confidence that the observed phenomenon actually is the phenomenon identified by ConceptTermGUID.

Table 32. Fields in Sample Table. Table for identifying material samples representative of some location in the Earth. Samples may be located by a Station, SectionIntercept, SectionInterval, or AreaOfInterest, depending on how they are collected.

Field	**Description**
FieldID	String; identifier assigned to sample by original collector.
RockName	String; name given to rock type by original collector.
GeologicUnitTermGUID	Foreign key to GeologicUnit; term identifies geologic unit from which specimen was collected.
StandardLithologyTermGUID	Foreign key to StandardLithology; term classifies kind of material sampled.
CollectionTrackingSysGUID	Foreign key to Tracking table; specifies metadata for sample collection event, as opposed to standard tracking that has to do with origin of information and processing to get in database.
CollectionDate	Date/time; date sample was collected (should be same as date in associated Collection tracking record).
CollectionLocationSysGUID/ CollectionLocationEntityGUID	Compound foreign key; specifies a spatial object that represents location at which sample was acquired; may be a Station, SectionIntercept, SectionInterval, or AreaOfInterest. EntityGUID specifies the Entity that contains the data instance identified by SysGUID.
StationLocationDescription	Free text; collector's description of actual outcrop to assist in relocating the exact site.
Area	String; collector's designation of area in which sample was collected, e.g., Harquahala Mountains, Bullfrog Canyon...
UTME	Floating point number; UTM easting coordinate for sample location that may be represented as a point. In the case of point sample locations, coordinate location is redundant with information linked through CollectionLocationSysGUID, but is included in the sample table to make the table portable and to provide insurance against losing sample locations….
UTMN	Floating point number; UTM northing coordinate for sample location.
UTMZone	String; specifies UTM zone for UTME and UTMN, also specifies datum used (e.g. 12N_NAD27).
Oriented	Boolean; true if sample orientation is marked on the material collected.
LocationProblemFlag	Boolean; true if actual sample location is problematic.

Description Tables (see also Table 7 in text):

Table 33. GenericDescription. Physical base table that implements abstract description class as a physical table. This is a convenience for the ESRI CASE tool, so subtype integers are defined over generic descriptions only, and subtypes of other description types with other physical base tables (EarthMaterialDescription, GeologicUnitDescription, Geologic-StructureDescription) may have their own ESRI subtype domains. Instances in this table identify descriptions of various types, identified in Geodatabase by ESRISubType, and whose attributes are defined by EntityProperty correlations for the entity (specified by DescriptionEntityGUID) associated with the description.

Field	Description
ConceptTermGUID/ ConceptEntityGUID	Compound foreign key to vocabulary entity specified by ConceptEntityGUID; term specifies what is being description; identifies the 'topic' of the description. Should be the most specific subsuming term from a controlled vocabulary. EntityGUID identifies the Vocabulary entity that contains the term identified by TermGUID.
ContextSysGUID/ ContextEntityGUID	Compound foreign key to entity specified by ContextEntityGUID; specifies a non-spatial scope for the description—a project, a workspace, an organization, a person... SysGUID identifies the data instance and EntityGUID identifies the containing entity. For some descriptions, the context may be another description, e.g., particle size description is implemented using generic description, and the owner is the EarthMaterialDescription or FractionalPartRelationship (Earth material constituent) to which the particle size applies.
DescriptionExtentSysGUID	Foreign key to Extent table; specifies the spatial domain of validity for the description.
DescriptionPurposeTermGUID	Foreign key to ScienceLanguage; term specifies the purpose of this description. DescriptionPurpose distinguishes between default description, necessary property descriptions, identifying property descriptions, and instance descriptions.
DescriptionEntityGUID	Foreign key to entity table; defines the attribute collection and domains for this description instance. Essentially defines the 'kind' of description. For entities implemented as logical (soft) tables, this value will be associated with a unique ESRI subtype value.
ESRISubtype	Integer; used to distinguish different description types in geodatabase; redundant with DescriptionEntityGUID, but geodatabase requires integer for typing.

Table 34. GeologicAge. Base table for geologic age description. Different specification details may be used through AttributeRelationship links based on the type property. Derived classes (identified by ESRISubtype / DescriptionEntityGUID) represent age specification in different ways: time instant (a number Ma before present, which may be inferred from 1 to many geochron observations...), a named era (geologic time scale—e.g. Miocene), or a range specified by lower and upper bounds that may be instants, named eras, or geologic events. These different specifications are unified in this table with a best guess numerical minimum and maximum time coordinate (for analysis), and a DisplayName that summarizes the interpretation (for a data browser).

Field	Description
MinimumAge	Floating point number; best guess time coordinate for younger bound of age range or uncertainty envelope for age.
MaximumAge	Floating point number; best guess time coordinate for older bound of age range or uncertainty envelope for age.

Table 35. EarthMaterialDescription. Base table for descriptions of rocks and non-consolidated compound Earth material. Identifies description instances with a GUID and a display name, and provides values for specifying properties common to all compound Earth Materials. Other description attributes are linked through AttributeRelationship instances. Includes all fields from GenericDescription (above).

Field	Description
DegreeOfCrystallinityTermGUID	Foreign key to ScienceLanguage; term specifies degree to which particles in an igneous or metamorphic rock are bounded by crystal faces. e.g., holocrystalline, holohyaline... Range is CompoundMaterial—not applied to individual constituents. For sedimentary materials specify granular. Useful to distinguish crystalline from granular rocks (Struik et al., 2000).
ConsolidationDegreeTermGUID	Foreign key to ScienceLanguage; term specifies the degree to which an aggregation of EarthMaterial particles is a distinct solid material. Not to be confused with induration. Consolidated materials may have varying degrees of induration, which relates to the hardness of the aggregated material.
GrainDiscernibilityTermGUID	Foreign key to ScienceLanguage; term specifies the degree to which the individual grains of the material may be distinguished visually (unaided eye, or with hand held magnifier); generally falls into three groups: phaneritic—all grains discernible; aphanitic—all grains non-discernible, and 'phaneritic and aphanitic'—some grains discernible.
RepresentativeSize	Floating point number; specifies 'unit cell' diameter for representative sample of the described material. How big a piece of the material must be sampled to obtain all the characteristics of the material in general.

Table 36. GeologicUnitDescription. Geologic unit description specifies properties common to all geologic units; ESRI subtypes are used to apply rules for specific kinds of geologic units that have different combinations of properties and property value domains. Includes all fields from GenericDescription (above).

Field	Description
GeologicAgeSysGUID	Foreign key to GeologicAge table; specifies default age object from identifying, necessary, or default description for a geologic unit unless this description pertains to a more specifically constrained age for the unit, in which case reference is to a unique GeologicAge object.

Table 37. GeologicStructureDescription. Base table for description of geologic structures. The attributes and subtypes for these descriptions remain to be spelled out explicitly. This table includes only properties common to all geologic structures. Includes all fields from GenericDescription (above).

Field	Description
PervasivenessTermGUID	Foreign key to ScienceLanguage; term specifies degree to which a structure is continuous throughout a body of rock, considered at the dimension specified by the CharacteristicDimension property.
GeometricAspectTermGUID	Foreign key to ScienceLanguage; term specifies if structure is essentially planar, linear, or irregular.
CharacteristicDimension	Floating point number; specifies order of magnitude length scale (in meters) for size of the structure. e.g. mica schistosity might have a scale factor of cm (.01); a bedding fabric might have a scale of 10 m (10); a large scale fold might have scale km (1000).

Relationship Tables (see also Table 8 in text):

Table 38. AttributeRelationship. Represents observation/quantification/specification of the value of some property that is part of a description. Nature of value type is specified by Domain attribute of EntityProperty instance associated with the AttributeRelationship instance.

Field	Description
OwningItemSysGUID/ OwningItemEntityGUID	Compound foreign key to entity specified by OwningItemEntityGUID; specifies the base data instance for the description. The AttributeRelationship instance is owned in that if the base instance is deleted, associated attribute relationship instances are also deleted. SysGUID identifies the data instance and EntityGUID identifies the containing entity.
SequenceNo	Integer; orders attributes associated with a particular description instance. In a Genesis description, this establishes the order of events included in the genesis.
EntityPropertySysGUID	Foreign key to EntityProperty table; specifies the semantics of the attribute (what property is specified), the domain for the attribute in this entity, as well as other validation constraints.
FrequencyTermSysGUID	Foreign key to ScienceLanguage; term specifies the degree to which a property is present in all parts or instances of the described object.
PropertyValueSysGUID/ PropertyValueEntityGUID	Compound foreign key to entity specified by PropertyValueEntityGUID; specifies the value for the property identified by EntityPropertySysGUID in this attribute. SysGUID identifies the data instance and EntityGUID identifies the containing entity.
IntensityTermGUID	Foreign key to ScienceLanguage; term specifies the intensity or degree to which a property is developed. Properties like fabric may be present, but developed to different degrees. For example, schistosity may be weakly developed in a rock in which only 10-20% of tabular mineral grains are aligned (quartzo feldspathic gneiss with 20% mica in schistosity), but strongly developed in a rock in which all mineral grains are tabular and aligned in schistosity.
ConfidenceTermGUID	Foreign key to ScienceLanguage; term specifies the data source agent's confidence in assigning this value to the property in this instance.
MeasureProcedureTextSysGUID	Foreign key to TextDescription; specifies the procedure used to assign the attribute value.
EvidenceTextSysGUID	Foreign key to TextDescription; discussion for evidence in assigning value to this property in this instance.
ESRISubtype	Integer; identifies relationship instances for different attributes. Used by geodatabase implementation to define business rules for cardinalities and table links.

Table 39. FractionalPartRelationship. Correlation that aggregates parts into a whole to represent parts explosion for EarthMaterial and GeologicUnit.

Field	Description
OwningItemSysGUID/ OwningItemEntityGUID	Compound foreign key to entity specified by OwningItemEntityGUID; specifies the base data instance that represents the 'whole' whose parts are being enumerated by the part relationships. The FractionalPartRelationship instance is owned in that if the base instance is deleted, associated attribute relationship instances are also deleted. SysGUID identifies the data instance and EntityGUID identifies the containing entity.
PartSysGUID/ PartEntityGUID	Compound foreign key to entity specified by PartEntityGUID; specifies the data instance that represents a part in the 'whole' being described. SysGUID identifies the data instance and EntityGUID identifies the containing entity.
SubClassEntityGUID	Foreign key to Entity table; specifies different fractional part types defined based on the kind of associated part (e.g., GeologicUnit, EarthMaterial part; EarthMaterial, StandardMineral part). Redundant with ESRI subtype in geodatabase to specify cardinalities and valid OwningItemEntity, PartEntity.
PartTypeTermGUID	Foreign key to ScienceLanguage; term specifies the nature of the units of each constituent in an aggregation. Because the constituents are identifiable parts, the part units must have some definition. For example clast, crystal. In a GeologicUnit part type distinguishes parts with identity—like members, from parts that do not have identity—like a lithosome (fining-upward sequence) that is used as a building block for the unit description, but is generalized to define a 'kind' of geologic unit part; EarthMaterial constituents likewise do not have identity in a geologic unit; these will typically have roles in the geologic unit aggregation—e.g., vein, inclusion.
PartRoleTermGUID	Foreign key to ScienceLanguage; term specifies the role a constituent or part plays in a composition aggregation. In a compound Earth material, the same EarthMaterial may occur as more than one constituent, playing different roles. For example, feldspar may be present as groundmass and as phenocrysts within a single igneous rock. Geologic unit parts that are geologic units (have identity, extent) have role that is inherent in the part; parts that do not have identity ('xenolith', 'fining upward sequence') have roles.
ProportionDisplayString	String; conveys the proportion information for this constituent to users.
QuantityTypeTermGUID	Foreign key to ScienceLanguage; term distinguishes quantities specified by a value range, an average value with symmetric uncertainty, a value with asymmetric uncertainty, etc. for proportion quantity of this constituent.
TypicalProportion	Floating point number; single value to that best represents the proportion that this part makes in the whole aggregation, by volume, for use in analyses where a single value is required; determination of value is based on quantity type.
MinProprotion	Floating point number; lower numerical bound for proportion value, may be limit of uncertainty envelope or lower bound of value range (determined by QuantityTypeTermGUID).
MaxProportion	Floating point number; upper numerical bound for proportion value, may be limit of uncertainty envelope or upper bound of value range (determined by QuantityTypeTermGUID).
ProportionBasisTermGUID	Foreign key to ScienceLanguage; term specifies the basis for assigning proportion values—published data, compilers summary of published data, original data by data set author…
MeasurementMethodTermGUID	Foreign key to ScienceLanguage; term specifies the method used to determine the proportion—e.g., single point count, field estimate, average of multiple point counts.

Table 40. SimpleRelationship. Simple correlation table that associates two objects with no attributes; relationship type specifies semantics and links to a vocabulary term; this attribute is redundant with ESRI subtype integer used to specify valid relationship targets in geodatabase implementation.

Field	Description
SequenceNo	Integer; for ordering groups of relationship instances.
FirstItemSysGUID/ FirstItemEntityGUID	Compound foreign key to entity specified by FirstItemEntityGUID; identifies first item role filler for relationship. SysGUID identifies the data instance and EntityGUID identifies the containing entity.
SecondItemSysGUID/ SecondItemEntityGUID	Compound foreign key to entity specified by SecondItemEntityGUID; identifies second item role filler for relationship. SysGUID identifies the data instance and EntityGUID identifies the containing entity.
RelationshipTypeTermGUID	Foreign key to ScienceLanguage; term specifies relationship semantics, include to have a globally unique, implementation independent specification. Redundant with ESRI subtype in geodatabase to specify valid FirstItemEntity and SecondItemEntity and cardinalities.

Table 41. ObservationRelationship. Correlation table to establish science relationships between objects; the related objects have a lifetime independent of the observation relationship. Semantics determined by RelationshipTypeTermGUID.

Field	Description
SequenceNo	Integer; for ordering groups of relationship instances.
RelationshipTypeTermGUID	Foreign key to ScienceLanguage; term specifies the semantics of the relationship. In this table, this field is not redundant with ESRI subtype, since relationships with different semantics might involved the same Entities in the From and To roles.
FromSysGUID/ FromEntityGUID	Compound foreign key to entity specified by FromEntityGUID; identifies the source role filler for relationship. SysGUID identifies the data instance and EntityGUID identifies the containing entity.
ToSysGUID/ ToEntityGUID	Compound foreign key to entity specified by ToEntityGUID; identifies the target role filler for relationship. SysGUID identifies the data instance and EntityGUID identifies the containing entity.
ConfidenceTermGUID	Foreign key to ScienceLanguage; term specifies the data source agent's confidence in assigning this value to the property in this instance.
BasisTermGUID	Foreign key to ScienceLanguage; term specifies the basis for asserting the relationship.
EvidenceTextSysGUID	Foreign key to TextDescription; identifies text that provides more in depth discussion of evidence for relationship.
ESRISubtype	Integer; used by geodatabase to define business rules for cardinalities and table links. Specifies valid From and To object classes (entities).

Table 42. MetadataRelationship. Simple correlation table for associating metadata objects. Semantics determined by RelationshipTypeTermGUID.

Field	Description
SequenceNo	Integer; for ordering groups of relationship links.
ItemSysGUID/ ItemEntityGUID	Compound foreign key to entity specified by ItemEntityGUID; identifies a data instance that has associated metadata information. E.g., a Tracking record (Item) that has associated Citation instances (metadata), or an Activity that has multiple associated Person-Organization instances (metadata). SysGUID identifies a data instance in the entity specified by EntityGUID.
MetadataSysGUID/ MetadataEntityGUID	Compound foreign key to entity specified by MetadataEntityGUID; identifies metadata object to attach to a data instance (e.g. Tracking or Activity). E.g., a Citation (metadata) may be attached to a Tracking record (item). SysGUID identifies a data instance in the entity specified by EntityGUID.
RelationshipTypeTermGUID	Foreign key to ScienceLanguage; term specifies relationship semantics. Redundant with ESRI subtype, but included so relationship type has globally unique identifier.
ESRISubtype	Integer; used by geodatabase to identify object classes that are allowable ItemEntity and MetadataEntity entities, and set business rules for cardinality.

Photogrammetry Methods at the Utah Geological Survey: From Field Mapping to Published Map

By Kent D. Brown

Utah Geological Survey
1594 W. North Temple
Salt Lake City, Utah 84116
Phone: (801) 537-3300
Fax: (801) 537-3400
e-mail: kentbrown@utah.gov

INTRODUCTION

Since 1991 the Geologic Mapping Program of the Utah Geological Survey (UGS) has used photogrammetry to produce geologic maps from field mapping drawn on aerial photographs. Since the science of geology is three-dimensional it makes sense to do mapping in 3-D. Photogrammetry can be defined as applying "the sciences of geometry, mathematics, and physics that use the photographic image of a 3-D scene on a 2-D piece of paper to reconstruct a reliable and accurate model of the original 3-D scene" (James R. Williamson, Ph.D., Photogrammetry website http://www.123photogrammetry.com/photogrammetry.html).

Traditional mapping methods involve duplication of work as the geology is hand drawn three or more times before the map is published: first on the aerial photographs in the field, then hand transferred to either a topographic base map or an orthophoto quad, then redrawn by the cartographer, either by hand or digitized. When using photogrammetry, the geologic map is compiled in stereo as a digital file and there is no longer a need for redrawing the map, thereby saving many hours of labor and reducing the potential for introducing errors. Other benefits include increased precision; ability to view orthographically corrected photographic images in 3-D, and precise software functions for calculating three-point problems.

There are three types of photogrammetric equipment: analog, analytical, and digital (softcopy). The analog type is entirely mechanical, the analytical type is a computer-controlled mechanical system with digital output, and the digital type is free of mechanical devices; it uses only a computer, 3-D glasses, and software.

In analog and analytical systems the aerial photographs are placed in the instrument and three orientations are performed to correct for lens distortion, Earth curvature, aircraft altitude changes between photos, aircraft orientation, changes in ground surface elevation between photos, etc. In a digital system the photographs are scanned and then the same three orientations are performed on the photographic images using software.

The UGS purchased an analytical system in 1991 and has used it in publishing more than 35 geologic maps. In May of 2003 we began using a digital system in which geologists digitize geologic map features directly on the on-screen 3-D photographic image. This paper describes the use of our analytical stereoplotter and our digital photogrammetry workstation and how we use them to compile geologic maps.

OVERVIEW OF METHODS

Any successful photogrammetry project begins with the acquisition of suitable aerial photographic coverage of the project area. The geologist then delineates geologic features on the photographs in the field.

Photogrammetry was developed to accurately reproduce the three-dimensional surface of the Earth from two-dimensional photographs. Pairs of overlapping aerial photographs, called stereo pairs, are required to achieve this. Typical image overlap between photographs is 60 percent.

Cameras designed for aerial photography must be precisely calibrated. In the United States, aerial cameras are calibrated by the United States Geological Survey (USGS), National Mapping Division. Camera calibration reports are made for each camera as certification of precision. This report is critical in the process of photogrammetric setups.

The film plane of each camera includes a set of four or eight reference marks known as fiducials (Figure 1). These fiducial marks are precisely measured positions in the exposure frame of each photograph. Fiducial marks are used to locate the position of the photograph in the stereoplotter. The fiducial measurements, along with the calibrated focal length of the lens, are included in the camera calibration report.

In order to recreate the original Earth surface from

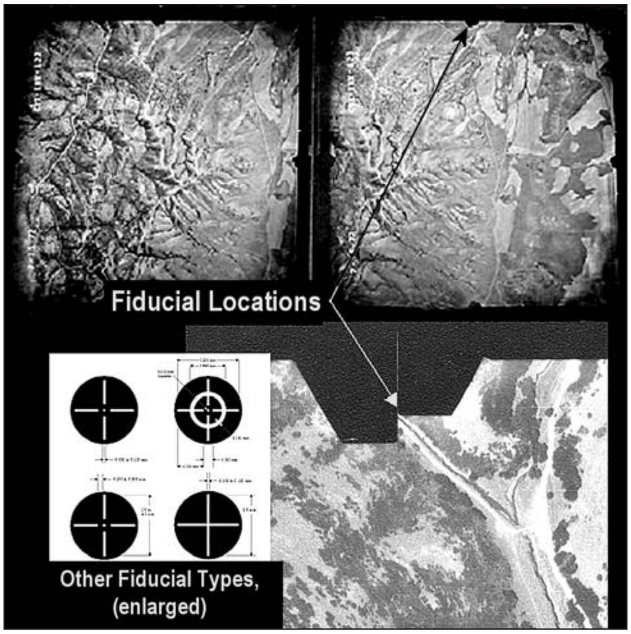


Figure 1. Typical aerial camera fiducial marks.

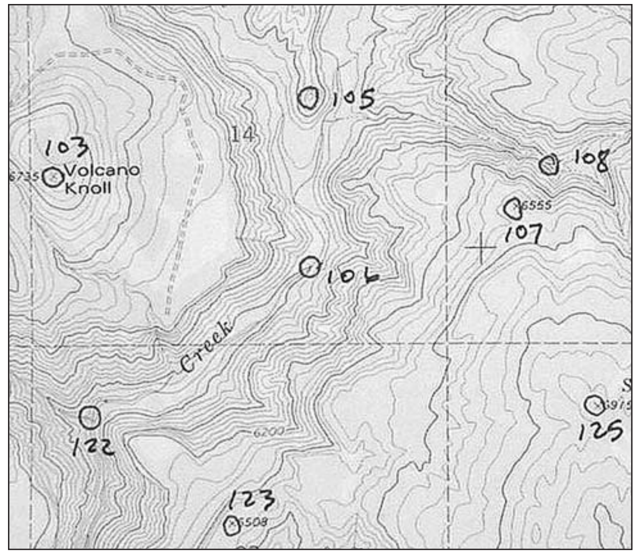


Figure 2. USGS topographic map showing selected control point locations.

the photography, three-dimensional ground control is required. Traditionally at the UGS, we prepared the ground control using the USGS 7.5-minute topographic maps (Figure 2). We viewed the stereo pair through a stereoscope and selected control points by correlating features that were visible on both the aerial photographs and the topographic map. We marked these control point positions on the topographic map and interpolated the point elevations from the contours. The x-y-z coordinates for the ground control points were established by digitizing them from the topographic map and manually entering the interpolated point elevations.

Currently we use a newer method of creating the ground control using a USGS digital orthophoto quadrangle (DOQ). A DOQ is a mosaic of photogrammetrically corrected photographic images that meets National Map Accuracy Standards for 1:24,000-scale mapping. Similar techniques of point selection are employed with this method; however, the points are placed on the DOQ (Figure 3) using ESRI ArcGIS software. This method extends the possibilities for control point locations since almost any point visible on the photograph can be used for control. Next, the x-y points are draped over a USGS 10-meter-resolution digital elevation model (DEM) from which the point elevations can be obtained. The x-y-z coordinates are then extracted and exported to a ground control file used by orientation software to establish real-world coordinates for the stereo model.

Analytical Photogrammetry

In 1991, the UGS purchased an Alpha 2000 analytical stereoplotter (Figure 4) manufactured by International Imaging Systems, a now-defunct corporation. Analytical stereoplotters are computer-controlled instruments and most of them use an x-y-z motion device called a "puck" as a user interface. The photogrammetry software we use is VrOne by Cardinal Systems L.L.C.

A pair of photographs properly set up in the stereoplotter is called a stereo model. To begin the setup, a stereo pair of photographs is placed between the glass plates of the instrument (Figure 5). The glass plates move independently in the x and y directions to maintain stereovision during model setups and map compilation. The photogrammetrist creates a stereo model by performing three critical orientations using the system software functions.

The first orientation is called "inner orientation" and is necessary to establish the position of the aerial photographs in relation to the glass plates. The positions of the glass plates in the stereoplotter have been calibrated to within 2 μm. Using the puck, each fiducial mark is measured by locating its x-y position and digitizing it. These positions are cross-referenced by the orientation software to the fiducial measurements in the camera data file. The camera data file is created using information from the camera calibration report.

The second orientation is called "relative orientation" and is performed to eliminate parallax in the stereo model. Parallax is defined as "the apparent displacement of an observed object due to a change in the position of the observer" (Random House Webster's Electronic Dictionary and Thesaurus, College Edition, Version 1.0). It commonly occurs in a set of overlapping photos. Depending

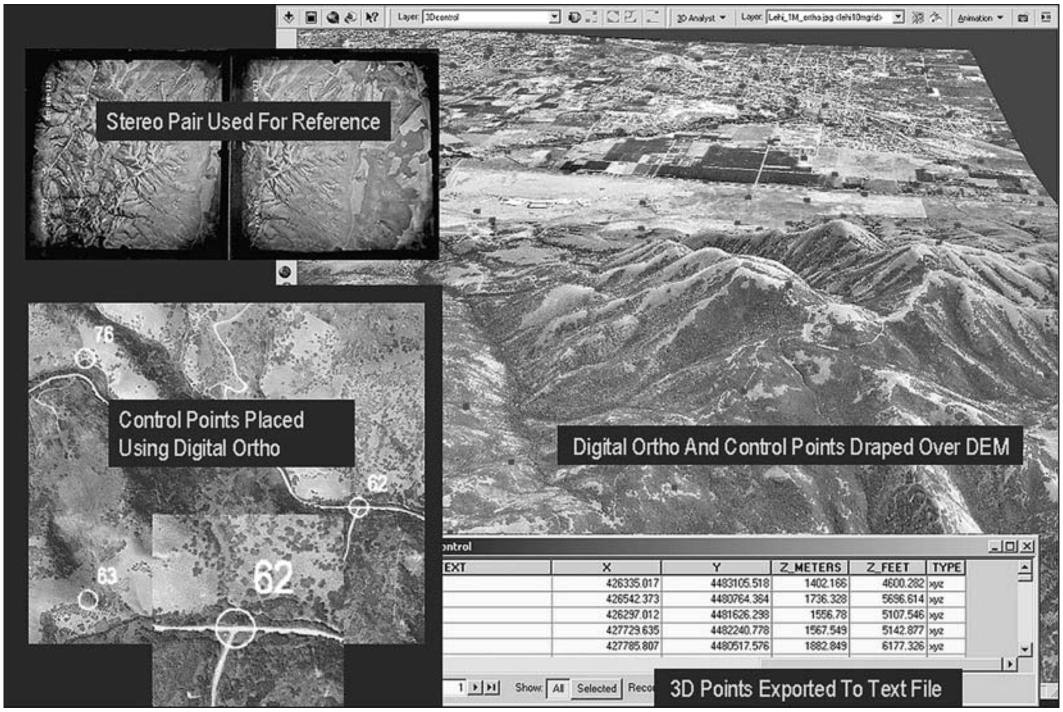


Figure 3. Suitable ground control is prepared by comparing the stereo pair to the digital orthophoto and then choosing correlative photo points. Points are placed on the digital orthophoto using ESRI ArcMap and saved as a shapefile. Using ESRI ArcScene, the points and digital orthophoto are draped over a 10-meter digital elevation model, from which the x, y, and z values of the points are calculated and added to matching fields in the shapefile table.

on the software, six or seven matching photo positions are measured (Figure 6) for each photo pair and the orientation software uses these positions to eliminate parallax and achieve stereovision.

The third orientation is called "absolute orientation."

Figure 4. Alpha 2000 analytical stereoplotter.

In this step, the software accesses the ground control file and prompts the photogrammetrist to locate two ground control points within the stereo model area. After the first two points are located and digitized, the software searches the ground control file for points located within the dimensions of the stereo model and drives the glass plates of the stereoplotter to each of the remaining control point positions. These point positions are then digitized and the software triangulates between all control points. This step establishes three-dimensional geometry, and real-world coordinates are applied to the stereo model.

When all three orientations are completed the stereo model is ready for the geologist to compile geologic features delineated on the aerial photographs. The geologist views the stereo imagery through the binocular eyepiece. Both the left and right views show a small illuminated dot called a measuring mark, which represents the ground surface. The puck has a z-motion thumbwheel that, when rotated, shows the two measuring marks either converging or diverging until they appear to fuse, indicating that the x-y position is on the ground surface. Moving the puck, the geologist travels in the x and y directions to follow the geologic feature being compiled while adjusting the z position with the thumbwheel to keep the measuring marks converged on the ground surface. These motions are recorded as streams of x-y-z coordinates and are trans-

Figure 5. "Under the hood" showing the glass plates of the Alpha 2000.

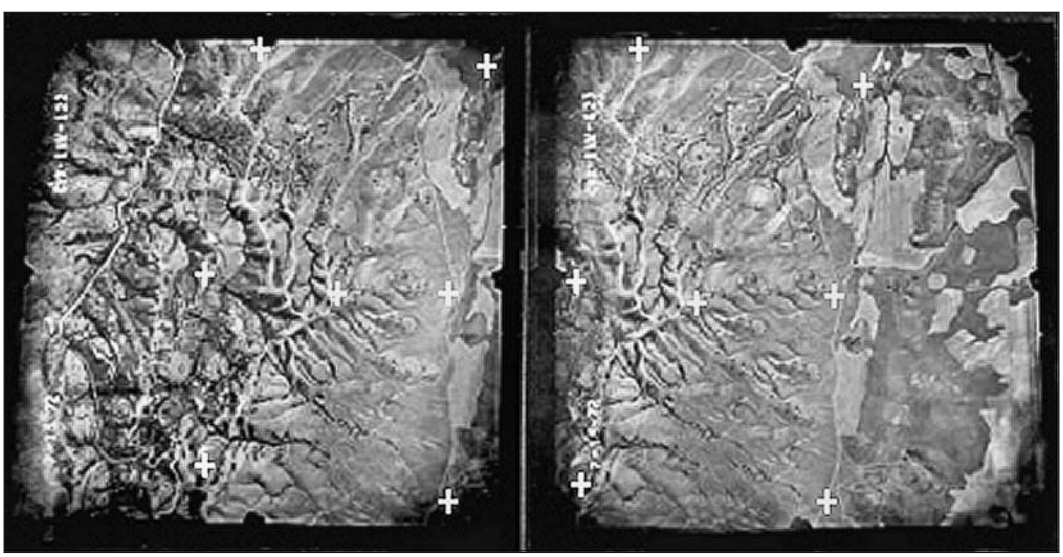


Figure 6. "Relative orientation" is done by selecting correlative photo points: this eliminates model parallax.

ferred to a computer hosting the VrOne photogrammetry software.

A valuable component of VrOne is a function designed to precisely calculate three-point problems. Using this function the geologist is able to map fault, bedding, and foliation planes more accurately, especially where exposures are poor, or in areas that are difficult or even impossible to reach in the field.

All x, y, and z motions are recorded as three-dimensional lines with each vertex carrying an x-y-z coordinate. In this manner all contacts, faults, fold axes, point data, symbols, and other geologic features are compiled (Figure 7). This compilation results in a georeferenced vector file that is imported into ArcGIS, where a Geographic Information System (GIS) database is created.

Digital Photogrammetry

The digital, or soft copy, photogrammetry systems (Figure 8) are much simpler in design than the analytical systems; they consist of a computer with a stereo-capable graphics system, 3-D glasses with electronic shutters, and a "3-D mouse" as a user interface. The 3-D mouse is a reconfigured optical mouse with x, y, and z motion control and several user-configurable buttons. All other hardware of an analytical system has been replaced with software programming, alleviating the problems with mechanical devices wearing out or needing adjustments to stay within close tolerances. For digital photogrammetry, the UGS uses VrTwo software, also from Cardinal Systems L.L.C.

Digital systems require that the aerial photographs be scanned as high-resolution images (1000+ dpi) that are then spatially corrected using the same three orientations described previously. During map compilation the photographic images are seen in 3-D on the screen using the special glasses, and all geologic features are drawn directly on this virtual ground surface using the 3-D mouse (Figure 9).

Digital systems also support the creation of digital terrain models (DTM's) to enhance the mapping project. A DTM is an elevation model of the Earth's surface derived from photogrammetry. Using a DTM, a triangular irregular network (TIN) can be created. From a TIN, digital orthophotos and automatically generated topographic contours can be produced.

PUBLISHING

UGS publishing methods are quite detailed so only a brief overview is given here. The completed geologic map compilation is exported as a series of drawing exchange format (DXF) files. DXF is an industry standard export format and the files are readable by virtually any mapping or publishing software. The DXF files are imported into ArcInfo and coverages are created. After further editing, these arcs are used to build polygon topology and then all arcs and polygons are attributed to show what type of geologic features they represent. The geologic map is

Figure 7. Screenshot showing geologic map created with VrOne photogrammetry software.

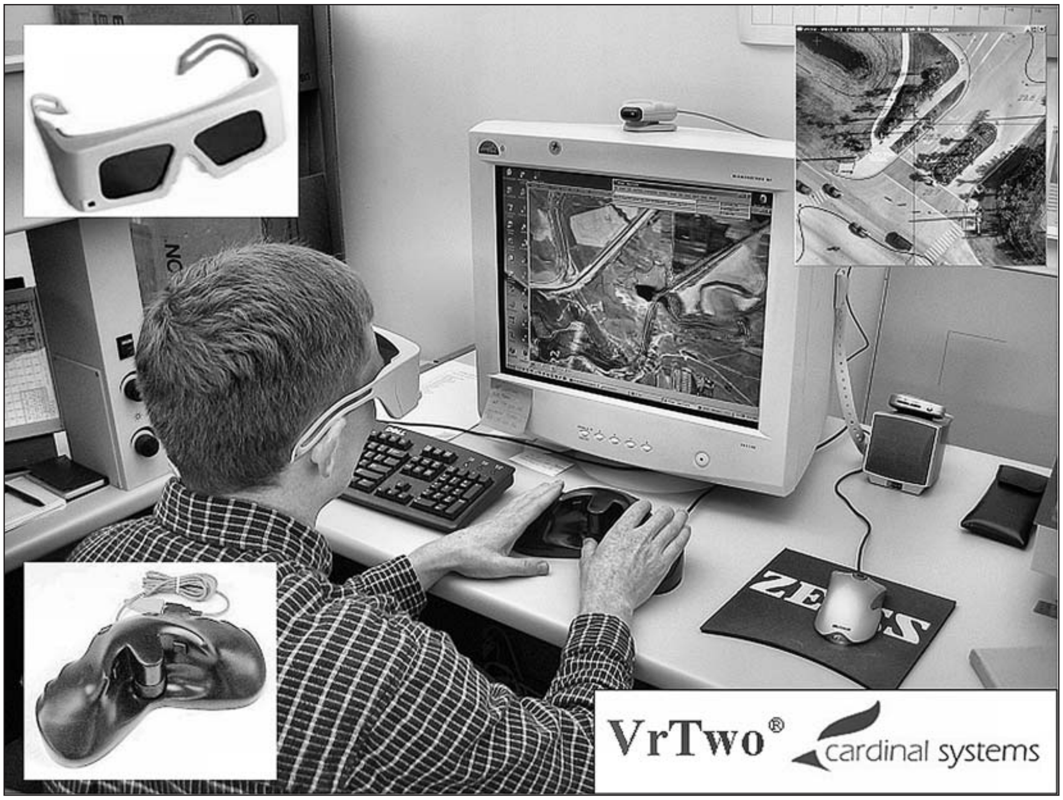


Figure 8. Digital or "soft copy" photogrammetry workstation using VrTwo software, showing stereo glasses and "3-D mouse".

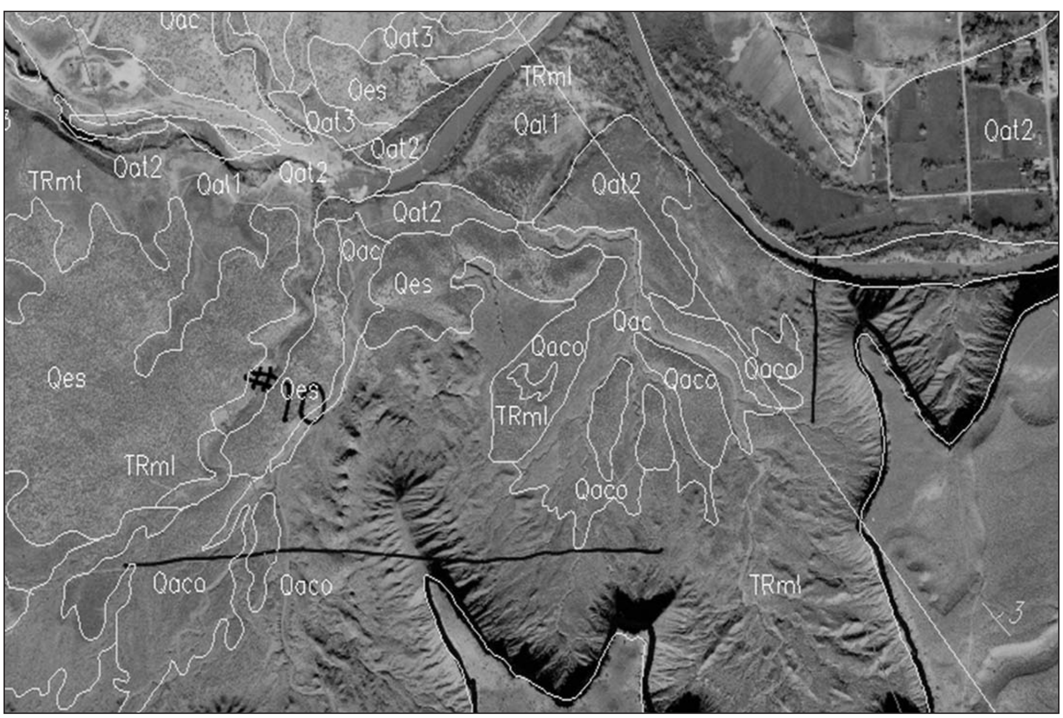


Figure 9. Screenshot showing geologic features digitized on the 3-D photographic surface in VrTwo software. This image is a simulation since true 3-D using VrTwo is not possible in this document.

then plotted on a large-format inkjet plotter for scientific review. After review comments are made and any corrections completed, the coverages are converted to shapefiles and are forwarded to our editorial group for publication.

The UGS editorial group employs two cartographers who use Illustrator, by Adobe Systems Inc., and MAPublisher, by Avenza Systems Inc., to import shapefiles and base map images. MAPublisher is designed as a set of plug-ins for Illustrator and will import georeferenced GIS data with full access to all feature attributes. The cartographers can select geologic features from the attribute tables and then use the publishing tools of the Illustrator software to create the final geologic map layout. Illustrator is also used to create the map explanation sheet. At this stage the maps can be converted to Adobe portable document format (PDF) or MrSID (LizardTech Inc.) image files for distribution over the Internet, or inkjet plots are made for distribution to libraries and to be sold in the Utah Department of Natural Resources bookstore. Some UGS geologic maps are sent as Adobe Illustrator files to a printing company for high-quality offset printing. Graphically, offset printing yields a superior paper copy with a longer useful lifetime than the inkjet plot and remains the UGS's preferred method of publishing paper maps.

CONCLUSIONS

The use of photogrammetry for geologic mapping has benefited the UGS mapping program in many ways. These techniques have resulted in a much higher level of detail and precision than can be achieved with other methods. Using software designed for precise calculation of three-point problems, geologists can map or measure fault, bedding, and foliation planes more accurately, especially where exposures are poor or in areas that are difficult to reach in the field. Our geologic maps are compiled in stereo as digital files and using photogrammetry and ArcGIS nearly eliminates most duplication of effort. No longer do geologists redraw the map several times, thereby saving many hours of labor and reducing the potential of introducing errors.

SUGGESTED READING

Avery, E. T., 1968, Interpretation of Aerial Photographs, Second Edition: Burgess Publishing Company, 324 p.

McGlone, Chris, editor, 2004, Manual of Photogrammetry, Fifth Edition: American Society of Photogrammetry, 1168 p.

Raisz, Erwin, 1962, Principles of Cartography: McGraw-Hill, 315 p.

HARDWARE AND SOFTWARE VENDORS

Adobe Systems Inc., 345 Park Ave., San Jose, California 95110-2704, accessed at http://www.adobe.com/.

Avenza Systems Inc., 124 Merton St., Suite 400, Toronto, Ontario, CANADA M4S 2Z2, accessed at http://www.avenza.com/.

Cardinal Systems L.L.C., 175 Lehigh Avenue, Flagler Beach, FL 32136 USA, (386) 439-2525, Fax: (386) 439-1553, Email: mike@cardinalsystems.net, accessed at http://www.cardinalsystems.net/.

ESRI, 380 New York St., Redlands, CA 92373-8100 USA, (909) 793-2853, Fax: (909) 793-5953, Email: info@esri.com, accessed at http://www.esri.com.

LizardTech, The National Building, Suite 200, 1008 Western Ave., Seattle, WA 98104, (206) 652-5211, Fax: (206) 652-0880, accessed at http://www.lizardtech.com/.

Digital Cartographic Production Techniques Using Airborne Interferometric Synthetic Aperture Radar (IFSAR): North Slope, Alaska

By Christopher P. Garrity

U.S. Geological Survey
National Center, MS 956
12201 Sunrise Valley Drive
Reston, VA 20192
Telephone: (703) 648-6426
Fax: (703) 648-6419
e-mail: cgarrity@usgs.gov

INTRODUCTION

In 2003, the Bureau of Land Management, in cooperation with various agencies including the U.S. Geological Survey (USGS), began preparing an Environmental Impact Statement in preparation for potential oil and gas development in the northeastern portion of the National Petroleum Reserve–Alaska. To aid decision-makers in assessing the placement and design of permanent facilities, the USGS was assigned the task of digitizing a series of engineering geologic maps originally published in 1979. Resultant digital maps were incorporated into a Geographic Information System (GIS) project designed to facilitate analysis among federal and industry project managers. A secondary study was to determine the feasibility of using remotely-sensed data to map surficial geology in areas where existing geologic maps were not available. Sensors implemented in the study included Interferometric Synthetic Aperture Radar (IFSAR; http://www.intermap.com), Landsat-7 Enhanced Thematic Mapper Plus, and Advanced Spaceborne Thermal Emission Reflection Radiometer. This paper focuses on various cartographic production techniques related to data collected by the IFSAR system.

DATA DESCRIPTION

IFSAR data were collected by the STAR-3i airborne synthetic aperture radar system. STAR-3i is a high-resolution, single-pass, across-track IFSAR system, which uses two apertures to image the surface. The path length difference between the apertures for each image point, along with the known aperture distance, is used to determine the topographic height of the terrain. The IFSAR system is capable of collecting data with a vertical accuracy of <1 m and a horizontal accuracy of <3 m.

Data are delivered as three core products: orthorectified radar images (ORRIs), digital surface models (DSMs), and digital terrain models (DTMs). ORRIs are 8-bit grayscale GeoTIFF images that show the radar reflectance intensity of various earth surface materials (Figure 1). These images are commonly used to identify and extract drainage networks and cultural features such as pipelines, roads, and buildings. The ORRIs used in this study had a pixel size of 1.25 m and a horizontal accuracy of 2.5 m. The DSMs, or "first-return" elevation data, display the first surface on the ground that the radar strikes (Figure 2). These images consist of all measured points collected by the sensor, including the z-values of structures (e.g., building and towers) and vegetation (e.g., trees and crops). These elements are removed from the DSM through filtering techniques to create a DTM (Figure 2). The DTMs, or "bald-earth" elevation data, are similar to Digital Elevation Models (DEMs) in that non-terrain elements are absent. However, unlike the regular array of elevation values that are characteristic of a DEM, a DTM defines topographic elements by irregularly spaced break-lines, or abrupt changes in surface smoothness, such as shorelines, roads, streams, and slope breaks. The result is a more accurate depiction of the terrain, which is useful for contouring, TIN calculations, and other terrain modeling.

STUDY AREA

The study area is located in the Alaskan North Slope along the western edge of the Colville River delta. Bounded by the Beaufort Sea to the north, the study area encompasses nearly 825,000 acres, 95% of which lies in NPRA. The area is completely within the Arctic Coastal Plain, a terrain characterized by a flat to gently rolling tundra-covered surface (Carter, 1985). Toward the coast, the terrain is dominated by periglacial features, includ-

Figure 1. ORRI of Nuiqsut area, North Slope, Alaska. Although ORRIs resemble grayscale aerial photographs, each pixel value actually represents the magnitude of sensor signal return.

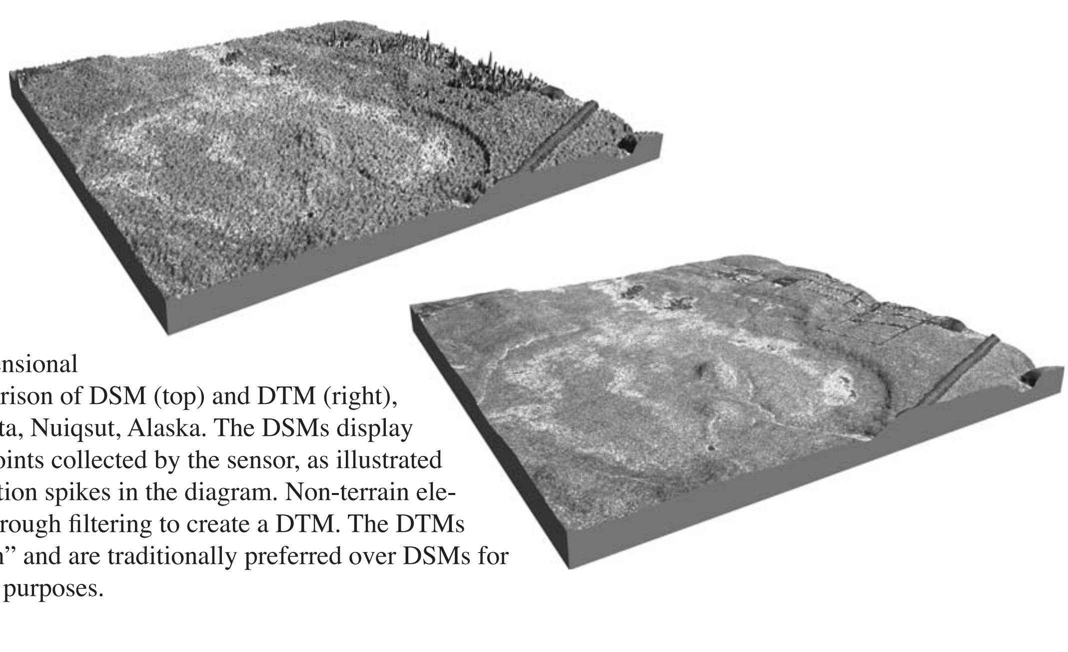


Figure 2. Three-dimensional block diagram comparison of DSM (top) and DTM (right), draped with ORRI data, Nuiqsut, Alaska. The DSMs display all of the measured points collected by the sensor, as illustrated by the irregular elevation spikes in the diagram. Non-terrain elements are removed through filtering to create a DTM. The DTMs simulate a "bald-earth" and are traditionally preferred over DSMs for topographic mapping purposes.

ing thousands of northeast-trending thaw-lakes. These thaw-lakes are created when ponding of water occurs at the surface, as a result of melting of ground ice followed by surface subsidence. The distinctive elongated shape of the lakes is caused by wave and channel erosion from prevailing winds (Carter, 1985). Inland, the thaw lakes are replaced by meandering streams and treeless areas of tundra vegetation. Elevation in the study area gradually increases in a southerly direction from about 1 to 7 m near the coast, to a maximum of about 50 m at an inland area of large stabilized dunes.

SHADED RELIEF PRODUCTION

Computer-based analytical hillshading has become a widely used tool to visualize three-dimensional topography on a two-dimensional surface. The use of aesthetically pleasing shaded relief images that realistically show topography continues to increase as the quality of rendering software and elevation data improve. Still, it is difficult, if not impossible through purely automated methods to achieve the overall quality of a well-drafted shaded relief map. Unlike manually-produced shaded relief maps, analytical hillshade images often reveal unsightly imperfections in the elevation data used to render the image. The IFSAR data was no exception, and contained flaws in some areas because excess motion in the aircraft caused visible ripples in the dataset. A regular banding pattern was apparent along sensor swath boundaries when elevation data were viewed at small-scales. This aside, the DSMs and DTMs derived from the IFSAR system proved to be an excellent data source for generating shaded relief images.

Large-scale (1:20,000) shaded relief images were generated from DTM data to identify the location of surficial materials with greater accuracy. Initial image rendering revealed potential challenges related to the portrayal of surface features in an area devoid of any significant relief. When rendered with no vertical exaggeration, the shaded relief image resembled a flat, gray rectangle. When a modest amount of vertical exaggeration (5-10X) was applied to the image, shadows appeared blocky and generally were unsightly at the desired map scale. As described below, two graphical techniques (bump mapping and shade softening) were applied to alleviate these aesthetic concerns and to give the landscape images a more natural appearance.

Bump Mapping

Bump mapping is a technique that simulates surface texture through shading. The technique essentially adds detail to a surface without modifying the surface itself, and is widely used in both graphic design and computer gaming. Bump mapping uses a "height map" to simulate surface texture. A height map, which is analogous to a DEM, is a grayscale image that dictates the amount of vertical displacement, relative to adjacent cells, observed in the rendered bump map. Lighter grayscale pixels create a sense of maximum relief, whereas darker pixels have the opposite effect (Figure 3). Bump mapping is especially useful when rendering small-scale to medium-scale 3-D scenes with low resolution DEM data (Figure 4).

For the project, DTM surfaces were slightly bump-mapped to generate a hillshade image with a more natural appearance. Because the images were used for surface delineation, a random height map was unsuitable for fear of compromising data integrity. Instead, the corresponding DSM was used as a height map to generate the bumped surface. The natural "roughness" of the DSM gave the hillshaded DTM a subtle texture without obscuring the DTM surface with unnecessary detail. Similar results were achieved in Adobe Photoshop by multiplying DTM and DSM hillshade images with varying vertical exaggerations and layer opacities. However, the bump mapping method proved less time-consuming and provided greater functionality for texture manipulation.

Shade Softening

Shade softening is a technique used to alleviate jagged shades that sometimes appear in automated hillshade images, particularly when vertically exaggerated at large scales. The technique is similar in principle to shadow softening, a computer graphic technique that is applied to 3D objects, rendering object shadows with both an umbra and penumbra(s). In both cases, these softening techniques give the final image or scene a more realistic appearance (Figure 5).

For the North Slope project, shade softening was accomplished by generalizing the raw data grid (which smooths, or softens, the irregularities), isolating the shaded pixels, and then merging the softened shade layer with the original grid. To generalize the raw data grid, softening techniques were applied in ArcInfo Workstation by means of various focal functions. The amount of

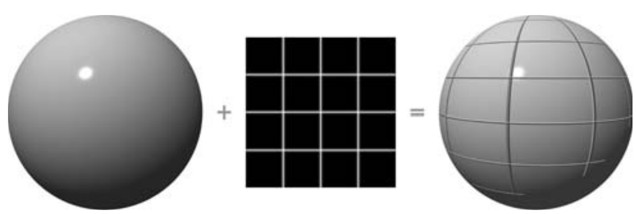


Figure 3. The original surface (left), when combined with a height map (middle), results in bump map surface (right). Bump maps are similar to DEMs in that lighter grayscale pixels create a sense of maximum relief, whereas the darker pixels have less of an effect.

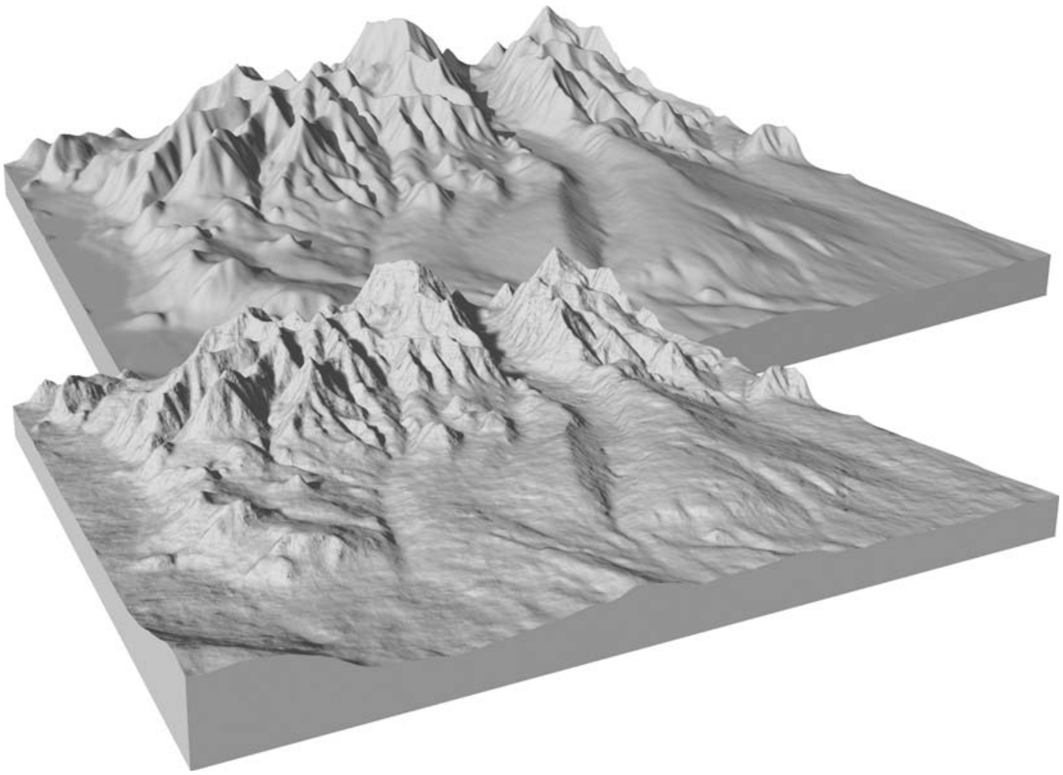


Figure 4. Surface model comparison of the foothills area, rendered with 60 m DEM data, Brooks Range, Alaska. The original model (background) was slightly textured by bump mapping (foreground). Bump mapping gives the impression of a more detailed surface through subtle shading.

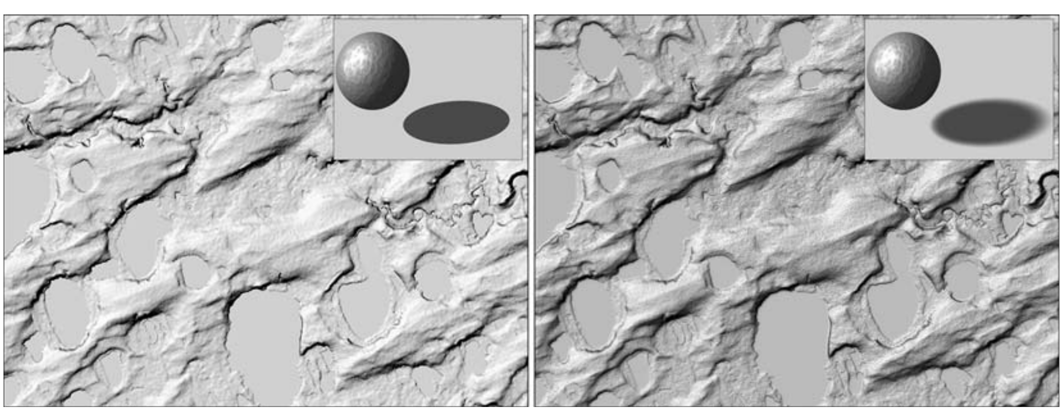


Figure 5. Hillshade images rendered without shade softening (left) and with shade softening (right). The shade softening technique smooths blocky shading that is sometimes visible in large-scale automated hillshade images. The technique is similar in principle to 3-dimensional object shadow softening, where shadows are rendered with penumbra effects (right inset), giving them a more realistic appearance than hard shadows (left inset).

softening was regulated by adjusting the neighborhood configuration (shape, size) in the focal command. In all instances, a very small neighborhood was used and care was taken to not over-generalize the softened grids. The softened grids then were hillshaded in ArcInfo, and were converted to images for editing. By using the “Curves” command in Photoshop, the darkest shades of the softened images were isolated, and all other pixels were converted to white. The adjusted softened hillshade and original hillshade were merged by multiplying the two images in Photoshop with the “multiply” feature in the layers palette. By multiplying the images, the converted white pixels of the softened image were dropped out, resulting in a softening effect restricted to only the darkest shades of the merged image. The resulting hillshade image had a more natural appearance, devoid of any jagged shading.

3-DIMENSIONAL VISUALIZATIONS

3-dimensional visualizations using DTM and ORRI data were generated within the study area to create realistic perspective images and animated fly-throughs. A variety of visualization software packages were used, including Visual Nature Studio (VNS; http://www.3dnature.com/index.html), ArcGIS 3D-Analyst (http://www.esri.com), Bryce (http://bryce.daz3d.com), and Terragen (http://www.planetside.co.uk/terragen). Although each application has unique capabilities, the majority of the visualizations were rendered using VNS. It combines powerful 3-dimensional scene-rendering abilities with traditional GIS functionality, allowing users to import a variety of geospatial data and to work in projected coordinate space. VNS fractalizes terrain objects, giving them more apparent detail by increasing fractal depth. This technique is useful when rendering realistic scenes at large scales, especially in conjunction with bump mapped textures.

The ORRI image data were draped over DTM elevation data in scene-renderings and fly-throughs. The ORRIs were desirable for texture mapping because of their high resolution (2.5m) pixel dimensions. The only limitation was their inherent grayscale color model. To overcome this limitation, the ORRIs were colorized in Photoshop using the “Magic Wand” tool. Grayscale pixels with similar values were isolated and colorized based on the surficial features that they represented. A palette for groundcover colors, derived from colors on digital field photos, was used when colorizing the images. This colorizing method proved to be an inexpensive alternative to purchasing high-resolution color imagery, and the method mimicked true landscape images surprisingly well.

ANAGLYPHS

Grayscale and value-added red-blue anaglyph stereo pairs were created to enhance visualizing and analyzing IFSAR surface data. The anaglyphs were generated using ORRI imagery in conjunction with DTM data to create two slightly differing perspectives of a single image or scene; with one perspective intended for the left eye and the other for the right eye. Upon processing, each ORRI pixel was shifted in the red band depending on the underlying elevation value of the DTM. The processed output was a RGB multi-band image with a shifted red band and identical non-shifted green and blue bands. The processed color image appears gray with purple overtones to the naked eye when a grayscale ORRI is used. When viewed through red-blue glasses, the same image displays a vertically-exaggerated surface (user defined) in three dimensions.

Value-added color anaglyphs (Figure 6) were more difficult to produce due to limitations in color choices when colorizing the ORRIs. Colors with high red, blue, or green content were not complementary to anaglyph processing, and they tended to interfere with the three dimensional visualization effect. In addition, the anaglyph glasses themselves revealed false-colors due to the red-blue film, further limiting desirable color choices. Adjusting the intensity of the anaglyph red channel in Photoshop minimized the color interference, but did not eliminate the interference problem altogether.

REFERENCES

Carter, D.L., and Galloway, J.P., 1985, Engineering-geologic maps of northern Alaska, Harrison Bay quadrangle: U.S. Geological Survey Open File Report 85-256, 47 p., scale 1:250,000.

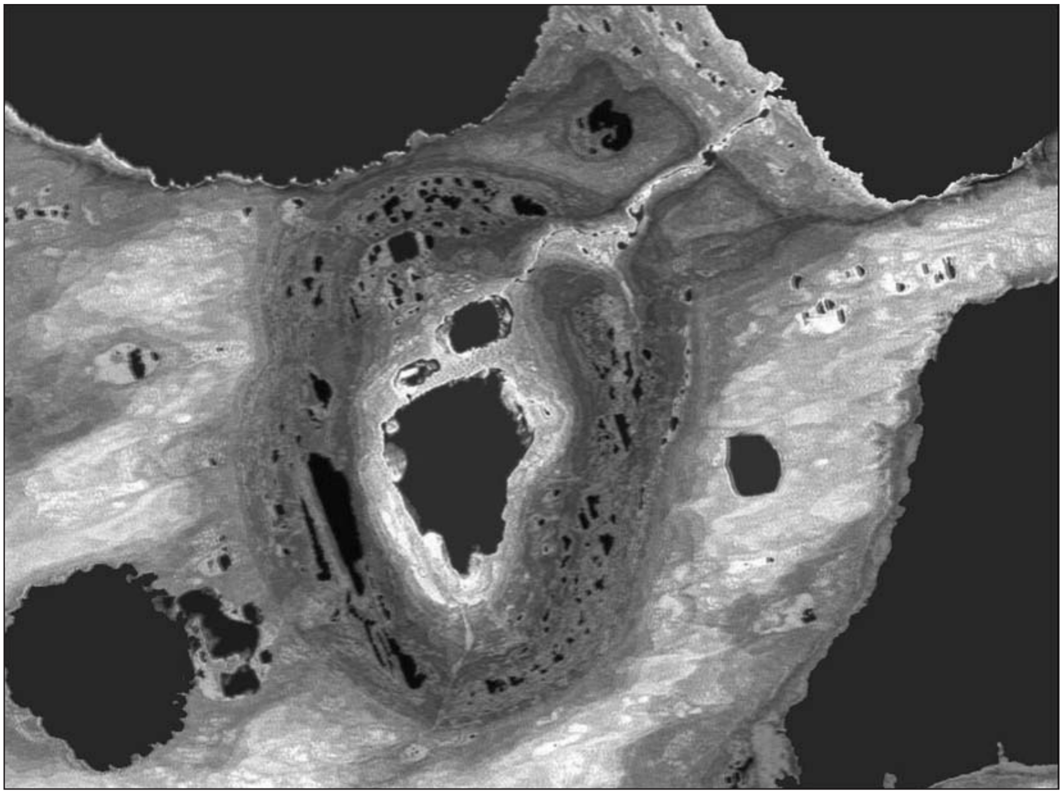


Figure 6. Color anaglyph of some of the numerous thaw lakes in the North Slope of Alaska. When viewed through red-blue glasses, the topography is seen in three-dimensions (online version only). The ORRI was value-added with an elevation color ramp, and the underlying DTM was exaggerated to accentuate surface features. This figure is represented here in black & white, but is available in color in the online version of this publication.

The Use of Topology on Geologic Maps

By Ronald R. Wahl

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

ABSTRACT

Topology, as a term used in the context of a digital representation of spatial data in two dimensions, pertains to the relationships among feature geometry objects such as points, lines and polygons. Topology helps define as well the functional relationships among geologic map objects (feature classes or sub-themes, and themes or feature data sets), the feature geometries of which provide the framework for these representations. Geologic maps contain abundant data about the locations of different kinds of rock materials, as well as relational information about the contacts between contrasting units and the nature of these contacts (for example, depositional contacts, intrusive contacts, faults, etc.). Various tools have been developed to aid in creating geologic map databases that contain the complex information and relationships expressed in geologic maps, but not all tools have been equally successful.

ESRI's ArcGIS software represents one of the latest developments in this process, and it differs from earlier tools in that a topological framework is optional. Topology in ArcGIS is expressed in the form of rules that govern the allowable relationships among geometric entities and objects, which if carefully followed, lead to digital database structures that reduce the likelihood of errors and inaccuracies and that facilitate automated analysis. The main point of this paper is that the application of topological structure in ArcGIS is beneficial to converting geologic maps into digital databases that are both rigorous and efficient.

I recommend the following kinds of topology for making geologic maps with ArcGIS:

1. A boundary line must not cross over itself or lie on top of itself.
2. Polygons must not overlap one another, must not have gaps between one another, and map unit polygon edges must be covered by boundaries (contacts and faults).
3. Points that are measurements on a fault or other structure represented by a line must be covered by the line with which they are associated.

INTRODUCTION

Dave Viljoen (Geological Survey of Canada) in 1997 presented a paper about geologic map topology at the first Digital Mapping Techniques workshop. He said, "Topology is the [expression of] relationships that can exist among the geometries of spatial features…(and) is a major step towards minimizing the number of points, lines, and polygons to digitally capture and manage a geologic map." Data accuracy and validation are enhanced by the proper use of topological rules. These concepts are still valid.

In addition to the making of geologic maps, topology enhances analysis of the features on a geologic map in the following general cases:

- Polygons that outline a part of an image
- Point to polygon (is a point on the inside, outside, or on the boundary of the polygon?)
- Polygon adjacency
- Buffer distance from a line (or arcs of a type). An example of this is the following query from a geologic map: "Show as a polygon the region where every point on the constructed polygon boundary is ten miles from the San Andreas fault system."

HISTORY OF TOPOLOGY IN GEOLOGIC MAPS

Long before geographic information systems (GIS), geologic mappers and cartographers were less concerned with questions about topology than they are today. For example, mappers colored early geologic maps by hand mostly with watercolors. These maps had topologic problems. Lines bounding the areas on the maps representing rock outcrops did not necessarily close. Application of color with a brush meant that the color sometimes went outside these lines. The only topological question that was asked was "Are any adjacent polygons of the same color?"

In the 1950s, the basic geologic map-making process was the same, but by then, the materials used to make geologic maps had changed. Geologic maps were made using many scribe coats and peel coats each of which contained information primarily separated out by color. Mappers and cartographers proceeded with great care to insure that all of the layers of mylar that were used were accurately registered with every other layer. Minor misregistrations meant that the old maps weren't accurate under close inspection. Contacts and faults features overshot or undershot other lines on the final geologic map, and this was particularly true for lines in other map themes like hydrography. Rock outcrop color fills went across contact and fault lines.

In the late 1970s and early 1980s, an Israeli company, SCITEX, brought equipment to the United States that originally was used to make patterns for wallpaper and cloth. I attended the 11th Annual Conference on Computer Graphics and Interactive Techniques, at Anaheim, CA in April 1984, where SCITEX showed equipment that was able to scan and then print maps. These scans could be digitized and subsequently edited to make map separates that regular printing presses could use. The USGS soon adopted this technology to make maps, including geologic maps. The scanned map was digitized within the SCITEX system. This process forced the map makers to ensure line closures to make closed areas (or polygons) for rock outcrops and to ensure that contacts and faults aligned precisely with rock outcrop boundaries.

Selner and others (1986) wrote a software package named GSMAP for the IBM PC environment to digitize and then print geologic maps and diagrams on Hewlett-Packard Graphics Language pen plotters. The USGS (author unknown) developed a translator to change GSMAP files to SCITEX format. Now, USGS geoscientists could digitize their own maps and deliver their digital maps to the SCITEX. However, the SCITEX installations were so complex they required many technicians to operate it and it was a "black box" for many geologists. In general, geologists still had no knowledge about the topology requirements for their maps.

In the mid-to-late 1980s, ARC/INFO from ESRI was slowly being introduced to USGS geologists. ESRI, convinced of the useful of the SCITEX system, wrote a series of commands to exchange files with the SCITEX. In the process, USGS geologists discovered intersecting lines with dangles (both over and under shoots), pseudo-nodes, and other new problems identified by the topology of an ARC/INFO coverage. Geologic maps improved from the cartographic standpoint because of the enforced topology. These improvements included:

- no over or undershoots when lines intersect, leading to
- no unclosed polygons so that polygon fills are exact, and
- line sharing is exact among themes.

While ARC/INFO was the old name for ESRI's premier product, the name ArcInfo Workstation, included with Arc GIS, will be used below.

For ArcView 3.x, theme means one shape file set and the associated attributes. Theme as used in ArcGIS is a synonym for a feature data set. The basic idea for both ArcView and ArcGIS is that lines used for contacts or faults are not exactly the same lines that are stored for polygon edges. ArcInfo Workstation coverages can have more than one attribution (or theme) for a line or polygon so that the actual storage of lines is shared among the various map themes.

ArcGIS introduces a new view of topology. Successful implementations of this topology can lead to geologic map databases that are more internally consistent and better prepared for analysis and comparison.

TOPOLOGY

The word "topology" will be used in the following discussion to mean a topologic paradigm. For ArcGIS, a topology is: "…a set of integrity rules and software tools that defines the behavior of spatially integrated geographic features and feature classes" (ESRI, 2002a). While old ARC/INFO and the Geographic Resources and Support System (GRASS http://grass.itc.it) each had a topologic paradigm, ArcView, through version 3.3 did not.

Arc-Node Topology

The topology model used in the Spatial Data Transfer System (SDTS) is essentially the topology model used in ArcInfo Workstation. SDTS topology including the topological Vector Profile (TVP) is extensively described in USGS, 1991. In ArcInfo Workstation, the following simplified definitions are compared to those from SDTS TVP.

ArcInfo Workstation Vector or GT-Complete Chains:

- Have a line with a node at each end
- Have zero or more vertices
- Have direction—"From Node" and "To Node"
- Can form polygons. If polygons are constructed, each line is attributed with an entry that shows the polygons to the left and right of the line as the line is traversed in the direction of the "To Node" from the "From Node". A line with a sense of direction is properly called a vector.

ArcInfo Workstation Polygons or GT-Polygons:

- Are a closed figure made of one or more arcs or GT-complete chains
- Support islands (polygons completely internal to a larger polygon).

Label Points:

- ArcInfo Workstation can have label points in a polygon coverage,, but otherwise it has no point topology.

Open GIS Consortium Geometry and Topology

The Open GIS Consortium (OGC), of which ESRI is one of many members, has agreed upon standards for the collection of digital data describing spatial objects called features (OGC, 1999a) and has specified geometries for these features in OGC, 1999b. ESRI, in implementing ArcGIS, has built upon the geometries in OGC, 1999b.

I first present a comparison of the geometries of ArcInfo Workstation to ArcGIS, and between the OGC and ArcGIS geometries (Table 1).

Topology—OGC and ArcGIS

OGC geometries

Topology is implicit in the OGC geometry descrip-

Table 1. Comparison of geometries for OGC, ArcInfo Workstation, and ArcGIS.

ArcInfo Workstation, and ArcGIS:

ArcInfo Workstation	ArcGIS	Notes
Coverage feature classes	Geodatabase feature classes	Each geodatabase can contain one or more feature data sets. Each feature data class can contain one or more feature classes
Tic	Point*	
Bnd	*	*The spatial extent and coordinate system are defined as part of the dataset's spatial reference. The tics could optionally be managed as a point feature class.
Arc	Polyline	
Node	Point**	**With topological relationships to other line and
polygon feature classes.		
Point	Point	
Polygon	Polygon	
Polygon label	Point***	***A label location with a topological association to its polygon.
Route System	Line with measures	
Region	Polygon	
Annotation	Annotation	
NA	Dimension	
NA	Custom features****	****The best strategy is to use generic feature classes whenever possible to ensure support for open, multipurpose, standards-based data. However, the geodatabase is extensible and it is possible to add custom feature classes.

From: ESRI, 2002a

OGC geometries and ArcGIS:

OGC Geometries (OGC, 1999b)	ArcGIS Geometries (ESRI, 2002a)
Point	Point
Line, LineString	Polyline
LinearRing	Polygon Boundary
Polygon (one or more linear rings)	Polygons with islands

tions but no basic topology model exists. Topologic relationships are determined from the Dimensionally Extended Nine-Intersection Model (DE9-IM) shown in Table 2.

The matrix in Table 2 is difficult to implement directly in a GIS, so the OGC has specified the following methods (or program functions) based on DE9-IM model for use in computer programming languages. They are:

- Equals
- Disjoint
- Intersects
- Touches
- Crosses
- Within
- Contains
- Overlaps
- Relate

ArcGIS—Topology in a Geodatabase

ESRI has embraced OGC geometries and has developed topological rules and software to implement them as a part of the geodatabase structure. In a geodatabase, spatial relationships among feature classes in a feature dataset are defined by topology, and use specific rules for each basic geometry type. A map maker can choose to create topology for feature classes in a feature dataset. The primary spatial relationships that one can model using topology are adjacency, coincidence, and connectivity.

There are three types of topology available in the geodatabase: geodatabase topology (over 20 topology rules), map topology, and geometric network topology. Map topology is a dynamic topology derived from the creation of polygons from one or more polyline feature classes in a feature data set. Network topology uses a feature data set with one or more feature data classes that have intersecting lines, and this topology allow the tracing of a designated path through these lines. An example of network topology would be a bus route through a city road network. I will concentrate only on geodatabase topology in this paper. For a complete discussion of the topology rules under ArcGIS, see ESRI, 2002a and ESRI, 2002b,and ESRI's on-line documentation.

Line Sharing

When geodatabase topology is appropriate, geodatabases create a place for the implemented topological rules in a separate feature class in a feature data set. When the chosen topology rules are in force, an individual feature in a feature class can be chosen separately even though line work is shared among a number of feature classes. This is quite different than the ArcInfo Workstation model. In ArcInfo Workstation, only one set of lines exist in a coverage no matter how many attributes each line may carry. Even though lines can apparently be broken into pieces when a route is built from a line coverage (using dynamic segmentation), lines are inviolate in such a case and extra attribution is added in additional tables to contain the route construct.

Polygons in a Geodatabase

Unlike ArcInfo Workstation, only closed linear rings form polygons under the OGC definitions (see above and OGC, 1999b). This means that if two polygons are adjacent, they do not share adjacent edges, which in turn means that the polygons can overlap (one polygon is drawn over an adjacent polygon) or have gaps between the two nearly adjacent edges. This gap looks like a polygon between two other polygons, but this "gap polygon" is not directly known as a polygon by the geodatabase or ArcGIS. An island (one polygon wholly contained in an-

Table 2. The OGC Dimensionally Extended Nine-Intersection Model.

	Interior	**Boundary**	**Exterior**
Interior	$dim(I(a) \cap I(b))$	$dim(I(a) \cap B(b))$	$dim(I(a) \cap E(b))$
Boundary	$dim(B(a) \cap I(b))$	$dim(B(a) \cap B(b))$	$dim(B(a) \cap E(b))$
Exterior	$dim(E(a) \cap I(b))$	$dim(E(a) \cap B(b))$	$dim(E(a) \cap E(b))$

From OGC (1999b) Table 2.1—The DE-9IM

Where:
a is a geometry, and b is another geometry
I is the interior of a geometry
B is the boundary of a geometry
E is the exterior of a geometry
dim is the dimensionality of the resulting geometry
$\cap$ is the symbol for intersection.

other polygon requires, in theory, two linear rings for the containing polygon. The outer polygon tracks clockwise and the inner polygon tracks counter-clockwise. A third linear ring tracking clockwise is then necessary for the inner polygon. Then, if polygons have edges that need to carry attribution (as in geologic mapping), another feature class (rock outcrop boundaries for example) should cover the polygon edges exactly. Topological rules to accomplish these tasks are among those provided in ArcGIS, (ESRI, 2002a, and ESRI, 2003). These rules are summarized on a PDF poster (ESRI, 2002b).

OTHER GIS SYSTEMS WITH OGC TOPOLOGY

A number of other GIS systems and RDBMs are equipped to implement geometry and (perhaps) topology for vector features. I list some of these below.

GIS Packages

- Vivid Solutions—JUMP (Java Unified Mapping Platform) http://www.vividsolutions.com/products.asp?catg=spaapp&code=jump.
- Vivid Solutions—JTS library (Java Topology Suite) http://www.vividsolutions.com/products.asp?catg=spaapp&code=jts.
- PostGIS (Based on PostgreSQL and JTS) http://postgis.refractions.net and http://www.postgresql.org.
- GRASS5.7—Based on PostGIS.

RDBMS Packages with Geometry

The Relational Database Management Systems (RDBMS) below can store geometry of the OGC type, but do not always have a topology software suite for use with geometry in the RDBMS.

- Oracle Spatial can be modified with the Laser-Scan Radian topology software http://www.oracle.com/index.html and http://www.laser-scan.com/technologies/radius/radius_topology.
- Informix (DB2)—Data Blades can store geometry http://www-306.ibm.com/software/data/informix.
- MySQL—Stores geometry http://www.mysql.com.
- PostgreSQL—Stores geometry and is an Object-RDBMS.

CONCLUSIONS

Topology greatly aids the validation of spatial data, which leads to increased accuracy and rigor among feature classes. It also means easier geodatabase maintenance tasks, and that the geodatabase is physically smaller on a disk. The speed and capability with which analyses can be made is enormously increased.

I want to write a word about splined curves and other interpolations to give smooth curves. Mathematically, splines revolve around the process of taking derivatives of a function. This is an inherently "noisy" process. By noisy, I mean that the result of adding together derivatives of a function is an inaccurate process, the higher the derivative order the less accurate the result. Further, if the same vertices are interpolated from opposite sides, the direction of curvature from these opposite sides will be in the opposite direction. The net result is that vectors made from vertices tracked from opposite directions along the same path will not lie on top of one another. As a consequence, spline-generated curves for adjacent polygon edges cannot be subject to topological rules.

Recommendations

I recommend that the geologic community study implementation of topology in ArcGIS for use with geologic mapping and recommend minimal standard topology rules. I further recommend that this study consider in addition: network topology, and relationship classes.

Recommended Topology Rules

I have found the following topology rules to be useful and suggest that these rules form a basis for standard usage of topology rules:

1. Polygons
   a. Adjacent polygons must not partially overlay (or overlap) each other.
   b. Adjacent polygons must not have gaps between them.
2. Arcs (Edges)
   a. An line must not self-overlap or cross over itself—like a freeway off-ramp.
   b. Polygon boundaries must be covered by contacts and faults.
3. Points
   Structure measures associated with a contact or fault must be covered by the appropriate line, modified from ESRI, 2002b.

REFERENCES

Note: Some of the ESRI documents referenced below are best found on a computer with ArcGIS installed. This insures that the reader will have the version that coincides with the version to which the reader has access. The top of the installation is the directory (folder) named "arcgis".

ESRI, 2003, ArcGIS: Working with geodatabase topology: An ESRI White Paper, Environmental Systems Research Institute, Inc., Redlands, CA, 23p., accessed at http://support.esri.com/index.cfm?fa=knowledgebase.whitepapers.viewPaper&PID=43&MetaID=524.

ESRI, 2002a, ArcGIS 8.3 geodatabase topology rules: Environmental Systems Research Institute, Inc., Redlands, CA, accessed at http://arcgis/arcexe83/Documentation/topology_rules_poster.pdf.

ESRI, 2002b, What is ArcGIS?: What_is_ArcGIS.pdf, Document file with ArcGIS, Environmental Systems Research Institute, Inc., Redlands, CA, 80p., accessed at http://arcgis/arcexe83/Documentation,.

Open GIS Consortium (OGC), 1999a, The OpenGIS Abstract Specification Topic 5: Features, Version 4: OpenGIS Project Document Number 99-105r2.doc, accessed at http://www.opengis.org/docs/99-105r2.pdf.

Open GIS Consortium, 1999b, OpenGIS Simple Features Specification For SQL, Revision 1.1: OpenGIS Project Document 99-049, accessed at http://www.opengis.org/docs/99-049.pdf.

Selner, G.I., Taylor, R.B., and Johnson, B.R., 1986, GSDRAW and GSMAP version 3.0; prototype programs for the IBM PC or compatible microcomputers to assist compilation and publication of geologic maps and illustrations: U.S. Geological Survey Open-File Report 86-447-A, 53 p.

USGS, 1991, Logical Specifications in SDTS_standard_oct91, accessed at http://mcmcweb.er.usgs.gov/sdts/SDTS_standard_oct91/part1.html#p23.

Viljoen, David, 1997, Topological and thematic layering of geological map information: Improving efficiency of digital data capture and management, in: D.R. Soller, ed., Digital Mapping Techniques ‘97—Workshop Proceedings: U.S. Geological Survey Open-File Report 97-269, p. 15–22.

Geology, Stewart River Area (Parts of 115 N/1,2,7,8 and 115-O/2-12), Yukon Territory

By S. P. Gordey, J.J. Ryan, and R.B. Cocking

Geological Survey of Canada
101-605 Robson Street
Vancouver, B.C., Canada V6B 5J3
Telephone: (604) 666-2116
Fax: (604) 666-1124
e-mail: sgordey@nrcan.gc.ca

INTRODUCTION

Geological mapping in the Stewart River area (NTS 115 N, O, in the Yukon along the Alaska border) began as part of the Ancient Pacific Margin NATMAP Project. Initiated by the Geological Survey of Canada, Yukon Geological Survey, and British Columbia Geological Survey Branch, the NATMAP Project seeks to improve understand of the composition, relationships, and metallogeny of terranes lying between the ancestral North American margin and those known with more certainty to be tectonically accreted (Thompson and others, 2000; Colpron and others, 2001). The Stewart River component (Figure 1) focuses on the Yukon-Tanana terrane (Mortensen, 1990 and 1992), comprising complexly deformed, mostly (?) Paleozoic meta-igneous and metasedimentary rocks. The final two years of the Stewart River Project are funded under the Geological Survey of Canada's Northern Resources Development Program.

The objective of the Stewart River Project is to investigate the stratigraphic, structural, and tectonic history, and the economic framework, of this large tract of Yukon-Tanana terrane by mapping about 2/3 of the area over a four year period. Geology is being interpreted in light of new geophysical data collected in this area under the Targeted Geoscience Initiative (Figure 2; Shives and others, 2002). Concurrent surficial geological studies were aimed at understanding the Quaternary history and setting of the numerous placer gold deposits in the region (e.g. Jackson and others, 2001, 2002; Rotheisler and others, 2003).

In summer 2003, gaps in the previous mapping were bridged and the mapping extended to cover about eleven 1:50 000 scale map areas (see Figure 2). These new data and previous work in surrounding areas (e.g. Bostock, 1942; Tempelman-Kluit, 1974; Mortensen, 1996) will be synthesized into a new geological map of the Stewart River map area (1:250,000 scale; 115N-O).

Access into the heart of the Stewart River area is afforded by boat along the Yukon and Stewart rivers and by truck on placer mining roads, many of which extend south from Dawson. Fieldwork in 2000-2003 included foot traverses from small camps mobilized along these routes and from helicopter or fixed-wing supported camps in more remote areas. All-terrain vehicles were used on placer mining access roads along Thistle, Kirkman, Henderson, Black Hills and Maisy Mae creeks and the Sixty Mile River. Helicopter spot checks were used to fill in widely separated outcrops in the southwest part of the map area where foot traverses or fly camps were impractical. Bedrock mapping is hampered by a deep (~1 m) soil veneer, thick gravel, and loess deposits in valley bottoms, and by dense cover of forest, moss, and lichen. The detailed aeromagnetic and gamma-ray surveys (Shives and others, 2002) are an effective aid to bedrock mapping in this poorly exposed, unglaciated terrain.

GEOLOGICAL FRAMEWORK

The Stewart River area is underlain by twice-transposed, amphibolite-facies gneiss and schist of mostly (?) Paleozoic age. These are intruded by younger plutonic rocks and overlain by Upper Cretaceous volcanic rocks and local occurrences of Lower Cretaceous conglomerate. The reader is directed to Ryan and Gordey (2001a, b; 2002a, b) and Ryan and others (2003) for a more comprehensive description of the geology.

Metasiliciclastic rocks are widespread, and dominated by psammite and quartzite, with lesser pelites and rare conglomerate. These were thought to be as old as late Proterozoic (e.g. Tempelman-Kluit, 1974); however, preliminary detrital zircon geochronology and geochronology for plutonic rocks suggest a middle Paleozoic age (M. Villeneuve, Geological Survey of Canada, in preparation). Intermediate to mafic composition amphibolite interdigitates with, and lies stratigraphically above, the siliciclastic rocks. Although intensely tectonized, heterogeneous compositional layering and local vestiges of primary textures in the amphibolite, such as breccia clasts

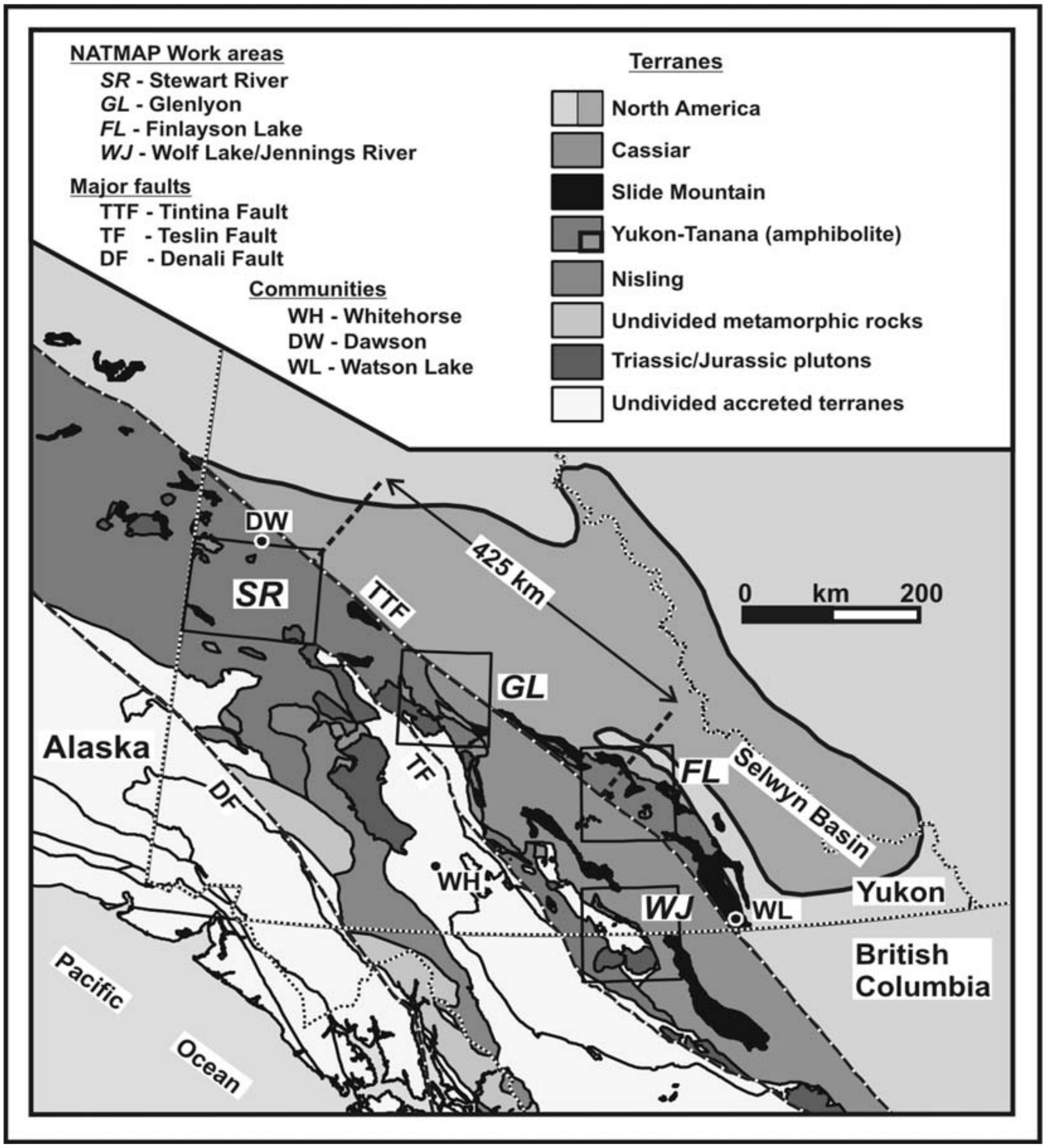


Figure 1. Regional tectonic setting of the Stewart River area.

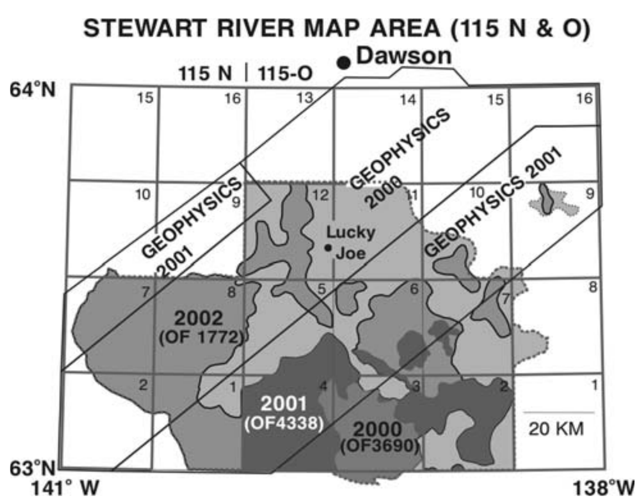


Figure 2. Cumulative progress of bedrock mapping in the Stewart River area (reference to previous Geological Survey of Canada open file reports are in parentheses). The grid outlines 1:50 000 NTS boundaries. Helicopter-borne detailed aeromagnetic and gamma-ray surveys (gray areas denoted as "Geophysics <date>") were flown in 2000 and 2001.

and pillow selvages, indicate derivation from volcanic and volcaniclastic rocks. Marble horizons occur within the amphibolite, as well as the siliciclastic rocks. In turn, dark carbonaceous quartzite, metapelite and minor marble of the Nasina assemblage, markedly sparse in volcanic-derived material, lies structurally above and/or may be partly equivalent to the aforementioned metaclastic rocks. Abundant orthogneiss bodies with diorite, tonalite, granodiorite, monzogranite and granite protoliths, intrude the above assemblages. Some are Devono-Mississippian in age, whereas others are known to be Permian. For many others, the age is probably one of these, yet remains unclear. The tectonic significance of ultramafic and gabbroic rocks that lie near the boundary between the siliciclastic and metavolcanic-metavolcaniclastic successions is also unclear.

An extensive area of likely Permian, low to medium grade muscovite-quartz and chlorite-quartz schist in the western part of the map area, correlated by Tempelman-Kluit (1974) with the Klondike Schist (McConnell, 1905) is derived from a combination of volcanic, volcaniclastic and plutonic rocks. Southeast of the White River this succession may lie beneath a low-angle fault. To the northwest, contact relationships are uncertain. East of Ladue River these rocks are overlain by relatively unstrained, chlorite-altered intermediate to mafic volcanics, of unknown but possibly Permian age.

In summary, the extensive metaplutonic and metavolcanic rocks represent two periods of arc activity. The older arc, built upon a siliciclastic foundation, largely comprises Devono-Mississippian amphibolite associated with coeval widespread tonalitic orthogneiss that formed its subvolcanic intrusive complex. A Permian arc, built upon the previous, is represented by granitic orthogneiss and coeval metavolcanics.

STRUCTURE

The Paleozoic rocks in the field area exhibit a regional foliation (ST), characterized by high-strain transposition of layering in the gneiss and schist, with abundant intrafolial isoclinal folds that are commonly rootless. The intensity of strain within the regional foliation locally grades to mylonite. Primary compositional layering (S0) in metasedimentary rocks, unit contacts, and a pre-existing foliation (S1) can be traced around closures of the transposition folds, indicating that they are at least F2 structures. F2 deformation appears to accompany the regional metamorphism, and preliminary geochronological results indicate that this happened during the mid-Permian (M. Villeneuve, Geological Survey of Canada, in preparation). The F2 folds are generally recumbent to shallowly inclined, close to isoclinal, long-wavelength structures. They commonly lack an axial planar foliation, and their axes parallel a regional extension lineation (L2). This relationship helps distinguish F2 and F3 folds, which can have very similar style. The latter are open, moderately inclined (but varying from shallow to steep), shallowly plunging structures, that have weak axial-planar fabric where developed in schistose layers, and have no associated extension lineation. The map area is also affected by faults of varying significance. Most of these could not be observed directly, but are interpreted from changes in rock type and/or structural grain; some are also well delineated by prominent physiographic and aeromagnetic lineaments. Locally, fault breccia and slickensides provide direct evidence of fault contacts.

ECONOMIC GEOLOGY

One of the more significant findings is that parts of the area are dominated by a mid-Paleozoic volcano-plutonic arc (?) complex with implied potential for VMS type mineralization. In the Finlayson Lake area (Figure 1), originally contiguous with the Stewart River area (allowing for 425 km of late Mesozoic-Tertiary dextral offset), correlative mid-Paleozoic strata host massive sulphide mineralization in both felsic (e.g., Kudz Ze Kayah and Wolverine Lake deposits; Murphy (1998, and references therein), Piercey and others, 2001) and mafic (Fyre Lake deposit; Foreman (1998)) metavolcanic sequences. It should be noted that primary geochemical (e.g., alteration), structural, and lithological signatures may be strongly modified by the amphibolite facies metamorphism and high state of strain in the Stewart River area.

The Lucky Joe occurrence was explored in 2003 by Kennecott Exploration. Two large strong parallel geochemical trends defined by high soil values of Cu and Au, with associated Mo and Ag, have been identified (see press release at http://www.copper-ridge.com). The origin of the occurrence is obscured by complex structure and metamorphism. Cu-Au porphyry, Fe-oxide Cu-Au, or sediment-hosted Cu deposit models have all been suggested. A metallogenic study now underway (Jan Peter (GSC)) is aimed at identifying the deposit type and its origin. The Lucky Joe represents a new type of potentially large occurrence within Yukon-Tanana terrane.

In Yukon and Alaska, mid-Cretaceous (105-90 Ma) and Late Cretaceous (70-65 Ma) plutons and their country rock are prospective targets for intrusion-related gold deposits (e.g., Hart and others, 2000). Undeformed granite-syenite stocks, such as near Mt. Stewart, possibly of Cretaceous or Tertiary age, could be prospective. Although perhaps of less significance, Early Jurassic plutons are known to host Au±Cu rich shear zones, stockworks and skarns in Alaska (Newberry, 2000) as well as central Yukon (e.g. Minto deposit, Tafti and Mortensen, 2004). Other plutonic bodies show evidence of significant strain, are all pre-Early Jurassic (Paleozoic) in age, and regionally unproductive. The source of gold leading to significant placer deposits in many drainages (e.g. Thistle, Kirkman, Barker, Scroggie, Black Hills, Maisy Mae and Henderson creeks) remains enigmatic. For example, Dumula and Mortensen (2002) suggest undiscovered intrusion-related gold as a placer source within the Thistle basin, on the basis of placer gold composition. However, Mesozoic plutonic rocks are rare within this drainage. They also indicate that as yet undiscovered sources for placer gold in the Eureka Dome or Henderson Dome area are of epithermal origin. Rotheisler and others (2003) suggest two separate, as yet unidentified lode gold occurrences as sources for placer deposits in the Scroggie Creek basin.

CARTOGRAPHY

Cartography for this map was done in ArcInfo Workstation version 8.3, with CorelDraw 11 and Excel primarily supporting the work. Map information was processed, as follows:

- DEM—the DEM was supplied by the Yukon Government as individual 1:50,000 scale sections, and then compiled into a single dataset. The DEM was converted into a hillshade to accentuate the topographic features. The DEM data is available at http://www.renres.gov.yk.ca/pubs/rrgis/data/data_desc/90m_dem.html.
- Geology—the map area was mapped at a scale of 1:50,000 and scaled to 1:100,000. This dataset was compiled over 4 years of extensive fieldwork. New geological lines were digitized each year, then symbolized. Point data was entered into PDAs in the field and imported first into Excel, and then into ArcInfo. In many cases the point data was far too dense for the plotting scale and was weeded significantly. Mines data was extracted from the Yukon Minfile database (Deklerk, 2003).
- Contours—this National Topographic Data Base dataset, as was all the topographic data, was purchased from Geomatics Canada (http://www.cits.rncan.gc.ca) at a scale of 1:50,000 and compiled into a single 1:100,000 dataset. Every second contour line was unsymbolized to prevent the background data from obscuring the main theme, the geology.
- Surround—the legend was generated using a custom set of AMLs designed to incorporate the cartographic standards of the Geological Survey of Canada. Many other features on the map such as the border, symbology, logos, location maps, scalebar, and titlebar (among others) were codified in these AMLs. These AMLs called GEMS (**Ge**ological **M**apping **S**ystem) can be downloaded at http://www2.nrcan-rncan.gc.ca/ess/carto/gems_e.asp.

REFERENCES

Bostock, H.S., 1942, Ogilvie, Yukon Territory: Geological Survey of Canada, Map 711A, scale 1:250,000.

Colpron, M., and Yukon-Tanana Working Group, 2001, Ancient Pacific Margin - an update on stratigraphic comparison of potential volcanogenic massive sulphide-hosting successions of Yukon-Tanana Terrane, northern British Columbia and Yukon, *in* Yukon Exploration and Geology 2000: Exploration and Geological Services Division, Yukon, Indian and Northern Affairs Canada, p. 97–110.

Deklerk, R. (compiler), 2003, Yukon Minfile 2003 - A mineral occurrence database: Yukon Geological Survey, CD-ROM.

Dumula, M.R. and Mortensen, J.K., 2002, Composition of placer and lode gold as an exploration tool in the Stewart River map area, western Yukon, *in* Emond, D.S., Weston, L.H., and Lewis, L.L., eds., 2001, Yukon Exploration and Geology: Exploration and Geological Services Division, Yukon Region, Indian and Northern Affairs Canada, p. 87-102.

Foreman, I., 1998, The Fyre Lake project 1997: Geology and mineralization of the Kona massive sulphide deposit, *in* Yukon Exploration and Geology 1997: Exploration and Geological Services Division, Yukon Region, Indian and Northern Affairs Canada, p. 105-113.

Gordey, S.P. and Ryan, J.J., 2003, Geology, Stewart River area (parts of 115 N/1, 2, 7, 8 and 115 O/2 - 7, 12), Yukon Territory: Geological Survey of Canada, Open File 1772, scale 1:100,000.

Hart, C.J.R., Baker, T. and Burke, M., 2000, New exploration concepts for country-rock-hosted, intrusion-related gold systems: Tintina gold belt in Yukon, *in* Tucker, T. and Smith, M.T., eds., The Tintina gold belt: concepts, exploration and discoveries, Special Volume 2: British Columbia and Yukon Chamber of Mines, Cordilleran Roundup, January, 2000, p. 145-172.

Jackson, L.E., Shimamura, K., and Huscroft, C.A., 2001, Late Cenozoic geology, Ancient Pacific Margin NATMAP Project, report 3: a re-evaluation of glacial limits in the Stewart River basin of Stewart River map area, Yukon Territory: Geological Survey of Canada, Current Research 2001-A3, 17 p.

Jackson, L.E., Froese, D.G., Telka, A.M., Westgate, J.A., Preece, S.J., Storer, J.E., and Huscroft, C.A., 2002, Late Cenozoic geology, Ancient Pacific Margin NATMAP Project, report 5: paleoecology and proxy climatic change records, south Klondike placer region, Yukon Territory: Geological Survey of Canada, Current Research 2002-A2, 16 p.

McConnell, R.G., 1905, Report on the Klondike Gold Fields: Geological Survey of Canada Annual Report, v.XIV, part B.

Mortensen, J.K., 1990,Geology and U-Pb geochronology of the Klondike District, west-central Yukon Territory: Canadian Journal of Earth Sciences, v. 27, p. 903–914.

Mortensen, J.K., 1992, Pre-mid-Mesozoic tectonic evolution of the Yukon-Tanana terrane, Yukon and Alaska: Tectonics, v. 11, no. 4, p. 836-853.

Mortensen, J.K., 1996, Geological compilation maps of the northern Stewart River area, Klondike and Sixtymile Districts (115N/15, 16, 115O/13, 14 and Parts of 115 O/15, 16): Indian and Northern Affairs Canada, Northern Affairs, Yukon Region, Open File 1996-1(G), scale 1:50,000.

Murphy, D.C., 1998, Stratigraphic framework for syngenetic mineral occurrences, Yukon-Tanana terrane south of Finlayson Lake: a progress Report, *in* Yukon Exploration and Geology, 1997: Exploration and Geological Services Division, Yukon Region, Indian and Northern Affairs Canada, p. 51–58.

Newberry, R., 2000, Mineral deposits and associated Mesozoic and Tertiary igneous rocks within the interior Alaska and adjacent Yukon portions of the 'Tintina gold belt': a progress report, *in* Tucker, T. and Smith, M.T., eds., The Tintina gold belt: concepts, exploration and discoveries, Special Volume 2: British Columbia and Yukon Chamber of Mines, Cordilleran Roundup, January, 2000, p. 59-88.

Piercey, S.J, Paradis, S., Murphy, D.C. and Mortensen, J.K., 2001, Geochemistry and paleotectonic setting of felsic volcanic rocks in the Finlayson Lake volcanic-hosted massive sulfide district, Yukon, Canada: Economic Geology and the Bulletin of the Society of Economic Geologists, v. 96, no. 8, p. 1877-1905.

Rotheisler, P.N; Jackson, L.E, Jr. and Hicock, S.R., 2003, Late Cenozoic geology, Ancient Pacific Margin NATMAP Project, report 6: glacial limits, Middle Pleistocene sediments, and placer gold in the Scroggie Creek basin, south Klondike placer region, Yukon Territory: Geological Survey of Canada, Current Research, no. 2003-A1, 8 p.

Ryan, J.J. and Gordey, S.P., 2001a, New geological mapping in Yukon-Tanana terrane near Thistle Creek, Stewart River map area, Yukon Territory: Geological Survey of Canada, Current Research 2001-A2, 18 p.

Ryan, J.J. and Gordey, S.P., 2001b, Geology of the Thistle Creek area (115-O/3), Yukon Territory: Geological Survey of Canada, Open File 3690, scale 1:50,000.

Ryan, J.J. and Gordey, S.P., 2002a, Bedrock geology of Yukon Tanana terrane in southern Stewart River map area, Yukon Territory: Geological Survey of Canada, Current Research 2002-A1, 11 p.

Ryan, J.J. and Gordey, S.P., 2002b, Geology, southern Stewart River area, Yukon Territory (Parts of 115 J/14, 115-O/ 2,3,4,5,7): Geological Survey of Canada, Open File 4338, scale 1:100,000.

Ryan, J.J., Gordey, S.P., Glombick, P., Piercey, S.J. and Villeneuve, M.E., 2003, Update on bedrock geological mapping of the Yukon-Tanana terrane, southern Stewart River map area, Yukon Territory: Geological Survey of Canada, Current Research, no. 2003-A9, 7 p.

Shives, R.B.K., Carson, J.M., Ford, K.L., Holman, P.B., Gordey, S.P., and Abbott, G., 2002, Airborne multisensor geophysical survey, Stewart River area, Yukon Territory, Phase 1 and 2 (Parts of 115 N, O and 116 B): 120 digital images of 1: 50,000 (110) and 1:250,000 (10) scale colour interval maps, on CD-ROM, in Portable Document Format (PDF): Geological Survey of Canada, Open File 4311 (also Yukon Exploration and Geological Services Division Open File 2002-17D).

Tafti, R. And Mortensen, J.K., 2004, Early Jurassic porphyry(?) copper (-gold) deposits at Minto and Williams Creek, Carmacks copper belt, western Yukon, *in* Emond, D.S. and Lewis, L.L., eds., 2003, Yukon Exploration and Geology: Yukon Geological Survey, p. 289-303.

Tempelman-Kluit, D.J., 1974, Reconnaissance geology of Aishihik Lake, Snag and part of Stewart River map-areas, west-central Yukon: Geological Survey of Canada, Paper 73-41, 97 p. (including maps).

Thompson, R.I., Nelson, J.L., Paradis, S., Roots, C.F., Murphy, D.C., Gordey, S.P., and Jackson, L.E., 2000, Ancient Pacific Margin NATMAP Project, year one: Geological Survey of Canada, Current Research, 2000-A1, 8 p.

Surface Terrain of Indiana—a Digital Elevation Model

By Robin Rupp

Indiana Geological Survey
Indiana University
611 N. Walnut Grove
Bloomington, IN 47408
Telephone: (812) 855-3801
Fax: (812) 855-2862
e-mail: rrupp@indiana.edu

INTRODUCTION

The surface terrain dataset of Indiana was created using part of the U.S. National Elevation Dataset (NED) and newly-created digital elevation models (DEMs). New digital elevation data was processed into new DEMs and served to replace 275 of the 710 7.5-minute quadrangles within the Indiana portion of the NED.

This report describes the processes of 1) creating new DEMs, 2) merging the new data and the NED into a new dataset for Indiana, and 3) creating TIFF images from the final dataset. The revised digital elevation model for Indiana and parts of the surrounding states is available on a CD-ROM (Brown and others, 2004) as raster data in ESRI Grid format. This report also provides the background for the Indiana Geological Survey Poster05 (Berry and others, 2004; Figure 1) and the forthcoming terrain image series.

NATIONAL ELEVATION DATASET

The National Elevation Dataset (NED) is a relatively new raster product created and maintained by the U.S. Geological Survey (USGS). The data and information about the NED are available at http://gisdata.usgs.net/ned and http://mac.usgs.gov/isb/pubs/factsheets/fs14899.html, respectively. It provides seamless elevation data over the entire United States at approximately 30-meter resolution. The USGS envisioned this dataset as a vehicle to allow users to focus on analysis rather than data preparation (http://mac.usgs.gov/isb/ pubs/factsheets/fs14899.html). They created the NED from a variety of digital elevation model sources that had various horizontal datums, map projections, elevation units, and quality. In 1999, the Indiana Geological Survey (IGS) acquired the NED for Indiana and parts of Illinois, Michigan, Ohio, and Kentucky; at that time, it was the best available DEM of Indiana at 30-meter resolution.

REVISING THE INDIANA DEM

The majority of the NED is high-quality (Level 2) DEM data, however, a part of the data is lower-quality (Level 1) data. Level 1 DEMs were created with older

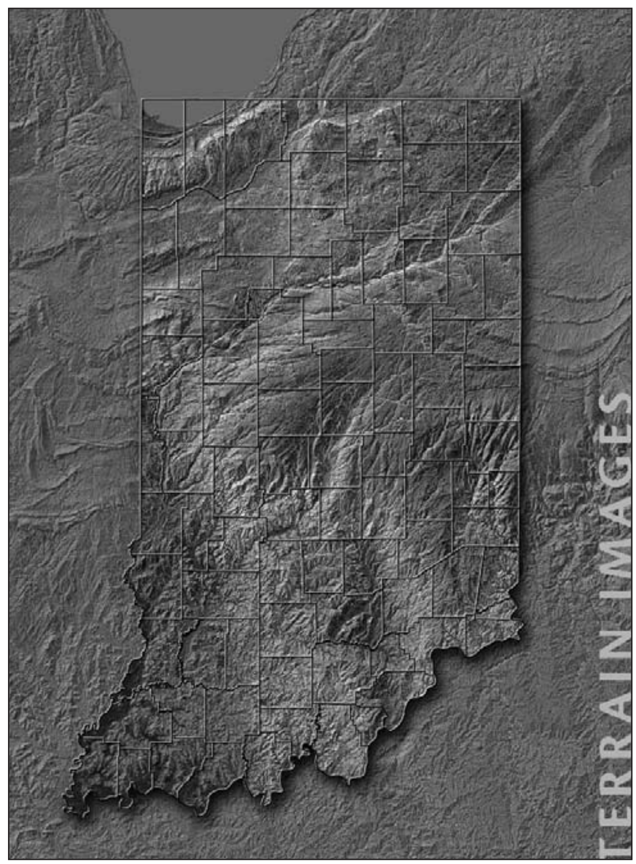


Figure 1. Surface Terrain of Indiana from IGS Poster 5. In this image, the revised digital elevation model has been exported from ESRI ArcView and brought into Adobe Illustrator/Avenza MapPublisher (http://www.avenza.com/main.html) and Adobe Photoshop for enhancement.

methods and have lower accuracy standards, while Level 2 DEMs were created using more modern methods and have higher accuracy standards (Figure 2). The IGS discarded the Level 1 data in the greater Indiana area of the NED and replaced it with newly created digital elevation models (DEMs). In the process of making these DEMs, it was necessary to reprocess some of the high-quality (Level 2) quadrangles owing to their location within blocks of Level 1 data (Figure 3). Of the 275 7.5-minute quadrangles replaced in the Indiana dataset, 216 quadrangles were Level 1 and 59 quadrangles were Level 2. Other revised areas (shown on Figure 3) are: 1) a part of western Ohio, including the eastern portion of Allen County, Indiana, which was processed by the Ohio Division of Geological Survey; 2) sixteen 7.5 minute quadrangles in and adjacent to Berrien County, Michigan; and 3) a 1000-meter buffer zone, extending just beyond the block of replaced quadrangles, which was added in order to create a more seamless splice at the edge of the newly created quadrangles.

The processing steps to create the DEMs in ArcInfo 8.2 are listed in the Appendix. This processing included the following components:

- digital hypsography—contour data provided the elevation values for the DEM. The USGS completed Indiana hypsography at 1:24,000 scale in 2003.
- digital waterbody boundaries—waterbody polygons are needed to ensure that lakes appear flat in the resulting DEM. These were obtained from the U.S. Census Bureau 2000 TIGER/Line files (1:100,000 scale) and the USGS/EPA National Hydrography Dataset (1:100,000 scale).
- digital hydrography—oriented stream data provide the direction of the slope, and ensure the correct placement of the stream valleys. These data were obtained from both the U.S. Census Bureau 2000 TIGER/Line files (1:100,000 scale) and the USGS/EPA National Hydrography Dataset (1:100,000 scale).

ELEVATION DISCREPANCIES

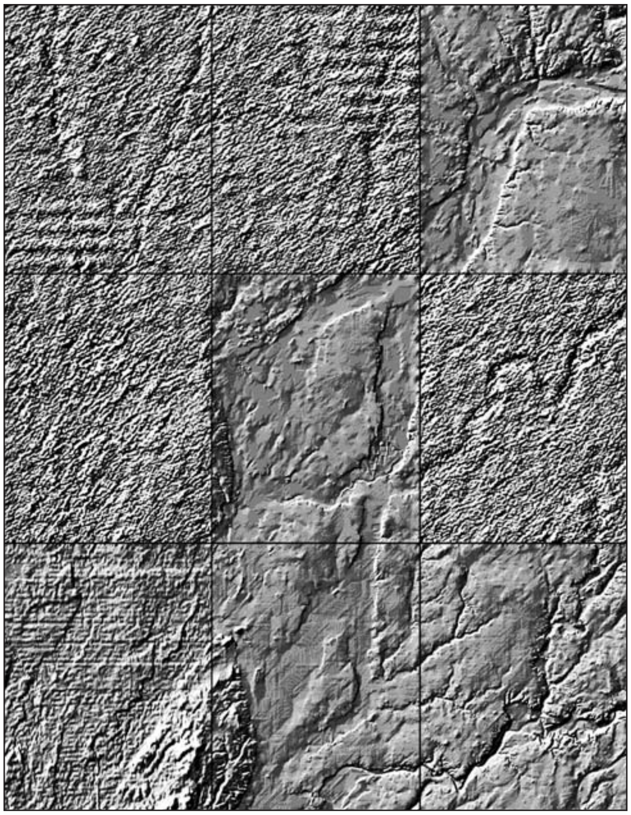
Despite our processing, elevation discrepancies may exist in the DEMs for some of the reprocessed 7.5-minute quadrangles for the following reasons:

- Some of the digital hypsography contains incorrect elevation data. These were not corrected because there appear to be very few of these errors and the datasets were large.
- The lake polygons and stream paths from the TIGER dataset adhere to a (USGS) 1:100,000 scale map accuracy standard and the digital hypsography data adhere to a (USGS) 1:24,000 map accuracy standard. Because of the original scale differences, the boundaries of the lakes and streams do not exactly coincide with the 1:24,000 contours, and this occasionally causes incorrect elevations in the grid near some of the lake boundaries.
- Corrections using "sinks," a process step available in ArcInfo, were not made, and this caused relatively flat areas and areas with karst topography to have slightly erroneous elevation values.
- Contour lines within old and currently-operating strip mines located in southwestern Indiana commonly were absent from the hypsography data. These areas therefore have some incorrect elevation values within the perimeters of the mines and spoil piles.
- Some areas of the Ohio River have inaccurate elevation values owing to the steps taken to make the river have some width as it does on a topographic map, instead of a simple line or arc. The process of applying a waterbody shapefile to the river valley in TOPOGRID disrupted some of the true elevation values of the Ohio River in the southwestern portion of the state.

RESULTANT IMAGES

The four Tagged Image File Format (TIFF) images shown in Figure 4 are samples of graphical images that were created from the new statewide DEM using ArcView 3.3 and desktop publishing software. To create these images the following procedure was used. First, the grid was imported into ArcView using the Spatial Analyst extension and a color ramp was applied. Second, the hillshade was created with the surface\hillshade option. Finally, the graphical images were created and exported using an extension called Image Conversion-Georeferencing2, adapted from a script by Kenneth McVay (downloadable from ESRI's ArcScripts website at http://arcscripts.esri.com/details.asp?dbid=10603). This extension exports images of several file formats while maintaining original raster resolution and also creates a world file. Images may be exported as color DEMs, hillshades, or a combination of both DEM and hillshade.

Once the images are exported as TIFFs, they may be imported into graphics software programs such as Adobe Illustrator with the Avenza MAPublisher filter, where georegistered TIFFS may be added as layers with their world files. The MAPublisher filter allows for on-screen data analysis and for creating new GIS information. Further enhancements can be accomplished in Adobe Photoshop where various adjustments can illuminate textures and patterns not usually visible. Resultant images can be returned to any GIS application, with their original world files.



A

B

C

Figure 2. The digital elevation model of nine central Indiana 7.5-minute quadrangles before (A) and after (B and C) the editing process.

A. Original NED, hillshaded. Five quadrangles are Level 1 (lower-quality) data and four quadrangles are Level 2 (high-quality) data. In the editing process of this area, the IGS replaced all the Level 1 quadrangles and two Level 2 quadrangles (the northeast and central quadrangles) with new digital elevation models made by the IGS. The south-central and southeast quadrangles were outside the revised area.

B. Revised DEM, hillshaded. There is a slight difference in topographic "texture" between the areas processed by the IGS and the areas processed by the U.S. Geological Survey (southeast and south-central quadrangles).

C. Revised DEM, with a color ramp and the hillshade. This is the same edited DEM as in B., but with color applied according to elevation combined with the hillshade.

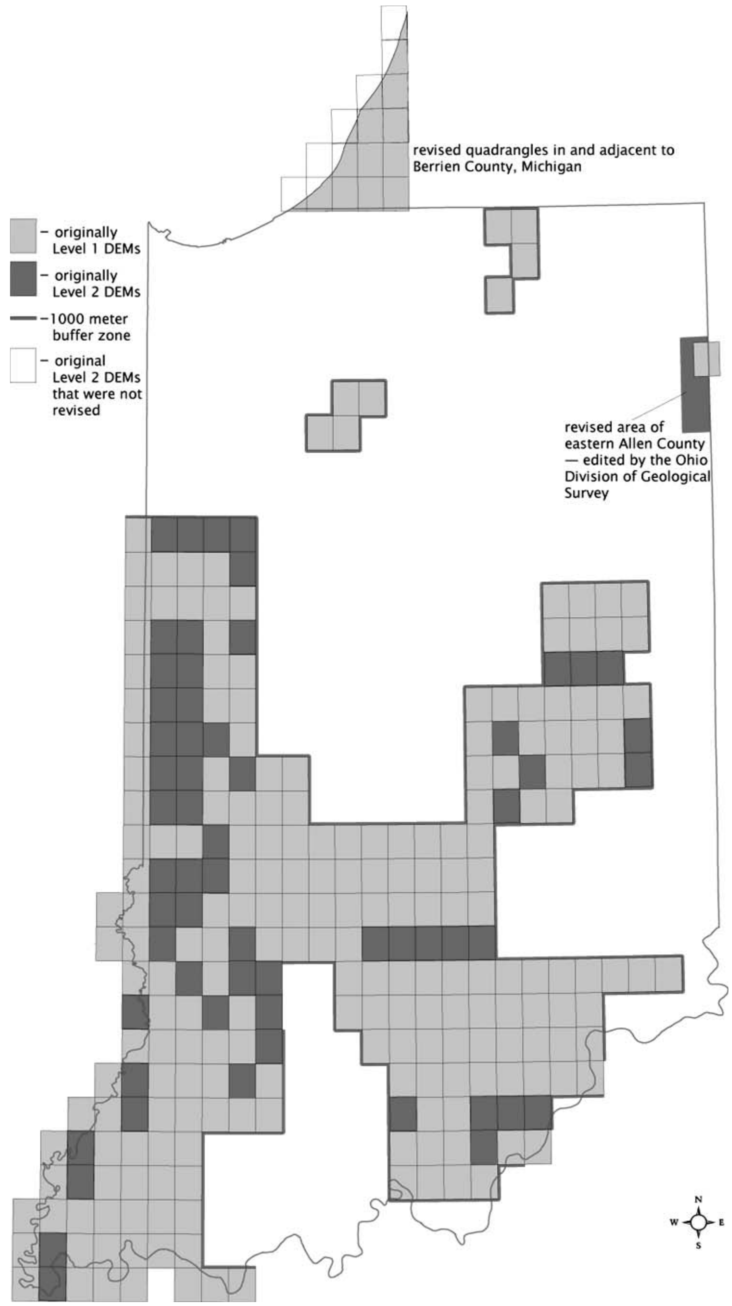


Figure 3. Revisions in the DEM dataset of Indiana.

CREDITS

The IGS purchased the NED from the USGS with funds from a 319 grant from the Indiana Department of Environmental Resources (IDEM). The 1:24,000-scale hypsography was provided to the IGS by the Indiana Department of Natural Resources (INDR), which received them from the USGS. Chris Dintaman, IGS geologist and GIS analyst, standardized a set of commands and variables within ArcInfo to produce the new DEMs

REFERENCES

Berry, G.M., Bleuer, N.K., Brown, S.E., Dickson, M.L., Dintaman, C., Olejnik, J., and Rupp, R., 2004, Surface Terrain of Indiana: Indiana Geological Survey Poster, POSTER05, accessed at http://igs.indiana.edu/survey/publications/searchPubsDetail.cfm?Pub_Num=POSTER05.

Brown, S.E., Bleuer, N.K., Dickson, M., Dintaman, C., Olejnik, J., and Rupp, R., 2004, Digital Elevation Model of Indiana—Revised: Indiana Geological Survey Open-File Studies, OFS04-01, accessed at http://igs.indiana.edu/survey/publications/searchPubsDetail.cfm?Pub_Num=ofs04-01.

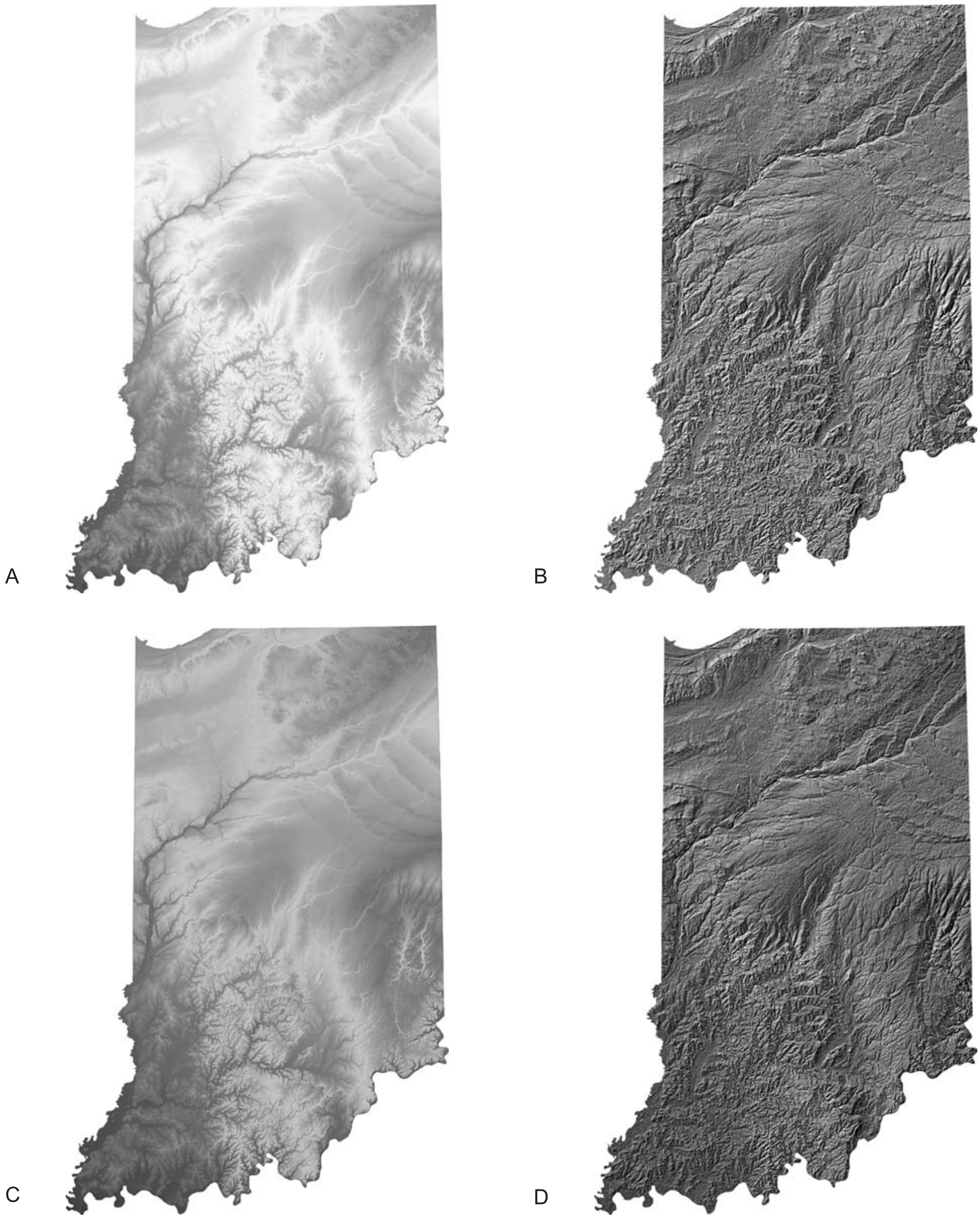


Figure 4. Thumbnails of TIFF images made from the new DEM—available on CD-ROM by Brown and others (2004).

A. Image of the colored DEM only. A color ramp was applied to the DEM in ArcView 3.3.

B. Image of the grayscale hillshade only.

C. Combined image of the colored DEM and the hillshade, exported from ArcView as one image.

D. Combined image of the colored DEM and the hillshade which were exported from ArcView individually (Figure 4A and 4B). The two images were placed into Adobe Illustrator as separate layers where they were combined and enhanced.

APPENDIX

Steps for creating the IGS digital elevation models (DEMs)

1. Obtain 1:24,000 DLG hypsography from the USGS.
2. Check each attribute table in each hypsography file for the number of major and minor fields; there can be 4 pairs or more of major/minor fields depending on the amount of information conveyed. The elevation values are located in the minor fields. In ArcInfo, use the APPEND command to group the DLGs by similar number of major and minor fields (group those with major3 and minor3 together, those with major4 and minor4 together, etc.).
3. In ArcInfo Tables, use ADDITEM command to add a field called ELEV to the attribute table for each DLG. Then populate the ELEV field with the elevation values from the minor fields. Use the DROPITEM command to delete the data columns labeled major and minor.
4. Delete unnecessary “boundary” arcs—those with ELEV values of 0. Also, delete arcs with ELEV values of 5, 10, 15, etc, if these are obviously not elevations. These are usually lake-bottom depths.
5. Append all the edited DLG coverages into one coverage. Build this coverage in ArcInfo for lines, but not for polygons.
6. Obtain USGS hydrography, if available. If USGS hydrography is not available, obtain U.S. Census Bureau lake and stream data.
7. Use the geoprocessing wizard in ArcView to merge all the stream shapefiles and the waterbody shapefiles together to make a seamless shapefile of the entire area. Delete all double-edged streams and very small lakes in the waterbody shapefile (if using TIGER files). Delete all arcs around all lakes in the stream shapefile (there should be no lakes on the stream coverage). Redigitize any lines that are missing after deleting the lakes in stream valleys. All double-edged streams should be redigitized down the center of the valley.
8. Convert the waterbody shapefiles and the stream shapefiles to coverages in ArcInfo. Then build the water body coverage for polygons and the stream coverage for lines.
9. Edit the stream flow directions in ArcEdit so that each segment of each stream has the proper flow direction. This process can be labor intensive, as many of the stream flow directions are reversed.
10. Run the TOPOGRID command in ArcInfo GRID using the following list of subcommands and parameters:

    <Arc command prompt>: topogrid grid_name 30 *(the grid to be created and the cell size, in map units, of the output grid)*
    TopoGrid: contour coverage_name elev *(keyword and parameters for input of a line coverage representing elevation contours; this command causes ArcInfo to prompt the user for the keywords and parameters)*
    TopoGrid: datatype contour *(the primary type of input data)*
    TopoGrid: enforce on *(enforce trends of drainage)*
    TopoGrid: iterations 30 *(maximum number of iterations for the resolution)*
    TopoGrid: outputs sink_1 drain_1 diag_1 *(optional outputs providing information that can be used to evaluate the output elevation grid)*
    TopoGrid: stream stream_name *(keyword for input of a line coverage representing streams)*
    TopoGrid: lake lake_name *(keyword for input of a polygon coverage representing lakes)*
    TopoGrid: tolerances 2.5 1 0 *(a set of tolerances used to adjust the calculations of the interpolation and drainage enforcement process)*
    TopoGrid: end

11. Clip the resulting grid to a buffered shapefile (the outline of the new DEMs with a 1,000-meter buffer).
12. Merge the final clipped grid or grids back into the state DEM using the ArcInfo MERGE command in GRID—grid: final_name = merge (1stpriorityclip, 2ndprioritystate, etc).
13. Change the type of grid from a floating point grid to an integer grid. This creates a much smaller file.

Cartography at the Montana Bureau of Mines and Geology

By Susan Smith

Montana Bureau of Mines and Geology
1300 West Park Street,
Butte, Montana 59701-8997
Phone: (406) 496-4173
Fax: (406) 496-4451
e-mail: ssmith@mtech.edu

This paper shows two of the Montana Bureau of Mines and Geology's most recent publications. Each map was originally created in ArcInfo7 then transferred to FreeHand10 for final layout.

The Lewis and Clark map in its final form has 10 layers of information. Each layer was exported as an .ai file from ArcInfo and imported into FreeHand. Tic marks common to each layer were used to register the layers. The line weights/colors, and font style/color were set in FreeHand and the relief map was added. City, mountain range and stream names, and the photographs, captions, main text, legend, and collar information were added in FreeHand.

The Butte mine map with over 65 layers was imported layer by layer into FreeHand using MAPublisher5. The line weights and colors were set as each layer was imported. The headframes, text on the map, legend, graphics, main text file, and collar information were added in FreeHand.

MAP 1

Entitled "The Route and Campsites of Lewis and Clark in Montana: A Geologic Perspective", by Robert N. Bergantino and Kenneth L. Sandau, this publication features a large map of Montana showing a detailed depiction of the Expedition's route and campsites (Figure 1). There is also a sidebar of text explaining the geology of the Expedition, and nine photographs of specific landmarks as Lewis & Clark would have seen them, with appropriate quotes from their journals.

MAP 2

Entitled "Butte, Montana, Richest Hill on Earth; 100 Years of Underground Mining", by Ted Duaime, Patrick J. Kennelly, and Paul Thale, this map is a compilation of previously unpublished historical information about the underground mines in Butte, as mapped by the Anaconda Copper Mining Company (Figure 2). The mine workings were plotted over an existing City of Butte base map showing existing streets, the outline of the Berkeley Pit, and the historical location of the Silver Bow Creek stream channel. The names and locations of seventy-four of the major mines are shown on the map. These mines are 1000 feet or more in depth and constitute over 80 percent of the total depth of the Butte mines. Also shown on the map are locations of the fourteen remaining headframes.

Please see our website (http://www.mbmg.mtech.edu/) for other publications. Many of them are available as PDF files.

REFERENCES

Bergantino, R.N., and Sandau, K.L., 2004, The route and campsites of Lewis and Clark in Montana: a geologic perspective: Montana Bureau of Mines and Geology Special Publication 116, scale 1:1,250,000.

Duaime, Ted, Kennelly, P.J., and Thale, Paul, 2004, Butte, Montana, richest hill on Earth; 100 years of underground mining: Montana Bureau of Mines and Geology Miscellaneous Contribution 19, scale 1:9,000.

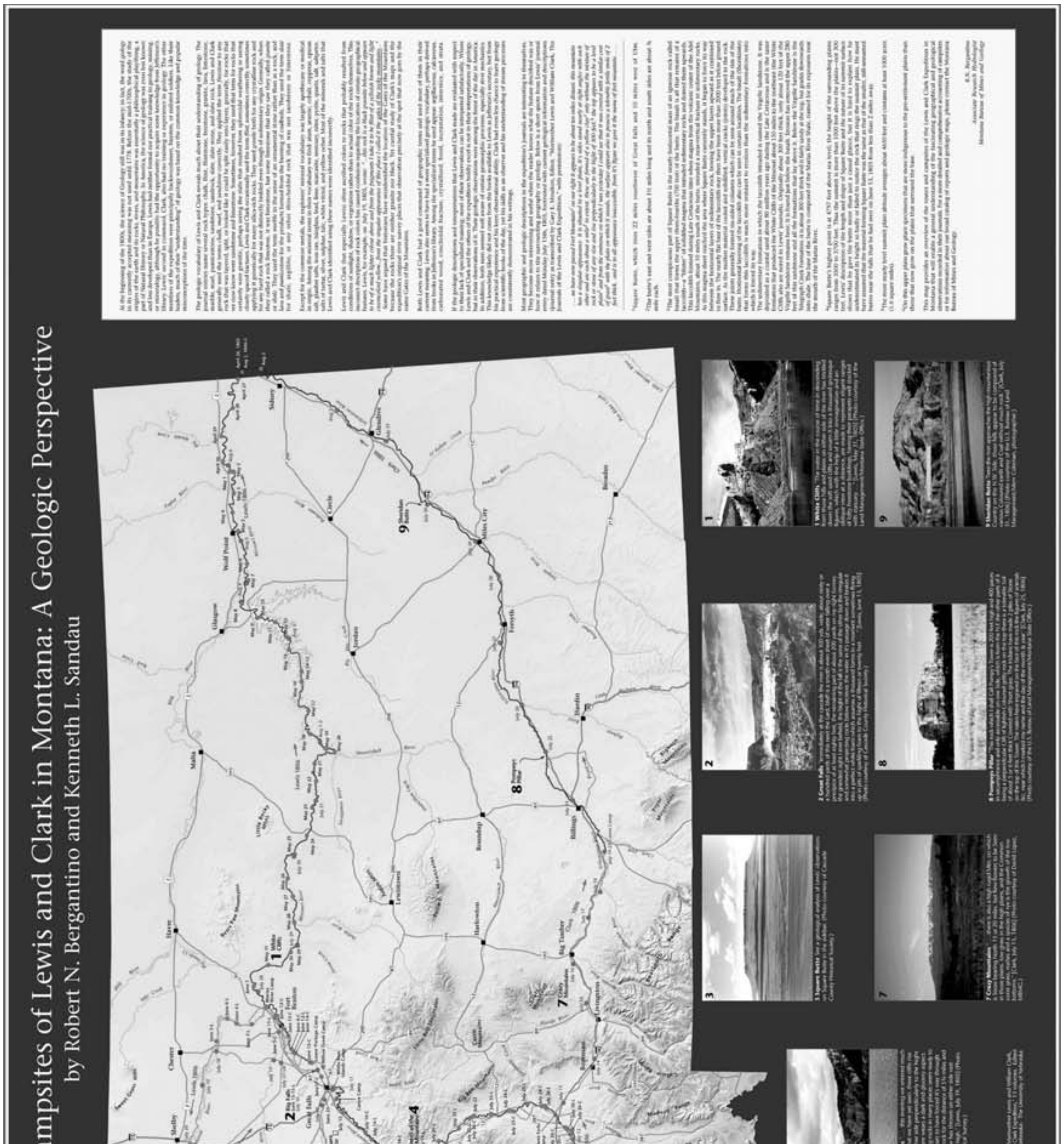


Figure 1. Map 1, The Route and Campsites of Lewis and Clark in Montana.

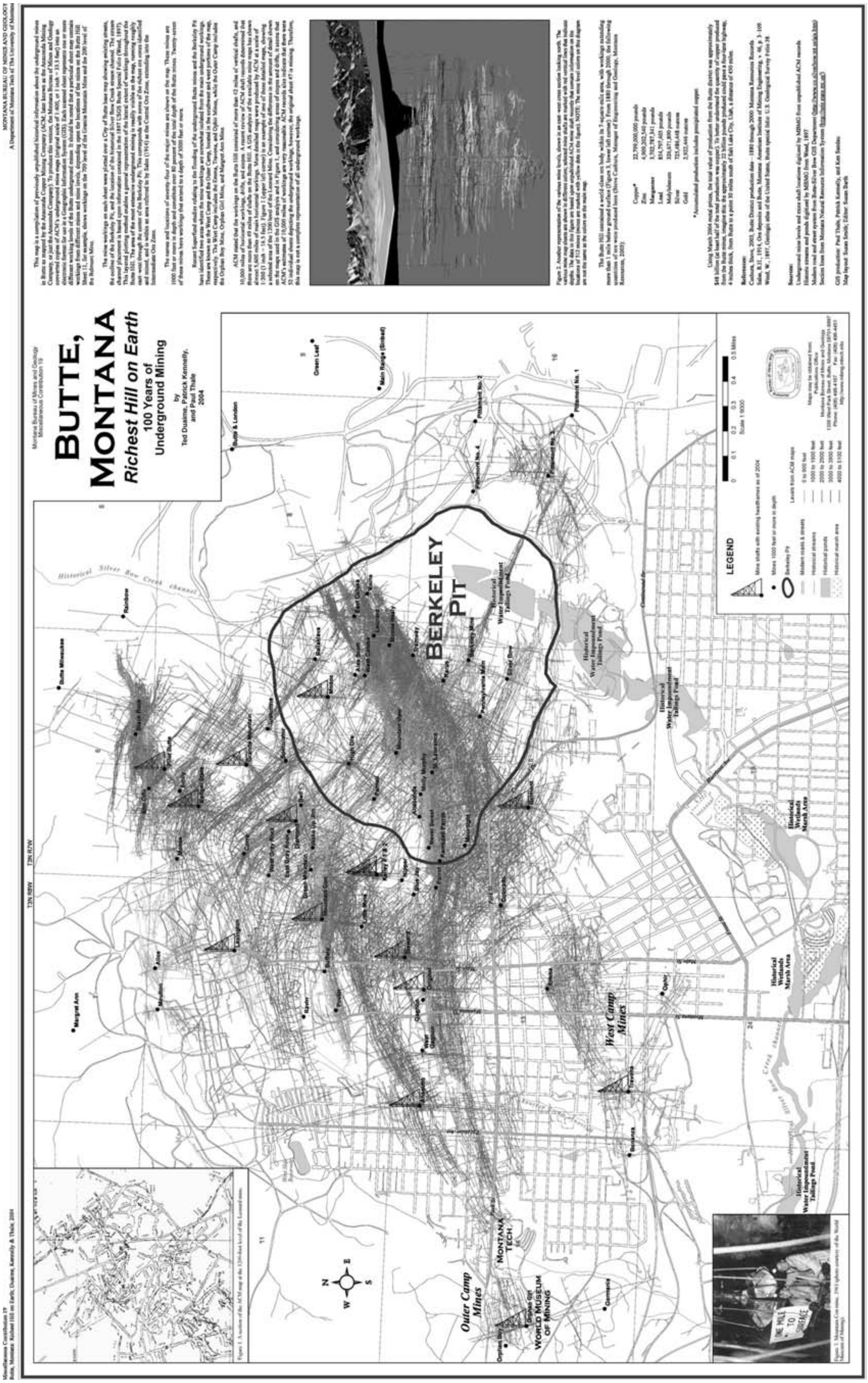


Figure 2. Map 2, Butte, Montana, Richest Hill on Earth.

Conversion of Surficial Geologic Maps to Digital Format in the Seacoast Region of New Hampshire

By Derek S. Bennett, Frederick H. Chormann, Jr., Carl Koteff, and David R. Wunsch

New Hampshire Geological Survey
29 Hazen Drive, PO Box 95
Concord, NH 03302-0095
Telephone: (603) 271-4087
Fax: (603) 271-3305
e-mail: dbennett@des.state.nh.us

ABSTRACT

Over the past 20 years the Seacoast region of New Hampshire has experienced population growth that far exceeds other parts of the state. Figure 1 (Sundquist, 2000) illustrates population trends and projected growth for the state of New Hampshire. This increased population means more homes, buildings, pavement, and other impervious surfaces, which ultimately affect groundwater recharge. Serious questions have been raised about the sustainability of groundwater resources, as demand for the resource continues to rise, whereas groundwater recharge is most likely decreasing.

The New Hampshire Geological Survey (NHGS), in cooperation with the New Hampshire Department of Environmental Services (NHDES), New Hampshire Office of Energy and Planning (NHOEP), and the U.S. Geological Survey New Hampshire/Vermont district (USGS), have entered into an agreement to estimate the availability of groundwater resources at a regional scale in the Piscataqua River / Coastal drainage basin. The goal of the project is to provide southeastern New Hampshire communities with the necessary tools to make informed water resource decisions.

The thickness and distribution of surficial materials play an integral role in determining groundwater recharge, storage, and availability. Therefore, a better understanding of surficial deposits in the region is essential in order to make accurate availability estimates. NHGS has focused its mapping efforts to complete surficial mapping in the few remaining unmapped quadrangles in the area, and to convert existing maps to digital format (Figure 2).

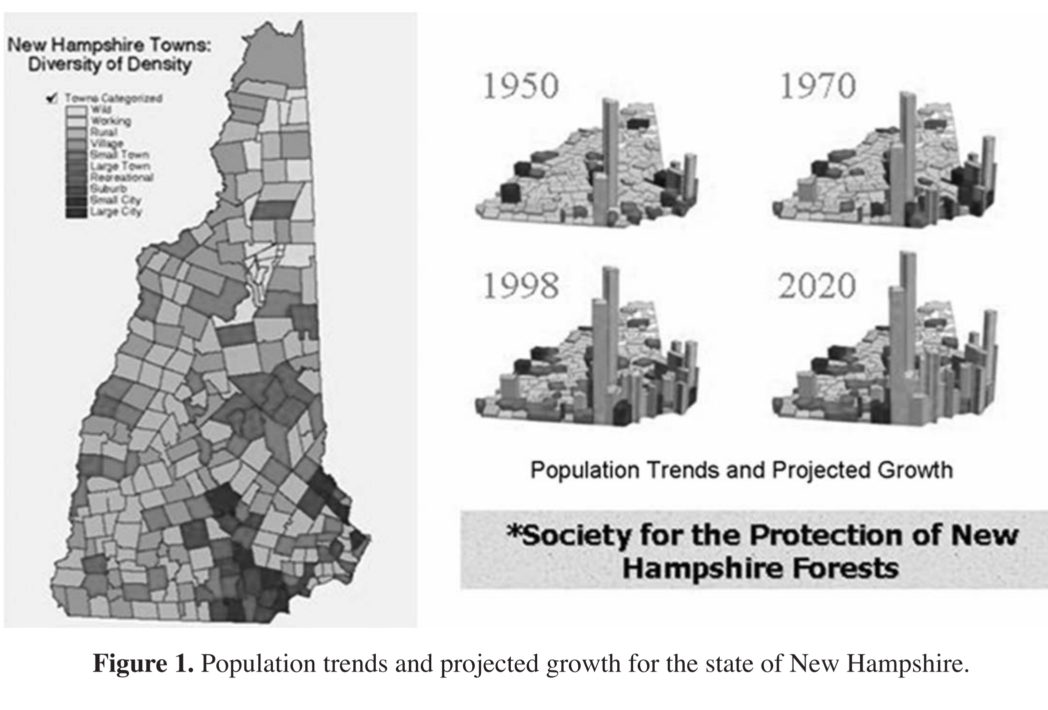


Figure 1. Population trends and projected growth for the state of New Hampshire.

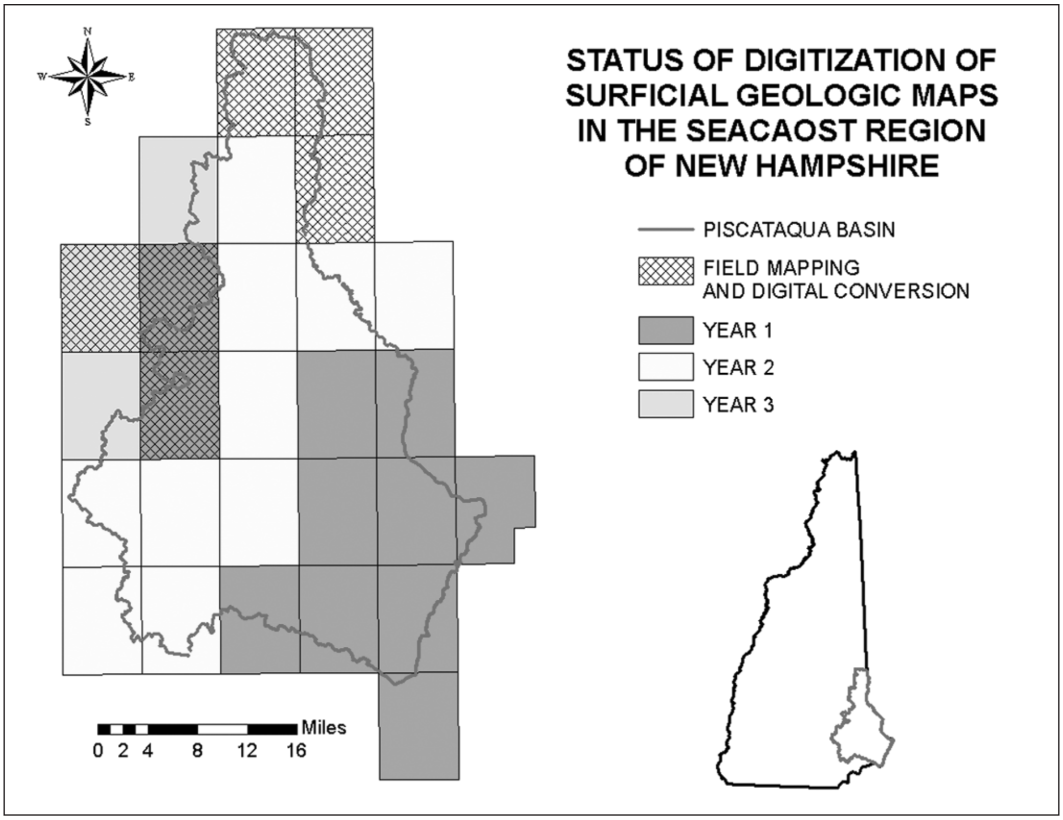


Figure 2. Status of surficial mapping in relation to the Seacoast project area.

Over a three-year period the NHGS will be converting twenty-one published surficial geologic maps to digital format, as well as mapping and digitizing six new surficial geologic maps. In year one (2003), the NHGS converted the following 7.5-minute quadrangles to digital format: Dover East (Larson and Goldsmith, 1989), Dover West (Koteff and others, 1989b), Exeter (Goldsmith, 2001), Hampton (Koteff and others, 1989a), Kingston (Koteff and Moore, 1994), Kittery (Larson, 1992), Newburyport East (Koteff and others, 1989a), Newmarket (Delcore and Koteff, 1989), and Portsmouth (Larson, 1992). The Northwood (Brooks, 2004) and Parker Mountain (Koteff, 2004) quadrangles were mapped as well as digitized. In year two (2004), the NHGS digitized the following thirteen quadrangles: Baxter Lake (Goldsmith, 1993), Barrington (Goldsmith, 1990a), Candia (Gephart, 1985a), Derry (Gephart, 1985b), Epping (Goldsmith, 1990b), Farmington (Goldsmith, 1994), Mount Pawtuckaway (Goldsmith, 1997), Rochester (Koteff, 1991), Sandown (Gephart, 1987), and Somersworth (Koteff, 1991). Three new surficial mapping projects that will include digitization will be conducted in year two: Sanbornville, Great East Lake, and Milton. In year three (2005), the NHGS will digitize the remaining quadrangles covering the Piscataqua River / Coastal drainage basin: Alton (Goldsmith, 1995) and Gossville (Goldsmith, 1998). The Pittsfield quadrangle will be mapped and digitized during year three.

LOCATION AND GEOLOGIC SETTING

The Seacoast region of New Hampshire contains a complex system of sand and gravel deposits of mostly glaciomarine origin, silty facies of the glaciomarine Presumpscot Formation, glaciolacustrine and glaciofluvial deposits, at least two ages of glacial till, and locally thick eolian deposits. Glacial cover exceeds 90 percent in most areas. Tills average 15 feet in thickness but in drumlins can exceed 100 feet. There are two distinct types of till in the region. The upper or ablation till (late Wisconsinan age) is fairly sandy and slightly weathered. Till found within the drumlins (Illinoian age) is more compact and silty, and is deeply oxidized. Deltaic deposits can be as much as 150 feet thick. Most of the sand and gravel deposits in this region have been extracted, but very sizeable amounts remain. The bedrock of the region, which is the source rock for the glacial deposits, consists of metasandstone, phyllite, schist, gneiss, and granite.

APPROACH FOR CONVERTING MAPS TO DIGITAL FORMAT

The conversion of paper maps to digital format presents numerous challenges that are unique in nature. In order to develop a useful and seamless dataset, criteria for map standards, data organization, and attribution need to be established. NHGS utilized the ArcInfo coverage datamodel and organized the surficial units and textures into region subclasses (Figure 3).

Throughout the history of the geologic mapping program, the NHGS has contracted with many different mappers. This presents a challenge from a cartographic perspective, as it is often difficult to reconcile even slight differences between maps without losing integrity in the original map.

A wide variety of coding conventions had been used to describe similar units in different quadrangles. An effort was made to standardize the codes based on mappers' descriptions. New codes that were developed as a result of this exercise will be used in future mapping projects (Figures 4 and 5). Undoubtedly because of geologic setting, nonconforming units will be encountered during future mapping projects. In these instances, new codes will be adopted and added to the geologic database.

Texture codes also were standardized. Figure 6 contains texture descriptions provided by mappers. The descriptions were consolidated into more broadly described texture classes and coded accordingly. However, the original descriptions were maintained in the attribute table in order to preserve the detail recorded by the mapper.

Matters are further complicated by differing interpretations of geologic setting and depositional environment. Although differing opinions are among the driving forces of science, lack of consensus can be very problematic from the cartographer's perspective. This problem is clearly exposed along quadrangle boundaries where one mappers' work adjoins another. Many mis-matches in unit boundaries, such as the example in Figure 7, can easily be resolved. Considering the scale at which the geology is mapped (1:24,000), it usually is appropriate to split the difference between polygons. However, some discrepancies in interpretation are often difficult to resolve and usually require additional field work, such as the example in Figure 8.

As noted above, New Hampshire has also adopted a texture region subclass describing the surficial units where appropriate. By utilizing region subclasses, textures may

REGION	ITEM NAMES	ATTRIBUTES	CODES
SURMA	CODE (7,7,C)	new, standardized code	~~~
			bedrock
			water
	CODE_OLD (7,7,C)	original map unit code	~~~
		null value	-9999
	BEDROCK (1,1,C)	surficial material less than 10 ft thick, or bedrock exposed	y
		not bedrock	n
	DEP_ENVIRON (30,30,C)	environment/mode of deposition	Glaciomarine
			Glaciolacustrine
			Glaciofluvial
			Glaciolacustrine/Glaciofluvial
			Glacial Till
			Marine
			Lacustrine
			Palustrine
			Alluvial
			Colluvial
			Eolian
			Anthropogenic
			-9999
	AGE (30,30,C)	age of deposit	Holocene
			Pleistocene
			Holocene/Pleistocene
			-9999

Figure 3. Attributes associated with the surficial material (SURMA) region subclass. All polygons are assigned a new standardized "code" while retaining the original code assigned by the mapper ("code_old"). The "~~~" in these fields represents the new code for each polygon. Thin deposits receive a "y" in the bedrock column. Polygons are also assigned a depositional environment and age.

DOVER WEST	DOVER EAST	NEWMARKET	PORTSMOUTH	KINGSTON	EXETER	HAMPTON
al	al	al	Qal	Qal	al	
t	t	t	Qt	Qt	t	t
Qw	sw	sw	Qs	Qs	sw	sw
sm	sm	sm	Qsm			sm
mn		mn		Qmw	mn	mn
gs	gs	gs	Qg	Qgs	gs	gs
ms	ms	ms	Qms	Qps		ms
msc	msc	msc	Qm	Qpc		msc

•Qal: ALLUVIUM
•Qt: TILL
•Qw: FRESH-WATER WETLANDS DEPOSITS
•Qsm: SALT MARSH DEPOSITS
•Qmw: WAVE-FORMED DEPOSITS
•Qmwd: WAVE MODIFIED MARINE DELTA DEPOSITS (formerly Qg, Qgs, gs)
•Qps: PRESUMPSCOT FORMATION: SANDY FACIES
•Qpc: PRESUMPSCOT FORMATION: CLAYEY SILTY FACIES

Figure 4. Surficial unit codes used for 7 different quadrangles. Each column represents a different 7.5 minute quadrangle, while each row represents a different type of surficial unit. The original surficial unit codes varied from quadrangle to quadrangle, but have been standardized using the codes below the table (for example, all codes along the bottom row have been converted to standard code "Qpc"). The codes originally used to describe the surficial unit in row 6 were changed to Qmwd throughout the seven quadrangle area.

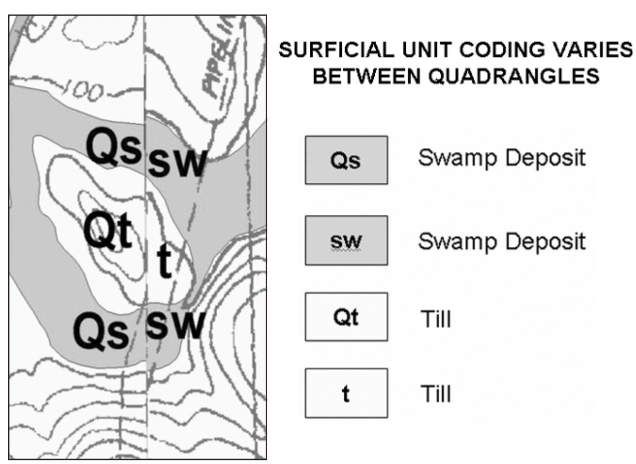


Figure 5. Example of surficial unit coding discrepancy.

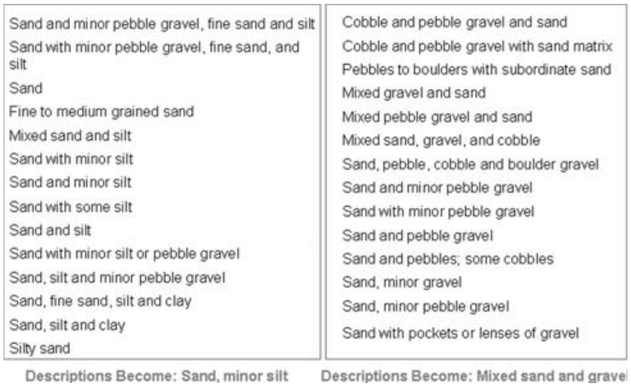


Figure 6. Examples of texture descriptions provided by surficial mappers. Descriptions were generalized into the terms below each box. However, the original, more detailed description used by the mapper also is maintained in the database.

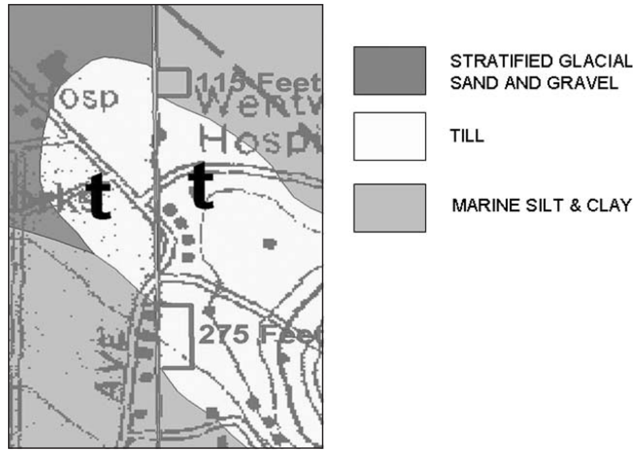


Figure 7. Relatively minor surficial unit discrepancy along quadrangle boundary.

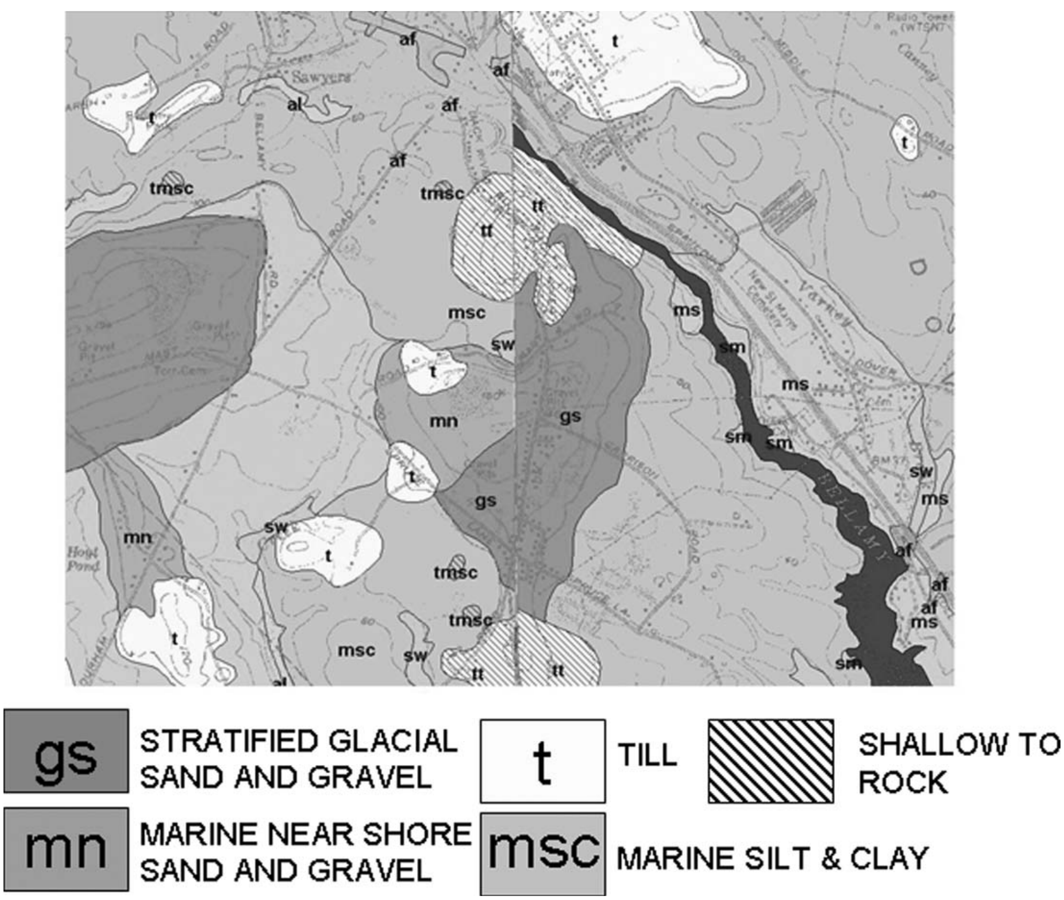


Figure 8. Example of surficial unit discrepancy along quadrangle boundary which requires additional field work.

be associated with the surficial units. Region subclasses also help to ensure that texture and surficial unit boundary arcs are edited simultaneously. As with the surficial unit boundaries, textures also need to be edgematched across quadrangle neatlines as illustrated in Figure 9.

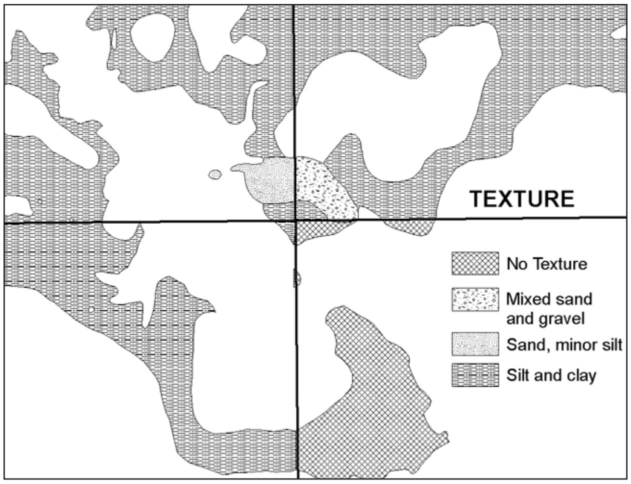


Figure 9. Illustration of four adjoining quads that have different polygon textures needing to be resolved. White areas represent surficial units that do not have texture values while polygons with the "no texture" should have a texture and need to be reconciled with their neighboring quadrangle.

DESKTOP WELL INVENTORY

A desktop procedure for rapidly georeferencing wells has been used to generate data that assist with the resolution of mapping errors. Figure 10 shows georeferenced well data; these give relatively accurate information on overburden thickness and gross material textures, which may provide insight in areas where discrepancies between maps exist.

Since 1984, water well contractors have been required by statute to submit a well completion report for any new water well constructed in the state. From that time, the focus has been on digital data storage/retrieval and georeferencing to enable the data to be used in a geographic information system (GIS) environment. However, the labor-intensive effort to field-locate each reported well, initially with traditional map and compass techniques and later with global positioning satellite (GPS) technology, has failed to keep pace with the rate of new well construction. As a result, only 31% of the 93,000+ reported wells have geographic coordinate values.

A decline in staff resources available to georeference the growing backlog of reported wells, combined with a growing demand for georeferenced well data, have provided impetus for developing an alternative approach to locating these wells. Since 1999, the NHGS has been working to develop, test, and refine a desktop GIS well inventory procedure utilizing digital tax maps and digital orthophotography. The procedure is currently being used

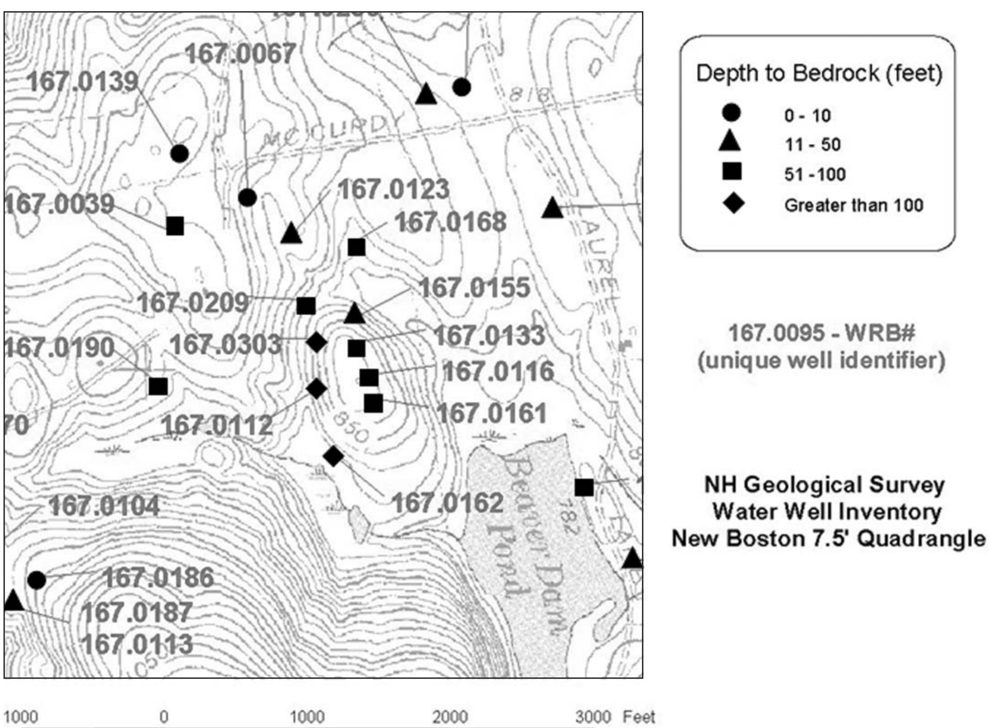


Figure 10. Well locations coded by depth to bedrock ranges. The depths are provided by drillers, on well completion reports. The seven digit WRB# (Water Resources Board Number) is a unique identifier assigned by NHGS when the well completion report information is entered into a database. This unique identifier is used to link the georeferenced well location with well construction details in the well database.

in a "production mode" to georeference wells in the Seacoast region of the state in order to provide basic data on hydrogeologic conditions.

Tax map and parcel information is obtained from local government officials and is matched to well completion reports. A GIS coverage of map and parcel boundaries is draped over digital orthophotography, and well locations are plotted on housetops with the assumption that the well is in fairly close proximity to the residence (Figure 11).

Location data obtained via the desktop method usually are quite accurate as shown in the chart in Figure 12. In this case, 28% of the wells located do not meet the desired accuracy of 100 feet. However, this population shrinks to only 6% when a 250 foot error is deemed acceptable (Chormann, 2001). Errors are reduced even further if the method is selectively applied to domestic wells and smaller parcels.

Over the course of only a few months, NHGS has successfully identified over 1000 well locations to assist with the Seacoast groundwater availability project and the digitization of surficial geologic maps (Figure 13).

CONCLUSIONS

The conversion of paper maps to digital format is a labor intensive process requiring close scrutiny of the maps to be digitized. Identification and documentation of ALL existing features and coding conventions is necessary if standardization is to be applied to the digital data. However, we also must preserve original map content. The integrity of a map may be lost if a feature is changed simply to conform with a standard. Therefore, it is critical to document changes where they occur and to maintain original data within the new map's database.

Converting maps to digital form creates a much more usable and dynamic product, as the data may be used in conjunction with a wide variety of other datasets. Geologic maps are often a work in progress and by maintaining a product digitally the burden of editing is eased considerably as more information becomes available at a specific location. For example, private lands once closed to entrance may open allowing a geologist to perform field investigations where accessibility was once a problem. New wells may be drilled that provide insight into an area

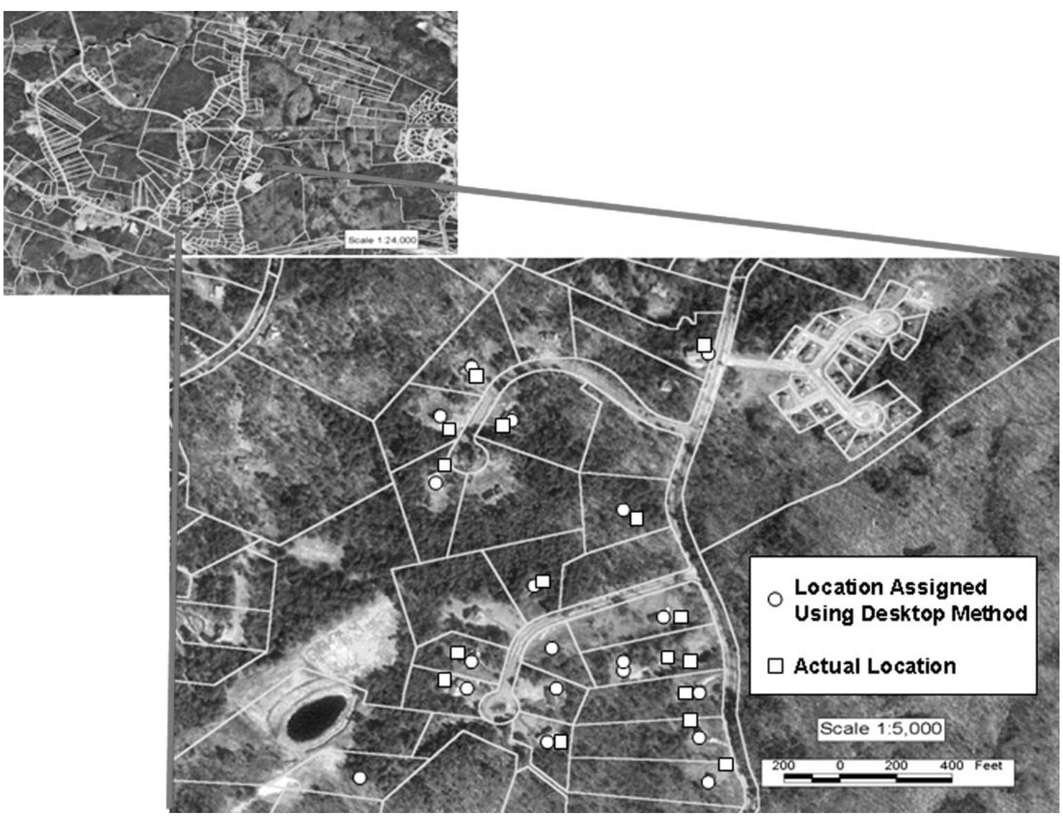


Figure 11. Comparison of actual well location and desktop well inventory procedure for georeferencing well locations. The procedure utilizes town map and parcel boundaries draped over orthophotography.

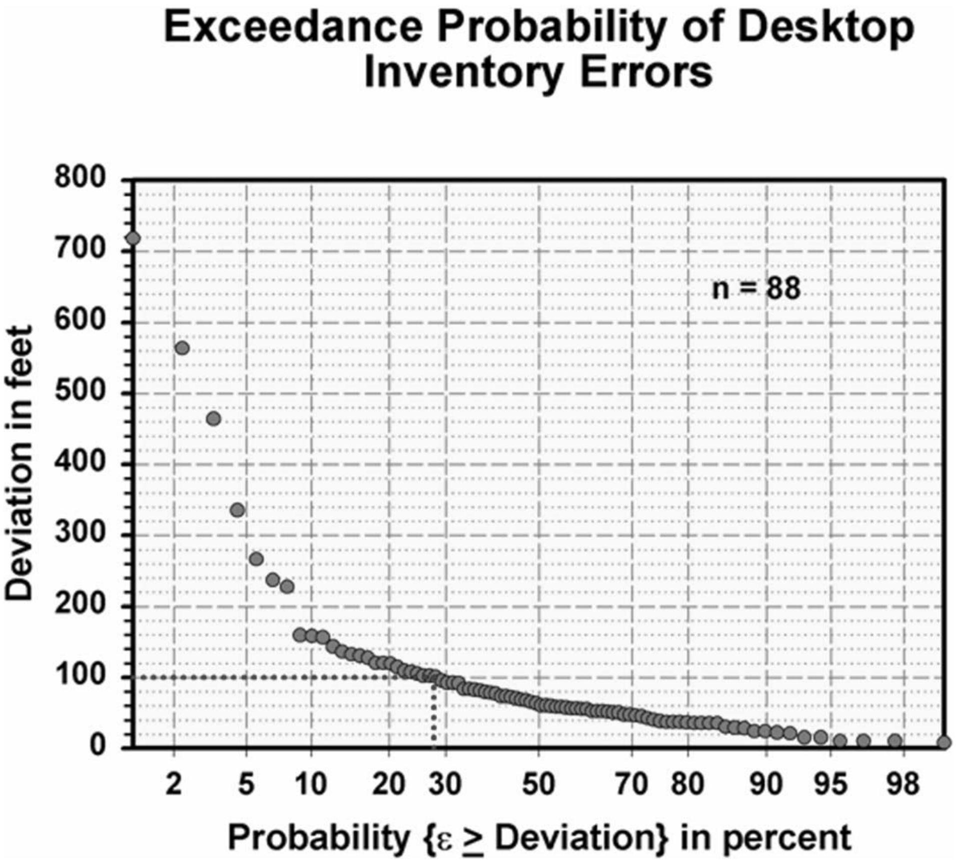


Figure 12. Exceedence probability of desktop inventory errors. The y-axis is the deviation (in feet) between the location assigned using the desktop method and the actual location of the well (see Figure 11). The x-axis is the probability (in percent) that a given deviation will be exceeded.

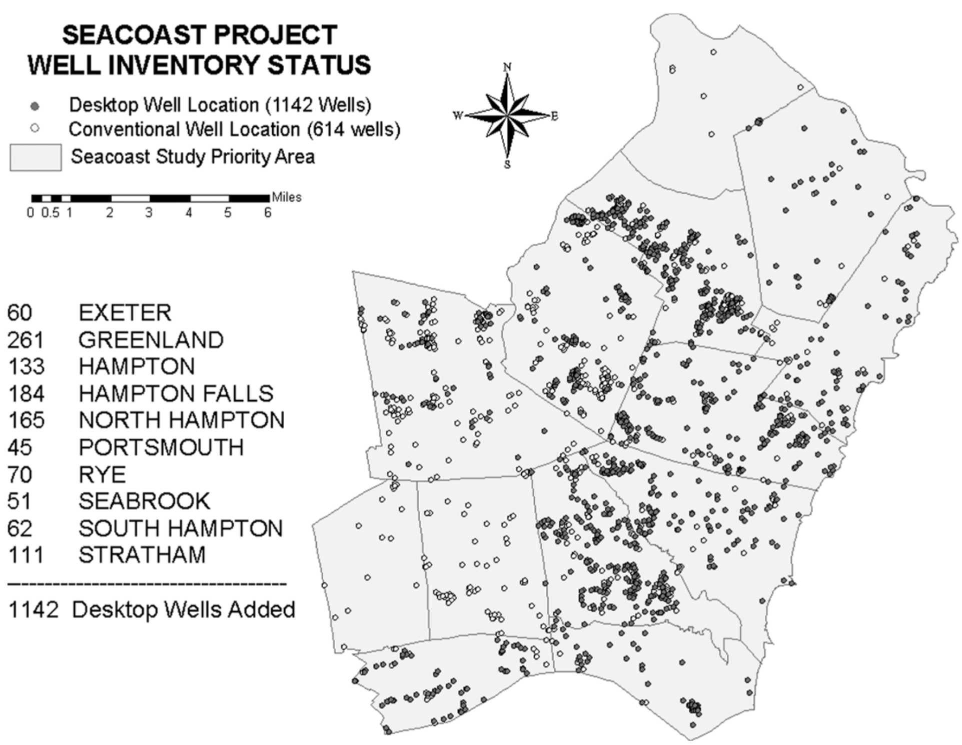


Figure 13. Georeferenced well locations added to the Seacoast project area utilizing the desktop inventory procedure.

where data was not readily available at the time of map publication. Digital products also provide an easy way of tracking changes to maps so comparisons can be made over time.

Looking into the future, it is important to employ standards in new mapping. With standards in place, mappers may reference specific criteria such as unit coding and descriptions during data collection. By utilizing these criteria before map production, a great deal of time and energy may be saved as new maps can easily and quickly be converted to digital form. Promoting the use of standards will help to ensure that geology is seamless across mapping boundaries and mappers alike.

REFERENCES

Brooks, John, 2004, Surficial Geologic Map of the Northwood Quadrangle, Rockingham and Strafford Counties, New Hampshire: New Hampshire Geological Survey, Geo-166, scale 1:24,000.

Chormann, F.H., 2001, The New Hampshire Water Well Inventory (abstract): 53rd Annual National Ground Water Association Convention and Exposition, Nashville, TN, December 7-9, 2001.

Delcore, Manrico and Koteff, Carl, 1989, Surficial Geologic Map of the Newmarket Quadrangle, Rockingham and Strafford Counties, New Hampshire: U.S. Geological Survey Open-file Report 89-105, scale 1:24,000.

Goldsmith, Richard, 2001, Surficial Geologic Map of the Exeter Quadrangle, Rockingham County, New Hampshire: New Hampshire Geological Survey, Geo-126, scale 1:24,000.

Goldsmith, Richard, 1998, Surficial Geologic Map of the Gossville Quadrangle, Belknap and Strafford Counties, New Hampshire: New Hampshire Geological Survey, Geo-169, scale 1:24,000.

Goldsmith, Richard, 1997, Surficial Geologic Map of the Mt. Pawtuckaway Quadrangle, Rockingham County New Hampshire, U.S. Geological Survey, Open File Map OF97-1, scale 1:24,000.

Goldsmith, Richard, 1995, Surficial Geologic Map of the Alton quadrangle, Belknap and Strafford Counties, New Hampshire: New Hampshire Geological Survey, NH 95-1, scale 1:24,000.

Goldsmith, Richard, 1994, Surficial Geologic Map of the Farmington Quadrangle, Carroll and Strafford Counties, New Hampshire: New Hampshire Geological Survey, Geo-112, scale 1:24,000.

Goldsmith, Richard, 1993, Surficial Geologic Map of the Baxter Lake Quadrangle, Strafford County, New Hampshire: New Hampshire Geological Survey, NH93-01, scale 1:24,000.

Goldsmith, Richard, 1990a, Surficial Geologic Map of the Barrington Quadrangle, Rockingham and Strafford Counties, New Hampshire: New Hampshire Geological Survey, NH 90-02, scale 1:24,000.

Goldsmith, Richard, 1990b, Surficial Geologic Map of the Epping Quadrangle, Strafford and Rockingham Counties, New Hampshire: New Hampshire Geological Survey, NH 90-01, scale 1:24,000.

Gephart, G.D., 1987, Surficial Geologic Map of the Sandown Quadrangle, Rockingham County, New Hampshire: New Hampshire Geological Survey, GEO-91, scale 1:24,000

Gephart, G.D., 1985a, Surficial Geologic Map of the Candia Quadrangle, Rockingham County, New Hampshire: New Hampshire Geological Survey, GEO-90, scale 1:24,000.

Gephart, G.D., 1985b, Surficial Geologic Map of the Derry Quadrangle, Rockingham County, New Hampshire: New Hampshire Geological Survey, GEO-89, scale 1:24,000.

Koteff, Carl, 2004, Surficial Geologic Map of the Parker Mountain Quadrangle, Belknap, Strafford, Rockingham, and Merrimack Counties, New Hampshire: New Hampshire Geological Survey, Geo-139, scale 1:24,000.

Koteff, Carl, 1991, Surficial geologic Map of parts of the Rochester and Somersworth Quadrangles, Strafford County, New Hampshire: U.S. Geological Survey, Map I-2265, scale 1:24,000.

Koteff, Carl, Gephart, G.D., and Schaefer, J.P., 1989a, Surficial Geologic Map of the Hampton and Newburyport East 7.5 Minute Quadrangles, New Hampshire and Massachusetts: U.S. Geological Survey, Open-file Report 89-430, scale 1:24,000.

Koteff, Carl, Goldsmith, Richard and Gephart, G.D., 1989b, Surficial Geologic Map of the Dover West quadrangle, Strafford County, New Hampshire: U.S. Geological Survey, Open-file Report 89-166, scale 1:24,000.

Koteff, Carl, and Moore, R.B., 1994, Surficial Geologic Map of the Kingston Quadrangle, Rockingham County New Hampshire: U.S. Geological Survey Geologic Quadrangle Map GQ-1740, scale 1:24,000.

Larson, G.J., and Goldsmith, Richard, 1989, Surficial Geologic Map of the Dover East Quadrangle, New Hampshire: U.S. Geological Survey, Open-file Report 89-97, scale 1:24,000.

Larson, G.J., 1992, Surficial Geologic Map of the Portsmouth and Kittery Quadrangles, Rockingham County, New Hampshire: New Hampshire Geological Survey, Map SG-6, scale 1:24,000.

Sundquist, Daniel, 2000, New Hampshire Towns—Diversity of Density: Society for the Protection of New Hampshire Forests, Concord, NH.

Brownfield and Contaminated Site Remediation: Historic Fill Series

By Michael W. Girard

New Jersey Department of Environmental Protection
New Jersey Geological Survey
29 Arctic Parkway
P.O. Box 427
Trenton, NJ 08625-0427
Telephone: (609) 292-2576
Fax: (609) 633-1004
e-mail: mike.girard@dep.state.nj.us

The "Brownfield and Contaminated Site Remediation Act" (P. L. 1997, c. 278) requires the New Jersey Department of Environmental Protection (the New Jersey Geological Survey's parent agency) to map regions of the state where large areas of historic fill exist, and make this information available to the public. The legislation defines "historic fill material" as "large volumes of non-indigenous material… used to raise the topographic elevation of a site" (N.J.S.A. 58:10B-12(h)1). Fill is mapped on USGS 7.5-minute topographic quadrangle maps (1:24,000-scale) by analyzing landforms using stereo aerial photography, and by comparing areas of swamp, marsh, and floodplain shown on archival (ca. 1840-1880) topographic and geologic maps to their extent as shown on the most recent map. In a few places, fill is mapped from field observations and from driller's logs of wells and borings. Most areas of fill larger than about 5 acres can be identified and mapped using this method. The Historic Fill Map Series has been digitized in ESRI's ArcEdit software at a scale of 1:24,000 and compiled with ESRI's ArcInfo and Adobe Illustrator software. Historic Fill of the Elizabeth Quadrangle (HFM-52 and Figure 1) was mapped by S.D. Stanford of the New Jersey Geological Survey in 2001.

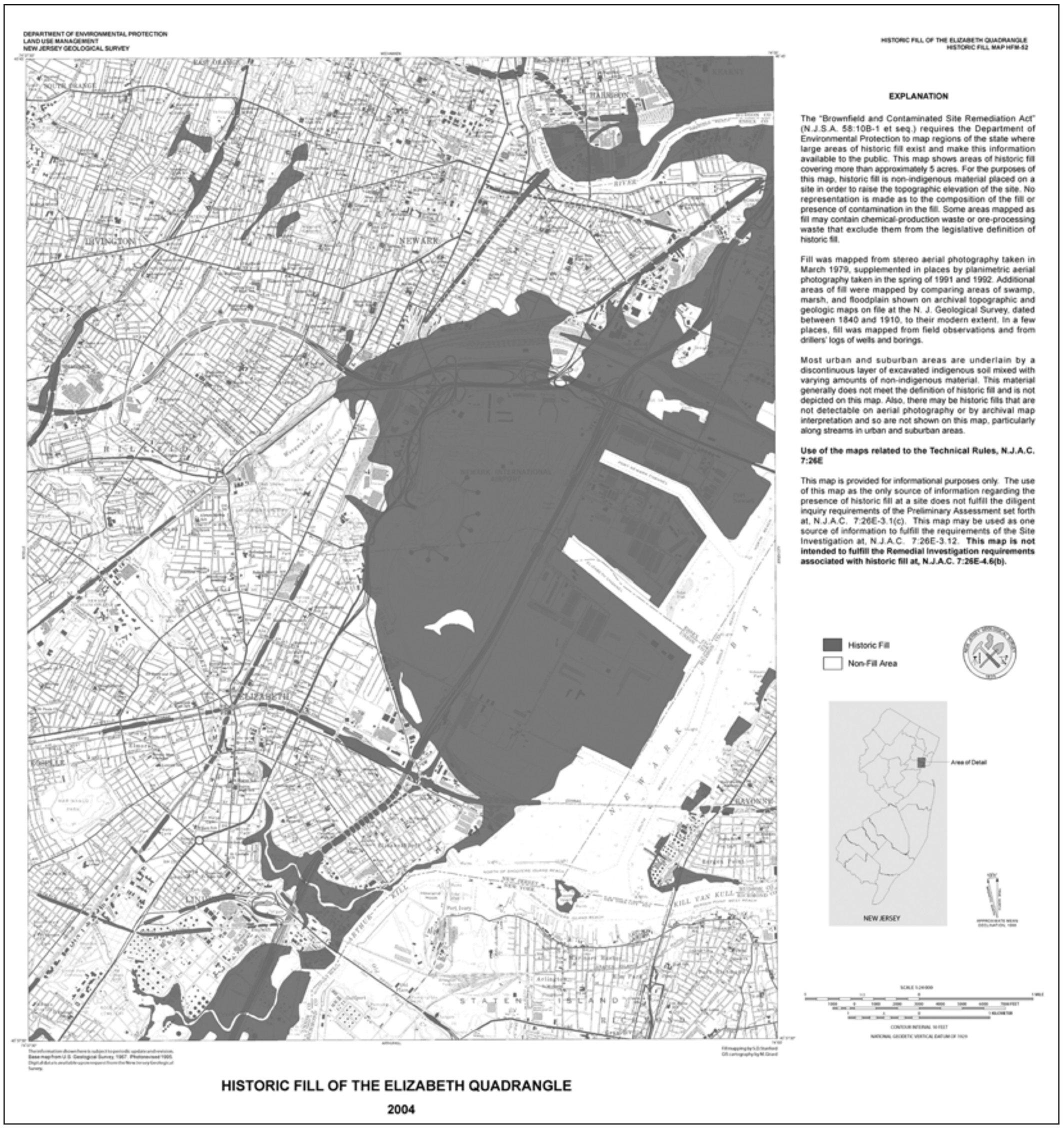


Figure 1. New Jersey Geological Survey Map HFM-52, Historic Fill of the Elizabeth Quadrangle.

Exploring Shaded Relief: The Shaded Drift-Thickness Map of Ohio

By Donovan M. Powers

Ohio Division of Geological Survey
2045 Morse Road, Building C-1
Columbus, Ohio 43229-6693
Telephone: (614) 265-6591
Fax: (614) 447-1918
e-mail: donovan.powers@dnr.state.oh.us

INTRODUCTION

Shaded relief images have become a popular and useful means for improving the "readability" of geologic maps. Some important considerations when evaluating the possibility of using shaded relief for a given map application are: availability of suitable digital elevation data, extent to which available digital elevation data will need to be corrected/enhanced for the intended application, and how other cartographers and Geographic Information System (GIS) professionals are currently using the dataset. The intended end-uses of the shaded relief could have an effect on the methods and formats you use in deriving the data. GIS professionals may use a variety of different software packages and your final output may need to conform to a standard format in order to accommodate various users. Maintaining communication with other professionals could save you some time and effort, and help you to standardize your products.

Digital elevation data generally is available for any area of the United States at no cost. However, if the available data is not sufficiently detailed for your intended application, you may need to create your own dataset. Sometimes, available data must be heavily edited and corrected in order to be suitable for your intended use. In most instances, conscientious cartographers will want to edit and enhance any publicly available digital elevation dataset before using it on their mapping product. This can be a very detailed and time consuming process, but the final product will give you great satisfaction and provide you with an excellent elevation dataset that can be used for many other map applications. This paper describes some techniques used to create a statewide shaded-relief image (Powers and Swinford, 2004; Figure 1).

GETTING THE DATA

When creating an ArcGrid elevation dataset, you will want to select a dataset or create one with a cell size appropriate for your scale or resolution of mapping. These datasets, or Digital Elevation Models (DEMs), are available from the U.S. Geological Survey (USGS). The most complete coverage is available through their National Elevation Dataset (NED); this dataset has a 30-meter spacing. Higher-resolution data (e.g., 10-meter grid) and lower-resolution data (e.g., 90-meter grid) also are available for selected areas of the U.S. The NED and other DEMs can be obtained from the USGS Earth Resources Observation Systems (EROS) Data Center (USGS EROS Data Center, 47914 252nd Street, Sioux Falls, SD 57198; http://edc.usgs.gov/products/map/dlg.html). NED data is useful for cartography at scales of about 1:100,000, or smaller (e.g., 1:250,000). If your map is a larger scale, then you will need to create or obtain a higher resolution (smaller cell size) dataset. If the higher-resolution DEM's are not available for your area of interest, you can follow the process outlined by Haugerud and Greenberg (1998) to create your own. This process, presented at the 1998 Digital Mapping Techniques workshop, utilizes ESRI's ArcInfo Workstation to create an ArcGrid file based on DEM or Digital Line Graph (DLG) layers available through the EROS Data Center.

The Drift Thickness Map of Ohio (Figure 1) was created from the Ohio Division of Geological Survey' bedrock topography contours and converted to a surface using Haugerud and Greenberg's methods. This surface was then subtracted from the DEM (earth's surface) to yield a thickness of drift. From the thickness layer a hillshade layer was created, converted to an image, and cleaned in a graphics program.

CLEANING THE SHADED RELIEF

Many GIS professionals already have some type of shaded relief data on hand that has not been 'cleaned' (i.e., edited and corrected)—probably for lack of time. Nonetheless, a proper cleaning is essential before an elevation dataset is used in a published product.

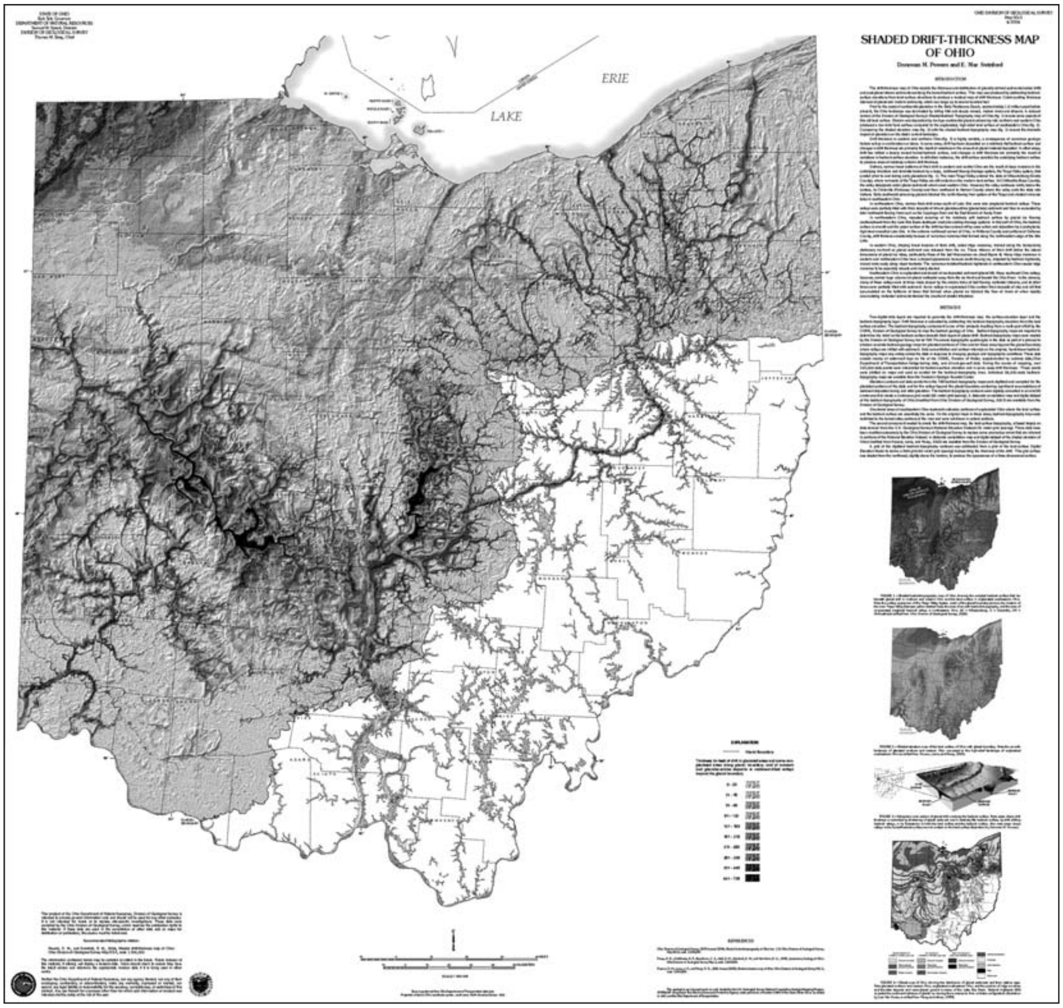


Figure 1. The *Shaded Drift-Thickness Map of Ohio*.

Upon close inspection of the unclean data, you may notice artifacts or "terracing" (Patterson, 1997) in your shaded relief layer that correlate to the topographic contours from which your DEM was created. A slight bias in the interpolation algorithm that Arc Workstation uses to create surface grids causes input contours to have a stronger effect on the output surface than areas between the contours. This bias can result in a slight 'flattening' of the output surface as it crosses the contour. This may result in misleading results when calculating the profile curvature of the output surface. Recently, a new version of the algorithm has been released claiming to address this issue. The algorithm could reduce the amount of time it takes to clean a DEM. It is available as a standalone application with a graphical user interface. The version ESRI uses is substantially older and does not incorporate the additional functionality. It is unknown if or when ESRI will implement this version (Hutchinson, 2000; 2004).

Using tools such as Arc Workstation or Adobe Photoshop, you can accomplish a significant amount of data cleaning without compromising data integrity. In Arc Workstation, the FOCALMEAN command can be used to smooth out unwanted artifacts. FOCALMEAN includes in its calculations a specified cluster of cells (a "neighborhood") around each cell. You specify the size and shape of the neighborhood, and FOCALMEAN calculates the average of the value of the neighborhood. In many cases, you can use a circle with a 5 or 7 cell radius to smooth the raster, depending on the extent and resolution of your data. This doesn't always take out all of the artifacts but it will smooth them for a cleaner look. As you get familiar with the commands, you might find other ways to clean your data. Presently, this seems to be the best approach to eliminate a majority of the artifacts, but I am eagerly awaiting a more comprehensive method.

Should you want to further clean the data, you can use a graphics program like Photoshop. By editing the appearance of a hillshade image created by ArcGrid rather than editing the elevation data, you can greatly enhance your final output. Convert your hillshade grid into a grayscale tiff and edit it in your graphics/photo program using the filters and tools to smooth out the "plateau" or "terracing" effects. This step requires practice and time. Be careful not to overuse the tools. The rubber stamp, dodge, burn, blur, and airbrush tools are the most useful. Some 3rd party add-on filters might also be helpful.

Another option to better display your hillshade is to enhance the peak and valley tones (Figure 2). There is a tendency to lose data on either end of the tonal scale—light values (less than 5 percent) usually disappear completely and dark values (greater than 90 percent) fill in and appear solid. For example, steep slopes perpendicular to the illumination direction appear washed out, whereas shadowed slopes become solid black (Patterson and Hermann, 1997). Blending two separate renderings of the hillshade images can emphasize the subtle details that would likely be lost using the typical software application. Blending uses two copies of the hillshade image overlapping each other. The uppermost copy can be rendered with a lower percentage of transparency, roughly 60% depending on the strength desired, and the histogram, or the contrast can then be adjusted to allow details to appear in the darker and lighter areas. The layers can be merged into one once the desired effects have been achieved.

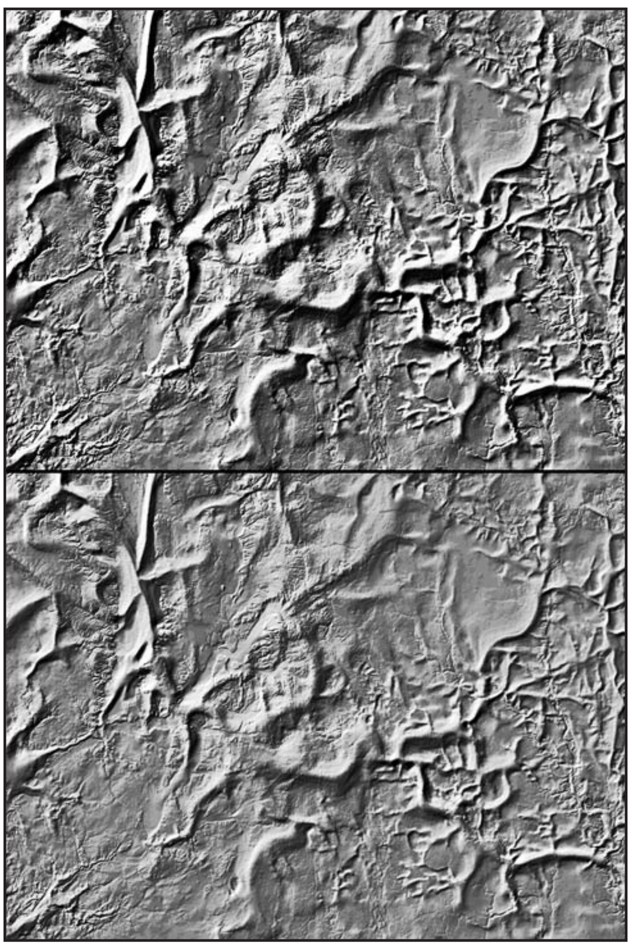


Figure 2. A portion of the *Shaded Drift-Thickness Map of Ohio*, before (top image) and after (bottom image) value-enhancing the tonal qualities to improve the depiction of topographic relief.

USING THE SHADED RELIEF

Cartographers use several methods to create shaded relief images, and the correct choice for you depends on the kind of map you are producing and the software package with which you intend to produce the map. Many cartographers use ESRI products for data creation, and export the data into graphics software such as Adobe Illustrator or Macromedia's Freehand for production of a final product. The best approach is to work with various programs since no single software package seems to be able to adequately handle all of these steps.

If you are making a geological map and wish to add shaded relief to emphasize topographic features, you will want to make the geology somewhat transparent and place that layer over your hillshade. You could do the opposite and make the geology the topmost layer and give it transparency; I generally place the hillshade as the topmost layer with transparency, for a few reasons. While creating the map, I often turn the hillshade off for faster digital rendering. Having the geology with no transparency allows printing without hillshade for geology editing during early inspections. The use of layers and transparencies can be implemented in nearly all of the software packages you might be using. The percentage of transparency depends on your needs, the software you are using, and how you want your map to look. ESRI's products use percentages of transparency whereas Adobe's products use percentages of opacity—a slightly different approach, but the same effect can be accomplished.

If you are making a shaded relief map with elevation intervals shown in different colors, you can do this by either of two methods. The first option is to use two separate layers, one for the colors of the elevations, and another for the hillshade. In this approach, a transparent hillshade layer is draped over a separate layer of color-coded elevation polygons to achieve a textured elevation map. The other option is to merge these two layers together using ArcInfo Workstation, ArcMap, or Photoshop, depending on the format of your data. I prefer to keep the elevation or geology separate from the hillshade in the event the map will be sent out for printing. The colors will be divided into color separates and having the hillshade on its own layer as a separate color will allow you to keep a similar appearance to what you have on-screen during digital compilation. When merged, the shading becomes part of the underlying colors of geology or elevation and the final print colors will not be what you anticipated. Another reason I don't merge the hillshade is simply to have them available for use in other projects as individual layers. You may find it beneficial to merge the two layers in certain situations. One merged layer would have a smaller file size and use less hardware resources than working with two separate layers would. When you compile a final layout this small detail could be the difference between making a print and crashing your computer while attempting to print. Minimizing the size of your digital file also reduces the processing time.

REFERENCES

Haugerud, R., and Greenberg, H.M., 1998, Recipes for Digital Cartography: Cooking with DEMs, *in* Soller, D.R., ed., Digital Mapping Techniques '98-Workshop Proceedings: U.S. Geological Survey Open-File Report 98-487, p. 119-126, accessed at http://pubs.usgs.gov/of/of98-487/haug2.html.

Patterson, T., and Hermann, M., 1997, A Desktop Approach to Shaded Relief Production: North American Cartographic Information Society (NACIS), Cartographic Perspectives, no. 28, p.38.

Patterson, T., 1997, Creating value-enhanced shaded relief in Photoshop: NACIS Cartographic Perspectives, no. 28, accessed at http://www.shadedrelief.com/value/value.html.

Hutchinson, M.F., 2000, Optimising the degree of data smoothing for locally adaptive finite element bivariate smoothing splines: Australian & New Zealand Industrial and Applied Mathematics Journal, v. 42(E), p. C774-C796, accessed at http://anziamj.austms.org.au/V42/CTAC99/Hutc/Hutc.pdf.

Hutchinson, M.F., 2004, The Australian National University, Centre for Resource and Environmental Studies, ANUDEM, accessed at http://cres.anu.edu.au/outputs/anudem.php#3.

U.S. Geological Survey, 1990, Digital Elevation Models: U.S. Geological Survey, Data Users Guide 5, 51 p.

Powers, and Swinford, 2004, Shaded Drift-Thickness Map of Ohio: Ohio Department of Natural Resources, Division of Geological Survey, Map SG-3, 1:500,000 scale, accessed at http://www.dnr.state.oh.us/geosurvey/pdf/sg3map.pdf.

Conversion of a Digital Raster Graphics Image from Color to Black & White

By Clark A. Niewendorp

Oregon Department of Geology and Mineral Industries
800 NE Oregon Street, Suite 965
Portland, OR 97232
Telephone: (503) 731-0283
Fax: (503) 731-4066
e-mail: clark.niewendorp@state.or.us

ABSTRACT

A digital geologic map usually, but not always, includes a base map such as a Digital Raster Graphics (DRG) image created by scanning a U.S. Geological Survey quadrangle map. The Oregon Department of Geology and Mineral Industries (DOGAMI) uses JASC Software's Paint Shop Pro (PSP) to convert the DRG from its color version to black & white. It is rendered as a 1-bit image with no shades of gray; only black for topographic and physical features, and white for the background. In MapInfo Professional (MI), we adjust the percentage of translucency of the DRG, which allows the geology layer (opaque polygons) below the DRG to become visible through the DRG's background.

At DOGAMI, not every cartographer, geologist, or map is best served by MI's capabilities, and so Adobe Illustrator (AI) is sometimes used for the final cartography. But the effects of AI's transparency controls (opacity and blending modes) applied to the DRG may or may not produce a desirable map for presentation. However, if the DRG's white background is saved with PSP's transparency before placement in AI, the use of AI's transparency controls is avoided altogether. As a result, the DRG is a semi-transparent image in AI through which you can "see" the underlying geology.

INTRODUCTION

In a Geographic Information System (GIS), the geologic data gathered through fieldwork is usually, but not always, registered to a Digital Raster Graphics (DRG) image, and attribute information that further explains each of the geologic features is added. A DRG is a raster image file. DRGs are generated from scanned U.S. Geological Survey quadrangle maps, and usually are in TIFF format. A DRG, like a paper quadrangle map, presents the horizontal and vertical positions of features by the use of lines, symbols, and various colors or shading. When it comes to displaying natural and cultural features, there are some commonly-held conventions: water bodies are shown in blue, vegetation in green, relief features in brown, cultural features in brown, and urban areas in red.

We use the DRG as a semi-transparent overlay through which you can "see" the underlying geology layer. For this approach, the DRG works best if it is rendered with a white background, no shades of gray, and all features shown in black – a 1-bit, two-color image. In this paper, the procedure to quickly convert the DRG from its color version to a two-color (actually a black & white) version is discussed. A number of challenges are then addressed, each related to representation of a black & white DRG in Adobe Illustrator (AI). The procedures below are described only to the extent needed for an understanding of the steps. Also, the following text is freely drawn from material already published in instruction manuals and other similar sources.

DECREASING THE COLOR PALETTE

Any image-editing software that allows editing and enhancing of TIFF raster images can successfully transform a DRG into a black & white version. Our choice is JASC Software's Paint Shop Pro (PSP), a graphics and photo editor. We first used PSP over 10 years ago when it was introduced as shareware.

PSP's "color correction" tools are easy to master, and can be used to automatically decrease a color range or to selectively replace individual colors. With a single command the color depth of the DRG is decreased to a single bit, thereby saving the trouble of manually replacing colors. But there is a trade-off with this one-step approach—a loss of image quality. The challenge is to reduce the color depth, save time and effort in editing, but to maintain the best possible image quality. A good compromise is to decrease the color depth to a 4-bit image. The resulting image has 16 colors, a manageable number of colors to manually edit, and more pixels are retained that a one-step reduction would otherwise erase. Editing each of the 16 colors is possible through the "Edit Palette dialog box" using the eyedropper tool.

The remaining editing steps are manual, with the first being the replacement of the green color tint that represents the vegetation. Far more tedious editing is required to delete the water body fill (blue color) and other patterns (e.g. tailings, hachured or stippled urban areas) with the erasing tool. It's a matter of personal choice whether the water fill and patterns are erased from the DRG. For presentation purposes, the DRG certainly looks better without these areas converted to blackened shapes. The final editing step is to return to the "Edit Palette dialog box" and finish replacing the colors that are left in the image to either black or white using the eyedropper tool.

COLOR TRANSPARENCY

Now the DRG contains only two possible pixel values: 0 and 1—black or white. Black pixels represent topographic and physical features, and white pixels represent the background. With this configuration, the DRG is ready for display in MapInfo Professional (MI), our desktop mapping tool. In MI, you can adjust the percentage of translucency of the DRG, which allows the geology layer (opaque polygons) below the DRG to become visible through the DRG's background.

Not every cartographer, geologist, or map is best served by MI's capabilities, and so AI is sometimes used for the final cartography. This part of the paper will focus on the representation of a black & white DRG in AI.

First, adjust the transparency of the DRG in order to see the geology in the layer below the DRG. AI offers two transparency controls – opacity and blending – that can adjust the degree of transparency of the DRG layer, geology, or both. However, applying one or both controls to either the DRG layer or geology can present particularly knotty problems. Some of these problems are:

- A "ghost-like" map may result when adjusting the opacity of a layer(s) or objects.
- There is a color change when a blending mode (e.g. multiply) is applied. The blending mode compares the color values from overlapping pixels and processes the two pixels into a different color amount depending on color and brightness.
- Before output to a plotter, AI performs a process called flattening to those objects modified by its transparency controls. In this process, AI automatically determines whether to retain the transparent objects and overlapping objects as vectors or convert them into a raster image. As the geology map becomes more complex (mixing images, vectors, type, spot colors, and so on), so does the flattening process and the time it takes to spool a print file.

Fortunately, there is a work-around that avoids the use of AI's transparency controls. The work-around is possible because PSP's transparency tag is compatible with AI's format. Before the DRG is placed in AI, we simply apply PSP's "set color transparency" function to the DRG's background (white pixels) for transparency. A tag is added to the DRG indicating that the white pixels have no color. AI does not assign a color value to them. As a result, the geology beneath the white pixels is visible through the DRG's background. The DRG's black pixels remain opaque.

Finally, a comment regarding the contrast between the black & white DRG and the geology is needed. The DRG's black pixels can be more prominently visible than is desired, especially for urban areas, steep topographic relief, or a combination of both. To reduce the intensity of the black pixels, we can use AI's color palette. With the layer containing the DRG selected, bring the fill box in the toolbar to the front. Click on the fill box and open the color palette menu. As a general rule, we set the K% value (black) to between 50 and 70. This range corresponds to a gray value. As a result, the DRG becomes slightly less prominent, allowing a clearer view of the geology beneath.

SOFTWARE CITED

Adobe Illustrator, Adobe Systems Inc., 345 Park Ave., San Jose, CA 95110-2704, accessed at http://www.adobe.com/.
Jasc Paint Shop Pro, Jasc Software Inc., 7905 Fuller Road. Eden Prairie, MN 55344-2697, accessed at http://www.jasc.com/.
MapInfo Pro, MapInfo Corporation, One Global View, Troy, NY 12180-8399, accessed at http://www.mapinfo.com/mipro/.

Oregon Statewide Geologic Map Data: A pilot project where digital techniques changed the geologic map compilation process and product

By Mark L. Ferns, Ronald P.Geitgey, Margaret D. Jenks, Lina Ma, Ian P. Madin, Vicki S. McConnell, and Paul E. Staub

Oregon Department of Geology & Mineral Industries
800 NE Oregon Street, Suite 965
Portland, OR 97232
Telephone: (503) 731-4100
Fax: (503) 731-4066
e-mail: paul.staub@state.or.us
Website: www.oregongeology.com

The Oregon Department of Geology and Mineral Industries (DOGAMI) has begun a six-year project to digitally compile geologic data for the entire state. This effort brings together the best available geologic mapping from state and federal sources, student thesis work, and consultants. The project will create a new statewide digitally-compiled geologic map coverage that will become a component of the Geoscience Theme within the Oregon Framework Themes (Figure 1). It will also improve the only digital statewide coverage, the 1:500,000-scale state geologic map.

To accomplish this project, DOGAMI is working in partnership with the USGS National Cooperative Geologic Mapping Program's STATEMAP component and the Oregon Geographic Information Council (OGIC). This partnership shares the funding burden of this ambitious effort, and provides a review process to ensure that the resulting data is consistent in structure, fully documented, and serves the greatest number of potential users.

COMPILATION METHODOLOGIES

The following list of steps defines the usual way in which new small scale geologic compilation maps have been produced in the past. This process is referred to as "the conventional method".

- Give the best available geologic maps to a geologist or team of geologists.
- The geologist(s) then compiles a new, seamless map by drawing new linework at a particular scale and assigning new unit labels to each polygon.
- The new unit linework is then digitized and, based on the newly written explanation of map units, the unit information is entered into a database.

The DOGAMI compilation team decided to use a different method to compile a new statewide digital map. This process is referred to as "the Oregon Pilot method." Many of the differences between this method and the conventional method were driven by the expanded opportunities created by providing a digital-only version of the map. The compilation team used the following list of steps to make the statewide digital product, using the Oregon Pilot method.

- Digitize the original polygons/units for each of the best available source geologic maps (Figures 2 and 3).
- Enter the information from the source map author's explanation of units, into a relational database (Figure 4).
- Rank the maps in terms of the quality of the geological mapping and then decide on the order of supersedure for appending the maps together. In this ranking, a newer, 1:24,000-scale map by a professional geologist would replace an area or part of an older map at a smaller scale or one created by a graduate student.
- Put the best available polygonal/spatial information together into a single layer, primarily using the more detailed or better quality maps, but retaining the less detailed or poorer quality maps in areas where no other coverage is available. This process creates an "appended" map that contains all of the best geologic unit polygons (Figure 5).
- Create new compilation merge unit labels for each of the original source map unit polygons that have been appended together in the map. These new labels effectively merge the units from all of the different source maps into a coherent stratigraphic

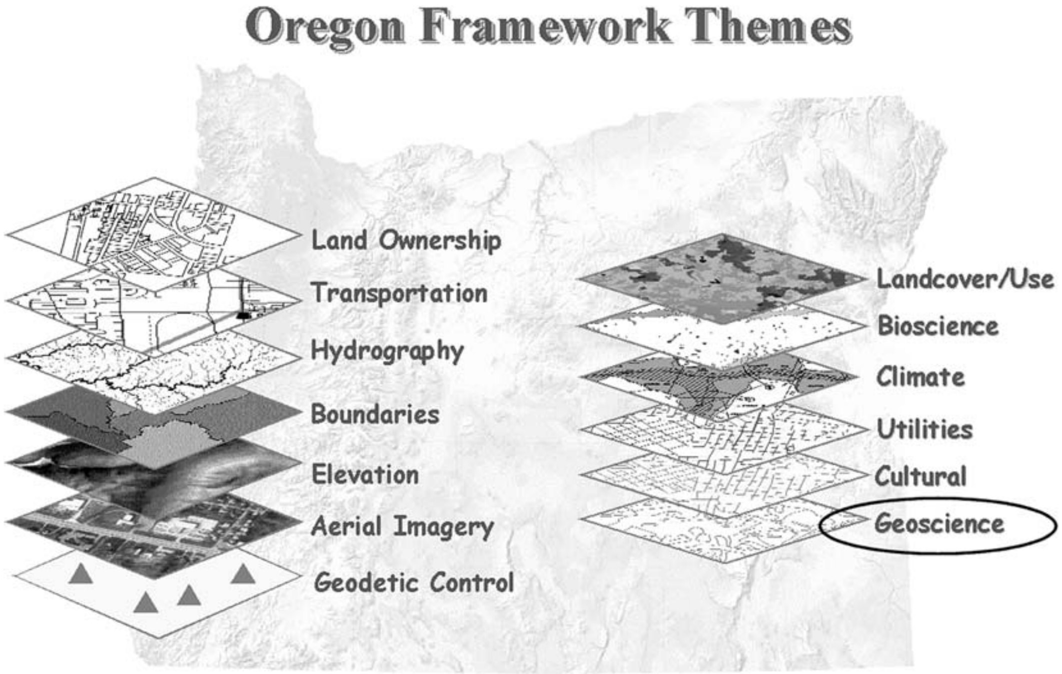


Figure 1. The Oregon Geographic Information Council has identified for statewide development thirteen Framework Themes. A workgroup for each theme is charged with developing a content standard and implementation plan. Geoscience members are from state and federal natural resource and transportation agencies, as well as academia. The Geoscience Theme presently consists of Geology and Soils layers.

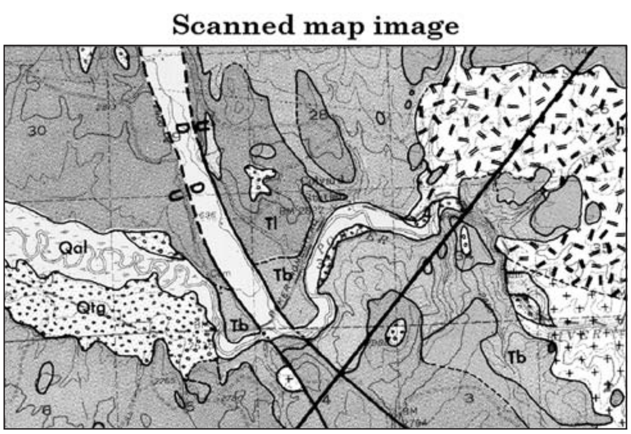


Figure 2. Image of part of a published geologic map. Raster scanning yields a high resolution image which then is georeferenced and projected prior to vectorizing the linework (image projection performed using Blue Marble Geographics software).

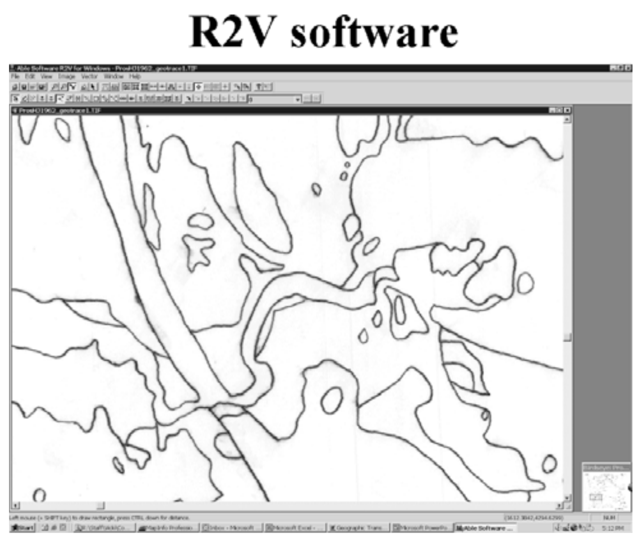
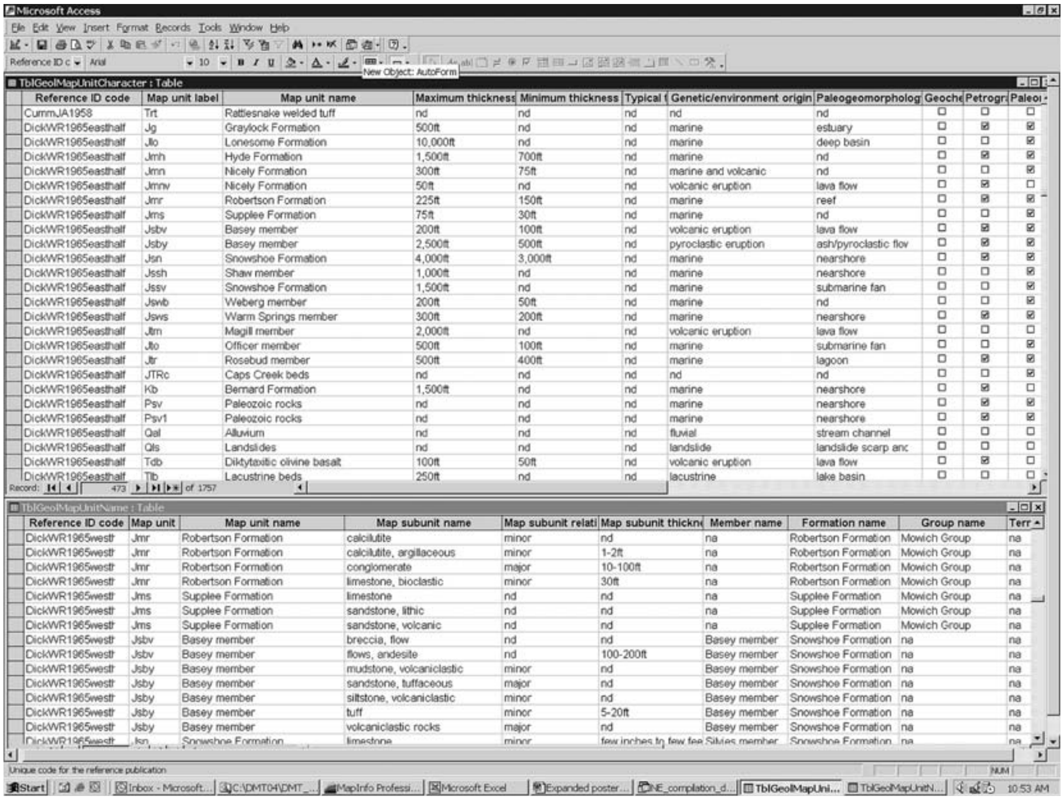


Figure 3. Image of the traced/vectorized linework from the geologic map shown in Figure 2. Conversion to vector format is done through on-screen digitizing or through use of R2V software (Able Software Corp.).



Reference ID code	Map unit label	Map unit name	Maximum thickness	Minimum thickness	Typical t	Genetic/environment origin	Paleogeomorpholog	Geoche	Petrogr	Paleo
CummJA1958	Trt	Rattlesnake welded tuff	nd	nd	nd	nd	nd	☐	☐	☐
DickWR1965easthalf	Jg	Graylock Formation	500ft	nd	nd	marine	estuary	☐	☑	☑
DickWR1965easthalf	Jlo	Lonesome Formation	10,000ft	nd	nd	marine	deep basin	☐	☐	☑
DickWR1965easthalf	Jmh	Hyde Formation	1,500ft	700ft	nd	marine	nd	☐	☑	☑
DickWR1965easthalf	Jmn	Nicely Formation	300ft	75ft	nd	marine and volcanic	nd	☐	☐	☑
DickWR1965easthalf	Jmnv	Nicely Formation	50ft	nd	nd	volcanic eruption	lava flow	☐	☑	☐
DickWR1965easthalf	Jmr	Robertson Formation	225ft	150ft	nd	marine	reef	☐	☑	☑
DickWR1965easthalf	Jms	Supplee Formation	75ft	30ft	nd	marine	nd	☐	☐	☑
DickWR1965easthalf	Jsbv	Basey member	200ft	100ft	nd	volcanic eruption	lava flow	☐	☑	☑
DickWR1965easthalf	Jsby	Basey member	2,500ft	500ft	nd	pyroclastic eruption	ash/pyroclastic flov	☐	☑	☑
DickWR1965easthalf	Jsn	Snowshoe Formation	4,000ft	3,000ft	nd	marine	nearshore	☐	☑	☑
DickWR1965easthalf	Jssh	Shaw member	1,000ft	nd	nd	marine	nearshore	☐	☐	☑
DickWR1965easthalf	Jssv	Snowshoe Formation	1,500ft	nd	nd	marine	submarine fan	☐	☐	☑
DickWR1965easthalf	Jswb	Weberg member	200ft	50ft	nd	marine	nd	☐	☐	☑
DickWR1965easthalf	Jsws	Warm Springs member	300ft	200ft	nd	marine	nearshore	☐	☑	☑
DickWR1965easthalf	Jtm	Magill member	2,000ft	nd	nd	volcanic eruption	lava flow	☐	☐	☐
DickWR1965easthalf	Jto	Officer member	500ft	100ft	nd	marine	submarine fan	☐	☐	☑
DickWR1965easthalf	Jtr	Rosebud member	500ft	400ft	nd	marine	lagoon	☐	☑	☑
DickWR1965easthalf	JTRc	Caps Creek beds	nd	nd	nd	nd	nd	☐	☐	☑
DickWR1965easthalf	Kb	Bernard Formation	1,500ft	nd	nd	marine	nearshore	☐	☑	☐
DickWR1965easthalf	Psv	Paleozoic rocks	nd	nd	nd	marine	nearshore	☐	☑	☑
DickWR1965easthalf	Psv1	Paleozoic rocks	nd	nd	nd	marine	nearshore	☐	☑	☑
DickWR1965easthalf	Qal	Alluvium	nd	nd	nd	fluvial	stream channel	☐	☐	☐
DickWR1965easthalf	Qls	Landslides	nd	nd	nd	landslide	landslide scarp anc	☐	☐	☐
DickWR1965easthalf	Tdb	Diktytaxitic olivine basalt	100ft	50ft	nd	volcanic eruption	lava flow	☐	☑	☐
DickWR1965easthalf	Tlb	Lacustrine beds	250ft	nd	nd	lacustrine	lake basin	☐	☐	☐

Reference ID code	Map unit	Map unit name	Map subunit name	Map subunit relati	Map subunit thickn	Member name	Formation name	Group name	Terr
DickWR1965westh	Jmr	Robertson Formation	calcilutite	minor	nd	na	Robertson Formation	Mowich Group	na
DickWR1965westh	Jmr	Robertson Formation	calcilutite, argillaceous	minor	1-2ft	na	Robertson Formation	Mowich Group	na
DickWR1965westh	Jmr	Robertson Formation	conglomerate	major	10-100ft	na	Robertson Formation	Mowich Group	na
DickWR1965westh	Jmr	Robertson Formation	limestone, bioclastic	minor	30ft	na	Robertson Formation	Mowich Group	na
DickWR1965westh	Jms	Supplee Formation	limestone	nd	nd	na	Supplee Formation	Mowich Group	na
DickWR1965westh	Jms	Supplee Formation	sandstone, lithic	nd	nd	na	Supplee Formation	Mowich Group	na
DickWR1965westh	Jms	Supplee Formation	sandstone, volcanic	nd	nd	na	Supplee Formation	Mowich Group	na
DickWR1965westh	Jsbv	Basey member	breccia, flow	nd	nd	Basey member	Snowshoe Formation	na	na
DickWR1965westh	Jsbv	Basey member	flows, andesite	nd	100-200ft	Basey member	Snowshoe Formation	na	na
DickWR1965westh	Jsby	Basey member	mudstone, volcaniclastic	minor	nd	Basey member	Snowshoe Formation	na	na
DickWR1965westh	Jsby	Basey member	sandstone, tuffaceous	major	nd	Basey member	Snowshoe Formation	na	na
DickWR1965westh	Jsby	Basey member	siltstone, volcaniclastic	minor	nd	Basey member	Snowshoe Formation	na	na
DickWR1965westh	Jsby	Basey member	tuff	minor	5-20ft	Basey member	Snowshoe Formation	na	na
DickWR1965westh	Jsby	Basey member	volcaniclastic rocks	major	nd	Basey member	Snowshoe Formation	na	na
DickWR1965westh	Jsn	Snowshoe Formation	limestone	minor	few inches to few fee	Silvies member	Snowshoe Formation	na	na

Figure 4. Screenshot of two of the Oregon Pilot method's Microsoft Office Access database tables, showing the typical data entry method and language.

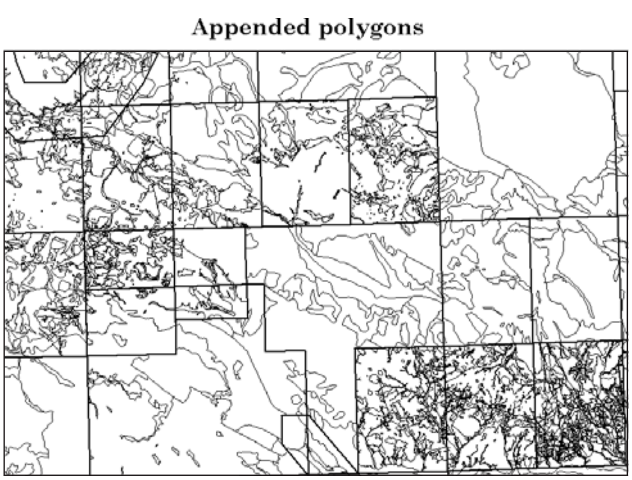


Figure 5. Screenshot of polygons appended from various source maps into the draft digital compiled map. Bold, rectangular lines are the neatlines of original geologic maps. Fine lines are appended polygon boundaries.

or lithologic framework, thus creating "logical seamlessness" for the map. Logical seamlessness occurs when a number of source datasets are integrated into one and the resulting disjointed features are not edgematched (modified geometrically) to fit together. Instead, these features are associated through attributing (FGDC, 1995). The merge unit labels are based on the current understanding of the geologic history of the area, as well as any new geochemical analyses. Professional geologists, who have been working in and have a broad understanding of the geology of a particular area of the state, assign the compilation merge unit labels. Several types of compilation merge unit labels can be made or modified to suit the end user's needs, but the DOGAMI product includes labels for geology, lithology and general geologic type (sedimentary, volcanic, intrusive, etc.).

GOALS OF THE OREGON PILOT METHOD

Several of the compilation team members have experience in making compilation maps using the conventional methodology. Out of these experiences grew the goals of the compilation project and therefore of the Oregon Pilot methodology:

- New geologic information is always becoming available. Therefore, the process must break free of a methodology that requires recompiling the state's geologic information from scratch every few years.
- As the complexity of management decision-making increases, the need to factor in geologic information becomes more widespread among different governmental agencies and non-governmental organizations. The method must create a product that is readily understood and can be easily used by a wide range of disciplines, not just by geologists.
- The new Oregon Framework Themes process provides statewide coverage of the best available information for each of the themes. Thus, the geologic layer in that Framework must be capable of being constantly updated, in order to provide the most current or "best" geologic information. The Oregon Framework themes will serve not only state decision-making processes, like the Oregon Department of Transportation's siting of a new aggregate pit, but possibly will also be useful to federal and local efforts.
- Geological science uses complex, often difficult-to-understand vocabulary. To minimize confusing terminology, we have limited the amount of non-mnemonic language in the database information.
- The compilation map also must refrain from re-interpreting the original authors' map polygons or the explanation of units. Thus, maintaining a linkage to the source maps and to their authors' original interpretations is a very important part of the methodology.
- The task of putting together a new statewide layer, especially one that is accompanied by complex, descriptive information, is arduous. Therefore, the method must take advantage of the efficiencies of relational databases, i.e., the descriptive geologic information in the original source map explanation is entered into the respective database tables only one time for each unit.

ADVANTAGES AND DISADVANTAGES

Each of these two methods has advantages and disadvantages. Of course, we chose to use the "new" Oregon Pilot method because we felt that its advantages outweighed its disadvantages. The advantages and disadvantages of each method are listed below.

Conventional method

Advantages

- It provides seamless, uniform coverage without "map faults" between the different geologic interpretations and/or map scales.
- It is simple to use because it contains only a single set of descriptive data for every compiled map unit.
- It speeds up and simplifies the process of entering the digital information because it requires entering only a single set of polygons and a single set of map unit descriptors.

Disadvantages

- The map is not updateable. It is a snapshot of the best geologic information available at a particular time. New mapping cannot be added to it. Every new compilation project must start with the original source maps.
- The varying quality of the geologic information is not transparent. The compiled map does not necessarily retain either the original author's polygons or unit descriptions. The seamless coverage at a single, small scale masks the areas of poor quality mapping.
- The final product is not flexible. The compiled data and map unit labels result only in a single stand-alone, conventional geologic map, not providing for other derivative map products.
- It is not scaleable. The map is a single, fixed scale, and does not retain the larger-scale, detailed information that may have been available in some of the original source maps.

Oregon pilot method

Advantages

- It is updateable. New mapping can be added as soon as it is completed, and any of the compilation merge unit labels can be easily changed to reflect the new information and interpretation. To make a new version of the compilation map, the method fits the new source geologic maps into the previously appended mapping. Thus, the statewide compilation map can be continuously modified/updated.
- It is transparent. The author's original polygons and descriptive explanation are in the database, and are always available to the end-user. Digital versions of the original geologic maps, both as scanned images and converted vectors, are part of the compiled map package. The compiled map also clearly conveys the areas of lesser-quality, usually small-scale mapping (see Figure 5). The obvious differences between areas of detailed versus reconnaissance mapping can direct future mapping efforts to those parts of the state with the most critical management issues, which may require mapping of higher qual-

ity and resolution than currently available.
- It is flexible. Derivative geologic or other types of maps can be made for any purpose. Users can easily modify the compilation merge unit labels which DOGAMI geologists have assigned, to fit their own stratigraphic or lithologic interpretations. Because the map is intended for use as a digital product, the compilation merge unit labels are not restricted to the length of typical geologic map unit labels (i.e., Qal). Thus, more information, like lithology, formation, age, etc, can also be conveyed in the compilation merge unit labels. Using period delimiters in the merge unit labels allows them to be parsed into individual themes that can then be made into their own derivative geologic maps.
- It is scaleable. Because it retains the original source map's information, those areas of the state that contain detailed (1:24,000) information from the original map can be used at that scale; while the compilation merge unit labels create maps that are more appropriate for intermediate-scale (e.g., 1:100,000) and small-scale (1:250,000 or greater) usage.

Disadvantages

- It produces a seamed coverage with obvious "map faults", or seams, between areas of differing original geologic interpretations and/or source map scales. Edgematching among the units of the original source maps is only addressed by the compilation merge unit labels.
- It is not a static product, so at any point in time there is no single, official "Geologic Map of the State of Oregon". Rather, versioned databases will be periodically released to keep the state's digital geologic coverage as up-to-date as possible.
- It is more difficult for the casual, non-professional audience to access and use the information. The digital seamed coverage requires that the user be capable of choosing the type of derivative map product that they want to produce, as well as the map scale displayed.
- It is not easily printable in its entirety. Local and regional land and resource management projects are the intended audience for the digital product. The entire statewide layer is too large and detailed to be printed at a single scale, and on a single sheet of paper.
- The final digital product varies in quality from one area of the map to another. The older source maps, and their explanations of map units that are entered into the database, often contain information that is from previous, now discarded, generations of geologic interpretation. However, they are still used in the compiled map because they are the best available information for that particular area.
- A large volume of information must be digitized at the beginning of the compilation process. The final digital product is a patchwork of many geologic maps instead of a single coverage; many sets of unit descriptions are attached to the merged polygons instead of a single set of unit descriptions.

ROLE OF DIGITAL CONCEPTS

As noted earlier, the digital concepts and techniques that are available now, such as raster-to-vector conversion (R2V) and relational databases, were a driving force behind our ability to create the Oregon Pilot method. Our choice to produce a digital-only statewide compilation product changed the way that we looked at compilation mapping and therefore led to the differences between the Oregon Pilot method and the conventional method. Some of those conceptual and methodological changes are listed below.

1. Digital maps do not have to be made at a particular scale and do not have to be printable on standard paper sheet sizes. Thus, they can include a range of different-scaled mapping.
2. Digital techniques make it easy to convert maps individually into digital products and then splice or append them together to make the final single layer of polygons. This simplicity allows the Oregon Pilot method to carry along, unchanged, the original source map linework and unit descriptions. Without the digital methodology, the compilation work would be forced to revert to the old method of drawing completely new linework and writing a new explanation of units.
3. Compact digital storage media (e.g., DVDs) now have sufficient capacity to store scanned and digitized original maps as well as a final, single, appended statewide digital map layer. Thus, the original source maps, which may be out-of-print or difficult to access, are more easily available to the end-user.
4. Most federal, state, and local governments use GIS systems to make management decisions. These entities need a digital geologic coverage that is as detailed as possible, and that can be easily understood by non-geologists. The appended source maps provide the best available spatial information at the largest possible scale and the greatest detail, while the new compilation merge unit labels provide the most current geologic interpretation.
5. Digital geologic data can be layered with other digital spatial themes to provide a more complete understanding of a project area or a management issue. Thus, the digital product makes both the

original source maps, and the compiled and merged data, more accessible to the end-user.

REFERENCES

Federal Geographic Data Committee, 1995, Development of a National Digital Geospatial Data Framework, section 5.2 Technical Context, accessed at http://www.fgdc.gov/framework/framdev.html.

SOFTWARE VENDOR CONTACT INFORMATION

Able Software Corp., 5 Appletree Lane, Lexington, MA 02420-2406

Blue Marble Geographics, 345 Water Street, Gardiner, Maine 04345

Microsoft Corporation, One Microsoft Way, Redmond, WA 98052-6399

Going From Greenline Mylar to Digital Geologic Map

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, the finished (but not finalized) project consisted of a text report and one or more clear or greenline mylar quadrangle maps. A finished 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 project is one that has had a more formal review and has met all the standards necessary for formal publication. Geologic contact lines, isochors, bedrock outcrop ledges, etc. were drafted directly onto the 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 lines drafted directly on them. Scanning these maps and separating the drafted line from the background was very difficult and time consuming. Some semi-automatic digitizing was tried, but most of the digitizing had to be done by hand.

The problem was studied and, after many trials, a solution emerged. New and improved scanning techniques and software solved the problem of digitizing lines that had been drafted directly on the greenline mylar. The drafted lines could be separated 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). Much improved auto-tracing software that allowed interactive image editing was also a great step forward. Now, we include many different georeferenced and attributed data layers in the digital map products released in our open-file series.

THE PROCESS

Heads-up digitizing of a scanned image is generally a straightforward process. 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 tracer will follow ones or zeros, but not the number ranges associated with color designations. Producing such an image from a greenline mylar can be a difficult task.

During the scanning operation, a threshold setting determines the sensitivity of the scanner. The threshold sets the values used 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 the black object lines to be dropped out in another area. Setting it too low will increase noise (speckling) and will pick up unwanted background lines. A "happy medium" can be elusive.

Many of the new scanning interfaces have an automatic thresholding feature. During the scanning process, the scanner will analyze small sections of the scanned object map and determine the optimal threshold setting for that section within a variability range that is set at the interface. By 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 and older maps that tend to degrade to a yellowish color, it worked very well for us in dropping the greenline background from our mylars and keeping the black contact lines.

We used a Vidar Titan II scanner. 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.

Obtaining a good scan also depends on an effective contrast setting. The scanner illuminates the mylar being scanned but, because the mylar is translucent, random background noise is picked up from the light bouncing back from the paper hold-down bar of the scanner. To correct this, we made a sheath by folding a large piece of clear acetate in half. We then placed a clean sheet of plotter paper behind the mylar and placed both inside the acetate sheath. This gave us a very clean, white background behind the mylar, and protected the mylar during scanning. This method is also very useful when scanning delicate maps and papers. The acetate sheath holds loose or tattered pieces in place and protects them from damage that the scanner rollers may cause.

Georeferencing the scanned image is an essential step in this process. Before scanning, we had to ensure that the reference tic marks were clearly visible on the mylars. The reference tic marks on the greenline mylar also were green. We had to change them to black by either hand drafting or by rub-on transfer so they would be detected by the scanner.

Once the scanning was completed, we brought the images into ArcMap 8.3 and georeferenced them to a 2.5-minute georeferenced point and line grid. Once the images were georeferenced, we prepared them for vectorization by converting them into grids.

Vectorization of most of the linework was accomplished using the ArcScan module extension of ArcMap 8.3. ArcScan allowed us to do interactive raster editing and clean-up while previewing how ArcScan was going to vectorize the lines. Raster line intersections have always been one of the hardest things for ArcScan to interpret. "T" intersections would commonly have a deep "V" in them where the tracer would move to the center of the pixel cluster in the middle of the "T" before continuing down the pixel line. Also, lines intersecting at low angles often have pixel in-fills between the lines as they approach the actual line intersection. The tracer often interprets the line intersection short of its actual location and at a larger angle than intended. Interactive raster editing allows you to remove the in-fill pixels between the low angle lines and clean up a "T" intersection, while observing how ArcScan intends to interpret and vectorize the intersection. This allows you to obtain a vector trace that is truer to the actual intersection, and avoid time-consuming clean-up of the vector lines.

Once the raster editing was done, we used the ArcScan automatic vectorization feature to vectorize the entire scan. The results were very good, but some final clean-up was necessary. Discontinuous lines, points, and other features were digitized by hand. Line and point placement were checked for accuracy and corrected where necessary. Individual data layers were attributed where appropriate and checked for accuracy. Author review of the vector files in some cases resulted in changes and clarifications. Editing was easily accomplished on the digital files.

The overall goal of this project was to quickly create digital data and release it to the public without the delays involved in creating a formal publication. Caveats apply until formal publication, but the data is quickly available. Nevertheless, many in the user community prefer hard copy paper maps. For this reason, a generic map template was created in which any of the thirty open-filed quads can be placed, and with minimal editing and adjustment, printed within several hours. The map document is saved as a PDF file and can be reprinted at any time.

For those using the digital data in ArcGIS, we include the ArcMap MXD (ArcMap document) file in our open-file data release. We also include a PMF (ArcPublisher created) file for use with ArcReader. ArcReader, a free download from ESRI, is a limited version of ArcGIS that allows the user to view the data, do some limited data manipulation, and print the map in the same template that was used to create it in ArcGIS.

A text write-up of formation descriptions, stratigraphy, structure, geologic settings, etc. also is available as a separate document for each quad. By not including much text in the marginalia, map production time is greatly reduced.

A 1:100,000-scale composite map template also was constructed for regional studies. As digital conversion of the 1:24,000-scale quads is completed, their digital files are referenced in the 1:100,000-scale map MXD template. When printed, the 1:100,000-scale map appears to be a composite, but is actually a mosaic of the many digital 1:24,000-scale quadrangles. Printing this map as a mosaic instead of a composite adds to the processing overhead. In a composite map, polygons of the same formation along quad boundaries are merged together. In a mosaic, like polygons along the quad boundaries are not merged, and remain separate, thereby increasing the amount of data to be processed for a print file.

The composite map is also self-correcting. Because the 1:24,000-scale quadrangle data sets are referenced by the 1:100,000-scale composite map MXD file, edits and corrections are reflected in the composite map as they are completed on the 1:24,000-scale data. Because this map is a preliminary, open-file release, caveats are included in the map template margin and in the metadata. These caveats warn of possible edgematching problems along quad borders and the changeable nature of the data.

CONCLUSION

One of the uses of the Open File Report series has been to disseminate maps and data that are incomplete or are finished and waiting for formal publication. Usual distribution methods included photocopied maps, map composites, notes, etc., where the user could obtain the data, with various caveats. The increasing demand for digital data, however, made distribution of analog paper reports and maps less desirable.

Using some simple techniques, we were able to vectorize various map layers into a digital format. Development of a generic map template allows us to "make a map" by dropping the various map layers into the template.

In northeastern Pennsylvania, we have more than 30 USGS 7.5-minute quadrangles already mapped in analog form. Each quad has from 3 to 5 mylar overlays along with a greenline quad base. Using the techniques described above, we were able to complete a digital rendering of a quad area in about 3 to 5 man-days.

These maps are simple but functional. The digital data is now available for those who don't want to wait for formal publication, with the typical open file caveat that the data is subject to change. Because it is digital, the data now has more utility than it had when it was available only as composite photocopies.

Geoenvironmental Map of the Christmas Point Quadrangle, Texas Gulf Coast

By Thomas A. Tremblay, William A. White,
Rachel L. Waldinger, and Edward W. Collins

The University of Texas at Austin
John A. and Katherine G. Jackson School of Geosciences
Bureau of Economic Geology
10100 Burnet Rd.
Austin, TX 78758-4445
Telephone: (512) 475-9537
Fax: (512) 471-0140
e-mail: tom.tremblay@beg.utexas.edu

INTRODUCTION

The Bureau of Economic Geology (BEG) has been mapping wetlands and aquatic habitats on Texas coastal barrier systems since 2001. Approximately one-third of the Texas coast has been mapped, and we expect to complete the entire coastal area by 2006 (Figure 1). The maps provide a recent status map of wetland and aquatic habitats and a basis for measuring wetland change over time. Wetland maps serve an important function but are limited in their use for other coastal applications. Therefore, BEG proposes a geoenvironmental map series that incorporates wetland information into a more general-purpose coastal map. Our present example of the Christmas Point, Texas, quadrangle combines barrier-island wetland habitats with bay and delta-margin geologic units. Map units are coded based on primary, well-defined mapping criteria. Wetland units are coded according to a wetland classification, and areas outside the barrier-island system are assigned traditional geologic descriptions. The database design retains original unit coding but also identifies like units within the "geoenvironmental" attribute code.

METHODS

Since 2001 the Texas General Land Office (GLO) has flown aerial photography missions in areas of particular interest along the Texas coast. The photography program produces high-resolution (1-meter), color-infrared, digital aerial photographs that are georeferenced to a 1995–1996 digital orthophoto quadrangle (DOQ) base. BEG scientists use these photographs within the GIS environment to map wetland boundaries. Wetland unit codes are based on the National Wetlands Inventory (NWI) classification of Cowardin et al. (1979).

Bay and delta-margin mapping was conducted in a like manner using the 1995–1996 DOQ base. The bay and delta-margin units are a modification of those found in the *Environmental Geologic Atlas of Texas* (Fisher et al., 1972). Marsh delineation was aided through consultation of wetland mapping conducted by White et al. (1986) in the *Submerged Lands of Texas* series.

Wetland, bay, and delta-margin mapping were merged into a seamless data set and assigned codes on the basis of newly-devised mapping criteria. Map units that exhibit similar geoenvironmental characteristics, which are based on location or physical properties, were assigned a cross-classification "geoenvironmental" code.

Surface faulting was mapped by White and Morton (1997) in a paper examining wetland losses related to fault movement. The fault data were gathered from existing GIS data sets contained within the *Coastal Hazards Atlas of Texas—Volume 1* (Gibeaut et al., 2000).

CONCLUSION

The proposed "Geoenvironmental Map" series will utilize existing barrier-system wetland and habitat mapping in conjunction with generalized geologic mapping to produce a multipurpose coastal-map series. Digitally based, the series will adapt to incorporate the best and most timely data sets available. The GIS-based map is seamless not only in spatial extent but in data content as well. Highly detailed wetland information, where available, will be retained within the database design. This aspect of the database provides additional information not found in traditional coastal-environmental maps. The BEG Coastal Studies Group owns and operates a LIDAR system that is used in various shoreline studies. Shoreline position and geomorphic features such as washover chan-

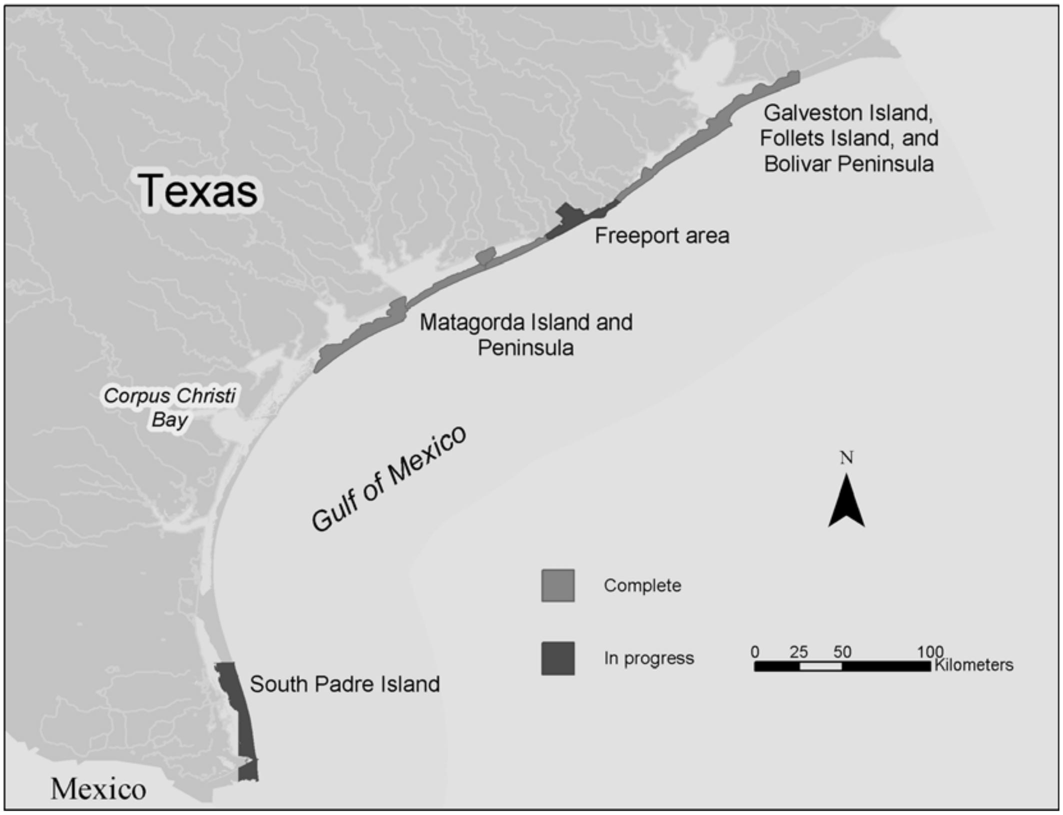


Figure 1. Status of barrier-system wetland mapping on the Texas Gulf Coast.

nels, obtained from LIDAR, will enhance the information content of future versions of the geoenvironmental map. With some regional modifications, the geoenvironmental map model can be applied to any coastal region.

ACKNOWLEDGMENTS

Wetlands mapping is funded through grants provided by the National Oceanic and Atmospheric Administration and the Coastal Coordination Council,

administered by the Texas General Land Office. The STATEMAP component of the National Cooperative Geologic Mapping Program, administered by the U.S. Geological Survey, also provided funding for this study.

REFERENCES

Cowardin, L.M., Carter, V., Golet, F.C., and LaRoe, E.T., 1979, Classification of wetlands and deepwater habitats of the United States: U.S. Department of Interior, Fish and Wildlife Service, Washington, D.C., USA, 131 p.

Fisher, W.L., McGowen, J.H., Brown, L.F., and Groat, C.G., 1972, Environmental geologic atlas of the Texas Coastal Zone—Galveston–Houston area: The University of Texas at Austin, Bureau of Economic Geology, 91 p.

Gibeaut, J.C., Tremblay, T.A., and White, W.A., 2000, Coastal hazards atlas of Texas: a tool for hurricane preparedness and coastal management—volume 1, the southeast coast: The University of Texas at Austin, Bureau of Economic Geology, final report prepared for the Texas Coastal Coordination Council, pursuant to National Oceanic and Atmospheric Administration Award No. NA770Z0202, 69 p., accessed at http://www.beg.utexas.edu/coastal/hazard_atlas1.htm.

White, W.A., Calnan, T.R., Morton, R.A., Kimble, R.S., Littleton, T.G., McGowen, J.H., Nance, H.S., and Schmedes, K.E., 1986, Submerged lands of Texas, Brownsville–Harlingen area: sediments, geochemistry, benthic macroinvertebrates, and associated wetlands: The University of Texas at Austin, Bureau of Economic Geology Special Publication, 138 p.

White, W.A., and Morton, R.A., 1997, Wetland losses related to fault movement and hydrocarbon production, southeastern Texas coast: Journal of Coastal Research, v. 13, no. 4, p. 1305–1320.

Process of Converting Geologic Maps from Coverages to Geodatabases

By Karen Wheeler, Philip Dinterman, Russell Evarts, and Ray Wells

U.S. Geological Survey
Western Earth Surface Processes Team
345 Middlefield Rd., MS-973
Menlo Park, CA 94025
Telephone: (650) 329-4935
Fax: (650) 329-4936
e-mail: kwheeler@usgs.gov

INTRODUCTION

The Pacific Northwest Urban Corridor Geologic Mapping project is providing digital 1:24000-scale geologic mapping of areas in western Oregon and Washington. Our databases are online, with documentation describing the datasets at http://geology.wr.usgs.gov/wgmt/pacnw**.**

Online releases from 1995-2002 include:

1. Database package—ArcInfo coverage export (.e00), grids, AMLs, symbolsets, readme, and metadata
2. PostScript package—Encapsulated PostScript files for printing
3. PDF package—Adobe Acrobat PDF files for viewing and printing.

In 2002, we began using ArcGIS for database management and map creation. The data structure for our databases followed the Alacarte standard (Fitzgibbon, 1991; Fitzgibbon and Wentworth, 1991; Wentworth and Fitzgibbon, 1991).

Our project has converted almost all our pre-existing coverages from Workstation ArcInfo to ArcGIS geodatabases.

Current packages produced for the website include:

1. PDFs of the geology map and pamphlet
2. Geodatabase, metadata, and ArcMap document for viewing and querying
3. Shapefiles exported from Geodatabase.

We now are adding more data to our geodatabases, improving our collection of data with ArcPad, developing styles for ArcMap which make the maps more uniform, streamlining the map production with templates, and improving our metadata. A downloadable PDF of our DMT '04 poster is available at http://ngmdb.usgs.gov/Info/dmt/docs/wheeler04.pdf (15MB).

CONVERTING COVERAGES TO GEODATABASES

A typical conversion of one geology map from coverages to a Geodatabase takes about 10 minutes or less.

1. Use ArcCatalog to create a Personal Geodatabase

A. Specify a New Personal Geodatabase (Figure 1a).
B. Add Feature Datasets such as geology, locations, and structure (Figure 1b).
C. Create a Feature Class in that Feature Dataset or import coverages into a Feature Dataset as Feature Classes (Figure 1c).
D. If desired, specify Domains for tagging in ArcMap (Figure 1d). A Domain is the range of values allowed for a column in a database.

2. Use ArcMap to symbolize a Geologic Map of the Geodatabase

A. Create a new Map Document (Figure 2a).
B. Add data from newly created geodatabase (Figure 2b).
C. Use Styles or a previous Layer file for symbology (Figure 2c).

3. Exporting from ArcMap

A. Export the map at a specified scale, to EPS or to Adobe Illustrator (Figure 2d). Place exported EPS of map in an Illustrator map template for printing.

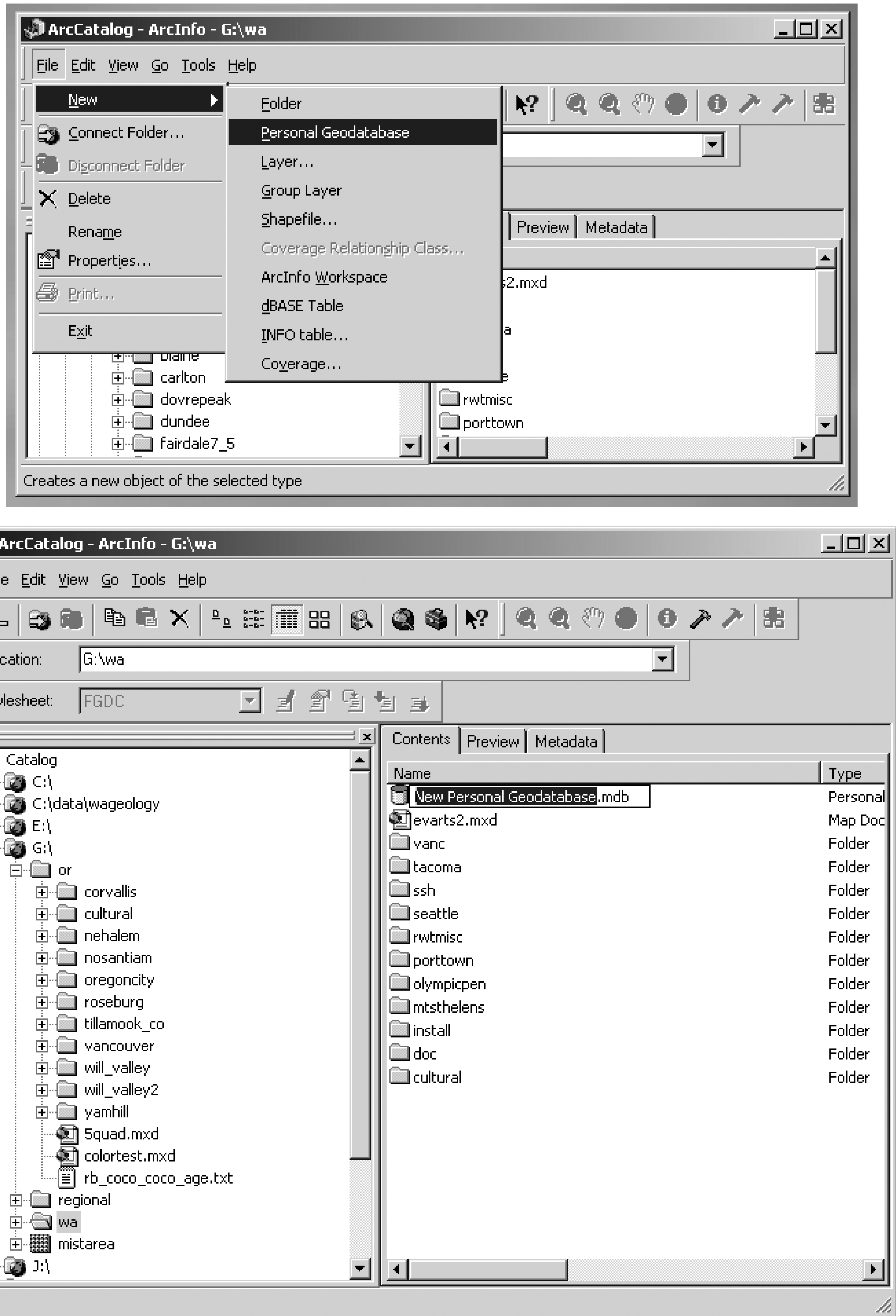


Figure 1a. Open ArcCatalog and create a Personal Geodatabase. Rename the geodatabase with a name that you or your organization prefers.

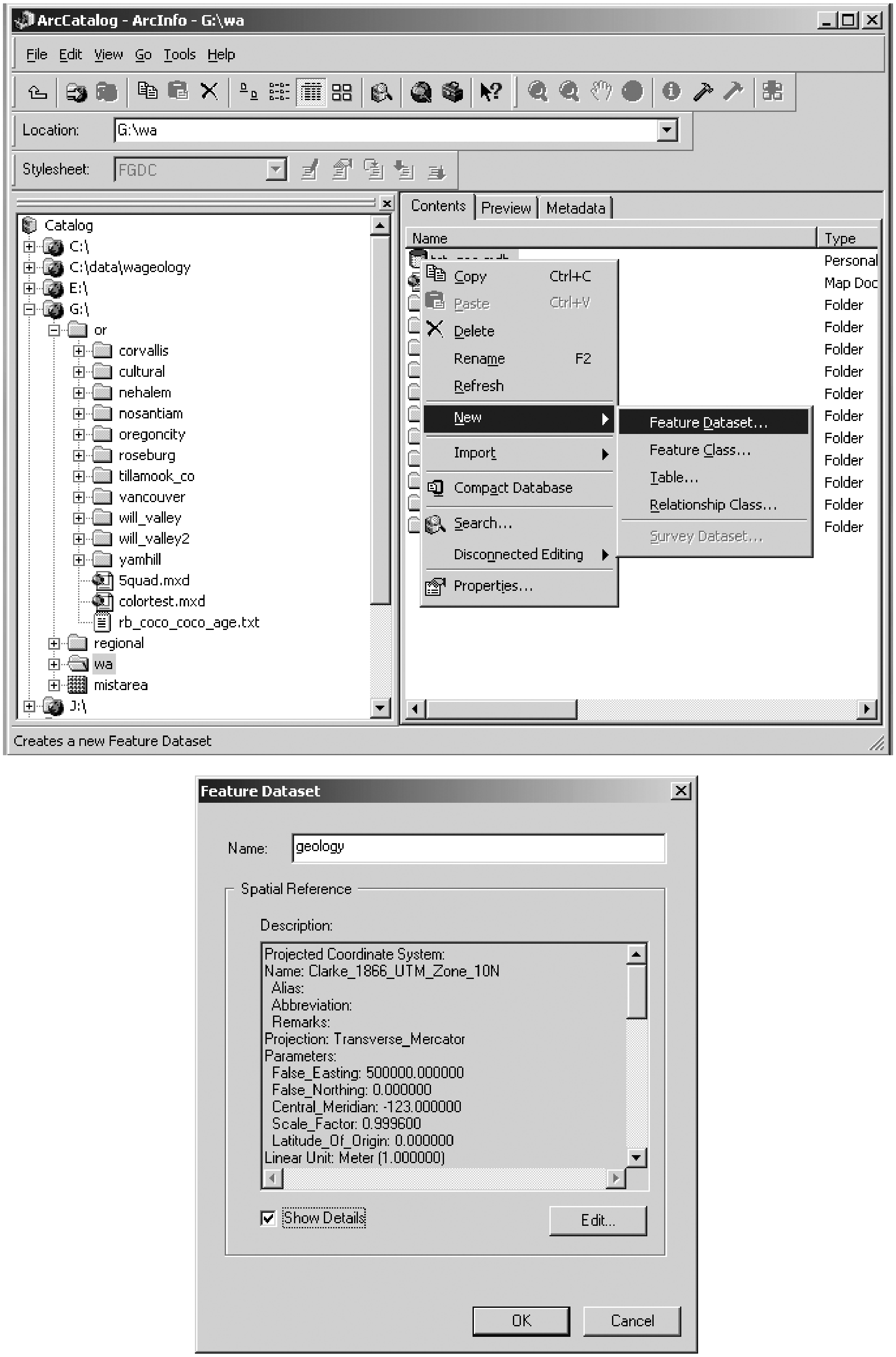


Figure 1b. Add Feature Datasets (such as geology, locations, and structure) in ArcCatalog. Create a Feature Dataset in that geodatabase (right click the name of the geodatabase). Name the Feature Dataset and click the Edit button to choose a projection.

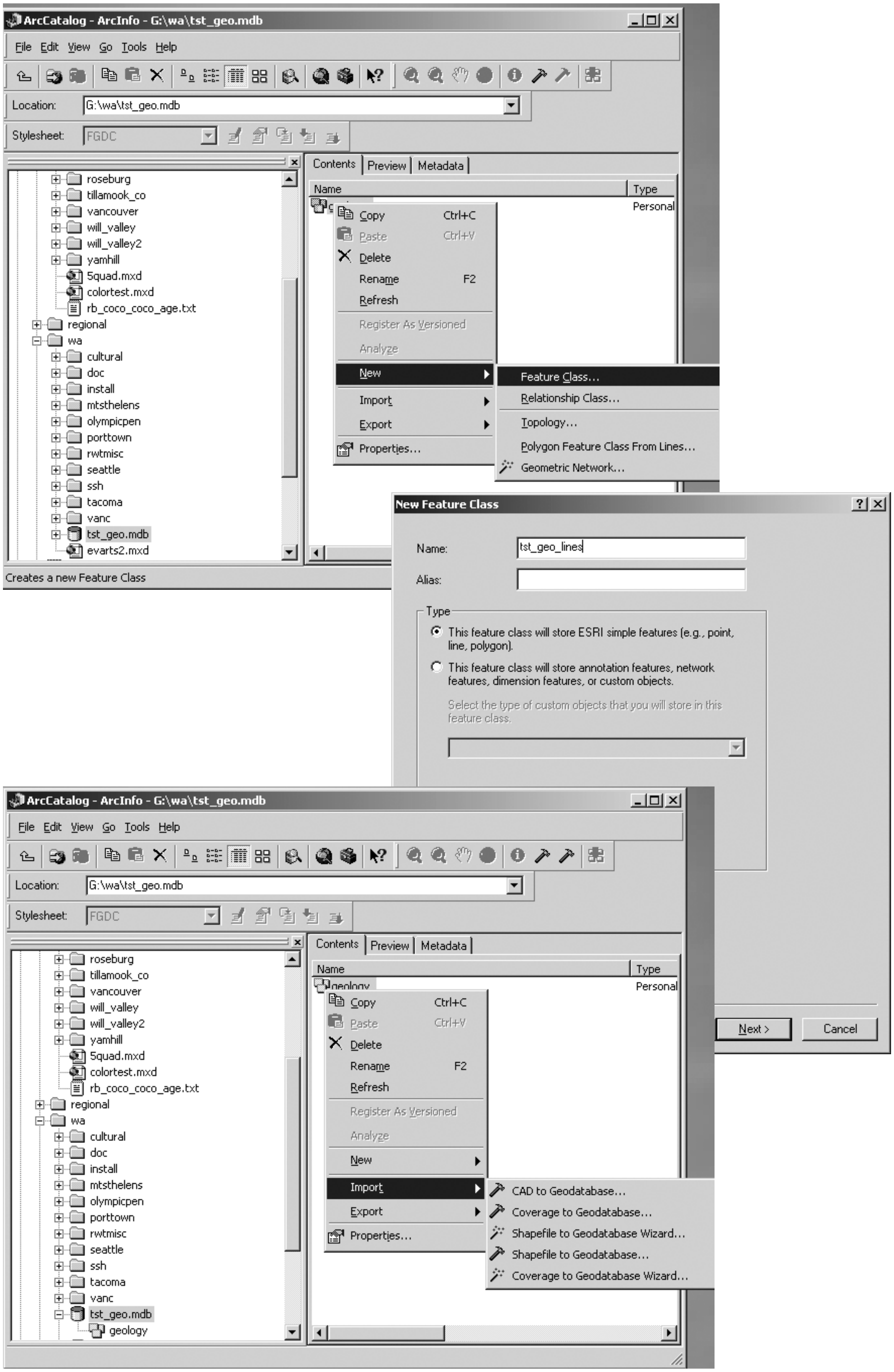


Figure 1c. Create a Feature Class in that Feature Dataset or import coverages into the Feature Dataset as Feature Classes in ArcCatalog (right click the name of the dataset). Type in the Name of the new Feature Class, then click which Type of Feature Class, and click Next button. Or you might import preexisting coverages or shapefiles to the Feature Dataset as Feature Classes.

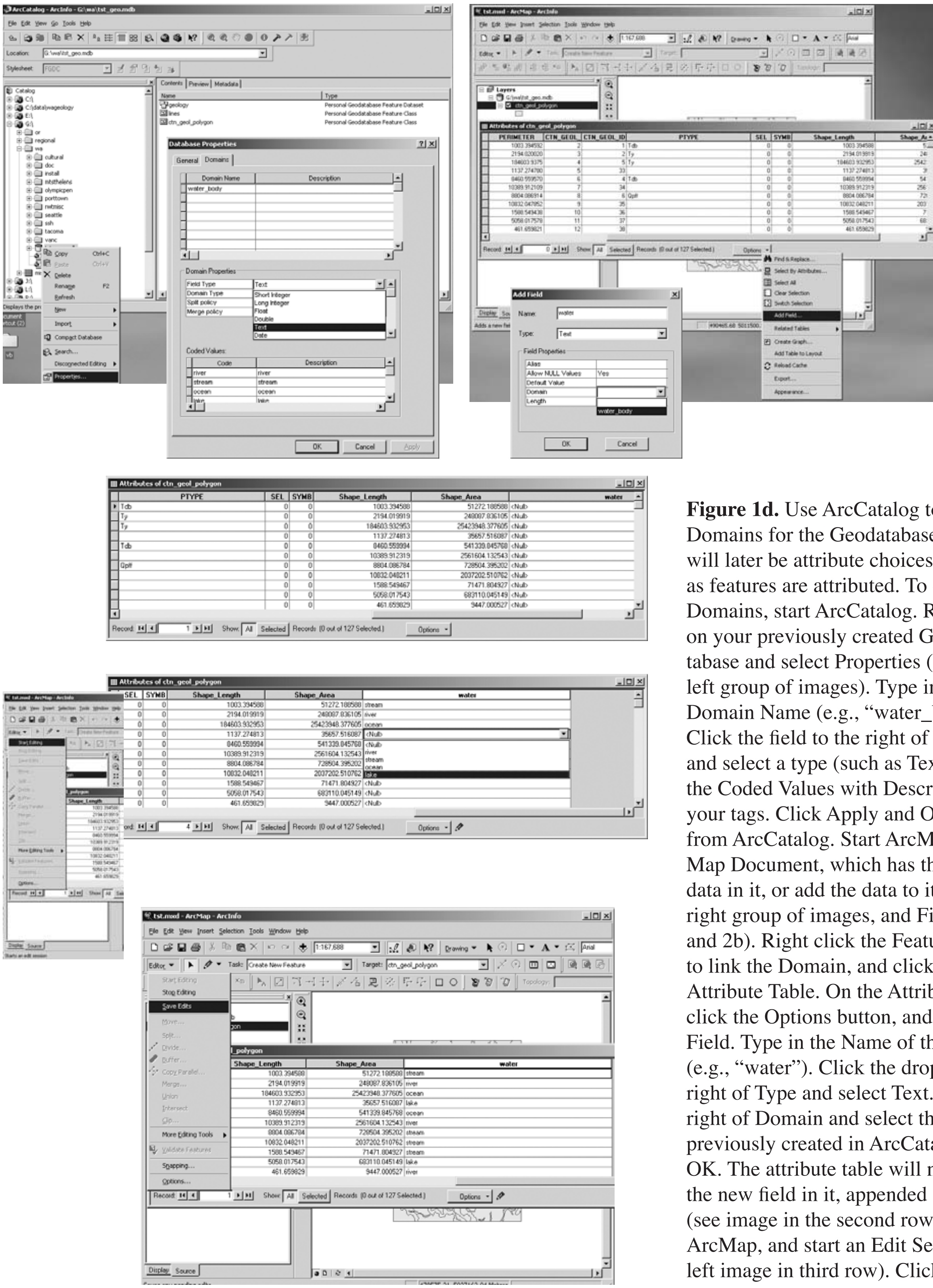


Figure 1d. Use ArcCatalog to specify Domains for the Geodatabase which will later be attribute choices in Arcmap as features are attributed. To create Domains, start ArcCatalog. Right-click on your previously created Geodatabase and select Properties (see top left group of images). Type in a new Domain Name (e.g., "water_body"). Click the field to the right of Field Type and select a type (such as Text). Enter the Coded Values with Description for your tags. Click Apply and OK, and exit from ArcCatalog. Start ArcMap, open a Map Document, which has the geologic data in it, or add the data to it (see top right group of images, and Figures 2a and 2b). Right click the Feature Class to link the Domain, and click Open Attribute Table. On the Attribute Table, click the Options button, and select Add Field. Type in the Name of the new field (e.g., "water"). Click the drop arrow to right of Type and select Text. Click to right of Domain and select the Domain previously created in ArcCatalog. Click OK. The attribute table will now have the new field in it, appended at the end (see image in the second row). Save ArcMap, and start an Edit Session (see left image in third row). Click in a record to tag, select a tag from the drop menu, and tag feature (e.g., as "lake"; see right image in third row). Save edits (see image at bottom). Save ArcMap Document.

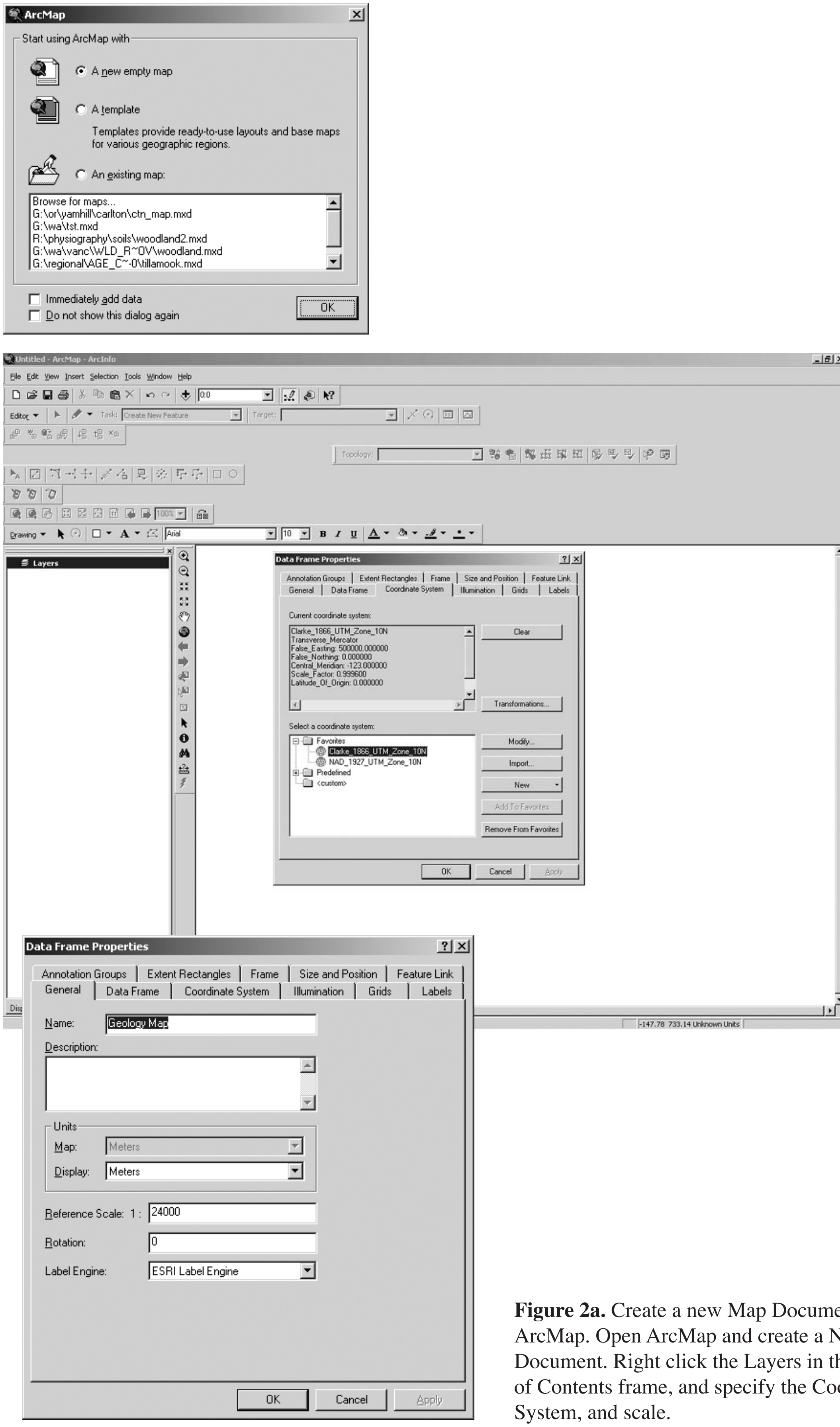
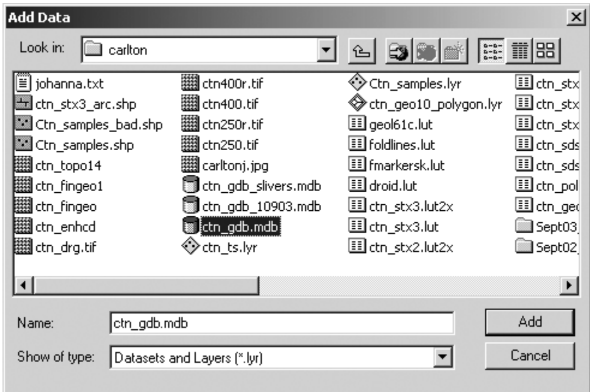
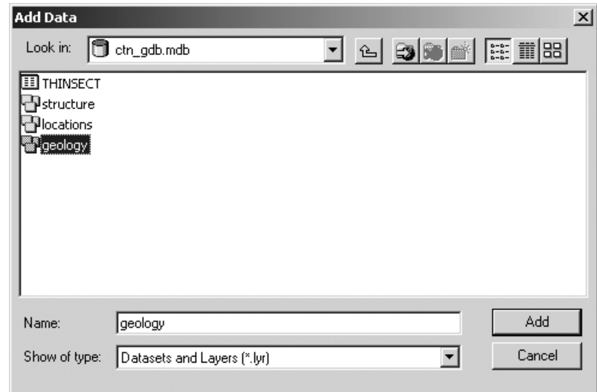
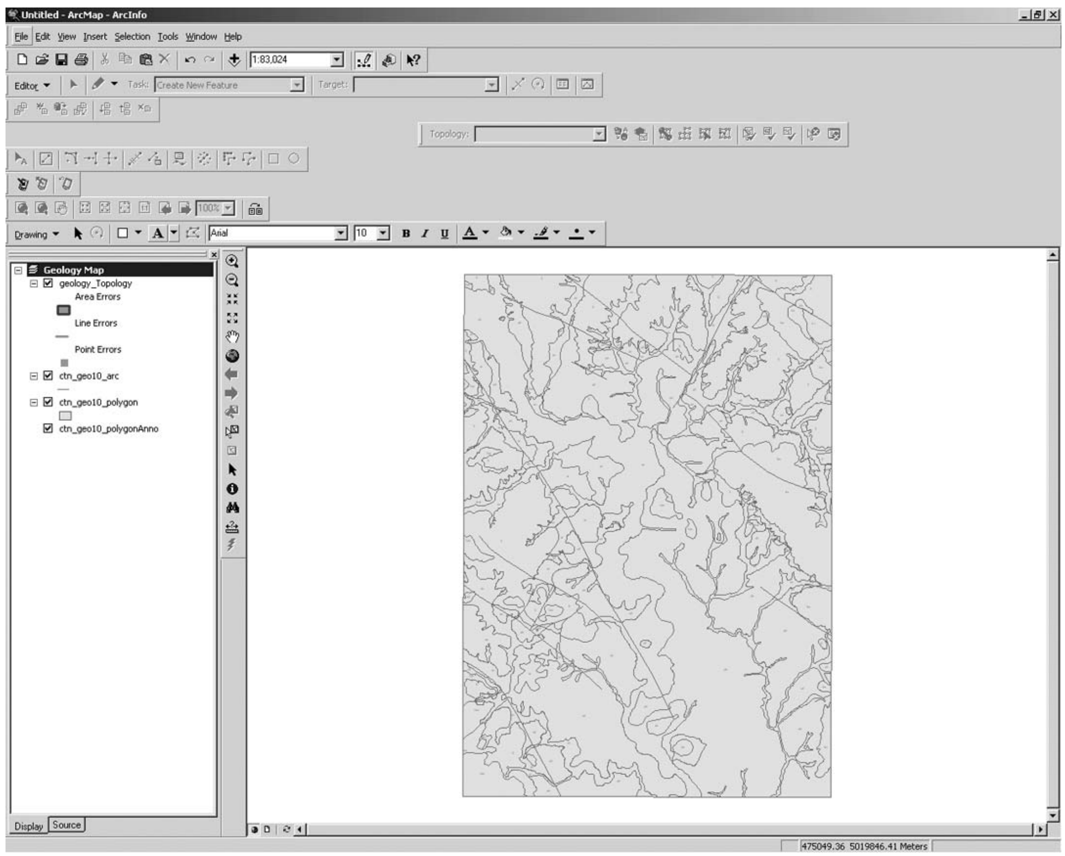


Figure 2a. Create a new Map Document in ArcMap. Open ArcMap and create a New Map Document. Right click the Layers in the Table of Contents frame, and specify the Coordinate System, and scale.







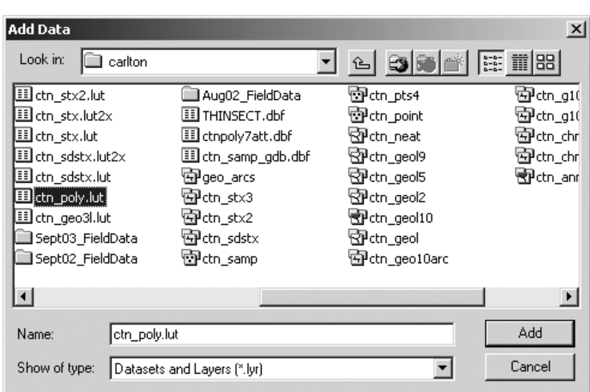
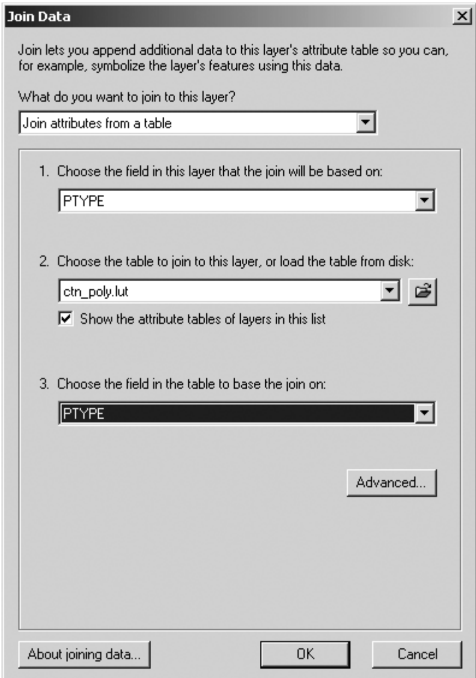
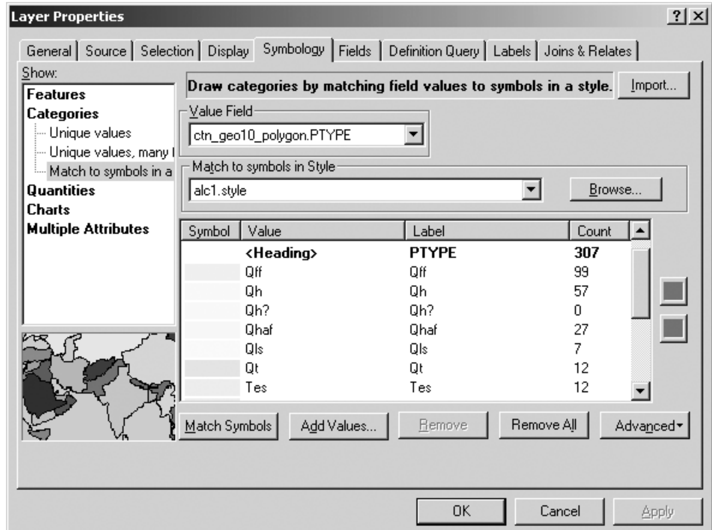
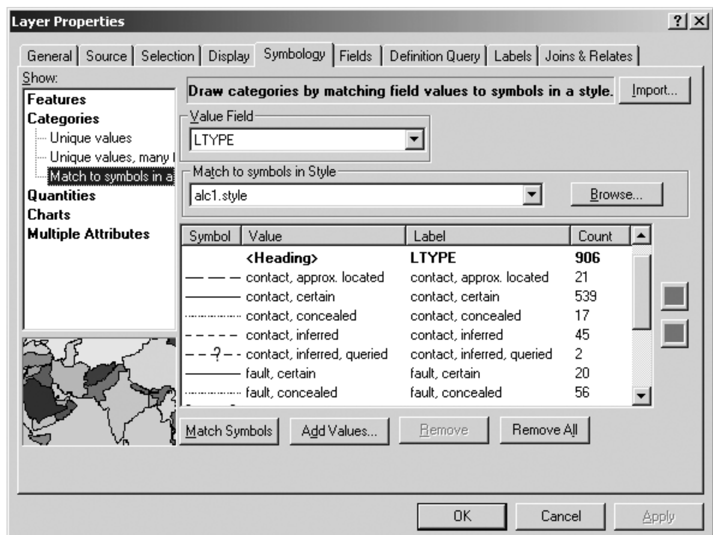


Figure 2b. Add data from newly created geodatabase in ArcMap. Add data (the yellow button with plus symbol), and select the Geodatabase previously created with Dataset (see two images in top row). It plots with default symbols (see middle image), but you can easily change that with a lookup table (see bottom image), with a style, or by specifying each one individually in the Layer Properties > Symbology tab.







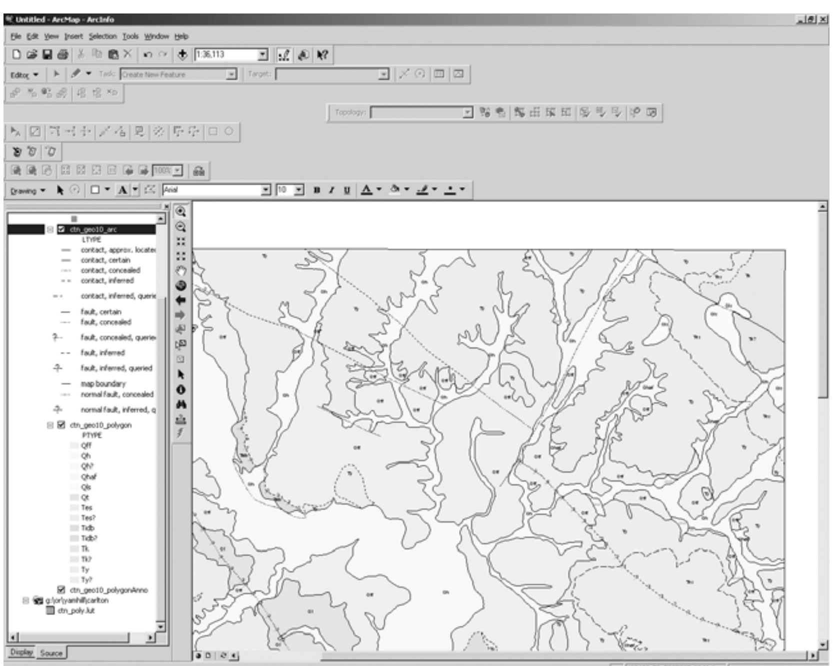
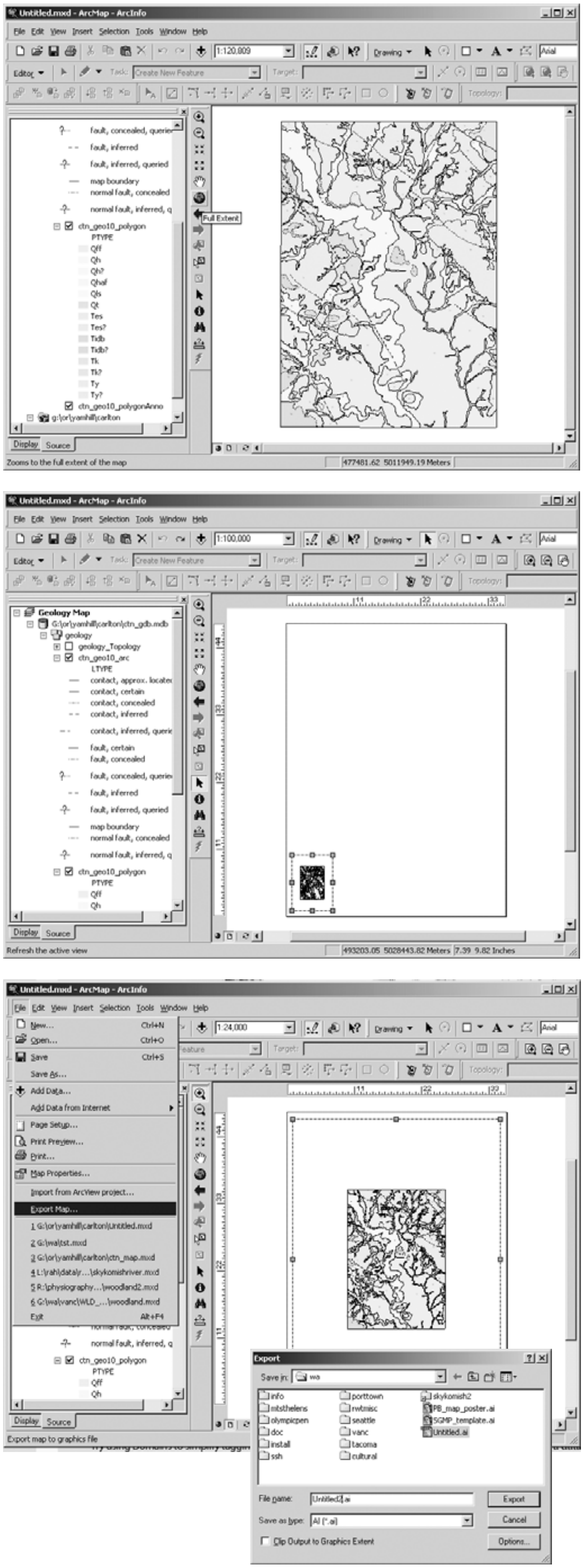


Figure 2c. Use Styles or a previous Layer file for symbology in ArcMap. To symbolize with the Lookup Table, join the table to the polygons (top left image). Bring up the Layer Properties for the Polygons, Symbology Tab (top right image), and choose “Categories > Match Symbols in a Style”. Select Value field, then the Style, and click the Match Symbols button (middle image). Styles are edited using the Style Manager in ArcMap, and may be shared among ArcGIS desktop users. Select “Tools > Styles > Style Manager” to chose a Style that you’ve received from another ArcGIS user. Right click on the Name, copy and paste, edit new symbol by Properties. The bottom image shows lines symbolized with a style, not a lookup table.



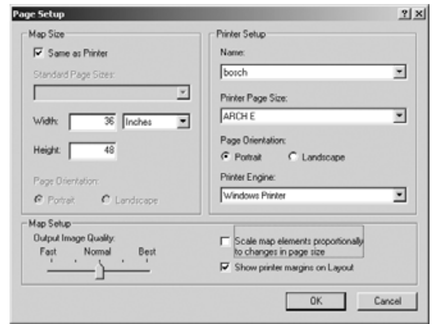


Figure 2d. Export the map from ArcMap, at a specified scale, to EPS or to Adobe Illustrator. In the Data View (top left image), click the Full Extent (Globe button) so that map is zoomed out. Then click the Layout View (second row left image). Click "File > Page Setup". Choose plotter, paper size, and orientation of page (upper right image). Also deselect the Scale map to page size box, or the map will not be a true scale. If the frame of the map is small, click the corner with the arrow tool and drag it to the larger size (bottom row left image). Then specify map scale in the box at the top. Click "File > Export", and name the file for exporting (bottom row right image). From the drop list, choose the type of file (here, Illustrator). Click the Export button. The file will be created, and should open in Adobe Illustrator. If this doesn't work well, try other formats for exporting.

PUBLISHED GEODATABASES

Our latest geologic map, published as a geodatabase, was presented on our poster at the Workshop. It was originally compiled on mylar and greenlines, scanned, and digitized into Workstation ArcInfo. It was imported into a Personal Geodatabase, edited in ArcMap, and exported to a graphic file for Illustrator.

Our project's first geodatabase, *Maps showing Inundation depths, ice-rafted erratics, and sedimentary facies of the Late Pleistocene Missoula Floods of the Willamette Valley, Oregon*, was published in 2003 and is online at http://geopubs.wr.usgs.gov/open-file/of03-408. It consists of shapefiles, a geodatabase and metadata, PDFs of maps, and a readme file.

Our most recent geodatabases are best found by visiting our website and clicking the "What's New" tab or at http://geology.wr.usgs.gov/wgmt/pacnw/new.html. We keep it updated with the latest publications.

FUTURE DIRECTIONS

We are testing ArcPad on an IPAQ-brand Personal Digital Assistant (PDA) with GPS (see paper by Thoms and Haugerud, in these Proceedings). For base maps in ArcPad, we use ArcCatalog to export the DRGs to MrSID compressed-format images**.** Forms developed in ArcPad are used to collect for each point our data for sample locations, structural measurements, geologic unit, date, time, and also general field notes. Collecting data for points in ArcMap allows us to directly use the information as Shapefiles, and saves time digitizing the points from paper field sheets as was previously done. Once we created a geologic map in ArcGIS, the Geodatabase was exported to ArcPad and used in the field on the IPAQs to add more information; later the file was reimported to the Geodatabase to update the geologic map.

On Windows CE running on Compaq IPAQs, the file navigation software was awkward, so we found it necessary to purchase an inexpensive utility called PE File (HeavenTools software, http://www.heaventools.com/), to better view filename lists with extensions. Our cost has been minimal for these IPAQs and the PE File software.

We are now trying to use ArcSDE for multiuser editing of Geodatabases and for managing large image catalogs, but this is only in the testing phase. It is simple to convert the ArcGIS Personal Geodatabases to ArcSDE databases. We are have migrated from ArcGIS 8.3 to ArcGIS 9, and are adopting a new ArcGIS geologic map database structure and symbolsets.

ADDITIONAL TIPS

- Convert SymbolSets from Workstation Arc to ArcGIS Styles, with the SymbolConverter.exe available online at http://arcobjectsonline.esri.com (full URL is: http://arcobjectsonline.esri.com/default.asp?URL=/ArcObjectsOnline/Samples/ArcMap/Symbology/Symbols/AI%207x%20Symbol%20Conversion/SymbolConverter.htm).
- Try using Attribute Domains to simplify and constrain values when tagging in ArcMap. Domains are the range of values allowed for a field in a table, feature class or subtype in a Geodatabase. You can use ArcCatalog to create Domains in a Geodatabase. A tutorial on creating Attribute Domains is available at http://www.esri.com/news/arcuser/0400/files/stutorial.pdf.
- If printing from ArcMap is problematic, export the Layout View of the map to Illustrator or EPS and then print it.
- ArcCatalog, ArcMap, or ArcToolBox work on imported versions of your coverages and shapefiles, so it is possible to export the Geodatabase and work on the files in the Workstation ArcInfo or ArcView if needed.
- Save often in ArcMap, as it can be unstable.
- Samba is an Open Source/Free Software suite that provides seamless file and print services to SMB/CIFS clients. It allows for interoperability between Linux/Unix servers and Windows-based clients. For more information on Samba see http://us1.samba.org/samba/. If your data resides remotely, on a host computer on which Samba is running, first map the network drives in Windows to the data, then start ArcCatalog and establish the Connection to Database. Close ArcCatalog and open ArcMap to work on the data.
- Take the classes online from the ESRI Virtual Campus, at http://campus.esri.com/, and do the exercises.
- Practice these procedures and teach them to someone else, and be patient. Remember how long it took to learn Arc Workstation? You will learn ArcGIS desktop much faster.

REFERENCES

Fitzgibbon, T.T., 1991, ALACARTE installation and system manual (version 1.0): U.S. Geological Survey Open File Report 91-587B, 66 p.

Fitzgibbon, T.T., and Wentworth, C.M., 1991, ALACARTE user interface - AML code and demonstration maps (version 1.0): U.S. Geological Survey Open File Report 91-587A, 10 p.

Wentworth, C.M., and Fitzgibbon, T.T., 1991, ALACARTE user manual (version 1.0): U.S. Geological Survey Open File Report 91-587C, 268 p.

APPENDIX A

List of Workshop Attendees

[Grouped by affiliation]

Arizona Geological Survey
Stephen M. Richard

Colorado State University/National Park Service
Stephanie O'Meara
Trista Thornberry-Ehrlich

County of San Luis Obispo, CA
Lewis I. Rosenberg

Diamondex Resources
David McKee

Dynamic Graphics, Inc.
Skip Pack

Earthfx, Inc.
Dirk Kassenaar

ESPRI—University of Arizona/USGS
Jonathan A. Craigue
Douglas Hirschberg
Harry McGregor
G. Stephen (Bear) Pitts

ESRI, Inc.
Alon Yaari

GeoReference Online, Ltd.
Clinton Smyth

Geological Survey of Canada
Eric Boisvert
Andrée Bolduc
Guy Buller
Robert Cocking
Peter Davenport
Parm Dhesi
Victor Dohar
Dave Everett
Terry Houlahan
Sheila Hynes
Jamel Joseph
Roger MacLeod
Andrew Moore
Marianne Quat
Barbara Szlavko
Scott Tweedy
Carol Wagner
Stephen Williams

Hewlett-Packard Company
Randy Heilbrunn

Idaho Geological Survey
Jane Freed
Kurt Othberg
Loudon R. Stanford

Illinois State Geological Survey
Sheena K. Beaverson
Jane Domier
Robert Krumm

Indiana Geological Survey
Robin Rupp

Kentucky Geological Survey
Warren H. Anderson

Louisiana Geological Survey
Robert Paulsell

Montana Bureau of Mines and Geology
Susan Smith

National Park Service
Gregory S. Mack
Anne R. Poole

National Resources Canada
Norah Brown

Nevada Bureau of Mines and Geology
Christine Arritt
Gary Johnson
Jennifer Mauldin

New Hampshire Geological Survey
Derek Bennett

New Mexico Bureau of Geology
Sandra Azevedo

Noranda-Falconbridge
Pierre St-Antoine

North Dakota Geological Survey
Fred Anderson

Ohio Geological Survey
Thomas M. Berg
James McDonald
Donovan M. Powers

Oregon Department of Geology and Mineral Industries
Margaret Jenks
Vicki McConnell
Mark Neuhaus
Clark Niewendorp
Paul Staub

Pennsylvania Geological Survey
Thomas G. Whitfield

Portland State University, Geology Department
David Percy

South Carolina Geological Survey
Erin E. Hudson

Tennessee Division of Geology
Elaine Foust

U.S. Forest Service
Andrew H. Rorick

University of Alabama
Douglas Behm

University of Texas Bureau of Economic Geology
Tom Tremblay

U.S. Geological Survey
Debra Block
Pamela M. Cossette
Alex Donatich
Christopher Garrity
Bruce Johnson
Jonathan C. Matti
David R. Soller
Nancy R. Stamm
Will Stettner
Evan Thoms
Ronald R. Wahl
Rob Wardwell
Ray Wells
Karen Wheeler

USGS/UCSB
Jordan T. Hastings

Utah Geological Survey
Kent Brown

Washington Division of Geology and Earth Resources
Charles Caruthers
Anne C. Heinitz
Sandi McAuliffe
Karen D. Meyers
Rebecca Niggemann
Jaretta M. Roloff

West Virginia Geological Survey
Jane S. McColloch
Gayle H. (Scott) McColloch

Wisconsin Geological and Natural History Survey
Michael L. Czechanski
Deborah Patterson
Peter Schoephoester

Wyoming State Geological Survey
Joseph Huss
Phyllis Ranz