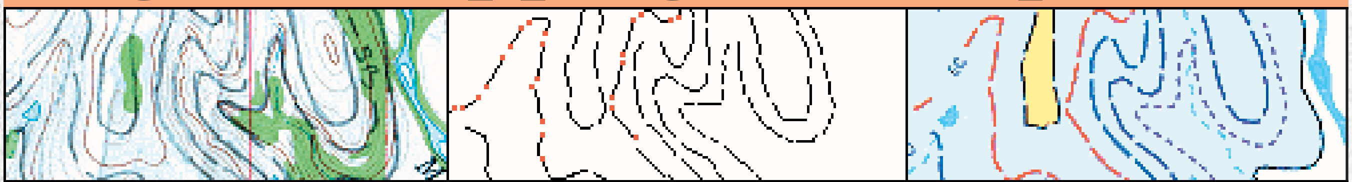
Digital Mapping Techniques ‘02



Association of American State Geologists **United States Geological Survey**

Digital Mapping Techniques ‘02—Workshop Proceedings

May 19-22, 2002
Salt Lake City, Utah

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

Hosted by the
Utah Geological Survey

U.S. Geological Survey Open-File Report 02-370

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

Digital Mapping Techniques '02— Workshop Proceedings

Edited by David R. Soller

May 19-22, 2002
Salt Lake City, Utah

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

Hosted by the
Utah Geological Survey

U.S. GEOLOGICAL SURVEY OPEN-FILE REPORT 02-370
2002

CONTENTS

Poster Presentations

Introduction

By David R. Soller

U.S. Geological Survey
908 National Center
Reston, VA 20192
Telephone: (703) 648-6907
Fax: (703) 648-6937
e-mail: drsoller@usgs.gov

The Digital Mapping Techniques '02 (DMT'02) workshop was attended by 101 technical experts from 43 agencies, universities, and private companies, including representatives from 25 state geological surveys (see Appendix A). This workshop was similar in nature to the previous five meetings, held in Lawrence, Kansas (Soller, 1997), in Champaign, Illinois (Soller, 1998a), in Madison, Wisconsin (Soller, 1999), in Lexington, Kentucky (Soller, 2000), and in Tuscaloosa, Alabama (Soller, 2001). This year's meeting was hosted by the Utah Geological Survey, from May 19 to 22, 2002, on the University of Utah campus in Salt Lake City. As in the previous meetings, the objective was to foster informal discussion and exchange of technical information. When an attendee adopts or modifies a newly learned technique on the basis of discussions at the workshop, the workshop clearly has met that objective. Evidence of learning and cooperation among participating agencies continued to be a highlight of the DMT workshops (see example in Soller, 1998b, and various papers in this volume).

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

At the 2002 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) techniques for scanning already published maps and managing and delivering them on the Web; (3) continued development of the National Geologic Map Database; and (4) progress toward building a standard geologic map data model. In addition, special presentations were provided on building a statewide GIS council, incorporating geology as a NSDI Framework layer, and resolving the roles of surveyors and GIS professionals.

ACKNOWLEDGMENTS

I thank the Utah Geological Survey (UGS) and their Director and State Geologist, Rick Allis, for hosting this meeting. In the tradition of past DMT meetings, it was quite productive and enjoyable. I especially thank Grant Willis and Kent Brown (UGS), who coordinated the meeting, provided excellent support for the attendees, offered an entertaining range of technical and social activities (e.g., a pre-meeting field demonstration of GPS-GIS tools), and managed the meeting's Web site (see Appendix B). Thanks also to Cheryl Ostlund, Jo Lynn Campbell, and Mike Wright (UGS), Jan McEwan, Roni Whittle, and Ryan Halstenrud (Huntsman Cancer Institute), Lesle Wells (The Point Restaurant), and Wes Christianson (University of Utah), for helping with the meeting planning, logistics, audiovisual presentation, vehicles, and meals. For the impressive tour of the Kennecott Mine, I thank Geoff Bedell, Stan Nelson, and Tracy Smith (Kennecott Utah Copper Corporation).

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 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 advice in planning the workshop's content.

I warmly thank Merrianne Hackathorn (Ohio Geological Survey) and Mindy C. James (Wisconsin Geological and Natural History Survey) for their technical editing of each manuscript and their support through the production process, and Lisa Van Doren (Ohio Geological Survey) for electronic layout of the Proceedings. 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 25 oral presentations. Nearly all are supported by a short paper contained in these Proceedings. Some presentations were coordinated with discussion sessions, described below. The papers represent 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."

The papers are generally organized by topic. Information about the software and hardware referred to in these Proceedings is provided in Appendix C.

POSTERS

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

DISCUSSION SESSIONS

To provide the opportunity to consider a topic in some detail, special discussion sessions are held at the DMT workshops. This year there were three sessions: (1) scanning, delivery, and archiving of existing maps, (2) archiving of digital field notes, and (3) information delivery. Session 1 included numerous oral presentations and open discussions. Sessions 2 and 3 were held on the final day of the meeting and produced several ideas and recommendations that will be discussed by the Data Capture Working Group and likely implemented at DMT'03. These sessions highlight an important aspect of the DMT workshop series—it provides a unique venue for sharing technical information and experience for those in the geologic and GIS disciplines.

THE NEXT DMT WORKSHOP

The seventh annual DMT meeting will be held in late spring 2003 in Pennsylvania. Please consult the Web site <http://ncgmp.usgs.gov/ngmdbproject/standards/datacapt/> for updated information. While planning for that event, the Data Capture Working Group will carefully consider the recommendations offered by DMT'02 attendees.

REFERENCES

Soller, D.R., ed., 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., <http://ncgmp.usgs.gov/pubs/of97-269/>.

Soller, D.R., ed., 1998a, Digital Mapping Techniques '98—Workshop Proceedings: U.S. Geological Survey Open-File Report 98-487, 134 p., <http://pubs.usgs.gov/openfile/of98-487/>.

Soller, D.R., 1998b, Introduction, *in* Soller, D.R., ed., Digital Mapping Techniques '98—Workshop Proceedings: U.S. Geological Survey Open-File Report 98-487, p. 1-3, <http://pubs.usgs.gov/openfile/of98-487/intro.html>.

Soller, D.R., ed., 1999, Digital Mapping Techniques '99—Workshop Proceedings: U.S. Geological Survey Open-File Report 99-386, 216 p., <http://pubs.usgs.gov/openfile/of99-386/>.

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

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

The Value of Geologic Maps and the Need for Digitally Vectorized Data

By James C. Cobb

Kentucky Geological Survey
228 Mining and Mineral Resources Bldg.
University of Kentucky
Lexington, KY 40506-0107
Telephone: (859) 257-5500
Fax: (859) 257-5500
e-mail: cobb@kgs.mm.uky.edu

INTRODUCTION

The costs of geologic mapping are challenging to justify. Geologic mapping is supported almost entirely by public funding, but understood by very few people in policy-making positions. The cost is justified for the most part only by anecdotal evidence, such as the discovery of valuable resources, which, when produced, return to society in tax dollars many times the cost of the mapping, or the locating of bridges or landfills so that societal problems are avoided. Public policy-makers want and often demand a cost-benefit analysis to assist in the appropriations process. Only rigorous economic analyses can provide the quantitative data needed to determine a cost-benefit ratio.

What is the value of a 1:24,000-scale geologic quadrangle map? A study based on a survey of professional geologists estimated $43,527! This represents the average amount of money a company would spend to collect the information provided by a published geologic quadrangle map (Bhagwat and Ipe, 2000a, b).

Many geologists and organizations currently involved in geologic mapping are not aware of the Bhagwat and Ipe (2000a) economic analysis. As a result, very compelling data are not being used to justify the cost and purpose of geologic mapping. By citing the enormous economic advantages of having geologic maps available, we can help funding authorities better understand the costs and benefits of geologic mapping. Whenever possible, the value of a geologic map should be cited, as well as the cost-benefit ratio of geologic maps as "public goods."

Few economic studies of the value of geologic maps have been published. Bhagwat and Ipe's (2000a) detailed analysis of the Kentucky geologic mapping program used data from 440 questionnaires and was based on the theory of public goods. Bernknopf and others (1993) also addressed the fundamentals of geologic maps for society, a cost-benefit model for valuing geologic maps, and the economic issues to assess geologic maps as public goods. A study by McGrain (1979) gave examples of the uses of geologic maps, quoted the number of maps sold, and gave anecdotal information about the value of geologic mapping in Kentucky. Cressman and Noger (1981) cited many facts and figures about the 18-year geologic-mapping experience in Kentucky, including history and origin of the program and technical, scientific, and personnel challenges. Although they alluded to economic benefits of the Kentucky geologic mapping program, they did not perform a rigorous economic analysis.

DATA COLLECTION

Subhash Bhagwat and Viju Ipe, economists with the Illinois State Geological Survey, focused their study on Kentucky because it has been completely mapped geologically at a scale of 1:24,000 for 25 years. They worked with staff of the Kentucky Geological Survey to create a questionnaire that would elicit responses suitable for quantitative analysis of the value of geologic maps. Because the questionnaires were being sent to registered professional geologists, other information about the uses of the maps and preferences of the users was requested. The questionnaire contained 14 questions that asked for information such as:

- What are the maps worth to the user?
- What are the maps worth to the state?
- How are the maps used?

The questionnaire was sent to 2,200 geologists registered in Kentucky. This pool of 2,200 included actual and potential users of geologic maps. The response rate of 20 percent (440 responses) provided a representative sample of the user population.

Data were extracted from the questionnaires into a database and for qualitative and quantitative analyses. A complete description of the data and the analytical methods is given in Bhagwat and Ipe (2000a).

Value of a Geologic Quadrangle Map

Bhagwat and Ipe (2000a, b) determined the value of a 1:24,000-scale (7.5-minute) geologic quadrangle map to be $43,527. The respondents to the questionnaire said they saved this amount, on the average, because the maps were already available and therefore they did not have to collect the data themselves. Similarly, respondents reported that to collect only a minimum amount of information for a credible job would have cost an average of $27,776. Because the questionnaires were directed to Kentucky registered professional geologists and were concerned with the Kentucky geologic mapping program, these value are specific to Kentucky GQ's, but could certainly be extrapolated to GQ's in other states, especially those with similar geology.

Value of Geologic Maps to Kentucky

Bhagwat and Ipe (2000a, b) calculated maximum and minimum values for the statewide mapping program by multiplying the number of maps sold (81,000) by the maximum ($43,527) and minimum ($27,776) values per map. This results in a maximum value to the state of Kentucky of $3.53 billion and a minimum value of $2.25 billion. These values are 25 to 39 times the cost of the mapping, giving a tremendous return in value to the State for the cost. A cost of $90 million (1999 dollars) resulting in a benefit of up to $3.35 billion is a remarkable return on the taxpayers' investment!

The public has been extremely well served by the mapping program, as demonstrated by this cost-benefit analysis. Even if you have never purchased a geologic quadrangle map, you still benefit from the availability of the maps to society. This is because economists consider GQ's "public goods," much the same as roads, dams, and reservoirs are—in fact, GQ's make it possible to build better roads, dams, and reservoirs, and build them more economically. And the public will continue to reap the benefits of the maps, because the information they contain will continue to be used for many more decades.

Reported Uses of Maps

The responses indicated a wide variety of uses for the maps, some of which could not have been anticipated at the time the mapping program began (Table 1). Some of the most common uses were:

- Exploring for and developing ground-water resources
- Cleaning up environmentally damaged sites
- Avoiding karst hazards
- Designing foundations and engineering
- Making zoning and city planning decisions
- Locating waste-disposal facilities
- Evaluating property

The user responses indicated that GQ maps are used in nearly all sectors of the economy to ensure environmental safety, to prevent hazards to manmade structures, and to delineate and develop natural resources such as ground water, minerals, and fuels. The use of GQ maps improves the quality and credibility of work and saves money. Most important, geologic mapping generates

Table 1. How people use Kentucky geologic maps (from Bhagwat and Ipe, 2000a).

Category	Map use	Percent of respondents
Exploration and development	Coal	30
	Oil and gas	32
	Industrial minerals	32
	Ground water	73
Environmental consulting	Pollution prevention	53
	Industrial applications	41
	Site clean-up	68
Hazard prevention and protection	Landslides	33
	Earthquakes	14
	Karst problems	54
	Subsidence	40
Engineering	Buildings and foundations	37
	Roads and highways	35
	Railroads	16
	Pipelines	30
	Utilities	26
	Dams, dikes, and locks	27
City planning	Zoning decisions	18
	Landscape planning	11
	Building codes	8
Regional planning	Waste disposal	45
	Transportation	16
	Industrial permits	38
Property valuation	Property tax assessment	11
	Land acquisition	35

knowledge—a public good vital to the economy, public safety, and public health. This knowledge would not be produced if left to private enterprise and has not been produced by private enterprise elsewhere, except on a site-specific basis or under contract to a public agency. In such cases, the resulting maps remain in proprietary or private hands and are not available to the public. Map users indicated the desirability of maps showing lithology, structural features, formation contacts, and cultural features. The most desired map scale was 1:24,000.

Intangible Benefits

A section of the cost-benefit study was devoted to intangible benefits derived by map users. They include such vital benefits as increased credibility in reports and studies prepared by map users, time saved in project completions, and the value of unbiased information in maps that were prepared by scientists without a vested interest. These kinds of intangible benefits often outweigh the monetary value of public goods. Such benefits are especially important in the case of public goods that create and deliver scientific knowledge, in contrast to public goods that provide physical facilities of economic or recreational value, such as parks, roads, or bridges.

KENTUCKY EXPERIENCE

In the 160-year history of the Kentucky Geological Survey, its most valuable accomplishment has been the geologic mapping of the state at a scale of 1:24,000. The 707 geologic quadrangle maps are the Survey's greatest assets and were the result of a 20-year cooperative program with the U.S. Geological Survey. The mission of the Kentucky Geological Survey has been and continues to be investigating the geology and minerals of the Commonwealth for the benefit of its citizens. The geologic mapping program, which started in 1960 and finished in 1978, not only advanced this mission, but also contributed to all future work by the Survey and other agencies involved in mineral resources, water, geologic hazards, environment, construction, and land-use planning. The Kentucky Geological Survey owes a great debt of gratitude to the U.S. Geological Survey for its cooperation while the program was underway. The successful completion of this program is a tremendous testimonial to the planning, foresight, geologic and administrative effort, and cooperation of these two organizations. Demand for the maps has been strong. More than 5,000 geologic quadrangle maps are sold to the public each year, and the initial printing of a number of the GQ's has sold out completely. A number of remarkable benefits from these maps are not readily apparent.

1. The cost of the mapping program in Kentucky was justified, if only for the economic development of oil, natural gas, coal, and minerals. The economic development of these natural resources was made possible by the valuable information these maps contain. Kentucky's mineral economy rose dramatically in the 1970's and 1980's, especially in the areas of industrial minerals, coal, oil, and gas. What could not have been anticipated in 1960 when this mapping program began was that the use of these maps 30 years later for the management of land, water, and the environment would surpass their use for mineral development. In a society where landowners are responsible for their land and water, making information readily available for the prudent use of those resources is not only important, it is essential.
2. The geologic maps provide knowledge about the land and geology for a broad cross section of users in society (e.g., researchers, engineers, miners, urban planners, and hikers). In fact, there are so many diverse users that listing them all is almost impossible.
3. If a picture is worth a thousand words, then certainly a geologic map is worth a million words. For the tens of thousands of requests that the Kentucky Geological Survey receives each year from the public about land, water, minerals, and hazards, the geologic maps are sufficient to respond to a great number of them. For requests for which more detailed information is needed, the geologic maps provide a context or base of understanding for more detailed data and analysis.
4. Following the completion of the geologic mapping program in 1979, a new state geologic map published at a scale of 1:250,000 became a popular map for statewide analysis and study. A geologic map published in 1988 at a scale of 1:500,000 also became popular for regional resource assessments. The publication of both of these maps was made possible by the existence of the original detailed geologic maps at a scale of 1:24,000.
5. Currently, the 707 geologic quadrangle maps are being converted into digital format for use in a wide variety of computer applications. Scanned versions of the maps are available for the public to view and print from the Kentucky Geological Survey Web site, and digitally vectorized geologic quadrangle data are made available on CD-ROM. Vectorized and attributed geologic quadrangle maps will be available in the near future for use in geographic information systems. This will make detailed geologic information available for every office and home over the World Wide Web, 24 hours a day, 7 days a week, for use with decisions requiring geologic information.

What started 40 years ago as a program to spur the economic development of the mineral and fuel industries of Kentucky has proven to be enormously valuable in many other ways in both the public and private sector. The forethought of the Tenth Kentucky Geological Survey to commit itself to that challenge and to complete the geologic mapping program is a legacy whose value should never be underestimated. I cannot describe in stronger terms what a valuable resource the geologic maps have been and continue to be for the Kentucky Geological Survey and the State.

DIGITAL GEOLOGIC MAP PRODUCTS

Many states are digitizing old and new geologic maps so that they can be used in computer programs and plotted on demand or transferred electronically as needs arise. Bhagwat and Ipe's (2000a) research indicated broad support for the creation of digital products. An overwhelming 82 percent of respondents agreed that digital geologic maps are valuable to them. The user community also needs help learning how to use the associated computer programs. To the extent possible, every state engaged in the production of digital geologic products (including maps) should help the user community learn applications that manipulate and use the data.

The Kentucky Geological Survey is publishing digitally vectorized geologic quadrangle data (DVGQ's). The DVGQ's are the line, point, and attribute data from a 1:24,000-scale geologic quadgrangle map. The philosophy driving the DVGQ's is that potential users will fall into two categories: they will either need a paper map, or they will need specific parts of the map in digital format to be cut and pasted into their own work. A user who wants a paper map can request one, and a user who needs the computer data can get them from the DVGQ. The DVGQ gives the user all the files from which to select a specific element from the map. This eliminates the need for the user community to vectorize geologic map data and risk introducing errors. Still, making the user community aware of the DVGQ's and teaching them how to use the data is a big job. For this reason, the Kentucky Geological Survey, and other geologic groups and societies, host workshops and short courses on the use of DVGQ's.

CONCLUSION

A total of 46 states had geologic mapping projects under the STATEMAP section of the National Cooperative Geologic Mapping Program in 2001. The total funding for this national program in 2001 was $6.7 million, which is roughly equivalent to what was spent per year, adjusted for inflation, for geologic mapping in Kentucky in the 1960's and 1970's. Clearly, work still needs to be done to increase funding for a program as vital as STATEMAP. The outstanding benefits of this program, when compared to the cost, make it an important public investment.

REFERENCES

Bernknopf, R.L., Brookshire, D.S., Soller, D.R., McKee, M.J., Sutter, J.F., Matti, J.C., and Campbell, R.H., 1993, Societal value of geologic maps: U.S. Geological Survey Circular 1111, 53 p.

Bhagwat, S.B., and Ipe, V.C., 2000a, The economic benefits of detailed geologic mapping to Kentucky: Illinois State Geological Survey Special Report 3, 39 p.

Bhagwat, S.B., and Ipe, V.C., 2000b, What are geologic maps worth: Geotimes, December 2000, p. 36-37.

Cressman, E.R., and Noger, M.C., 1981, Geologic mapping of Kentucky—a history and evaluation of the Kentucky Geological Survey–U.S. Geological Survey Mapping Program, 1960-1978: U.S. Geological Survey Circular 801, 22 p.

McGrain, P., 1979, An economic evaluation of the Kentucky geologic mapping program: Kentucky Geological Survey, ser. 11, 12 p.

Compilation of a 1:24,000-Scale Geologic Map Database, Phoenix Metropolitan Area

By Stephen M. Richard and Tim R. Orr

Arizona Geological Survey
416 W. Congress #100
Tucson, AZ 85701
Telephone: (520) 770-3500
Fax: (520) 770-3305
e-mail: {Steve.Richard,Tim.Orr}@azgs.az.gov

The Arizona Geological Survey (AZGS) has embarked on a program to compile and release digital geologic map data for the Phoenix metropolitan area (Figure 1). This project is motivated by the fact that the Arizona geologic mapping advisory committee voted overwhelmingly in favor of detailed (1:24,000-scale) geologic data production as opposed to 1:100,000-scale compilation. More detailed geologic information is required for engineering and environmental applications in the rapidly urbanizing metropolitan area. The goal of the 1:24,000-scale compilation is to provide a geologic map database that is as accurate and up-to-date as possible, based on a consistent set of geologic map units across the entire compilation region.

In conjunction with existing 1:100,000- and 1:1,000,000-scale digital geologic data, our goal is to make geologic map data available at several scales through a single ESRI ArcView 3.2 project. Initially, the thematic geology data included in the database are very simple, with geologic unit information designated using a single geologic-unit identifier associated with each polygon. Lines are classified in a simple scheme recognizing various kinds of faults and nonfaulted contacts between units. Users must refer to published source maps for a

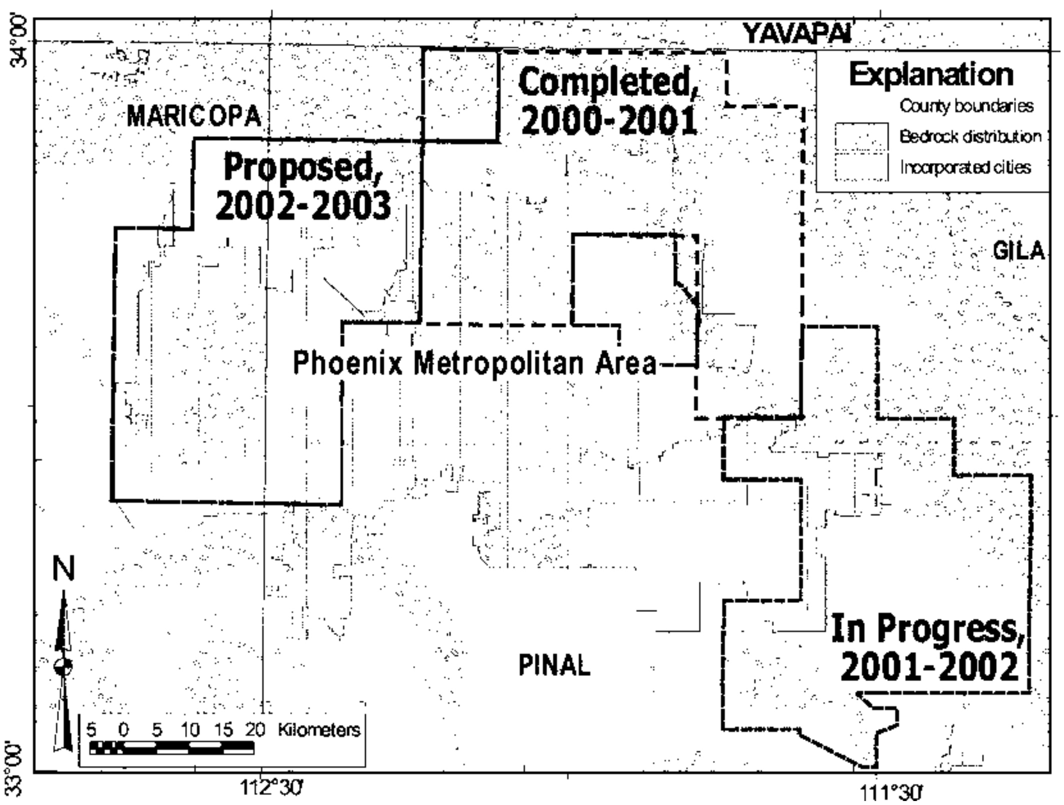


Figure 1. Compilation plan, Phoenix metropolitan area.

more in-depth description of the geologic features. As standardized data models and software tools evolve (see Boisvert and others, Progress Report: North American Geologic Map Data Model Design Team, this volume), we plan to progressively expand the thematic geologic content of the database to allow more sophisticated applications that tailor the geologic map presentation to the needs of the user.

COMPILATION MECHANICS

Geologic map source data for much of the compilation area are available in original mechanical form (typically ink drafted on mylar) at the AZGS, because much of the mapping has been completed over the last 15 years by Survey geologists under the auspices of the U.S. Geological Survey–Association of American State Geologists National Cooperative Geologic Mapping Program. Digital map production at the AZGS has been implemented only in the last three years, after completion of mapping in the Phoenix area. Depending on the nature of the material available, source maps are either scanned and digitized on the computer screen ("heads up"), or digitized using a large digitizing table. Heads-up digitizing is done in operator-assisted mode using Able Software's R2V program if mylar separates containing only the linework are available. The output from R2V is converted to shape files, imported into Arc/Info coverages, and final editing, cleaning, and building of polygon topology is done using Arc and ArcEdit. If the original mechanical material consists of drafted linework on a screened topographic base map, the scans are digitized heads up using ArcEdit because the benefit of the autovectorization functionality of R2V is lost. Table digitizing is completed using ArcEdit to directly construct an Arc/Info coverage with topology. Data fields for the AZGS data structure (Richard and Orr, 2001) are added to the PAT and AAT tables in Arc. Line and polygon attributes for topologically complete coverages are then assigned using ArcView 3.2. Point structure data are digitized using a custom Arc Macro Language interface that automatically calculates the strike indicated by a structure symbol when the digitizer locates the ends of the symbol's strike line. Metadata associated with each point, line, and polygon, referred to as the feature-level tracking, record the provenance of all data.

The boundary of each source map is preserved as a polygon-bounding line, allowing the polygons derived from each source map to be shown with the source map bounding line. When the compilation is symbolized as one map, these source-map boundary lines are not symbolized.

Supporting metadata and thematic geologic data are stored in a Microsoft Access database. The structure of this supporting database system is described in Richard and Orr (2001). The key features of the database implementation are separation of the user interface ("front end") from the data tables ("back end"). The back-end data are stored in several thematic infrastructure databases, including metadata infrastructure, geoscience infrastructure, and cartography infrastructure. These infrastructure databases contain data elements used by all AZGS geologic map databases, including metadata such as person, organization, and project identification, feature-level metadata for infrastructure data, basic geoscience concept terminology for describing geologic occurrences, and description of standard symbology used for AZGS maps. Geologic data and metadata specific to the Phoenix area database are stored in a single back-end database specific to this project. The front-end database consists of a collection of forms, queries, and Visual Basic code for the user interface. Back-end tables are linked to the front-end database. This allows evolution of the user interface independent of the data.

COMPILATION ISSUES

Overlapping maps

In some parts of the compilation area, more that one source map is available. Following the standard procedure for map compilation, the compilers reconcile the various sources. This was simplified in many cases because the original geologists who created the source maps were not available to participate in the process. If, as part of this reconciliation process, the line work has been modified from the original source maps, the feature-level tracking for the edited lines notes the modifications made.

Reconciliation of Boundary Discrepancies

At the boundaries between source maps, geologic lines (contacts and faults) that clearly represent the same feature commonly do not precisely meet. In these cases the line locations have been adjusted by the compiler, and the location uncertainty of the lines (an attribute of each line) is updated to reflect the location discrepancy between the maps. Such discrepancies provide an operational indication of the actual precision of the source maps. A less common boundary problem is a contact that trends subparallel to the map boundary and, although in places it clearly must cross over into the adjacent map sheet, it is not mapped there. This requires the addition on the adjacent map of a geologic contact to close the polygon, which introduces a "sliver" polygon bounded by a geologic line on one side and a source-map boundary line on the other.

Inconsistent Map Units

At the boundaries of some of the source maps, the unit designations between the two maps is so incompatible that no reconciliation is possible. In these cases, the source-map boundary serves as a geologic polygon boundary. Table 1 provides examples of some of the incompatible geologic units for adjacent polygons across source-map boundaries.

Table 1. Comparison of unit name incompatibilities across map boundary.

Map 1	Map 2
Andesite	Chalk Canyon Formation
Andesite	Subaqueous lava complex
Argillite, siltite, and argillaceous sandstone	Metasedimentary rocks
Chert	Metasedimentary rocks
Chert	Felsic metavolcanic rocks
Felsic metavolcanic rocks	Green argillite
Green argillite	Metasedimentary rocks
Mafic (basalt-ultramafic)	Metavolcanic rocks

New Unpublished Mapping

In several parts of the compilation area, AZGS geologists have completed new geologic mapping that is as yet unpublished. Because our goal is to provide the most complete, accurate, and up-to-date information, we have used the new mapping where available, and identified the source using the feature-level tracking record.

Multiple Map Views

We have used this compilation of the Phoenix area as a pilot project to develop methods to record multiple geologic-map visualizations based on the same database. Specifically, we wished to record, as nearly as possible, the original source geologic map. Most of the source maps consist only of black-and-white lines, with no color fill. All the maps used a compatible collection of fault and nonfaulted geologic-contact classifications, so only a single collection of line-type classification and symbolization is necessary. On the other hand, each source map includes a collection of geologic units, some of which are identical with those on adjacent maps, some of which may be correlated with those on adjacent maps, and some of which are unique. Thus, two visualizations are recorded in the database, one based on the legend from the original source map and one based on the compilation legend for the whole extent of the database. Both of these use the same symbolization for point structure data and geologic linework.

Map visualizations from the database are specified in the MapViewDefinition table (Figure 2). Each record in this table defines a map view through links to correlation tables that associate spatial objects with classification concepts and symbols (see Richard and Orr, 2001). The ViewSchemeTypeID defines which of several methods are used to select the spatial features to display and assigns symbols to the displayed features. The MapLegendID is used as an identifier to filter the relationship table to select links between the symbols used in the map view and classification concepts that define their meaning. ClassSchemeID identifies links between spatial objects and classification concepts that specify the geologic thing represented by the spatial object. SymbolSchemeID identifies links between spatial objects and symbols that directly specify how a particular feature is to be symbolized. The CatalogLinksDS identifies a data set that contains links of type "MapView components," linking the MapViewID with data set identifiers for all data sets required to construct the MapView. This structure provides a very general mechanism for representing the variety of approaches to composing a map visualization from the database.

The AZGS database design includes a field to assign a default cartographic object to each spatial object in the database (Figure 2). Only one symbolization scheme is used for point structure data and geologic linework. For the northeast Phoenix metropolitan area database, we have assigned default symbols for geologic polygons based on the compilation map legend. The database includes a GeologicUnit table (a kind of ClassificationConcept table represented in Figure 2) that contains definitions of the geologic units from all the original maps, including the full text description of the unit. That table also includes definitions of all the compilation map units. The classification of geologic polygons on the original source maps is recorded in a classification scheme correlation (implemented in the AttributedRelationship table) that associates each polygon with its appropriate geologic unit. This correlation serves to select for display in the map view the polygons that compose the source map. A MapLegend correlation for each source map assigns CartographicObjects (symbols) for each geologic unit, based on the unit assignment for the polygon in the classification scheme for that map view. The points and lines to display in the view are selected by a Map View-Spatial Object correlation, implemented by instances in the AttributedRelationship table.

A geologic map includes many annotative features that convey the meaning of graphical elements on the map. The location of these annotative features is not determined by the location of a corresponding occurrence on the Earth, but rather by the requirements of esthetics and cartographic clarity. These include map-unit labels, name labels for named features, and a variety of symbols associated with lines that convey classification of the line (e.g., anticline symbol along fold hinge surface trace, teeth on thrust-fault trace). The location of such

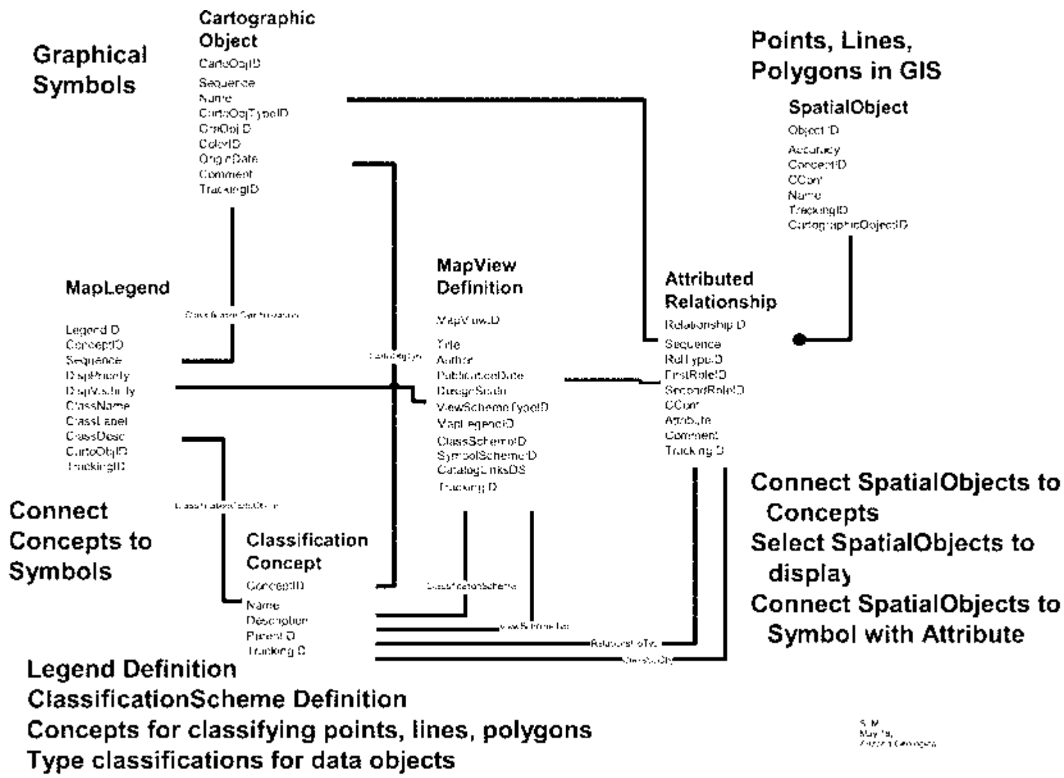


Figure 2. Map visualization represented by general relationship links between cartographic, classification, and spatial objects. Cartographic objects represent graphical elements (symbols) used to depict geologic features. Classification objects represent kinds of things (geologic units, faults, contacts, structure) and terminology used to describe geologic things. Spatial objects represent descriptions of geometry (location).

annotative features is recorded in the database in separate spatial data sets. Spatial objects in these data sets carry a CartographicObject identifier that specifies the graphical element located by that object. The northeast Phoenix area database currently does not include specification of these cartographic features for the map views that are represented.

INFORMATION DISTRIBUTION

Geologic map database products from the AZGS are distributed on CD-ROM, with an image that provides a rudimentary view of the map. The database publications are found in the Digital Information publication series, and each of these includes a version number. At present, the version numbers are in the form "m.n." The "n" number is incremented when data are added or updated. The "m" number is incremented when the data structure is modified by adding tables or redefining tables.

Some databases have one or more associated, fully decorated, digital map layouts as separate publications. These are titled "Geologic map of . . . ," and are included in the Digital Geologic Map publication series. The maps are subject to revision, indicated by including the word "revised" as a suffix on the publication date of any revised version of the map. The publication date on the map indicates the date of the most recent revision. The text associated with the map includes the revision history. The Digital Geologic Map series is treated differently from the traditional Geologic Map series because the authoritative, archival version for Digital Geologic Map Series maps is an Adobe Acrobat (.pdf) file on an archival CD-ROM, as opposed to a paper copy of the printed map in the archive.

IMPLEMENTATION PLAN

The first digital geologic map databases were released by the AZGS in 1997, using a data structure described in Richard and Thieme (1997). In that structure, metadata were presented as a separate data set, following the Federal Geographic Data Committee format. The thematic geologic information consisted simply of identification of geologic concepts for lines (contacts, faults of various sorts) and the map units represented by polygons. Development of the database design has continued, in conjunction with the evolution of ideas for a standard geologic map data model (Johnson and others, 1998; Brodaric and Hastings, 2000; Data Model Design Team, this volume).

We now have implemented a robust feature-level metadata tracking scheme, while developing a logical structure to allow more in-depth representation of geologic knowledge (Richard and Orr, 2001). The framework for representing map visualization (cartography) has been developed and implemented, but we are only now beginning to populate databases to record this information. Because of the rapidly evolving nature of the standard framework for database representation of lithology, map units, geologic structure, and geologic history, these aspects have not been implemented for publication databases.

Current database development and implementation at the AZGS has been focused on the inclusion of field data produced in conjunction with mapping projects to archive a more complete record of the data collected in the field. Priorities for implementation in the near future include (1) implementation of the lithologic description, which will require development of standard terminology for many aspects of lithologic description (see <http://geology.usgs.gov/dm/steering/teams/language/charter.shtml>), as well as tools to allow data entry for structured descriptions of the complex relationships inherent in lithologic description; (2) integration of stratigraphic lexicon and geochronology into the information system as separate, but linked databases, and (3) development of tools for more flexible and user-friendly cartographic design by geologists. These tools are needed to accelerate development of publication-quality layouts, and for data exploration and discovery—it's time to start actually using the information that's going into the databases!

REFERENCES

Brodaric, B., and Hastings, J., 2000, Evolution of an object-oriented, NADM-based data model prototype for the USGS National Geologic Map Database Project (abs.): Annual Conference of the International Association for Mathematical Geology, IAMG2001, Cancun, Mexico.

Johnson, B.R., Brodaric, B., and Raines, G.L., Hastings, J.T., and Wahl, R., 1999, Digital Geologic Maps Data Model, v. 4.3: U.S. Geological Survey and Association of American State Geologists, <http://geology.usgs.gov/dm/model/Model43a.pdf>.

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, <http://pubs.usgs.gov/of/2001/of01-223/richard2.html>.

Richard, S.M., and Thieme, J.P., 1997, Data structure for Arizona Geological Survey digital geologic maps: Arizona Geological Survey Open-File Report 97-5, 15 p.

Distributed Spatial Databases—The MIDCARB Carbon Sequestration Project

By Gerald A. Weisenfluh,[1] Nathan K. Eaton,[2] and Ken Nelson[3]

[1]Kentucky Geological Survey
228 Mining and Mineral Resources Bldg.
University of Kentucky
Lexington, KY 40506
Telephone: (859) 257-5500
Fax: (859) 257-1147
e-mail: jerryw@kgs.mm.uky.edu

[2]Indiana Geological Survey
611 North Walnut Grove
Bloomington, IN 47405
Telephone: (812) 855-7636
Fax: (812) 855-2862
e-mail: neaton@indiana.edu

[3]Kansas Geological Survey
1930 Constant Ave.
Lawrence, KS 66047
Telephone: (785) 864-3965
Fax: (785) 864-5317
e-mail: nelson@kgs.ku.edu

INTRODUCTION

The MIDCARB Project

The state geological surveys of Illinois, Indiana, Kansas, Kentucky, and Ohio have formed a consortium to investigate the potential for sequestering carbon dioxide from significant emitters within their regions. The multi-year Midcontinent Interactive Digital Carbon Atlas and Relational dataBase (MIDCARB) project is funded by the U.S. Department of Energy's National Energy Technology Laboratory. The goals of MIDCARB are to (1) develop and organize scientific information related to CO_2 sources (primarily electricity-generating facilities) and potential sequestration sites in the five-state area; (2) develop the information technology needed to access, query, analyze, display, and disseminate natural-resource data related to carbon management; and (3) make this information accessible to users of the World Wide Web. Each of the participating states has petroleum and coal-fired electricity-generating facilities that produce CO_2 as well as potential for sequestration in petroleum reservoirs, unmineable coal beds, and saline aquifers.

Project Strategy

Each state survey has well established relational and spatial databases that can be used to evaluate potential sequestration sites. Each of these databases has different data elements and design characteristics, however, and are continuously updated by the respective state organizations. It was clear from the onset of the MIDCARB project that compiling information from each state into a centralized database would present serious maintenance issues, especially because each database is frequently updated. The project strategy was, therefore, to use the existing distributed database structure for tabular data, to create a similar distributed system for spatial information, and to develop software tools to integrate and interact with the data in a Web environment. Figure 1 shows the

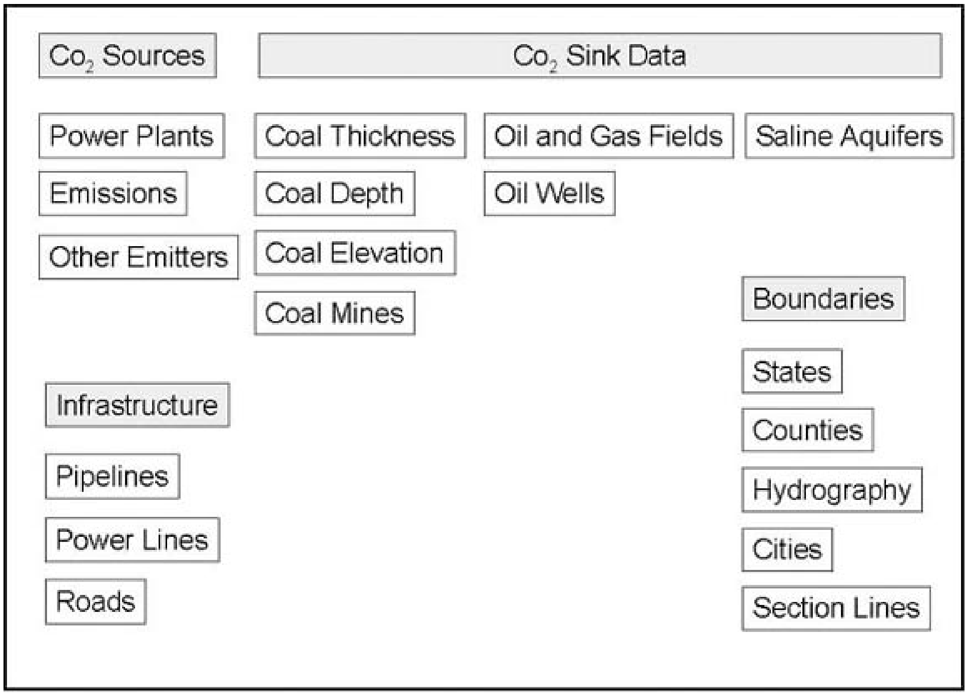


Figure 1. Categories and types of geographic databases used for the MIDCARB ArcIMS project.

spectrum of individual state geographic databases that are being used in the Arc Internet Map Server (ArcIMS) project.

Tabular databases from the participating organizations currently use either Oracle or SQLServer relational database management systems (RDBMS). Environmental Systems Research Institute's (ESRI) Spatial Database Engine (ArcSDE) software was chosen so that the spatial data sets could also be managed in an RDBMS. ArcSDE software is installed on each state survey's RDBMS and acts as an interface between GIS software and the underlying relational database (e.g., Oracle). SDE allows GIS data to be managed by a traditional enterprise-scale relational database and manages requests for information from a variety of ESRI applications, including ArcMap, ArcCatalog, and ArcIMS. Staff of the project elected to use this software solution because each of the organizations was using some or all of the ESRI products and because of ESRI's diverse off-the-shelf application support. ArcIMS was chosen as the platform for integrating all the distributed information in a single Web application. Figure 2 shows the logical architecture for the MIDCARB data integration system.

ARCSDE DATABASE DESIGN

The Kentucky Spatial Database

The MIDCARB project coincided with efforts at the Kentucky Geological Survey (KYGS) to construct its own ArcSDE database. Several database design considerations evolved during this development process that relate to data storage and user access.

The KYGS maintains numerous large spatial data sets that cover a diverse spectrum of natural resource themes. These include general-purpose GIS data, such as base maps and geologic maps, as well as specialized themes such as the carbon-sequestration layers developed specifically for the MIDCARB project. The KYGS decided to maintain all its spatial data in a single ArcSDE database, using sub-tables for organizing the data thematically, rather than creating separate ArcSDE databases (Figure 3). This approach alleviated the necessity for users to make the many database connections required by a multiple database scenario. Data layers were prepared in a single coordinate system and datum (NAD83, decimal degrees) to simplify data integration. All data were added to ArcSDE as simple, unregistered feature classes—no complex geodatabase functionality, such as feature editing, custom object behaviors, or versioning, has yet been enabled.

The KYGS point databases (oil, gas, coal, and water well/sample locations) that are maintained in a relational database as tabular datasets were spatially enabled by adding their location and basic descriptive information to ArcSDE feature classes. This greatly simplifies the task of using these data in mapping applications (e.g., eliminating the need to add data files to ArcView for event theme creation). Attribute data for well locations are still managed by the relational database system; however, updating location information is more problematic. At the present

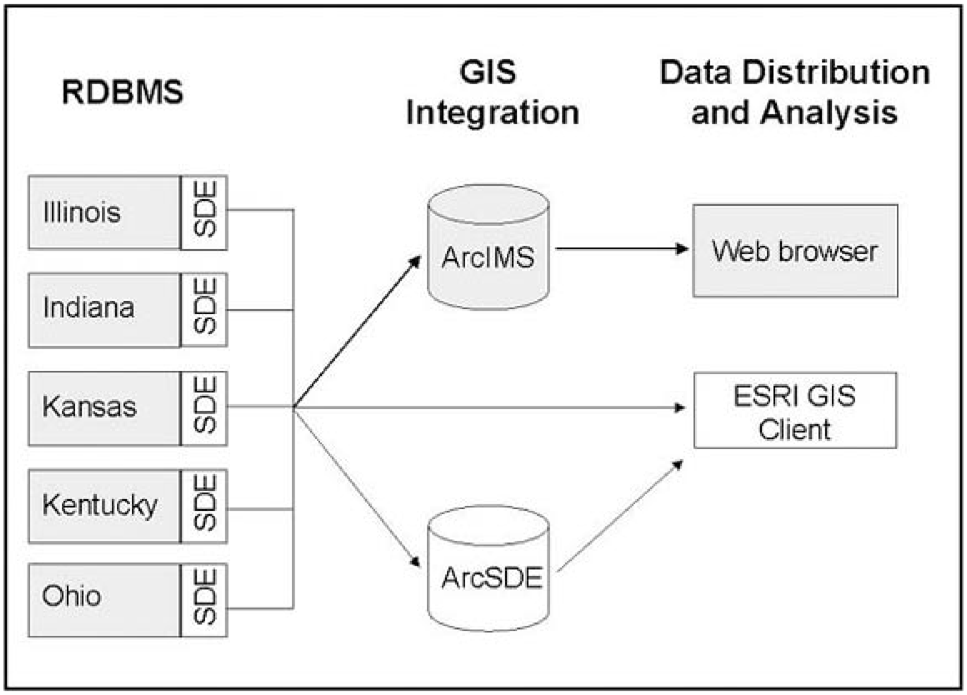


Figure 2. Architecture of the MIDCARB distributed database system. Gray objects show existing configuration that uses a Web browser to interact with the data through ArcIMS. White objects show alternative data pathways where ESRI clients (e.g., ArcView or ArcExplorer) could access spatial data directly from an SDE database.

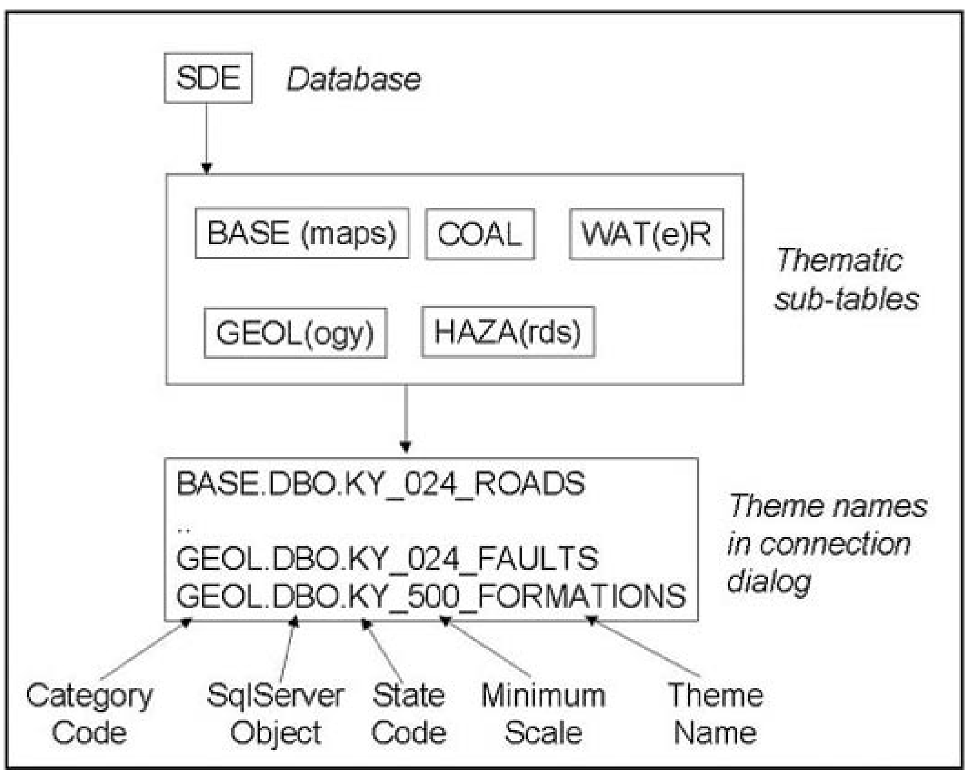


Figure 3. Configuration of the Kentucky ArcSDE database, showing feature-class naming scheme. Thematic tables represent the sub-databases within ArcSDE. Theme names are formatted to easily identify feature classes and their characteristics within an application's database connection dialog (e.g., ArcView).

time, there is no convenient procedure to update ArcSDE locations that are maintained in a relational table. The best solution would probably be to manage location information in ArcSDE and create a method for updating the coordinate attributes stored in the relational database.

Using a single ArcSDE database presents logistical challenges for user access. The large number of ArcSDE themes listed in "Add Theme" dialogs (e.g., in ArcView) can confuse the user. One solution to this problem is to manage groups of feature classes with permissions. For example, all coal (and some related) themes could be made accessible only to a default "coal" user. Such users would not see other unrelated feature classes. KYGS implemented a table-naming scheme as an alternative solution to this problem. Thematic groupings are added to ArcSDE tables named with a four-character code (e.g., COAL, GEOL). Feature classes within the groups are named with a two-letter state prefix, followed by a three-character scale integer, and then a meaningful theme name (Figure 3). This format results in a connection list that is sorted first by theme type, and then by data scale and name, with each designation nearly in vertical alignment. Most relational databases limit such table names to approximately 32 characters.

Another challenge for managing ArcSDE data relates to the potentially large size of statewide databases. For example, the KYGS 1:24,000-scale geologic map database, when complete, will comprise 707 detailed vector data sets that have been edge matched and joined into statewide feature themes. Although ArcSDE does support spatial queries using tiling methods, they will not be effective with such databases because many of the merged features can cover as much as 30 percent of the state. ArcSDE's tile methods return all features that touch the tiles within the current view extent. The KYGS staff decided to pre-intersect these complex feature classes with commonly queried geographic extents (i.e., county and quadrangle outlines). This not only facilitates faster queries, but simplifies the process of preparing finished map layouts by eliminating the need for clipping for common map extents.

Future ArcSDE Work

Constructing maps of custom areas using ArcSDE query methods is effective, but if many themes are involved, the query needs to be issued separately for each theme. This can be a tedious process. To simplify the ArcSDE query process, a tool could be constructed for each ESRI application (e.g., ArcMap, ArcView) that collects the query criteria from the user, then iterates through selected themes.

Large, seamless databases, such as the Kentucky 7.5-minute geologic map formations, present challenges for feature symbolization because the number of distinct map units is very large. Moreover, each of the ESRI applications uses a different method for storing symbol styles. Such maps should be rendered in a standard way, irrespective of the application, and map legends should be constructed so that they contain only styles for features in the current selection. This calls for a database solution that stores symbol definitions in a generic format (e.g., RGB or CMYK), with functions that obtain the required symbols for a selected feature set and construct a custom legend.

Finally, most of the current MIDCARB feature themes are relatively simple—point locations and simple geographic outlines. Serving complex and large spatial databases from distributed locations will require extensive testing for efficiency and development of methods for filtering the data that are returned to the user.

DATA INTEGRATION

The MIDCARB Web Site

The MIDCARB databases are integrated using a single ArcIMS service at the Kansas Geological Survey. This map service is accessed through the MIDCARB Web site, <http://www.MIDCARB.org>. Site development is based on a standard HTML template customized with additional HTML and JavaScript code. Spatial data on the Web page are integrated in the ArcIMS AXL file. Connections to each remote ArcSDE database are made with a WORKSPACE reference (example 1 below) and attachment to a feature theme is specified with the DATASET reference (example 2 below).

Example 1. Example workspace reference in ArcIMS AXL file used to specify a connection to a remote data server.

```
<WORKSPACES>
<SDEWORKSPACE name="sde_ws-48"
  server="kgsdata" instance="port:5151" database=""
  user="jerryw_sde" encrypted="true" password="PKOT
  JKSWGTKNGMKR" geoindexdir="c:\tmp\" />
</WORKSPACES>
```

Example 2. Example dataset reference in ArcIMS AXL file used to specify specific feature layer from a remote database connection.

```
<DATASET name="SEQUESTER.DBO.IB_500_
  SPRINGFIELD_COAL_OVR" type="polygon"
  workspace="sde_ws-48" />
```

It is transparent to the user, both from a design and efficiency perspective, that map layers are being loaded from more than one location. All data layers relevant to the five-state area are viewable from a single ArcIMS page. Figure 4 shows an example of a custom map view that integrates Illinois and Indiana oil and gas fields. Users control the themes to be displayed by activating layers in the ArcIMS Web page table of contents. The amount of data returned to the Web page from the various servers also can be limited by use of the zoom function, which uses ArcSDE's database tiling capabilities. All themes can be queried for attribute information using tools provided in the standard template. Because of the large number of themes provided in the MIDCARB map service, the standard table of contents required customization to clarify and simplify the user legend.

Table of Contents Customization

The first solution for legend simplification was to group themes by subject categories. This was accomplished by adding subject headings at the top of the table of contents with hyperlinks to appropriate parts of the legend (Figure 5A). Clicking on a subject hyperlinks to the list of themes related to that subject. To further simplify the user interface, the legend view (symbolization of feature types) was combined with the table of contents (list of features). These interface elements are typically shown separately in the standard ArcIMS template. Combining these functions saves screen space that can be used for the map view. Legends are displayed only for active themes selected by the user. The legends for this map service were preformatted as GIF images and are inserted dynamically when a theme is activated. This provides custom control over the appearance of the legend that is more readable than dynamic legends generated by the ArcIMS application (Figure 5B).

Tabular Data Integration

In addition to integrating spatial data, the MIDCARB site has linked tabular databases related to the map themes from each of the state's repositories. The standard ArcIMS template displays feature attributes (using the identity tool) in a horizontal frame at the base of the map. If the number of attributes is large, or field sizes are long, this frame must be scrolled to view the data. For the MIDCARB site, customized reports were prepared to more clearly summarize attribute information. Macromedia's

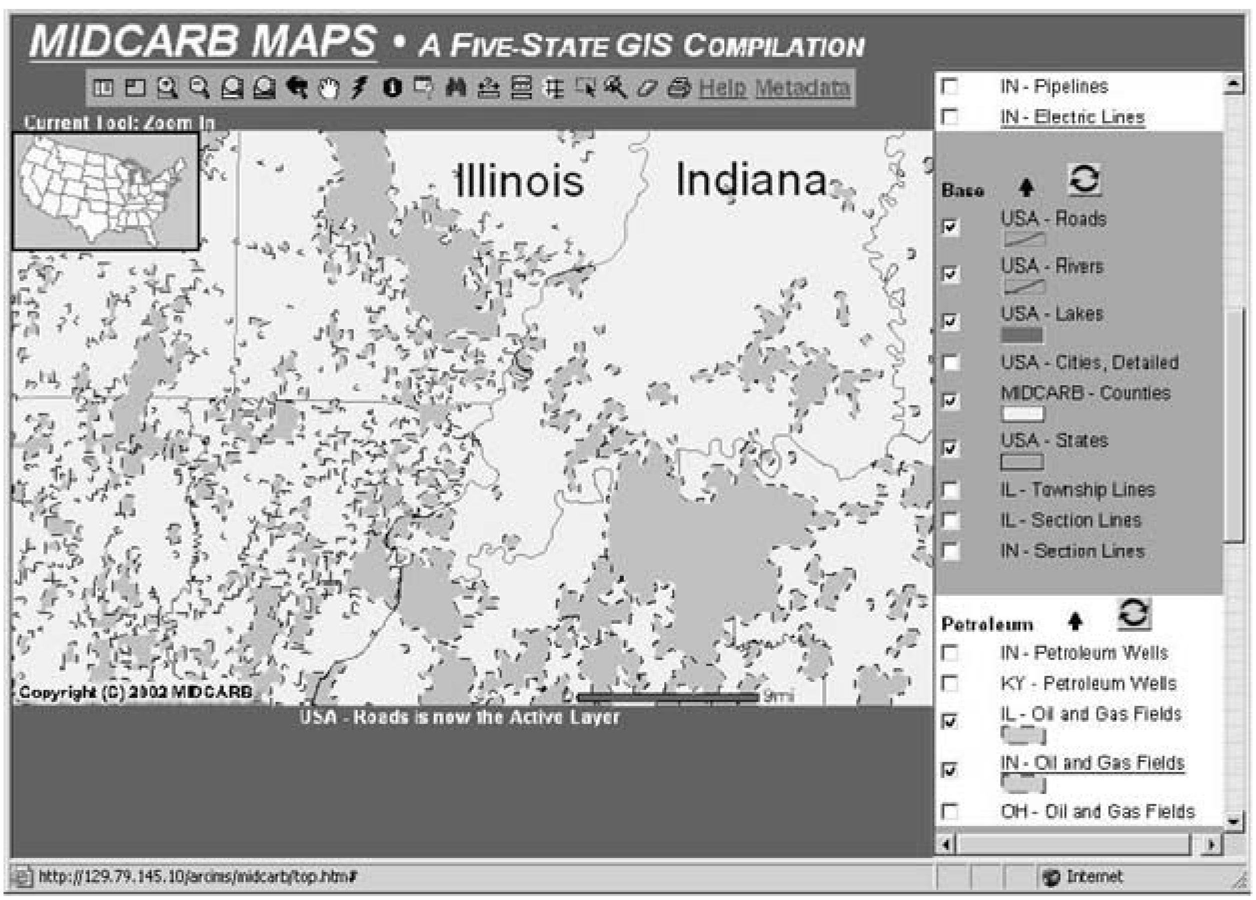


Figure 4. Example view of the MIDCARB ArcIMS site, showing oil and gas fields accessed from the Illinois and Indiana ArcSDE databases.

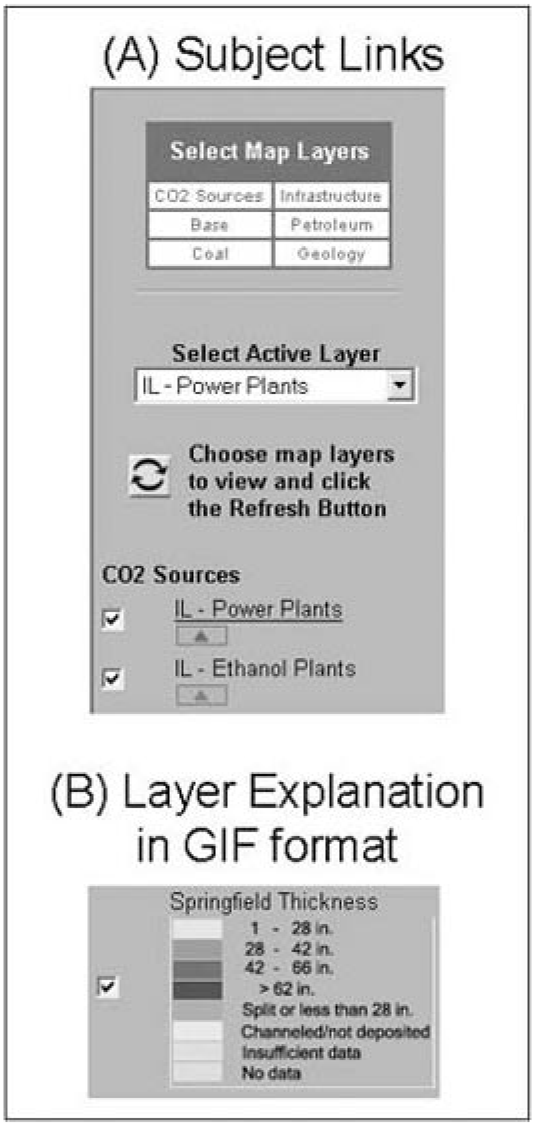


Figure 5. ArcIMS legend customization. **A.** Subject categories (Select Map Layers) provide hyperlinks to respective parts of the legend. Clicking on CO_2 Sources results in scrolling to that part of legend. **B.** Active feature classes (checked items in legend) expand to custom explanations that are inserted as GIF images.

ColdFusion software is being used to interpret browser requests for data, retrieve those data from the appropriate state databases, and return a formatted report to the user. The ColdFusion processor resides at the same location as the ArcIMS server; however, this is not required. Figure 6 shows the architecture of the data pathways used to prepare a ColdFusion report of Illinois database information for a hypothetical user request. This request is initiated by the ArcIMS hotlink tool when a user clicks on a feature. ColdFusion reports are returned to the browser in HTML format. Some data, like power plant emissions, are more clearly represented in graphical form. In these cases, the ColdFusion server passes the database information to a Java program for graph preparation. Both these functions are efficiently processed in a matter of seconds.

Future ArcIMS Development

The distributed nature of the data in MIDCARB is relatively transparent to the user. The fact that related themes are coming from different feature classes (on different servers) is not, because individual database themes have a separate entry in the ArcIMS table of contents (e.g., each state's oil and gas wells are displayed as a separate legend item). Showing only one legend item for the collection of related feature themes from each state would cause less confusion for the user. One approach to this problem would use custom programming to consolidate the legend. An alternative would be to create an ArcSDE database view of the related themes so that ArcIMS would only need to connect to a single database. The latter method may also remedy a problem related to multiple ArcSDE connections. ArcIMS may malfunction if one or more attached ArcSDE services becomes unavailable. Consolidating remote ArcSDE databases into a single database view may prevent these malfunctions, but could also affect performance.

The number of potential themes available to MIDCARB users is large and continues to grow. Much work needs to be done to simplify the user interface so that only desired themes are shown. This simplification could be accomplished with a query interface to collect information from users about what they want to see. The results of each query would be used to construct a customized view of the MIDCARB data.

The current implementation of the MIDCARB ArcIMS service uses version 3.1 software. New capabilities for accessing metadata from an ArcSDE database will greatly enhance the functionality of the service when version 4.0 is implemented.

CONCLUSIONS

The integration of spatial and tabular information in a Web environment has clear advantages for organizations that are collaborating with other institutions in research and public service programs. The principal benefits are that each agency can continue to maintain their own data and ensure that the Web service provides up-to-date information. ArcSDE and ArcIMS appear to be efficient environments for achieving this goal. The MIDCARB project has also demonstrated interesting opportunities for institutional data sharing using direct connections among ArcSDE databases with other ESRI applications.

The technological challenges for implementing a distributed data site are considerable, but the greatest challenge is designing an interface that clearly communicates the function of the site and how to use it.

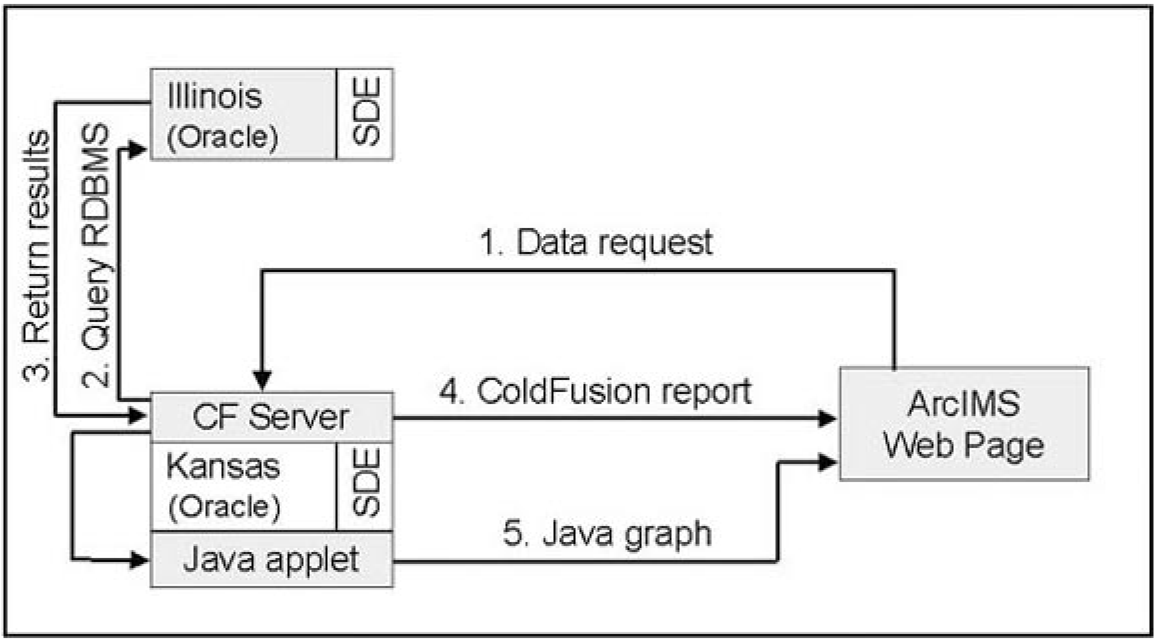


Figure 6. Data pathways for a typical tabular data request. **1.** A data request is sent from the Web browser to the ArcIMS server that passes it to the ColdFusion Server. **2.** The ColdFusion server sends a formatted SQL request to the appropriate databases. **3.** The RDBMS returns the requested data to ColdFusion. **4.** The ColdFusion Server prepares a formatted HTML report and returns it to the user's browser. **5.** For some requests, data are transferred to a JAVA program that prepares a formatted graph and returns it to the browser.

Developing a Working Database for Mapping and Modeling in Illinois

By Barbara Stiff

Illinois State Geological Survey
615 East Peabody Drive
Champaign, IL 61820
Telephone: (217) 244-2520
Fax: (217) 333-2830
e-mail: stiff@isgs.uiuc.edu

INTRODUCTION

The Illinois State Geological Survey (ISGS) has focused on development, production, and archiving of geologic maps for nearly a century. Central to development of geologic maps are the observations, measurements, analyses, and interpretations of research scientists working in the field. Historically, these field-acquired data and associated research materials were recorded, compiled, and maintained on paper by individual scientists. At the completion of the study, field maps and notebooks were archived in the ISGS library.

Currently, digital maps and models of geologic materials in both two and three dimensions are being developed by interdisciplinary teams of scientists. The transition to digital methods and team-based research requires new methods for recording and managing field-acquired and field-verified data (field data). Ideally these methods need to provide all team members with simultaneous desktop access to the most current data.

A working database of customized tables is being developed to accept field data, including site information, observations, and interpretations. Data interchange is via write-access forms that allow researchers to query, append, and edit the data housed in the working database. All ISGS staff members have read-only access to these data as they are developed. The data are maintained as part of the ISGS database system.

DIGITAL BACKGROUND

The transition to digital technology began in 1968 with tabular entry of well drillers' logs into a system called "3-card." A computer-based mapping system to help with the presentation of research data followed (Swann and others, 1970). These data were converted to a geographic information system (GIS) in the 1980's, and the tabular well data were subsequently extracted to form the core of the ISGS master well database (archival database). With advancements in digital technology, acquisition of geologic data, associated research, and presentation of research results have "gone digital" (Stiff, 1997; Krumm and others, 1997; Stiff and others, 1998; Luman and others, 1998; Stiff and Hansel, 1999; Abert and others, 2000; Hansel and others, 2001).

For each research study, ISGS field-acquired data, both old and new, provide detailed information for the maps and models. These data may include descriptions from study-specific drilling, geologic samples, and outcrops and exposures; aerial and terrestrial photographs; shallow seismic reflection and refraction profiles; wireline and natural gamma logs from both new and old boreholes; various other well data; field notes from ISGS library archives and previously published studies; and results of chemical and laboratory analyses. The data are compiled from various combinations of digital, analog, and hardcopy sources and converted to the digital format according to the needs of individual members of the research team. The resulting digital database is a diffuse aggregation of individually maintained word-processing, spreadsheet, and image files plus GIS coverages, shape and geodatabase files, and associated tabular data.

The ISGS maintains an extensive well-records library that contains both digital and paper records. The digital, tabular data reside in an ISGS archival well database that presently contains records from 488,347 sites, including water wells, oil and gas wells, and various test and engineering borings. It was designed primarily as a digital repository for well and boring records and includes location, driller descriptions, and well-completion information. All data entry and editing are performed by the Geologic Records Unit (GRU) staff to protect database integrity. Data in the archival database are entered as they appear on paper records submitted by drillers and/or site observers. Location information is initially recorded as Public Land Survey System (PLSS) section, township, range, and quarters.

A new, Web-based, working database is being developed to accept the input of field data. Its purpose is to facilitate retrieval, interpretation, and analysis for geologic mapping, and modeling. It is based on relational database management principles using existing Oracle and Arc/Info GIS software.

Field scientists may edit, append, and delete the data in the working database. Geographic descriptions, attribute information, and source-reference data may be tailored to specific project needs, are adaptable to the particular methodologies of individual scientists, and are flexible to allow for future extensions. The working database will provide an interactive level of quality assurance/ quality control (QA/QC) not currently available, and increase efficiency of data input and archiving. Data for each mapping project is monitored by a project-specific data specialist, who oversees timely input, maintenance, and accuracy. The field database also accepts legacy field data from previous research, publications, and reports. Archival well records that require change as the result of field-verification may be merged into the working database when necessary.

Data entry is adaptable to the needs of the researchers. Geologists may input study data directly into the working database tables via Web-enabled forms. These forms also permit database query and edit functions for existing records in either the working database or the archival database. Edits are written only to the working database. Alternatively, data may be entered into formatted spreadsheets that are then batch-loaded into the working database using Structured Query Language (SQL) Loader.

WORKING DATABASE TABLES AND INPUT FORMS

A specific “grammar” was developed for field data following the Digital Geologic Map Data Model, version 4.3 (Johnson and others, 1999). The data are defined as singular objects obtained by direct observation. The GIS objects are treated as point data. Each point has x/y coordinates and a unique identification code (SYS_ID). This identification code is used to establish the relationship to all associated tabular data including location information and geographic description (FIELD_NOTES_HEADER table) and geologic observations and interpretations (FIELD_DESCRIPTION and DESCRIPTION tables).

Input and data editing are accomplished by means of Web forms called managers. The ID manager is used to query the database for available data. The header manager is used to enter and update site information. The description manager is used to enter and edit field data in prose that preserves verbatim field notes and observations. The interpretation manager permits interpreted descriptions to be entered in the archival database. The interpretation manager uses a standardized format and geologic vocabulary that expedites transfer to modeling software but also accepts observations and comments that reflect the special preferences of individual scientists.

ID MANAGER

The ID manager lists the various identification codes that have been assigned to a particular site (Table 1). The API-Number, Field-ID and Mapper-ID are compound relational identifiers (Figure 1). The API-Number is the unique identifier assigned by the GRU to all well records that reside in the archival well-data tables. These data may only be entered and edited by GRU staff in order to protect the integrity of these data. The Field-ID and Mapper-ID are compounded of study, scientist, year, site type, and sequential number codes so that users may query by all or any part of these ID’s depending upon their needs. The Site-ID records identifiers used by individual mappers on field maps, in field notebooks, and in publications. The Site-ID is particularly important for historic data. The Legacy-ID is used as a bibliographic reference and may include library identification codes. The layout of the ID manager is shown in Figure 2.

HEADER MANAGER

Header data include site and location information (Table 2). Many of these data are housed in the archival

Table 1. Structure of the Field-ID table.

DB_ITEM_NAME	FORM	EXPLANATION
SYS_ID	N(12)	sequential, system-generated identifier
API_NUMBER	C(12)	assigned 12-character database identifier (e.g., 12097004500)
FIELD_ID	C(10)	study-specific id,field code,sequential numbers (e.g., IL29X00045)
MAPPER_ID	C(12)	study-specific id,field code,sequential numbers (e.g., JAL1953X0015)
SITE_ID	C(10)	mapper-specific identifier (e.g., EDM-1, JL-176, etc.)
SAMPLE_SET	C(5)	laboratory-assigned sample-id
LEGACY_ID	C(100)	may add several mapper-id’s from the literature, archives, etc.

Oracle DB - API number				**Working DB - Field_ID**			**Mapper_ID**			
state	county	co_number	workover	study	type	number	initials	year	type	number
xx	xxx	xxxxx	xx	xxxx	x	xxxxx	xxx	xxxx	x	xxxx
120970013400				WI02X0013			JAL1953X0013			

Figure 1. Components of the compound relational identifiers. The type codes within the Field_ID and Mapper_ID include: X, outcrop or section; E, engineering test; H, highway/bridge boring; I, ISGS test hole; A, hand auger; etc.

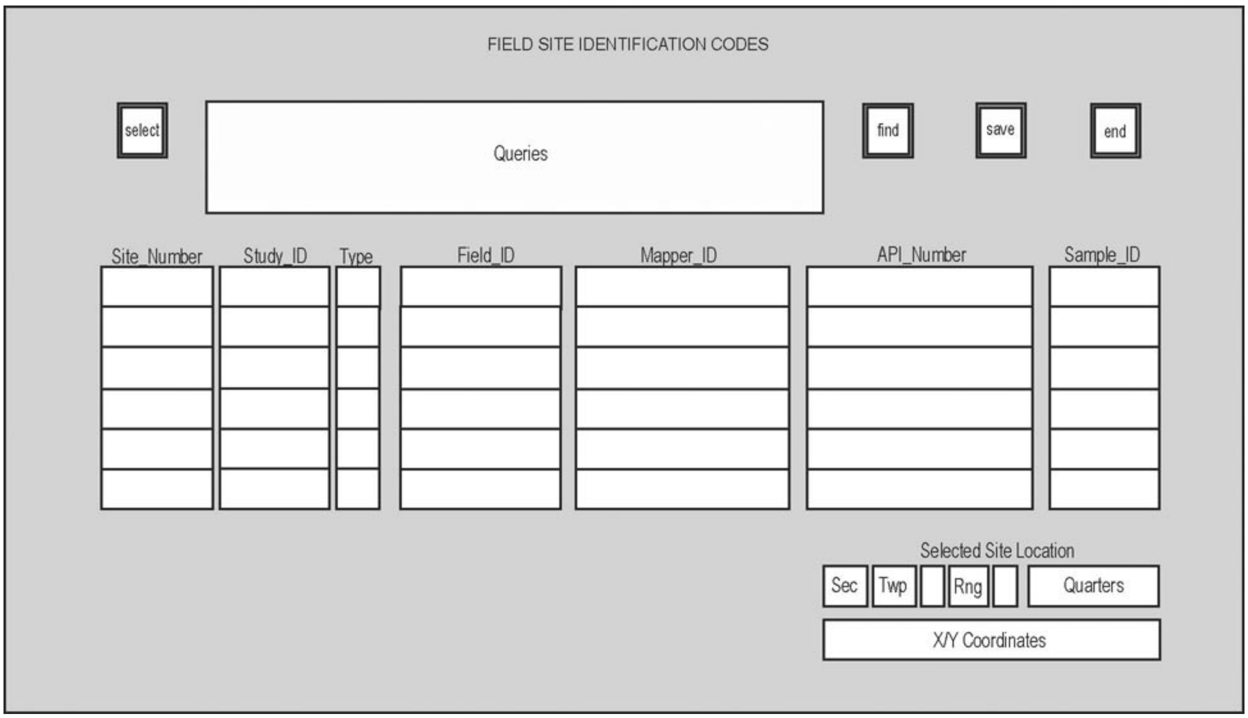


Figure 2. The ID manager. To access data about a particular site, the Field_ID, Mapper_ID, or API_Number is entered into the appropriate box and the select button is clicked. The database is queried and the remaining boxes are populated from both the archival and working databases. Additional ID's may be added, or existing data will be changed in the working data set. Archival data (e.g., API_Number and Sample_Set fields) are for informational purposes and may not be changed or added. Date and authorship are automatically recorded for additions and edits.

data tables. If the scientist has more accurate or more complete data, this table allows the scientist to make appropriate entries or edits. It also contains fields for site information requested by the field scientists. Figure 3 is a graphic representation of the Web-form layout of the header manager.

DESCRIPTION MANAGER

The detailed observations made by the field scientist are critical to a field-notes archive. Although standardization helps with mapping- and modeling-software compatibility, it is essential that the ideas and perspective of the individual scientist be respected/recorded. This table contains fields relating to the particular needs of Quaternary geologists describing surficial materials (Table 3). The fields were developed by researching field notebooks and conferring with individual scientists. Figure 4 is a graphic representation of the Web-form layout of the description manager.

The description manager provides a form by which data entry staff may enter field notes verbatim, without the restriction of having to interpret what the scientist wrote. It also allows entry of historic descriptions from key stratigraphic sites that have been published in Field Guidebooks and other ISGS publications.

Table 2. Structure of the Field-Notes-Header table.

DB_ITEM_NAME	FORM	EXPLANATION
SYS_ID	N(12)	sequential, system-generated identifier
FIELD_LOC_VERBOSE	C(500)	verbose field/site description
FIELD_WHO	C(15)	initials of field scientists
FIELD_DATE	D	field date of the description (e.g., 01/05/2002 or JAN-05-2002)
FIELD_SITE	C(30)	west side of county road, Illinois River bluff, etc.
FIELD_TYPE	C(20)	gravel pit, road cut, borrow pit, etc.
FIELD_CONDITION	C(30)	pit inactive, covered, wet and slippery, etc.
FIELD_NAME	C(30)	Cottonwood School Section, Rattlesnale Hollow Cut I, etc.
ELEV	N(5)	elevation
SEC_ELEV_SOURCE	C(3)	GPS, DEM, DLG, TOP(topo map), IDO(IDOT supplied), etc.
SECTION_ELEV_REF	C(1)	T (top) or B (bottom)
DATA_UNITS	C(1)	F (feet), M (meters), I (inches), C (centimeters)
SEC_LOC_SOURCE	C(3)	GPS, DRG, DOQ, DLG, TOP, TAX, PBK, PMT, DRL, LOG
UTMEASTING83	N(14,6)	meters (up to 6 decimal places)
UTMNORTHING83	N(14,6)	meters (up to 6 decimal places)
LATITUDE83	N(10,6)	decimal degrees (6 decimal places required)
LONGITUDE83	N(10,6)	decimal degrees (6 decimal places required)
LAMFEETX27	N(10,2)	feet (up to 2 decimal places)
LAMFEETY27	N(10,2)	feet (up to 2 decimal places)
SECTION	N(2)	section number
TOWN	N(2)	township number
TDIR	C(1)	N or S
RANGE	N(2)	range number
RDIR	C(1)	E or W
QUARTER	C(11)	up to 4 quarters (e.g., NE NW NW SE)
NSFOOT	N(5)	measurement in feet from reference corner
NSDIR	C(1)	direction of measurement
EWFOOT	N(5)	measurement in feet from reference corner
EWDIR	C(1)	direction of measurement
CORNER	C(2)	corner from which footages were measured
OWNER_NAME	C(30)	name of the property owner
COUNTY	C(3)	FIPS county code
QUAD	C(25)	7.5-minute quadrangle name
TOTAL_MEASURED	N(10,2)	total thickness of the section
ENTERED_BY	C(3)	initials of individual entering data
ENTERED_DATE	D	date of data entry (e.g., 01/05/2002 or JAN-05-2002)
SAMPLE_SETS	C(1)	samples taken, Y or N
CO_NAME	C(20)	county name
QUAD15	C(15)	15-minute quadrangle name
CITATION	C(35)	bibliographic references for site
FIELD_COMMENTS	C(300)	miscellaneous comments regarding field site

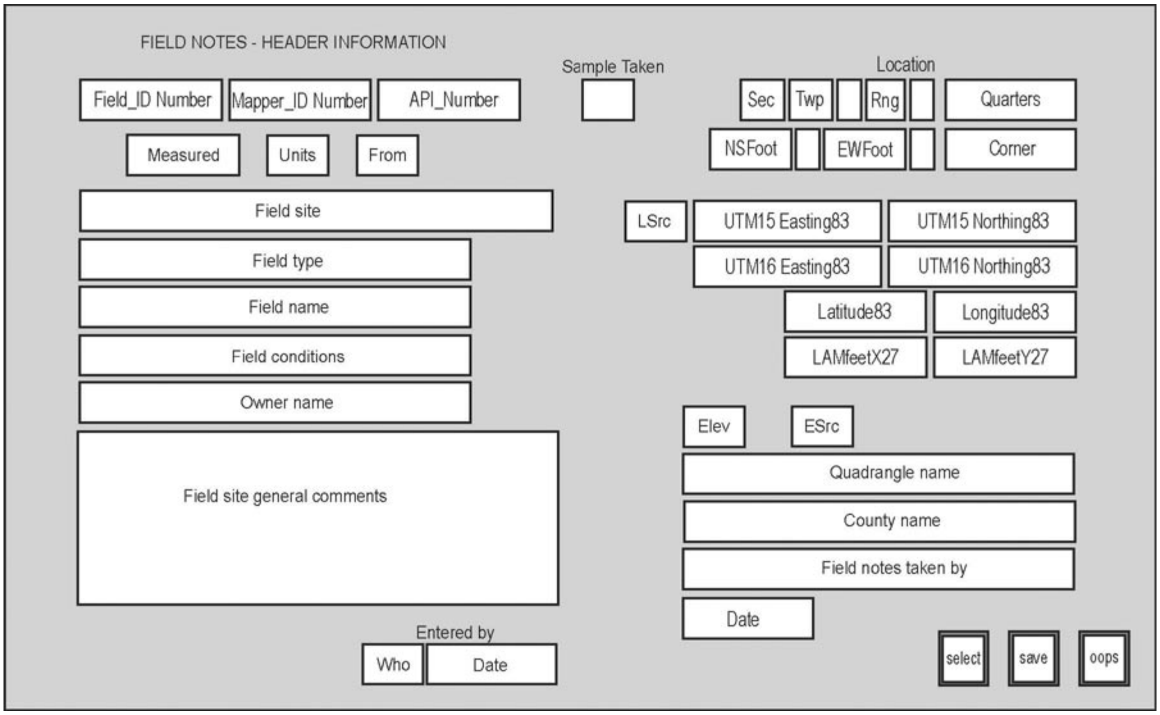


Figure 3. The header manager. To access data about a site, the Field_ID, Mapper_ID, or API_Number is entered into the appropriate box and the select button is clicked. The database is queried and the remaining boxes are populated from available data. Additional information may be added, or existing data may be changed. Date and authorship are automatically recorded for additions and edits.

Table 3. Structure of the description fields form the Field-Descriptions table.

DB_ITEM_NAME	FORM	EXPLANATION
DESCRIPTION	C(2000)	verbose unit description
MATERIAL_VERBOSE	C(200)	verbose material description as recorded in field notes, etc.
TEXTURAL_VERBOSE	C(100)	verbose textural description as recorded in field notes, etc.
COLOR_VERBOSE	C(100)	verbose color description as recorded in field notes, etc.
CLAST_VEBOSE	C(100)	verbose clast description as recorded in field notes, etc.
LITHOLOGIC_VERBOSE	C(300)	verbose lithology as recorded in field notes, etc.
PEDOLOGIC_VERBOSE	C(300)	verbose soils description as recorded in field notes, etc.
GEOMORPHIC_VERBOSE	C(300)	verbose site characterization as recorded in field notes, etc.

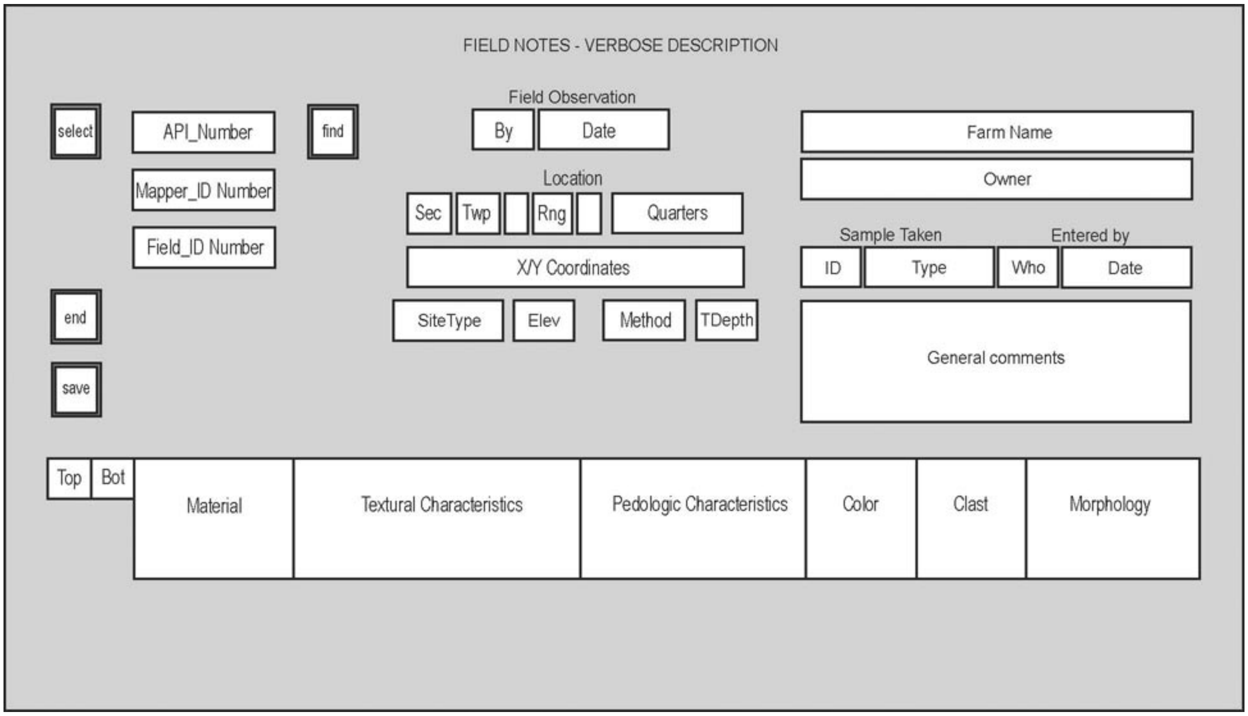


Figure 4. The description manager. To access data about a site, click the select button, enter the Field_ID, Mapper_ID, or API_Number and click the select button. The database is queried and the remaining boxes are populated from available data. Additional descriptive information may be added, or existing data may be changed. Text added to the general comments, material, textural characteristics, pedologic characteristics, color, clast, and morphology boxes wraps to accommodate the "prose" format of the data. Date and authorship are automatically recorded for additions and edits.

INTERPRETATION MANAGER

The classes and types of descriptive information used in the data-model structure (Johnson and others, 1999) do not include fields for the detailed information gathered by mappers of surficial materials in Illinois. Field notes and published descriptions by ISGS scientists were used to develop a preliminary set of descriptive criteria. These criteria were distributed to mappers for their comments and input. The current structure of the descriptions table (Table 4) contains fields that are diagnostic for the glacial sediments at the surface in Illinois. In addition to lithologic, soils, chemical, and engineering properties and geomorphic fields listed, several classification codes are included in the data table. The various classification codes embed data classifiers from other digital data sets (e.g., USDA soils classifications, AASHTO engineering codes, etc.) in the ISGS data set.

ALTERNATIVE DATA ENTRY

Previously, scientists kept digital records of field notes in word-processing software. They are being encouraged to enter data into spreadsheet software formatted for efficient import into the working database using SQL Loader when online forms are not available for data entry. Fields from the digital files as used to define and describe the columns in the Microsoft Excel spreadsheet.

CONCLUSION

Though designed parallel to the archival data sets, the field-data sets differ in several key areas. Because these data sets are "open," they have both benefits and risks. Benefits include: timely and direct data entry, centrality, multi-access, data-location reference maintained at discretion of project team (in UTM15NAD83, UTM16NAD83, LambertNAD27, or Latitude/Longitude), interactive QA/QC performed by the team involved with the data, reduction/simplification of load on data entry staff, and a fixed data format understood by everyone. The risk is the open access to the data. All authorized users will be able to make changes to the data. The Input forms permit ISGS scientists to enter data directly into the ISGS Oracle database from wherever they access the Internet.

Table 4. Structure of the Descriptions table grouped by category.

DB_ITEM_NAME	FORM	EXPLANATION
SYS_ID	N(12)	sequential, system-generated identifier
WHO	C(3)	initials of describing individual
INTERP_DATE	D(11)	field date of the description (e.g., 01/05/2002 or JAN-05-2002)
TOP	N(7,2)	top of the unit (up to 2 decimal places)
BOTTOM	N(7,2)	bottom of the unit (up to 2 decimal places)
Lithology		
MATERIAL	C(20)	Dm, silt, sand, gravel, etc.
MATERIAL_COMMENT	C(100)	descriptive details relating to material
TEXTURE	C(50)	loam, silty clay loam, etc.
TEXTURE_COMMENT	C(100)	descriptive details relating to texture
FRAMEWORK	C(20)	matrix supported, etc.
FRAMEW_COMMENTS	C(30)	descriptive details relating to framework
COLOR_DESCRIPTIVE	C(30)	light yellowish brown, etc.
COLOR_MUNSELL	C(25)	5YR3/4, 5YR3/4-2.5YR5/6, etc.
COLOR_COMMENT	C(50)	descriptive details relating to color
COLOR_CONDITIONS	C(5)	wet, dry, etc.
Clasts and Inclusions		
CLAST_TYPE	C(20)	limestone, exotic, granitic, slate, etc.
CLAST_SIZE	C(15)	pebble, cobble, 5 mm, etc.
CLAST_QUANTITY	C(15)	few, many, etc.
CLAST_ROUNDING	C(15)	angular, subangular, etc.
CLAST_DESCRIPTION	C(40)	striated, etc.
CLAST_COMMENTS	C(30)	descriptive details relating to CLASTS
Soils		
ORGANICS	C(20)	numerous rootlets, etc.
BURROWS	C(20)	krotovina, etc.
COATINGS	C(20)	clay skins, iron stains, etc.
STRUCTURE	C(30)	blocky, etc.
STRUCT_COMMENTS	C(30)	descriptive details relating to structure
PEDOLOGIC_NAME	C(20)	Sangamon Soil, etc.
PEDOLOGIC_COMMENT	C(100)	descriptive details relating to PEDOLOGIC_NAME
SOIL_HORIZON	C(10)	A, B2, etc.
Chemical Properties		
REACTIVITY	C(20)	calcareous, leached, etc.
REACT_COMMENTS	C(30)	descriptive details relating to reactivity (violent, slight, etc.)
REDOX_FEATURES	C(1)	Y or N
REDOX_COMMENTS	C(30)	descriptive details relating to redox
Engineering Properties		
MOISTURE	C(10)	wet, moist, etc.
CONSISTENCY	C(15)	firm, sticky, plastic, etc.
PLASTICITY	C(10)	slight, very, etc.
ODOR	C(15)	petrochem, musty, etc.
FRACTURES	C(1)	Y or N
FRACT_COMMENTS	C(30)	descriptive details relating to fracture features
FRAC_VERT_INCLINE	N(3,2)	in degrees (up to 2 decimal places)
FRAC_HORIZ_INCLINE	N(3,2)	in degrees (up to 2 decimal places)
Stratigraphy		
UNIT_NUMBER	N(9)	ISGS numeric code (used in generating digital maps)
UNIT_NAME	C(20)	Wadsworth, Lemont, etc.
UNIT_COMMENTS	C(50)	descriptive details relating to the unit
Depositional		
CONTACT	C(15)	clear, gradational, etc.
CONTACT_COMMENTS	C(30)	descriptive details relating to the contact
STRATIFICATION	C(1)	Y or N
STRAT_COMMENTS	C(30)	descriptive details relating to stratification
DEPOSITIONAL_ENVIRONS	C(30)	till, loess, lacustrine, debris flow, etc.
Classification		
USCS_CLASS	C(5)	Unified Soil Classification System code (CL-ML, GM, etc.)
USDA_CLASS	C(8)	U.S. Department of Agriculture soils designation code (127C2, 8E2, etc.)
IDOT_CLASS	C(20)	Illinois Department of Transportation soil texture designation (loam, etc.)
AASHTO_CLASS	C(10)	engineering (highway) codes (A-7(5), A-7-5, A-7,6(15), etc.)
THICKNESS	N(7,2)	thickness of the unit (up to 2 decimal places)
FORM_CODE1	C(1)	primary material classification (see attached ESCODES)
FORM_CODE2	C(3)	descriptive classification (see attached ESCODES)
FORM_CODE3	C(2)	secondary descriptive classification (see attached ESCODES)

ACKNOWLEDGMENTS

"Recommended guidelines for the development of the Illinois State Geological Survey digital geologic map database" (Illinois State Geological Survey, 2002) were modified from a Request For Proposals (ASPS 2002-1000-2669) issued by the Alaska Department of Geological and Geophysical Surveys (see Freeman, 2001). The Alaska document provided a framework for ISGS discussions regarding database development.

REFERENCES

Abert, C.C., Weibel, C.P., and Berg, R.C., 2000, Three dimensional geologic mapping of the Villa Grove quadrangle, Douglas County, Illinois, *in* Soller, D.R., ed., Digital Mapping Techniques '00—Workshop Proceedings: U.S. Geological Survey Open-File Report 00-325, p. 125-129, <http://pubs.usgs.gov/of/of00-325/abert.html>.

Freeman, L.K., 2001, A case study in database design: the Alaska Geologic Database, *in* Soller, D.R., ed., Digital Mapping Techniques '01—Workshop Proceedings: U.S. Geological Survey Open-File Report 01-223, p. 31-34, <http://pubs.usgs.gov/of/2001/of01-223/freeman.html>.

Hansel, A.K., Stiff, B.J., and Pugin, A., 2001, Creating a regional 3-D model of Quaternary deposits for mapping projects in northeastern Illinois: Geological Society of America Abstracts with Programs, v. 33, no. 6.

Johnson, B.R., Brodaric, B., Gaines, G.L., Hastings, J.T., and Wahl, R., 1999, Digital Geologic Map Data Model, version 4.3: AASG/USGS Geologic Map Data Model Working Group, <http://geology.usgs.gov/dm>.

Krumm, R.J., Abert, C.C., Nelson, D.O., and Hester, J.C., 1997, Review of digital mapping techniques: the Illinois experience, *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. 5-8, <http://pubs.usgs.gov/openfile/of97-269/krumm.html>.

Luman, D.E., Smith, L.R., and Stohr, C., 1998, Remote sensing inputs to a geologic mapping program for Illinois, *in* Soller, D.R., ed., Digital Mapping Techniques '98—Workshop Proceedings: U.S. Geological Survey Open-File Report 98-487, p. 19-25, <http://pubs.usgs.gov/openfile/of98-487/luman.html>.

Illinois State Geological Survey, 2002, Recommended guidelines for the development of the Illinois State Geological Survey digital geologic map database and development of guidelines for the ISGS digital geologic map database and recommended follow-on activities: Report of Working Group 2 of the ISGS Mapping Steering Committee.

Stiff, B.J., 1997, Use of raster imagery and vector data in support of a geologic mapping program, *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. 23-26, <http://pubs.usgs.gov/openfile/of97-269/stiff.html>.

Stiff, B. J., Beaverson, S.K., and Krumm, R.J., 1998, Streamlining quadrangle map production for on-demand publication, *in* Soller, D.R., ed., Digital Mapping Techniques '98—Workshop Proceedings: U.S. Geological Survey Open-File Report 98-487, p. 53-56, <http://pubs.usgs.gov/openfile/of98-487/stiff.html>.

Stiff, B.J., and Hansel, A.K., 1999, Utility of digital geologic data in addressing outreach requests: Geological Society of America Abstracts with Programs, v. 31, no. 5.

Swann, D.H., DuMontelle,P.B., Mast, R.F., and VanDyke, L.H., 1970, ILLIMAP—a computer-based mapping system for Illinois: Illinois State Geological Survey Circular 451, 24 p.

Surveyors and GIS Professionals Reach Accord

By Bruce A. Joffe, AICP

GIS Consultants
1615 Broadway, Suite 415
Oakland, CA 94612
Telephone: (510) 238-9771
Fax: (510) 654-0196
e-mail: GIS.Consultants@joffes.com

INTRODUCTION

After 13 months of negotiation, representatives from five Surveyor professional organizations and two GIS organizations reached agreement on changing the National Council of Examiners for Engineering and Surveying (NCEES) Model Law that defines the practice of surveying for which licensure is required. The NCEES comprises representatives from each state's Board of Registration and provides guidelines for state laws concerning professional licensure.

BACKGROUND

The American Society for Photogrammetry and Remote Sensing (ASPRS) organized a multi-association task force[1] to consider the current NCEES policy on the responsibility of Surveyors to create and maintain information stored in geographic information systems (GIS). The ASPRS-sponsored Task Force met 32 times by teleconference in a conscious effort by all members to understand and appreciate the varying perspectives on the issues and practices of the Surveyor and GIS disciplines. Over 650 professional hours were invested. The result of these negotiations is a broad-based consensus on a series of recommendations for NCEES concerning the legal responsibilities of Professional Surveyors with respect to the use of GIS and land information systems (LIS).

The GIS-related concerns included a general perception that the language of the current NCEES Model Law on Surveying can be interpreted to over-reach the legitimate professional jurisdiction of the practice of surveying with regard to the creation and maintenance of maps and databases in GIS. Surveyors' concerns, recognized by all the Task Force members, were that GIS/LIS tools are potentially being used by nonlicensed practitioners in activities that clearly fall within the long-established responsibility of the licensed surveyor.

The goal of the Task Force was to recommend modifications to the NCEES Model Law that would clearly identify those activities requiring the services of a registered professional, in order to safeguard the public health, safety, and welfare. The resulting recommendations have gained the support of each of the seven participating associations. During the fall and winter of 2001-2002, a subcommittee of the NCEES reviewed the recommendations; the recommended modifications of the Model Law were approved by the NCEES at its summer 2002 meeting.

CURRENT LAWS

All 50 states have professional licensing laws that define the "Practice of Survey." Their definitions vary, but generally include the creation, preparation, and modification of certain types of data that require licensure of the person in responsible charge. The data referred to include the contour of the Earth's surface; the position of fixed objects thereon; the elevation of fixed works embraced within the practice of civil engineering; the location of property lines or boundaries of any parcel of land, rights-of-way, easement, or alignment; and the position of any monument or reference point which marks a property-line boundary. Such data exist in most public agency GIS "framework" layers.

A literal interpretation of many such laws would conclude that agencies with GIS base maps that are not supervised by Licensed Surveyors are in violation. GIS Professionals regard these laws as exclusionary—prohibiting them from doing the work they have been conducting historically.

Traditional survey map products such as subdivision plats, legal records of parcel boundaries, or construction grading plans are clearly the Surveyors' purview. But

[1]The seven participating organizations were: American Congress on Surveying and Mapping (ACSM), American Society of Civil Engineers (ASCE), American Society for Photogrammetry and Remote Sensing (ASPRS), Management Association for Private Photogrammetric Surveyors (MAPPS), National Society of Professional Surveyors (NSPS), National States Geographic Information Council (NSGIC), and Urban and Regional Information Systems Association (URISA).

what about commercially available road maps that show the location of “fixed works” (streets, bridges, etc.), assessors’ tax maps that show the boundaries of parcels, or watershed drainage maps showing contours of the Earth’s surface? These maps, and many similar maps, are being created and used in GIS for inventory and analysis. They are not used to define the authoritative location of boundaries or fixed works.

Many Surveyors concede that the law ought to apply only to “survey products” (which these examples are not); nevertheless, the wording of many state laws, and the national model law, do not indicate such flexibility of interpretation. This is the reason for the Task Force’s assembly and recommendations.

GIS base maps are referential. They are not the legal record of original survey measurements. They are representations or reproductions of information taken from original documents. As such, GIS maps do not carry legal authority to determine a boundary or the location of fixed works, and, therefore, they need not be supervised or regulated as survey products.

RECOMMENDED PRINCIPLES

The Task Force debated at great length the difference between the licensure of practice and the control of the use of tools utilized in a practice. As is true with many sophisticated techniques and technologies, a layperson and a licensed practitioner may be able to accomplish what appear to be similar functions utilizing a common tool set, and often the purposes for those activities may appear to resemble each other. Historically the guiding principle to determine whether an activity or function must be restricted to a licensed practice is if the public health, safety, or welfare is at stake. Thus, the GIS/LIS-related functions were carefully analyzed to determine whether such practice restrictions should apply, not based on the tool or technique used, but rather upon the service, product, or advice delivered. The criteria the task force used to distinguish between the use of GIS technology for survey purposes versus uses of GIS-based techniques for other purposes included the following:

1. A distinction must be made in the use of electronic systems between making or documenting original measurements in the creation of survey products, versus the copying, interpretation, or representation of those measurements in such systems.
2. A distinction must be made according to the intent, use, or purpose of measurement products in electronic systems between determining a definitive location versus using those products as a locational reference for planning, infrastructure management, and general information.
3. GIS databases and maps prepared to be simply referential, representational, or diagrammatic portrayals of existing source documents (many of which were compiled by licensed professionals and are a matter of public record) should not automatically fall under the requirement for supervision by licensed professionals, unless the use of the databases and/or maps is intended to serve as authoritative public records for geographic location.
4. GIS-based databases and maps that are intended to be used as the authoritative document to describe or determine the location of parcels, fixed works, survey monuments, elevation measurements, etc., must be compiled under the responsible charge of a Professional Surveyor or Land Surveyor.
5. Because geospatial technologies are changing very rapidly, references to specific technologies should be removed from the NCEES Model Law and state professional codes. The language of the Model Law should concentrate on the practices to be covered regardless of the technologies employed.

These principles, along with many explicit examples of GIS-related activities requiring the supervision of licensed Surveyors (“inclusions”), as well as examples of GIS-related activities that do not require the supervision of Licensed Surveyors (“exclusions”), may be found in the complete report from the ASPRS Task Force on the following Web site: <http://www.asprs.org/asprs/news/ncees_frame.html>. The file name is “GIS/LIS Addendum to the Report of the Task Force on the NCEES Model Law for Surveying.”

The actual text of the Task Force’s recommendation for the “Preamble” section of the NCEES Model Law is:

The term “Practice of Surveying or Land Surveying” within the intent of this Act shall mean providing, or offering to provide, professional services involving ***both*** *(1) the making of geometric measurements of, and gathering related information pertaining to, the physical or legal features of: the earth, improvements on the earth, the space above the earth, or any part of the earth;* ***and*** *(2) utilization and/or development of these facts into survey products such as graphics, digital data, maps, plans, reports, descriptions, and/or projects.*

Under the recommended changes to the Model Law, these two conditions must be fulfilled to require supervision of a Licensed Surveyor: (1) making original measurements of parcels, fixed works, topography etc., and (2) compiling the measurements into a survey product. “Survey product” is a specific term which the Surveyors on the Task Force were not able to confine to a definition. Instead, the Task Force agreed on “The creation of maps and geo-referenced databases representing **authoritative locations** for boundaries, the location of fixed works, or topography . . . " as stated in paragraphs 1 and 2 of the "Inclusions" section of the Task Force's recommendations. The key phrase is "authoritative locations."

OTHER CONSIDERATIONS

Neither Surveyors or GIS Professionals have yet developed a systematic and consistent methodology for creating and maintaining areawide base maps. Surveyor Lee Hennes (a member of the Task Force) calls this "macro surveying" and acknowledges that it is very different from traditional surveying of individual parcels or tracts. Apocryphal stories abound in the Surveyor community recounting damage that results from the inappropriate use of maps. How can the public be protected from such a threat? GIS Professionals offer a number of recommendations:

- GIS mapped features should explicitly refer to the source documents from which they were compiled. Such linkage could be achieved by carrying a source document identifier in the database record of each mapped feature or linking to scanned images of those source documents.
- GIS mapped features should contain explicit and easy-to-understand metadata. The public can be reasonably assumed to be protected if they are informed about the locational accuracy, currency, and method of compilation (lineage) of the data in a GIS.
- GIS maps and data should contain an explicit statement of intended use and disclaimer from other uses. Specifically, a disclaimer should state: "This is not a survey product."[2] (Note that the context refers to the product of Land Surveyors.)
- GIS maps that have been adjusted (rubbersheeted) to create consistent, coherent display maps should retain the original mapped coordinates as feature attributes, as well as metadata describing the transformation adjustments that were made.

While considering these "base-map certification" issues, one also might consider the implications of a "certified base map." Would such certification usurp some of the legal authority for determining land ownership that currently resides with subdivision plats, deeds, and similar source documents? If so, a government-controlled GIS base map would change the legal basis of boundary determination in this country. Such a change must be decided upon by explicit political expression, not simply as a technical consequence.

Certification of GIS Professionals also raises the implication of liability and responsibility. What liability would a Licensed Surveyor or Licensed GIS Practitioner be willing to accept for potential "damages" caused by GIS data errors, or by the inappropriate use of GIS data? These questions remain open for your consideration.

If you have comments, please use the eForum sponsored by URISA at <http://www.URISA.org/gispolicy.htm> (item #11), or contact the author. Additional information may be downloaded from <ftp://joffes.com>.

---

[2]My personal favorite is the San Diego Water Company's "Caution: Objects in the GIS may be closer than they appear."

Field Description of the Scientific and Locational Accuracy of Geologic Features (a Part of the *Draft* FGDC Geologic Map Symbolization Standard)

By David R. Soller,[1] Taryn A. Lindquist,[2] and Jonathan C. Matti[3]

[1]U.S. Geological Survey
908 National Center
Reston, VA 20192
Telephone: (703) 648-6907
Fax: (703) 648-6937
e-mail: drsoller@usgs.gov

[2]U.S. Geological Survey
Menlo Park, CA 94025
e-mail: tlnquist@usgs.gov

[3]U.S. Geological Survey
Tucson, AZ 85719
e-mail: jmatti@usgs.gov

BACKGROUND

A national standard for the digital cartographic representation of geologic map features is being prepared for approval by the Federal Geographic Data Committee's (FDGC) Geologic Data Subcommittee. Anticipated for approval in 2003, this standard will be applicable to all geologic map information and geologic map databases published by the Federal Government and its Federally funded contractors and collaborators. Nonfederal agencies and private firms that produce geologic map information also are urged to adopt the standard. When approved, the standard will be available at <http://ncgmp.usgs.gov/fgdc_gds/>; at present, background information and the draft standard are found there; the subject of this paper is new material that is not found in that draft.

The standard is intended to support the Nation's producers and users of geologic map information by providing line symbols, point symbols, and colors and patterns that can be used to portray the various features on geologic maps. The objective of the standard is to aid in the production of geologic maps and related products, as well as to help provide geologic maps and products that are more consistent in their appearance and their underlying database content. The imperative for the clear communication of geologic map information to a diverse audience was outlined early in the history of the U.S. Geological Survey (USGS) by then-Director John Wesley Powell, who stated that "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 available alike to the theoretic physicist or astronomer, the practical engineer or miner, and the skilled agriculturist or artisan" (Powell, 1888, p. 229).

The consistent, unambiguous expression of geologic map information is even more critical now because such information increasingly is compiled, stored, manipulated, and exchanged in digital files and geospatial databases. In these files, the cartographic representation of each feature on a geologic map must have a unique and explicit meaning, which also must be compatible with the feature's attributes in the geologic map database. To that end, the preparers of the standard reviewed existing formal and informal USGS geologic map symbolization standards and adapted them for implementation with modern digital mapmaking systems and geospatial databases.

This standard is not intended to be used inflexibly or in a manner that will unduly restrict a geologist's ability to communicate the observations and interpretations gained from geologic mapping. On the contrary, the standard recognizes that, in certain situations, an existing symbol or its use might need to be modified to fit a particular geologic situation or setting. Likewise, the standard recognizes that a new symbol or set of symbols may need to be created to more fully express local geologic conditions or to keep pace with evolving geologic mapping concepts and practices. Accordingly, such new or modified symbols, if found to be of wide applicability, will be incorporated into the standard through planned, periodic revisions.

The standard will contain updated, precise guidance on the selection of symbology for geologic features for which scientific interpretation is uncertain and/or for which locational accuracy is diminished by poor exposure, an inadequate base map, or other factors. At the DMT ‘02 workshop, the section of the standard addressing these issues was introduced and discussed, and guidance and suggested revisions were sought. From that discussion and from comments by the newly formed Map Standards Committee,[1] that section of the draft FGDC standard was slightly revised. In this paper we excerpt a preliminary version of that section to expose the issues to the broader geoscience community.

ACKNOWLEDGMENTS

The draft FGDC standard owes its existence to the well-established history and traditions of geologic map cartography by the USGS. In particular, the editors and compilers of this standard wish to thank the many cartographers, editors, and geologists who contributed to the informal USGS “Technical Cartographic Standards” volume. We especially wish to thank the many members of the USGS Geologic Discipline’s Western Publications Group and the Map Standards Committee, who have made substantial contributions to the design and preparation of this standard. We also thank the DMT ‘02 attendees for their helpful comments, and Nancy Stamm (USGS) for suggestions on the document’s concepts and organization and for creating the figures.

SCIENTIFIC CONFIDENCE AND LOCATIONAL ACCURACY OF GEOLOGIC MAP FEATURES

An important concept in geologic mapping is the geologist’s level of confidence in the interpretation of features observed in the field. Many factors can adversely affect a geologist’s level of confidence when mapping; in the field, interpretation of a feature may be in question, as indicated by the following examples:

- A planar feature is well exposed in outcrop, but it is not easily identifiable as either a contact or a fault.
- A contact is clearly exposed in a roadcut, but its location cannot be followed away from that roadcut.
- A fault is obscured by vegetation, and so both its location and its sense of offset cannot be definitively determined.
- A fault’s location is completely concealed beneath valley fill.

As these examples show, uncertainties can exist in either the scientific interpretation or the mapped location of a feature, or in both. Therefore, not only is it important to communicate to the map user the level of confidence in each geologic map feature, but also which type of uncertainty—scientific and/or locational—may be associated with that feature.

Traditionally, geologic maps have used a system of solid, dashed, dotted, or queried line-symbol styles (for example, Ridgway, 1920, plate 2) to show levels of locational accuracy of planar and linear geologic features observed in the field. This convention followed USGS Director Powell’s 1888 policy that stipulated: “. . . fault lines (particularly when they are formation boundaries) shall be indicated when actually traced by somewhat heavy full lines in black; and when not actually traced, by similar broken lines” (Powell, 1890, p. 76). More guidance was provided in 1956 by USGS Chief Geologist W.H. Bradley, who, in a memorandum to USGS personnel regarding geologic map standards, stated, “The accuracy of location of faults and contacts should be shown by appropriate symbols . . . . Solid lines should be used to indicate accurate locations of features that are geologically identifiable within the plottable limits of the map . . . . Features that are only approximately located should be shown by long dashed lines; those that are indefinite or inferred, by short dashed lines; and those that are concealed, by dotted lines” (W.H. Bradley, written commun., 1956). To further encourage the use of such symbology, Bradley added, “The use of many dashed contacts or faults on a map is not to be construed as a detraction from the quality of the map, and for many maps, it may be undesirable or impossible to achieve sufficiently accurate locations to permit use of solid lines. The quality of the map is not impaired so long as the reader can interpret the accuracy of location.”

In conjunction with these traditional line symbol styles, geologists have used terms such as “known,” “probable,” “certain,” “uncertain,” “accurately located,” “approximately located,” “inferred,” “projected,” “concealed,” and “queried” to express the various levels of confidence of planar and linear geologic features. However, these terms and their associated line-symbol styles have not been used consistently from region to region or from map to map, in part owing to different geologic conditions, mapper’s preferences, and available time,

[1]Members are David R. Soller, coordinator, U.S. Geological Survey; Thomas M. Berg, State Geologist, Ohio; Robert D. Hatcher, Jr., University of Tennessee, Knoxville; Mark A. Jirsa, Minnesota Geological Survey; Taryn A. Lindquist, U.S. Geological Survey; Jonathan C. Matti, U.S. Geological Survey; Jay B. Parrish, State Geologist, Pennsylvania; John C. (Jack) Reed, U.S. Geological Survey; Stephen J. Reynolds, Arizona State University; and Byron D. Stone, U.S. Geological Survey.

funding, and purpose for the mapping. For example, some distinct inconsistencies in the meaning of a solid-line geologic contact have persisted (Figure 1). It is not always clear whether these terms reflect uncertainty in a feature's scientific interpretation, its mapped location, or both. As noted above, this standard endeavors to address these inconsistencies through guidance and standard terminology.

Scientific Confidence

Scientific confidence expresses the geologist's level of certainty regarding the nature, origin, geometry, identity, and even the existence of a geologic feature. The characteristics of the geologic materials and structures, the number of outcrops, and the availability of subsurface or geophysical data directly affect the level of scientific confidence in any area. Experience and resources available to the geologist also affect scientific confidence. These fundamental characteristics of geologic features can be grouped into two distinct but related concepts: *identity* and *existence*.

Identity expresses whether or not the observations and data support the stated nature, origin, or geometry of a mapped geologic feature (for example, a contact versus a fault, or a normal fault versus a thrust fault). On the geologic map, the feature is identified and described using standard symbology; uncertainty in its identity may be included in the explanation of the symbol or the description of the map unit. In the geologic map database, the attribute describing the confidence in a feature's identity is specified as either *certain* or *questionable*.

Existence expresses whether or not the observations and data support the continuity or existence of a concealed geologic feature (for example, a mapped fault versus a postulated subsurface fault). On the geologic map, the feature is identified and described using standard symbology; uncertainty in its continuity or existence may be included in the explanation of the symbol or the description of the map unit. In the geologic map database, the attribute describing the confidence in a feature's existence is specified as either *certain* or *questionable*.

Levels of Scientific Confidence

A geologic map must indicate the level of scientific confidence associated with each mapped feature (both its identity and its existence). This information also should be specified as attributes in the geologic map database, contained in two attribute fields: *identity* (permissible values = certain; questionable), and *existence* (permissible values = certain; questionable).

For some types of geologic map features, the level of scientific confidence also is communicated cartographically. To facilitate the cartographic communication of the two concepts of identity and existence, this standard sets forth the following new terminology to express clearly yet concisely the levels of scientific confidence of the features that are shown on a geologic map:

- *Identity and existence certain*. The identity and the existence of a feature can be determined using relevant observations and scientific judgment; therefore, one can be reasonably confident in the scientific credibility of this interpretation. These criteria are met, for example, when a geologist reasons, "*I am certain that the planar feature I see in this outcrop is a fault*." This is the default condition for all geologic map features unless otherwise stated on the geologic map or in the geologic map database.
- *Identity or existence questionable*. Either the identity or the existence of a feature cannot be determined using relevant observations and scientific judgment; therefore, one cannot be reasonably confident in the scientific credibility of this interpretation. These criteria are met, for example, when a geologist reasons, "*I can see some kind of planar feature in this outcrop, but I cannot be certain whether it is a contact or a fault*," or "*My interpretation requires that a thrust fault be present to account for incongruities in the stratigraphy of these rocks, but I can't be certain because I haven't yet seen one here*."

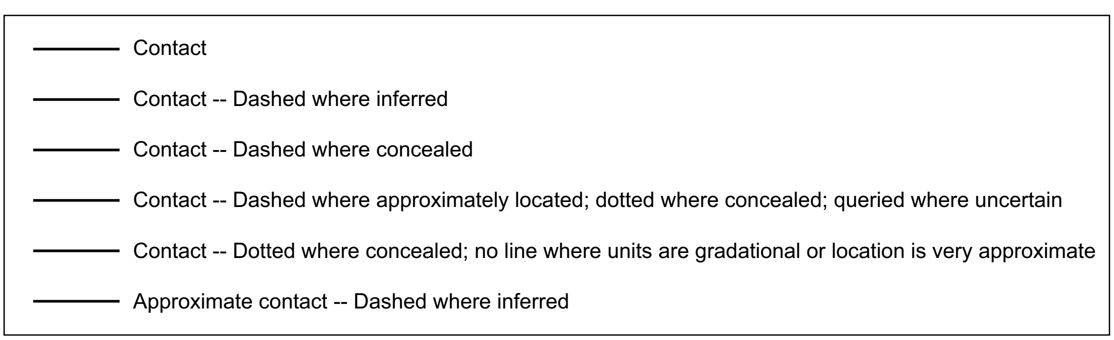


Figure 1. An informal perusal of map-unit explanations on geologic maps published in the past 30 years reveals widely varying meaning for a solid-line geologic contact, depending on the geologic conditions; mapper's preferences; and available time, funding, and purpose for the mapping.

This new terminology is intended to be used for choosing a particular style of symbol to represent a feature on a geologic map (for example, Figure 2 shows the symbolization for geologic contacts), as well as for describing that feature on the map. If a feature is symbolized or described as “identity or existence questionable,” the map user should consult the geologic map database for more complete information.

Locational Accuracy

Locational accuracy involves the interplay between two distinct but related factors: the *locatability* of a feature in the field, and its *positioning* on the map.

Locatability expresses whether a geologist can clearly observe a feature *in the field*, as indicated by the following examples:

- A planar or linear feature is observable in several outcrops along its trace.
- A planar or linear feature is observable in only a few outcrops along its trace, but its physical characteristics permit locating it between outcrops by indirect methods.
- A planar or linear feature is not defined by a distinctive physical trace and so is not observable beneath either vegetation, a thin veneer of unmapped geologic material (colluvium, eolian deposits, or residual soil), or manmade features; therefore, its location must be inferred by indirect means.
- A planar or linear feature is not observable because it is concealed by an overlying geologic map unit, although it may be observable nearby (for example, a thrust fault is visible on both sides of a glacial valley, but its location within the valley is concealed by ice), and so its location must be projected beneath the overlying map unit.

In the explanation of geologic map symbols and in the geologic map database, the attribute describing the confidence in a feature’s locatability is specified as either *observable*, *inferred*, or *concealed* (Figure 3).

Positioning expresses the relative degree of accuracy with which a geologic feature is plotted *on the base map*. Commonly, a feature can be accurately plotted on the map because the base-map information is accurate, detailed, and distinctive. However, in some field situations, a feature cannot be confidently plotted on the map because the topographic, drainage, or cultural information on the base map is insufficiently detailed for the feature to be accurately located relative to features on the map, as indicated by the following examples:

- A feature is observable, but its position on the map cannot be placed accurately because topographic contours, drainage lines, or cultural information on the base map are insufficiently detailed for the feature to be confidently located relative to the various base-map features (for example, a contact is observable in outcrop, but its location in relatively featureless terrain prevents its position from being plotted accurately on the base map).
- A feature is observable, and its geographic coordinates can be determined in the field using a global positioning system (GPS) device or in the laboratory using a georeferenced aerial photograph stereopair; however, the geographic relation between these coordinates and the topographic or cultural setting shown on the base map is not compatible (for example, a feature was mapped on a hillside, but the GPS-derived coordinates, when plotted on the base map, place its position in a valley bottom).

In such situations, either a feature can be plotted relative to the indistinct or incompatible base-map features,

Symbol	Scientific confidence	Examples
——————— — — — — — — - - - - - - - - - - - -	Identity and Existence certain	*"I am certain that the planar feature I see in this outcrop is a fault."*
———?——— — — —?— — — - - -?- - -?- - -?- - - .....?.....?.....?.....	Identity or Existence questionable	*"I can see some kind of planar feature in this outcrop, but I can't be certain if it's a contact or a fault."*

Figure 2. Symbology and new standard terminology to express the level of scientific confidence in the identity and existence of a geologic feature.

Symbol	Scientific confidence	Locational accuracy	
	Identity / Existence	**Location (in the field)**	**Position (on the map)**
————	*certain*	***Observable***	***Within*** **zone of confidence**
——?——	*questionable*		
– – – – – – –	*certain*		***May not be within*** **zone of confidence**
– – –?– – –	*questionable*		
----------	*certain*	***Inferred*** between outcrops or beneath rubble or vegetation	
--?--?--?--	*questionable*		
···················	*certain*	***Concealed*** beneath overlying map unit, ice, or water	
···?···?···?···	*questionable*		

Figure 3. Diagram showing relationship between map symbol and scientific confidence and locational accuracy terminology.

or the locations of topographic contours or other base-map features can be adjusted. (The latter approach is not encouraged unless it is done systematically and is well-documented.) In either case, the inherent uncertainty in a feature's positioning must be conveyed cartographically and recorded in the geologic map database, as explained below.

Specifying the Positional Accuracy

Information about the locational accuracy of mapped features is important to all disciplines, even those in which mapped features commonly are directly observable and can be positioned with a significant degree of accuracy (for example, roads or utilities). It is especially critical in the natural sciences because many mapped features are either interpretive or not directly observable. To specify the locational accuracy of a mapped feature, the geologic mapper must weigh three factors: (1) the nature of the feature and its degree of exposure (for example, a gradational or sharp geologic contact that is poorly or well exposed along a hillside), (2) the quality of the base map (for example, whether the map shows cultural and topographic features positioned accurately according to the geologist's reckoning, triangulation, GPS readings, etc.), and (3) the geologist's confidence in accurately positioning the geologic feature on the base map. Together, these factors determine the mapper's confidence in the locational accuracy of each feature positioned on the base map. The first factor is characterized above as "locatability," and terminology is given for describing it. This section addresses the aspect of locational accuracy involved in positioning an identifiable and observable feature on the base map.

The process of locating a feature in the field and then positioning it on the base map is complex, and the locational accuracy of the mapped feature is not easily described and quantified. In the USGS, stringent policies for the accuracy with which a well-located feature can be *positioned* on the base map have been put forth in the past. For example, Chief Geologist W.H. Bradley's 1956 memorandum to the staff advocated a geologic map accuracy standard based on the United States National Map Accuracy Standards (NMAS) for topographic and other types of base maps. The geologic map adaptation of the NMAS stipulated that "features that . . . can be located from exposures or other evidence [should be positioned] within 1/25 inch [on the map] of their true map position" (W.H. Bradley, written commun., 1956; U.S. Geological Survey, 1995, Part 1, p. 1.0-4). These earlier efforts to quantify the locational accuracy of geologic features were not widely adopted by the geoscience community, likely in part because of (1) the difficulty in translating to geologic mapping a concept designed for topographic and other types of base maps, and (2) the impracticality of requiring that all geologic map information meet the same accuracy criteria uniformly across the Nation, in all types of geologic and topographic settings, and (3) the need to convert ground distance to publication-scale cartographic units before determining whether a feature can be plotted accurately on a base map. In contrast, the standard described in this document advocates a more flexible and conceptually simpler approach in which the accuracy criteria can be defined for each project so that the expressed positional accuracy reflects the character of the geologic setting and other factors.

When a feature is drawn or digitized onto the base map, the geologic mapper commonly has some level of confidence regarding whether the feature has been accurately positioned, on the basis of the three factors noted above. This confidence can be expressed by the likelihood that the feature actually lies within a certain distance from the location where it is positioned on the base map. This distance, extending outward from a feature's position on the map, is designated as the *zone of confidence*, and it characterizes the feature's positional accuracy as follows:

- For planar and linear geologic features, the zone of confidence borders the feature along both sides, forming what is described in GIS terminology as a buffer zone; it is specified as the distance in ground units (for example, in feet or meters) from the feature to the edge of the buffer zone (Figure 4). For geologic point features, the zone of confidence is concentric around the feature, forming a circle, the specified distance being the radius of that circle. The numerical value of the zone of confidence should be provided on the geologic map (in the symbol explanation and the map-unit description) and as an attribute in the geologic map database.

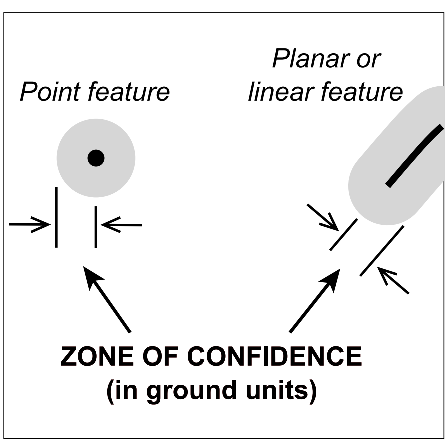


Figure 4. Diagram showing the zone of confidence for point, planar, and linear features.

- For each geologic feature on the map and in the database, the likelihood that it actually lies within the zone of confidence should be indicated (see Figure 3). On the geologic map, this likelihood is conveyed cartographically (for example, an observable and accurately located contact is shown as a solid line; an inferred contact is shown as a dashed line). In the geologic map database, this is conveyed by an attribute specifying that the feature's position is either "*within zone of confidence*" or "*may not be within zone of confidence*." Note that the standard does not stipulate that a feature for which positioning is specified as "may not be within zone of confidence" must *necessarily* be located outside the zone of confidence, but simply that it *may* be.

For any geologic map or mapped area, the numerical value of the zone of confidence will depend on a number of factors: the area's geology, landscape terrain, vegetation cover, and/or cultural features; the scale of mapping; the quality and nature of the base map used; and/or a particular project's allotted field-mapping time or other logistical constraints. Because the standard recognizes that the factors affecting the value of the zone of confidence will vary from region to region (and from map to map), and because different agencies have differing mapping needs and mandates, a universal value for the zone of confidence is not established in the standard. Instead, this standard advocates that the responsibility for setting the value of the zone of confidence for a particular geologic map or mapped area lies with each geoscience organization and each mapping geologist.

For many geologic maps or mapped areas, especially those that are defined by latitude and longitude (for example, quadrangle maps) or political boundaries (for example, state or county maps), one map may contain areas of vastly contrasting geology, topography, vegetation cover, and/or societal infrastructure. For example, a geologic map may include a mountain range underlain by sedimentary rocks and a broad alluvial valley. Clear distinction among the various sedimentary rocks as well as their high relief may provide the geologist with a significantly higher level of confidence in the position of contacts than in the adjacent valley, where few topographic landmarks or contours exist and where geologic contacts are gradational and/or obscured by vegetation and soil cover. In areas as diverse as these, the levels of confidence in positional accuracy will be different, and so the geologist has the following choices: (1) specify different values of the zone of confidence for the two areas, thereby permitting more differentiation of features within each area, or (2) express the differences in confidence solely by differences in line symbology (for example, use mostly solid lines in the mountain range and mostly dotted lines in the valley). The choice might depend on the magnitude of the difference between the areas, or on the geologist's confidence in the positional accuracy of features across the map area. Map compilations represent another example for which different positional accuracy criteria can exist within a single map; a map compilation may be composed of several source maps or mapped areas, each of which may have had a different value specified for the zone of confidence (or perhaps no value had been specified). The geologist may choose to preserve in the map compilation these various zones of confidence.

In the geologic map database, variations in the value of the zone of confidence can be readily accommodated because each feature is assigned (as an attribute in the database) the value of the zone of confidence that has been specified for a particular area. On the geologic map, the areas that have different zone of confidence values should be shown in an index map.

Levels of Locational Accuracy

A geologic map must communicate to the map user the level of locational accuracy associated with each mapped feature (its locatability in the field and its positioning on the base map) by specifying this information as attributes in the geologic map database. The information is contained in the following three attribute fields: (1) locatability (*observable, inferred*, or *concealed*), (2) the numerical value of the zone of confidence (for example, 5 meters), and (3) the likelihood that the feature actually occurs within the zone of confidence *(within zone of confidence, may not be within zone of confidence)*. In addition, the numerical value of the zone of confidence should be indicated on the geologic map, either in a general statement (if one value applies to the entire geologic map) or shown in an index map (if different values apply to different mapped areas); likewise, if a zone of confidence was not used during mapping or map compilation, this should be indicated.

For some types of geologic map features, the level of locational accuracy also is communicated cartographically by using specialized symbology. To facilitate the cartographic communication of the two concepts of locatability and positioning, the standard sets forth the following revised terminology to express clearly yet concisely the levels of locational accuracy of the features that are shown on a geologic map:

- *Location accurate*. A feature is observable, and its plotted position on the map is within the declared zone of confidence. These criteria are met, for example, when a geologist reasons, "*I can clearly see this contact in outcrop, and I can accurately plot its position on the map*." This is the default condition for all geologic map features unless otherwise stated on the geologic map or in the geologic map database.
- *Location approximate*. A feature is observable, but its plotted position on the map may not be within the declared zone of confidence. These criteria are met, for example, when a geologist reasons, "*I can see this contact in outcrop, but I can't tell exactly where it is located because I am surrounded by trees*," or "*I can see this contact in outcrop, but the poor quality of my base map prohibits me from accurately plotting its position*," or "*I can see that the width of the gradational contact between these two map units exceeds my value of the* zone of confidence*, and so, although my base map is of high quality, my confidence in the accuracy of its plotted position is not high*."
- *Location inferred*. A feature is not directly observable between outcrops or beneath rubble or vegetation, so its location must be inferred by indirect means; by definition, its plotted position on the map may not be within the declared zone of confidence. These criteria are met, for example, when a geologist reasons, "*I can see by the change in debris materials visible around these gopher holes that a contact runs through here, but I can't locate it very precisely.*"
- *Location concealed*. A feature is not observable because it is completely concealed beneath an overlying map unit or body of water or ice (although it may be observable nearby); by definition, its plotted position on the map may not be within the declared zone of confidence. These criteria are met, for example, when a geologist reasons, "*I can see that a contact is present on both sides of this lake, but I can't tell where it is located beneath the water.*"

This new terminology is intended for choosing a particular style of symbol to represent a feature on a geologic map (for example, Figure 5 shows the symbolization for geologic contacts) as well as for describing that feature on the map. This terminology also is intended to be used in the attributes of symbols in the geologic map database.

SUMMARY

In part because geologic map information is increasingly used in a digital (GIS) environment in concert with environmental and cultural information, precise terminology is essential for describing the scientific confidence and locational accuracy of geologic features. This section of the pending FGDC standard for geologic map symbolization provides a standard terminology, but retains some flexibility for describing locational accuracy. The standard is intended for periodic review and revision; therefore, the authors welcome your comments and guidance on the issues presented in this paper.

Symbol	Locational accuracy	Examples
	accurately located	*"I can clearly see this contact in outcrop, and can accurately plot its position on the map."*
	approximately located	*"I can see this contact in outcrop, but the poor quality of my base map prohibits me from accurately plotting its position."*
	inferred	*"I can see by the change in debris materials visible around these gopher holes that a contact runs through here, but I can't locate it very precisely."*
	concealed	*"I can see that a contact is present on both sides of this lake, but I can't tell where it is located beneath the water."*

Figure 5. Symbology and new standard terminology to express the locational accuracy of a geologic feature.

REFERENCES

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.

Powell, J.W., 1890, Tenth Annual Report of the United States Geological Survey 1888-89, Part I—Geology: Washington, D.C., U.S. Government Printing Office, 774 p.

Ridgway, J.L., 1920, The preparation of illustrations for reports of the United States Geological Survey: Washington, D.C., U.S. Government Printing Office, 101 p.

U.S. Geological Survey, 1995, Cartographic and digital standard for geologic map information: U.S. Geological Survey Open-File Report 95–525, 257 p.

Progress Report: North American Geologic Map Data Model Design Team

Team members, in alphabetical order:

Eric Boisvert (Geological Survey of Canada, eboisver@nrcan.gc.ca)
Boyan Brodaric (Geological Survey of Canada, brodaric@nrcan.gc.ca)
Jordan Hastings (University of California, Santa Barbara, jordan@geog.ucsb.edu)
Bruce R. Johnson (U.S. Geological Survey, bjohnson@usgs.gov)
James McDonald (Ohio Geological Survey, jim.mcdonald@dnr.state.oh.us)
Stephen M. Richard (Arizona Geological Survey, Steve.Richard@azgs.az.gov)
Peter N. Schweitzer (U.S. Geological Survey, pschweitzer@usgs.gov)
Ronald R. Wahl (U.S. Geological Survey, rwahl@usgs.gov)
Gerald A. Weisenfluh (Kentucky Geological Survey, jerryw@kgs.mm.uky.edu)

The Data Model Design Team (DMDT) was established in 1999 by the North American Data Model Steering Committee (NADMSC) with the purpose of drafting a geologic map data model for consideration as a standard (see <http://geology.usgs.gov/dm/steering/teams/design/charter.shtml>). The team includes experts in GIS, data modeling, databases, and geologic mapping. The team has been holding meetings twice yearly at the Geological Society of America national meeting in the fall and at the Digital Mapping Techniques conference in the spring. Working groups within the team meet as required.

The team has developed a requirements analysis for the model and has documented variant data models that have been developed since the North American Data Model (NADM) version 4.3 (Johnson and others, 1999). On the basis of this input, the team has embarked on a new course of action, developing a geoscience-based conceptual data model (ontology) that will be proposed as a standard basis for the exchange and sharing of information associated with geologic mapping and maps.

REQUIREMENTS ANALYSIS

After reviewing its existing work, the team determined that a documented requirements analysis was necessary as a basis for data model revision. The purpose of the analysis was to determine how geologic map databases might be queried, and to document requirements for database content. The analysis was conducted in conjunction with the NADM's Science Language Technical Team (SLTT) by requesting lists of questions that potential users wanted to be able to answer using a geologic map database. The user community submitted approximately 760 questions. The questions were manually categorized and generalized by SLTT and DMDT members to produce a distillation that consists of 30 types of queries as well as lists of descriptions, classifications, and relationships that must be represented in the database and operations that must be available for analysis of the data. The compiled question lists, the categorizations by various team members, and a summary document are available at <http://geology.usgs.gov/dm/steering/teams/design/>. Table 1 summarizes the representative queries; Table 2 summarizes the lists of descriptions, classifications, and relationships.

VARIANT DOCUMENTATION

The release of the NADM version 4.3 model of Johnson and others (1999) marked the beginning of a phase of experimental implementations. The initial intent was that each would implement the logical model presented by Johnson and others, but in fact each implementation evolved to varying degrees from the NADM 4.3 model. It thus became apparent that another effort was necessary to generalize from the existing implementations to a model that could serve as a standard. The technical team asked the various groups implementing databases related to or derived from NADM 4.3 (referred to as "variants") to document their implementation, comparing and contrasting the implementation against NADM 4.3. This documentation serves as a resource for the revision of NADM 4.3.

A template for the variant documentation was outlined at the technical team meeting held in Boston (November 2001) in conjunction with the Geological Society of America annual meeting. The documentation is to include an abstract, a diagram (schema), a compari-

Table 1. A set of representative and general queries that could be posed to a geologic map information database system.

1 Metadata
- 1.1. For selected object, report metadata information
- 1.2. Given metadata criteria, find all spatial objects that meet specification
- 1.3. Report all citations for classifications and descriptions related to this map view
- 1.4. Report all maps and/or other documents related to an area

2 Classification (concepts)
- 1.5. Select classifications (e.g., map units) of specified name
- 1.6. Select occurrences (e.g., structural features, drill holes) of a given type
- 1.7. Report the description of a class
- 1.8. Given a description, report the classes that satisfy that description (e.g., all Devonian map units)
- 1.9. Report the classes related to a specific descriptions (e.g., map units in the hanging wall of the Bozo thrust)

3 Description
- 1.10. Select map units that contain >10% of lithology X (or subtypes)
- 1.11. Select normal faults within XYZ fault system
- 1.12. Select polygons for sedimentary rocks containing bedding measurements dipping >50 degrees
- 1.13. Report description
  - 1.13.1. Report all bedding orientations from map unit Z (to file, to screen)
  - 1.13.2. Report all figures and images associated with this area, with unit X, etc.
- 1.14. Select location points for samples with U-Pb zircon geochronologic data
- 1.15. Select all hornblende bearing plutonic rocks (requires default descriptions for rocks that do not have explicit mineralogy description but do have lithology classification)
- 1.16. Fractional analysis descriptions
  - 1.16.1. Report most abundant constituent
  - 1.16.2. Boolean—is constituent present
  - 1.16.3. Boolean—is constituent present within fraction range (e.g., 10-40%)
- 1.17. Select structural horizons equivalent to the base of the Chattanooga Shale

4 Relationship
- 1.18. Class relations
  - 1.18.1. Identify map units by hierarchical relations (e.g., all members of X; units containing X)
  - 1.18.2. Identify map units by spatial relations (e.g., overlies X)
  - 1.18.3. Identify map units by temporal relations (e.g., younger than X)
  - 1.18.4. Identify map units by semantic relations (e.g., correlated with X)
- 1.19. Class-Description relations
  - 1.19.1. Select map units that have arkosic sandstone as a protolith
- 1.20. Description relations
  - 1.20.1. Select stretching lineation measurements that have associated foliation measurements for the foliation that contains the lineation

5 Map symbolization and cartography
- 1.21. Reclassify map units based on description (e.g., dominant lithology, stratigraphic age)
- 1.22. Reclassify map units based on relations with occurrences (e.g., containing planar measurements with certain dip ranges or magnitudes)

6 Standard spatial queries
- 1.23. Buffer spatial occurrence
- 1.24. Select occurrences that intersect selected occurrences (e.g., lines, polygons, etc.)
- 1.25. Select points within polygon
- 1.26. Select polygons adjacent to selected polygons
- 1.27. Select contacts that are truncated at faults
- 1.28. Report the total length of selected lines, and their average length
- 1.29. Select polygons with area >50 hectares

Table 2. Descriptions, classifications, and relationships.

Descriptions. Each item represents a geologic feature that must be describable in the database system to address some aspect of one of the queries. The list is not comprehensive, but representative. Descriptions can be thought of as sentences that use a specific vocabulary to describe particular geologic features.

2.1. Geologic age
  2.1.1. Older/younger than
  2.1.2. Intrusive age
  2.1.3. Metamorphic age
  2.1.4. Depositional age
  2.1.5. Surface age
  2.1.6. Fault movement age
  2.1.7. Landslide movement age
  2.1.8. Stratigraphic age
  2.1.9. Absolute age
2.2. Contact
  2.2.1. Thickness
  2.2.2. Geometry
  2.2.3. Exposure
  2.2.4. Confidence
  2.2.5. Relations to map units
2.3. Alteration
2.4. Chemical composition
2.5. Default (standard) description
  2.5.1. Mineralogy for lithology class
  2.5.2. Chemistry for lithology class
2.6. Lithology component in aggregated unit
2.7. Physical property
  2.7.1. Density
  2.7.2. Magnetic susceptibility
  2.7.3. Magnetization direction
2.8. Lithology description
  2.8.1. Grain size
  2.8.2. Mineral composition
  2.8.3. Sorting
  2.8.4. Color
  2.8.5. Fabric
  2.8.6. Sedimentary structure
  2.8.7. Biologic structure
2.9. Rock unit
  2.9.1. Weathering color
  2.9.2. Surface morphology
  2.9.3. Deposit genetic structures
  2.9.4. Tectonic setting
  2.9.5. Constituent lithology
  2.9.6. Thickness
  2.9.7. Diagnostic features
2.10. Continuous variation (isopleth) description
  2.10.1. Magnetic field
  2.10.2. Gravity field
  2.10.3. Stratigraphic unit thickness
  2.10.4. Depth to horizon (top unit, fault, unconformity)
2.11. Fossils
  2.11.1. Collection location
  2.11.2. Taxonomic class
2.12. Structure orientation
  2.12.1. Planar (bedding, foliation, joints, etc.)
  2.12.2. Linear (lineation, axes, etc.)
  2.12.3. Dip magnitude
  2.12.4. Azimuth
2.13. Metadata
  2.13.1. Classification confidence
  2.13.2. Measurement accuracy
  2.13.3. Measurement precision
  2.13.4. Completeness
  2.13.5. Spatial (e.g., projections)
  2.13.6. Temporal (e.g., creation date, publication date)
  2.13.7. Historical (e.g., data processing, inferences, causal processes)

Classifications. Each item represents a classification system (kinds of) or concept that is necessary to standardize descriptions. Classifications can be thought of as scientific vocabulary. Listed below are types of vocabulary identified in the examined queries. The list is representative, not exhaustive.

- 2.14. Surface morphology (independent of map units) (e.g., sinkhole, scarp, ridge crest, terrace, plateau)
- 2.15. Lithology (fabric, mineralogy)
  - 2.15.1. Hand sample
  - 2.15.2. Outcrop
  - 2.15.3. Map unit
  - 2.15.4. Chemical classification
- 2.16. Geologic surfaces (or boundaries?)
  - 2.16.1. Depositional contacts
  - 2.16.2. Intrusive contact
  - 2.16.3. Fault
- 2.17. Rock body
  - 2.17.1. Geometry (pluton, dike, sill)
  - 2.17.2. Genetic origin
- 2.18. Alteration types
- 2.19. Geologic age; e.g.,
  - 2.19.1. Geologic time scale
  - 2.19.2. Orogenies (e.g., Nevadan, Laramide)
  - 2.19.3. Stratigraphic sequences (e.g., Sauk sequence)
  - 2.19.4. Magnetostratigraphic age
- 2.20. Formal stratigraphic units
  - 2.20.1. Lithostratigraphic
  - 2.20.2. Chronostratigraphic
  - 2.20.3. Biostratigraphic
- 2.21. Tectonic setting

Relationships. Each item is a type of relationship between geologic entities or is a significant example of a relationship between such entities. The list is representative, but not exhaustive.

- 2.22. Rocks in hanging wall of fault
- 2.23. Stratigraphic order
- 2.24. Stratigraphic relationships
  - 2.24.1. Between constituents in map unit
  - 2.24.2. Between map units
  - 2.24.3. Rank relations (contains, equivalent to, composes)
- 2.25. Class-description relationship (e.g., proportion; is 10% of class X)
- 2.26. Description-description relationship
  - 2.26.1. Fault-age
  - 2.26.2. Feature-orientation
- 2.27. Metamorphic rock-protolith
  - 2.27.1. Evolution (e.g., protolith B into rock C)
  - 2.27.2. Process (e.g., process A changed protolith B into rock C)
- 2.28. Stacking of mapping horizons (top Precambrian below top of Paleozoic below base of Quaternary)
- 2.29. Relative age (older than, younger than, overlaps older, overlaps younger, overlaps contained, overlaps covers, meets older, meets younger)

---

son with NADM 4.3, and an example of values used to represent a standard rock description. To date, documentation for four variants is posted on the DMDT Web site at <http://geology.usgs.gov/dm/steering/teams/design/>. These variants are CORDLink v5.2 (Geological Survey of Canada, Pacific), Idaho Geological Survey (IGS), Arizona Geological Survey (AGS), and NGMDB/Kentucky (USGS, KGS). Additional variants are also being documented and will be posted to the Web site as they become available.

DATA MODEL REVISION

Currently, the DMDT is working on a proposal for the next generation standard data model. This effort is based on the requirements analysis, variant documentation, and other standardization efforts, technology developments, and community data-sharing initiatives in the geoscience and computer-science communities. These include Environmental Systems Research Institute, Exploration and Mining Markup Language (XMML), Geoinformatics Initiative, Open GIS Consortium, International Standards Organization Geographic Information/ Geomatics (TC 211), and others.

The team has redefined the standard model to be a technology-neutral conceptual model (e.g., ontology), which together with a Web-based interchange format using evolving information technology (e.g., XML), is intended to allow information sharing independent of local logical and physical implementation. Complete definition of the standard will require the development of reference data sets for document conformance testing. The level of conformance at any particular agency will depend on which elements of the conceptual model are implemented and mapped into the interchange format.

To facilitate implementation of the conceptual model, the team plans to present one or more reference logical models and implementations for particular software environments. These will enable users who do not wish to design their own conformant database and interchange middleware to use off-the-shelf packages.

A working group consisting of Boyan Brodaric, Bruce Johnson, Stephen Richard, Peter Schweitzer, and Gerald Weisenfluh has been developing a draft conceptual model for review and evolution by the entire team. Figure 1 is a generalized version of the concept hierarchy framework currently under consideration. This basic framework will be fleshed out over the next several months with schema for description of the various entities in the model and definition of relationships between objects. Development of the interchange format cannot begin in earnest until the conceptual model is determined to be stable. Progress on conceptual model development will be reviewed at the next DMDT meeting, to be held during the October 2002 Geological Society of America annual meeting in Denver.

REFERENCE

Johnson, B.R., Brodaric, B., and Raines, G.L., Hastings, J.T., and Wahl, R., 1999, Digital Geologic Maps Data Model, v. 4.3: U.S. Geological Survey and Association of American State Geologists, <http://geology.usgs.gov/dm/model/Model43a.pdf>.

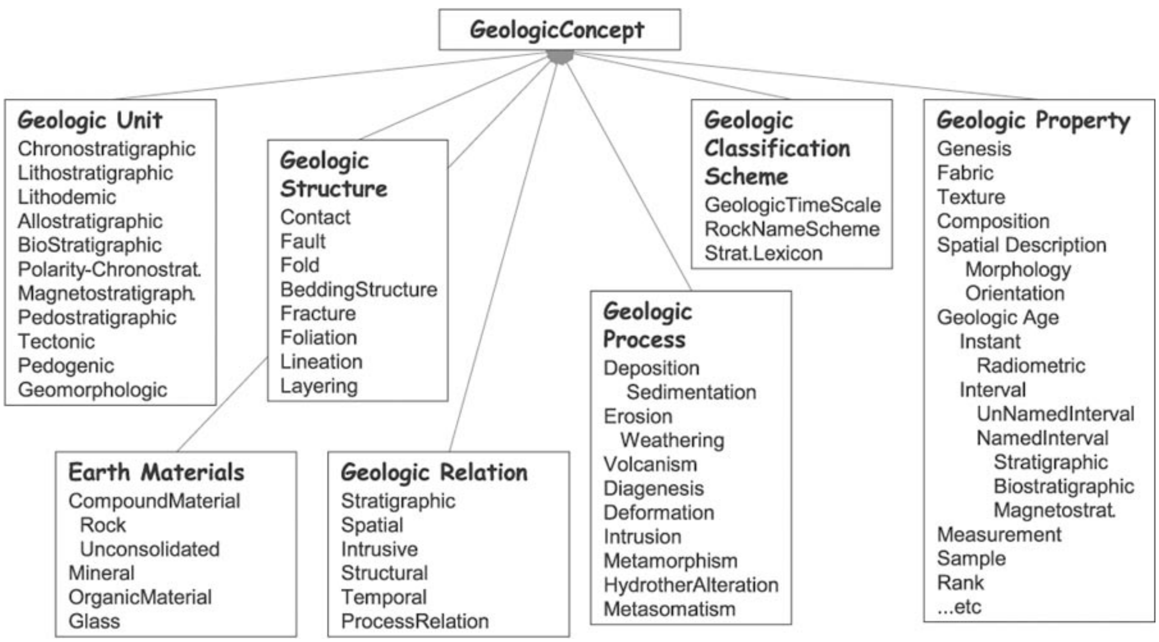


Figure 1. Draft high-level geologic concept hierarchy.

A Scalable, Digital Map Database of Bedrock Geology for Canada: A Progress Report

By Peter Davenport,[1] Eric Boisvert,[2] Marianne Quat,[3] Andy Okulitch,[4] Boyan Brodaric,[5] Stephen Colman-Sadd,[6] Larry Nolan,[7] Bert Struik,[8] Don MacIntyre,[9] Ping Tzeng,[10] David Scott,[11] Celine Gilbert,[12] Grant Abbott,[13] Andrea Bassan,[14] Murray Journeay,[15] Jodie Francis,[16] and Terry Houlahan[17]

[1]Geological Survey of Canada
3303 33rd St. NW
Calgary, AB, Canada, T2L 2A7
Telephone: (403) 292-7141
Fax: (403) 292-7071
email: pdavenpo@nrcan.gc.ca

[2]Geological Survey of Canada, eboisver@nrcan.gc.ca
[3]Geological Survey of Canada, mquat@nrcan.gc.ca
[4]Geological Survey of Canada, aokulitc@nrcan.gc.ca
[5]Geological Survey of Canada, brodaric@nrcan.gc.ca
[6]Geological Survey of Newfoundland and Labrador, scs@zeppo.geosurv.gov.nf.ca
[7]Geological Survey of Canada, lrn@zeppo.geosurv.gov.nf.ca
[8]Geological Survey of Canada, BStruik@nrcan.gc.ca
[9]Geological Survey Branch, British Columbia, Don.McIntyre@gems1.gov.bc.ca
[10]Geological Survey of Canada, ptzeng@nrcan.gc.ca
[11]Canada-Nunavut Geoscience Office, dascott@nrcan.gc.ca
[12]Canada-Nunavut Geoscience Office, cgilbert@nrcan.gc.ca
[13]Yukon Geology Program, abbottg@nrcan.gc.ca
[14]Geological Survey of Newfoundland and Labrador, aab@zeppo.geosurv.gov.nf.ca
[15]Geological Survey of Canada, mjournea@nrcan.gc.ca
[16]ESS Info, Natural Resources Canada, jfrancis@nrcan.gc.ca
[17]ESS Info, Natural Resources Canada, houlahan@nrcan.gc.ca

INTRODUCTION

Canadian geological surveys at the federal and provincial/territorial levels are collaborating to make their collective holdings of geoscience data and information accessible through the Internet, under the umbrella of the Canadian Geoscience Knowledge Network at <http://CGKN.net>. As part of this collaboration, work has started on organizing sets of existing digital geological maps (Figure 1) as a database that can be searched and reclassified consistently to prepare customized maps to meet various user needs. This project has adopted a variant of the North American Data Model (<http://www.cgq-cgc.ca/hydrolink/public/NADM_Documentation/en/nadm52en.html>) and is adapting and developing the software tools, standards, and protocols that are required to deliver bedrock-geological map data from several geological surveys in a nationally consistent form. The current partners are the geological surveys of British Columbia and Newfoundland and Labrador, the Canada-Nunavut Geoscience Program office, the Yukon Geology Program, and Natural Resources Canada (Geological Survey of Canada and Earth Sciences Sector).

PROJECT OBJECTIVE

The ultimate goal of the project is to produce from a collection of the most recent geological maps available for the various parts of Canada a composite geological map "layer" in a database from which elements can

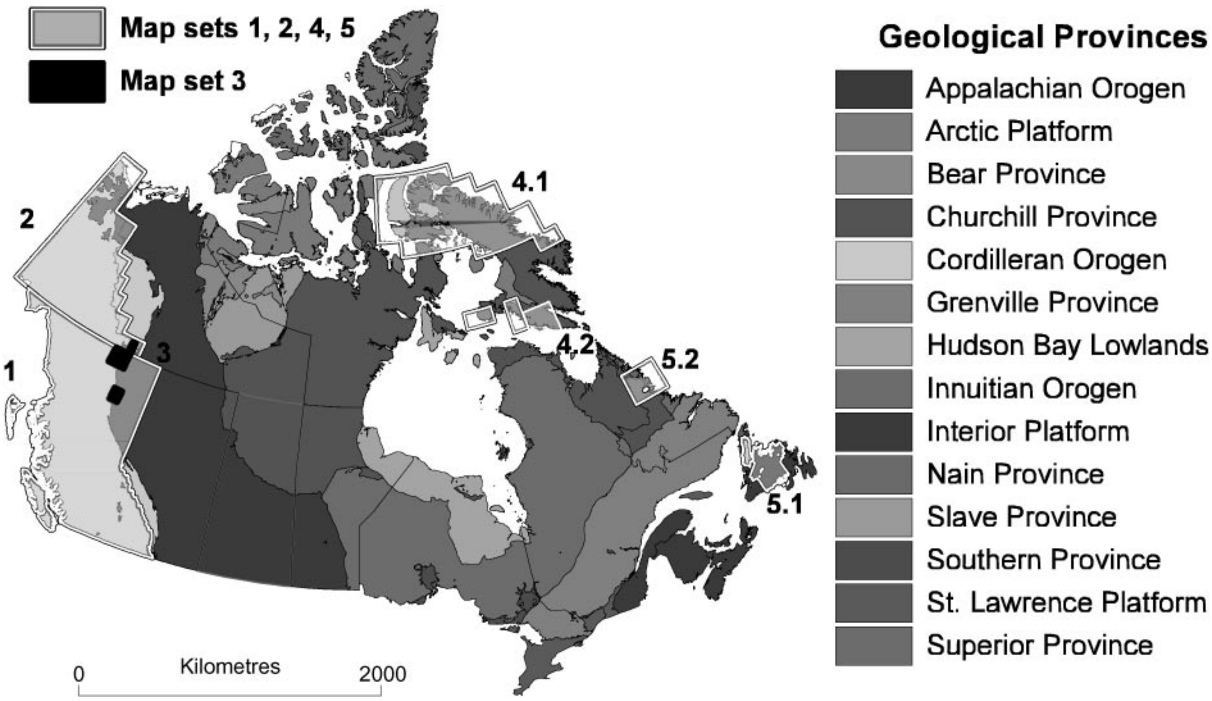


Figure 1. Location and geological setting of map sets used in project to develop a bedrock geology map database for Canada.

be selected by area, age, and lithological properties to produce “new” maps to meet specific user requirements. This map database also will allow generalizations of the geological information using the classifications from regional compilations at various resolutions, permitting its display at scales ranging from that of the original mapping to broad regional maps at 1:5,000,000 or less. This first phase of the project will produce a working example of a distributed geological-map database and a system to access it that will form the foundation for bedrock geology in CGKN. At first, it will have limited functionality and will certainly be incomplete, but its design will be both extensible and adaptable.

MAP SETS

Several digital maps sets from different agencies and a variety of geological domains are being used, each representing some of the “best and most current” published information for bedrock geology that is available for each region (Figure 1). Scale of mapping is mainly 1:250,000 or more detailed, from mostly provincial, territorial, and federal survey publications, but with contributions from academic and industry sources. There are some differences in the way these digital map sets were created, reflecting differing approaches adopted by individual agencies. These differences have to be accommodated in building the national geological database, and their salient features are described below.

Set 1. This map set comprises 88 map tiles compiled by the Geological Survey of British Columbia (BCGS) at a scale of 1:250,000 as a base for a mineral potential assessment of the entire province. Map sheets have been edge-matched, and a provincewide legend applied to the whole set.

Set 2. The Yukon map set includes approximately 90 maps whose original line work was digitized and synthesized into a digital compilation that was released on CD-ROM for the whole territory by Gordey and Makepeace (1999). In contrast to set 1, map features remain linked to their original map sources, and scale of mapping ranged from 1:250,000 (or less in a few cases) to 1:50,000. The compilers concentrated on fitting individual source legends into a regional legend for the whole area, and adjusting some map features across map boundaries. The varying resolution of the source maps resulted in variable levels of detail in the final compilation. Little original information has been lost in the compilation process, and the compilers’ changes are identified as such. Together, map sets 1 and 2 cover the entire Canadian Cordillera and the western margin of the Interior Platform.

Set 3. This map set consists of nine 1:50,000 digital geological maps and one 1:250,000 map that was digitized and released by the Geological Survey of Canada in GIS (ArcView) format as part of the ongoing Central Forelands NATMAP project (http://www.nrcan.gc.ca/gsc/calgary/majorprojects/central_e.html). The 1:50,000 maps are the most recent detailed mapping for this part of northern British Columbia and Yukon and will be used to update the BCGS and Yukon map sets.

Sets 4.1 and 4.2. There are two map sets for Nunavut, all for Baffin Island, which lies within the Churchill Province and Arctic Platform. Set 4.1 is being digitized by the Nunavut Geoscience Office. These 36 existing 1:250,000 paper maps were published by the Geological Survey of Canada as part of a project to create a digital geological map for the new territory. In addition, seven contiguous 1:100,000 maps (set 4.2) have been prepared by ESS Info (the Earth Science Sector information agency for the Geological Survey of Canada) in ArcInfo format in a pilot project to create a data warehouse of its published geological maps.

Sets 5.1 and 5.2. The final two map sets are from the geological-map database system Geolegend (Colman-Sadd and others, 1997) developed and maintained by the Geological Survey of Newfoundland and Labrador (Colman-Sadd and Crisby-Whittle, 2001). These map sets were digitized from the original source maps at their original scales. Set 5.1 comprises 82 maps for the Island of Newfoundland at scales ranging from 1:15,000 to 1:250,000. These lie within the Appalachian Orogen. Set 5.2 comprises 15 maps for Labrador at scales ranging from 1:50,000 to 1:500,000, covering parts of the Churchill and Nain Geological Provinces. For any particular area, the most detailed and recent information has been used, and the goal is to have provincewide coverage. The database also is maintained to reflect the results of new mapping and other research (such as radiometric dating) as they are published.

THE DATA MODEL

The variant of the North American Data Model (NADM 5.2) used in the CORDLink digital library <http: //cordlink.gsc.nrcan.gc.ca> is the starting point for this project. Minor modifications have been made as the project progressed to accommodate specific requirements. Ways are being explored to use the COA tree (Compound Object Archive, perhaps more easily understood as Concept tree Archive) and attribute tables together to organize map-unit descriptors in several logical and linkable hierarchies. A combination of COA architecture and logical science language will be the key to producing customized geological maps that can be displayed not only using a "conventional" legend format (i.e., map units ordered by age and identified as lithostratigraphic/ lithodemic entities), but also other classifications based on different combinations of concepts (e.g., composition, age, genesis, tectonic association, etc.). Initial databases were assembled in Microsoft Access, which allowed easier testing of ideas on table design. As the design has stabilized the databases have being moved to Oracle 8.17 implementations.

Geomatter II (Boisvert and others, 2001) has been used as a graphical user interface for populating and organizing the NADM tables. Some tables are also loaded directly from spreadsheets, but Geomatter is used to verify the results of the bulk loading and to make any minor corrections. Geomatter II has been modified during the project as the database structure has evolved and works with both the Access and Oracle versions of NADM.

The multi-agency nature of CGKN dictates that the data and information of each participating survey is maintained locally, while the need to deliver consistent information across agencies requires the adoption of common standards and protocols for coordinating the way it is organized. This distributed nature of the project poses problems in coordinating concept trees between several databases, especially since there will be differences in the details of the information to be stored in each. A way to allow local diversity while maintaining overall consistency has been proposed (Boisvert and others, 2001) through a central database (or registry) of concepts that are identified as being nationally important. This central concept registry will mediate between clients and the distributed databases to present a nationally consistent interface. The distributed databases will be able to accommodate local needs with a minimum of external constraints and allow them to be accessed directly.

DATA ORGANIZATION FOR SCALABILITY AND INTEROPERABILITY

In tackling these issues, the Yukon map set (Gordey and Makepeace, 1999) has been used extensively for both designing the approaches to the problem and testing the results of the designs. The other maps sets have been used to test the general applicability and effectiveness of these designs, and this testing process is continuing.

Scalability (Varying Map Resolution)

The initial plan was to use a map-unit hierarchy as depicted in Table 1 to allow the aggregation of detailed source units into progressively more general groupings.

Table 1. Map-unit levels as an idealized hierarchy, and appropriate display scales for each level.

Map-unit type	Approximate display scale
Geological Province	≤ 1:25,000,000
Tectonic Terrane (Cordillera)	1:10,000,000
≅ Tectonic Zone (Appalachians)	
Tectonic Assemblage	1:1,000,000
Supergroup (Super Suite)	≥ 1:250,000
Group (Suite, Complex)	≥ 1:250,000
Formation (Lithodeme)	≥ 1:250,000
Member	≥ 1:250,000
Source-map unit (informal)	variable

There are fundamental problems with this idealized hierarchy as an entity, however, in particular between the upper part (Geological Province, Tectonic Terrane) and the lower part (the Lithostratigraphic/Lithodemic hierarchy). "Geological province" is defined as "an extensive region characterized throughout by similar geologic history or by similar structural, petrographic, or physiographic features" (Jackson, 1997). A lithostratigraphic unit is "a defined body of sedimentary, extrusive igneous, metasedimentary, or metavolcanic strata that is distinguished and delimited on the basis of lithic characteristics and stratigraphic position" (Jackson, 1997), while a lithodemic unit is "a defined body of predominantly intrusive, highly deformed, and/or metamorphosed rock, distinguished and delimited on the basis of rock characteristics" (North American Commission on Stratigraphic Nomeclature, 1980). Thus, Geological Provinces are defined and distinguished not only by the bodies of rock they contain, but also by their structural history.

Tectonic Terranes, like Geological Provinces, "are parts of the earth's crust which preserve a geological record different from those of neighbouring terranes" (Gabrielse and others, 1992), and are thus defined on more than one criterion. Tectonic Assemblages are comparable in most respects to lithostratigraphic/lithodemic units. A Tectonic Assemblage is defined as a grouping of lithostratigraphic units that is "commonly bounded by regional unconformities or by faults [and] represents a specific depositional or volcanic setting and/or response to one or more tectonic events"; "each tectonic assemblage reflects a specific tectonic and/or depositional environment regardless of its place of origin." A specific assemblage may belong to two or more Terranes that differ in their history of deformation. Some source units in the Yukon map set also have been split among two or more assemblages.

The variation in criteria used for each classification type leads inevitably to multiple inheritance problems when one attempts to impose a simple hierarchy. Instead of defining a hierarchy for the map-unit types as a COA tree, the spatial classification, classification object, and classification scheme tables are used to classify the source-map units into the various levels by individual spatial feature (polygon). Specifically, each source-map polygon must be individually assigned to each of the higher levels in the spatial classification table (as Gordey and Makepeace, 1999, had done in their compilation). For other maps, this will entail more work during loading because with a simple hierarchy, each unit at the lowest level (i.e., the original units on the source maps) has only to be related to its immediate parent to be placed correctly in the hierarchy.

The above discussion applies mainly to the stratified rocks that form the lithotectonic framework of a region. Plutonic rocks, which are typically considered apart from the stratified rocks, generally can be grouped into major pulses of magmatism. Metamorphic rocks may either be included in the lithostratigraphic units, or, in some cases where they are of uncertain affinity, they may be grouped into an "unassigned" category. Finally, postorogenic lithostratigraphic units commonly may be grouped into larger assemblages based on criteria such as regional unconformities. These groupings of plutonic, metamorphic, and postorogenic lithostratigraphic units provide ways to aggregate them into more general categories that serve as the equivalents to Tectonic Assemblages and Tectonic Terranes for the purpose of map scaling.

Test of Map-Unit Hierarchy

One problem encountered very quickly was the asymmetry of the lithostratigraphic/ lithodemic hierarchy, as illustrated for the northeastern Yukon in Table 2. The Supergroup/Supersuite and Group/Suite/Complex levels are seldom used, and in fact for the Yukon map set, almost 60% of the units were informal. Furthermore, these informal units vary in their apparent rank equivalency from Group to Formation to Member. To overcome this problem, for the Yukon compilation Gordey and Makepeace (1999) developed a set of "regional compilation" units[1] and subunits that they used to better group the source-map units for display at different resolutions, as illustrated for a single quadrangle in Figure 2. In Figure 2a, all of the 58 source-map units were informal, with no name assigned. In the database, for these unnamed units, a provisional name has been created by concatenating its label with the name of the regional compilation unit to which they have been assigned by Gordey and Makepeace (1999). In Figure 2b, some of the units are represented by small polygons even at the Tectonic Assemblage and Terrane levels, and GIS functions will be required to dissolve these to allow a satisfactory display at small scales.

Interoperability at the Map-Unit Level

Legend entries for units on bedrock geology maps are almost always characterized by a chronostratigraphic age (or age range) and a lithological description. Together these provide the common elements for associating related units from different maps. The way both age and lithology are described is quite variable, however, and usage is seldom explicitly defined, making it difficult to correlate map units from one map source to another from legend information alone. A less rigorous approach is proposed to define "related" map units that exhibit varying degrees of similarity based on these characteristics (cf.

[1]A regional compilation unit is an informal map unit used in a geological compilation to group formal and informal lithostratigraphic and lithodemic map units from original maps by several authors.

Table 2. Hierarchy developed for map units in the portion of the Yukon map set northeast of the Tintina Fault. "Source units" are original map units that are uniquely identified by their original label and a source-map ID. "Regional unit" and "Regional subunit" are groupings of source units developed by Gordey and Makepeace (1999).

COA_Name (map-unit type)	Number of classes
Geological Province	2
Tectonic Terrane	17
Tectonic Assemblage	42
"Regional unit"	99
(Supergroup, Supersuite)	2
"Regional subunit"	159
Group, Suite, Complex	43
Formation, Lithodeme (+ informal equivalents)	161
Member (+ informal equivalents)	21
"Source unit"	2082

lithodemic units). Ultimately, the user will need to review the original map-unit descriptions to decide whether the units identified as related do indeed have the association needed for the task at hand.

The chronostratigraphic interval assigned to a map unit may be from one of a number of regional systems, and the particular system used in map legends is rarely, if ever, specifically referenced. Despite this uncertainty, map-unit ages can be reconciled in a general way by reference to their currently accepted absolute age ranges. For this purpose in Canada, the geologic time scale compiled and updated periodically by Okulitch (2002) is being employed.

Lithological nomenclature presents similar, but more complex challenges for correlating map units from different sources. There are two main problems. Firstly, rock names are based on one or more of the following properties: genesis, composition, texture, fabric and degree of consolidation or induration; that is, rock names are "multidimensional." Secondly, several common rock names imply different rock properties to different geological communities, and as in the case of chronostratigraphy, usage of these names is rarely defined explicitly. For both these reasons, rock names themselves do not provide a reliable basis for either querying or defining relationships among map units on geological maps from different sources.

The issue of the multidimensionality of rock names has been addressed by attempting to break down each rock name into its implicit properties, building on the proposals of Weisenfluh (2001) and Struik and others (2002). A single rock name such as "siltstone" implies a *genesis* (a clastic, sedimentary rock), a *texture* (a sorted rock composed of silt- and clay-sized grains), and an *indurated* material. A "shale" likewise implies a similar *genesis*, *texture,* and *degree of induration*, but also a *fabric,* as a shale is implicitly a rock with planar laminations that impart fissility. In some cases, additional information about rock properties is provided by qualifiers; for example, "marine siltstone" provides more information about the rock's genesis, "foliated sandstone" provides information about fabric that is not implicit in the rock name itself.

Classifications for Lithological Characteristics

Initially, each of the four characteristics—genesis, composition, texture, and fabric—was considered as a single, simple theme. For each, a single hierarchical classification was attempted to allow the rock name and associated qualifiers to be indexed with a degree of precision appropriate to the information in the map legend. It became apparent that several of these themes are in fact composite, and these have been further broken down so that the classifications are both independent and simpler. In most cases, they are shallow hierarchies (2-3 levels) to allow characterization with different levels of precision and to allow searching at various levels of generalization. In addition, "mechanical" properties such as degree of induration and parting characteristics have been added to the set of parameters used for classification. In some cases, a rock name may imply more than one genesis, composition, fabric, etc.; multiple values are allowed as required. The purpose of these rock-property classifications is to provide an effective mechanism for searching inconsistent and loosely defined information, to be used as sets of controlled keywords. Furthermore, an attempt has been made to make the terms generic and descriptive, avoiding specialized jargon as much as possible.

Mechanical Properties

Degree of *induration* is a primary criterion used to distinguish "bedrock" geology from "surficial" deposits. The Science Language Technical Team of NADM is working on a definition to separate "consolidated" from "unconsolidated" for sedimentary rocks (Matti, 2002). For this project, all igneous and metamorphic rocks are classified as consolidated, and a simple, qualitative, threefold classification is used for sedimentary rocks: (1) unconsolidated, (2) poorly consolidated, and (3) consolidated (Table 3).

The way a rock *parts* is quite commonly described in map legends, such as rock's fissility or blockiness, as well as more specific jointing characteristics. For this preliminary scheme, a simple set of categories has been set up that distinguishes partings along one set of planar surfaces from rocks with two or more sets, and lithologies that are explicitly stated to be structureless.

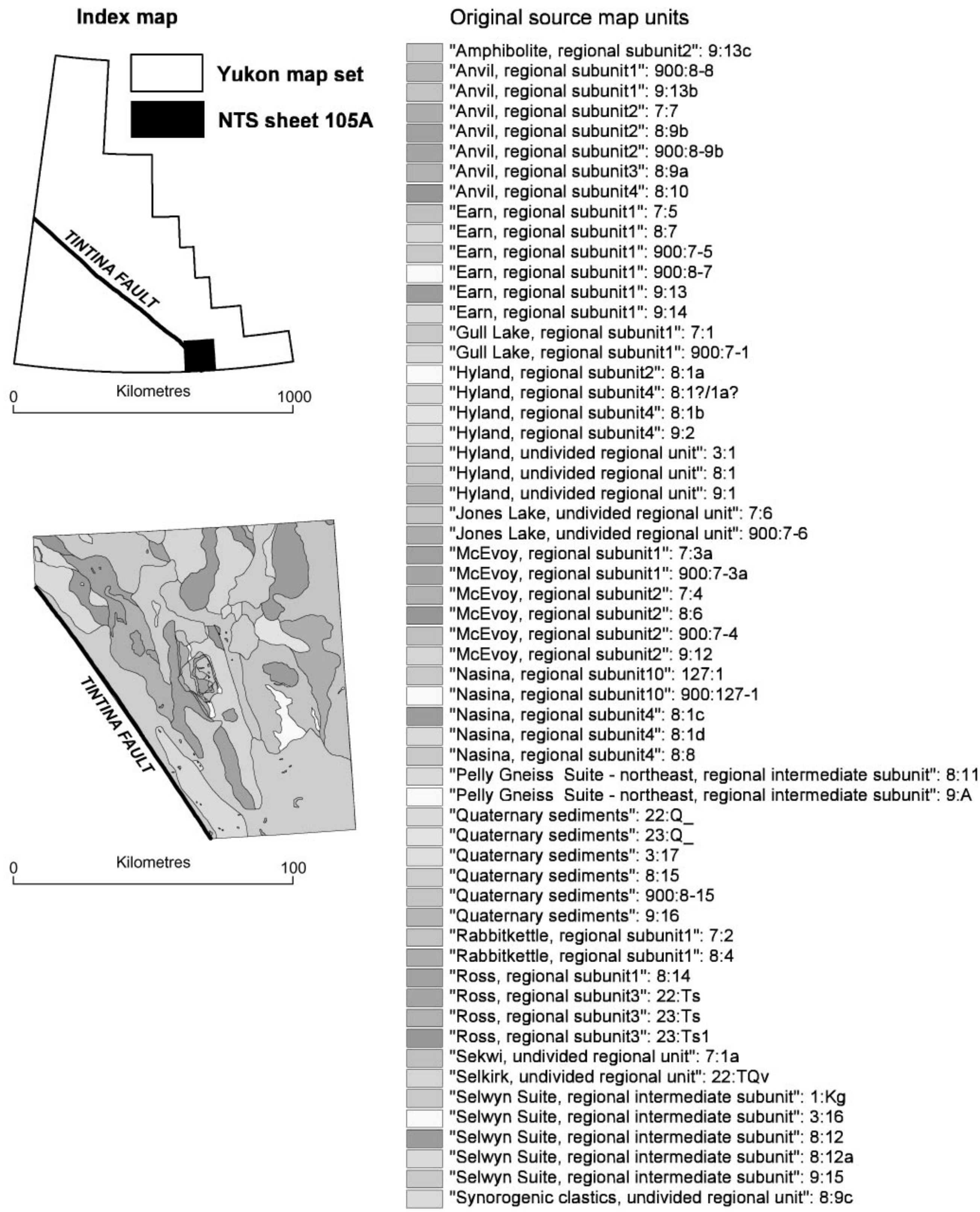


Figure 2a. Examples of changing map resolution based on map-unit level for NTS sheet 105A, NE of Tintina Fault. Original units from source maps, and index map (see Figure 1, map set 2 for location).

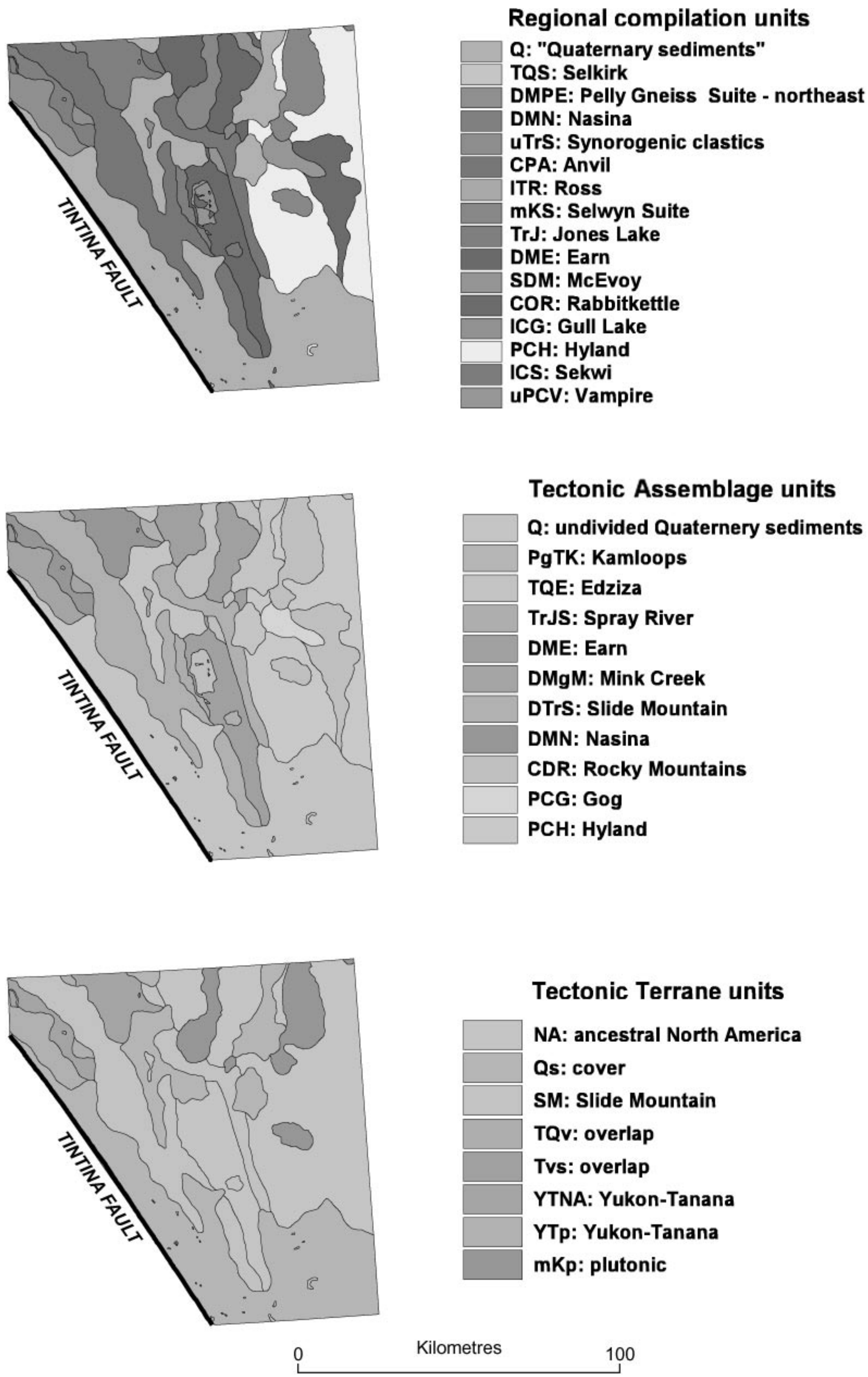


Figure 2b. Examples of changing map resolution based on map-unit level for NTS sheet 105A, NE of Tintina Fault. Top to bottom, the 58 source-map units have been reclassified into 16 regional compilation units, 11 Tectonic Assemblages, and 8 Tectonic Terranes.

Table 3. Classification of mechanical properties as implied by rock names used to describe map units. This scheme is extensible based on the properties of actual map legends.

Property	Category	Subcategory
Induration		
	1 unconsolidated	
	2 poorly consolidated	
	3 consolidated	
Weathering		
	1 resistant	
	2 recessive	
Parting		
	1 multiplanar	
		1.1 columnar
		1.2 blocky
	2 fissile (single plane)	
		2.1 platy
		2.2 flaggy
		2.3 slaty

Resistance to weathering is another mechanical property that commonly is used in map-unit descriptions. Its performance as a standard rock-property index has yet to be tested.

Composition

The information about the composition of map units is almost always descriptive and qualitative, so these attributes are reflected in the classification proposed here. For composition, a simple two-level system is suggested (Table 4). The primary grouping is based mainly on the main anion type for each lithology: silicate, oxide (nonsilicate), carbonate, sulphate, sulphide, halide, phosphate, nitrate, and borate, as well as a native element class and a carbonaceous class for rocks that are predominantly hydrocarbon. The second level is based in a general way on the dominant cation or cation group. For silicates, the terms high silica (felsic or acidic), intermediate silica (intermediate), low silica (mafic or basic), and very low silica (ultramafic or ultrabasic) are usually applied to igneous rocks, but they can in some cases be applied in a descriptive way to sedimentary and metamorphic rocks if the rock name contains sufficient information (e.g., an orthoquartzite would be classified as "silicate, high silica"). The numeric ranges for silica content are only an indication of the typical values for each category and should not be used for quantitative modeling.

As assemblages of earth materials, rocks commonly comprise two or more of the first-level compositional groupings (e.g., a calcareous quartz sandstone is predominantly a high-silica silicate rock, but has a lesser, but noteworthy, carbonate component). Thus, a rock can be assigned to one dominant and one or more subordinate compositional categories, all based on a single, simple, qualitative classification.

Table 4. Simple two-level classification of composition as implied by rock names used to describe map units. This scheme is extensible based on the properties of actual map legends.

Composition category	Subcategory
1 Silicate	
	1.1 high silica > 65% SiO_2
	1.2 intermediate silica 53-65% SiO_2
	1.3 low silica 44-53% SiO_2
	1.4 very low silica < 44% SiO_2
2 Oxide (nonsilicate)	
	2.1 ferruginous
	2.2 manganiferous
3 Carbonate	
	3.1 calcic
	3.2 magnesian
	3.3 barium
	3.4 iron
4 Sulphate	(subcategories by major cation)
5 Sulphide	(subcategories by major cation)
6 Halide	(subcategories by major cation)
7 Phosphate	(subcategories by major cation)
8 Nitrate	(subcategories by major cation)
9 Borate	(subcategories by major cation)
10 Native element	(subcategories by element/ polymorph)
11 Carbonaceous	

Genesis

Initially, a single classification for the "genesis" of each rock was attempted, but it became apparent that two main themes, genetic process and site of formation, were commonly used in genetic classifications. These two themes are therefore used as separate classifiers named "Genetic Process" and "Environment of Formation."

Genetic Process: The primary level of subdivision for this classifier (Table 5) recognizes the traditional categories Igneous, Sedimentary, and Metamorphic, although the term "metamorphic" is used in its most general sense to mean a protolith that has undergone a mineralogical and/or compositional change, including metasomatism and pedogenesis. This sense follows the preliminary approach of the metamorphic subgroup of the Science Language Technical Team of NADM (Richard, 2002). The second level allows further description of the process in a very general way. The third level has been developed only partially, but other equivalent categories may be added to accommodate the lithological information to be classified in actual map legends.

Table 5. Simple three-level classification of genetic process as implied by rock names used to describe map units. This scheme is extensible based on the properties of actual map legends.

Category	Subcategory	Sub-subcategory
1 Igneous		
	1.1 explosive	
	1.2 passive	
2 Sedimentary		
	2.1 clastic	
	2.2 chemical precipitation	
		2.2.1 evaporitic
		2.2.2 nonevaporitic
	2.3 biogenic	
3 Metamorphic		
	3.1 dynamic (high strain)	
	3.2 regional (dynamothermal)	
	3.3 contact	
	3.4 metasomatic	
		3.4.1 hydrothermal
		3.4.2 deuteric
		3.4.3 pyrometaso-matic
	3.5 pedogenic	
	3.6 impact	

Environment of Formation: The other information about a rock's genesis that is commonly implicit is its place of formation relative to the upper surface of the crust. This information is indexed using the three-level classification in Table 6. The first level distinguishes rocks formed on or above the surface of the crust—"supra-crustal"—from those formed below—"crust" or "mantle" (it could also be extended to other astronomical bodies using a level above this). The second and third levels provide more detail where this is available. As in other classifications, the third level is not completely developed at present.

Rock names may imply a genetic history rather than a single genesis, so a rock name may be classified against more than one genetic process, each of which may have an associated environment of formation. Thus, a slate would be classified as both "sedimentary, clastic" and "metamorphic, dynamothermal" in terms of genetic process, and "supracrustal, subaqueous" and "crust" as the environments of formation associated with each respective genetic process. No attempt has been made to capture the order of the genetic events.

Table 6. Simple three-level classification of environment of formation as implied by rock names used to describe map units. This scheme is extensible based on the properties of actual map legends.

Category	Subcategory	Sub-subcategory
1 Supracrustal		
	1.1 subaerial	
	1.2 subglacial	
	1.3 subaqueous	
		1.3.1 marine
		1.3.2 freshwater
		1.3.3 intertidal
2 Crust		
	2.1 shallow	
	2.2 deep	
3 Mantle		

Physical Properties: Texture, Fabric, and Structure

Initially two classifications, one based on texture and a second on fabric were tried until it was realized that both are multidimensional concepts (especially texture), and also somewhat overlapping in the way they are used. For example, texture is defined by Jackson (1997) as "the general physical appearance or character of a rock, including the geometric aspects of, and the mutual relationships among, its component particles or crystals, e.g. the size, shape, and arrangement of the constituent elements of a sedimentary rock, or the crystallinity, granularity, and fabric of the constituent elements of an igneous rock." The definition of fabric for deformed rocks from the same source includes textural properties as well as the orientation of their constituent physical elements. Even if the properties to be included under each heading are decided, the concepts of both texture and fabric remain multidimensional. Structure as a theme presents similar problems. For this reason, rather than attempting to construct two or three hierarchies, two sets of categorical classes have been established: Texture and Physical Form. Texture includes five classifications based on properties of the particles that constitute a rock (Table 7), and Physical

Table 7. Five categories for texture as implied by rock names used to describe map units. This scheme is extensible based on the properties of actual map legends.

Textural Property	Category	Subcategory	Sub-subcategory
Grain intergrowth			
	1 crystalline		
	2 granular		
Grain-size variability			
	1 homogeneous		
	2 heterogeneous		
	3 gradational		
Grain size			
	1 size class 1 <0.05 mm		
	2 size class 2 0.05-0.2mm		
	3 size class 3 >0.2 mm		
		3.1 size class 3.1 0.2-2 mm	
		3.2 size class 3.2 2-4 mm	
		3.3 size class 3.3 4-64 mm	
			3.3.1 size class 3.3.1 4-16 mm
			3.3.2 size class 3.3.2 16-64 mm
		3.4 size class 3.4 64-256 mm	
		3.5 size class 3.5 >256 mm	
Grain morphology			
	1 rounded		
	2 subrounded		
	3 angular		
	4 irregular		
Large particle to matrix proportions (heterogeneous rocks only)	1 matrix dominant		
	2 matrix subordinate		

Form includes two classifications based on the spatial orientation of a rock’s constituent physical elements (Tables 8a, b).

The classification for texture is:

grain intergrowth of a rock’s component particles; crystalline is used for rocks where intergrowth is complete or nearly so; granular for rocks whose grains are not interlocking and, therefore, potentially have pore space.

grain-size variability; homogeneous, heterogeneous, and gradational.

grain size; three general categories are proposed: rocks whose grains are all microscopic (size class 1), rocks containing grains that are just discernable to the naked eye or by using a hand lens (size class 2), and rocks that contain clearly visible grains (size class 3). Size class 3 is further subdivided to accommodate size ranges commonly implicit in rock names or qualifiers. Approximate size ranges in millimeters are given for each class (Table 7), together with commonly used equivalent terms for sedimentary and igneous/metamorphic rocks. The numeric ranges are only an indication of the typical values for each category and should not be used for quantitative modeling. The use of descriptive terms such as fine, medium, and coarse is avoided because they have been defined in a several conflicting ways. Many rocks exhibit a range in grain size that spans even these general size classes (i.e., heterogeneous or gradational rocks), so two grain-size fields are used; one for the matrix (or finer particles), and a second for the coarser grains or clasts. Grain size for homogeneous rocks is entered in the “matrix” field. Clearly this scheme cannot accommodate the full range of particle sizes found in some rocks, but it seems adequate for the qualitative information contained in most map unit descriptions.

grain morphology; based on the relative degree of rounding of grains or crystals.

proportions of coarser to finer elements; this category is for rocks with heterogeneous grain-size distributions, where the rocks can be classified as being either matrix/groundmass dominated, or matrix/groundmass subordinate.

Table 8a. Classification of physical form—external habit as implied by rock names used to describe map units. This scheme is extensible based on the properties of actual map legends.

Category	Subcategory
1 Tabular	
	1.1 thin
	1.2 medium
	1.3 thick
	1.4 very thick
2 Lenticular	
3 Equant	

Table 8b. Classification of physical features—internal structures as implied by rock names used to describe map units. This scheme is extensible based on the properties of actual map legends.

Category	Subcategory	Sub-subcategory
1 Surfaces		
	1.1 laminated	
	1.2 cross-stratified	
	1.3 wavy	
	1.4 mudcracked	
	1.5 foliated	
		1.5.1 continuous (schistose)
		1.5.2 discontinuous (gneissose)
	1.6 flaser	
	1.7 sheared	
2 Volumes		
	2.1 amygdaloidal	
	2.2 boudinaged	
	2.3 concretionary	
	2.4 miarolitic	
	2.5 nodular	
	2.6 orbicular	
	2.7 vesicular	
	2.8 vuggy	

Aspects of structure and fabric are treated under physical form. External habit captures the overall form of the lithological entity in a very general way—features seen at outcrop and larger scales—whereas structures and fabrics at the outcrop and smaller scales are categorized under internal structures. External habit (Table 8a) is subdivided broadly on geometric rather than genetic criteria. Thus, tabular form includes beds, dykes, veins, and volcanic flows (except where the last are explicitly described as having a different shape such as a pillow). The other two forms recognized are lenticular (e.g., bioherms and pillows) and equant (roughly equidimensional rock bodies such as plugs, stocks, and pinnacle reefs).

Internal structures (Table 8b) are grouped under surfaces (both primary and secondary) and volumes (more equidimensional features). Some “lithology strings” may imply more than one surface type or volume type.

Summary of Classifications

Although geological principles have been used in their design, the main purpose of these classifications is to allow effective selection of map units from a geological-map database based on consistent, thematically based lithological attributes. Each thematic classification is organized as a shallow hierarchy or simple set of categories, so that keywords may be selected at a level of detail that corresponds to the level of information available. Further, because the number and nature of themes that are implicit in any particular rock name are variable, the use of multiple keyword sets organized under thematic headings allows a rock name to be indexed by as many (or as few) as appropriate. The design is preliminary and has been tested against a digital map database derived from about 60 individual source maps, using both the original map-unit descriptions, and the more general regional legends. As the map database grows, the classifications will be extended and modified as necessary to accommodate new types of lithological information as required.

RESULTS OF A TRIAL APPLICATION

Procedure of Indexing Lithological Information from the Map Legend

The classification schemes described above were developed to use as much information as possible from “typical” map-unit descriptions. These initial schemes were developed and tested on 2,082 map-unit descriptions from the Yukon Geology CD (NE of the Tintina Fault) that were taken by Gordey and Makepeace (1999) directly from the source maps. To these were added the map-unit descriptions from the 223 regional compilation units, the 41 Tectonic Assemblages, and the 19 Tectonic Terranes to determine how the lithological indexes would work with these more generalized descriptions.

The first step was to parse the unit descriptions from the map legends to identify a set of specific rock names for each map unit (here termed the “root lithology”), together with any associated qualifiers that may influence the way each instance of that rock name is classified, their relative proportions (Table 9), and rock colour(s) if given. Colour and proportion were captured only at the source-unit and regional-unit levels, because there was very little

Table 9. Qualitative classification of the proportions of lithologies in map units from legend descriptions of map units.

Proportion label	Description
All:	map unit contains only a single lithology
Major:	lithologies explicitly described as the main component, or implied to be major where the all other components are stated to be minor
Significant:	lithologies are named in the main part of the text without qualification as to their importance
Minor:	lithologies explicitly described as minor (or rare) components of the map unit

information for these attributes at the Tectonic Terrane and Assemblage levels. In this initial trial, information about weathering characteristics was not captured, although it is common (especially weathering colours). The authors’ rock names are retained as far as possible (e.g., dolomite was not renamed dolostone), although the order of the rock names and their associated qualifiers was standardized, so that in general the rock name is first (a “root” lithology), followed by modifiers for composition, texture, fabric, and genesis, in that order, to create a “lithology string.” The reason for this ordering is to facilitate their subsequent classification as efficiently and consistently as possible, not to try to develop a set of “preferred” rock names (or qualifiers).

In addition to the qualifiers explicitly expressed in the map legend, in some cases an additional qualifier may be inferred from the context of the map unit. For example, the rock name “quartzite” may refer to either a sedimentary or a metamorphic rock, and it is commonly possible to determine which applies from the associated lithologies within the same map unit. The aim is to make each “rock name and associated qualifier” contain all the information that will be needed to classify it as precisely as possible against the generic classifications without further reference to the map-unit description. This speeds up the classification process and ensures that similar strings from different sources are classified identically.

Extracting lithological information from the legend in this way is the slowest step of the process and requires geological knowledge and experience. It is critical to the success of the whole exercise and probably cannot be automated. Although the classifications will be refined over time, if this step is done well, it should not have to be repeated.

When the list of rock names and associated qualifiers was complete for a map set, a list of all the unique combinations was compiled, spelling errors corrected, and “trivial synonyms” caused by variations in spelling or order of qualifiers removed. This edited list was then classified against the thematic indexes, and these classifications related back to the map units. The legend of the next map set was taken through the same process, and the list of unique rock names and qualifiers compared to the first, edited as before, and any new strings added to the set to build up a master list. This list grew rapidly at first, but more slowly as more map sets were processed. This list also allowed an analysis of historical usage of rock names that can be used to suggest a better approach to the process of describing map units for geological map databases in the future.

Results of Legend Parsing, Yukon Test Data Set

From the 2,311 map-unit descriptions, 1,947 unique “lithology strings” were parsed out. These consist of about 160 “root names” (Table 10), and combinations of some 640 qualifying words or phrases (each “root name” having between 0 and 8 distinct qualifiers). The “root names” for lithology can be subdivided into three broad categories; 125 common names that are quite specific (e.g., andesite, metaquartzite), 28 more general “rock class” names (e.g., igneous rocks, organic deposits), and 8 mineral names that are used to label a significant, monomineralic component of a map unit. One approach to standardizing the diversity of language would be to map the actual usage to a “standard” rock-classification scheme. This task would be very time consuming and would involve many assumptions. Furthermore, it would have to be done both for the “root names” and their qualifiers. Two further drawbacks are the current lack of a “standard” for rock nomenclature that is well accepted by Canadian geologists, and the multidimensionality that persists in most of the draft rock-naming schemes that have been proposed or are in preparation. This feature of rock names will continue to hamper the construction of simple queries for map units based on lithology and the production of simple, derivative maps based on a single theme.

The colour of each lithology also was captured for the legends of the source maps and regional compilation units (colour is little used as a descriptor at the Tectonic Assemblage and Terrane levels). The frequency of association of colours with lithologies decreased from 28% for the source map units to 24% and 17% for the compilers’ regional subunits and units, respectively. On closer examination, there are 321 unique colour combinations, ranging from single, simple names (e.g., black, red, white), to those with little specific meaning (e.g., dark or varicoloured), to multi-hued assemblages to describe a single lithology (e.g., light grey to black, greenish grey to

Table 10. "Root" rock names extracted from 2,311 NE Yukon map-unit descriptions.

Common rock names				
agglomerate	dolomite	leucogranite	nepheline-syenite	sandur
amphibolite	dolosiltite	lignite	orthoconglomerate	schist
andesite	dolosiltstone	limestone	orthogneiss	serpentinite
aplite	dolostone	marble	orthoquartzite	shale
arenite	dunite	marl	packstone	silt
argillite	eclogite	marlstone	paraconglomerate	siltrock
arkose	felsite	metachert	paragneiss	siltstone
ash	gabbro	metaconglomerate	pegmatite	skarn
basalt	gneiss	metadiorite	pelite	spiculite
bentonite	grainstone	metagabbro	peridotite	subgreywacke
biocalcarenite	granite	metagreywacke	phyllite	syenite
blastomylonite	granodiorite	metagrit	phyllonite	syenodiorite
calcarenite	gravel	metaporphyry	porcellanite	tholeiite
calcirudite	greenstone	metaquartzite	porphyry	till
chert	greywacke	metarhyolite	protomylonite	tillite
clay	grit	metasandstone	psammite	tinguaite
clayrock	harzburgite	metasiltstone	pyroxenite	trachyandesite
claystone	hornblendite	metatuff	quartz-diorite	trachybasalt
clinopyroxenite	hornfels	monzodiorite	quartzite	trachyte
coal	ignimbrite	monzogranite	quartz-monzonite	tufa
conglomerate	iron-formation	monzonite	quartz-syenite	tuff
dacite	ironstone	moraine	rhyodacite	turbidite
diabase	jaspilite	mudrock	rhyolite	ultramylonite
diamictite	latite	mudstone	sand	wacke
diatomite	leucogabbro	mylonite	sandstone	wackestone
diorite				
Mineral names				
anhydrite				
barite				
chalcedony				
gypsum				
hematite				
limonite				
pyrite				
quartz				
Rock class names				
alluvium	metasedimentary rocks			
breccia	metavolcanic rocks			
calc-silicate rocks	organic deposits			
carbonate rocks	phosphate rocks			
clastic rocks	plutonic rocks			
colluvium	pyroclastic rocks			
glacial drift	quartz-carbonate rocks			
igneous rocks	quartzose rocks			
intrusive rocks	sedimentary rocks			
laminate	sediments			
lime-silicate rocks	siliciclastic laminates			
megabreccia	ultramafic rocks			
metacarbonate rocks	volcanic rocks			
metamorphic rocks	volcaniclastic rocks			

turquoise). In short, the descriptions of colours as used in this test map set are too varied to organize in a single database field, although in retrospect some could have been captured in the "lithology string" as they provide information on genesis or composition (e.g., red sandstone, black shale). Weathering colours are more frequently described than the colour of fresh surfaces but suffer from similar inconsistencies.

Few map units at either the source or regional map-unit levels consist of a single lithology (5% and 3%, respectively); most include two or more distinct lithologies. Very few source-unit descriptions contain any quantitative information on the relative proportions of each rock type within the map unit as a whole. For this reason, a qualitative measure of proportions (Table 9) was used, and the results are summarized in Figure 3. The "significant" category really means no information could be gleaned about the proportion of a rock type in a map unit from the legend; this was the case for about 70% of the rock types named in map units at both levels. At the source-map level, a greater number of map units contain, or are dominated by, a single lithology than at the regional level, and the frequency of minor rock types is less. These changes are to be expected in the process of grouping source units into regional compilation units.

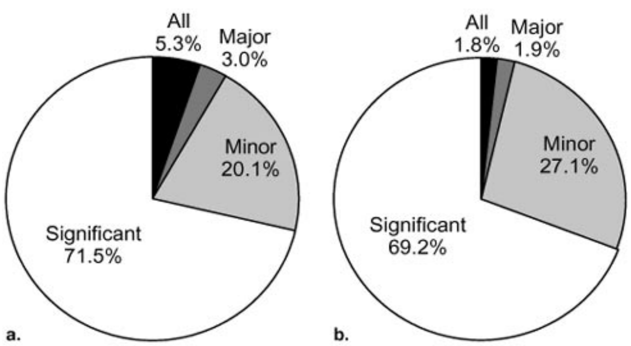


Figure 3. Frequency of occurrence of rock-type proportion categories (see Table 2). a. source-map descriptions. b. regional-compilation-unit descriptions.

Analysis of the Generic Classification Schemes

To be effective as criteria for querying map units on their lithological characteristics, each classification scheme should address the following:

- they should apply to a significant number of "lithology strings"
- the majority of these "lithology strings" should be classifiable on several of the schemes
- there should be a reasonably even distribution of values for any classification if it is to discriminate among the various "lithology strings," and
- the classification schemes should be implicit in the "root lithologies."

The number of classifiers associated with particular "lithology strings" is illustrated in Figure 4. The mode is 9 classifiers per unique "lithology string," and the number ranges from 2 to 15. An analysis of each group of classifications by theme shows that some are more effective discriminators than others.

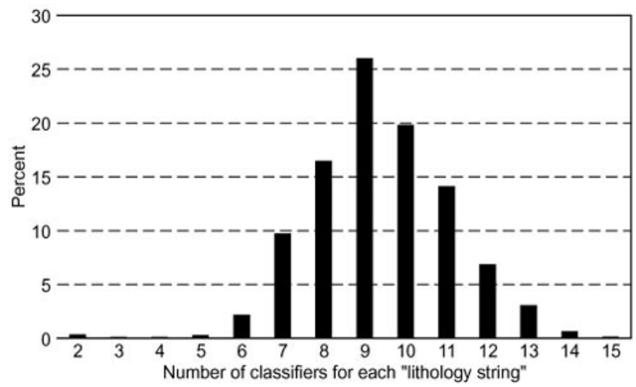


Figure 4. Histogram showing the range and frequency of the number of generic classifications implicit in each unique "lithology string."

The frequencies with which composition can be applied to the test set of lithologies is illustrated in Figure 5. A composition could be inferred for more than 99%, and there is a reasonably good breakdown among the 14 categories represented. The dominance of the undivided "silicate" category (32.4% of lithologies, Figure 5) is likely due to the predominance of terrigenous sedimentary rocks in this particular map set. Subordinate composition is associated with only 17% of lithologies, but 12 categories are represented.

The potential of genetic process as a discriminator is illustrated in Figure 6. Again, more than 99% of the test "lithology strings" could be assigned at least one genetic process, and about 17% a second process. Clastic (sedimentary) was the most common category, due to the particular map set used for this exercise, but, in total, 15 different values for genetic process were assigned in the first instance, and 11 for lithologies where a second process was implied.

Environment of formation could be assigned to more than 98% of lithologies (Figure 7). For the 17% of "lithology strings" where a second genetic process was recognized, a corresponding environment of formation could be assigned in almost all instances. As in the case of composition and genetic process, the frequency of occurrence of the various processes reflects the nature of the geological terrane that forms the basis of this test map set.

Texture was classified under five separate schemes. Over 92% of lithologies could be assigned to either a granular or crystalline style of grain intergrowth, with a somewhat even split between the two categories (Figure

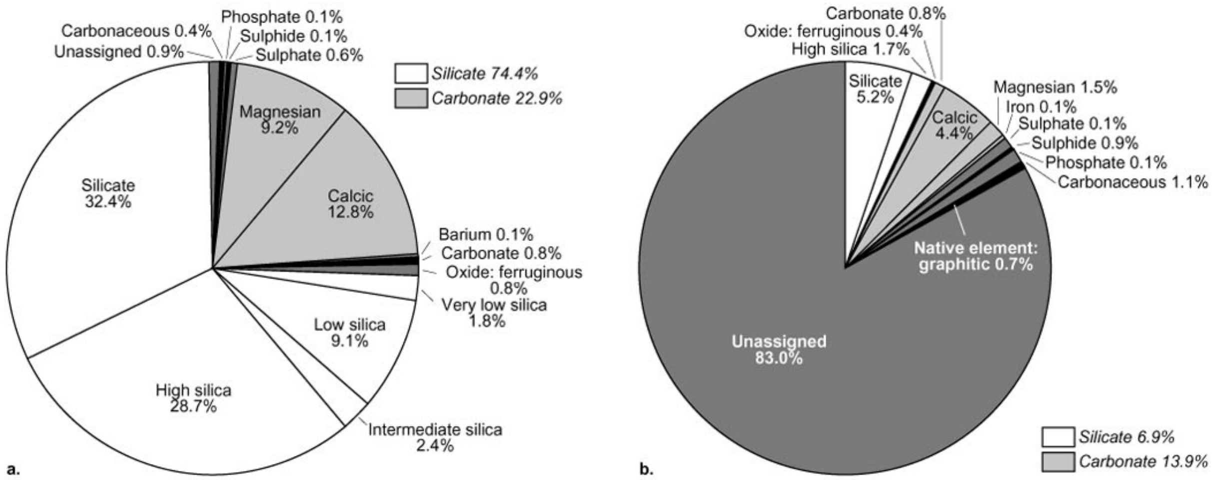


Figure 5. Breakdown of lithologies by their implicit composition. a. dominant composition. b. subordinate composition.

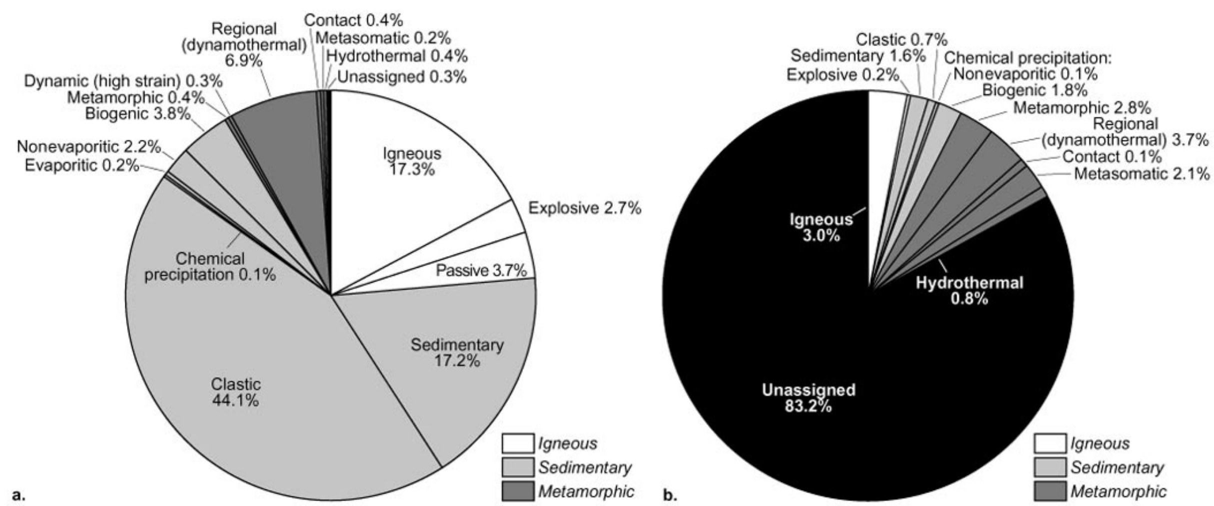


Figure 6. Breakdown of lithologies by their implicit genetic process. a. at least one process. b. second process.

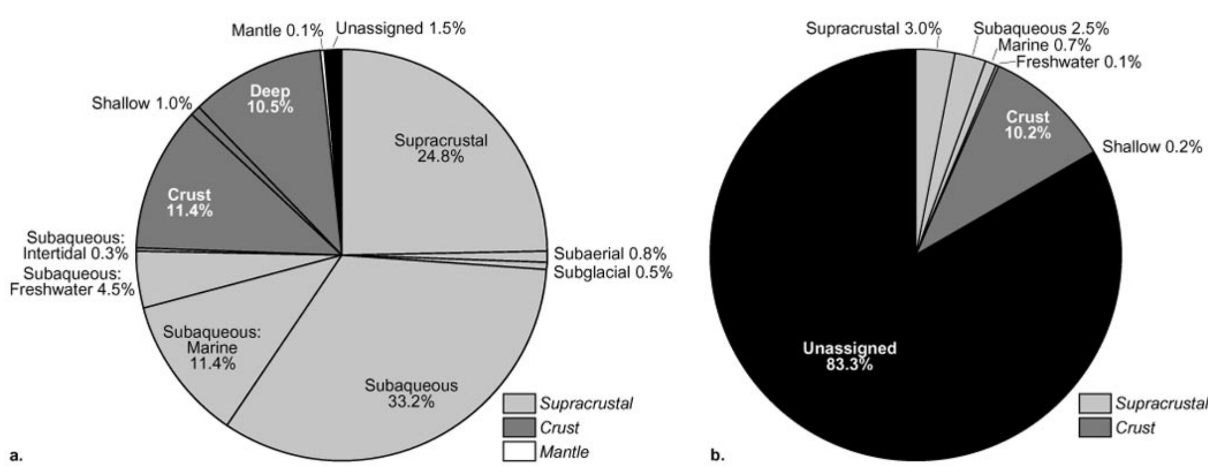


Figure 7. Breakdown of lithologies by their implicit environment of formation. a. first genetic process. b. second genetic process.

8a). Grain-size variability (Figure 8b) could be classified for about half of the "lithology strings," and was evenly split between homogeneous and heterogeneous, with about 2% classified as gradational. Grain size for the matrix (or all particles for homogeneous rocks) could be classified for two-thirds of the "lithology strings" (Figure 8c), with roughly one-third each in size class 2 (0.05-0.2 mm) and size class 3 (>0.2 mm), and less than 3% assigned to size class 1 (<0.05 mm). Although 24% of the "lithology strings" imply that the rocks are heterogeneous in grain size (Figure 8b), for less than 8% could the size of the coarse clasts or crystals be classified (Figure 8d), and almost all of these fell into size class 3. Grain-size morphology (Figure 8e) could be classified for only about 13% of the "lithology strings," and the majority of these fell into the "rounded" category. Finally, the proportion of matrix or groundmass to coarse clasts or crystals (rocks with heterogeneous grain size only) could be assigned to less than 5% of "lithology strings" (Figure 8f).

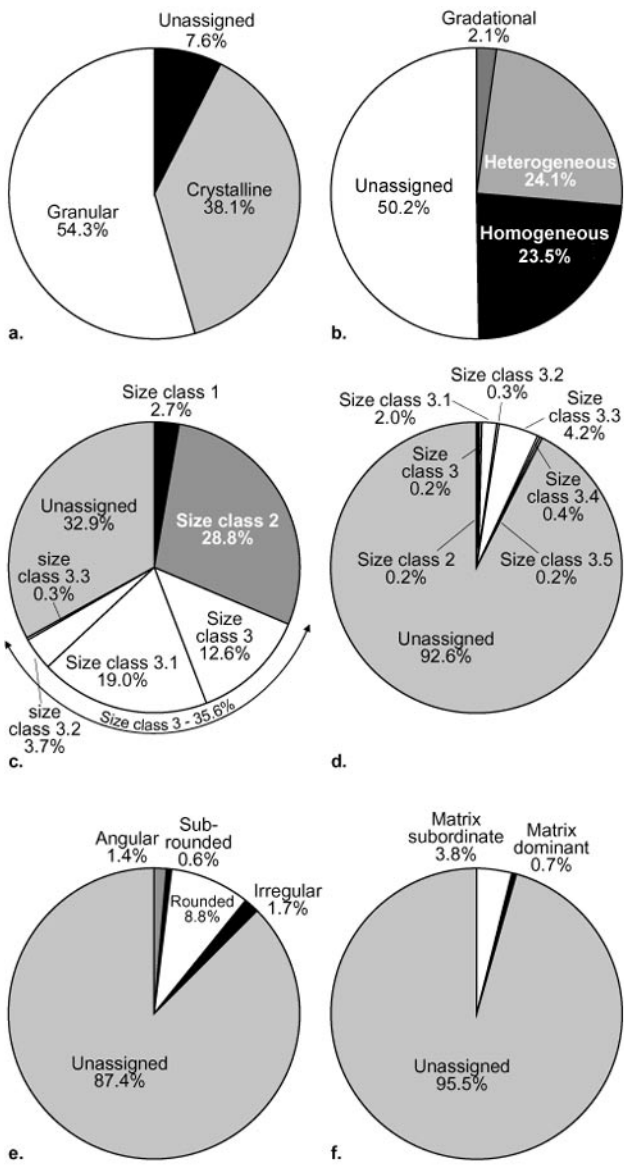


Figure 8. Breakdown of lithologies by their textural characteristics. a. grain intergrowth. b. grain-size variability. c. grain size, matrix. d. grain size, clasts/megacrysts. e. grain morphology. f. matrix abundance.

The last major category of indexes was based on physical form, both external habit and internal structure or fabric. External habit could be inferred for only about 25% of the "lithology strings" (Figure 9a), and nearly 23% of these were tabular, reflecting the predominance of stratified rocks in this map set. Internal structures were divided into two categories on the basis of their geometric dimensions: surfaces (2-D) and volumes (3-D). Because very few lithologies (less than 0.5%) were qualified by a linear (1-D) structure, this information has not been used. Almost 21% of "lithology strings" implied a planar fabric, either primary or secondary (Figure 9b), while less than 5% implied a 3-D internal structure (Figure 9c).

The final property that influenced the power of these classification schemes to select map units on specific rock properties was how many are implicit in the "root" rock names themselves, so that queries such as "make a map of all the units that contain carbonate minerals" can be effectively and efficiently constructed. Table 11 contains the percentage of the 161 rock names from Table 10 that intrinsically contain information for each of the main classification schemes.

The results are not unexpected—composition, genetic process, and environment of formation, together with some textural properties, are implicit in most simple rock names. Similarly, a rough indication of the degree of consolidation can be inferred for most materials. Physical form was not very useful, and most of the textural classifications were useful for less than half of the simple rock names.

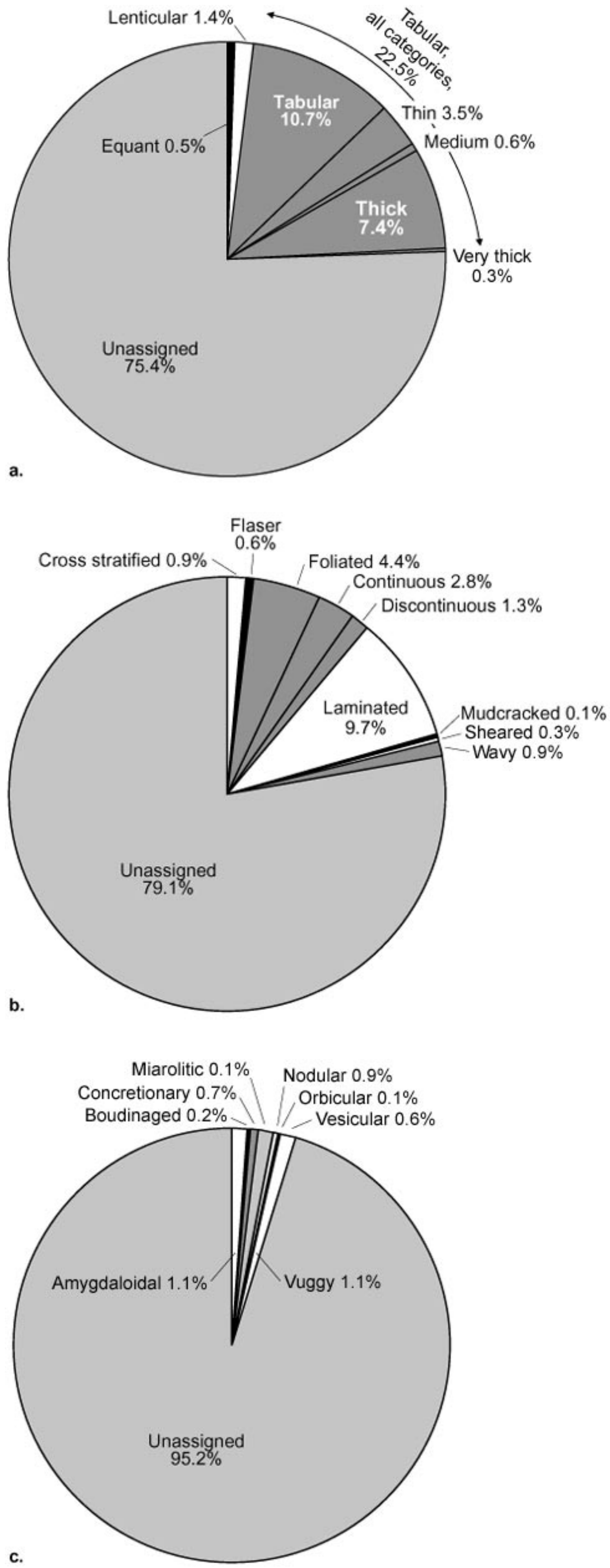


Figure 9. Breakdown of lithologies by their physical form. a. external habit. b. internal structure (fabric), surfaces. c. internal structure, volumes.

Table 11. Proportions of "root" rock names that implicitly allow classification against 17 schemes developed to systematically index lithological information from bedrock map units.

Classification scheme	Proportion of classified
Composition, dominant	93%
Composition, subordinate	2%
Genetic process, at least one	97%
Genetic process, second	19%
Environment of formation, first genetic process	94%
Environment of formation, second genetic process	19%
Texture, grain intergrowth	86%
Texture, grain-size variability	32%
Texture, grain size, matrix	44%
Texture, grain size, clasts/megacrysts	<1%
Texture, grain morphology	11%
Matrix proportion	4%
Physical form, external habit	0%
Physical form, internal surfaces	7%
Physical form, internal volumes	0%
Mechanical properties, induration	99%
Mechanical properties, partings	1%

DISCUSSION

Example of a Map Produced from a Simple Query on Lithological Properties

The true test of these classifications is their effectiveness in constructing simple queries to return specific subsets of map units. A single example will be described, based on the Yukon map set. The goal was to make a map of all units that contained volcanic rocks at both the source-unit and regional-compilation-unit levels of resolution. The query was constructed around the statement that a unit contain a lithology whose "genetic process = igneous" and "environment of formation = supracrustal," for either the first or second genetic process. The query was run on the whole database, but the results (Figure 10) are shown only for the same area as in Figure 2 (i.e., NTS sheet 105A, NE of the Tintina Fault). Figure 10 shows the outlines of all map units at each level; those contain-

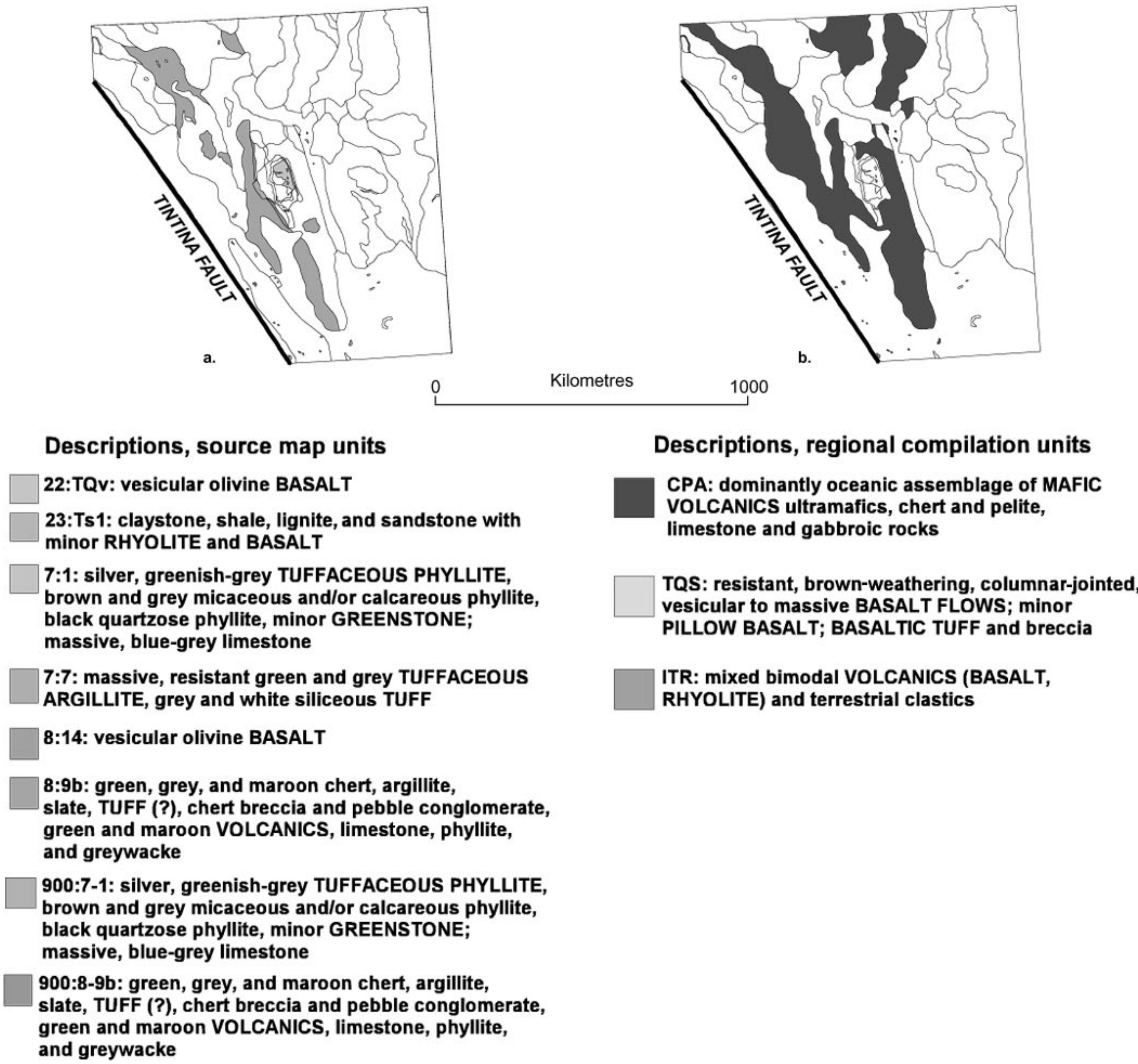


Figure 10. Results of a database query to select map units that contain volcanic rocks from the Yukon map set, shown for NTS sheet 105A only (see Figure 2). Left, source-map units; right, regional map units. The outlines of all units are shown, with those containing “volcanic” lithologies shaded, and their legend descriptions listed (“volcanic” lithologies capitalized).

ing volcanic rocks are shaded. The legend descriptions for the selected units are listed, with the lithologies that triggered their selection capitalized. The variety of these rock names is quite striking even within this small area. It is important that the descriptions of the selected units be scrutinized to ensure that they meet the purpose of the initial query. If not, the query should be refined.

Indexing Map Units

The generic indexes for lithology not only aid in the creation of derivative thematic maps as above, but also allow lithologically equivalent map units to be selected. Each map unit has in effect a “profile” of these generic classifications that is the aggregate of the classifications of its component lithologies. This profile can be used to search for other map units with the same profile. Because the classifications are broken down into several independent themes, and many of these classifications are shallow hierarchies, we envision developing a tool that allows the user to selectively modify the search profile by dropping or adding search parameters, or relaxing or tightening classification criteria. For example, if a unit used as the basis for a search was described as a “calcareous marine sandstone with minor concretionary marine shale," this would be parsed as:

lithology	**proportion**
sandstone, calcareous, marine	major
shale, concretionary, marine	minor

and classified as in Table 12. The profile in Table 12 could be modified to relax or tighten the search constraints. For example, if the "Physical form, internal surfaces = laminated" and "Mechanical properties, partings = platy" criteria for the second lithology were dropped, then units with minor mudstone, siltstone, and other fine-grained marine, clastic, concretionary rocks would be selected. These lithological criteria can be combined with lithological proportions and the age range of the target map units.

CONCLUSIONS

Scalability of geological maps can be achieved by applying the map-unit classifications from the legends of regional compilations to group the map units on more detailed maps used for most practical geological applications (typically at scales ≥1:250,000). Apart from national compilations at scales ≤1:5,000,000 (Wheeler and others, 1997), there are at present no nationally applied Canadian regional legend schemes that can be used at intermediate scales (i.e., 1:2,000,000 to 1:1,000,000). The Tectonic Terrane (Wheeler and others, 1991) and Tectonic Assemblage (Wheeler and McFeely, 1991) classifications developed for the Cordilleran Orogen are suitable approaches to fill this gap, and equivalent classifications are being developed for the Newfoundland and Nunavut regions. The various levels of regional legend do not belong to a simple hierarchy of map units, which means that instances of detailed map units on each source map must be assigned to their correct regional group at each level of generalization by compilers who have a thorough knowledge of the geology of the various regions of Canada.

These generic classification schemes for map units based on their lithological characteristics are preliminary, but we believe it is the type of classification scheme needed for a functional geological-map database. They are informed by geological principles, but do not purport to be anything more than categorical classifications. If we can make them truly generic we can unbundle the

Table 12. Generic lithological classification of a map unit described as "calcareous marine sandstone with minor concretionary marine shale." The actual values for each category are stored in an attribute table as COA ID's from the conceptual classifications in the COA table.

Classification scheme	sandstone, calcareous, marine	shale, concretionary, marine
Composition, dominant	silicate	silicate
Composition, subordinate	carbonate>calcic	
Genetic process	sedimentary>clastic	sedimentary>clastic
Genetic process, second	nul	nul
Environment of formation, first genetic process	supracrustal>subaqueous>marine	supracrustal>subaqueous>marine
Environment of formation, second genetic process	nul	nul
Texture, grain intergrowth	granular	granular
Texture, grain-size variability	homogeneous	homogeneous
Texture, grain size, matrix	size class 3.1	size class 2
Texture, grain size, clasts/megacrysts	nul	nul
Texture, grain morphology	nul	nul
Matrix proportion	nul	nul
Physical form, external habit	nul	nul
Physical form, internal surfaces	nul	laminated
Physical form, internal volumes	nul	concretionary
Mechanical properties, induration	yes	yes
Mechanical properties, partings	nul	platy

multidimensional nature of geological nomenclature. Such a classification scheme will greatly assist in making "derivative" maps, and perhaps in making geological map information less cryptic to the nonspecialist. Further, by keeping the number of "concepts" in the COA table to a reasonable size, the task of documenting the concepts (and translating the documentation into other languages) can be kept to a manageable size. Finally, by placing much of the detail (i.e., instances of concepts) in attribute tables, these can be managed locally, making the task of coordinating the "global concepts" more tractable. These generic classifications can also co-exist with more interpretative, thematic classifications at a regional map unit level such as that proposed by Struik and Quat (2002) for Tectonic Assemblages.

The project so far has addressed the map units of existing geological maps—the historical map information. For new maps, authors must be asked to place map units into predetermined higher levels, explicitly provide proportions of lithologies within map units, and complete the lithological indexing for map units. This last task will become much easier if geologists move toward the use of standard science language for rock names and qualifying information, as these standard terms would be already linked to generic keywords. Work on linear and point features from geological maps will be a task for the future.

REFERENCES

Boisvert, E., Morin, A., Lauzière, K. and Lebel, D., 2001, Using the proposed North American Data Model in a distributed database environment, *in* Soller, D.R., ed., Digital Mapping Techniques '01—Workshop Proceedings: U.S. Geological Survey Open-File Report 01-223, p. 35-44; <http://pubs.usgs.gov/of/2001/of01-223/boisvert.html>.

Colman-Sadd, S.P., Ash, J.S., and Nolan, L.W., 1997, GEOLEGEND: a database system for managing geological map units in a geographic information system: Computers and Geosciences, v. 23, p. 715-724.

Colman-Sadd, S.P., and Crisby-Whittle, L.V.J., compilers, 2001, Partial bedrock geology data set for the Island of Newfoundland: Newfoundland Department of Mines and Energy, Geological Survey, Open File NFLD/2616 version 3.0.

Gabrielse, H., Monger, J.W.H., Wheeler, J.O., and Yorath, C.J., 1992, Part A. Morphological belts, tectonic assemblages, and terranes, *Chapter 2 of* Gabrielse, H., and Yorath, C.J., eds., Geology of the Canadian Cordilleran Orogen in Canada: Geological Survey of Canada, Geology of Canada No. 4, p. 15-28.

Gordey, S.P. and Makepeace, A.J., compilers, 1999, Yukon digital geology: Geological Survey of Canada Open File D3826 and Exploration and Geological Services Division, Yukon, Indian and Northern Affairs Canada, Open File 1999-1(D) CD-ROM.

Jackson, J.A., 1997, Glossary of geology (4th ed.): Falls Church, Virginia, American Geological Institute, 769 p.

Matti, J., 2002, Classification of sedimentary materials, with attributes for their description and interpretation: Unpublished report of the Sedimentary Subgroup, Science Language Technical Team (SLTT), North American Data Model Steering Committee, August 2002.

North American Commission on Stratigraphic Nomeclature, 1983, North American Stratigraphic Code: American Association of Petroleum Geologists Bulletin, v. 67, p. 841-875.

Okulitch, A.V., 2002, Geological time scale 2002: Geological Survey of Canada Open File 3040 (National Earth Science Series—Geological Atlas) REVISION.

Richard, S.M., 2002, Classification of metamorphic rocks and other composite-genesis rock types, including hydrothermally altered, impact-metamorphic, mylonitic, and cataclastic rocks: Unpublished draft report of the Subgroup on Metamorphic Rocks, Science Language Technical Team (SLTT), North American Data Model Steering Committee, May 13, 2002.

Struik, L.C., and Quat, M.B., 2002, Hierarchical classification of tectonic assemblages for use with thematic computer-based query systems: a preliminary scheme: Geological Survey of Canada Current Research No. 2002-E18, 2002; 5 p.

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: Geological Survey of Canada Current Research No. 2002-D10, 9 p.

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; <http://pubs.usgs.gov/of/2001/of01-223/weisenfluh.html>.

Wheeler, J.O., Brookfield, A.J., Gabrielse, H., Monger, J.W.H., Tipper, H.W., and Woodsworth, G.J., compilers, 1991, Terrane map of the Canadian Cordillera, Geological Map of Canada: Geological Survey of Canada Map 1713A, scale 1:2,000,000.

Wheeler, J.O., Hoffman, P.F., Card, K.D., Davidson, A., Sanford, B.V., Okulitch, A.V., and Roest, W.R., compilers, 1997, Geological Map of Canada: Geological Survey of Canada Map D1860A, CD-ROM.

Wheeler, J.O., and McFeely, P., compilers, 1991, Tectonic Assemblage map of the Canadian Cordillera and adjacent parts of the United States of America, Geological Map of Canada: Geological Survey of Canada Map 1712A, scale 1:2,000,000.

A Geologic Gazetteer of the Lake Tahoe Region

By Jordan Hastings and Ada Otter

Department of Geography
University of California Santa Barbara
Santa Barbara, CA 93106
Telephone: (805) 893-3663
Fax: (805) 893-8617
e-mail: jordan@geog.ucsb.edu

ABSTRACT

The Lake Tahoe Region, straddling California and Nevada, presents a wealth of cultural, ecological, scientific, and scenic values; it has been inhabited for at least 8,000 years. The Lake Tahoe Region is within a relatively young, large, and very deep graben. The region's ecosystem is actively shaped by its geology, which includes strong tectonism and a history of recent landslides/tsunamis around the lake. Archaeological, historical-survey, and recent scientific mapping activities document on numerous maps the region's evolution. Over 3,400 names for topographic and geologic features appear on these maps.

In this paper, we describe the development of a geologically sophisticated gazetteer service, the Tahoe Regional Gazetteer (TARGA), which interrelates feature names with geologic maps. In conjunction, TARGA has built an inventory of 69 data sets, including 16 geologic maps, for the LTR, accumulated into a standardized repository. All three of TARGA's component subsystems—inventory, repository, and gazetteer—are Web accessible and Web mapped, providing convenient answers to such questions as: What geological maps exist for cultural and/or physical feature(s) "X"? In addition to its online capabilities, TARGA has accumulated a valuable database for future research on the geology of the LTR, and for geologic-map data-management systems in general.

INTRODUCTION

The Lake Tahoe Region[1] (LTR), straddling the states of California and Nevada (Figure 1), was designated a "national concern" under Presidential Order #13057 (Clinton, 1997), citing its "extraordinary natural, recreational, and ecological resources." This order also mandated the development of a comprehensive management plan for the LTR. In turn, this plan has spurred collection, digitization (where needed), and synthesis of a large number of geospatial information resources for the region within a geographic information system (GIS), specifically the Tahoe Environmental GIS (TEGIS). TEGIS was developed collaboratively by the U.S. Geological Survey (USGS) and the Tahoe Regional Planning Authority (TRPA) and is made available on the Lake Tahoe Data Clearinghouse Web site: <http://tahoe.usgs.gov/>.

A fundamental component of TEGIS, given Lake Tahoe's setting in a large, deep, and still tectonically active graben, is regional geology. Numerous geologic maps for the region exist (Cartier and others, 1994; Hess and Johnson, 1997; Bedford and others, 2002). These maps have been produced by different mappers at different times, in a variety of scales (1:24K, 1:62.5K, 1:250K, 1:500K [Nevada], and 1:750K [California]), projections (UTM10, UTM11, State Plane, Albers), and datums (NAD27, NAD83), as shown in Figure 2. The larger scale (<1:100K) and the smaller scale (≥1:250K) maps independently tile the region, thus overlapping each other and providing a welter of technical GIS challenges, in addition to scientific ones. A variety of USGS topographic base maps at 1:24K, 1:100K, and 1:250K also exist for the region.

Humans have occupied the environs of Lake Tahoe for at least 8,000 years (Forbes, 1982; Moratto, 1984), and in this period a large number of its cultural and physical features have been named (and renamed) in several languages. These names form the traditional basis for place identification, navigation, and wayfinding in the region. Some names are essentially point features (mountain peaks, springs, trail heads, forks, etc.), others are largely linear (creeks, roads, trails, etc.), and the majority are clearly areal (lakes and ponds, settlements, mine sites, wilderness areas, etc.). Most of the modern names appear on topographic maps, which the USGS Geographic Names Information System (GNIS, <http://geonames.usgs.gov/>) excerpts to point locations (often feature centroids, but sometimes simply georeferenced map-label positions). Other sources of georeferenced

[1]Herein defined as a 1° x 1° area, from 38.5°N, 120.5°W to 39.5°N, 119.5°W. This area includes the hydrologic closure of Lake Tahoe proper.

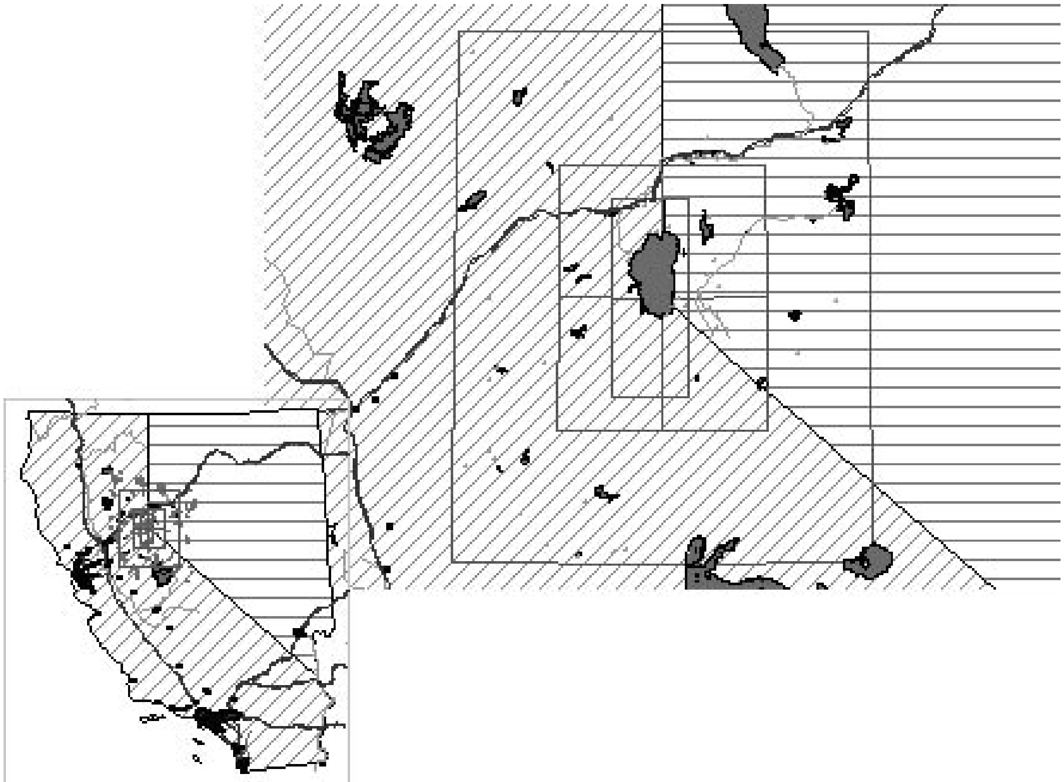


Figure 1. The Lake Tahoe region: 2° box surrounding 1° TARGA boundary, with TEGIS inset.

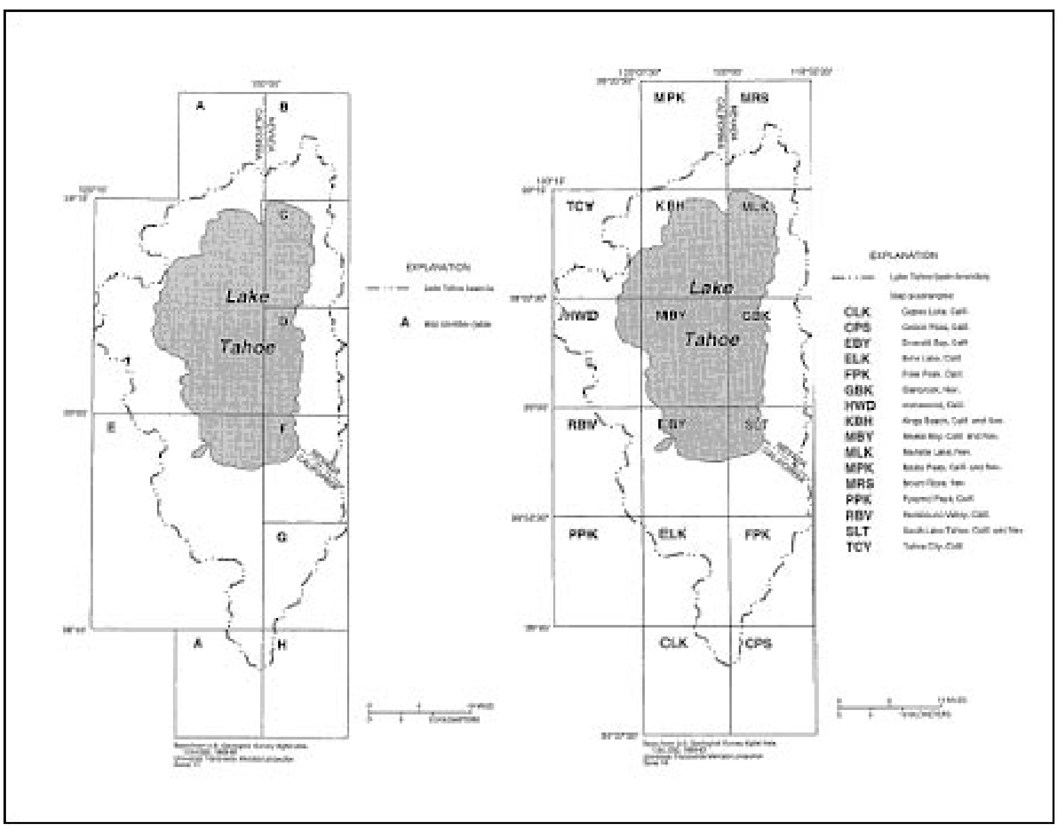


Figure 2. TEGIS geologic mapping by USGS and TRPA; 10 geologic maps cover 16 1:24K topographic quadrangles (from Cartier and others, 1994).

names include GIS/map data from commercial providers, the various Federal land management agencies, the states of California and Nevada (particularly their state geological surveys), local municipalities, historical documents, and, ultimately, anecdotal/common usage. Obviously, people can find their way to and among named features, whether or not these have been mapped, so undoubtedly more names exist than are documented in GNIS.

In this article, we describe a fusion of geological and topographic maps with a digital gazetteer[2] (Hill, 2001) for the LTR: the Tahoe Regional Gazetteer (TARGA). In the first instance, TARGA's purpose is largely geological, designed to answer such questions as:

- Where is/are feature(s) named "X"? The answer must take into account that one "X" may apply to multiple features.
- What geological maps exist for feature(s) "X"? Suitable maps may contain the whole or only part(s) of the features.
- How do I get to feature(s) "X" from some other feature or place "Y"?

Currently, the "X" and "Y" are restricted to cultural and physical geographic features; future extension to geological provinces and type sections is planned.

DEVELOPMENT

The TARGA system incorporates three major component subsystems: an Inventory, a Repository, and a Gazetteer. The development of each of these subsystems is briefly described below.

Inventory Subsystem

TARGA first developed an Inventory of the spatial datasets available for the LTR, including geologic maps, topographic maps (in both raster and vector forms), digital terrain models, and satellite imagery. For simplicity, this inventory was implemented in Microsoft (MS) Excel 2000. Altogether, 34 original data sets were identified, researched from various sources, and described in Excel by 18 attributes,[3] including name, source, type, scale and projection, and minimum bounding rectangle (MBR), as shown in Figure 3a. Visual Basic for Applications (VBA) scripts were written to assist with proofing the attributes as they were entered, and also to produce a descriptive Web page for each data set, complete with a thumbnail image of it (manually prepared), as shown in Figure 3b.

[2]A digital gazetteer is a "spatial dictionary" of named and typed features located in the environment.

[3]The full list appears in Appendix A.

Repository Subsystem

Source materials for the 34 original inventoried data sets were subsequently acquired in digital form, preprocessed in various ways, and stored in a companion repository for easy GIS access. Preprocessing tasks included format conversion, georeferencing, deprojection and/or reprojection, and clipping to the TARGA regional boundary. The result was a derived collection of standardized, conformable data sets in NAD27 geographic coordinates.[4] A total of 69 data sets were produced from the original 34 and incorporated in the TARGA repository, as summarized in Appendix A.

Gazetteer Subsystem

Each of the spatial data sets inventoried and processed covers an area of tens to hundreds of square kilometers. Within each area, numerous named sub-areas provide important context and/or foci for geologic map work. GNIS lists over 3,400 named places and features in the LTR, each identified by a point location derived from topographic maps at scales of 1:24K and smaller (i.e., less detailed). A commercial vendor, Geographic Data Technology (GDT), has documented about 1,000 of the larger features, complete with their polygonal geometry, predominantly taken from the same maps in digital form, (digital line graphs—DLG's). Typically, the GNIS locations fall within the GDT polygon boundaries for like-named features; however, there are exceptions. The same name may be given to multiple features; for example, there are three Frog Lakes and six Mud Lakes identified in the LTR. Also, a single feature may have several names; Lake Tahoe itself was formerly known as Lake Bigler. Redundant locations may exist for large features, particularly those that appear in portions on multiple maps. All such details were carefully resolved for each feature and recorded in an MS Access 2000 database, with the following principal fields: feature name, feature type, location(s), and MBR (computed to at least include all valid locations). The simple database schema appears in Appendix B.

APPLICATION

TARGA is intended to assist in finding maps, particularly geological maps, and other geospatial data sets relating to named features in the LTR. A prototypical

[4]NAD27 rather than NAD83 was chosen because the majority of topographic base maps and geologic maps, as well as all the GNIS gazetteer features, are reported in this datum.

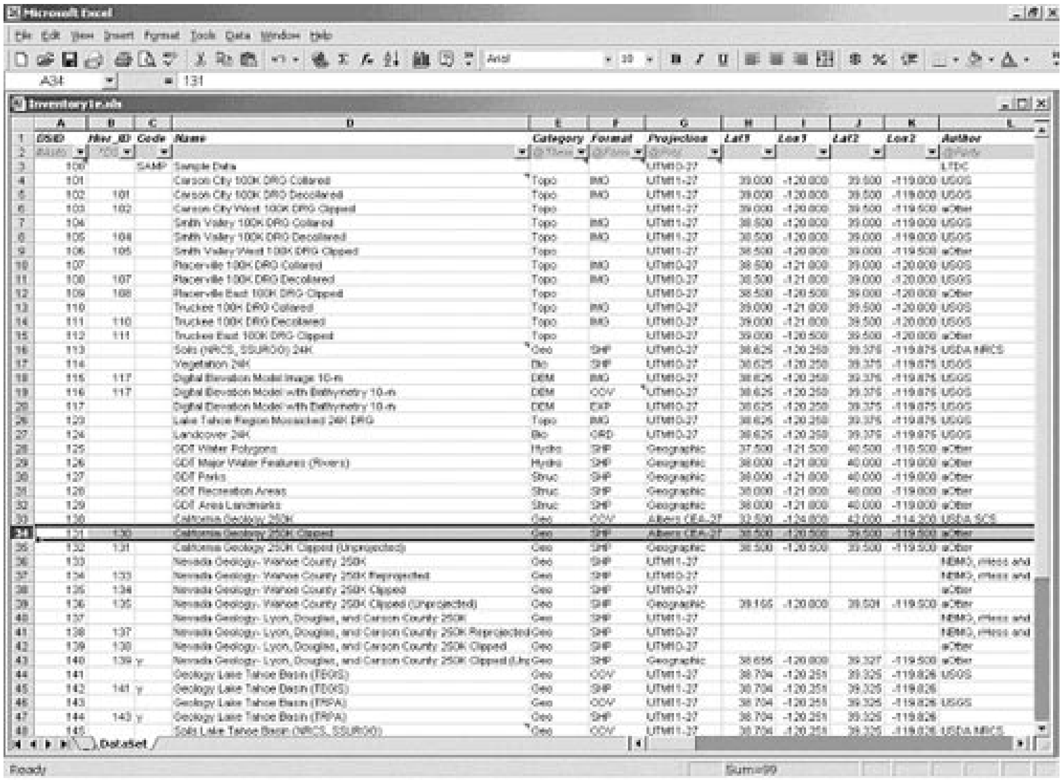


Figure 3a. TARGA Inventory workbook.

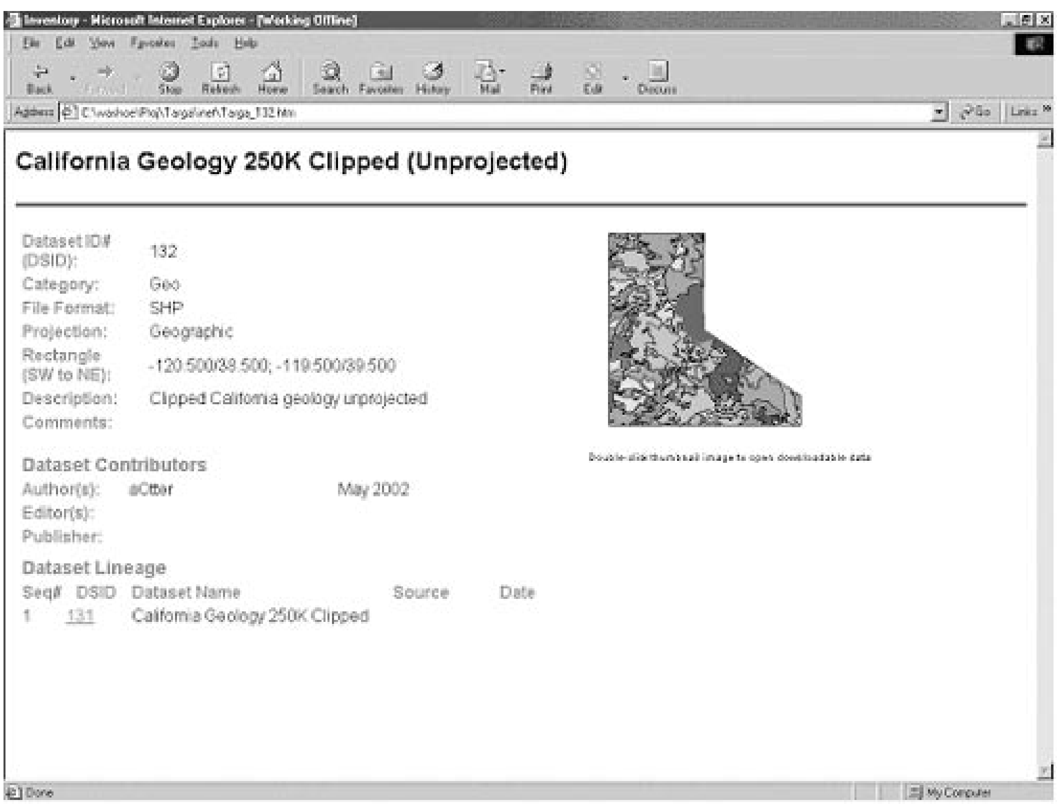


Figure 3b. TARGA Inventory Web page for one map product.

query is: Show me the maps on which geographic feature "X" appears. Answering this query involves integrating the three component subsystems described above: the inventory, the repository, and the gazetteer. The answer to the query is ideally shown in a GIS; tabular results also may be desired.

Specifically, the query process is: (1) the selected feature names/types are matched within the gazetteer, which returns their locations and/or MBR's; (2) the inventory is searched to determine data sets that contain the places (by point-in-rectangle and/or rectangle-on-rectangle overlay calculation, based on the data sets' MBR's); and (3) the qualifying data sets from the repository are shown in map view. In addition, a simple list of the qualifying data sets is made available outside the GIS. To increase the specificity of the system, and to reduce false retrievals, feature matches may be limited within a region of interest, indicated either by MBR (directly) or by name (indirectly, again using the gazetteer). For convenience, the system is Web accessible, which in turn mandates that the query interface be kept simple. A sample query/retrieval dialog from TARGA is shown in Figure 4.

Access to descriptive data about the data sets retrieved—metadata—also is important. Basic metadata for each of the datasets is provided by supporting Web pages, one per data set, automatically generated by VBA scripts from the inventory subsystem, as previously mentioned. A sample page is shown in Figure 3b. These metadata pages are explicitly linked in the list of retrieved data sets; they are also implicitly linked to the Web-mapping display (accessible by right-click). The entire collection of metadata pages is centrally available via the inventory subsystem.

TARGA is hosted on an MS Windows 2000 Server system, running MS Internet Information Server 5. Accordingly, the query interface is implemented in .asp scripts.

Basic GIS/Web-mapping support is provided by Environmental Systems Research Institute (ESRI) ArcView IMS software, running on the same system.

FUTURE WORK

TARGA has achieved its primary design objective: stimulating the collection and documentation of an integrated suite of geospatial data sets and maps for the Lake Tahoe region. Of particular note is standardization of the 16 geological maps, including one soils map, for the region in geographic (NAD27) coordinates.

Future work with TARGA will be to migrate its gazetteer database from Excel to Access, and also to modernize its Web mapping, using ESRI ArcIMS v4 in place of ArcView IMS. In conjunction with these changes, more sophisticated spatial operations, involving general polygon-on-polygon operations, as well as the present simple MBR tests, will be supported. Also, nongeographic reprojection of raster data sets together with vector data sets will be possible.

ACKNOWLEDGMENTS

The authors acknowledge the kind support of the California Geological Survey, Geographic Data Technology, the Nevada Bureau of Mines and Geology, and the U.S. Geological Survey, for both dataset access and

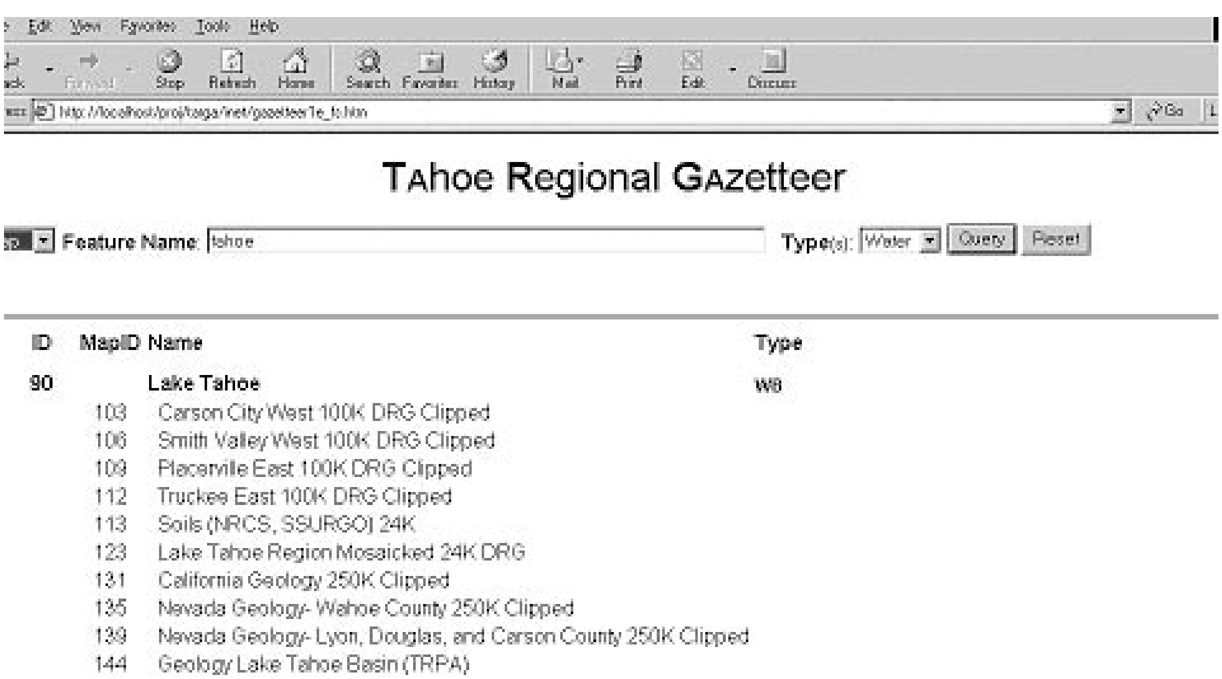


Figure 4. TARGA Gazetteer interface, showing maps on which Lake Tahoe appears.

supporting discussions during the course of this project. Collaborative discussions with the Alexandria Digital Library Gazetteer Project at the University of California Santa Barbara were also helpful.

REFERENCES

Bedford, D.R., Ludington, S., Nutt, C., Stone, P.A., Miller, D.M., Miller, R.J., Wagner, D.L., and Saucedo, G.J., 2002, Geologic database for digital geology of California, Nevada, and Utah—an application of the North American Data Model: U.S. Geological Survey Open-File Report.

Cartier, K.D., Peltz, L.A., and Smith, J.L., 1994, Development and documentation of spatial data bases for the Lake Tahoe Basin, California and Nevada: U.S. Geological Survey Water Resources Investigation Report 93-4182, 65 p.

Clinton, W.J., 1997, Federal actions in the Lake Tahoe Region: Office of the President, Executive Order #13057, dated July 26, 1997; <http://www.usda.gov/news/tahoe/ninteen.htm>.

Forbes, J.D., 1982, Native Americans of California and Nevada: Happy Valley, California, Naturegraph Publishers, 240 p.

Hess, R.H., and Johnson, G.L., 1997, County digital geologic mapping project: Nevada Bureau of Mines and Geology Open File Report 97-1, CD-ROM.

Hill, L.L., 2000, Core elements of digital gazetteers: placenames, categories, and footprints, *in* Borbinha, J., and Baker, T., eds., Research and advanced technology for digital libraries: Proceedings of the 4th European Conference, ECDL 2000 Lisbon, Portugal, September 18-20, 2000, Springer: Berlin, p. 280-290; <http://www.alexandria.ucsb.edu/~lhill/paper_drafts/ECDL2000_paperdraft7.pdf>.

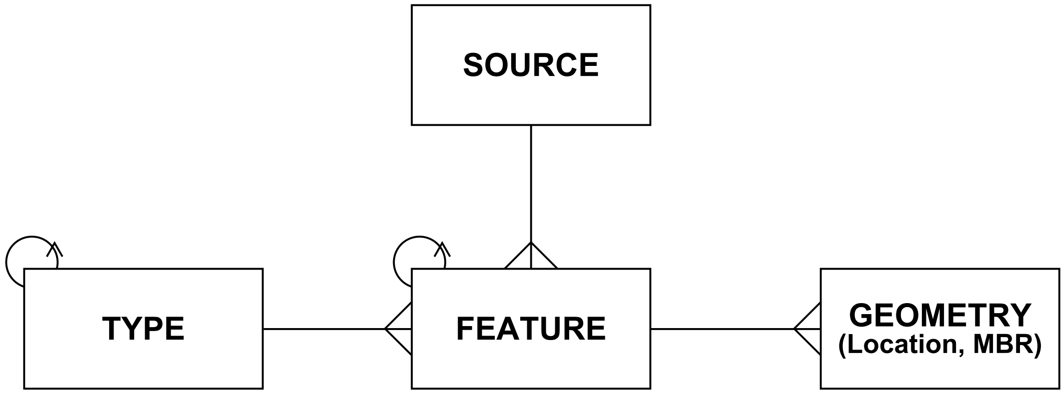
Moratto, M.J., 1984, California archaeology: Orlando, Florida, Academic Press, 757 p.

Appendix A. TARGA Repository Contents

DSID #Auto	Hier_ID *DSID	Name	Format @Form	Projection @Proj	SourceScale
100		Sample Data		UTM10-27	
101		Carson City Topo Collared	IMG	UTM11-27	1:100,000
102	101	Carson CityTopo Decollared	IMG	UTM11-27	1:100,000
103	102	Carson City West Topo Clipped	GRD	Geographic	1:100,000
104		Smith Valley Topo Collared	IMG	UTM11-27	1:100,000
105	104	Smith Valley Topo Decollared	IMG	UTM11-27	1:100,000
106	105	Smith Valley West Topo Clipped	GRD	Geographic	1:100,000
107		Placerville Topo Collared	IMG	UTM10-27	1:100,000
108	107	Placerville Topo Decollared	IMG	UTM10-27	1:100,000
109	108	Placerville East Topo Clipped	GRD	Geographic	1:100,000
110		Truckee Topo Collared	IMG	UTM10-27	1:100,000
111	110	Truckee Topo Decollared	IMG	UTM10-27	1:100,000
112	111	Truckee East Topo Clipped	GRD	Geographic	1:100,000
113		Soils (LTDC)	SHP	UTM10-27	1:24,000
114		Vegetation	SHP	UTM10-27	1:24,000
115	117	DEM Image 10-m	IMG	UTM10-27	
116	117	DEM with Bathymetry 10-m	COV	UTM10-27	
117		DEM with Bathymetry 10-m	EXP	UTM10-27	
118		Buildings	SHP	UTM10-27	1:24,000
119		Docks	SHP	UTM10-27	1:24,000
120		Hydrology	SHP	UTM10-27	1:24,000
121		Hypsography	SHP	UTM10-27	1:24,000
122		Roads	SHP	UTM10-27	1:24,000
123		Lake Tahoe (LTDC) Topo	IMG	UTM10-27	1:24,000
124		Landcover	GRD	UTM10-27	1:24,000
125		GDT Water Polygons	SHP	Geographic	
126		GDT Major Water Features (Rivers)	SHP	Geographic	
127		GDT Parks	SHP	Geographic	
128		GDT Recreation Areas	SHP	Geographic	
129		GDT Area Landmarks	SHP	Geographic	
130		CA Geology	COV	Albers CEA-27	1:250,000
131	130	CA Geology Clipped	SHP	Albers CEA-27	1:250,000
132	131	CA Geology Clipped	SHP	Geographic	1:250,000
133		NV Geology- Wahoe	SHP	UTM11-27	1:250,000
134	133	NV Geology- Wahoe Reprojected	SHP	UTM10-27	1:250,000
135	134	NV Geology- Wahoe Clipped	SHP	UTM10-27	1:250,000
136	135	NV Geology- Wahoe Clipped	SHP	Geographic	1:250,000
137		NV Geology- Lyon,Douglas,Carson	SHP	UTM11-27	1:250,000
138	137	NV Geology- Lyon,Douglas,Carson Reprojected	SHP	UTM10-27	1:250,000
139	138	NV Geology- Lyon,Douglas,Carson Clipped	SHP	UTM10-27	1:250,000
140	139	NV Geology- Lyon,Douglas,Carson Clipped	SHP	Geographic	1:24,000; 1:62,500; 1:250,000
141		Geology Lake Tahoe Basin (TEGIS)	COV	UTM11-27	1:24,000; 1:62,500; 1:125,000
142	141	Geology Lake Tahoe Basin (TEGIS)	SHP	UTM11-27	1:125,000
143		Geology Lake Tahoe Basin (TRPA)	COV	UTM11-27	1:24,000; 1:62,500; 1:125,000
144	143	Geology Lake Tahoe Basin (TRPA)	SHP	UTM11-27	1:24,000; 1:62,500; 1:125,000
145		Soils (TRPA)	COV	UTM11-27	1:24,000
147		LTDC Mapbounds	SHP	Geographic	n/a
148		TRPA,TEGIS Mapbounds	SHP	Geographic	n/a
149		1 Degree Mapbounds	SHP	Geographic	n/a
150		2 Degree Mapbounds	SHP	Geographic	n/a
151		3 Degree Mapbounds	SHP	Geographic	n/a
152		Hydrologic Units Map Great Basin	COV	Albers CEA-1866	1:250,000
153	152	Hydrologic Units Map Great Basin Clipped	SHP	Albers CEA-1866	1:250,000
154	153	Hydrologic Units Map Great Basin Clipped w/ 5 Mapbounds	SHP	Geographic	1:250,000
155		Hydrologic Units Map CA 250K	COV	Albers CEA-1866	1:250,000
156	155	Hydrologic Units Map CA Clipped	SHP	Albers CEA-1866	1:250,000
157	156	Hydrologic Units Map CA Clipped w/ 9 Mapbounds	SHP	Geographic	1:250,000
158		Streams- US	COV	Albers CEA-1866	1:2,000,000
159	158	Streams- CA,NV Clipped	SHP	Albers CEA-1866	1:2,000,000
160	101	Carson City Topo Collared	GRD	UTM11-27	1:100,000
161	101	Carson City Topo Collared Unprojected	GRD	Geographic	1:100,000
162	104	Smith Valley Topo Collared	GRD	UTM11-27	1:100,000
163	104	Smith Valley Topo Collared Unprojected	GRD	Geographic	1:100,000
164	107	Placerville Topo Collared	GRD	UTM10-27	1:100,000
165	107	Placerville Topo Collared Unprojected	GRD	Geographic	1:100,000
166	110	Truckee Topo Collared	GRD	UTM10-27	1:100,000
167	110	Truckee Topo Collared Unprojected	GRD	Geographic	1:100,000
168	103;106; 109;112	Tahoe Region Topo	GRD	Geographic	1:100,000
169	168	Tahoe Region Topo	IMG	Geographic	1:100,000

NOTE: Only selected attributes are shown; the full list of attributes for each data set is:
UniqueID #, Hierarchical/Parent ID#, Source Scale, File Size, Code (Web availability), Name, Category, Format, Projection, Mapbound extent (Lat1, Lon1, Lat2, Lon2), Author, Author Date, Editor, Editor Date, Publisher, Publisher Date, Description, Comments, Thumbnail Image, and Filepath.

Appendix B. TARGA Gazetteer Database Schema



SOURCE (ID, Name, …)

TYPE (<u>ID</u>, *Parent-ID*, Code, Name, …)

FEATURE (ID, *Parent-ID, SOURCE-ID, TYPE-ID*, Name, …)

GEOMETRY (*FEATURE-ID*, Location-Lat, Location-Lon, MBR-Lat1, MBR-Lat2, MBR-Lon1, MBR-Lon2, …)

Notes:

- Primary Keys are underlined, viz. <u>ID</u>
- Foreign Keys are italicized, viz. *Parent-ID, TYPE-ID*
- Both FEATURE and TYPE are hierarchically-structured, viz. *Parent-ID*, indicated by reflexive (circular) relationship in upper left-hand corner

The National Geologic Map Database: A Progress Report

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

[1]U.S. Geological Survey
908 National Center
Reston, VA 20192
Telephone: (703) 648-6907
Fax: (703) 648-6937
e-mail: drsoller@usgs.gov

[2]Ohio Geological Survey
4383 Fountain Square Dr.
Columbus, OH 43224
Telephone: (614) 265-6988
Fax: (614) 268-3669
e-mail: thomas.berg@dnr.state.oh.us

The National Geologic Mapping Act of 1992 and its reauthorizations in 1997 and 1999 (PL106-148) require that a National Geologic Map Database (NGMDB) be designed and built by the U.S. Geological Survey (USGS), in cooperation with the Association of American State Geologists (AASG) and other entities participating in the National Cooperative Geologic Mapping Program. The Act notes that the NGMDB is intended to serve as a "national archive" of geologic maps, to provide to a wide variety of people, from private citizens to professional geologists, the information needed to address various societal issues. The Act requires the NGMDB to also include the following related map themes: geophysics, geochemistry, paleontology, and geochronology. In this progress report, the term "geoscience" is used to refer to these five map themes.

In mid-1995, the general stipulations in the Act were addressed in the proposed design and implementation plan developed within the USGS and the Association of American State Geologists (AASG). This plan was summarized in Soller and Berg (1995). Because many maps are not yet in digital form and because many organizations produce and distribute geologic maps, it was decided to develop the NGMDB in several phases.

The first and most fundamental phase includes a comprehensive, searchable Catalog of all geoscience maps in the United States, whether in either paper or digital format. Figure 1a shows how the Map Catalog can be used to find a particular geologic map. Upon searching the NGMDB Catalog and identifying the needed map(s), the user is linked to the map data, the metadata, or to the appropriate organization for information about how to purchase the map. (The organization could be a participating state or federal agency, association, university, or private company.) The Map Catalog presently is supported by two databases developed under the NGMDB project: (1) GEOLEX, a searchable geologic names lexicon; and (2) Geologic Mapping in Progress, which provides information on current mapping projects, prior to inclusion of their products in the Map Catalog. In the coming year, an Image Library will be prototyped and made available to the public; this new initiative is described below. Plans for the prototype National Paleontology Database also are discussed below.

The second phase of the project focuses on public access to digital geoscience maps, and on the development of digital map standards and guidelines needed to improve the utility of those digital maps. The third phase proposes, in the long term, to develop an online, "living" database of geologic map information at various scales and resolution. Some functions of the planned online database, and its links to databases developed under Phase One, are shown in Figure 1b.

In late 1995, work began on Phase One. The formation of several Standards Working Groups in mid-1996 initiated work on Phase Two. Progress was summarized in Soller and Berg (1997, 1998, 1999a, 1999b, 2000, and 2001). At the Digital Mapping Techniques '98 through '02 workshops, a series of presentations and discussion sessions provided updates on the NGMDB and, specifically, on the activities of the Standards Working Groups. This report summarizes progress since the project's inception, but focuses on accomplishments since mid-2001. Further and more current information may be found at the NGMDB project-information Web site, at <http://ncgmp.usgs.gov/ngmdbproject>. The searchable

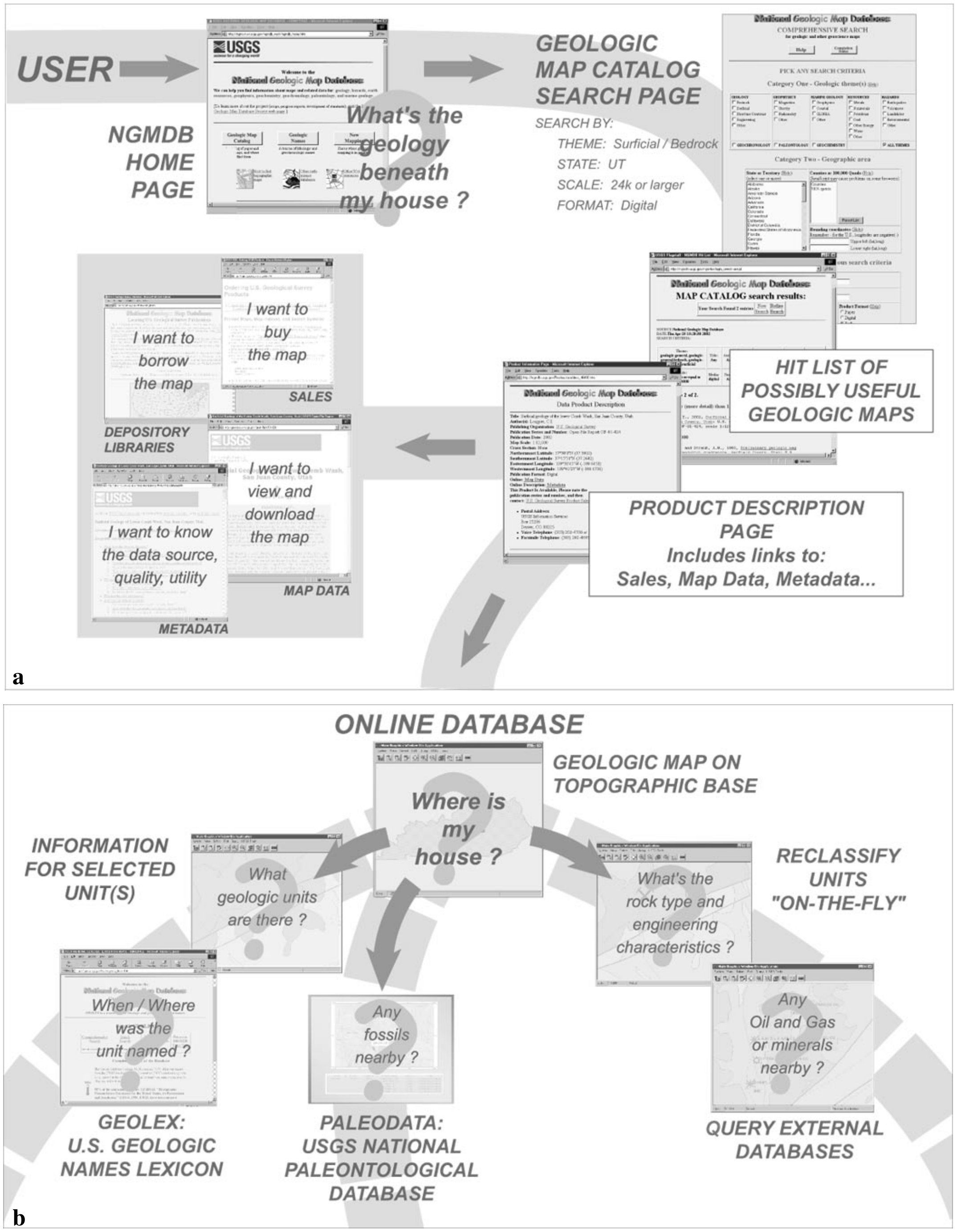


Figure 1. Diagram showing how a user might navigate the NGMDB Map Catalog and the online map database.

a. The user, interested in knowing something about the geology of an area (e.g., the land beneath his house), queries the Map Catalog, which returns a hit list of possibly useful maps. The user selects a map entry and, from the Product Description Page, obtains further information and can choose to either buy the map, view and download it, or inspect the metadata. The dark arrow toward the bottom of the figure points toward the online map database (Figure 1b).

b. The user queries the online map database. (Note: the paths along these queries are shown as dashed lines, to indicate planned development.) From the initial display showing the geology of the area surrounding his house, the user might choose to reclassify those units in order to derive a map showing engineering properties or query the geologic data in relation to external databases (here, an oil and gas database). Further, the user might be interested in the history of a particular geologic name or the availability of fossils; if so, the online map database would connect to GEOLEX or the National Paleontology Database.

database is available at <http://ngmdb.usgs.gov>.

To submit general comments about project scope and direction, please address the authors directly (see above). 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 (e.g., see "Your comments are welcome," at <http://ngmdb.usgs.gov/>).

PHASE ONE

The project opened its Web site to the public in January 1997, as a prototype intended to solicit comments regarding the Map Catalog. Since then, and with public access to GEOLEX and the Mapping in Progress databases beginning in 1998, Web site usage has gradually increased. Since the time when essentially all USGS maps were entered into the Map Catalog, Web-usage statistics indicate that many users are visiting the site several times a month. This trend suggests that the site is becoming a more useful resource. Additional increases in use are expected as the Catalog and GEOLEX become fully populated and as the other NGMDB databases come online. Figure 2 shows the number of people (actually the number of unique IP addresses or computers) who have used the NGMDB monthly since it opened to the public. (Note: Web "hits" are not shown in the figure, because they do not provide a realistic assessment of usage for sites like this. Further, the measure of unique IP addresses may in the future become a less reliable indicator, because increased security measures at some agencies are necessitating the use of dynamic IP addressing.)

The Map Catalog

The Map Catalog is designed to be a comprehensive, searchable Catalog of all geoscience maps of the United States, in paper or digital format. Entries to the Catalog include maps published in geological survey formal series and open-file series, maps in book publications, maps in theses and dissertations, maps published by park associations and scientific societies, maps published by other agencies, and publications that do not contain a map but instead provide a geological description of an area (for example, a state park). At the time of the DMT '02 conference, the Catalog contained a record for each of nearly 51,000 map products. Essentially all USGS maps have been recorded in the Catalog, and most state geological surveys are entering bibliographic records for all their maps and related maps (e.g., university theses). By the date of the DMT '02 meeting, geological surveys in 34 states were entering map records, as well as one University (Stanford). Maps by the Geological Society of America, the American Association of Petroleum Geologists, and numerous other publishers have been entered into the Catalog this year.

Soon after the DMT '02 meeting, the Map Catalog search page was extensively revised. It now addresses the diverse needs of our user audience through four search options (Figure 3). The easy-to-use Place Name Search is designed to address the needs of nongeologists who want to use a simple interface to find information about their home, town, or worksite, whereas the Comprehensive Search offers researchers a full range of search criteria.

Geologic Names Lexicon

The searchable, online, geologic-names lexicon (GEOLEX) contains roughly 90% of the geologic names found in the most recent listing of USGS-approved geologic names (published in 1996 as USGS Digital Data

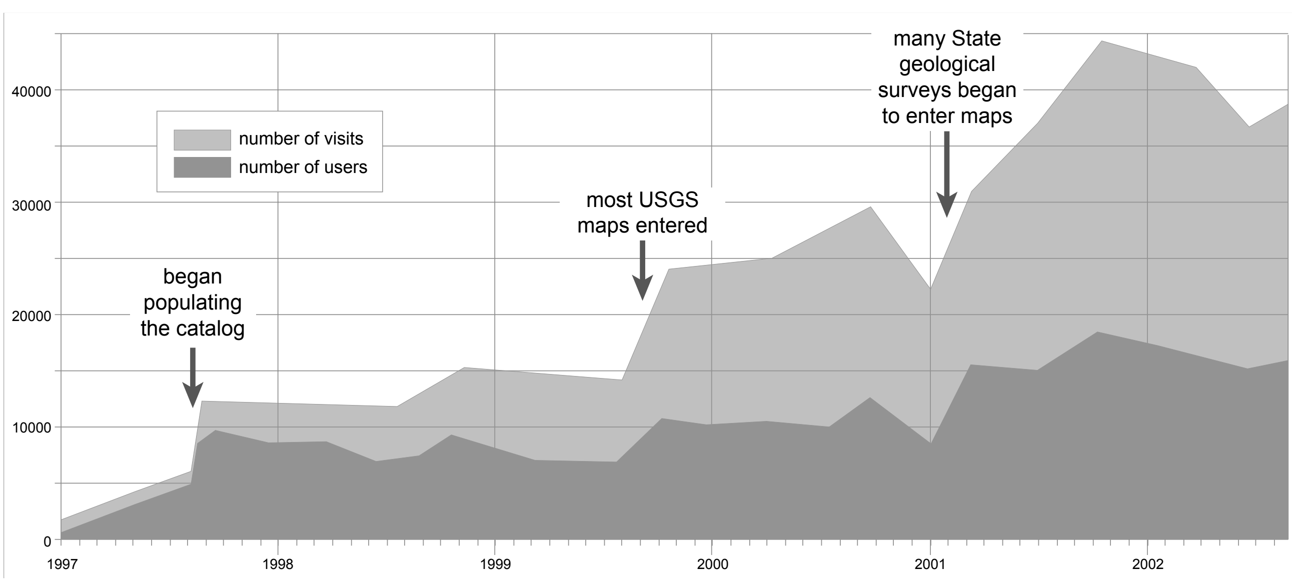


Figure 2. Monthly web-usage statistics for the Map Catalog, GEOLEX, and Mapping in Progress Databases. The Map Catalog accounts for roughly 75-80% of the usage.

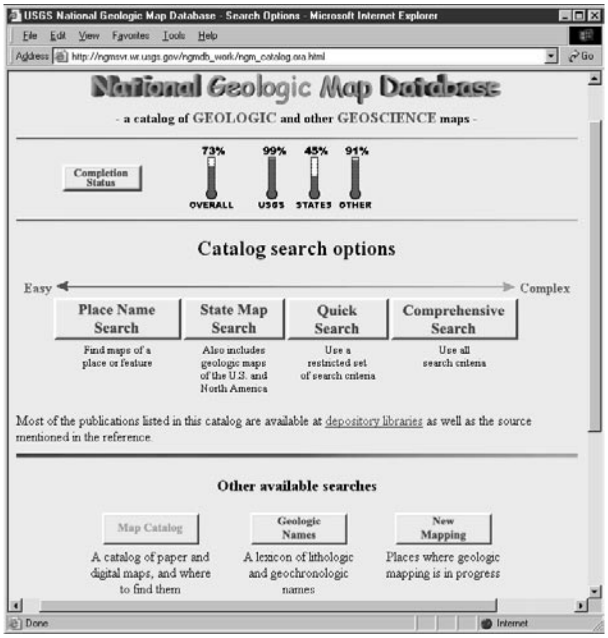


Figure 3. The new Map Catalog home page, showing the four search methods—Place Name, State Geologic Maps, Quick Search, and Comprehensive Search.

Series DDS-6, revision 3) and is estimated to contain roughly 75% of all geologic names in the United States. Prior to loading into GEOLEX, the information on DDS-6 was consolidated, revised, and error-corrected. Our work now focuses on resolving name conflicts and adding reference summary and other information for each entry and on preparing the database for USGS Director's Approval as a standing database. Work remaining includes incorporating geologic names not found on DDS-6 but recorded in the geologic-names card catalog at USGS headquarters and names approved by the state geological surveys but not yet in the USGS records.

GEOLEX is intended to be the comprehensive, authoritative listing of approved geologic names and is available as a resource for geologic mappers nationwide. Many state geological surveys have been registering new geologic names with the USGS for decades and are encouraged to continue this practice under GEOLEX, through a Web-based application form.

Geologic Mapping in Progress Database

To provide users with information about current mapping activities at 1:24,000- and 1:100,000-scale (1:63,360- and 1:250,000-scale in Alaska), a Geologic Mapping in Progress Database is available at <http://ngmdb.usgs.gov/MapProgress/MapProgress_home.html>. This database will be linked to the quadrangle and county search capability of the Image Library (see below).

Image Library

Through discussions with users and from comments received via our Web feedback form, it has become 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 others, it implies access to map images. We address the enormous task of developing a vector-based map database in Phase Three (see below). Here, we address the potential for providing map images to users.

If we view the project's Map Catalog and online map databases (Phase Three) as endpoints on a spectrum of complexity and ease-of-use, it is obvious that a significant gap exists between them. The Map Catalog exists today; it is relatively straightforward to use and it simply provides bibliographic information about each geoscience map product. In contrast, the online map database is still in its formative, prototype stage; when publicly available, it will provide full access to detailed geologic-map information, but many users may not be sufficiently familiar with geoscience concepts to comfortably use it. This is a serious concern, as this project seeks to address the needs of all users, of various backgrounds and interests. As noted below, the AASG and USGS are working together to build the foundation for the online map database, through development of the necessary technology, science concepts and data model, and a collection of widely distributed digital geologic-map coverage.

In the middle ground along this spectrum, we have the opportunity to provide users with geologic-map information in a raster image format, thereby allowing them Web access to the familiar paper map format we've known for generations (Figure 4). Therefore, we have begun a new project initiative, to build a library of geologic map images. The images will be managed in high-quality compressed format, using MrSID technology. To deliver this information, we are beginning to design a prototype database and Web site to allow users to find and view geologic maps (Figure 5). At present, we anticipate providing the capability for searching by quadrangle (e.g., 1:24,000, 1:100,000, 1:250,000-scale) and by county. Our efforts will focus on providing images of general-purpose bedrock and surficial geologic maps.

The Image Library will link directly to the Map Catalog and the Geologic Mapping in Progress Database. As evident in Figure 5, through these links we intend to direct users to the agencies producing the maps. We hope this initiative will further strengthen the cooperative relationship between the AASG and USGS.

Paleontology

The NGMDB project has designed and is planning to develop a national paleontology database (see Wardlaw

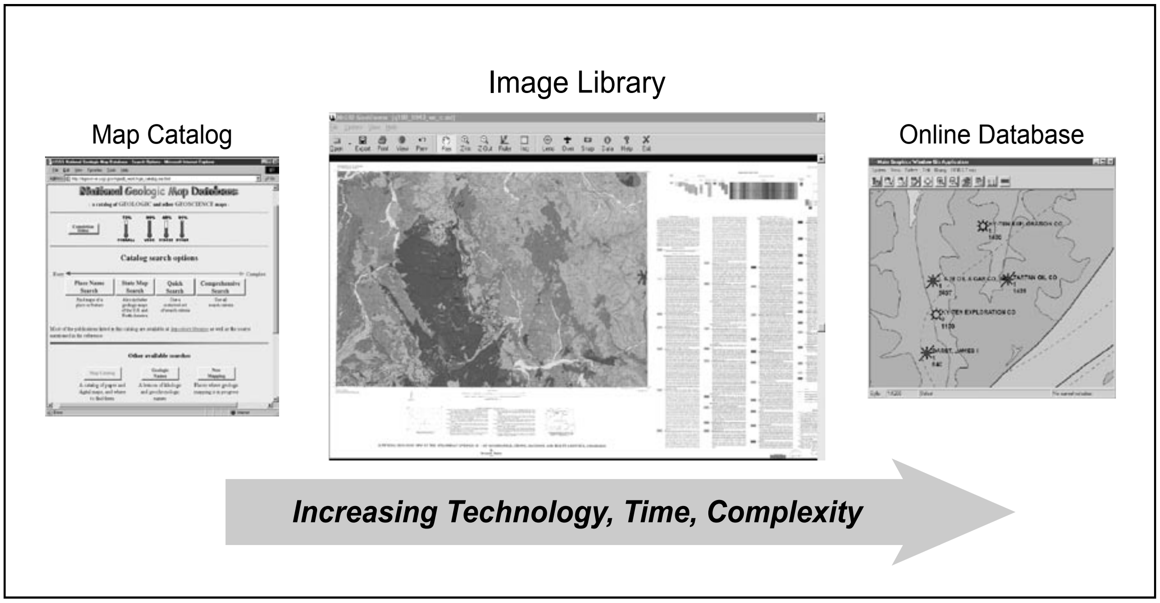


Figure 4. The proposed Image Library occupies a middle ground between the Map Catalog and the online map database in regard to technological complexity, ease-of-use, and related issues. It will provide users with geologic map information in a familiar image format.

and others, 2001). Our general plan is to build prototypes of this database in areas where geologic mapping is underway, so that we can work with mapping projects to design a database useful to science as well as to the public. A publicly accessible prototype is expected within the year.

PHASE TWO

Phase Two is directed mostly toward the development of standards and guidelines needed to help the USGS and state geological surveys more efficiently produce digital geologic maps and to produce those maps in a more standardized and common format among the various map-producing agencies. Significant progress has been made toward developing some of these standards and guidelines and to providing Map Catalog users with access to online products.

Standards Development

The following summaries concern activities of the AASG/USGS Standards Working Groups and their successors. General information about the Working Groups and details of their activities are available at <http://ncgmp.usgs.gov/ngmdbproject/standards/>.

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 in 1996 reviewed 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 was circulated for internal review in late 1997. 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://ncgmp.usgs.gov/fgdc_gds/mapsymb/> for the document itself and for information about the review process). This draft standard is described in some detail in Soller and Lindquist (2000). With assistance from a Standing Committee to oversee resolution of review comments and long-term maintenance of the standard, the document is being revised for submittal to FGDC for discussion and adoption as a Federal standard.

Digital Mapping

The Data Capture Working Group has coordinated six annual “Digital Mapping Techniques” workshops for state, Federal, and Canadian geologists, cartographers, managers, and industry partners. These meetings have been highly successful, and have resulted in adoption within agencies of new, more efficient techniques for digital map preparation, analysis, and production. The most

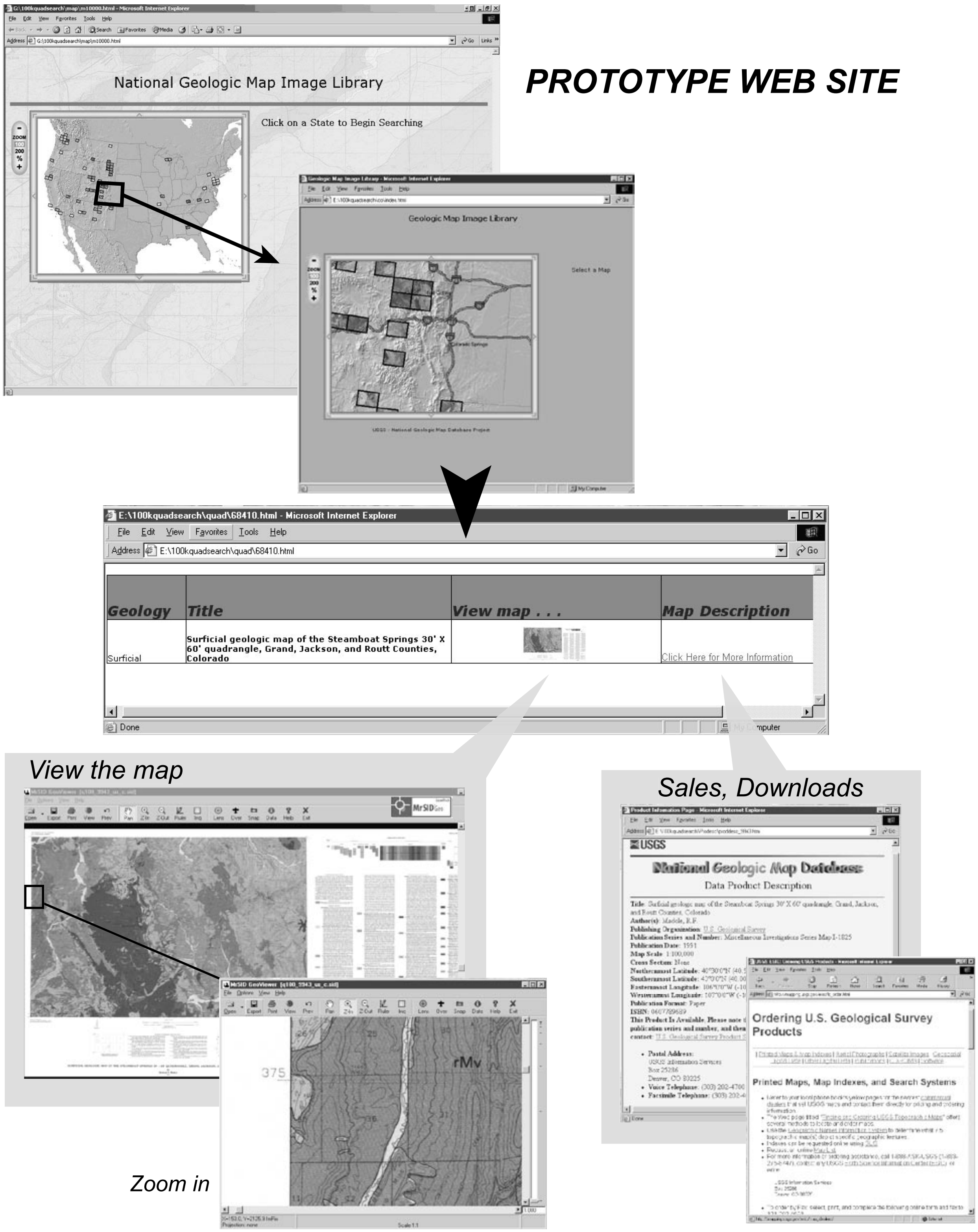


Figure 5. Diagram showing how users could access the proposed Image Library to view high-quality images of published bedrock- and surficial-geology maps. The user would select a state of interest and then a quadrangle within that state; that selection would generate an information table that listed the available map products. For each map product, the table would provide links to the image ("View map . . .") and to the map's record in the Map Catalog ("Map Description"). A user interested in purchasing a paper copy or downloading a digital version of the map can easily do so because the Map Catalog page includes the necessary information and links.

recent workshop, held in Salt Lake City, Utah, and hosted by the Utah Geological Survey, was attended by approximately 100 representatives of 42 state, Federal, and Canadian agencies and private companies. The workshop proceedings have been published (Soller, 1997, 1998, 1999, 2000, 2001, and in this volume) and are online at (<http://ncgmp.usgs.gov/pubs/of97-269>; <http://pubs.usgs.gov/openfile/of98-487>; <http://pubs.usgs.gov/openfile/of99-386>, <http://pubs.usgs.gov/openfile/of00-325>, <http://pubs.usgs.gov/openfile/of01-223>, and <http://pubs.usgs.gov/openfile/of02-370>. Published copies of the Proceedings may be obtained from Soller or Berg.

Map Publication Requirements

Through the USGS Geologic Division Information Council, Soller led development of the USGS policy "Publication Requirements for Digital Map Products," enacted May 24, 1999. A less USGS-specific version of this document was developed by the AASG/USGS Data Information Exchange Working Group and presented for technical review at a special session of the DMT '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.

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:// ncgmp.usgs.gov/ngmdbproject/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

State and USGS collaborators on the NGMDB continue to serve as representatives to the North American Data Model Steering Committee (NADMSC), assisting in the process of developing, refining, and testing the North American Geologic Map Data Model. The NADMSC has now formed various technical teams to conduct specific tasks within a one-year period and longer time-frames. If interested, please visit the NADMSC web site at <http://geology.usgs.gov/dm/>. More information is provided in these Proceedings in the paper by the NADMSC's Data Model Design Team.

Access to Online Products

As standards are developed under Phase Two and via other mechanisms, the products released by geological surveys increasingly are standardized in format and content. A principal goal of Phase Two is to provide links from the Map Catalog to the more standardized of these products. Through searches of the NGMDB Map Catalog, users are directed to Web sites for perusal of selected online products. This feature of the Map Catalog is now available for USGS products served on USGS Regional Publications servers and on the USGS Clearinghouse node, and for state geological survey products. At the time of the DMT '02 meeting, more than 900 links exist to online map products and their metadata.

PHASE THREE AND INTEGRATION WITH OTHER ASPECTS OF THE PROJECT

Over the past few decades, significant advances in computer technology have permitted complex spatial information to be stored, managed, and analyzed to the satisfaction of 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 database. Further, 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 of delivering useful information to the general public. Now, six years after the project's inception, there exists sufficient digital geologic map data; sufficient convergence on standard data formats, data models, digital mapping practices, and field data capture techniques; and sufficient technological advances in Internet delivery of spatial information to warrant a research effort aimed at building a prototype, online National Geologic Map Database.

To design an online database, project personnel have 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 is 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 (i.e., a "living" database),
4. standardized, adhering to a standard data model and with standard scientific terminology, and
5. available to users via Internet browsers and common GIS tools (e.g., ArcExplorer).

The NGMDB project has begun a series of prototypes designed to build this online, "living" database; an introduction to the design approach is given in Soller and others (2000). In 1999, we designed some basic requirements for a prototype geologic-map database and tested our concepts using some newly developed digital data for the Greater Yellowstone Area of Wyoming and Montana (Wahl and others, 2000). That first prototype was presented for discussion at the Geological Society of America annual meeting in October 1999. The prototype was well-received, and plans were begun for a second prototype, with a more complex set of tasks. That prototype, conducted in cooperation with the Kentucky Geological Survey, is summarized in Soller and others (2001). We anticipate further prototypes that will advance our understanding of the technical and management challenges to be addressed in development of the operational system.

The online map database is being designed to integrate with other databases developed under this 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 the history of the unit's geologic name. These user questions exploit the power and flexibility of the databases, and we anticipate building into the system the functionality diagrammed in Figure 6. As another example of the interaction of the various NGMDB phases, this diagram shows that a user might access the Map Catalog and identify a map of interest; the user might then purchase the map or link to a map server where the product can be downloaded. In the latter case, the arrow passing through "Standards Development" indicates that the NGMDB project's standards-development activities affect the content and format of products served.

The Geologic Map of North America

The NGMDB has supported development of the digital version of the Geological Society of America's (GSA) Decade of North American Geology Geologic Map of North America (GMNA). The project has 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 discussion regarding development of a GMNA map database. In the coming year, we will be prototyping the database design and beginning to populate it, in collaboration with GSA and interested national geological surveys.

ACKNOWLEDGMENTS

The authors thank the members of the NGMDB project staff and collaborators for their enthusiastic and expert support, without which the project would not be possible. In particular, we thank: Nancy Stamm (USGS, Reston, VA; GEOLEX database, and general support and guidance to the project), Ed Pfeifer, Alex Acosta, Dennis McMacken, Jana Ruhlman, and Michael Gishey (USGS, Flagstaff, AZ; Web-site and database management), and Chuck Mayfield and Nancy Blair (USGS Library; Map Catalog content).

REFERENCES

Soller, D.R., ed., 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., <http://ncgmp.usgs.gov/pubs/of97-269/>.

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

Soller, D.R., ed., 1999, Digital Mapping Techniques '99—Workshop Proceedings: U.S. Geological Survey Open-File Report 99-386, 216 p., <http://pubs.usgs.gov/openfile/of99-386/>.

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

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

Soller, D.R., and Berg, T.M., 1995, Developing the National Geologic Map Database: Geotimes, v. 40, no. 6, p. 16-18.

Soller, D.R., and Berg, T.M., 1997, The National Geologic Map Database—a progress report: Geotimes, v. 42, no. 12, p. 29-31.

Soller, D.R., and Berg, T.M., 1998, Progress Toward Development of the National Geologic Map Database, *in* Soller, D.R., ed., Digital Mapping Techniques '98—Workshop proceedings: U.S. Geological Survey Open-File Report 98-487, p. 37-39, <http://pubs.usgs.gov/openfile/of98-487/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., eds., Accreting the continent's collections: Geoscience Information Society Proceedings, v. 29, p. 47-55, <http://ncgmp.usgs.gov/ngmdbproject/reports/gisproc98.html>.

Soller, D.R., and Berg, T.M., 1999b, The National Geologic Map Database—a progress report, *in* Soller, D.R., ed., Digital Mapping Techniques '99—Workshop Proceedings: U.S. Geological Survey Open-File Report 99-386, p. 31-34, <http://pubs.usgs.gov/openfile/of99-386/soller1.html>.

Soller, D.R., and Berg, T.M., 2000, The National Geologic Map

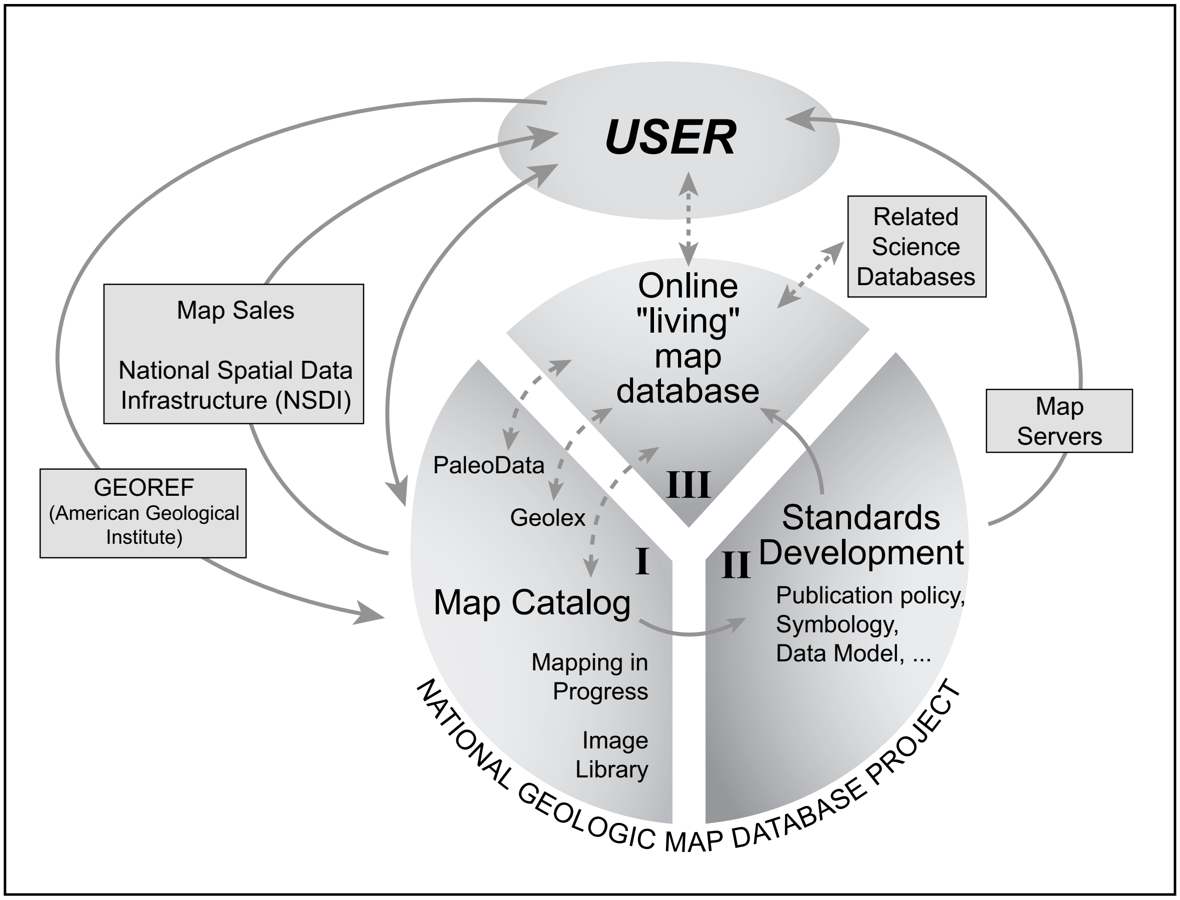


Figure 6. Diagram showing user access to the various components of the National Geologic Map Database (NGMDB) project and to related external databases and services. The three project phases and the relations among them are shown. Dashed arrows indicate planned relations.

Database—a progress report, *in* Soller, D.R., ed., Digital Mapping Techniques '00—Workshop Proceedings: U.S. Geological Survey Open-File Report 00-325, p. 27-30, <http://pubs.usgs.gov/openfile/of00-325/soller2.html>.

Soller, D.R., and Berg, T.M., 2001, The National Geologic Map Database—a progress report, *in* Soller, D.R., ed., Digital Mapping Techniques '01—Workshop Proceedings: U.S. Geological Survey Open-File Report 01-223, p. 51-57, <http://pubs.usgs.gov/of/2001/of01-223/soller1.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, <http://pubs.usgs.gov/openfile/of00-325/soller4.html>.

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, <http://pubs.usgs.gov/openfile/of99-386/soller2.html>.

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, <http://pubs.usgs.gov/openfile/of00-325/soller3.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, <http://pubs.usgs.gov/of/2001/of01-223/soller2.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, <http://pubs.usgs.gov/openfile/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., ed., Digital Mapping Techniques '01—Workshop Proceedings: U.S. Geological Survey Open-File Report 01-223, p. 59-70, <http://pubs.usgs.gov/of/2001/of01-223/wardlaw.html>.

The I-Team Initiative and Geology Framework Layers—Making Sure Geology Is Included

By Patrick J. Kennelly[1] and Grant C. Willis[2]

[1]Montana Bureau of Mines and Geology
Montana Tech of the University of Montana
Butte, MT 59701-8997
Telephone: (406) 782-8633
Fax: (406) 496-4451
e-mail: pkennelly@mtech.edu

[2]Utah Geological Survey
P.O. Box 146100
Salt Lake City, UT 84114-6100
Telephone: (801) 537-3355
Fax: (801) 537-3400
e-mail: grantwillis@utah.gov

INTRODUCTION

Most geological surveys are experienced in preparing geographic information system (GIS) data for their key geology-based customers. However, the GIS world is much bigger than just the geologic community. In fact, geologic data are not even considered one of the key layers in most federal, state, and local GIS databases. In our discussions with other digital geologic data producers, we find that geological surveys often fail to be an active part of broad-based GIS consortiums. This commonly results in missed opportunities to secure funding to produce and maintain geologic data as well as missed opportunities to assure that geology is properly included in government and land-manager GIS databases. These databases are generally the primary source of geographic information, including geology based, that are used in important policy and land-management decisions that significantly impact the general public, as well as the geologic community.

The U.S. President's Management Council recently identified 23 federal government initiatives that relate to Electronic Government (E-Gov). Of these, four have an important GIS component: Homeland Security, The National Map, Geospatial One-Stop, and Implementation Team (I-Team). In this paper, we briefly summarize the first three initiatives, but focus on the I-Team Initiative, which has the most application to the science of geology.

Homeland Security Initiative

The Homeland Security Initiative has the following goals:

- Interoperability of the systems that process homeland security information
- Commonality of the processes that collect, manage, and disseminate geospatial information
- Implementation of a comprehensive national security spatial data infrastructure

Much of the Homeland Security effort using GIS has been related to the group of cities referred to as "120 Cities for Domestic Preparedness." The U.S. Geological Survey (USGS) is currently involved in a broad effort to collect GIS data such as orthoimagery over these key locales.

The National Map Initiative

The goal of The National Map Initiative is a seamless, continuously maintained set of geographic base information across the United States. This base information would be similar to layers that currently appear on paper USGS topographic base maps. Whereas the average age of these paper base maps is 23 years, the vision of The

National Map Initiative is to provide updates within seven days of approving the change.

One current pilot project of The National Map Initiative has a geologic component. The Missouri Pilot Project is working in cooperation with the National Cooperative Geologic Mapping Program to support multihazard risk assessment, mitigation, and emergency planning. The study area comprises four USGS 7.5-minute quadrangles (scale 1:24,000) with flooding and seismic risk. Geologic data will be integrated with updated National Map Initiative base themes. For an example, see USGS Web site <http://pubs.usgs.gov/of/2001/of01-223/bradford.html>.

Geospatial One-Stop Initiative

The Geospatial One-Stop Initiative goals are to

- provide an interactive index to geospatial data holdings at the federal and nonfederal levels,
- initiate interaction between federal, state, and local agencies about existing and planned spatial data collections,
- provide an online access point to geospatial data, and
- provide standards and models for the geospatial framework data content

I-Team Initiative

The Implementation-Team (I-Team) Initiative, the focus of this paper, is envisioned as a "bottom-up" approach. Its goal is to implement the construction of key framework layers through interagency cooperation and partnerships. These layers will compose the National Spatial Data Infrastructure, and the contributing agencies will work in cooperation with the Federal Geographic Data Committee and the Federal Office of Management and Budget (OMB). I-Teams will be separate from, but work in cooperation with, technology advisory groups and financing solution teams (Figure 1).

Interaction Among These Initiatives

These four initiatives incorporate GIS into E-Gov with considerable overlap. For example, to create base layers for The National Map, the framework layers produced through I-Team activities will be needed. These layers in turn will affect standards, models, and distribution processes related to Geospatial One-Stop.

I-TEAMS

Team Makeup and Goals

Most I-Teams are constructed at the state level; however, they also can be constructed at a multistate level (for example, Rocky Mountain states), or around a major theme (for example, Colorado River users). The central concept of an I-Team is to assemble all GIS data producers, suppliers, and users who have an interest in the geographic area of concern. Team members commonly consist of representatives of state agencies (transportation, agriculture, public safety, governor's planning council, school trust land administrations, and geological surveys, among others); federal agencies (Forest Service, Park Service, Bureau of Land Management, military, and others); local governments (city and county units, water districts, law enforcement agencies, and others); academia (university geography, geology, and other departments; school districts; and others); and private interests (gas and electrical utilities, pipeline companies, communications companies, and others). There is probably an I-Team in your state. State geological surveys may need to actively seek out and join I-Teams. Most states have an agency that is designated as the coordinating agency for GIS information that should be aware of any I-Teams in your area.

The I-Team members first develop a list of key "framework" layers for their database and identify the most likely sources for each layer. The team then develops a plan to assemble and maintain the database, assigns or recommends duties to team members, and encourages the formation of cooperatives to streamline data assembly, coordinate efforts, and increase efficiency.

Generally, I-Teams begin with the seven framework layers identified by the Federal Geographic Data Committee: cadastral (public land surveys and ownership), elevation, geodetic control, government units, hydrography, digital ortho-imagery, and transportation. I-Teams then typically identify additional key framework layers. For example, because the Montana Geographic Information Council also recognizes geology, hydrologic units, soil, and land use/land cover as priority themes, these themes were included in Montana's I-Team Initiative. In the Utah Framework Implementation Plan, demographics, wetlands, geology, wildlife habitat, climate, ground cover, land use, soils, telecommunications infrastructure, critical facilities and infrastructure, and environmental hazards were identified as additional key themes.

Montana I-Team

In Montana, four statewide groups are concerned with GIS data. The Montana Geographic Information Council was created by the Governor in 1997 and provides policy-level direction and promotes efficient and effective use of geographic information. The Montana Interagency GIS Work Group acts as a forum for the exchange of information regarding the acquisition of new geospatial data, the existence of current geospatial data, and information relating to agency geospatial projects. The Montana Interagency GIS Work Group encourages

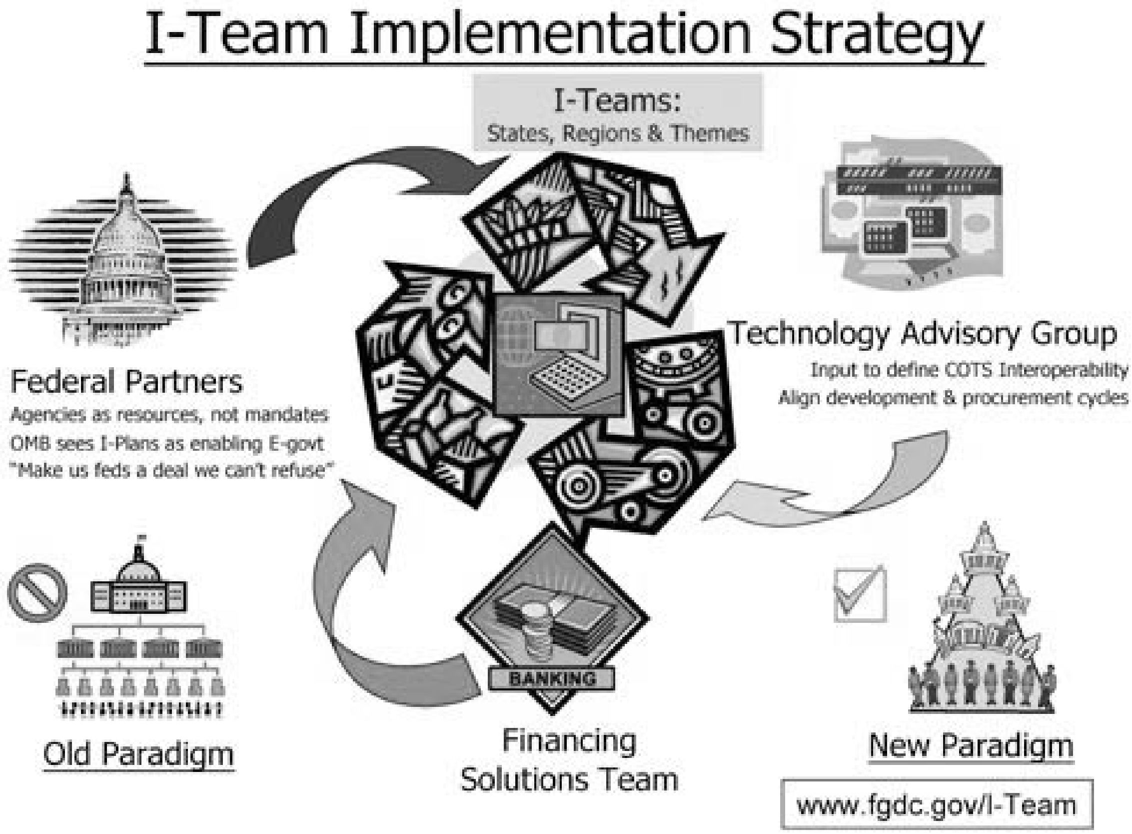


Figure 1. Schematic of I-Team implementation strategy (from <http://www.fgdc.gov/I-Team>).

agencies to minimize the duplication of digital data, implement transfer technologies for the exchange of data, develop data standards, and share resources in completing interagency projects. The Montana Local Government GIS Coalition was initiated by local government GIS practitioners in July 1995 to facilitate and advance the implementation and development of GIS technology in city and county government through communication and data sharing. The Montana GIS User's Group provides opportunities for education, training, and conferences.

Utah I-Team

In Utah, the Geographic Information Systems Advisory Council, which encompasses federal, state, and local government, academia, and the private sector, has led the statewide data coordination effort under a joint agreement forged in 1997. At that time the state and nine federal agencies signed a memorandum of understanding to cooperate in the production, acquisition, and dissemination of GIS data, creating the Utah Framework Implementation Team (Utah I-Team). Over the next two years, the Utah I-Team defined the 18 priority layers listed in Table 1, assigned responsibilities, set goals, developed budgets and schedules, and set standards. They also established the primary target scale for all data as 1:24,000. In the process of setting goals and working with the agencies they represent, they identified many of the State's most serious issues. The Utah I-Team then determined which data themes are required to successfully address each issue. This analysis is summarized in Table 1. The first three issues in the table represent the Governor's top priorities.

BUILDING THE GEOLOGY FRAMEWORK LAYER

In explaining how to build a geology framework, we will use the Utah plan as our example. The Utah Geological Survey (UGS) has been a member of the Utah I-Team from its inception, but did not play a major role until early 2001. At that time, as the Utah I-Team plan gradually took shape, the UGS was requested to develop a detailed model and plan for the geologic framework layer.

Prior to UGS involvement, the Utah I-Team viewed the geologic layer as simply consisting of one seamless data layer that they referred to as "surface geology" and assumed that it would be at a scale of 1:24,000. As the UGS staff became more involved, they pointed out that at current rates of mapping, it will require 100 years to complete 1:24,000-scale geologic map coverage of the state. The UGS instead recommended a more complex, but more practical system of three layers of geologic map

Table 1. Relationship between priority data layers and pressing ecological, political, and social issues in Utah.

UTAH ISSUES / **PRIORITY DATA LAYERS**	Geodetic Control	Digital Ortho-Imagery	Elevation	Transportation	Hydrography	Boundaries	Cadastral	Demographics	Wetlands	**Geology**	Wildlife Habitat	Climate	Ground Cover	Land Use	Soils	Telecommunications	Critical Facilities/Infrastructure	Environmental
Economic Development	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Education/Enrollment	X			X		X	X	X						X		X		X
E-Gov Service Delivery	X	X	X	X	X	X	X	X								X		
Olympics	X	X	X	X	X	X	X	X	X			X		X		X	X	X
Rural Economies	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Quality Growth	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Hazardous/Nuclear Waste	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X
Open Space/Agriculture	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X			X
Environmental Protection	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X			X
Emergency Management	X	X	X	X	X	X	X	X		X						X	X	X
Public Lands Management	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Law Enforcement	X	X	X	X		X	X	X								X	X	
Traffic/Transportation	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X
Redistricting/Census	X	X		X		X	X	X								X	X	
Epidemiology/Health Care	X	X		X		X	X	X								X	X	X
Social Services	X	X		X		X	X	X								X	X	
E911	X	X	X	X	X	X	X	X								X	X	

Table 2. The Utah I-Team Geology Framework Layer developed in early 2001.

Geologic maps
- Geologic map of Utah (scale 1:500,000; was complete in GIS format)
- Geologic maps of 30 x 60 minute quadrangles (scale 1:100,000; about 10% were in GIS format)
- Geologic maps of 7.5-minute quadrangles (scale 1:24,000; none were in GIS format)

Geologic hazards (most were in GIS format)
- Seismicity maps (includes database of historic earthquakes)
- Landslide maps (various scales)
- Quaternary fault map (scale 1:500,000, plus detailed maps of some areas)

Geologic resources
- Large georeferenced database of oil, gas, tar sand, oil shale, coal, metallic resources, nonmetallic resources, and data on various other commodities

data: 1:500,000 scale, which was already available in GIS format; 1:100,000 scale, which is the primary focus of their STATEMAP geologic mapping program; and 1:24,000-scale mapping in specific areas with high-priority needs (primarily centered around rapid urban growth). In addition, they recommended that geologic hazards and economic geology be added as parts of the core geology layer. Thus, the I-Team eventually decided to divide the geology layer into three parts (geologic maps, geologic hazards, and economic geology), with seven sublayers (Table 2). This plan did not include every type of geologic data, but instead struck a balance between the overly simplistic map first conceived and the endless "wish list" that a group of geologists could conjure up.

CONCLUSIONS AND RECOMMENDATIONS

The I-Team framework plan addresses the needs of the GIS community in most states, multistate associations, and geographic-oriented groups. I-Teams are commonly looked upon as the primary source of GIS data for most government policy-making and management decisions. Commonly, I-Teams are the loudest voice in legislative and other decision-making circles, influencing funding and setting local and state GIS standards and procedures. Although geology is only a small part of this much larger GIS community, we believe that it is important for geological surveys and other geology groups to actively seek out, join, and then promote geologic issues on I-Teams in their state or region.

In defining and creating geology layers for your I-Team framework plan, we recommend the following:

- Remember that you are part of a much larger group, most of whom do not understand geologic principles or purposes. You will need to patiently explain geology's role and importance many times, and at the same time listen supportively as others explain the role of their respective disciplines.
- Keep your framework layer simple. You can't include everything. The framework most likely will include only the major layers that have the broadest application. Certainly, most state geological surveys have many additional types of data (such as geophysical maps, alteration maps, and paleontologic data) that they may want to put in GIS format. Such information can be completed and released outside of the realm of the framework plan.
- Build on your existing goals, plans, and databases. The I-Team framework plan is not intended to make you change your focus. Instead, it is your opportunity to assure that geology is included in the larger multi-interest GIS community.

ACKNOWLEDGMENTS

We thank the members and staff of the National Cooperative Geologic Mapping Program who have advanced the development of geologic databases at local, state, and national levels. We also thank the many people who developed and worked behind the scenes to make the 2002 Digital Mapping Techniques Workshop in Salt Lake City, where this paper was first presented, a success. Finally, we thank our reviewers, Edmond Deal at the Montana Bureau of Mines and Geology, and Kent Brown, Robert Biek, and William Case at the Utah Geological Survey.

MRLC2000 Image Data and Geologic Mapping

By Ronald R. Wahl and David A. Sawyer

U.S. Geological Survey
Box 25046, Denver Federal Center, MS980
Denver, CO 80225
Telephone: (303) 236-1320, (303) 236-1021
Fax: (303) 236-5349
e-mail: {rwahl, dsawyer}@usgs.gov

Geologists traditionally have used aerial photography to help make geologic maps. Satellite imagery has been available to the geologic community only rarely, due to cost. Some investigations could afford to purchase such information but, commonly, good, clear, high-resolution satellite imagery was not available at any price. With the advent of the U.S. Geological Survey's Multi-Resolution Land Characteristics (MRLC) 2000 (now MRLC 2001) image data, these difficulties have been overcome for maps of 1:100,000 scale and smaller. These 30-meter-resolution data should also help at larger scales when sharpened by the included 15-meter-resolution panchromatic band.

BACKGROUND

Geologists from the U.S. Geological Survey (USGS) have for many years used imagery in one form or another to capture the information they thought interesting and important while investigating the geology of an area. Hayden (1883) in his 1878 report from the western part of the country used sketches and drawings as a way to portray the geology (including outcrops and fossils), vegetation, and the beginnings of Man's impact on the landscape (Figure 1). The image in Figure 1 has historical significance as well as scientific importance. The drawings in Hayden (1883) of the Pike's Peak, Colorado, area, Yellowstone National Park, and what would become Grand Teton National Park have great value in showing the state of the then-current environment as well as geologic phenomena.

At the first Digital Mapping Techniques workshop in Lawrence, Kansas, Wahl (1997) noted that the USGS was capturing geologic information by using aerial and oblique stereo photography in manual photogrammetry plotters. Since then, the USGS and others have acquired computer-based photogrammetry systems specifically for the capture of geologic information and are using imagery in these systems.

GEOLOGIC MAPPING AND IMAGERY

Geologic Information from Satellites

Geologic features are sometimes quite visible in satellite image scenes. Figure 2a is a small part of the *Geologic map of Yellowstone National Park* (U.S. Geological Survey, 1972) centered on Old Faithful geyser, and Figure 2b is approximately the same area from a Landsat 7 scene. The Landsat image is formed from bands 7, 4, and 1 (red, green, and blue). The hydrothermal alteration shows as various shades of blue (here as light gray to white). The intermediate gray patches are burn areas from the 1988 fire.

Geologic Map Production

Figure 3 shows a 7½-minute quadrangle produced from Landsat imagery. The geologic features were captured almost entirely from the image. More geologic maps are going to be compiled in the future using image data because the maps can be produced in a more timely manner and because the MRLC 2001 imagery will be able to show geologic features more clearly that other image formats, especially in semi-arid and arid regions.

Map Scales and Imagery

National Map Accuracy Standards state that the resolution of map features for 1:100,000-scale maps is about 50 meters. This means that because of line widths and the combination of thematic layers on a 1:100,000-scale map, features smaller than 50 meters in extent can't be drawn on such a map. Landsat image bands 1 through 5 and band 7 have a pixel size of 30 meters. Sharpened with a 15-meter-pixel panchromatic image, Landsat 7 data are quite usable at 1:100,000 scale. Larger scale maps may gain some advantage using such data, but feature resolution would be poorer.

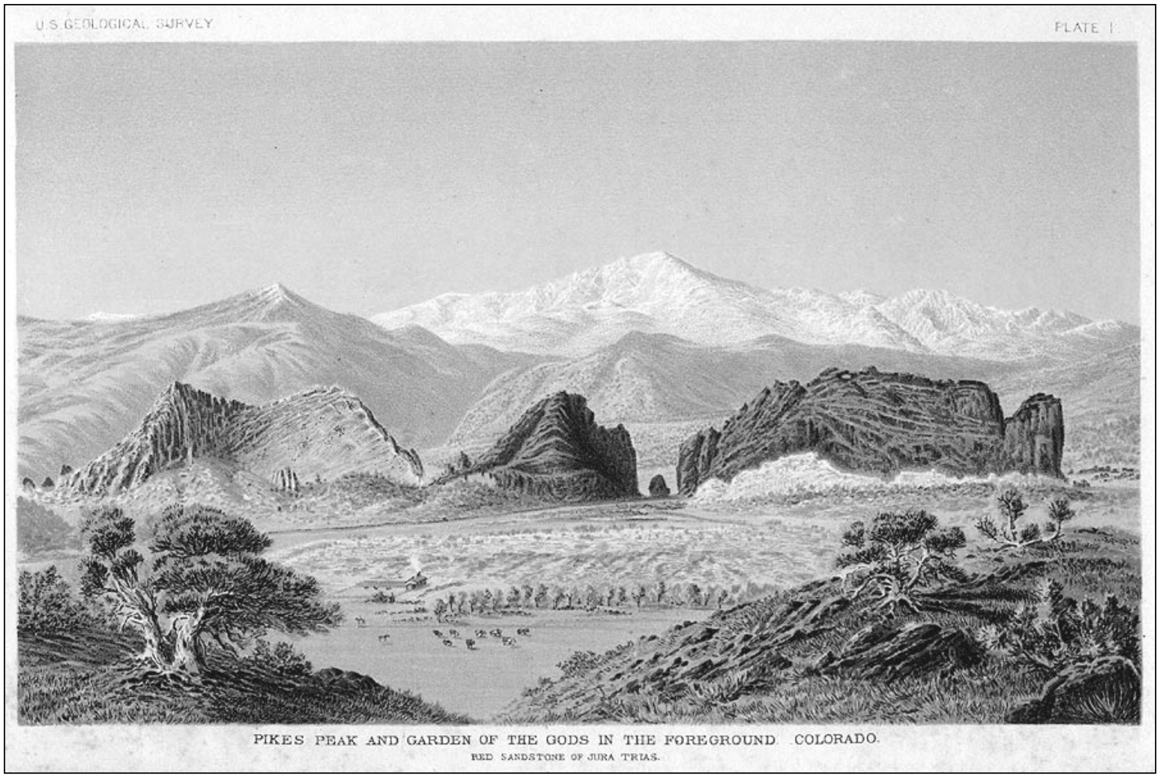


Figure 1. Pike's Peak and the Garden of the Gods, Colorado.

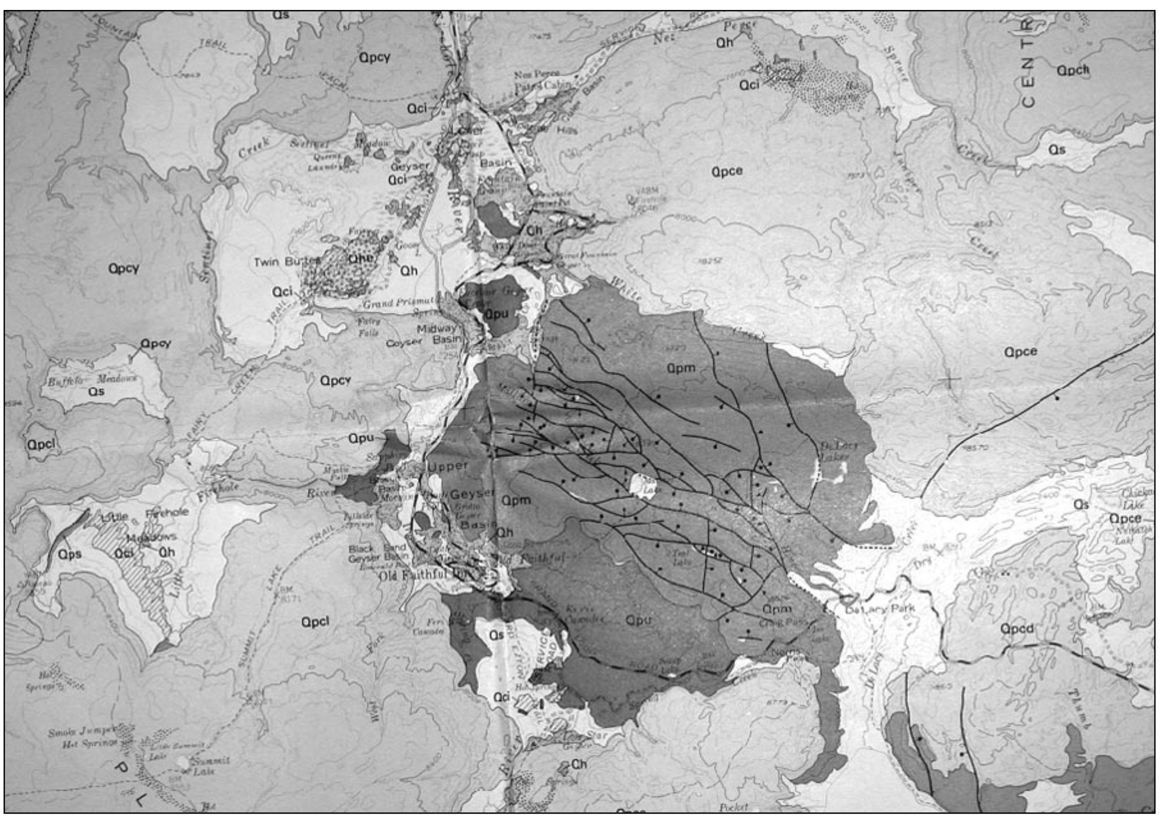


Figure 2a. A portion of the geologic map of Yellowstone National Park near Old Faithful Geyser.

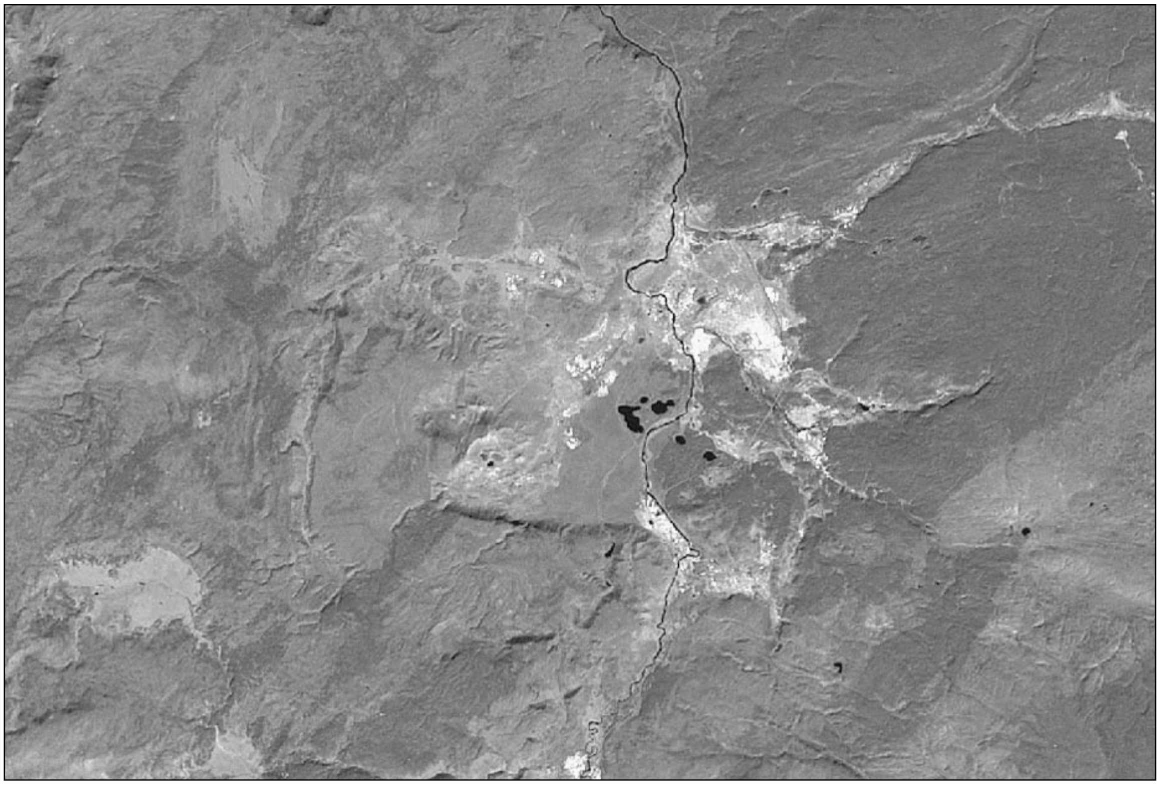


Figure 2b. A portion of a Landsat image of roughly the same area as figure 2a.

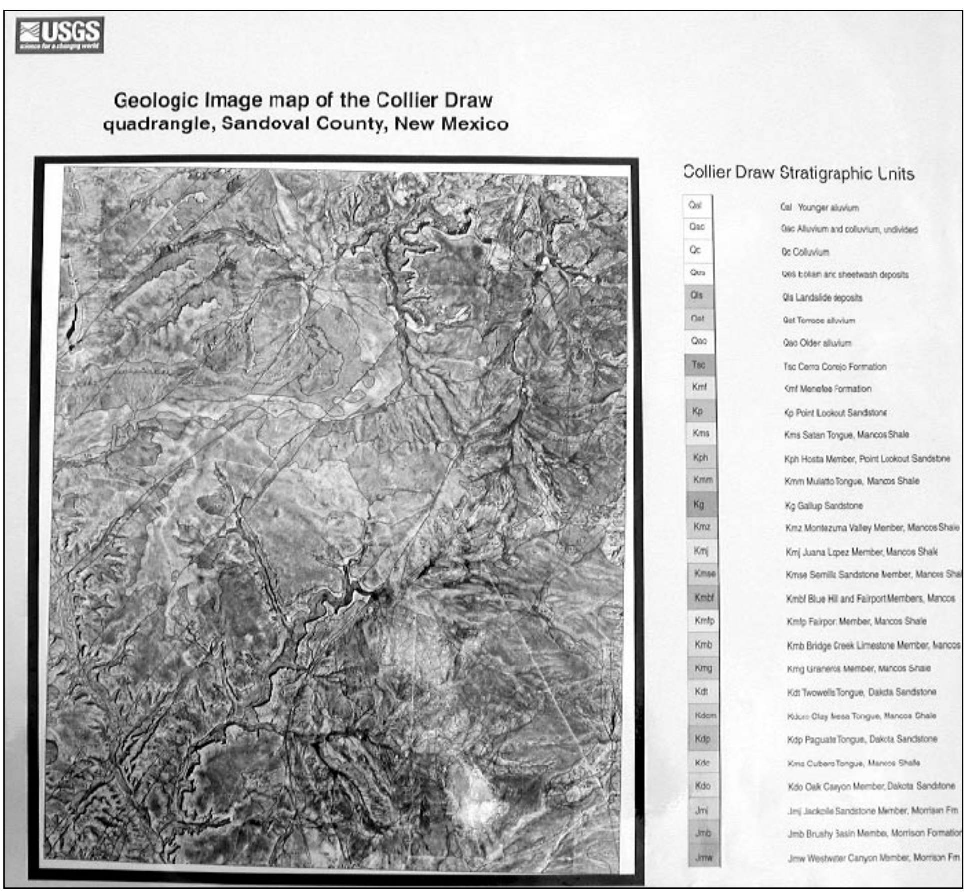


Figure 3. A geologic map produced primarily from Landsat imagery.

MRLC2001 DATA AND GEOLOGY

MRLC 2001 data are terrain corrected and georeferenced Landsat 7 data ("Landsat 7+") that contain the following information:

Bands 1-5 and 7 visible to long infrared, 30-meter pixels
Band 6 thermal, 60-meter pixels
Band 8 panchromatic (black & white), 15-meter pixels
Band 9 digital elevation model (DEM) data, 30-meter cells

The DEM (band 9) data that cover the image area were extracted from the National Elevation Dataset (NED) and were the basis for terrain correction. Band 9 is composed of 16-bit (2-byte) integers rather than the 8-bit (1-byte) integers in the other bands. Band 9 is recorded with the bytes in reverse order from the order that a PC would expect.

The following description is from the current fact sheet for the new radiance-corrected data sets on the MRLC 2001 Web page:

MRLC 2001 Terrain Corrected/Radiance Adjusted Dataset Description

Multi-Resolution Land Characterization 2001 (MRLC 2001) [formerly MRLC2000] *is a second-generation federal consortium to create an updated pool of nation-wide Landsat imagery, and derive a second-generation National Land Cover Database (NLCD 2001).*

One of the challenges to large-scale satellite based land cover characterization is consistent geometric correction and normalizing noise arising from atmospheric effect, changing illumination geometry, and instrument errors inherent when using multiple frames of imagery. For those reasons, we have created a product with additional processing from the Terrain Corrected scenes.

This product is designed for regional applications where scene mosaics can be substantially improved by converting DN [digital number] *to at-satellite reflectance to correct for Sun illumination angle effect. At-satellite reflectance images should be more appropriate for land cover and land cover change analysis than the original DN images. At-satellite reflectance also allows the development of a regionally applicable tasseled cap transformation using a physically based measurement. Overall, this method provides an important first-step to standardizing imagery, but users should realize that atmospheric, phonological, and topographic noise do remain. Atmospheric correction is not considered as a standard step in MRLC image processing . . . .*

MRLC 2001 Landsat 7 data are currently available to all users. The data are to be used or disseminated with the intent of use for scientific purposes and only for a non-commercial venture.

This imagery is available to users in the USGS and cooperating agencies (e.g., the state geological surveys that participate in STATEMAP or other projects). MRLC2001 data can be obtained via the MRLC Web site at <http://edc2.usgs.gov/lccp/mrlc2k/mrlc2k.asp>, or they can be located via a search engine such as Google by typing "MRLC2001 data."

The cooperating agencies for the MRLC 2001 data are: USGS, U.S. Environmental Protection Agency (USEPA), U.S. Forest Service (USFS), National Oceanographic and Atmospheric Administration (NOAA), National Aeronautics and Space Administration (NASA), National Park Service (NPS), National Resources Conservation Service (NRCS), and the Bureau of Land Management (BLM). Our understanding is that any federal, state, or local agency that works with one of the above cooperating agencies can obtain this imagery.

USES OF MRLC 2001 DATA IN THE GREATER YELLOWSTONE AREA

Currently, we are compiling the 1:100,000-scale geologic maps for the 30 x 60 minute quadrangles that cover Yellowstone National Park. Most of the field mapping was done over the past 35 years. This effort will tie together new 1:100,000-scale mapping across the park. We acquired the MRLC 2001 data for the area to help in this compilation. The ARC/INFO coverages for the current park map (Christiansen and Wahl, 1999), along with coverages from the current compilation, were plotted over the image data. Figure 4 shows the image data with the new compilation work in white and the Yellowstone National Park geologic map coverage in light gray. This figure clearly shows relationships between the geologic features on the maps and the imagery even though the vegetation is quite heavy. Study of the imagery with the map coverage should yield a better compilation product.

CONCLUSIONS

Why Use These Data?

Because MRLC 2001 image data are terrain corrected and georeferenced, it is like an orthophoto but with more information from the large number of image bands. Each band or combination of bands can be converted to a GeoTiff image and then used in a program such as Adobe Photoshop. In arid or semi-arid regions and when merged with the 15-meter band-9 data, MRLC 2001 imagery may be useful for 1:24,000-scale mapping.

The cost of the MRLC 2001 data is quite reasonable. Each scene is $45 with a $45 processing fee per order. MRLC 2001 data is available for most of the lower 48 states and more images are added each week.

Figure 4. ARC/INFO coverages on top of Landsat 7 imagery of Yellowstone National Park.

Difficulties in Using the Data

Several difficulties arise from trying to use Landsat 7 images as processed into MRLC 2001 data. Most significant is that they are available only in the USGS National Landsat Archive Production System (NLAPS) format. Not all image-processing software are able to process this data format yet, and significant effort is required to process the data into a useable format.

File sizes are quite large. All bands of one MRLC2001 scene are stored on two CD-ROM's. Data transfers of merged scene data (more that one image scene and all of the bands) may require either tapes or DVD's to keep the data on one piece of media. To use multi-scene images requires exceptional processing power from a PC.

The third problem is that this is a new area of exploration for geologists. The learning curve can be long and steep, but the new source of data will greatly add to the preparation of geologic maps and the result will be well worth the effort.

The Future of Imagery and Geologic Mapping

In the future, geologists will have many data and tools at their disposal, including more sensitive and higher resolution imagery. For example, one of the newer image data set types is ASTER (Figure 5). The following is quoted from the home page of the ASTER Web site at <http://asterweb.jpl.nasa.gov/>:

ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) is an imaging instrument that is flying on Terra, a satellite launched in December 1999 as part of NASA's Earth Observing System (EOS). ASTER is a cooperative effort between NASA and Japan's Ministry of Economy, Trade and Industry (METI) and the Earth Remote Sensing Data Analysis Center (ERSDAC). ASTER will be used to obtain detailed maps of land surface temperature, emissivity, reflectance, and elevation. The EOS platforms are part of NASA's Earth Science Enterprise, whose goal is to obtain a better understanding of the interactions between the biosphere, hydrosphere, lithosphere, and atmosphere.

REFERENCES

Christiansen, R.L., and Wahl, R.R., 1999, Digital geologic map for Yellowstone National Park, Idaho, Montana, and Wyoming and vicinity: U.S. Geological Survey Open-File Re-

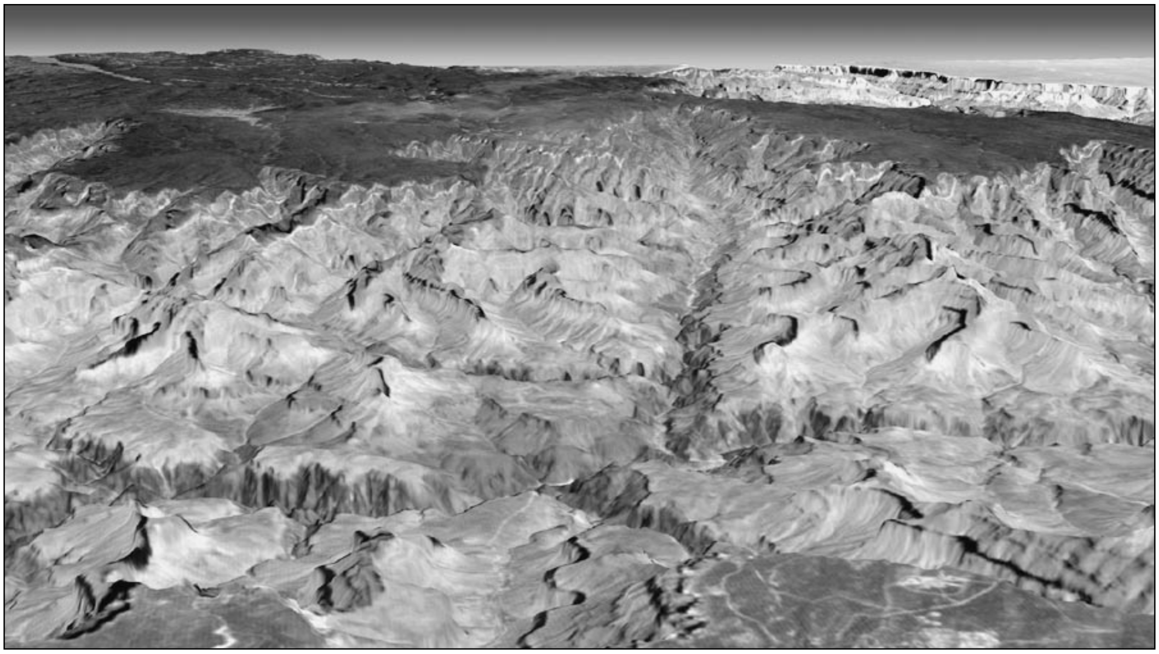


Figure 5. ASTER image of the Grand Canyon, Arizona, draped over a digital elevation model. The image has the rock units classified by lithology.

port 99-174, <http://pubs.usgs.gov/of/1999/ofr-99-0174/>.

Hayden, F.V., 1883, A report of progress of the exploration in Wyoming and Idaho for the year 1878: 12th Annual Report of the U.S. Geological and Geographical Survey of the Territories, pt. 1, 809 p., pt. 2, 503 p.

U.S. Geological Survey, 1972, Geologic map of Yellowstone National Park: U.S. Geological Survey Miscellaneous Geologic Investigations Map I-711, scale 1:125000.

Wahl, R.R., 1997, Digital geologic map data collection using photogrammetric and raster methods *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. 95-98, <http://pubs.usgs.gov/of/of97-269/wahl.html>.

The National Park Service Geologic Resources Inventory: An Update

By Tim Connors,[1] Stephanie A. O'Meara,[2] Trista L. Thornberry,[2] Joe D. Gregson,[3] Anne R. Poole,[4] and Bruce Heise[1]

[1] National Park Service–Geologic Resources Division
12795 West Alameda Parkway
Lakewood, CO 80228
Telephone: (303) 969-2093
e-mail: Tim_Connors@nps.gov; Bruce_Heise@nps.gov

[2] Colorado State University/National Park Service Cooperator
1201 Oak Ridge Drive, Suite 200
Fort Collins, CO 80525
Telephone: (970) 225-3584
e-mail: Stephanie_O'Meara@partner.nps.gov; Tthorn@cnr.colostate.edu

[3] National Park Service–Natural Resources Information Division
1201 Oak Ridge Parkway, Suite 200
Fort Collins, CO 80525
Telephone: (970) 225-3559
e-mail: Joe_Gregson@nps.gov

[4] National Park Service–Intermountain Region
Intermountain GIS Support
Bandelier Hall West, Room 203
Albuquerque, NM 87131
Telephone: (505) 346-2885 extension 260
e-mail: Anne_Poole@nps.gov

GEOLOGIC RESOURCES INVENTORY

Since 1998, the National Park Service (NPS) has been conducting a Geologic Resources Inventory (GRI) to document and evaluate the geologic resources of 273 National Park System units (national parks, monuments, recreational areas, historic sites, seashores, lakeshores, etc.). The GRI is a cooperative endeavor; cooperators include the NPS Geologic Resources Division, NPS Inventory and Monitoring (I&M) Program (Natural Resource Information Division), U.S. Geological Survey (USGS), individual state geological surveys, and academic institutions.

User-friendly geographic information system (GIS) tools have been developed in ESRI ArcView 3.x and ArcInfo 8.x formats for the digital geologic maps. Applications, including the NPS-developed ArcView Theme Manager v.2.01, (<http://www1.nature.nps.gov/im/apps/thmmgr/index.htm>), graphical cross-section viewer, and legend text-display tools are integrated with a standard geology GIS model to reproduce the components of a "paper" geologic map into a digital geologic database. The always-evolving geology GIS model is based on the Washington State ArcInfo GIS data model (Harris, 1998), which is being adapted for ArcView GIS and extended to include components of the North American Geologic Map Data Model (NADM), <http://geology.usgs.gov/dm/>. For more detailed information on the GRI than is presented here, please see Fryer and others (2001).

PRODUCTS OF THE GEOLOGIC RESOURCES INVENTORY

Geologic Bibliographies

"GRBIB," the bibliography of existing geologic maps and literature for 235 NPS units, is available on the Internet at <http://www.nature.nps.gov/im/apps/npbib/> and is also prepared as printable documents at <http://www2.nature.nps.gov/grd/geology/gri/products/geobib/>. Geologic index maps showing the location of associated geologic maps and their scales have also been prepared for these parks. In general, after map coverage for each park is determined, map products can be evaluated, and if needed, additional mapping projects identified and initiated.

Park Workshop Meetings

GRI Park Workshops (scoping sessions) have been conducted for 67 parks in Colorado, Utah, Idaho, North Carolina, California, Texas, New Mexico, Arizona, the Dakotas, and the National Capital area to evaluate each park's geologic resources. As a result of these workshops, park teams have evaluated existing published geologic maps to see whether they have been translated into digital format and to identify where data gaps exist in geologic mapping. New geologic mapping may be initiated on a case-by-case basis after careful evaluation of needs, costs, potential cooperators, and funding sources.

Geologic Mapping and Digitizing Projects

The NPS GRI Program has cost-shared with the following state agencies:

- Utah Geological Survey for new geologic field mapping at Zion National Park (NP) and Glen Canyon National Recreation Area (NRA);
- North Carolina Geological Survey for new mapping along Blue Ridge Parkway;
- Minnesota Geological Survey at Voyageurs NP; and
- South Carolina Geological Survey at Kings Mountain National Historic Park.

Additional field mapping projects have been initiated or completed for the geologic maps for Bent's Old Fort National Historic Site (NHS), Curecanti NRA, Florissant Fossil Beds National Monument (NM), Great Sand Dunes NP, Capitol Reef NP, Cedar Breaks NM, Golden Spike NHS, and Natural Bridges NM.

The NPS Geologic Resources Inventory is being actively developed with the formal cooperation of USGS and state geological surveys. However, many opportunities for project collaboration may exist that have not yet been identified, and effective communication among cooperators is a key factor for success of the inventory.

Another challenge of inventory planning is the development of digital map standards that are adaptable to diverse geological conditions, but still provide quality, uniform products and firm guidance for map developers. Indeed, the diversity of geologic resources found in the NPS will provide a continuing challenge for effective project management. The I&M Program and Geological Resources Division are developing an efficient inventory program to expedite the acquisition of digital geologic information for NPS units throughout the country. The NPS is attempting to align these digital standards with those of the USGS and the North American Data Model.

Summary Geologic Reports

Upon completion of an inventory in a park, the available geological literature and data from the NPS, USGS, state, and academic institutions will be documented in a summary report. The content, format, and database structure of such reports are still being developed.

REFERENCES

Fryer, Steve, Gregson, Joe, Connors, Tim, Poole, Anne and Heise, Bruce, 2001, The National Park Service Digital Geologic Map Model: Transformation from Paper to Digital, Featuring Legends, Cross Sections, Map Notes and Keyword Searchability, *in* Soller, D.R., ed. Digital Mapping Techniques '01—Workshop Proceedings, U.S. Geological Survey Open-File Report 01- 223, p. 87-100, <http://pubs.usgs.gov/openfile/of01-223/fryer.html>.

Harris, Carl F.T., 1998, Washington State's 1:100,000-scale geologic map database: an ArcInfo data model example, *in* Soller, D.R., ed., Digital Mapping Techniques '98—Workshop Proceedings, U.S. Geological Survey Open-File Report 98-487, p. 27-35, <http://pubs.usgs.gov/openfile/of98-487/harris.html>.

Scanning and Delivery of Historic Maps Over the Web: The Library of Congress Experience

By Colleen R. Cahill, Gene E. Roberts, and Diane Schug-O'Neill

Geography and Map Division
Library of Congress
101 Independence Ave, SE
Washington, DC 20740-4652
Telephone: (202) 707-8530
Fax: (202) 707-8531
e-mail: {ccah, grob, dschu}@loc.gov

IN THE BEGINNING . . .

In the early 1990's, the Geography and Map Division of the Library of Congress recognized that digital cartography would be a vital part of its collections. With very little digital-mapping proficiency, it became apparent that expert advice was mandatory for the Division to continue its growth in a rapidly changing cartographic environment. Help was sought from the James Madison Council, a private-sector group created to serve as the Library of Congress' primary philanthropic link to the business community. The Council awarded the Division a $30,000 grant to organize an advisory group focused on the issues of how to collect and maintain digital cartographic materials.

In January 1995, the Geography and Map Division's Center for Geographic Information was founded, composed of leaders in the cartography and computer industries. On advice from the Center partners, electronic mapping was divided into two broad categories: "born digital" materials and the digitization of historic items; this paper focuses on the latter. The Center for Geographic Information proved a rich source of not only advice but also material support. Several members loaned or donated equipment and software: among these was software from ESRI and Macromedia Corporation, computer hardware from Hewlett Packard, and a large flatbed scanner from Tangent Imaging (Figure 1). It took several months to set up work space, gather and train staff, and begin scanning.

SCANNING IS A LEARNING EXPERIENCE

The first item scanned was a map by George Washington titled *A plan of my farm on Little Huntg. Creek & Potomk*, drawn in 1766 (Figure 2). This scan and subsequent ones provided the parameters that were used for this project: maps are scanned at 300 dots per inch, in red-green-blue 24-bit color and the raw TIFF file is archived.

Figure 1. Tangent flatbed scanner.

This produces a file with enough data to be useful without being so large as to be unwieldy and is in keeping with the goal of the project: to make research-quality images, not facsimiles. On the Tangent, a full-bed scan is 24 inches by 36 inches and creates a file approximately 240 to 260 megabytes. If the map is larger than the scanner surface, several scans are joined into one file to create a digital image that is as close as possible to the original map.

Because of this early work with digital mapping, the Geography and Map Division was in a strong position to be a major player in the Library's World Wide Web effort, the American Memory project. In a major commitment to provide 5000 maps, the Division began scanning historic materials from its collections, which consist of 4.6 million maps, 60,000 atlases, plus globes, relief models, and more. Whereas the American Memory project focused on materials about the United States, the Division anticipated the need for materials to support presentations on immigration history and patterns, and so early in the project the Division began scanning maps on locations outside the United States.

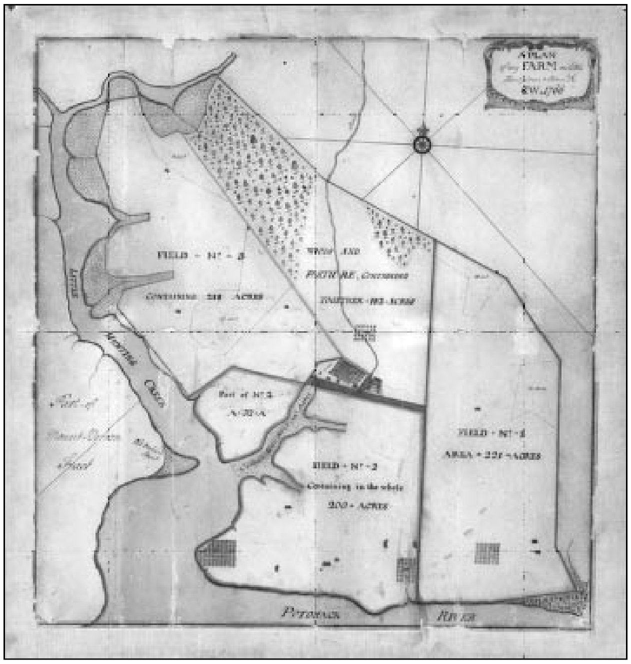


Figure 2. George Washington, 1766, *A plan of my farm on Little Huntg. Creek & Potomk.*

Once the tools for scanning were in place, a selection of what would be scanned was the next priority. There are several well-known cartobibliographies produced by the Division: the first set of maps scanned were those from the *Panoramic Maps of the United States and Canada*, compiled by John Hébert and Patrick Dempsey (1984). This scanning of published bibliographies was chosen as the criteria that are still used today: these are popular items, they are out of copyright, and they already have catalog records. The first two criteria are self explanatory; the third criterion reflects the fact that cataloging a map can become the most time consuming part of getting maps online.

Part of the learning curve in this project was setting up a smooth workflow, because putting a map on the Internet requires more than just pulling it from the collections and scanning it. It was quickly realized that the entire Geography and Map staff played a part in posting digitized maps on the Web. All staff provided advice on the selection of materials; the final decision was made by the Chief and collection curators. A collection technician would pull and prepare the item for scanning. The Digital Team then scanned the item, processed the image, moved the files to their appropriate location, gathered metadata, added or updated cataloging information to the Library’s database, and ensured all the pieces were in place before the image became available on the Internet. Collection curators created introductory descriptions for online presentations and assisted in determining Web navigational needs. This workflow slowly improved over several years. Even today, the process is being altered by changes in technology and the increased experience of the staff.

This entire process requires coordination and cooperation, not only within the Division, but with other staff in the Library. If a map needs repairs, it is sent to the Conservation Lab for expert treatment. The majority of the Web-page design is completed with the assistance of the National Digital Library staff, who create Web pages and graphics, edit text, and give technical support to make the Map Collections Web site a reality.

ONLINE AT LAST

The last hurdle to providing cartographic materials online was delivery through the Web. This problem was solved by LizardTech’s MrSID (MultiResolution Seamless Image Database) software. Users could view even the largest maps online without any additional plug-ins or software, as all the work is processed on the Library’s server. With MrSID in place and after several years of trial and testing, the first digitized maps from the Library of Congress became available on June 9, 1999, at <http://memory.loc.gov/ammem/pmhtml/panhome.html>. The site was so popular and received so many hits the first day that the server crashed. While the original offering of 26 maps was not a great deal of data, there was concern about users being able to find the map they wanted, especially as the online collection grew.

Wherever possible, the Division favors a graphic interface for navigation, as most users want to look at a certain country, state, or city (Figure 3). To help reduce the number of maps to be searched, several subsections were developed. There are seven general themes, giving the user a guide for finding a map (Figure 4): Cities and Towns, Conservation and Environment, Discovery and Exploration, Cultural Landscapes, Military Battles and Campaigns, Transportation and Communication, and General Maps. Under these themes are housed special presentations on such diverse subjects as Railroad Maps, Mapping the National Parks, and the American Colonization Society’s maps of Liberia. Each level has its own search capabilities and a map will appear in more than one theme, if appropriate.

After finding a map, the user is presented with the bibliographic record of the item and a thumbnail view to give some idea of what the map looks like. Clicking on the thumbnail retrieves the MrSID interface, showing a Zoom View and a Navigator View (Figure 5). The small Navigator View is especially useful for large files, allowing the user to know where they are in an image. MrSID also allows a user to explore the map in detail, print the various views, and even download the file. The Division was fortunate to have LizardTech as a partner in its Center for Geographic Information, as the MrSID software met all the requirements for online viewing of

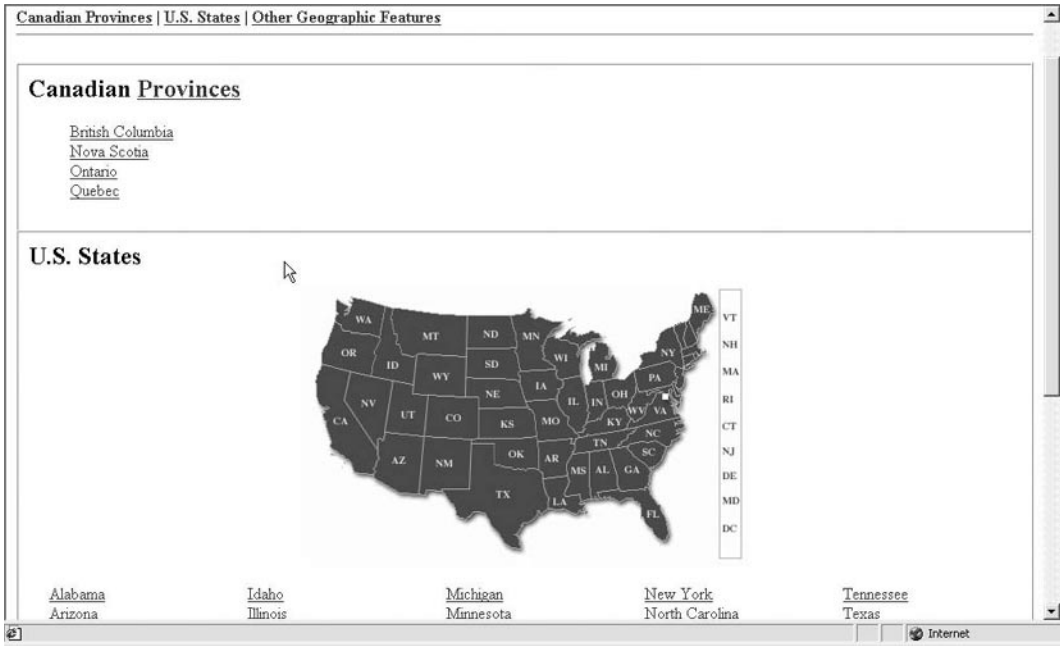


Figure 3. Graphic interface to aid in finding a map.

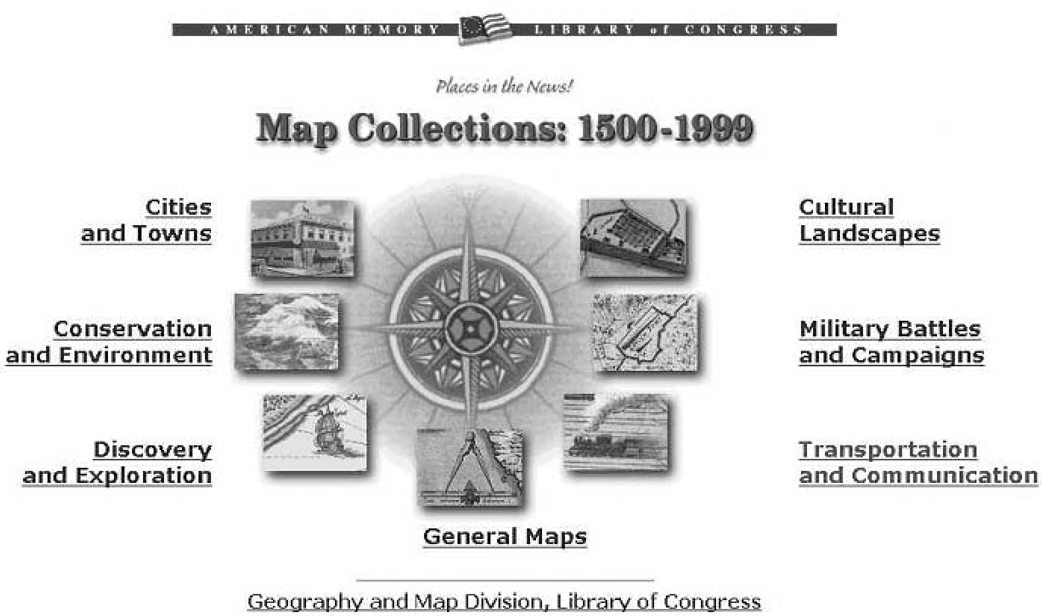


Figure 4. Map Collections Web page.

maps. The ability to zoom in and out and view such amazing detail, greater than possible with the naked eye, was a great advantage. All of the work to generate these images is accomplished at the Library of Congress server and no special software or plug-ins are required at the user's end. Even an older version Web browser, such as Netscape 3.0, will access these images just as easily as a graphic browser made today. The compressed files are small enough to allow for easy download if a user wants to view the image offline and, with free software from LizardTech, users can insert MrSID files into documents and databases.

The bibliographic record is one of two kinds of metadata supporting the online images. It allows for name, title, and subject searching and gives the user information about the original map. All of the Division's online maps have a catalog record in the Library's online database,

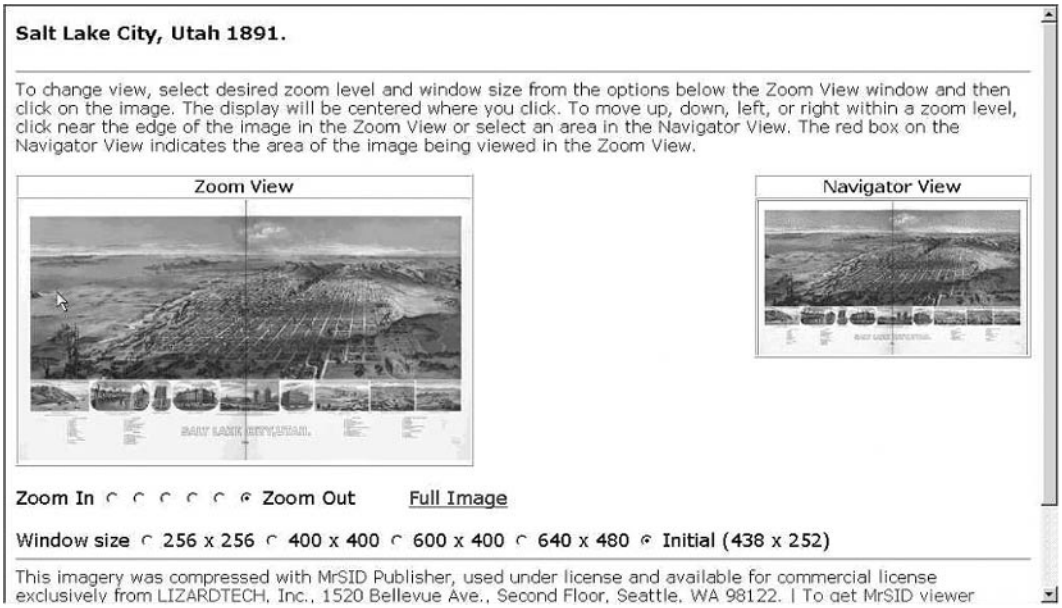


Figure 5. MrSID display.

at <http://catalog.loc.gov>. The records are available in American Memory and the Library of Congress online catalog; both of these databases allow you to search and retrieve digital maps. A link on the online cataloging record takes the user to the thumbnail and bibliographic information within American Memory. The Digital Team also maintains a Microsoft Access database of structural and administrative metadata. There are many uses for these data, including workflow tracking, error checking, and statistical reports.

CHALLENGES

As with every project, issues arose. Once individual maps were successfully placed online, the next step was dealing with an atlas. The Division has a 1570 edition of Abraham Ortelius' Theatrum Orbis Terrarum (Figure 6), which was disbound for an exhibit. Taking advantage of this, all the pages were scanned and we addresses the challenge of how to display the atlas. A page turner was used for the whole atlas, but the file names for maps were set up so that they could be displayed directly in contact sheets (Figure 7). The caption titles on these images were derived from the MS Access database, giving us additional value from these metadata.

With the first scanning project, a major challenge was encountered with the panoramic maps. *The Pictorial St. Louis* (Compton, 1876) was a panoramic atlas consisting of 110 pages. If laid out by the key map, these sheets would form a large wall map. Display was an issue, and the creative talents of David Woodward, a programmer on the American Memory project, found the solution. The key map became a base for a Web image map (Figure 8). An Access database identified for the system the relationship of each sheet to its surrounding pages. The user could select any map to start, and once the MrSID image loaded, a grid below the Navigation View allowed free movement to all adjoining maps (Figure 9). This tool allows the Division to manage related maps in a logical, user-friendly manner. Through the use of MrSID world aggregation files, which tile multiple TIFF files into one compressed image, there are no size limits on compressing files for Internet viewing. The Digital Team has MrSID files that represent more than 4 gigabytes of data.

INTO THE FUTURE

The Geography and Map Division continues to scan cartographic materials. There are currently over 5,000 items online and more are made available every year. The two current primary projects are the maps of the American Revolutionary era and the United States Civil War. In conjunction with the latter, the Division has a joint project with the Library of Virginia and the Virginia Historical Society to post to our Web site the scans from their collections. Farther on the horizon are manuscript maps from the Luso-Hispanic bibliography, city ward maps, Asian maps, and several other projects.

The present setup limits the Division to scanning flat items; however, in the near future we may try to scan bound atlases. Most of the atlases currently online were either disbound or scanned by other Library staff with a PhazeOne camera. The Division is investigating onsite scanning of books. Always looking ahead, the Division is investigating scanning of 3-D items, such as globes, to make all its holdings available on the Internet.

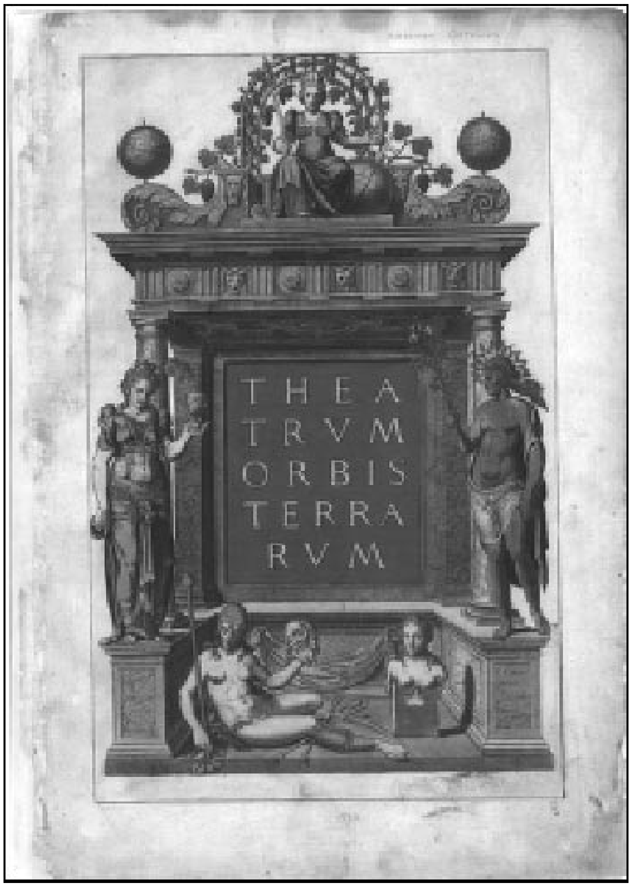


Figure 6. Abraham Ortelius, 1570, Theatrum Orbis Terrarum.

We hope you will visit the Map Collection web site at <http://memory.loc.gov/ammem/gmdhtml/gmdhome.html> and the Geography & Map Division Web site at <http://lcweb.loc.gov/rr/geogmap/gmpage.html> to view the online cartographic materials.

REFERENCES

Compton, R.J., 1876, Pictorial St. Louis: Compton & Co.

Hébert, J.R., and Dempsey, P.E. (revision author), 1984, Panoramic maps of cities of the United States and Canada: Library of Congress, 181 p.

Ortelius, Abraham, 1570, Theatrum orbis terrarum: Apud A.C. Diesth.

Washington, George, 1766, A plan of my farm on Little Huntg. Creek & Potomk.

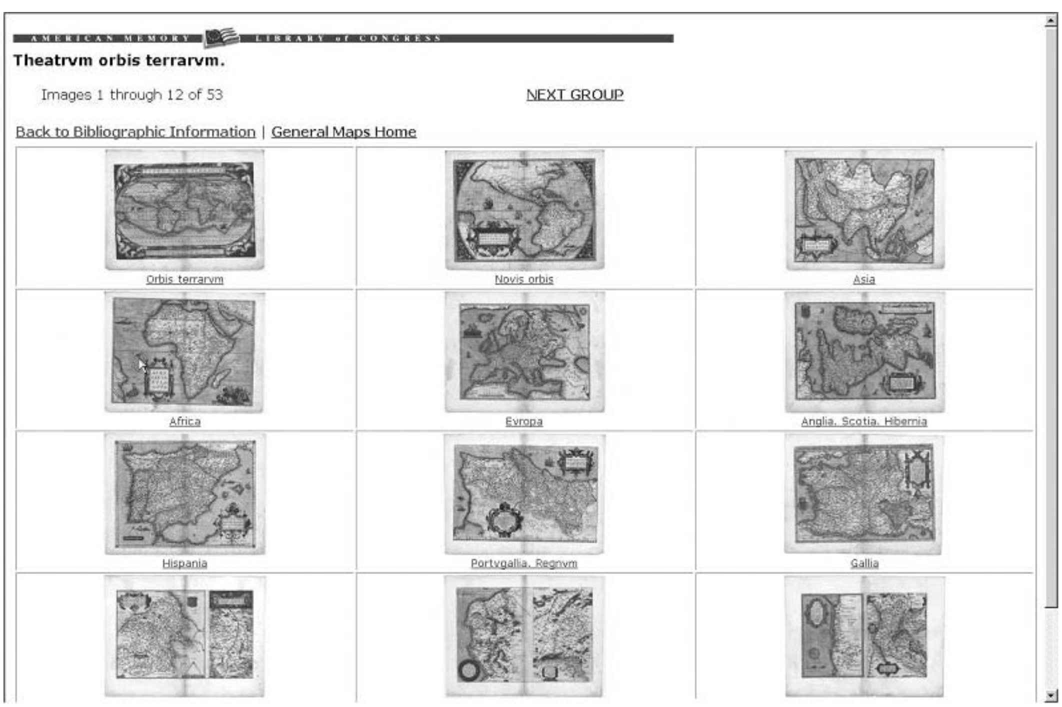


Figure 7. Contact sheet of images.

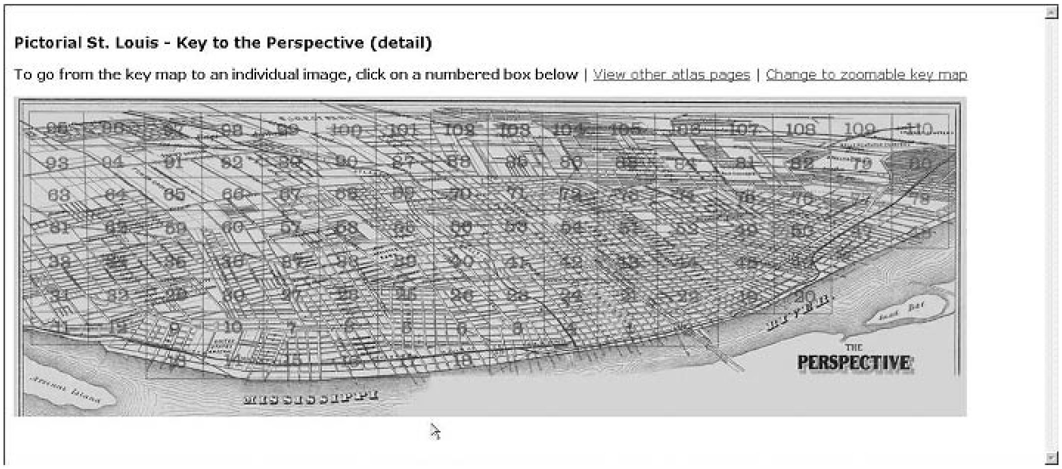


Figure 8. Image map made from key map of *Pictorial St. Louis*.

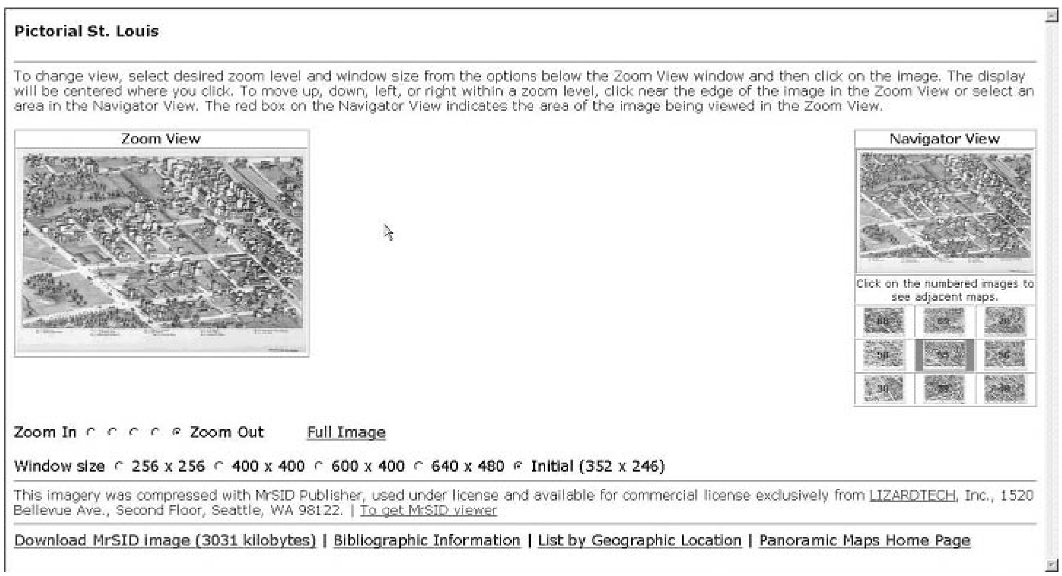


Figure 9. Grid for navigation of adjoining maps.

The Alaska DGGS Scanning Project: Conception, Execution, and Reality

By Gail Davidson, Lauren Staft, and E. Ellen Daley

Alaska Division of Geological & Geophysical Surveys
794 University Avenue, Suite 200
Fairbanks, AK 99709
Telephone: (907) 451-5006
Fax (907) 451-5050
Email: Gail_Davidson@dnr.state.ak.us

INTRODUCTION

The Alaska Division of Geological & Geophysical Surveys (DGGS) is engaged in an ongoing series of projects under the auspices of the Minerals Data and Information Rescue in Alaska (MDIRA) program. The mandate of this program is to recover and make easily available all mineral-related files, documents, and physical samples held in the public domain, in order to prevent this information from being lost through attrition or inaccessibility. Several different public entities, including the Alaska Department of Natural Resources (of which the Division of Geological & Geophysical Surveys is a part), the University of Alaska, the Alaska Resource Library Information System (ARLIS), the U.S. Geological Survey (USGS), and the U.S. Bureau of Land Management (BLM), are on the Liaison committee, along with private entities, including the Alaska Federation of Natives, the Alaska Miners Association, and various interested members of Alaska's mining community. Specific DGGS projects funded by the MDIRA program include the *Guide to Alaska Geologic and Mineral Information*, published by DGGS in 1998, Alaska Resource Data Files, Alaska State Agency Lithochemical Data; Alaskan Bedrock and Surficial Map Index; the partially completed DGGS-wide Geologic Database project; and the DGGS Scanning and Document Conversion Project, designed to scan and make available on the World Wide Web all DGGS publications. The Scanning Project began in 1999 and is virtually complete at this time.

CONCEPTION

The conversion of DGGS publications to an electronic format has a dual purpose. First, by converting these documents, we can make them available on our Web site for easy public access to our publications. In the past, people who requested information could come to our building in Fairbanks or to a library in Anchorage or Juneau, or they could wait for a publication to be mailed to them; now publications are available at a mouse click. Second, we wanted to provide electronic backups of irreplaceable documents. Some of our publications date back as far as 1903 and are extremely fragile. We formerly had to provide low-quality photocopies of these documents for distribution; now they can be viewed at any time, online, with no damage to the original and at better quality.

EXECUTION

Personnel

The following personnel have been involved in the project:

- Geologist V: Project oversight
- Geologist III: Project manager, in a long-term non-permanent position
- College intern: Map scanning
- College intern: Databasing and Web-page production

Management and Logistics

The first issue in this project involved the merits of scanning documents and maps in house or contracting them out. We decided to contract the document scanning for all but the most fragile of our 1,900 titles, totaling 67,000 pages. The contract required: (1) scanning and conversion of pages up to 11" x 17"; (2) conversion to Adobe Acrobat .PDF format; and (3) optical character recognition (OCR), because OCR'ed versions are smaller and more readable, as well as editable. Bindings were removed where necessary, and documents were shipped to the contractor in batches. The total time for completing the contract was about nine months. Upon their return to us, all files ran through a quality control filter. An Access database was employed to store data on documents sent and returned as well as on data quality.

The second issue concerned the advisability of contracting the scanning of approximately 3,000 oversized sheets, including maps, cross sections, tables, and other large-format published documents. Because we worried about scanning quality due to the large differences in color and quality of our hard copies, we ultimately decided to purchase a 36” scanner to scan the oversized sheets in-house. We then hired a college intern to do the actual scanning. Toward the end of the map-scanning portion of the project, we contracted scanning of 70 oversized sheets exceeding 36” in size to a local printing shop.

Next we needed to decide on an electronic storage method for the maps and documents that would allow archiving as well as delivery to the end user. Adobe Acrobat .PDF files were determined readable by most computer systems with a free viewer, <http://www.adobe.com/products/acrobat/readstep2.html>, both online or off line. This format had been used on our Web site in the past for both text and maps, so we knew it to be reliable. We decided to store in this format all documents scanned and converted to text by the contractor. Our experience led us to conclude that .PDF files of maps would be too big for Web delivery on a large scale, so we searched for a more compact, yet equally useful, format. We chose LizardTech’s MrSID because it is the most widely used compression format available, it provides very good resolution even when zoomed in, it is read directly by ArcInfo, and it has free readers, <http://www.lizardtech.com/download>, available to users for both online and off-line use. Figure 1 shows a MrSID-compressed map, originally made at 1:250,000, that has been zoomed on-screen several times.

The next issue involved deciding upon a means of delivering the scanned documents and maps to the public via our Web site. Because a Divisionwide geologic database was in the planning stages, we looked forward to using it as a means of access to the data. We found, however, that we needed to deliver the scanned products before the database was ready, so we wrote direct Web pages to do so.

The last issue is still in the process of being resolved. Information published since the digital age began is already in electronic format, so it can be included in our archive and Web presentation by moving it to the proper format. A simple means of adding these publications is still under construction.

TECHNICAL ASPECTS

Before purchasing a wide-format scanner, we sent a mylar with various graphics on it to several vendors for

Figure 1. MrSID map showing resolution when zoomed in.

testing. When the files were returned, we checked them for fuzziness, evenness of scanning over the entire map, and stretch. We then purchased a Widecom SLC 936C scanner. During the map-scanning phase of the project, this scanner broke down several times, and we had trouble getting parts for it. After about a year, we purchased a second scanner, a Contex FSC Color 36, which has served us well.

Metadata for both document scanning and map scanning were stored in an Access database (Figure 2). As each map was scanned, parameters on map corners, scales, and other such data were recorded for future entry into the National Geologic Map Database. Maps were scanned at a resolution of 400 dpi. Scanned files were archived on CD-ROM (documents in .PDF format and map files in .TIF format). The .TIF files were compressed to .SID files using parameters of c=30 and n=6.

Because we needed to deliver the scanned project to the public before it was complete, we chose to put all scanned documents on our Web site in December 2000, along with maps that had been scanned to that date. In the fall of 2001 we updated the pages to include all maps scanned. Maps and documents can be viewed directly from the Web server using free viewers. Three search methods are available at <http://wwwdggs.dnr.state.ak.us/pubs.html>: Quadrangle search, Publications Series search, and a keyword search that uses the Google search engine. The latest publications, which were prepared and published electronically, have not been added to the site at the date of this writing (May 2002).

Web pages listing all available publications were produced using a Visual Basic program that does much the same thing as a Microsoft Word mailmerge. The code reads a query that accesses data on all necessary variables from the database, punctuates and formats the results, and writes HTML code. An example is shown in the Appendix.

The Scanning Project database includes an index number that ties the scanned publications to a second Access database at DGGS, where data on authors, titles, and similar attributes of the publications of the Survey are stored. In order to query both databases simultaneously, late in the Scanning Project process we decided to combine them. This effort required changing some field names, but the combined database is useful for several purposes and will ultimately be uploaded to the planned Divisionwide Oracle database. In the future, the Web site will access the Oracle database for delivery of publications to the public.

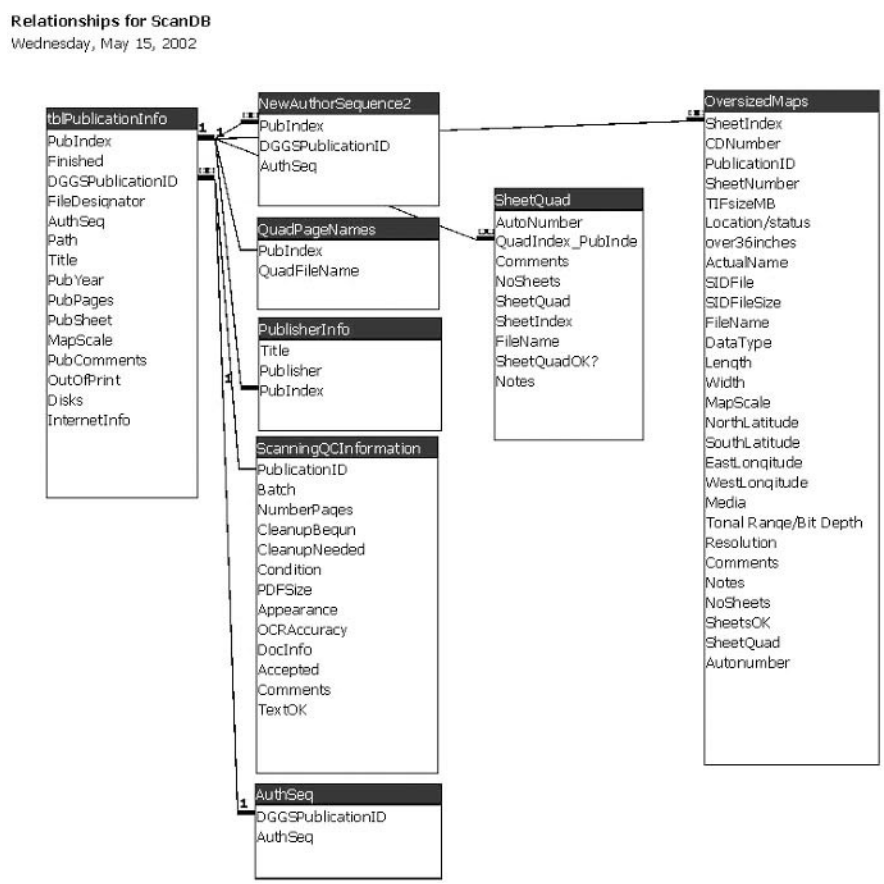
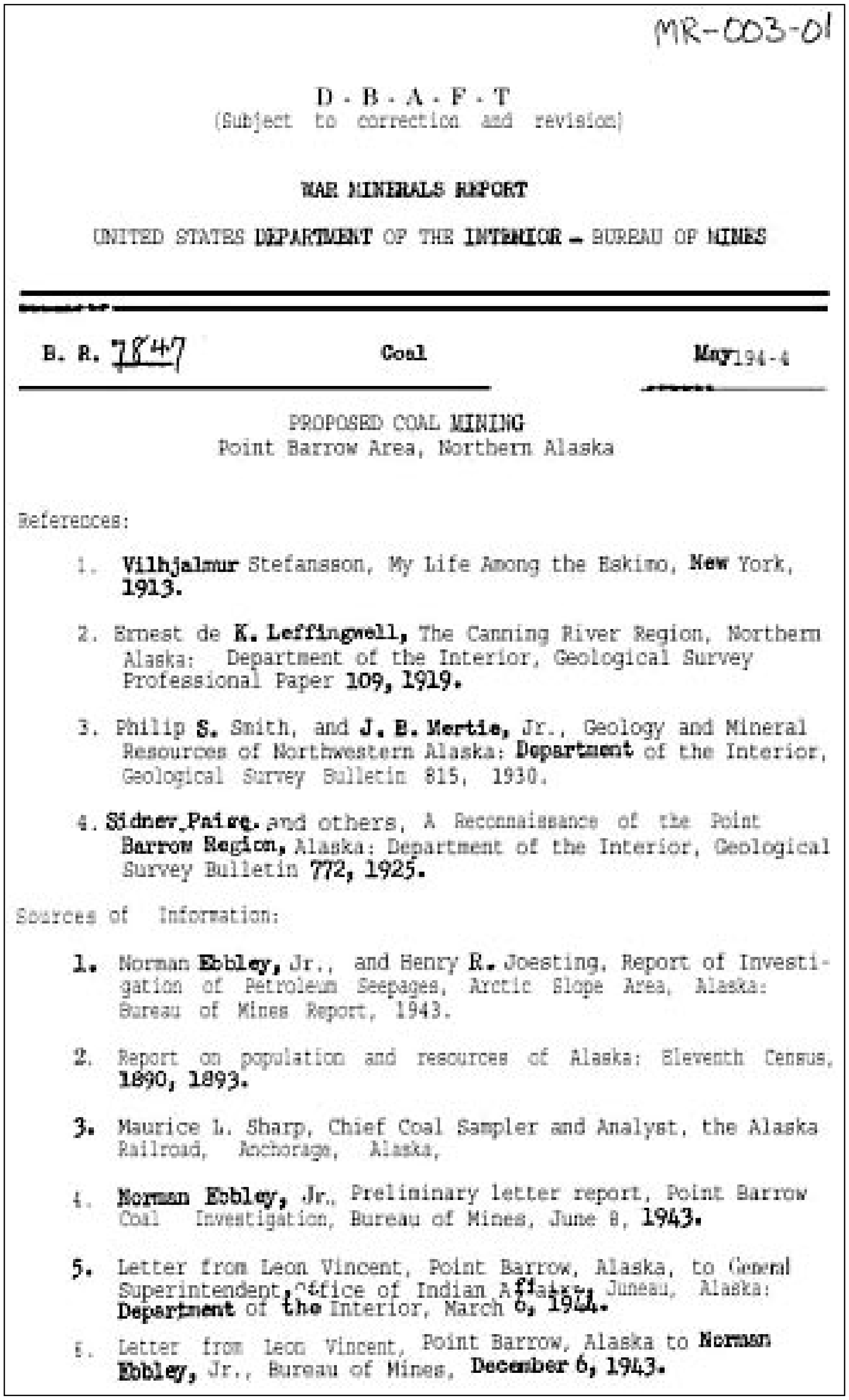


Figure 2. Relationship table for scanning database.

REALITY

Many different issues slowed the delivery of the Scanning Project. First was the time lag—approximately two months—between receiving the project money and hiring the project manager. This time lag is not uncommon in the hiring system of Alaska state government, but made it difficult to deliver products in a short time frame. The project manager left after document scanning was complete, but before map scanning was complete, leaving interns to complete the project.

Document scanning is not perfect technology. Even documents printed on a press may not scan and OCR perfectly; many DGGS publications were typed on manual typewriters and have handwritten notes on them (Figure 3). Although the project manager examined files returned from the contractor and made a note in the database as to the quality of each, the project timeline did not allow fix-



MR-003-01

D · R · A · F · T
(Subject to correction and revision)

WAR MINERALS REPORT

UNITED STATES DEPARTMENT OF THE INTERIOR – BUREAU OF MINES

B. R. 7847 Coal May 194-4

PROPOSED COAL MINING
Point Barrow Area, Northern Alaska

References:

1. Vilhjalmur Stefansson, My Life Among the Eskimo, New York, 1913.
2. Ernest de K. Leffingwell, The Canning River Region, Northern Alaska: Department of the Interior, Geological Survey Professional Paper 109, 1919.
3. Philip S. Smith, and J. B. Mertie, Jr., Geology and Mineral Resources of Northwestern Alaska: Department of the Interior, Geological Survey Bulletin 815, 1930.
4. Sidney Paige and others, A Reconnaissance of the Point Barrow Region, Alaska: Department of the Interior, Geological Survey Bulletin 772, 1925.

Sources of Information:

1. Norman Ebbley, Jr., and Henry R. Joesting, Report of Investigation of Petroleum Seepages, Arctic Slope Area, Alaska: Bureau of Mines Report, 1943.
2. Report on population and resources of Alaska: Eleventh Census, 1890, 1893.
3. Maurice L. Sharp, Chief Coal Sampler and Analyst, the Alaska Railroad, Anchorage, Alaska.
4. Norman Ebbley, Jr., Preliminary letter report, Point Barrow Coal Investigation, Bureau of Mines, June 8, 1943.
5. Letter from Leon Vincent, Point Barrow, Alaska, to General Superintendent, Office of Indian Affairs, Juneau, Alaska: Department of the Interior, March 6, 1944.
6. Letter from Leon Vincent, Point Barrow, Alaska to Norman Ebbley, Jr., Bureau of Mines, December 6, 1943.

Figure 3. Miscellaneous Report 003-01; note the OCR errors.

ing OCR errors. A decision was made early in the project to put the files on line "as is" and to fix them if time and budget allowed. In the course of using the scanned documents, we have found numerous mistakes, such as documents with every other page missing. We are fixing these as we go.

Although we have a large file of mylar originals of our maps, several are missing, and we had to scan paper copies of folded maps for this project. Because our intention was to make all published material available, we used the best copy we could find of each map. Scanner problems set us back several weeks on more than one occasion. We had trouble getting help and parts from the manufacturer.

One of the largest problems is the integration of newer publications with the ones that have been scanned. As we have published maps using GIS and drawing programs, and text using word processors, the files have become scattered and the methods keep changing. We envision that our coming Divisionwide database will alleviate these problems by keeping track of where the various pieces of publications reside. That database will also allow us to feed data to the web directly, using any sort of search imaginable. In the meantime, however, we are in the process of writing code to integrate these publications with those already on the Web site.

We have received very positive feedback on the availability of our publications online, in spite of the very slow Internet speed available between Fairbanks and Anchorage. That bandwidth is in the process of being upgraded now. Publication sales at DGGS have dropped dramatically due to online availability, but we find that net budget changes amount to very little because the lack of sales is balanced by the reduction in our reproduction costs.

APPENDIX

```
Option Compare Database
Sub QuadMailmerge()

' Dim CALLS VALUES FOR VARIABLES, SETS db AS ABBREVIATION FOR DATABASE, rs FOR RECORDSET AND
CALLS PubNumber AS AN INTEGER FIELD
Dim db As Database, rs As Recordset, PubNumber As Integer

Set db = CurrentDb
' SELECT QUERY TO EXTRACT DATA (SETS rs (recordset) AS NAMED TABLE, FOR EXAMPLE "NewQuadMailmerge")
Set rs = db.OpenRecordset("NewQuadMailmerge")

' CALL A FILE TO SEND THE TEXT TO; IN THIS CASE, TEXT IS SENT TO "C:\temp\Alaska.txt"
FileNum = FreeFile
Open "C:\temp\Alaska.txt" For Output As FileNum

' GO TO THE FIRST RECORD
rs.MoveFirst
With rs
  ' SET PUBNUMBER VALUE TO ZERO
  PubNumber = 0

    ' BEGIN LOOP. DOES NOT FINISH UNTIL END OF FILE IS REACHED.
    Do
      If PubNumber = 0 Then GoTo Top

        ' IF THE PUBNUMBER EQUALS THE SHEET INDEX NUMBER (I.E. IS A MAP BELONGING TO THAT
PUBLICATION)
      ' THEN SKIP TO THE MIDDLE OF THE LOOP AND PRINT ONLY SHEET INFO.
      ' INITIAL VALUE IS SET AS ZERO, SO THIS WILL NOT BE TRUE FOR THE FIRST RECORD AND WILL DEFAULT
      ' TO PRINTING PUBLICATION INFO
Here:    If PubNumber = rs!SheetIndex Then
        GoTo Middle
        End If
        ' FOR RECORDS WHERE SHEETINDEX DOES NOT MATCH PUBNUMBER, PRINT PUBLICATION INFO
Top:    PubNumber = rs!PubIndex
      Print #FileNum, "<BR>"
        ' PRINTS THE FILENUMBER, THE AUTHOR, THE PUBLICATION YEAR, THE TITLE, ETC. WHICH ARE ALL
FIELDS IN THE "NewQuadMailmerge"
      Print #FileNum, rs!AuthSeq & ", " & rs!PubYear & ", " & rs!Title & " " & rs!Publisher & ", " & rs!QuadFileName & ":<BR>"
      If rs!InternetInfo = "!" Then Print #FileNum, "<FONT COLOR='RED'>", rs!PubComments, "</FONT><BR>"
      If rs!TextOK = "!" Then GoTo Middle Else Print #FileNum, "<a href='../" & rs!Path & "/text/" & rs!FileDesignator &
".PDF'>Report</a>, " & rs!PubPages & " p., .PDF format (" & rs!PDFsize & " KB).<BR>"
      ' IF THERE ARE NO SHEETS, GOTO NEXT RECORD
      ' OTHERWISE PRINT THE SHEET INFORMATION
Middle:   If (IsNull(rs!NoSheets)) And rs!SheetQuad Like "*Alaska*" Then
      Print #FileNum, "<"; rs!SheetsOK & "a href='../" & rs!Path & "/oversized/" & rs!FileName & ".SID'>" & rs!FileName &
"</a>, " & rs!ActualName & ", "; rs!Comments & ", " & rs!MapScale & ", .SID format (" & rs!SIDFileSize & " KB).<BR>"
      End If
      ' GOTO THE NEXT RECORD
EndLp:   ' END LOOP, GOTO TOP
      .MoveNext
    Loop Until .EOF
End With
    ' CLOSE THE TEXT FILE
    Close #FileNum

End Sub
```

New Tricks for Old Dogs: A Digital Technique for Producing Mylar Base Maps for Geologic Mapping and Compilation

By Adam S. Read, David J. McCraw, and Geoffrey Rawling

New Mexico Bureau of Geology and Mineral Resources
New Mexico Institute of Mining and Technology
801 Leroy Place
Socorro, NM 87801
Telephone: (505) 835-5753
Fax: (505) 835-6333
e-mail: adamread@gis.nmt.edu

INTRODUCTION

There are many benefits to digital field-data-capture and map-compilation techniques, but for many geologists the older methods of mapping on paper and compiling on greenline mylar bases are much simpler, more flexible, and what they are most comfortable with. Mapping on paper is easy to modify, allows symbology to be defined and refined in the field, and is not reliant on cumbersome batteries or cabling. Likewise, compiling field mapping on mylar is typically much faster for geologists that are not GIS specialists. However, modern digital reproduction technologies have made photographic methods of mylar preparation largely obsolete and photoreactive materials for greenline production impossible to find. Furthermore, it has progressively become more difficult to find reprographic shops that still provide photographic services. Photographically enlarged mylar base maps have long been essential for geologists mapping in geologically complex areas. We typically use standard U.S. Geological Survey (USGS) 1:24,000-scale 7.5-minute topographic maps enlarged to 1:12,000 (Figure 1). However, the large size of such a mylar base is often beyond the capability of the reprographic shops that still do photographic work. Accepting that most of our experienced mappers prefer mylar for compilation and paper for field mapping, we devised an in-house, digital solution to mylar base-map production.

PROCESS

Overview

The basic process of digitally creating a mylar base is relatively simple—produce a digital version of the topographic map and plot a mirror image of the map onto double-matte mylar. In-house production of the digital base is a necessary step for final map compilation and production due to the often poor quality of low-resolution digital raster graphics available from the USGS or commercially (see discussion in McCraw, 1999). However, we have developed several refinements to this basic scheme that provide greater utility than base materials created using traditional methods.

Using a geographic information system (GIS) to produce the digital base makes it possible to add a collar of adjacent topographic maps around the map of interest. It is also possible to include previously mapped geology from these adjacent maps on the new base (Figure 2). This saves time in the field and helps eliminate map boundary mismatches during compilation. A digitally produced base map of a particular area of interest can also circumvent Murphy's first corollary of cartography (that the area of interest generally lies at the intersection of four maps). Also, a UTM grid can easily be added to the map for easier global positioning system navigation. In addition, you can change, remove, or retain the colors used in the original paper topographic base, which can aid in map element recognition (e.g., streams). You can also remove the distracting pattern screens representing forested areas or landownership printed on the original topographic maps. We choose colors that will reproduce using a blueprint machine and are easily differentiated from the black-inked linework. This, of course, greatly simplifies the digital data-capture process—after the map is compiled by hand, it can be scanned, converted to a paletted image, and rectified in a GIS where the compiled linework can then be turned off for easier digitization.

Process Details

The following is a method that works for us and uses the hardware and software we have available. There are certainly other, and probably better, ways to accomplish any given step and there may be better hardware, software, and media for this. In any case, this method works

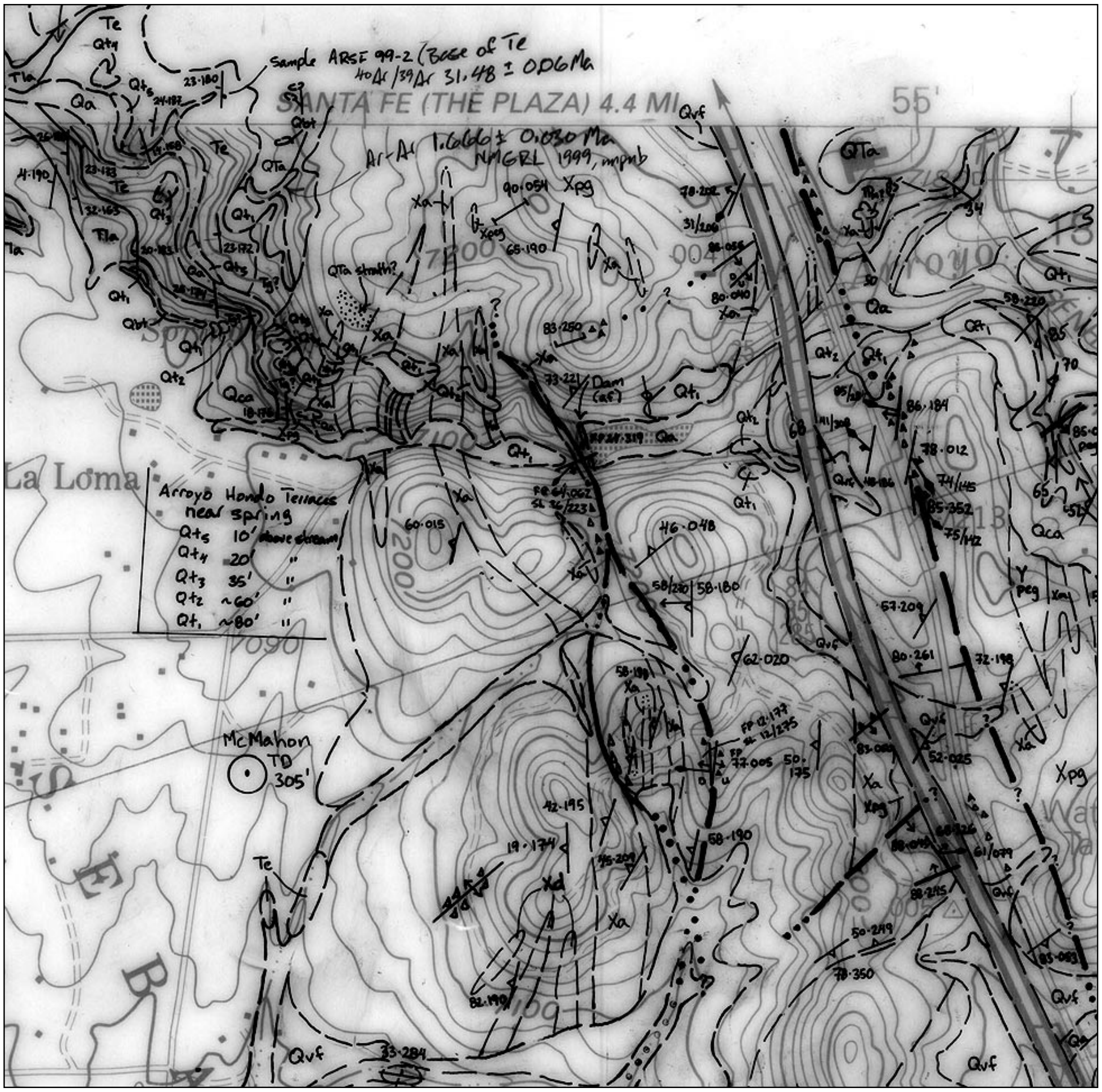


Figure 1. An example of geologic mapping carried out on a traditional photomechanical greenline made from a composite negative of the Seton Village 7.5-minute quadrangle (Read and others, 1999). In this case, the greenline had to be outsourced to a reprographic center capable of photographically enlarging the negative two times to a scale of 1:12,000, to facilitate mapping in an area with a high level of geologic complexity.

reasonably well for us now. There are many minor details not mentioned that improve the final product, but this description should help explain most of the problems we have encountered. We hope this will stimulate discussion and further experimentation.

1. Obtain a high-quality 24-bit color scan of the paper topographic base.

We scan at 400 dpi (the optical resolution of our scanner). We have tried scanning at higher resolutions, but file size becomes an issue and the quality at 400 dpi has been acceptable. A 1:24,000-scale quadrangle scanned at 400 dpi generates a red-green-blue (RGB) tiff file of around 275 Mb.

2. Clean up the image and touch up the green pattern screen of forested areas.

Adjust the contrast with the Auto Contrast command in Adobe Photoshop and adjust brightness if necessary. Select the green (forested) areas of the map with the magic wand tool in Photoshop with the tool's tolerance set to around 30, and the "contiguous areas," "anti-aliased," and "use all layers" boxes unchecked (see Figure 3). This process with the magic wand tool can take several iterations of holding the shift key down to add to the selection of all of the green pixels. Use a successively lower tolerance to avoid selecting colors representing other map elements. Finally, fill

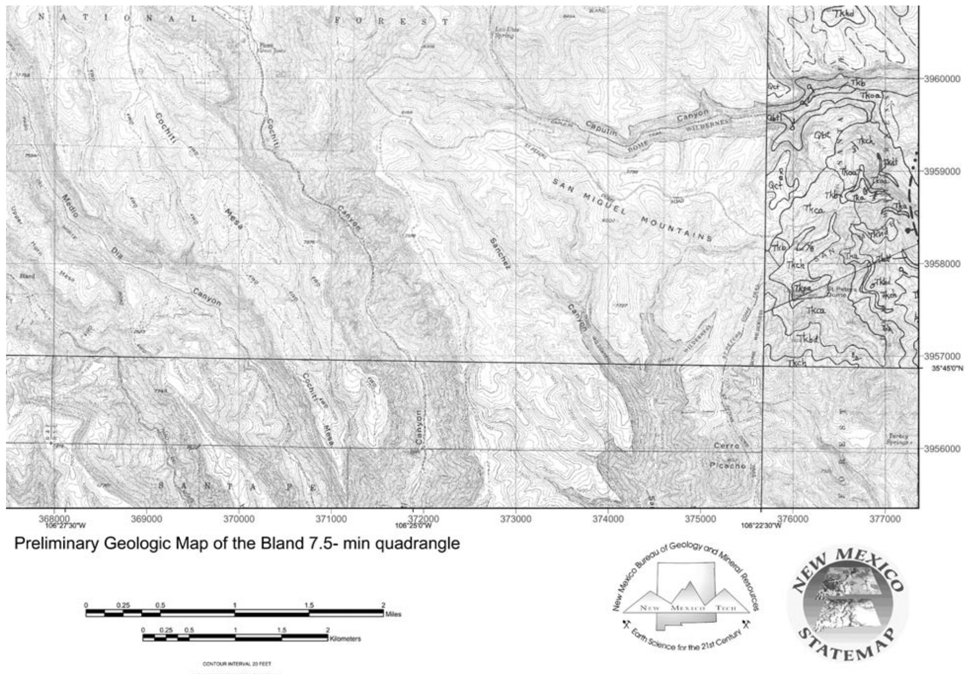


Figure 2. An example of a part of the digitally produced base map of the Bland 7.5-minute quadrangle to be utilized for future geologic mapping. Note that the topography of adjacent quadrangles and the hand-compiled geologic linework of the Frijoles quadrangle (Goff and others, 2001) have been incorporated with the base to facilitate edge-matching.

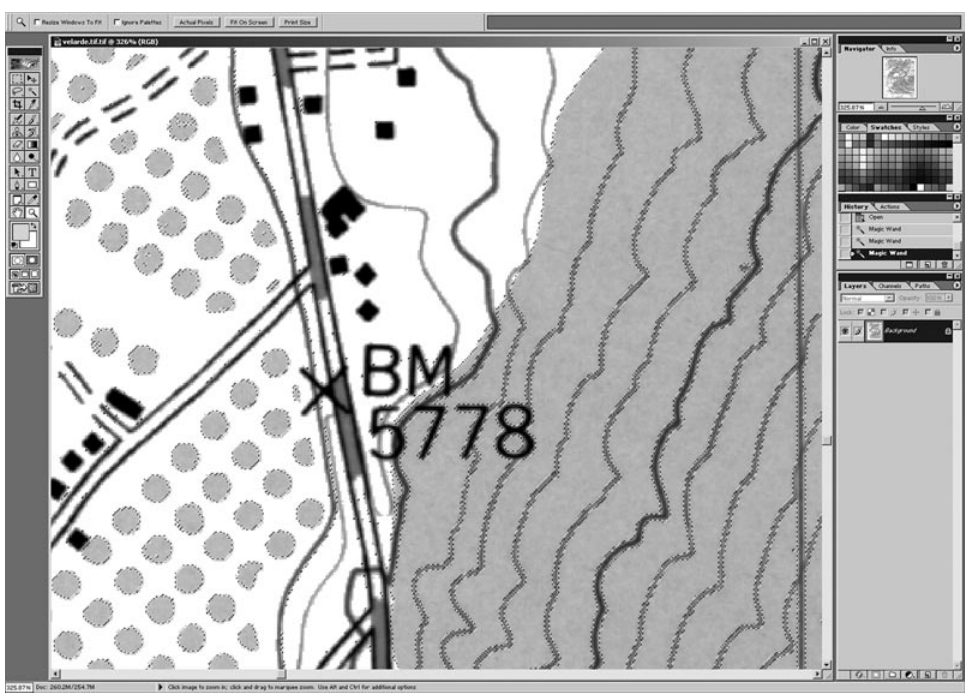


Figure 3. RGB scan of the Velarde 7.5-minute quadrangle showing the green screen of forest (gray patch) and orchard areas (large gray dots) selected with the magic wand tool in Photoshop. Note that not all of these pixels have been selected yet because both RGB scanning and color variations in CMYK printed paper maps introduce a range of green color values. Several successive selection iterations with the magic wand are often necessary because of the varying hues of green (see text). Once selected, they are filled with the standard USGS DRG green in Photoshop. This step forces the forested areas to be classified as green when converted to the 13-color USGS DRG palette.

the selected green regions with the standard USGS digital raster graphic (DRG) green. For us, this process has worked best in eliminating the moiré effect from the scanned pattern screen and in separating the green so that we can remove it from the final topographic base.

3. Convert the RGB image to the USGS standard DRG palette of 13 colors.

Any palette could be used, but a standard palette of 16 colors or fewer is much easier to work with. We also use clipped versions of the USGS DRGs for the topography of adjacent quadrangles; so using the USGS DRG palette is preferable. We use Adobe Photoshop to palette the image because, of the applications we have tried, it seems to work best. You can easily load this palette into Photoshop by opening a standard USGS DRG and saving the color table. To palette a RGB image, select your custom color table when prompted from the image>mode>indexed color menu. This generates a much smaller tiff file that is about 80 Mb uncompressed. There is apparently an in-house USGS paletting program that does this as well, but we have not been able to obtain it.

4. Rectify the image using ArcMap.

Rectify the image using all 16 latitude/longitude tics present on the quadrangle (Figure 4). At this point, we decide whether we want to show the topography or geology from adjacent quads on the base map. If so, clip the collar information from the map; if not, proceed to modify the color table.

5. Clip the map to the extent of the quadrangle to remove the collar information.

Reclassify the map to free up the "0" slot in the color table by using the reclassify function of the ArcGIS Spatial Analyst extension to move the black color from the "0" bin to the "13" bin (Figure 5). This step can be combined with the clipping operation if the analysis mask and extent in Spatial Analyst is set to a polygon shapefile or coverage of the quadrangle boundary (obtained from a statewide coverage of 7.5-minute quadrangles). The clipped grid will use the "0" bin for the "no data" area outside the quadrangle, but the clipped grid will no longer retain the original color-table information (Figure 6). Because of a bug in the ArcGIS 8.1.2, it is necessary to do the next step from the ArcInfo command line.

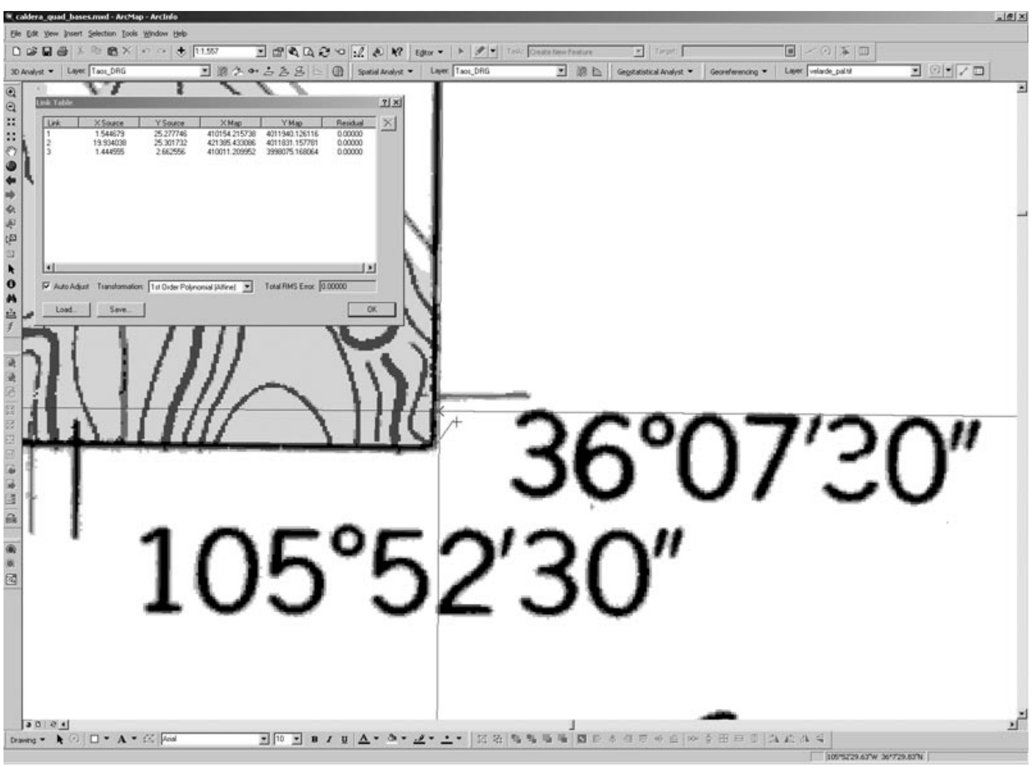


Figure 4. Rectification process step for the paletted image of the Velarde 7.5-minute quadrangle using ArcMap.

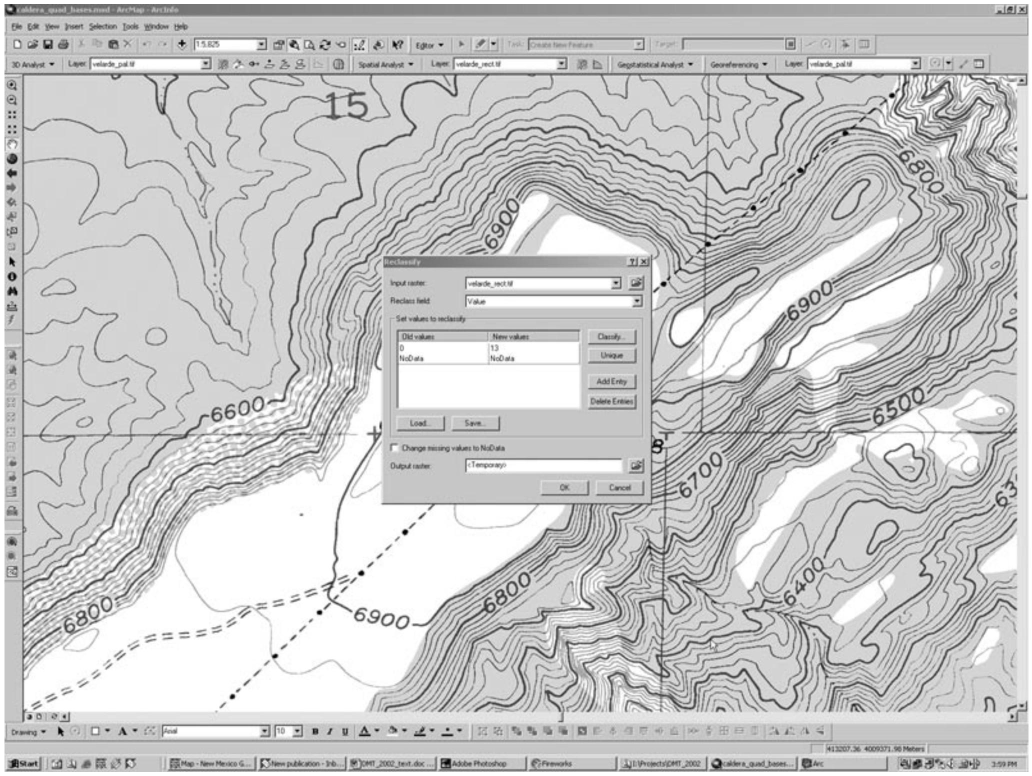


Figure 5. The reclassification process step for the Velarde 7.5-minute quadrangle in ArcMap. This removes the black color from the "0" bin to free it up for the "no data" area outside the map extent.

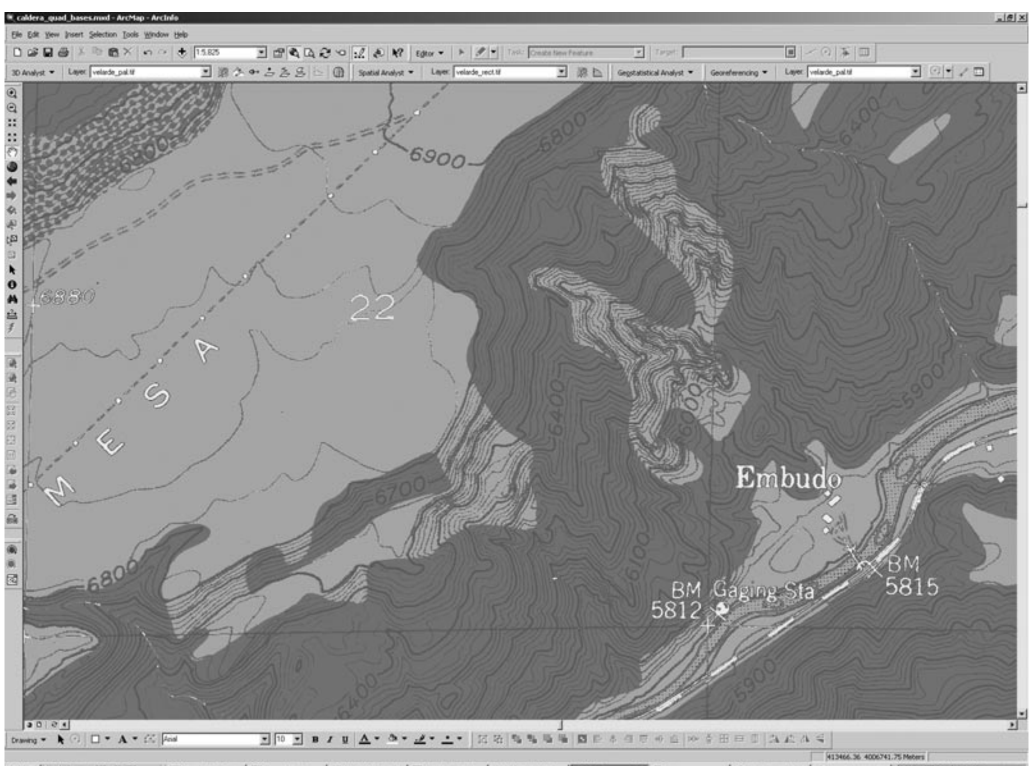


Figure 6. The clipped grid of the Velarde 7.5-minute quadrangle, which now has no data outside of the quadrangle extent, but has lost the original color-table information.

6. Reintegrate the color table with the clipped image using GRIDIMAGE.

The first time this step is done, you will need to use the ArcInfo command IMAGEGRID on a reclassified but not clipped image to save the color table to an ASCII file. Use the ArcInfo command GRIDIMAGE to convert the clipped grid to a geotiff raster using the predefined color table (Figure 7).

7. Modify the colors in the color table using ArcMap to make a “greenline” base.

- Change the black (13) to dark green—“leaf green” works well.
- Remove the forested areas’ green color (5).
- Change the contour-line color (4 and 12) to light green—“fern green” works well.
- Remove any colors that clutter the image—gray (11) is sometimes removed.
- Modify any other colors as necessary—each quadrangle will need custom color adjustment because the process-color inks used on the paper maps have changed over time.

8. If adjacent quadrangle data are to be included on the map, repeat the previous steps to add it. A similar process can be used to incorporate geology linework from a paletted scan of either a published map or from a hand-drafted mylar. All colors in the palette except the linework can be removed.

9. Add a latitude/longitude graticule and a UTM grid to the map along with scale bar, etc. A red UTM grid will reproduce on a blueprint machine; blue will not.

10. Print the map on paper as a test before trying to print to mylar (Figure 8). (These paper prints are great for field mapping, especially if thin matte mylar is glued to the surface first. We use 3M “Super 77” spray adhesive.)

11. Print a mirror image of the map onto double-matte mylar. This puts the base on the back of the mylar so that hand drafting can be done on the opposite matte side. We use the Postscript driver in ArcMap and set the emulsion side to “down.” You could also use the printer driver settings to make a mirror print.

TECHNOLOGY

Software

We use ESRI products including ArcInfo, ArcMap, and the Spatial Analyst Extension for most of the process of making mylar bases. We also use Adobe Photoshop 6.0.1 for cleaning up and paletting images.

Hardware

We have a Colortrac 5480 scanner from Action Imaging Solutions that is capable of scanning documents up to 54 inches wide at an optical resolution of 400 dpi. Our plotter is a Hewlett-Packard DesignJet 5000ps plotter capable of 600 x 1200 dpi plots up to 60 inches wide. We use the ultraviolet-resistant (pigment-based) ink system in the plotter to avoid rapid fading of the plots. This choice of ink systems made finding appropriate mylar media difficult.

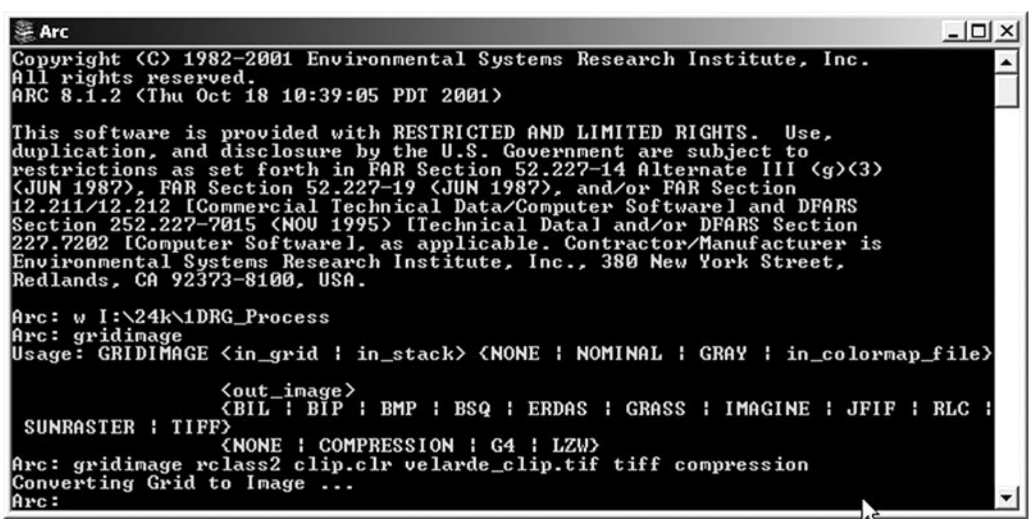


Figure 7. The ArcInfo GRIDIMAGE command is used to reintegrate the original color table with the clipped image file.

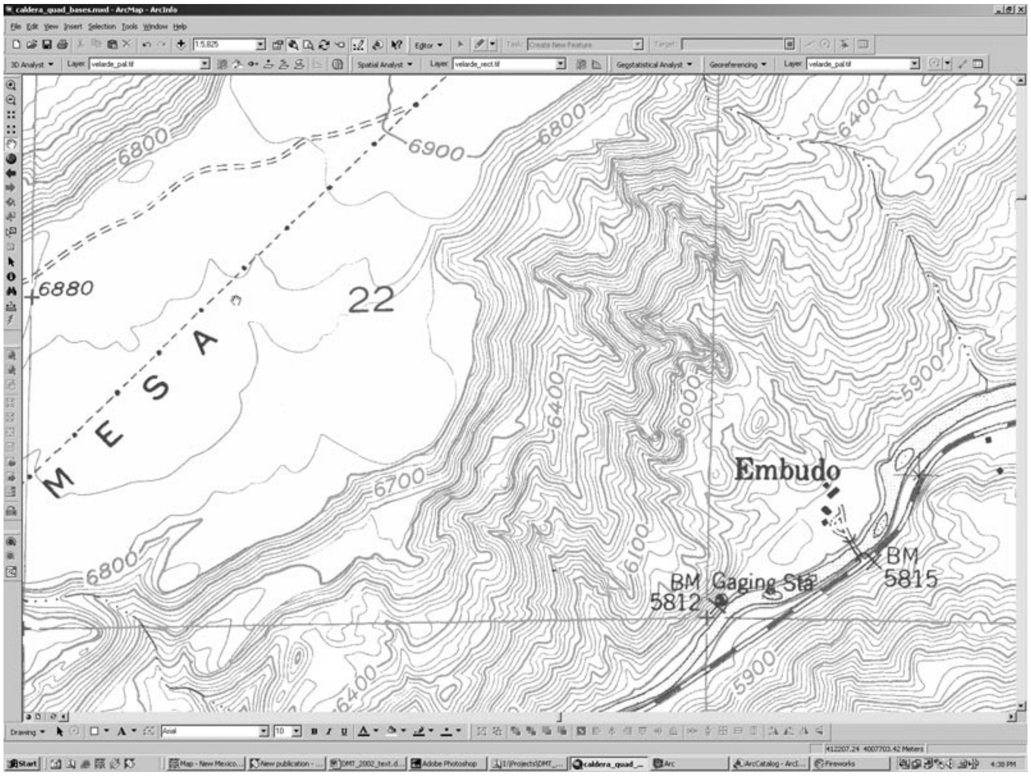


Figure 8. The digital greenline topographic base for the finished Velarde 7.5-minute quadrangle with the green forested-areas pattern removed and brown contour lines changed to green. The Embudo gaging station on the Rio Grande in the lower right is the oldest continuously active gaging station in the United States, installed by the USGS in 1889.

Media

After trying many types of mylar on our plotter, we have found the best results with Océ 4-mil double-matte film. According to the HP and Océ documentation, this film is not recommended for UV inks. It does work fairly well, however, as long as there are no large areas of solid color. HP 4-mil double-matte film appears to have identical specifications to the Océ media, but does not work well at all (and does not claim to be compatible).

Topographic maps generally look fine with this plotter/media combination. Occasionally, the ink will bleed where there are very close contours or otherwise cluttered areas (monochrome plots look better in these cases). There are many more film-media options if using dye-based inks, but we are concerned that maps plotted with dyes will not be archival documents. Ideally, we would prefer a thicker mylar, perhaps 7 mil, for our larger 1:12,000-scale bases because it would be somewhat more stable. However, no other media has worked as well as the 4-mil Océ media.

CONCLUSION

Digital geologic mapping is currently only the domain of true technophiles. Paper-based geologic mapping will be with us for a long time to come, as will the many technophobes who are experienced mappers. However, digital methods for making paper and mylar base maps can improve on the old photographic methods—even as those methods become a dying art.

REFERENCES

Goff, F., Gardner, J.N., and Reneau, S.L., 2001, Geology of the Frijoles 7.5-min. quadrangle, Los Alamos and Sandoval Counties, New Mexico: New Mexico Bureau of Geology and Mineral Resources Open-File Geologic Map OF-GM 42, scale 1:24,000.

McCraw, D.J., 1999, “Can’t see the geology for the ground clutter”—Shortcomings of the modern digital topographic base, *in* Soller, D.R., ed., Digital Mapping Techniques ‘99—Workshop Proceedings, U.S. Geological Survey Open-File Report 99-386, p. 21-26, <http://pubs.usgs.gov/openfile/of99-386/mccraw.html>.

Read, A.S., Rogers, J., Ralser, S., Ilg, B., and Kelley, S., 1999, Geology of the Seton Village 7.5-min. quadrangle, Santa Fe County, New Mexico: New Mexico Bureau of Mines and Mineral Resources Open-File Geologic Map OG-GM 23, scale 1:12,000.

Computer-Aided Structure-Contour Mapping in Support of the Ohio Division of Geological Survey Bedrock-Geology Program

By James McDonald

Ohio Geological Survey
4383 Fountain Square Drive
Columbus, OH 43224
Telephone: (614) 265-6601
Fax: (614) 447-1918
e-mail: jim.mcdonald@dnr.state.oh.us

INTRODUCTION

In 1997, the Ohio Division of Geological Survey (ODGS) completed a seven-year project to remap the bedrock geology of the state. The previous state bedrock-geology map dated from 1920 (Bownocker, 1920). While detailed bedrock-geology mapping has been ongoing since the mid-1960's, very little of the state had been completed by the mid-1980's. In recognition that detailed bedrock-geology mapping of the state would take over 100 years, reconnaissance bedrock-geology mapping was initiated in 1990 to accelerate the remapping of the state.

Computer-based technology was used to expedite completion of the reconnaissance-bedrock-geology mapping through digital generation of structure-contour maps of the units mapped. Computer techniques were developed to minimize edge effects during the gridding and contouring process. Other techniques were developed to assist in edge-matching of the maps produced by the geologists. Final output of the structure-contour maps was edited using CAD software that interfaced with the gridding and contouring software. The computer mapping software enabled the ODGS to complete the reconnaissance remapping of the bedrock geology of the state in the relatively short time period of seven years.

The 1920 state bedrock-geology map was inaccurate in many areas, used outdated stratigraphic nomenclature, was highly generalized, and was plotted on a now-outdated, 1:500,000-scale base map. Between 1918 and 1979, there were numerous changes in stratigraphic nomenclature and concepts, some of which dated the older style of mapping. For example, starting in the mid-1950's, the Survey began using lithostratigraphic units to describe formations, instead of chronolithostratigraphic terminology. In addition, there have been advances in the understanding of the bedrock topography of the state. While standard topographic maps can be used to depict bedrock contacts in the nondrift-covered areas, bedrock-topography maps are required to delineate the bedrock geology across the glaciated two-thirds of the state. Finally, new planimetric and topographic base maps have been made for Ohio by the U.S. Geological Survey, which allow the ODGS to more accurately depict the bedrock geology and cultural features in the state.

Mapping of the bedrock geology has been ongoing since the release of the 1920 bedrock geology map. From 1918 to 1979, the ODGS conducted county-level geologic mapping at the scales of 1:62,500 and smaller, on base maps constructed from the U.S. Geological Survey 15-minute-topographic quadrangles (see, for example, Sherman, 1933). Only 18 counties had been completed by the time the last county bulletin was published in 1977 (Collins and Smith, 1977). Starting in the late 1960's, a more detailed level of mapping was initiated. Between 1957 and 1963, a new topographic map series was created at a scale of 1:24,000 by the U.S. Geological Survey (Bernhagen, 1994). With the completion of that program, there were a number of initiatives, starting in the mid-1960's (DeLong, 1965), to perform detailed geologic mapping at the 1:24,000 scale. By the end of the 1980's, only 37 quadrangles, out of 788, had been completed, and it became apparent that the new detailed bedrock-geology mapping effort was going to take over 100 years to cover the entire state. With the appointment of a new state geologist in 1989, a new program was initiated at the ODGS to perform more rapid reconnaissance geologic mapping at 1:24,000 scale, using supplemental funds from U.S. Geological Survey STATEMAP grants, U.S. EPA Nonpoint-Source Pollution 319-A funds, and Ohio Department of Transportation grants. This program would allow the completion of a new statewide geologic map in a few years.

In order to complete mapping in such a short period, the use of computer mapping software was necessary. The software would be used to create structure-contour maps of the units being mapped and would accelerate the mapping in a number of ways. The software would reside on individual geologists' PC's, thereby allowing the geologists' easy access to the software and structure-contour maps at any time. The software would be programmed to minimize the problem of edge-matching within, and outside, the geologist's project area. Finally, multiple versions of the structure-contour maps could be created very easily, thereby allowing the geologist to easily correct problems and test multiple hypotheses. This paper describes the gridding algorithm and mapping technique used to automate the creation of structure-contour maps for the statewide mapping of the bedrock geology in Ohio.

GENERAL MAPPING PROCEDURE

To map the bedrock geology over a several-year period, the state was divided into six mapping phases (Figure 1) by area. Each phase was divided into 30 x 60 minute quadrangles, corresponding to the U.S. Geological Survey's 1:100,000-scale topographic maps. Each mapping geologist was assigned one or more 30 x 60 minute quadrangles per year and was responsible for collecting and interpreting all available near-surface bedrock data and creating structure-contour maps for the assigned 30 x 60 minute quadrangles. For the first phase of the mapping, the ODGS decided to map the central-western portion of the state. This first phase of mapping covered an area of a 1° x 2° block represented by the Ohio portions of the Lima, Piqua, Bellefontaine, and Marion 30 x 60 minute quadrangles. The original mapping procedure

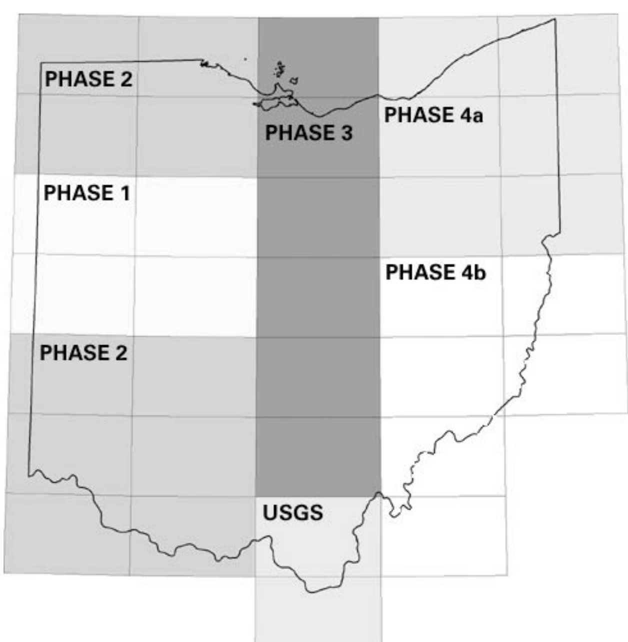


Figure 1. Map showing the different phases of the bedrock-geology mapping program.

for these four quadrangles required each geologist to grid and contour the entire assigned 1° x 2° project area(s). Unfortunately, the process of gridding and contouring that large an area was slow, especially when mistakes were discovered. PC technology was inadequate for the size of the area being mapped. Changes in procedures during Phases 2 through 4b increased efficiency in the processing time for the geologists. During Phase 1, all geologists had to progress together through each task in the mapping sequence. In all subsequent phases, each geologist mapped his areas at his own pace. Data points, pseudodata points (projected points), and grid values were shared between the project geologists to ensure that the resulting maps matched each other. The final products were structure-contour maps at 1:24,000 scale, which were cut from either the 1° x 2° block (Phase 1) or from the 30 x 60 minute quadrangles (Phases 2 through 4b) (Figure 2).

Structure-contour maps were then used in conjunction with bedrock-topography and surface-hypsography maps to produce 1:24,000-scale bedrock-geology maps. The structure-contour maps were overlain with bedrock-topography maps and surface hypsography. At the location where the structure-contour elevation intersected the surface hypsography or the bedrock topography, the unit contact was drawn (Figure 3). Depending on how many units were present in the area at the surface or subcrop, between one and eight different structure-contour maps were drawn per 7.5-minute quadrangle. By the end of the project, the geologists had created more than 1,840 structure-contour maps for the bedrock-geology mapping project. It was only through the use of the computer-mapping software that such a large number of maps could be created in such a short period of time.

SOFTWARE AND GRIDDING ALGORITHM

After a review of the then-available mapping software for the personal computer, the ODGS elected to use CPS/PC from the Radian Corporation to create the structure-contour maps. This software package had a number of features that made it attractive, including the ability to choose different gridding algorithms, the ability to handle faults, macro programs to change the functionality of the software, and an add-on CAD software program. The ODGS needed to evaluate the different gridding algorithms to see which one was most appropriate for its use. Therefore, the ODGS needed a software package that had different gridding algorithms. The software had to realistically handle faults. Another desired feature was the ability to modify the operation of the software. The geologists performing the mapping were novices in the use of computers and mapping software, so it was necessary to automate many of the computer-mapping operations. Finally, the output of the structure-contour mapping was

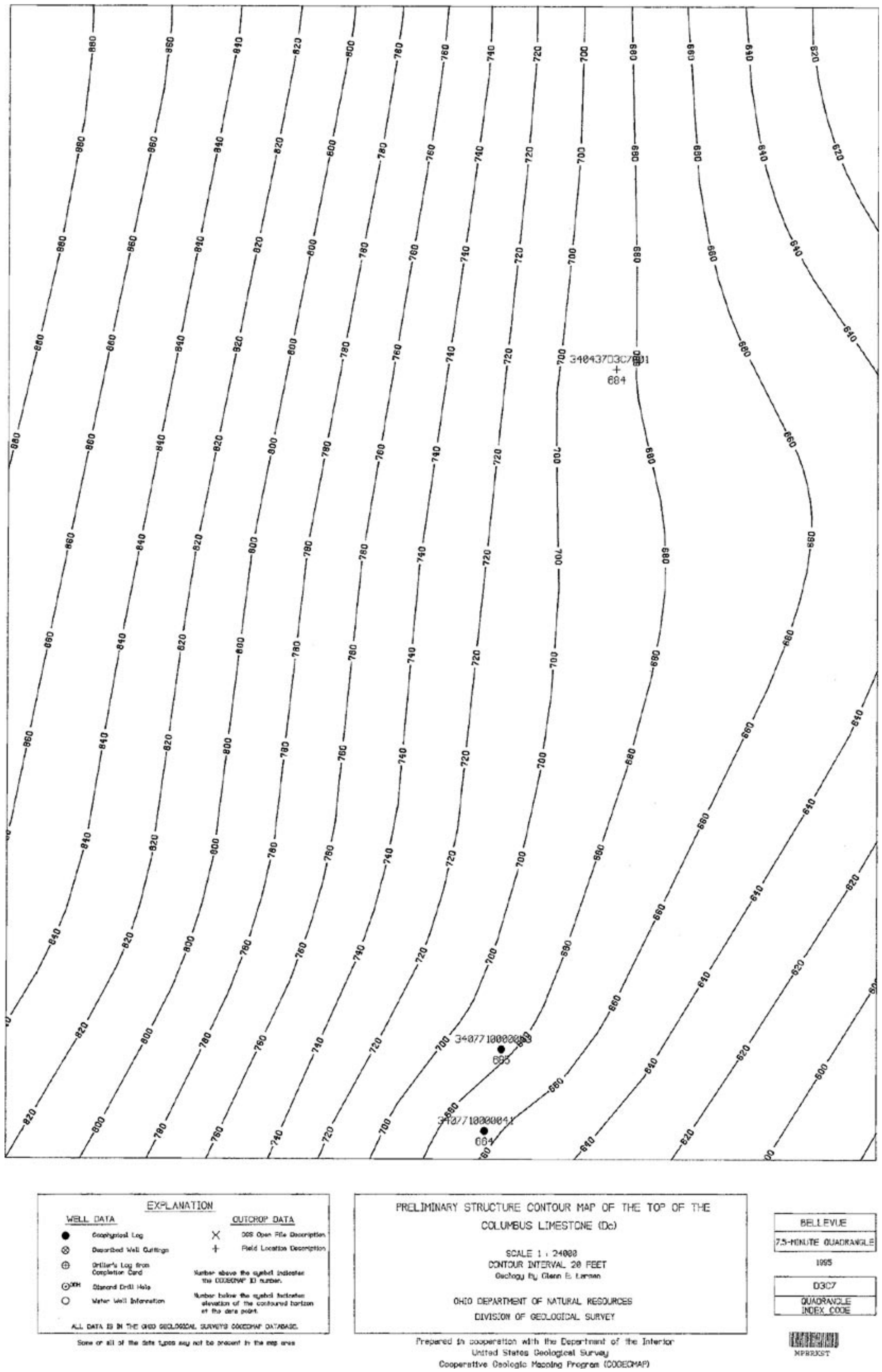


Figure 2. A computer-generated structure-contour map, as released in the ODGS informal, open-file series.

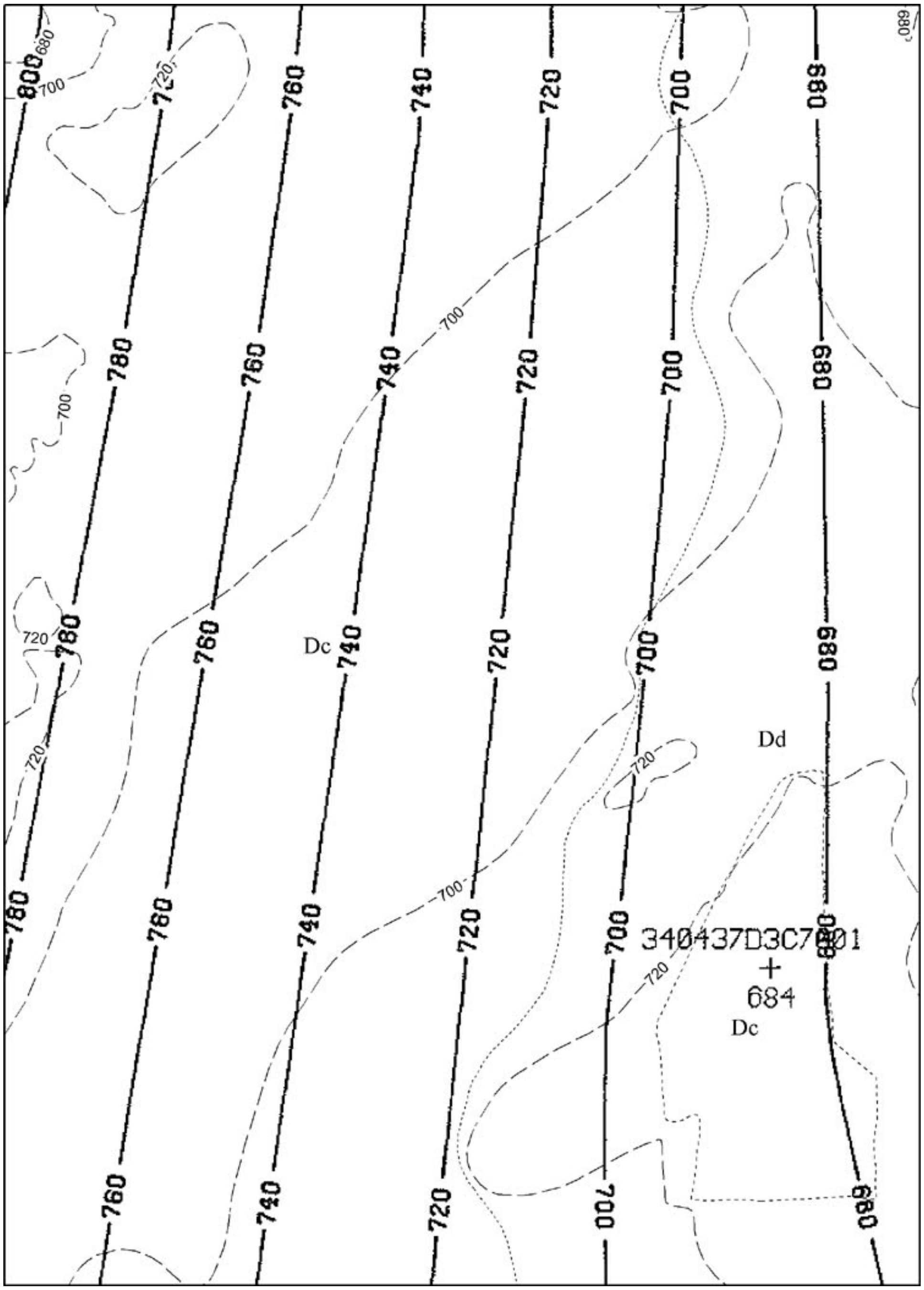


Figure 3. An open-file structure-contour map is overlain with the bedrock-topography map and surface hypsography. Where the structure-contour surface intersects the bedrock-topography surface, bedrock contacts are drawn. In this example, the solid lines are the structure contours, the dashed lines are the bedrock-topography contours, and the dotted line is the bedrock contact.

to be placed into the ODGS open-file map series. ODGS needed a feature to produce more advanced output from a software package than was typically available. The CPS/PC software package had an option to import graphics data from the software into EasyCAD. Each of these features made the CPS/PC software an easy choice.

On the basis of its own experience and upon discussions with the technical staff at the Radian Corporation, ODGS decided to use a proprietary gridding algorithm created by the Radian Corporation called convergent gridding (Haecker, 1992). This algorithm appeared to minimize the edge effects that are produced by other gridding algorithms, such as minimum curvature or least squares. When gridding parameters are properly set up, the gridding algorithm also would honor the data points. These aspects of minimizing the edge effects and honoring the data points greatly attracted ODGS to use the convergent gridding algorithm.

Other types of gridding algorithms, such as least squares or minimum curvature, produce a number of undesirable effects when creating a grid of calculated values from the original data. One of the undesirable effects is that most gridding algorithms do not honor the original data points. These algorithms may produce a grid of calculated values that conflict with the original data values. Another undesirable effect is edge effects. Most algorithms compute the first and second derivatives (slope and curvature) during the process of creating the grid. At the edge of the gridding area, critical information is not available and the derivatives cannot be correctly computed. Most software programs make assumptions about the first and second derivatives at the edge of the gridding area. The surfaces that they produce generally are wildly divergent from what is expected. The convergent gridding algorithm used in the CPS/PC program minimized these edge effects, but, as was discovered during the project, the algorithm did not totally remove them from the resulting calculated grid. For ODGS to further minimize the edge effects from the maps, ODGS implemented a number of different data-handling and processing techniques, described in a later section.

The convergent gridding algorithm was developed by the Radian Corporation to help honor the original data points and to help minimize the edge effects (Haecker, 1992). The first step in computing the grid is to assign the values of the data points to the initial grid that was set up as part of the gridding parameters. The initial grid values comprise the data point values, the slope, and the curvature of the nearest grid values. Each grid node undergoes a smoothing process using a biharmonic Taylor series. The next step is to divide the grid in half, in the X and Y directions. The data points and the previous grid-node values are then assigned to the nearest grid nodes. The biharmonic smoothing process is then applied again to the grid-node values. The process continues until the final grid-node-spacing value is reached. Also, at any point, additional data can be introduced during the gridding process. Once the gridding algorithm converges on a solution, contouring can begin (Haecker, 1992).

For the final maps at 1:24,000 scale, the geologist used a final grid interval of 2,000 feet in the X and Y directions. This grid interval was generally small enough to ensure that every data point would be honored by the structure contours. Once the final grid for each structure-contour map was created, a macro routine was used to cut out the grid nodes within a 7.5-minute quadrangle area, from either the 1° x 2° block (Phase 1) or from the 30 x 60 minute quadrangles (Phases 2 through 4b). In order for contouring to extend to the boundary of the quadrangle, the grid nodes up to 2,000 feet away from the quadrangle boundary also were cut out during the process. After contouring was completed, the contours were cut off at the quadrangle boundaries to create a publication-quality map.

MITIGATION OF EDGE EFFECTS

Even though the convergent gridding algorithm minimizes edge effects, it does not eliminate them. It was necessary for ODGS to devise techniques to minimize these edge effects. The techniques involve using data from outside the mapping area, gridding outside the mapping area, using pseudodata or projected points in order to constrain edge effects at the outcrops, and using isopach thickness as control on the overlying/underlying units. Each of these techniques had a positive result in minimizing edge effects.

The first two techniques, the use of data points outside of the mapping area and gridding outside the mapping area, used in conjunction with one another are well known in the mitigation of edge effects (Davis, 1986). Figures 4 and 5 show an example of the two techniques for the Marion 30 x 60 minute quadrangle and the 32 7.5-minute quadrangles within it. In the first technique, data points both within and outside the mapping area are used in the gridding process (Figure 4). The data points outside the mapping area provide control up to the edge of the area to be mapped. The second technique, gridding outside of the mapping area, is shown in Figure 5. Figure 5a shows what happens when gridding occurs up to the boundary of the intended mapping area. Typically, geologists using mapping software will create a grid only within the intended mapping area to allow the contours to stop at the boundary of the mapping area. Unfortunately, this method produces edge effects within the intended mapping area. By expanding the gridding area beyond the mapping area, the edge effects are moved outside the mapping area. In Figure 5b, the gridding was expanded out by one row of 7.5-minute quadrangles. By expanding the gridded area, the edge effects are moved away from the intended mapping area.

The next technique is the use of pseudodata points along and outside of the outcrop line. Serious edge-effect problems occur when the subsurface structure-contour mapping approaches the surface outcrop line. Because the unit being mapped does not exist beyond the outcrop, there is no way to control the subsurface structure-contour mapping close to the outcrop. The use of pseudodata points near the outcrop can minimize contouring errors by providing additional control for the mapping software (Figure 6). An elevation is supplied to the pseudodata point by a number of methods. If there is sufficient outcrop exposure, the geologist may assign an elevation to the point based upon the nearby outcrop elevation. Otherwise, an elevation is assigned to the pseudodata point using a preliminary structure-contour map. Projections are

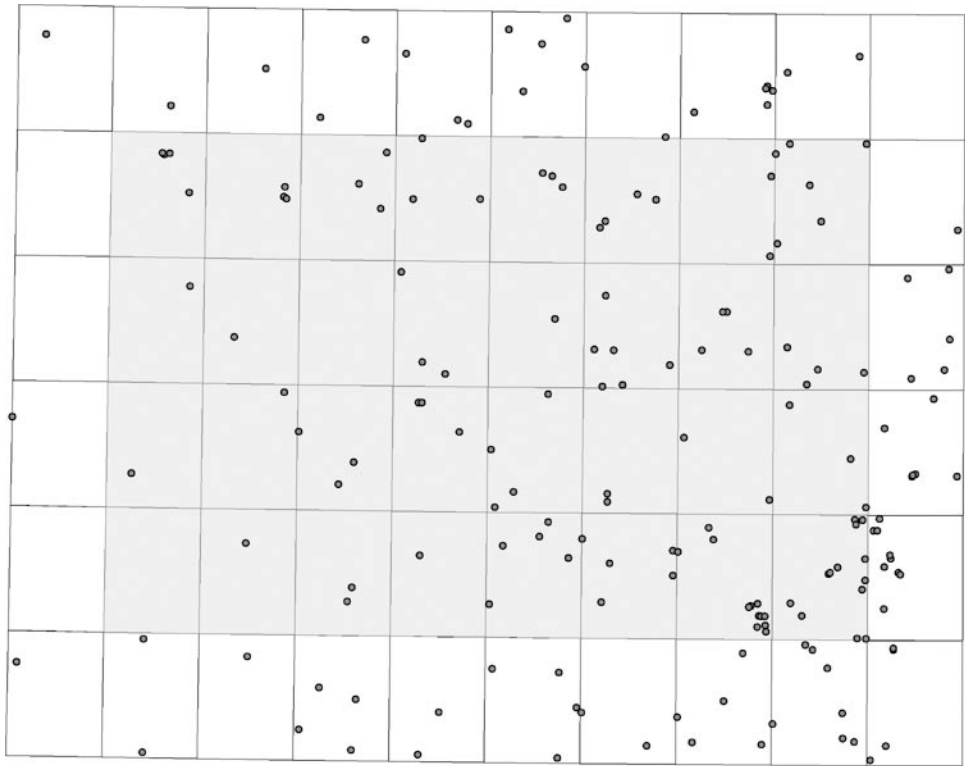


Figure 4. Example of a mapping area and the data points used to create the structure-contour map. The structure-contour map is created using data from both inside and outside the mapping area. This pulls the edge effects away from the project area. In this figure, the mapping area is the Marion 30 x 60 minute quadrangle.

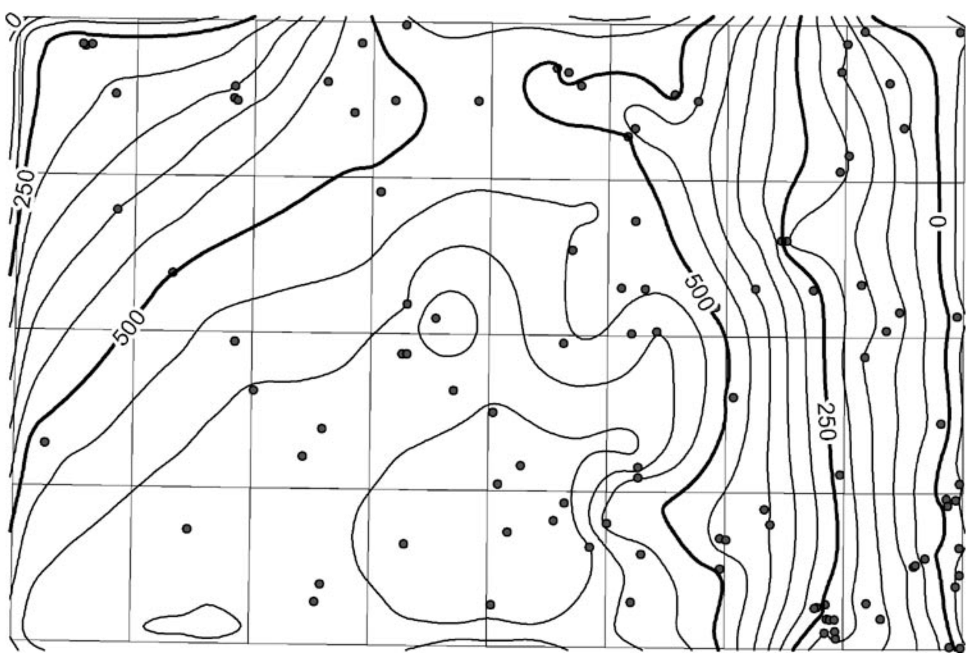


Figure 5a. Example of a grid area that is the same size as the mapping area. In this example of the Marion 30 x 60 minute quadrangle, the edge effects are most noticeable on the western boundary. Edge effects are less noticeable on the northern and southern boundaries. Because there is a well-pronounced dip into the Appalachian Basin on the eastern boundary of the Marion 30 x 60 minute quadrangle, the edge effects here are not noticeable.

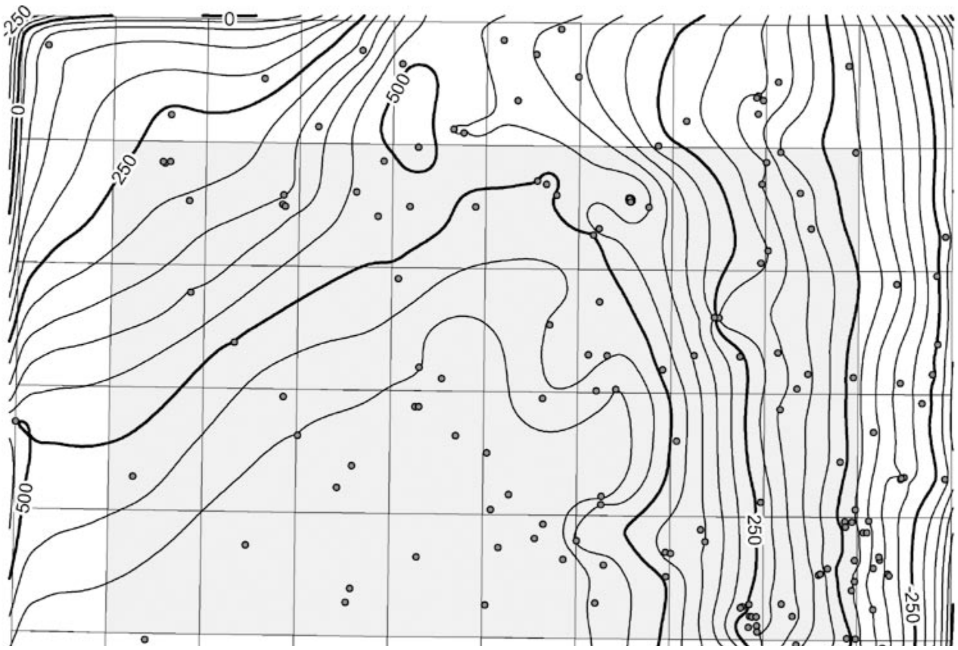


Figure 5b. Example of a grid that is larger than the mapping area. Edge effects are minimized because they are located outside the actual mapping area.

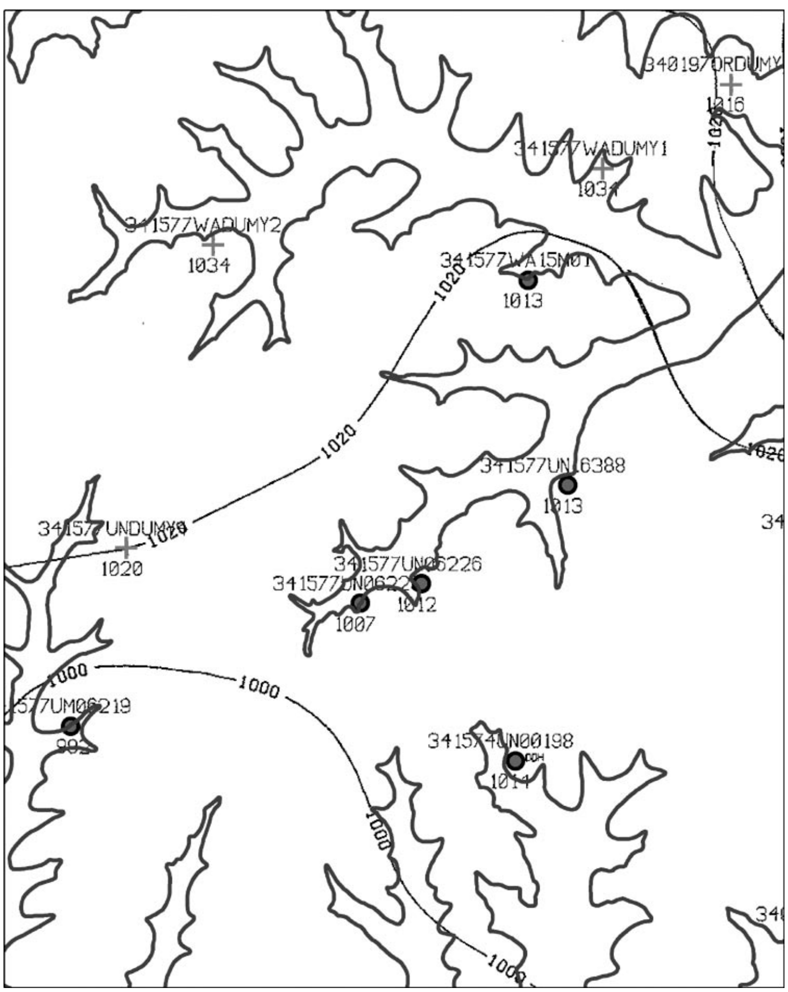


Figure 6. In this example, pseudodata points and data points are used to help create the structure-contour map. The light solid lines are the structure contours, the heavier solid line is the bedrock contact, the solid-filled circles represent the data points, and the crosses represent the pseudodata points.

made out into space, to the pseudodata point, using the preliminary computer-generated structure-contour map. The elevations of the pseudodata points are then adjusted upward or downward on the basis of a structure-contour map of an underlying unit and an isopach map of the unit being mapped. The modified pseudodata points are then included in the database and a second round of computer mapping is performed. Iterations of this technique continue until an acceptable structure-contour map is created.

EDGE-MATCHING

Edge-matching is one of the most difficult tasks for computer mapping software. The final products for the bedrock-mapping program were the individual quadrangles at 1:24,000 scale. In order to edge-match adjacent 1:24,000-scale quadrangles, each geologist needed to create structure-contour maps for individual units over entire 30 x 60 minute quadrangle areas. Because the mapping was continuous over these larger areas, edge-matching was eliminated for the 1:24,000-scale quadrangles within the 30 x 60 minute area. In order to edge-match adjacent 30 x 60 minute quadrangles, geologists used two different techniques. In the first technique, geologists used data points outside their mapping areas, similar to the technique used to minimize edge effects. Generally, geologists used data points from approximately two quadrangles away from the outer boundary of their study area as part of their data set. A geologist would get data points from other geologists working around their assigned areas and also from prior years' project areas. (See Figure 4 for an example of a project area and the data points.) A second method was to use the grid nodes from adjacent project areas as input points at the last stage of the gridding process. As explained previously, the convergent gridding algorithm can accept data at any point during the gridding process. By including the surrounding grid nodes in the last two to three passes of the gridding process, the algorithm allowed the structure contours to cross seamlessly from one project area to the next. The introduced grid nodes only affected the grid nodes closest to the edge of the mapping area.

In some cases, using data and grid nodes from two 1:24,000-scale quadrangles beyond the project areas was not enough to remove edge effects. In addition, the geologists did not always use the surrounding grid nodes as part of their gridding process. In either case, the contours did not match the previously created structure-contour maps. If that was the case, after the 1:24,000-scale structure-contour map was created, the geologist would edge-match the structure-contour maps by hand. While there were always instances where edge-matching between different mapping areas was done by hand, the techniques described above helped minimize the amount of edge-matching that had to be done during this program.

FINAL OUTPUT

The final step in creating the structure-contour map was to import the map into the EasyCAD software package. Once in the CAD software, geologists would add a predefined legend; add the appropriate author, date, and title to the title block; and finally adjust the line color and thickness of the contour lines. At this point, if a geologist did not like the interpretations that the contouring software had created, he would also have the option of modifying the structure-contour lines. Once the legend had been created and the edits made, the map could be printed out on mylar at 1:24,000 scale. The final map was then placed on open file (Figure 2) and used in the drawing of the bedrock-geology contact lines.

CONCLUSIONS

The process used during this project allowed for rapid structure-contour mapping throughout the state, which in turn allowed the geologists to more rapidly draw bedrock-geology maps using the bedrock topography and the surface hypsography. Project geologists estimated that by using the computer gridding and contouring software, they reduced the amount of time required to create each structure-contour map by a number of days to weeks. More than 1,840 structure-contour maps were created in seven years.

One of the most serious problems with using computer gridding and contouring software is that the software algorithms produce edge effects. The gridding algorithm used for this project, when assisted by the techniques described in this paper, tended to minimize the edge effects.

Techniques also were created to perform edge-matching using the computer. These techniques allowed for the individual project map areas to be merged together somewhat seamlessly, in addition to minimizing the amount of edge-matching along the border of the mapping areas. These techniques were generally successful, but edge-matching between the project areas by hand (in CAD) still had to be performed. While this was an inconvenience, the use of the computer to perform the contouring routines eliminated the need for edge-matching among the 1:24,000-scale maps within each project area and greatly sped up the mapping process.

The use of computer gridding and contouring allowed for multiple maps to be made quickly, seamlessly, without edge effects, and plotted in a professional-looking output. The geologists could generate multiple hypotheses and see which one allowed for the best geologic interpretation. If the parameters that were selected did not produce a geologically realistic structure-contour map, then other parameters could be selected rather easily.

REFERENCES

Bernhagen, R.J., 1994, A state geologist's memories of the Morrow County Boom, *in* Schafer, W.E., ed., Morrow County, Ohio "Oil Boom" 1961-1967 and the Cambro-Ordovician reservoir of central Ohio: Columbus, Ohio, Ohio Geological Society, p. 79-81.

Bownocker, J.A., 1920, Geologic map of Ohio: Ohio Division of Geological Survey Map 1, scale 1:500,000, 1 sheet.

Collins, H.R., and Smith, B.R., 1977, Geology and mineral resources of Washington County, Ohio: Ohio Division of Geological Survey Bulletin 66, 134 p.

Davis, J.C., 1986, Statistics and data analysis in geology (2nd ed.): John Wiley & Sons, 686 p.

DeLong, R.M., 1965, Geology of the Kensington quadrangle, Ohio: Ohio Division of Geological Survey Report of Investigations 55, scale 1:24,000, 1 sheet.

Haecker, M.A., 1992, Convergent gridding: a new approach to surface reconstruction: Geobyte, v. 7, no. 3, p. 48-53.

Sherman, C.E., 1933, Miscellaneous data: Ohio Cooperative Topographic Survey, Volume IV of the Final Report, 327 p.

Digital Geologic Field Mapping Using ArcPad

By Gary P. Edmondo

MinGIS
15540 Minnetonka Circle
Reno, NV 89521
Telephone: (775) 853-3933
Fax: (775) 853-3973
e-mail: g.edmondo@worldnet.att.net

ABSTRACT

Research into the practicality of digital mapping by Placer Dome Exploration identified hardware and software solutions to enhance the efficiency and accuracy of field work. The goal of the research was to find a lightweight hardware-software system that allows the user to build a digital map from field observations in much the same way as pen and paper methods.

The focus of the research was to minimize the size and weight of computer systems. Systems identified consist of a wearable PC or handheld computer (PDA) and lightweight GPS systems that can be used for a variety of tasks. These systems would incorporate ArcPad, a scalable, customizable field-based data-capture software package. ArcPad provides the user with the functionality needed to build a geologic map in the field and with data-capture methods to allow mapping to fit the needs of the project. Field testing has shown the hardware-software systems to be effective field-based solutions for the Earth sciences.

INTRODUCTION

Innovations in computer and global positioning system (GPS) technology have allowed for digital data capture and mapping in the field. Workers in the surveying and utilities industries have used this technology since the late 1980's to input data or to correct, modify, or build maps directly in the field, saving large amounts of time needed to postprocess or input hard-copy data into a digital format. Geologic mapping, however, has specific requirements. A geologic map must represent in two dimensions the four-dimensional relationships defining Earth history. To do so, spatial relationships and detailed notes regarding lithologies, alteration, structure, and stratigraphy must be recorded, often requiring very detailed observations. Currently, digital geologic maps are produced after the field geologist completes a map on paper and in notebooks, where the type of information to be captured is extremely varied.

Digital capture of geologic data in the field requires very complex graphical and database-management models. Early attempts at digital data capture were designed to input field data straight into database tables (Brodaric and Fyon, 1989) or into data collectors from early GPS units and other types of survey equipment (Walsh and others, 1999). Both of these methods required extensive processing in the office to produce a spatial map, negating any time savings gained by digital field capture. A mapping system was needed that captured and then displayed geologic features in real time.

Recent advances in computer technology and geographic information system (GIS) software have allowed for more complex data models to be captured digitally on personal computers (Kramer, 2000). Due to these innovations, Placer Dome Exploration began to research the practicality of digital field mapping and what is needed and currently available to accomplish the task. Placer Dome hired me as a consultant to review their efforts and make changes or recommendations as needed. This paper is a brief reporting of those results and changes.

MAPPING SYSTEMS

Placer Dome's answer was that digital geologic mapping is the integration of computer technology with GPS or other location devices for building a GIS-based geologic data set in real time. The research goal was to find a mapping system that is truly portable and capable of data storage, data processing, data capture, location, and digitizing, and a software package to allow the geologist to map field observations in much the same way as pen and paper methods. These parameters were identified in earlier tests of field-mapping equipment.

Hardware Systems

Field-mapping systems have been around for several years (Brodaric and Fyon, 1989; Kramer, 2000; Schetselaar, 1995). In the past, hardware systems (computer, cables, batteries, pens, backpacks, GPS, and possibly a laser rangefinder) weighed in excess of 7 kg and had limited battery life. The systems worked well when vehicles or infrastructure were available nearby to store additional batteries and equipment used during the day. Geologic mapping, however, requires long hikes, in many cases miles from a vehicle and hundreds of miles from the nearest town, thereby limiting the size and type of equipment a geologist could effectively carry in the field. Therefore, the ideal hardware system for any mapping project would be lightweight, have an easily viewable display, and yet be small enough to be readily handled. The unit must also be powerful enough to handle all data-storage and processing requirements, have full connectivity to a combination of different instruments, and use a minimal amount of battery power.

Portability is not the only factor. A geologist needs to be unencumbered during his work, so cumbersome systems (e.g., those with a lot of external cabling or large attachments) had to be removed from consideration. Additional criteria included internal vs. external batteries, battery type, and screen readability. Screen visibility is an important issue and is most apparent in bright-light conditions. However, depending on screen technology, visibility can also be an issue in low-light conditions as well (e.g., nonbacklit transflective screens). The best screen technology is a backlit transflective screen used in several computers, such as the Hewlett Packard iPAQ Pocket PC, and this technology is being applied by more and more hardware manufacturers on their units. Ruggedness also can be important, depending on the environment. Harsh environments (tropics, jungles, cold weather) may require the use of rugged machines; less harsh environments (deserts, subtropics, Mediterranean climates) allow the use of moderate to nonrugged equipment. Table 1 summarizes some of the critical factors in evaluating hardware.

Selecting a computer-GPS system for field work is as much a matter of personal preference as it is capability. With the evolution in computer technology, it is likely the number of available choices will increase every year, with units becoming smaller, lighter, and more powerful. Acceptable computers can be categorized into two basic processor classes. These are PC computers with x86-based processors running Microsoft Windows 95, 98, or 2000 as an operating system, and devices using various forms of RISC-based processors (PDA's, Web tablets) running Windows CE or Pocket PC operating systems. Three workable configurations exist for the field geologist. These are tablet PC's and PDA's, wearable PC's, and handheld (PDA) devices. PC-based systems provide their main benefit where large amounts of data will be generated and the need for storage and processing power are critical. It can be difficult to carry tablet-style PC or RISC-based units in the field, particularly when working outcrops. Also, battery life is typically poor on PC-based units. In contrast, handheld PDA devices do not yet have the processing power or storage capability to handle the very large data sets or large data-capture requirements in detailed mapping and project work. Nor do handheld PDA devices have sufficient screen area to manage interpretive mapping over a large physical region. RISC-based tablet and PDA devices provide their main benefit where data acquisition needs are minimal, such as small-scale mapping or sampling work in which point data and simplified maps are generated. Handheld PDA devices can be easily carried and used in the field. These smaller machines tend to have longer battery lives and are more cost effective. With the advent of secure digital storage cards, PDA devices can now hold a large amount of information. Recent hardware innovations have brought several lightweight, wearable PC-computer systems (such as the Xybernaut MA V) to the forefront of mapping technology. These smaller PC-based units are a good mix between PC tablets and handheld PDA units and offer the best combination of data capture and storage ability versus the weight, carrying ability, and battery life of the computer.

Table 1. Factors used to aid in hardware selection for digital field-mapping systems.

Requirement	Determining Factors
Size, morphology	Proximity to infrastructure, vehicles; carrying capability
Weight	Physical carrying restrictions
Battery type/life	Operating time frame, cost, added weight
Processor and data storage	Amount and type of background data to be used, and amount of data to be captured
GPS accuracy	Scale of mapping project
Connectivity	Number of peripheral devices to be used (GPS, laser rangefinder, etc.)
Display	Readability in low-light and bright-light conditions, available screen-display size
Ruggedness	Environment (tropics/desert)
Cost	Budget

GPS units fall into three broad categories based on DGPS (differential correction) accuracy: (1) survey-grade receivers with accuracies in the centimeter to millimeter range, (2) mapping-grade receivers that are accurate to the 1-m range, and (3) recreational-grade receivers that are accurate from 5 to 15 m. These accuracies, even for recreational-grade receivers, are typically more accurate than locations plotted by hand on a topographic map at scales of 1:12,000 and smaller. Like computers, GPS receivers also come in a variety of physical configurations. Most survey and mapping grade GPS receivers are configured either as a backpack system with receiver separate from antenna, or as integrated receiver-antenna domes (Trimble Pathfinder Power and Pro XR series). Recreational-grade units are configured as either handheld units (Garmin, Magellan, Trimble), low-power compact flash-card units, or as compact integrated domes.

The variety in computing power and morphology and GPS accuracy and morphology provide for a multitude of combinations available for use in field work. System design, however, should be tailored to the needs and requirements for the mapping task at hand. Computer selection should follow the guidelines as outlined in Table 1, with consideration for the environment to work in, amount and type of data sets to use, and area to be covered. GPS units need to be chosen with accuracy and mapping scale as the overall guiding factors. Except for certain tasks, survey-grade receivers need not be used. Mapping-grade receivers offer the best flexibility in terms of accuracy for detail work and weight for more regional projects. Recreational-grade receivers, without a differential correction, should only be used for mapping work at scales less than 1:6,000. Table 2 summarizes various scenarios, requirements, and recommendations on the type of system to use.

Software

Software is the communications manager between the user, the equipment, and the data. It is the key to building a digital geologic map in the field and office. In the past few years, there has been a significant expansion in available field data-collection software. Most packages have been developed for municipal governments or utilities industries to spatially map, monitor, and update equipment and infrastructure. As such, there are specific features missing from some software packages that render them not suitable for building geologic maps. These involve either the lack of sufficient computer-assisted drafting (CAD) tools (such as snapping or symbology support), or data-management tools (GIS functionality). To build a geologic map, one needs a mapping package with the functionality of CAD software and the database management of GIS.

CAD tools are necessary for the building of the graphical (map) part of the data set. Tools need to be

Table 2. Scenarios, requirements, and recommendations for field digital data-capture systems.

Requirements		Scenario	GPS	Computer
Accuracy	Centimeter	Claim staking, land surveying	Survey grade	Any (dependent on amount of base data and data capture)
	Submeter/meter	Detail mapping at scales of 1:6,000 or greater, detail sampling	Mapping grade	Wearable PC, handheld PDA
	>1 meter	Sampling, reconnaissance	Mapping/recreation	Handheld PDA, tablet PDA
Data sets	Large data set	Advanced project, development, mine site	Mapping/survey grade	Tablet/wearable PC
	Moderate data set	Advanced project, project work	Mapping grade	Tablet/wearable PC or PDA
	Minimal data set	Reconnaissance mapping, sampling	Recreation/mapping grade	PDA
Battery life	Long life	Inaccessible AC power/long distance from vehicle or office	Recreation/mapping grade	PDA/wearable PC
	Moderate life	Inaccessible power, long stay in field	Recreation/mapping grade	Wearable PC/PDA
	Short duration	Nearby power supply/vehicle	Survey/mapping grade	Tablet PC
Weight of system	Heavy (weight not issue)	Infrastructure/mine site	Survey/mapping grade, backpack style, integrated	Tablet PC
	Moderate (weight of moderate concern)	Mine site to detailed project work (with no infrastructure)	Mapping grade, either backpack or integrated	Wearable PC
	Light (weight of high concern)	Reconnaissance, early stage projects, sampling	Recreation grade/ mapping grade	PDA/wearable PC

flexible enough for a variety of different input methods, including GPS input, laser rangefinder input, or hand digitizing, yet be flexible enough such that a variety of methods can be used to capture a single piece of data (such as a fault or contact). CAD tools should also be easy to understand and use, so the field geologist spends less time thinking about what to use, and more time thinking about the geology. The ultimate desire is to provide the geologist the ability to map features as he would on pen and paper.

Data-management (GIS) tools are necessary for the input and storage of field observation notes as well as for project management and setup. Part of the efficiency in digital field mapping is the speed of data entry. Mapping software should support the use of form-based data entry to minimize the amount of typing needed to enter data. Form-based data entry also provides data validation and consistency in values used when data are input by more than one person. GIS tools also are helpful for integrating mapping skills with other data sets (layers) such as geophysics and geochemistry. Field observations and integration of other data sets form the basis for final map interpretation. The overall quality of that interpretation is dependent on the type, quantity, quality, and accuracy of the data being added into the database.

Geologic mapping software also needs to interface with the office software used to set up a mapping project and the software used to produce the final map product. Efficiencies gained in the field are lost if the geologist has to spend time formatting data to interface with the setup and production software. With the current trend of producing maps for Web-based use, integration with a data management and storage system also is necessary. The ideal is a seamless integration of data sets between all three aspects of map production.

Very few mapping packages are available that meet the needs of geologic mapping. Three packages that do meet most of these criteria are ArcPad 6.0 (ESRI), Penmap (Strata Software), and Geolink (Michael Baker Jr. Corp.). ArcPad was selected by Placer Dome because it met most of the user requirements and was a very cost effective solution.

ArcPad integrates near seamlessly with ArcView and ArcGIS as a front-end data-capture solution. Although not perfect (no software package is), ArcPad and the ArcGIS suite of software provide the field user with the flexibility needed to manage all aspects of a field-mapping project, from setup to final map production, and data storage. This integration provides for better data management, better project management, and more efficient data acquisition and manipulation (Figure 1). Geologically, using ArcView and ArcGIS to establish a mapping project in ArcPad provides significant benefits, including the ability to develop and use custom symbols, establish symbol rotation, and construct legends for data values. This capability allows for simple mid-project changes in symbology and/or data

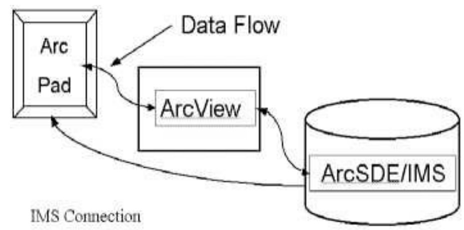


Figure 1. Diagram showing the relationship of ArcPad to the ArcGIS suite of software in a scalable data-management solution.

values, additions or deletions of data layers, and complex editing tasks.

In addition to ArcPad's integration with ArcView and the ArcGIS suite of software, ArcPad has additional strengths. ArcPad is fully customizable. Using ArcPad Studio, standard form files, standard layers, and customizable pick lists can be built into a template (Figure 2). The template structure allows for automation of the ArcPad command process and provides for more efficient data capture. Data can be captured in either two or three dimensions, and spatial and attribute value editing of previously existing data is possible. ArcPad supports PC (Windows 95, 98, 2000) and PDA (Windows CE, Pocket PC) operating systems, allowing for the same customization to be used on different field-unit platforms. This is beneficial when several people will work on the same project, but have different tasks to accomplish, or in projects that grow from small, regional-scale programs to advanced mapping projects.

ArcPad's capabilities for geologic mapping are best suited for fact-based observational mapping. It has been field tested on several mapping and sampling projects, including mine-pit mapping (Figure 3) and surface mapping (Figure 4). These two examples were built entirely within the ArcPad-ArcView-ArcGIS environment using a standard set of template files in a customized menu structure. They show some capabilities of ArcPad. Although commonly used in an observational (outcrop or fact-based) mapping environment by the mining industry, ArcPad is not limited to that function. Other uses for ArcPad include standard geologic mapping (interpretive), detail mapping at scales of 1:6,000 or less, geophysical surveys, geochemical surveys, environmental studies, and structural studies. Because of the customization environment, ArcPad can be used for any aspect of mapping required by the Earth sciences.

CONCLUSIONS

Research by Placer Dome identified portable computer and software systems that allow the geologist to map in a setting similar to pen and paper. This was only possible by recent technological advances that have brought to

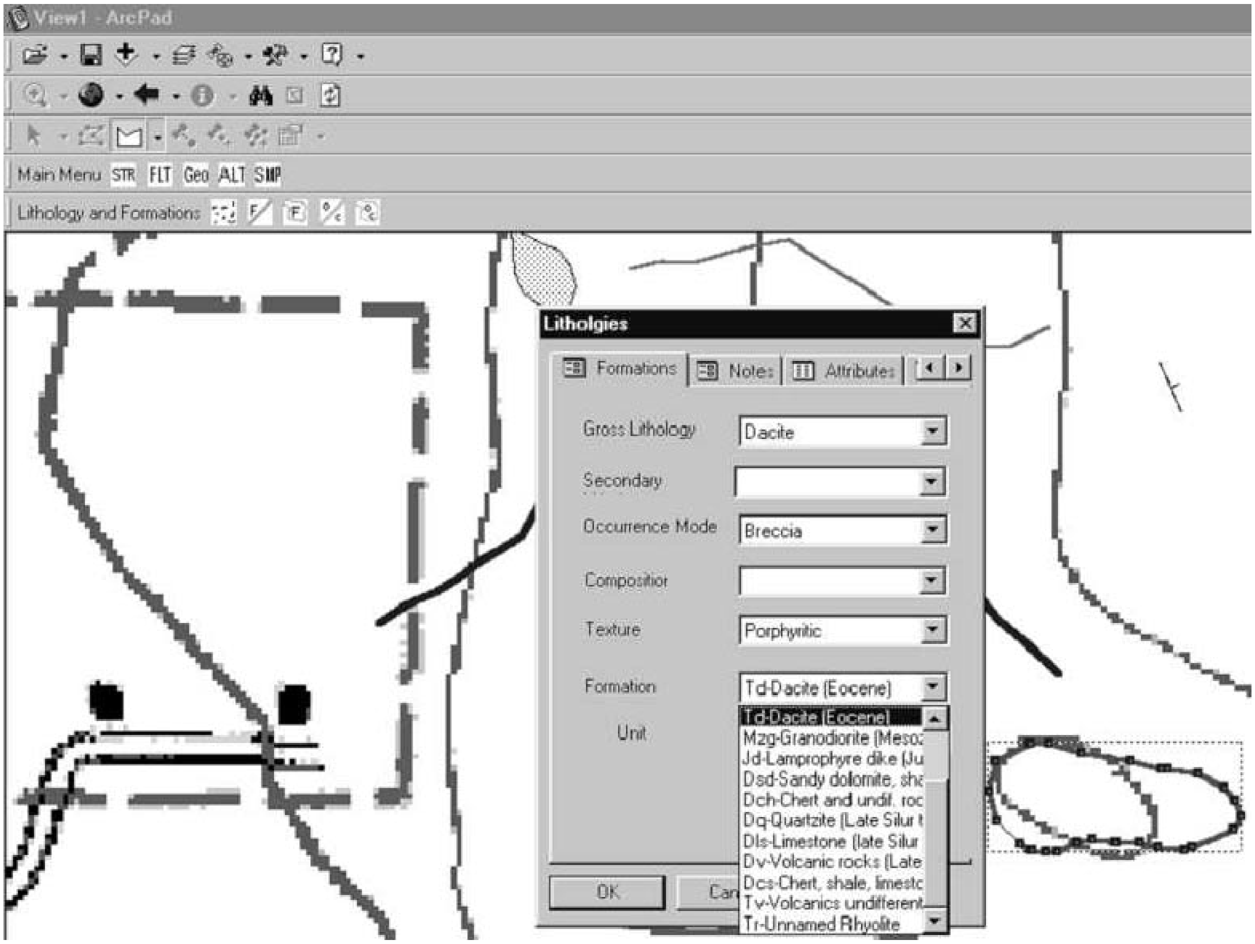


Figure 2. Features in a customized ArcPad environment.

industry smaller and more powerful computers, smaller and more accurate GPS units, and GIS data-capture software, such as ArcPad, that has CAD and GIS database functionality. This dual functionality allows for easy data capture and simple integration into GIS software, such as ArcView and ArcGIS. These developments have allowed real-time digital mapping for the geologist to become a reality.

GIS data-acquisition software is currently undergoing a technological evolution, with many new software packages coming to the market. This is in direct response to availability of hardware and the need for efficient data collection and upkeep for field-based industries. In the future, these new software developments, and the trend toward smaller, more powerful computer and GPS devices will continue to provide even greater enhancements to digital geologic field mapping.

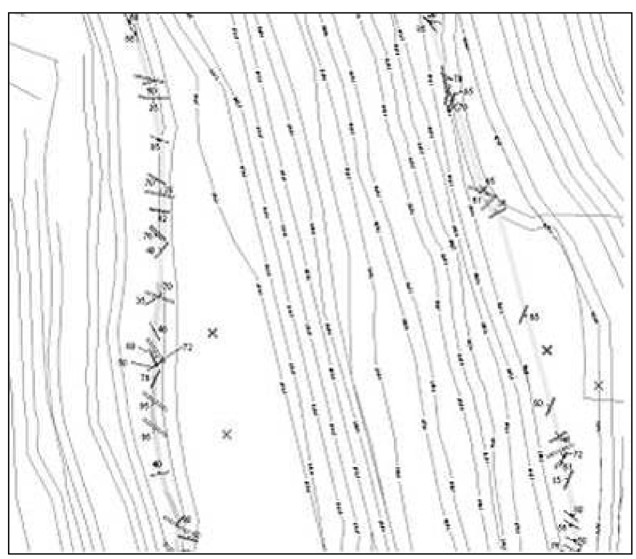


Figure 3. Example of a mine-pit map, created using a Xybernaut wearable PC, Pathfinder Power GPS, and ArcPad (courtesy of Placer Dome).

Figure 4. An example of surface observational mapping, created using a Compaq iPAQ and compact flash GPS with ArcPad (courtesy of Cortez Gold Mines).

REFERENCES

Brodaric, Boyan, and Fyon, J.A., 1989, OGS FIELDLOG: a microcomputer-based methodology to store, process and display map-related data: Ontario Geological Survey Open-File Report 5709, 73 p.

Kramer, J.H., 2000, Digital mapping systems for field data, *in* Soller, D.R., ed., Digital Mapping Techniques ‘00—Workshop Proceedings: U.S. Geological Survey Open-File Report 00-325, p. 13-19, <http://pubs.usgs.gov/openfile/of00-325/kramer.html>.

Schetselaar, E., 1995, Computerized field-data capture and GIS analysis for generation of cross sections in 3-D perspective views: Computers and Geosciences, v. 21, p. 687-701.

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., ed., 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>.

Techniques for Improved Geologic Modeling

By Donald A. Keefer

Illinois State Geological Survey
615 E. Peabody Drive
Champaign, IL 61820
Telephone: (217) 244-2786
Fax: (217) 244-2785
e-mail: keefer@isgs.uiuc.edu

INTRODUCTION

It can be a difficult process to create computer models of two-dimensional (2-D) surfaces and three-dimensional (3-D) geologic sequences. If mapping projects are not conducted with enough attention to the demands and constraints of the modeling process, excessive expenditures of time and money can result, and the model may not properly address the intended applications.

In this paper I present techniques for addressing some common 2-D and 3-D geologic modeling difficulties that can have a significant effect on the success of the project. I provide some perspectives on the selection of modeling objectives, the use of declustering as a tool for data management, and ways to control surface models so they agree with the interpretations of the mappers. Finally, I touch briefly on some technical issues on the horizon that will further affect how we approach geologic mapping on the computer.

CHOOSING MODELING OBJECTIVES AND DEFINING MODEL PARAMETERS

The first step to avoiding or minimizing difficulties is to define the modeling objectives properly. Following well-defined modeling objectives can help ensure that the final model output is compatible with its intended applications. At a minimum, modeling objectives need to address the following:

- stratigraphy and complexity of the geologic system,
- availability of relevant base map information,
- selected methods for addressing uncertainty within the data and the model,
- intended uses of modeling results,
- spatial distribution of the data,
- data quality and quantity,
- area or volume to be modeled,
- sizes and shapes of anticipated surface features,
- desired minimum size of feature that should be identified,
- grid spacing in each of the three axes (x, y, z),
- available computer hardware resources, and
- capabilities of available modeling and visualization software.

The first four considerations are beyond the scope of this paper. I will briefly address the remaining considerations.

The spatial distribution of data can dramatically affect the ability of an interpolation algorithm to define a realistic geologic surface, particularly a complex one. Typically, sources of geologic data (such as water-well logs) are not uniformly distributed throughout a map area, but instead are clustered. The location of the clusters may be strongly dependent on the reasons the data were collected, and these reasons may or may not be related to the occurrence of specific subsurface geologic deposits. For example, water wells in some locations may be clustered above channelized sand and gravel deposits. In other areas where ground-water resources are not spatially limited, the clustering of water wells may be more a reflection of urban-suburban development. The presence, density, and location of data clusters and their potential impact on surface models should be characterized in the early planning stages of the project.

Variations in data quality will affect the uncertainty of any resulting model. The sources of uncertainty, the mapping of changes in uncertainty, and the potential impacts of this uncertainty on the model need to be evaluated and documented. Following this evaluation, it is important to determine whether the intended applications can be sufficiently addressed by model results.

The area or volume to be modeled usually is well defined, but the boundaries of the formal mapping area may be irregular. Gridding software packages, however, typically use square or rectangular grids, which may require that a much larger area be included in the overall model. To define the distinction between the grid area and the mapping area, most modeling software provides some method for blanking out grid cells that fall outside the boundary of the desired mapping area. The grid coordinates and spacing and any blanking or map-area boundary need to be firmly established early in the project if more than one surface model is going to be developed. Changing grid coordinates or study-area boundaries after a project has started can create time delays and data losses.

Determining the minimum feature size that will be included in the model is important because it affects decisions about grid size and optimum data spacing. Geologic surface features are typically asymmetric, and data points cannot be expected to fall directly on the highest and lowest values of each feature. We can consistently describe the general shape of a feature if the data are spaced at approximately one-third of the shortest axis of that feature. If the first point lies on the edge of the feature, a data spacing that is one-third of the shortest axis length will result in four points along the direction of the shortest axis (e.g., point 1 = 0, point 2 = 1/3, point 3 = 2/3, point 4 = 3/3). For example, if the average separation distance between data points is 0.25 mile, then, on average, we will be able to identify features that are 0.75 mile wide in their shortest direction. Although more closely spaced data points will obviously identify smaller surface features, we still need at least four points along each of their principal axes to identify these smaller features. Ideally, the determination of the minimum size of a surface feature should be based on more than just the average data density. This determination should be based on the complexity of the geologic deposits (i.e., the sizes and shapes of anticipated subsurface features), the extent and location of data clusters, the variation in data spacing within these clusters, and the intended applications of the model.

Determining the minimum feature size that will be recognized by the model also is relevant to the selection of the grid spacing. Just as a minimum of four data points is needed in each of the two principal directions to characterize a surface feature, a minimum of four grid cells is needed in each of these same principal directions to adequately model or express each surface feature. It can be helpful to have a grid spacing that is several times smaller than the minimum data spacing. This will create surface models with smoother, more realistic surface morphologies. The interpolation algorithm chosen, however, may place a practical lower limit to the grid spacing. Some algorithms (i.e., splines, minimum curvature) will produce oscillations, or interpolation artifacts, in the surface model if the grid spacing is significantly smaller (i.e., about 10 times) than the data spacing. In this situation, the selection of the optimum grid size requires trial and error.

The total number of grid cells can be very large, especially if 3-D models are being developed, and the grid resolution may have to be coarsened for model computations to be completed within a practical time limit. Selection of grid spacing in the vertical direction should be based on the anticipated minimum thickness of the individual deposits and the total thickness of materials to be modeled. Thick sequences of deposits will probably require more generalization than is desired in the vertical direction because of the large number of grid cells that are generated by finer resolutions. For example, a 2-D grid that is 267 x 184 cells has a total of 49,128 cells. This is a manageable grid size for typical desktop computers. If we are modeling a 400-foot sequence of geologic materials and want to delineate units that are 5 feet thick and greater, this would require 80 cells in the vertical direction and produce a total model of 3,930,240 cells. Depending on how the specific software actually constructs the model, this may be too many grid cells for many desktop computers. A model of this size may even exceed the limitations of the modeling software. Large computer resources will definitely be required if significant visualization and slicing of very large models is desired.

A final consideration in defining the modeling objectives is identifying the software resources that will be used to address various data management, surface modeling, grid manipulations, 3-D model construction, model visualization, and final product development. Although high-end modeling and visualization software can address all these facets, the cost of these packages is too high for most modelers. Many moderately priced software packages available for desktop computers can accomplish one or more of these tasks. Careful evaluation of the available suite of software packages, their interoperability, and the specific roles that each will play may identify possible software incompatibilities before they cause project delays.

DECLUSTERING AS A DATA-MANAGEMENT TOOL

Understanding Declustering

One data-management practice that can save a lot of time and money in modeling projects is spatial declustering, a process of sorting through or reducing the number of data points to a level that is more efficient and effective for modeling. A data set can be considered clustered if either a number of grid cells have more than one data point or part of the map has data spaced more closely than necessary for identifying the minimum-size feature specified in the modeling objectives. The advantages of declustering and the methods available for it need to be evaluated early in the project to determine whether the process is worthwhile for the specific study area.

There are three major advantages to declustering a data set. The first is the savings in time and money. With any data set, time must be spent validating and correcting the locational information of each data point. This is a critical step in mapping or modeling because locational errors not only place the geologic data in the wrong horizontal location, but a significant vertical shift of the unit tops can occur if the elevation of the top of the borehole is significantly in error. In addition, lithologic descriptions, geophysical properties, or other properties encountered in the borehole must be reconciled within the local stratigraphic framework. Unless a data set is declustered prior to undertaking these steps, many more data points may be processed than can be effectively used in the modeling process, causing significant increases in time, effort, and expense. Establishing optimum data spacing prior to data evaluation can help reduce the number of data points that must be validated and correlated. In areas of complex sedimentary successions, some additional data points may be needed to clarify the stratigraphic framework and to increase the reliability of all stratigraphic correlations and the resulting model. However, these additional stratigraphic control points need not be included in the modeling data set unless they provide data quality that is superior to adjacent data points.

The second advantage to declustering relates to the way data are used in most interpolation algorithms. In any grid model, only one value can be assigned to represent each grid cell, even when two or more data points fall within a single grid cell. Unless some method is used to decluster this cell and define a single value, the value assigned to this cell will be some combination of all the clustered values. If the clustered data in the cell have a large range of values, the value assigned to the cell is likely to be skewed toward the most extreme value in the cluster. Thus, clustered data sets are inefficient because of the redundancy of information and inaccurate because of the inability to fit the surface to each observed value.

The third advantage to declustering is that many interpolation algorithms produce surface models that are severely biased by clustered data. This bias is likely to be more severe in areas where the clustered values vary widely and along the margins of data clusters where grid-cell calculations are unduly influenced by the large number of points in one area of the search neighborhood. Some interpolation algorithms handle clustered data better than others. Any algorithm that uses some form of curve fitting between calculated and observed values to determine the cell value should handle clustered data better than those that do not employ such curve fitting. Some algorithms that fall into this category include splines, minimum curvature, kriging, and other radial basis functions. Although these algorithms are not immune to complications from data clustering, they still can be strongly affected if the clustered data have high local variability. Algorithms that use a simple distance weighting function to assign a cell value generally do not work well with clustered data (e.g., algorithms that rely heavily on methods referred to as inverse-distance calculations).

In any mapping effort, there is a natural tendency to assume that more data are always better. In fact, there is always a limit at which additional data do not contribute significantly to the map or model. This limit can be determined by considering the minimum feature size desired for the model. Data that are spaced more closely than one-third of the short axis of the minimum-size feature will not contribute much useful information to a surface model. Additional data collected within this spacing limit can be considered to be too much data from a cost-effectiveness perspective. Additionally, if data can be grouped on the basis of probable reliability, the addition of clustered, poorer quality data may only reduce the influence of adjacent higher quality data, thereby reducing the reliability of the model. Where the natural variability of the surface at the local scale (i.e., small surface features) is similar to or smaller than the likely measurement error of the data, real features will be indistinguishable from data artifacts. These insights suggest that, for most purposes, no significant modeling advantage will be gained from the time and resources spent validating, interpreting, and modeling with too much data.

However, if the project is trying to address uncertainty of the modeling results, then it may be worthwhile to keep additional data, even highly clustered data, in the data set. Although clustered data may be helpful for some uncertainty evaluations, they are not effective or efficient for most surface-modeling efforts. The project objectives and uncertainty evaluations can be used to determine when sufficient data have been gathered. This determination can be made before the locational verification and correlation efforts are begun.

Choosing a Declustering Method

There are several good ways to implement declustering. If the data have inconsistent accuracy and information content, then it is likely that some data points provide more value to the mapping effort than others. In this situation, it makes most sense to integrate some type of data valuation into the declustering method. With the sophisticated capabilities of spreadsheet and database software, a customized data valuation can be conducted systematically on the entire data set.

Some data points may provide valuable information for only one or two surfaces; immediately adjacent points provide better information for other surfaces. For this reason, it may be advantageous to decluster the data set once for each stratigraphic surface being modeled. Logistically, this may be tricky for some model parameters (e.g., lithostratigraphic interpretation) because a data valuation might need to be conducted prior to the correlation of lithologic deposits within the stratigraphic framework. One viable solution is to initially conduct the data valuation for the entire borehole, without worrying about individual surfaces; this will allow the entire data set to be declustered. After declustering, the removed, or unused, data can be kept in a separate file. During the modeling process, if the data control for a specific surface is poor using the declustered set, the unused data set can be queried to identify possible alternative points.

One method of declustering a data set is to either pick one of the clustered data points within each grid cell or to calculate some statistic from the clustered values (e.g., mean, median, root mean square error) and use that single value as the representative datum for the grid cell. Some surface-modeling software programs provide options for this type of declustering and may refer to this type of clustered data as Aduplicate data.@ Many of these packages allow the user to define the distance within which data points will be considered duplicate, or clustered. They allow the user one or more of the above-mentioned methods for assigning a single value to represent these duplicates. Another declustering method is called cell-based declustering. In this method, a grid is overlaid on the data, and grid cells with multiple data values have these values transformed into a single value. This single value is typically the mean of the clustered values. Each software package is likely to differ slightly in the options it provides for treating clustered data. If the software package you are using does not provide any of these options, declustering can be done using either a geographic information system (GIS) or a spreadsheet.

Using a GIS, a grid of user-defined dimensions can be created. The data points then can be overlaid with the grid. Each data point is then assigned the identification number of the grid cell that bounds it. Determining the grid cells that have multiple points identifies data clusters. The user can then apply any criterion to decluster the data.

Using a spreadsheet, a grid can be defined in which each cell would correspond to a row in the spreadsheet. The spatial coordinates of each cell centroid would be defined by incremental changes in the x and y coordinate columns. If a data point is within half the distance to a cell centroid, the identification number of that cell can be assigned to the data point. As with the GIS example, clustered data will have the same grid cell-identification value, and one value per cell can be selected using any appropriate method.

INTRODUCING GEOLOGIC INSIGHT THROUGH USE OF SYNTHETIC DATA

Why Synthetic Data?

Typically, an initial surface model does not completely represent the geologist's interpretation of the surface. This is generally due to the shape constraints that the algorithm uses, to the presence of clustered or highly variable data, or to the lack of data in certain areas. All these problems have some remedies, including the following:

- adding individual synthetic data points,
- coarse to fine gridding,
- grid editing, and
- digitization of hand-drawn contour lines.

These methods basically involve the addition of synthetic, or hypothetical, data. Synthetic data are values not obtained from any type of observation or measurement; they are added to a data set by the modeler to help control the shape of a surface model. The values assigned to synthetic points are based on neighboring values and the project team's geologic knowledge and conceptual model of the surface. The synthetic values may be needed because the available data and interpolation algorithms cannot otherwise be used to acceptably express this conceptual model. It is important to identify and document these data to prevent them from being confused with real observations. This will help ensure that the procedure used to create the surface is easy to repeat and able to more readily accommodate any new and significant data that may be collected in the future.

Geologists who are unfamiliar with computer-based surface modeling sometimes argue about the validity of using synthetic data. In my experience, their concerns arise from either a poor understanding of how computer algorithms create surfaces, a use of concepts and terminology that is different from the modeler's usage, or from misunderstandings about the practice of assigning values to every point in the modeled space instead of only along contour lines. When creating a contour map by hand, the mapper makes assumptions about the continuity and value of the surface at points that have not been sampled. Although a specific value is not assigned to points between contour lines, the limited number of possible surface shapes that can occur between observed data points puts specific bounds on the otherwise unspecified values. This is no different from the use of synthetic points to constrain a surface—even if the synthetic points are not limited to contour lines.

Methods for Adding Synthetic Values

The addition of individual synthetic points can be a fairly simple procedure. Many software packages will allow you to either add points on-screen or to assign the coordinates for a new data point by using the mouse to point to a location on the screen. The geographic coordinates and a desired value can be recorded for the information inserted in the data file for re-interpolation, but synthetic values must be clearly distinguished from observed values.

This method of surface control can be effective for controlling the expression of slopes or bluffs along rivers or gullies. It also can be effective for suggesting the presence of a valuable deposit (such as an aquifer) in locations that are unsampled, but where the occurrence of such a deposit seems likely based on other geologic evidence.

The use of individual synthetic data points allows the modeler to add to a surface model detailed features that reflect significant geologic interpretations that are not evident from the real data alone. Synthetic values are relatively easy to update if new observations or interpretations become available. Using synthetic data points also offers the advantage of allowing you to add a range of values, even around intended break points, thereby preventing interpolation artifacts, such as flattening of the surface, that can occur when contour lines are digitized and used.

Coarse to fine gridding is another technique that can be used to constrain surface models in areas of low data density. This technique can be particularly helpful along model boundaries, where the data density is typically low and surface models may be unduly influenced by a few local points with extreme values (Jones and others, 1986). When used to stabilize a surface in more central parts of a map, coarse to fine gridding is most effective for algorithms that rely heavily on simple inverse-distance calculations and appears to have less benefit with curve-fitting algorithms (e.g., minimum curvature). Some algorithms implement an inverse-distance calculation and a curve-fitting calculation. The effect of coarse to fine gridding with these kinds of algorithms varies and should be tested on a few data sets.

To implement coarse to fine gridding, the data are initially gridded using a spacing that is at least five times larger than the desired final grid spacing. The resulting grid should show only the largest features in the modeled surface. The grid is then converted to an ASCII data file and integrated with the data set of observed values. To avoid having these new synthetic grid-based points overprinting a strong regional trend (i.e., generalizing) in areas with a good data density, the synthetic points must be deleted if they are within approximately three grid spacings of an observed data point. Once the unneeded synthetic points are deleted, the modified data set is then interpolated using the final grid spacing. This threshold-separation distance between the synthetic points and the original data points is somewhat arbitrary and can be adjusted on the basis of trial and error.

The coarse-to-fine-gridding method is reproducible if new data become available, thereby allowing the modeler to easily maintain the priority of observed data values and to create synthetic values that are easy to distinguish from the observed values. The disadvantage of this approach is that it only helps to maintain the regional character of a surface and cannot be used to add smaller surface details.

Grid editing can be a fast approach to constraining a model in situations that either require simple changes (e.g., changes to a small number of cells) or involve simple grids. For example, if a 3-D model shows one unit occurring over a larger extent than is desirable, the isopach grid of that unit can be edited so that areas where the deposit should be absent have a zero value. This situation would need additional grid editing to address the change in sediment occurrence, but it illustrates when grid editing might be feasible. This approach can be problematic for typical surface models because it is difficult to create smooth surfaces through editing of individual grid nodes. Also, grid editing can be a slow procedure, although tools may be available within individual software packages to simplify the editing process. The ability to create an acceptable surface through grid editing will depend on the grid resolution, grid complexity, and overall modeling objectives. Although this technique typically does not automatically record the grid changes, it is possible to save the values of only the changed grid nodes in a separate data file, with documentation to describe how and why they were used. The changed grid nodes can be identified by subtracting the original grid from the modified grid; the original values are then added back to all non-zero grid values. The subtraction of the grids will result in non-zero values only where the grid nodes were changed through the manual grid-editing process. These values will be the difference between the original and changed node. A record of the actual values from the final edited grid can be created by adding this difference to the original grid value for only the nonzero nodes.

Digitization of hand-drawn contour lines is another process for controlling surface models. In some desktop software, contour lines can be digitized on the screen using a mouse rather than with a digitizing board or tablet. Typically, these lines must be converted to points, the points must be added to the observed data set, and a new surface model then must be re-interpolated. Some software can use the contour lines directly without converting them to points. Other software packages also allow modelers to drag existing contour lines on-screen; the software automatically modifies the underlying grid model. If the on-screen, line-dragging approach is used and a grid is automatically created, it is important to record the changes to the initially interpolated grid. The procedure outlined in the grid-editing discussion also would work well in this situation. Digitization of contour lines allows the modeler to create any shape in the surface model that seems appropriate. As with all synthetic data techniques, the revised data should be preserved and identifiable from observed values. The time it takes to implement this technique will depend on the software options that are available. Editing a surface model can be accomplished very quickly if you can drag contour lines on-screen and automatically generate a new grid.

With many interpolation algorithms, the use of digitized contour lines or points from these lines creates surface models that have small flat spots, or benches, where the new contour data occur. This is because the use of hand-drawn contour lines typically results in hundreds of new synthetic data points, all with the same value; several of these new points typically fall within the local search neighborhood used to calculate nearby grid cells. This clustering of values with a single elevation has the effect of biasing the grid calculations for all cells adjacent to these clustered points. One way to reduce the expression of this benching artifact is to decluster the number of points used from the digitized contour lines. Ideally, the separation distance used for declustering the original data set would be the most appropriate distance for declustering these values. The time requirements of this technique should be considered for each project. For many projects it may be more efficient to manually add individual synthetic points.

GEOLOGIC MAPPING TECHNOLOGY ON THE HORIZON

The growth in computer modeling of geologic systems is providing geologists with an opportunity to better evaluate and characterize the quality of their data and to present information in ways that are otherwise impossible. This also has the effect of making geologic information much more accessible. Many technical issues will need to be addressed as this technology is embraced more completely. There are two particular technical issues on the horizon of computerized geologic mapping that I think are imminent and worth discussing briefly.

First, advances in surface and volume modeling software are coming at an amazingly fast pace. Desktop software priced less than $1,000 (e.g., Rockworks) provides a fairly robust 3-D geologic mapping and modeling environment; more sophisticated options are available in higher priced software for the Windows and UNIX environments. The availability of many packages makes it increasingly practical to incorporate computer modeling into any mapping effort.

With this growth of modeling, it is important to re-evaluate the focus of mapping projects. Traditional mapping projects focus on a set of map products or perhaps a set of computerized visualizations as the final goal of a project. With computerized modeling supporting the mapping, we can re-frame the goal of these projects to be the development of a set of models that provides a consistent interpretation. A suite of surface and volume models and associated documentation can be created; this suite would include the data and all resulting interpretations for the modeled geologic system. This goal of developing a consistent suite of surface and volume models will have the added benefit of allowing all the graphical products and visualizations to also be consistent. Currently, differences in the location of specific contacts and the geometry of surfaces can readily occur when each surface is created as a separate product. The consistency gained from developing all products from a single model will reduce the total uncertainty of the project's results by reducing the potential for inconsistent display of stratigraphic contacts. This reduction in uncertainty may be a significant benefit to end users who make a range of decisions from a suite of maps.

Second, parallel advancements in development of more sophisticated methods for producing computer-generated maps are leading to the creation of data models and object-based map construction (Hastings and Brodaric, 2001). We need to begin merging the data-model concept with surface and volume models to develop a more complete geologic data model.

REFERENCES

Hastings, J.T., and Brodaric, Boyan, 2001, Evolution of an object-oriented, NADM-based data model prototype for the USGS National Geologic Map Database Project: International Association of Mathematical Geology, Proceedings, Annual Meeting, Cancun, Mexico, September 9-12, 2001.

Jones, T.A., Hamilton, D.E., and Johnson, C.R., 1986, Contouring geologic surfaces with the computer: Van Nostrand Reinhold, Computer Methods in the Geosciences, 315 p.

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

By George H Brimhall, Abel Vanegas, and Derek Lerch

Earth Resources Center Digital Mapping Lab
Department of Earth and Planetary Science
University of California, Berkeley
307 McCone Hall
Berkeley, CA 94720-4767
Telephone: (510) 642-5868
Fax: (510) 643-9980
e-mail: brimhall@socrates.berkeley.edu

INTRODUCTION

Publication of the first colored hand-painted geological map of England and Wales by William "Strata" Smith in 1815 heralded the birth of modern geology and acceptance of its foundational principles including superposition (Winchester, 2001). Since then, geologists have drawn maps in the field with a paper and pencil technology largely unchanged for almost two centuries. Even long after the transition to digital map production in offices in most geological surveys in the U.S. and Europe (Jackson and Asch, 2002), field mapping by geologists still clings to the traditional paper technology of the early 19th century. Consequently, the geologists practicing their scientific profession in the field, mapping and interpreting the complex archive of Earth history contained in the rocks, generally do not derive appreciable benefit from the digital and information technological revolutions that have advanced other fields of science to unprecedented levels. Hence, the gap between the paper-based mapping technology of field geologists and portable real-time computing remains surprisingly wide. Why does this technology gap persist today? What capabilities would be necessary to motivate and support a cross-over technology to digital field methods?

Our contribution to this volume is a summary of the present status of our efforts to help bridge this technological gap by producing well-tested and robust mapping software called GeoMapper and GeoLogger which support real-time digital geological mapping and integration of digital base maps and mapping tools. With our programs, complete "paperless" maps with databases that have full compatibility with ESRI ArcMap for map production are made in the field while employing modern digital electronic tools to maximum advantage for positioning and ranging. In the DMT '01 volume, we described (Brimhall and Vanegas, 2001) GeoMapper functionality and use by geologists, including the scientific logic behind the design of the visual user interface, that support the complicated cognitive and reasoning processes of geological deduction. Many previous barriers to workflow have been removed through use of self-explanatory button arrays to access digital base maps and implement efficient mapping of structures (strike and dip, contacts, etc.), lithology, formations, mineralization, and alteration. Button control of GPS and lasers provides integration of base maps, positioning instrumentation, and mapping functions. In GeoMapper, Project Manager sets up mapping region files, and Legend Maker allows a user to easily customize the mapping legend to the local geology using only point-and-click techniques. In this volume, we present a series of applications of GeoMapper to real-world problems to illustrate the efficiency and versatility of GeoMapper and to introduce our new GeoLogger program for digital logging of drill holes.

While we present use of GeoMapper briefly here, we focus on the new features and refer interested readers to the DMT '01 paper (Brimhall and Vanegas, 2001) and to a Web site, <http://www.rubicondigital.com>, where GeoMapper and GeoLogger can be obtained commercially and support services can be accessed. For the first time, surface mapping, underground mapping, and drill-hole logging can be accomplished using digital methods implemented in two integrated and compatible mapping software systems, one for mapping and the other for logging. Map production is done using the most widely used GIS software or alternatively by printing directly from GeoMapper and GeoLogger.

Digital Mapping in the Office: Map Production

The GIS revolution and availability of commercial software that supports map digitization, processing, and printing rapidly propelled office map production well ahead of field data-capture methods. "Digital mapping" in this sense is a process of conversion of original paper-based maps made by geologists in the field to digital record form and for publication on paper, commonly with a digital medium also being available to end users. While many surveys (Jackson and Asch, 2002) and mining and environmental companies have adopted a range of software for their map production process, ESRI products (ArcInfo, ArcView and Arc Map) are the most widely used. MicroStation and MapInfo also are used but to a lesser extent. Given the wide use of ESRI products, we have designed GeoMapper to seamlessly export into ArcMap for production while retaining the well-established advantages of GeoMapper to support a geologist's field methods.

Remaining challenges in office digital mapping surround scientific and technical standards that reduce inconsistencies in the geological legends and in the creation of comprehensive relational databases that support the profound complexities of geology (Soller and Berg, 2001; Soller and others, 2001). Previously, paper geological-map series tolerated inconsistencies between map sheets, but GIS and related digital systems need a more stringent approach (Jackson and Asch, 2002).

Digital Mapping in the Field: Data Capture (Geological Mapping)

We have approached digital mapping technology from a standpoint not of a *conversion* of paper to digital formats but rather of the *creation* of maps directly in digital form with relational databases being created in accord with the mapping legend developed for a project area. This approach circumvents the use of paper altogether while producing a database that is compatible with the most common GIS systems for map production (e.g., ArcMap). Paperless digital mapping eliminates the still common intermediate step between the field geologist and the office GIS staff and hence simplifies map production; it also has the potential for improving productivity and reducing loss of scientific information, which is a recognized problem in industry.

GeoMapper uses Strata Software's PenMap as an underlying program and can be viewed as advanced extension of it with new programmed capabilities to organize projects and customize the mapping legends. PenMap is a powerful surveying program with extensive device drivers to many GPS and surveying instruments and provides the raw graphics elements of points, lines, areas, and symbols. Kramer (2000) described the GeoMapper extension of PenMap as the "most complete, field tested, and proven Windows-based software for creating geological maps in the field." Use of GeoMapper has been described for general geology (Brimhall, 1998), field classes for undergraduate and graduate students in science, engineering and planning (Brimhall, 1999), and in professional mining and exploration geology (Brimhall and Vanegas, 2000). In collaboration with the USGS Water Resources Division, GeoMapper has been linked with hyperspectral infrared spectrometers for identifying and mapping minerals exposed on abandoned mine dumps as part of site characterization for screening and remediation including mapping from a helicopter platform (Montero Sanchez and Brimhall, 1998, in press; Montero Sanchez and others, 1999, in press).

In the past, many challenges retarded advances in field data capture. Besides the lack of effective software, until recently the hardware systems proved less than desirable. Only in the last year have there been available truly daylight-readable ruggedized color-pen-tablet PC computers (running Windows '98 and its successors). These PC's use lithium ion batteries with 3 to 4 hours of battery life. Effective electronic tools that plug into the pen tablets have existed for several years. Portable GPS units using Omnistar or the Coast Guard beacon for differential corrections in real time function without the need for a local base station. These units offer 1-meter accuracy in Northing and Easting and 2.5 meters in elevation unless satellite reception is obstructed by steep topography or tall buildings. Laser range finders with built in digital tilt meter and compass function up to 300 meters from the user. GeoMapper provides ready access through PenMap to all these digital tools and, in addition, creates a visual user interface that supports efficient geological mapping, including addition of color infill for formations using snap nodes and shared databases along contacts. Creation of professional-grade colored geological maps is straightforward with GeoMapper. Visual Basic programs in GeoMapper provide database management and conversion for compatibility with ERSI ArcMap for seamless map production.

ADVANCES IN GEOMAPPER AND CREATION OF GEOLOGGER

From our perspective as scientific mapping software developers for industry and academia, the main challenges remaining in field-mapping software are in two areas: (1) providing ready access in digital form to supporting geological data such as local geology and stratigraphy for making legends while remaining as technically faithful as possible to established standards in nomenclature, and (2) the lack of effective digital drill-hole-logging software. Although paperless mapping has been possible with GeoMapper for three years, without a drill-hole-logging companion for GeoMapper, mining and exploration geologists were forced to rely on paper-log-sheet data entry, making complete transition to digital methods impossible. GeoMapper and GeoLogger can serve as a true cross-over technology from paper to complete digital mapping if implemented.

Enhancement of GeoMapper's Legend Maker: Direct Browser Access to the AAPG COSUNA Charts for the Entire U.S. in Digital Form on a Single CD-ROM

In any digital mapping project, a mapping legend is necessary to define mapping units that can be recognized and followed in the field by correlation. The legend must include all the discernible map units to be encountered in the region and also provide sufficient flexibility to be able to add new units if they are discovered. For this purpose, a wealth of carefully synthesized information exists in the American Association of Petroleum Geologists (AAPG) Correlation of Stratigraphic Units (COSUNA) charts (Childs and Salvador, 1985), which were published in digital form in 2002. The term "correlation" is used because the charts afford an opportunity to visually compare the stratigraphy from one column to another over an entire region, and recognize facies changes and lithotectonic classification (Muehlberger, 1996). The charts span the entire U.S. in 20 geographic regions including Alaska. These charts provide geological information from which effective mapping legends can be readily constructed for essentially anywhere in the U.S. All charts come in Adobe Acrobat PDF format on a single CD-ROM and show several thousand stratigraphic columns positioned in an index plan map within the 20 regions (Figure 1).

Within each regional AAPG chart, individual stratigraphic columns are presented in a plan map index (Figure 2) showing their geographic position and proximity to other stratigraphic sections in the same region, e.g., "GB" for Great Basin. The sections, including one example in Figure 3, are based both on drill-hole information and on surface geology. COSUNA charts show all stratigraphic sections in a region by their column number and geographic name. Formations are positioned vertically down through time with colored codes showing their dominant lithology. The COSUNA charts show the stratigraphic columns correlated with formal systems, series/stages, chronostratigraphic units, magnetic anomaly, planktonic foram zone, mammalian stages, molluscan stages, benthic foram zone, and absolute age in millions of years. With

Figure 1. General regions for Correlation of Stratigraphic Units Charts (COSUNA) of the American Association of Petroleum Geologists (AAPG) (Childs and Salvador, 1985). Abbreviations: NCA, Northern California; CCA, Central California; GB, Great Basin; etc.

this information presented graphically, construction of mapping legends is possible for a broad array of disciplines including general geology, paleontology, environmental geology, engineering geology, hydrology, petrology, oil and gas, coal, industrial minerals. and metals mining and exploration. The CD-ROM is available on the Web from the AAPG Bookstore at <http://www.aapg.org/datasystems/LibraryPricing.html>.

GeoMapper Project Manager and Legend Maker

Once defined and programmed for a specific field area, the button array expressing the stratigraphic section is the geologist's link with time and process and lays out the units to be recognized and mapped. To create a legend in GeoMapper we use Legend Maker, which is implemented when one clicks on Personalize Legend in the Project Manager (Figure 4). To personalize the legend, a user simply needs to use point-and-click skills to effect changes in the design of the formation and lithology buttons, select their area fill patterns and/or colors, and type in their descriptive names. Typically this process takes less than an hour even for a complicated legend.

Once the geological legend has been made, one clicks on Start Mapping on the Project Manager window (Figure 4). From this point on, GeoMapper's visual user interface shows arrays of buttons arranged so as to provide a logical, self-explanatory set of features used in mapping. In GeoMapper, the databases, layers files, and symbols are preprogrammed and linked so a user need not be concerned with database construction nor management. These

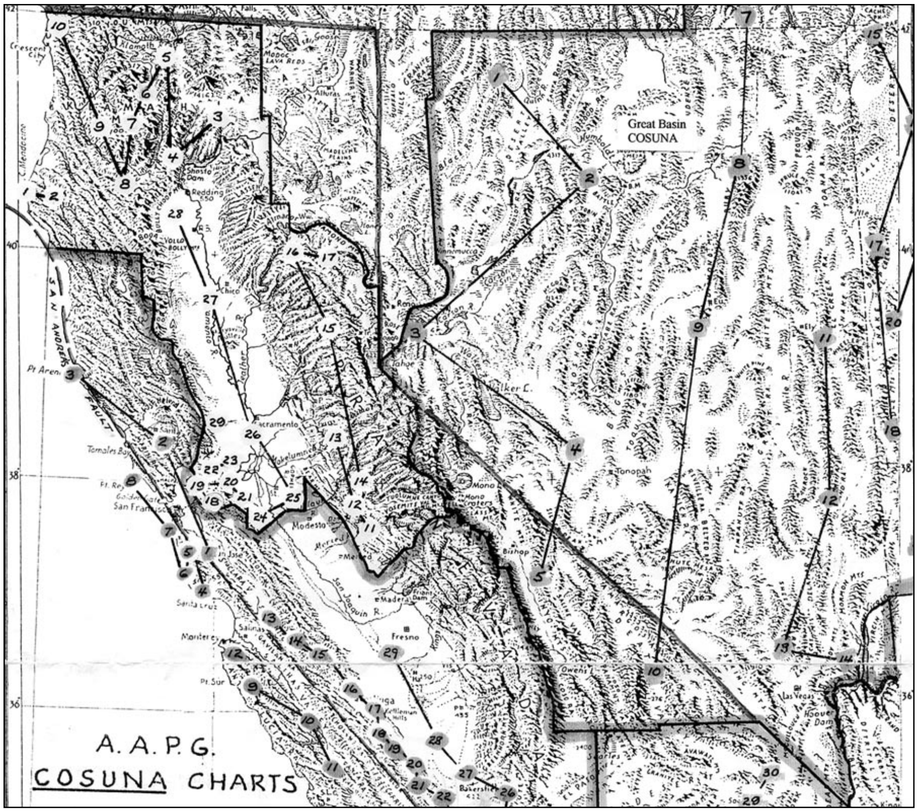


Figure 2. California and Nevada regions for AAPG COSUNA Charts. Numbers represent the position of specific stratigraphic columns.

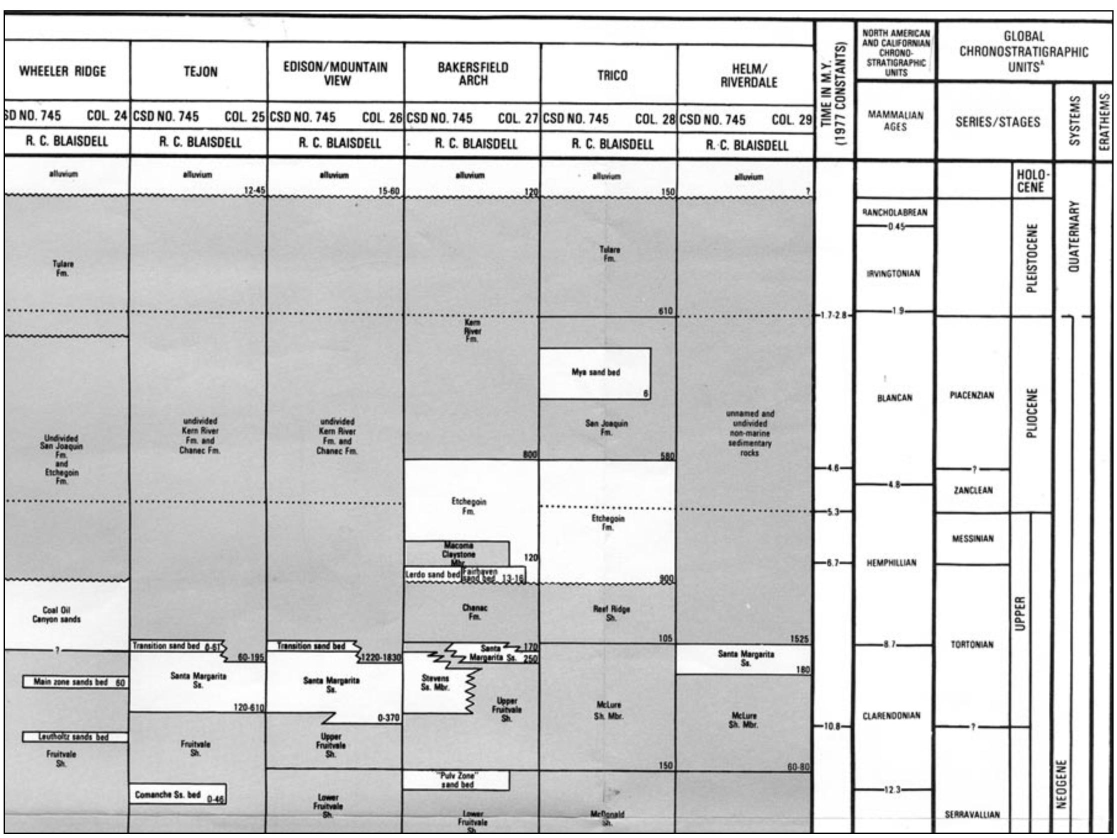


Figure 3. Example of part of a COSUNA chart for southern California.

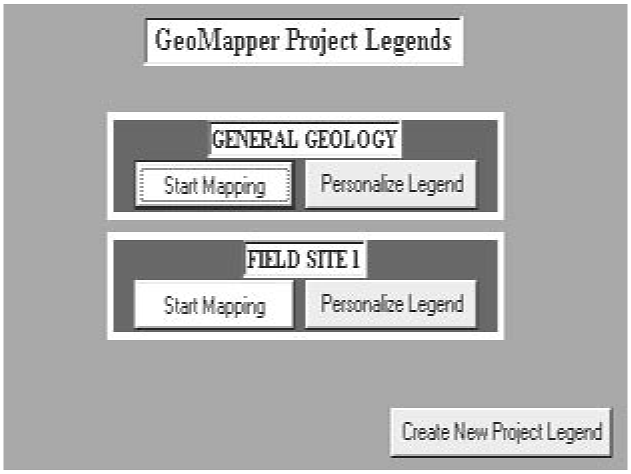


Figure 4. Project Manager in GeoMapper, showing button access to Legend Maker and other features.

tools are described in detail and illustrated in Brimhall and Vanegas (2001).

GeoMapper Tools Toolbar

This tool bar (Figure 5) includes specialized buttons for editing all types of graphics, sampling of rocks, soils, water, and infrared spectrometer sites. Export of all the digital mapped data as Shape files is done using the SHP button. The symbology of the graphics including areal patterns and infill colors is done with the button to the right of the SHP button.

Seamless Export from GeoMapper Into ESRI ArcMap: Open Architecture

From the Tools Toolbar, the finished database is exported so that ESRI ArcMap can be used for map production. We have made all the symbols such as strike and dip according to the USGS standard published in Open-File Report 95-525 (U.S. Geological Survey, 1995). Once in ArcMap, the Shape files are imported and our symbology is added to the Shape files, including points, lines, and areas (Figure 6). A map made in the field using GeoMapper is then imported into ArcMap with exactly the same line styles, colors, and patterns. We accomplish this seamless export to ESRI ArcMap and to other GIS programs as well (MicroStation, AutoCad) by using Microsoft Access as a personal, portable, relational database with a format amenable to ODBC (open database connectivity) clients. This open architecture provides many opportunities for using a variety of 2-D and 3-D GIS programs.

Final Map Production in ESRI ArcMap

With the mapping tools available in GeoMapper, complete geological maps made in the field can include color infill patterns for formations and patterns for outcrops if desired. Because we have designed GeoMapper from the standpoint of a geologist in the field, all the functionality desired by geologists is included. A key feature of GeoMapper is the ability to define the perimeters of colored polygons automatically by using so called “snap” nodes positioned along defining contacts so as to produce a professional-quality map. The level of certainty of contacts can be shown from solid black lines to dashed lines. Contacts offset by faults are easily shown as well. By showing outcrops in addition to formations, the primacy of the scientific record is preserved so that users of the map can find the same outcrops that the contacts are based upon. When a map is printed at a less detailed scale the outcrops disappear. Whenever desired, a working map can be printed directly from GeoMapper. Using the Shapefile export feature, the database can be exported and the compiled final map can be produced in ArcMap with the considerable advantages available in the ESRI environment, including compatibility with other ESRI products.

APPLICATIONS OF GEOMAPPER

To illustrate uses of GeoMapper we outline here several applications we have recently made. These applications include the fields of geology, astronomy, environmental geology, and tactile reality models for the disabled.

Mapping for Astronomical Purposes

The new Search for Extraterrestrial Intelligence (SETI) observatory will be located on the grounds of the Hat Creek Radio Telescope Observatory, in a remote area of northeastern California where radio background noise and artificial lighting are minimal. Available land for SETI’s 350 new radio telescopes is quite limited at the

Figure 5. The Tools Toolbar contains special features such as those necessary for mapping and exporting completed maps and numerical databases into ESRI ArcMap for seamless map production.

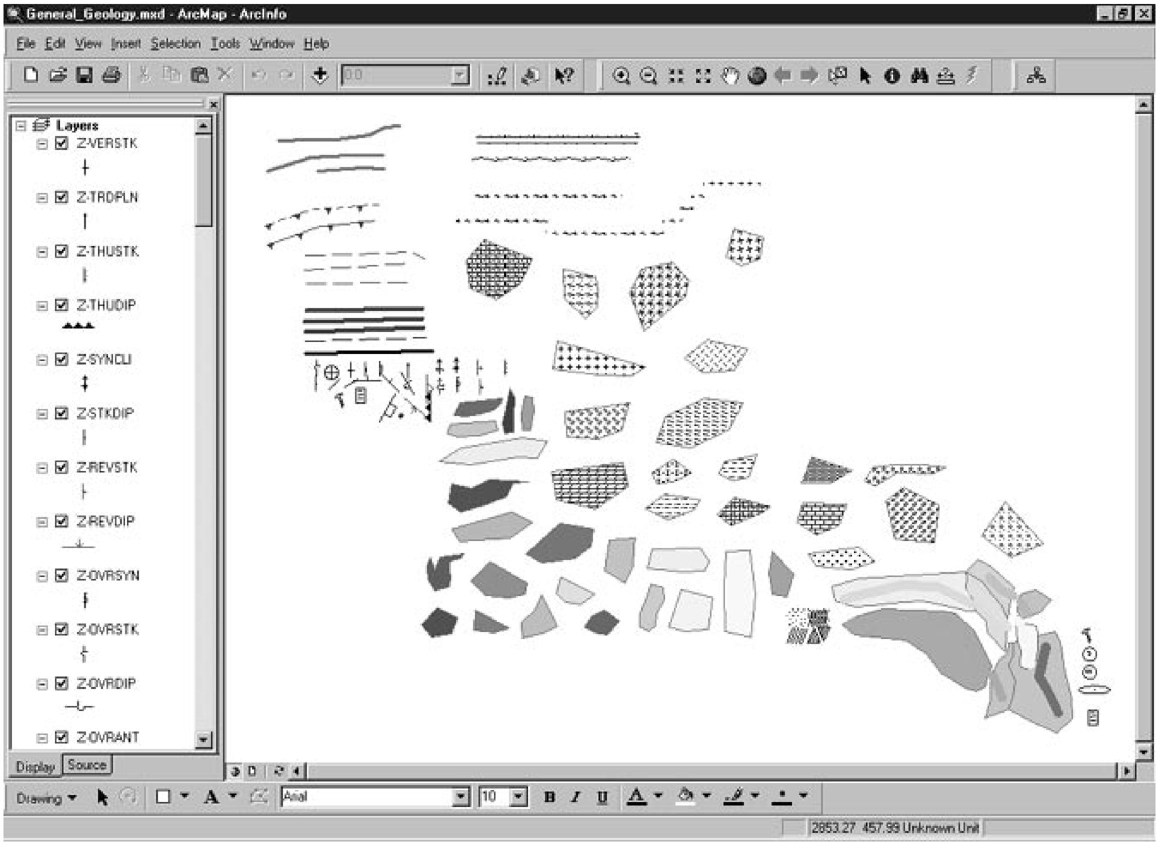


Figure 6. ESRI ArcMap rendering of lines, styles, symbols, formation color infills, and outcrop lithology patterns made in GeoMapper and exported directly into ArcMap.

Hat Creek site, necessitating a very closely spaced array, and hence, a very detailed topographic map. GeoMapper was used to make this map (Figure 7). The area is near the front of young basalt flows on valley-flow sediments. Land which is flat and free of lava tubes is required for the foundations of the new radio telescopes. Although we typically use GeoMapper in conjunction with a Trimble Ag-132 GPS that receives Omnistar differential corrections in real time with a point location of about 1 meter in plan and 2.5 meters in elevation, we needed higher accuracy for this project. By using a local base station (Figure 8) from which we computed diurnal drift correction, we corrected the x,y,z data for the four synchronized roving mapping systems used over a four-day period. Thus, by correcting our data in post-processing, we mapped at an accuracy of several decimeters in plan and 1 meter in elevation instead of the usual meter-level work. The data points were exported from GeoMapper and contoured in Surfer and output in plan view (Figure 9). Broad divisions of the available space for foundation sites for the new radio telescopes are based upon this map.

Environmental Remediation

Environmental remediation of abandoned mines requires a preliminary investigation of thousands of sites, and characterization and selection of a few sites most deserving of remediation. Hence, a sensitive and efficient screening method is paramount to success. We have integrated our GeoMapper system using a GPS and laser with a hyperspectral spectrometer to rapidly identify and map indicator minerals characteristic of pyrite oxidation and sulfuric acid generation (Montero Sanchez and Brimhall, 1998, in press; Montero Sanchez and others, 1999; in press) (Figure 10). The software/hardware combination works efficiently on the ground and from a helicopter using a reflectorless laser range finder (Figure 11). Oxidation indicator mineral maps show the regions of most intense acidification and show which mines and which areas warrant further investigation and water sampling.

Tactile Virtual-Reality Raised Models for the Blind and Disabled

GeoMapper has proven very effective in creating topographic maps and 3-D models. We have created a raised-relief scale model of the University of California, Berkeley campus 3 by 6 feet in size for use by blind and disabled students to learn their way around the campus (Figure 12). Managing the walking paths, avoiding traffic,

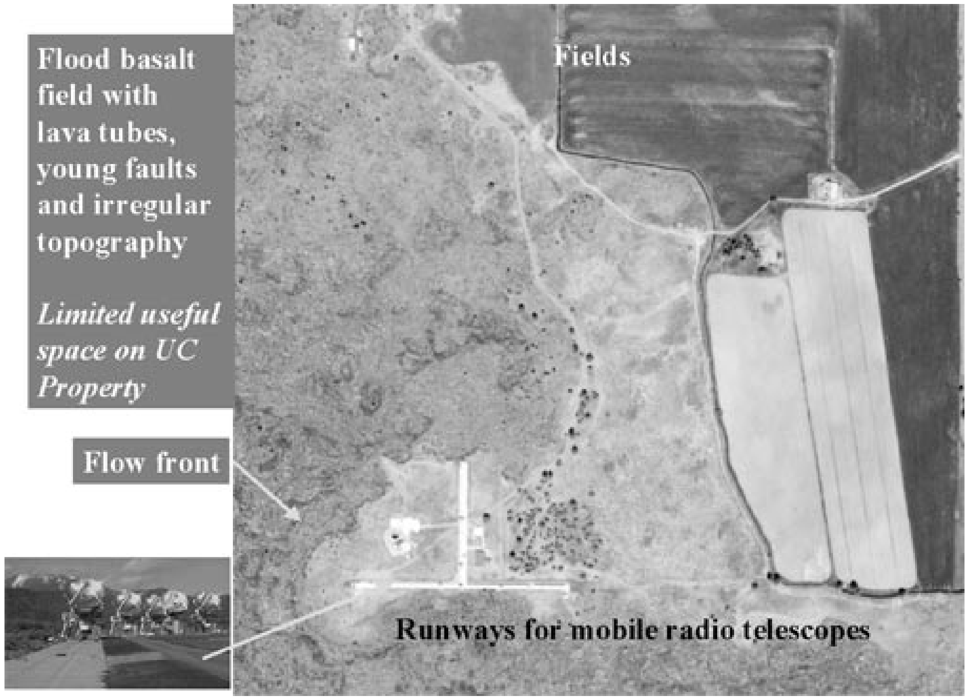


Figure 7. Aerial photograph of the Hat Creek Radio Telescope Observatory in northeastern California. Young basalt flows are present on the left side of the photo.

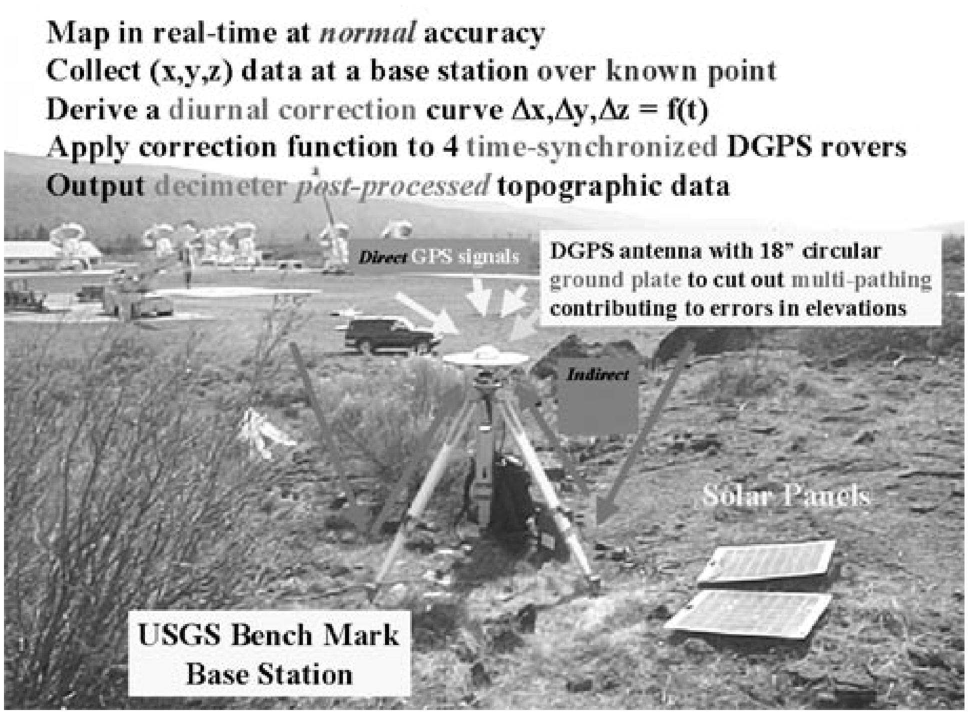


Figure 8. GPS base station set-up showing solar panels for power. Note radio telescopes in the background. Reference point is a permanent USGS survey monument.

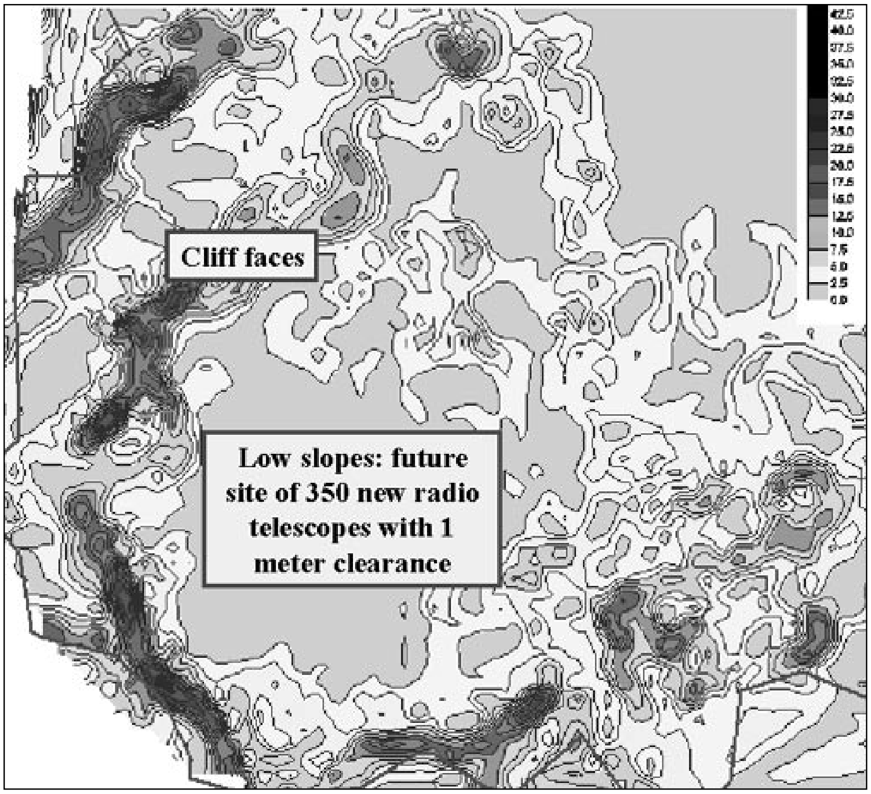


Figure 9. Hillslope angle in plan view. Dark areas are cliff faces of the flow terminus. Contour interval is 2.5 slope units.

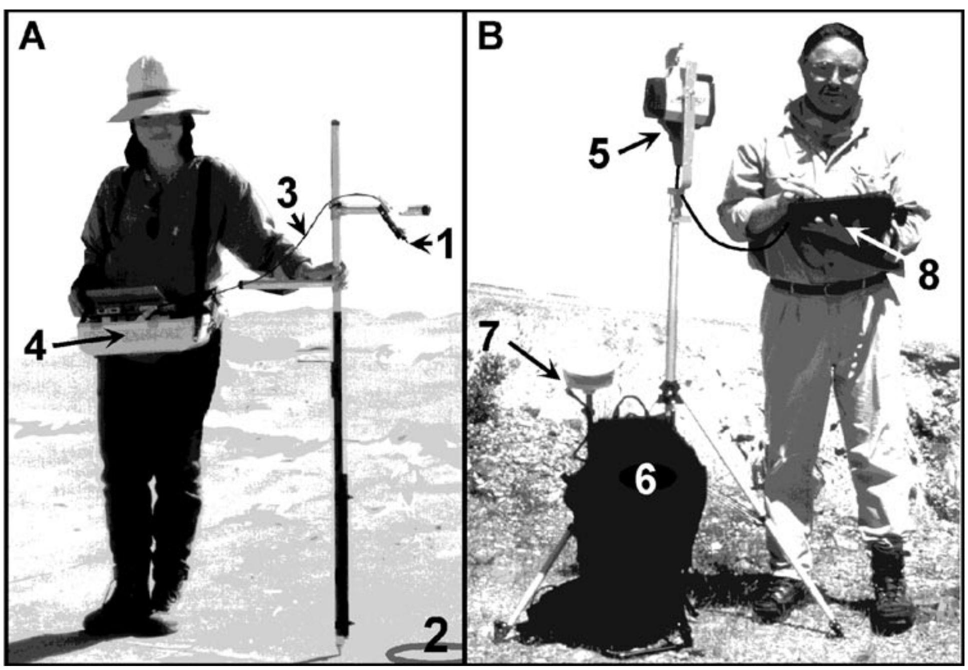


Figure 10. Irene Montero Sanchez (**A**) and George Brimhall (**B**) demonstrate use of digital mapping system (from Montero Sanchez and Brimhall, in press). Numbers identify components in the mapping system. **1**, probe holder for spectrometer's fiber optic probe; staff on which holder rests maintains the probe away from the operator at a constant height and angle above the ground. **2**, target on the ground. **3**, fiber optic cable transmitting light from the cable opening to the spectrometer. **4**, portable, battery-operated spectrometer. **5**, laser range finder with internal digital inclinometer and magnetic compass. **6**, portable differential GPS receiver (inside backpack). **7**, DGPS antenna. **8**, pen-tablet portable PC computer.

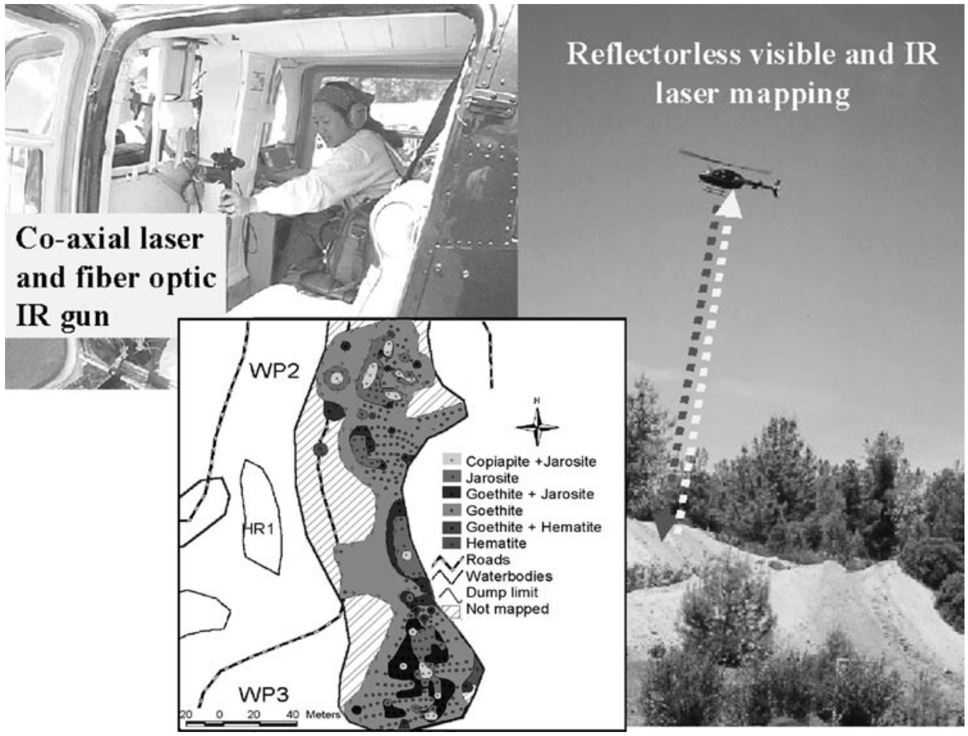


Figure 11. Helicopter-based digital mapping system combining hyperspectral visible and infrared (IR) spectrometer and pen-tablet-based control system running GeoMapper. GPS unit is in the tail section and automatically updates the UTM coordinates of the reflectorless laser. Map shows pixels classified by the dominant mineral identified by the spectrometer.

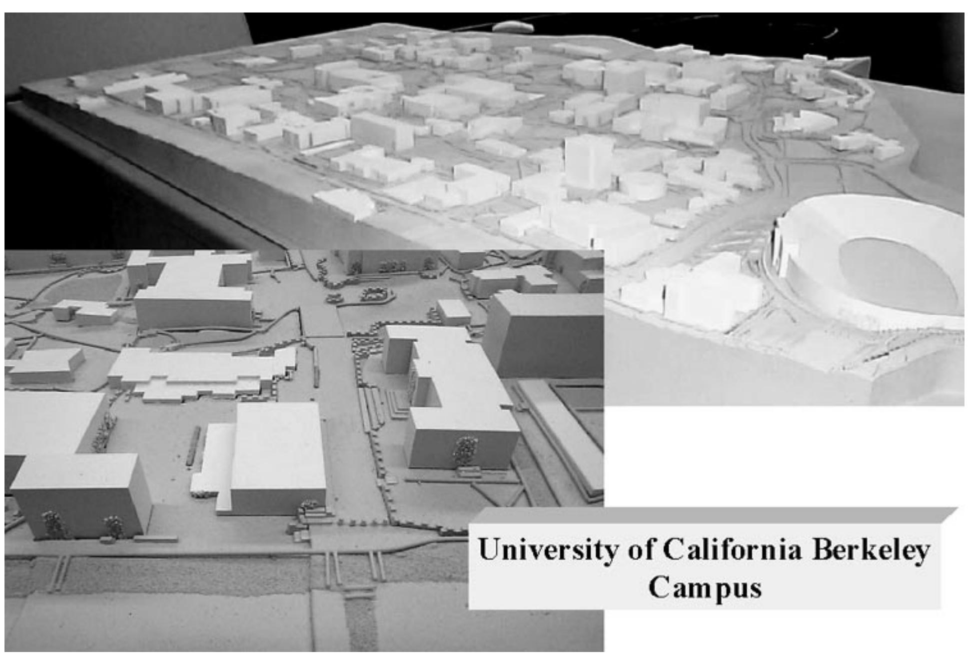


Figure 12. Raised-relief tactile-reality model of the University of California, Berkeley for blind students. Heights of buildings were determined by laser range finder from the ground. Model is made of polyurethane foam machined from a solid block.

and finding classrooms in buildings is aided by this tactile model. All entry points to buildings are shown, as are paths, roads, and obstacles. Paths and roads have textured surfaces classified by usage: roads where car traffic is expected, and paths where only pedestrians walk.

DRILL-HOLE LOGGING: FULL 3-D MAPPING AND MODELING

While the mapping described above is three dimensional in the sense that all nodes entered have 3-D coordinates which can be used to portray the Earth’s topography and the model buildings, digital mapping software is incomplete without drill-hole-logging capabilities. Our intention in implementing digital core logging as a companion for GeoMapper is threefold: (1) provide a pen-tablet-based logging system that captures data directly in real time, (2) provide useful mineralogical tools with predictive metallurgical capabilities, and (3) afford more time for geologists to engage in productive interpretation by eliminating the widely used paper log forms that, later, require digitization.

Geological Logging

Much like mapping, which in a stereological sense reduces four-dimensional space-time to a two-dimensional plane, geological logging captures essentially a one-dimensional sample of the Earth provided by drilling. Rather than carrying a map board, colored pencils, and log sheets, we implement digital mapping through a pen-tablet computer, a touch stylus, digital log sheets, and a visual user interface that provides all logging functions by touching the pen stylus to an array of buttons. Furthermore, the buttons are shown in the general sequence of their use so that scientific logic guides the selection of mapping tools. GeoLogger starts with a form for entry of drill-hole information (Figure 13). From there, a log sheet appears in digital form (Figure 14) showing geotechnical rock quality designation (RQD), structure, lithology, alteration, and sulfide mineralization. The format of the log sheet is designed in accordance with the needs of the users. The GeoLogger interface appears much like that of GeoMapper so that it is easy to learn and fast in execution. GeoLogger serves a variety of logging applications: base metals, precious metals, industrial minerals, and oil and gas.

The button color-coding uses the “traffic light” method with green, yellow and red phases of activity. Green buttons are the most commonly used buttons in geological mapping. Yellow buttons are procedures that are used only rarely (for example if you need to erase or undo the last work). Red buttons are procedures that are

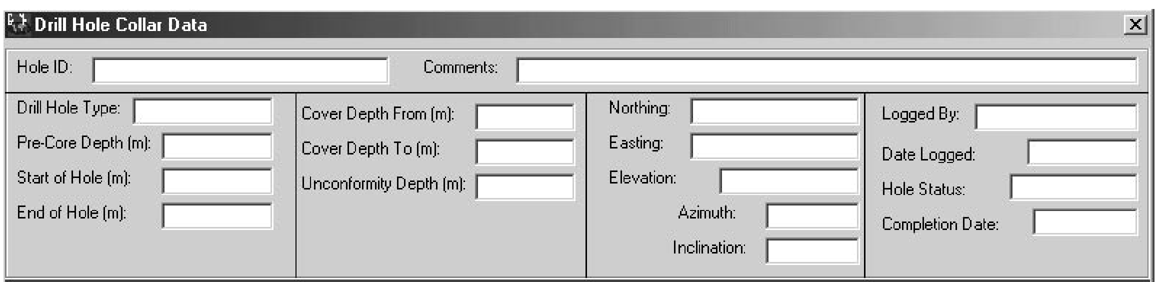


Figure 13. GeoLogger data-entry form for a drill hole.

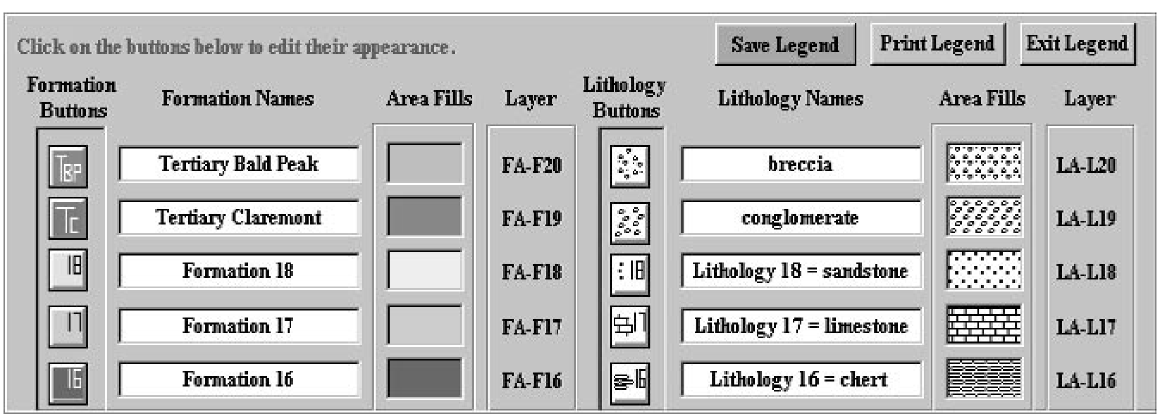


Figure 14. Digital log sheet. Point-and-click spatial resolution is 1 cm (features may be located to hole depths with an accuracy of 1 cm), stored, and retrieved from the digital log sheet and database.

essential to do before you stop mapping (for example, saving your files or exporting critical files). Other colors relate to special-use functions, such as light-blue buttons for selecting various-scale log sheets.

GeoLogger has an organized array of buttons representing different rock types, structures, mineralization, alteration, sampling, and geotechnical features, each set in a tool bar. To log a feature, one has only to click on the appropriate button and then touch the screen of the pen-tablet computer at the appropriate hole depth. If the feature occurs over an interval of core, then the top and bottom depths on the screen are touched. This point-and-click action automatically selects the right layer, database, and graphics and associates this information with the down-hole depth. The computer, stylus, visual user interface, and base log sheet provide an integrated system for the geologist. Digital images of core samples provide ready access to key features noticed while logging (Figure 15). Notes are entered using a stylus and digital keyboard. The final printout of the completed digital log appears identical to paper logs. Chemical data such as metal assays can be imported into the log database and displayed for purposes of correlating with observed geological features (Figure 16).

Pedagogical Advantages of a Preprogrammed Legend

The design of the mapping legend itself encourages attainment of high professional standards in a minimum of training time. Because the logging is done with a set of computer tools, standardization of the features is automatic thereby reducing the time necessary for training new geologists. Critical data-entry fields such as sample number are compulsory so that a logger cannot proceed without completing the data entry. The visual user interface provides a simple organization for lithologies, structures, mineralization, alteration, sampling intervals, and geotechnical information such as RQD.

CONCLUSIONS

Our experience in digital mapping has opened many new avenues of thought for us. It has convinced us that the central role of this powerful new technology is in creating and communicating in human terms a geospatial reality that conveys meaning and order about the physical world to many quite different audiences. Potential user communities abound, and geology remains the bedrock anchor to the solid Earth on which human endeavor is linked. The "map that changed the world" published in the early 19th century by William Smith gave birth to the field of geology and a host of scientific and economic benefits. Geologists today retain a practical acquaintance of space and time but now possess mapping tools so enhanced by the digital revolution that limitation seems unimaginable. Mapping opportunities abound to help guide the stewardship of the Earth and to improve the human condition.

REFERENCES

Brimhall, G., 1998, Direct digital field mapping using pen-based PC computers supported by differential global positioning systems and laser range finders: Geological Society of America Abstracts with Programs, v. 30, no. 7, p. A-256.

Brimhall, G., 1999, Evaluation of digital mapping classes in introductory and advanced field classes at UC Berkeley: Geological Society of America Abstracts with Programs, v. 32, no. 7, p. A-191.

Brimhall G.H., and Vanegas, A., 2000, Digital mapping of geology and ore deposits with GeoMapper: Geological Society of America Abstracts with Programs, v. 32, no. 7, p. A-514.

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 interface, in Soller, D.R., ed., Digital Mapping Techniques '01—Workshop Proceedings: U.S. Geological Survey Open-File Report 01-223, p. 103-114; <http://pubs.usgs.gov/of/2001/of01-223/brimhall.html>.

Childs, O.E., and Salvador, A., 1985, Correlation of Stratigraphic Units of North America (COSUNA): American Associa-

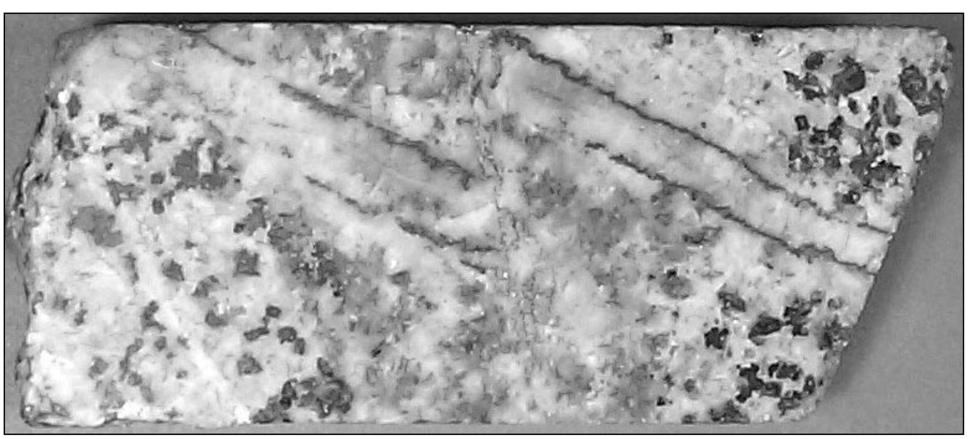


Figure 15. Digital image of diamond drill core.

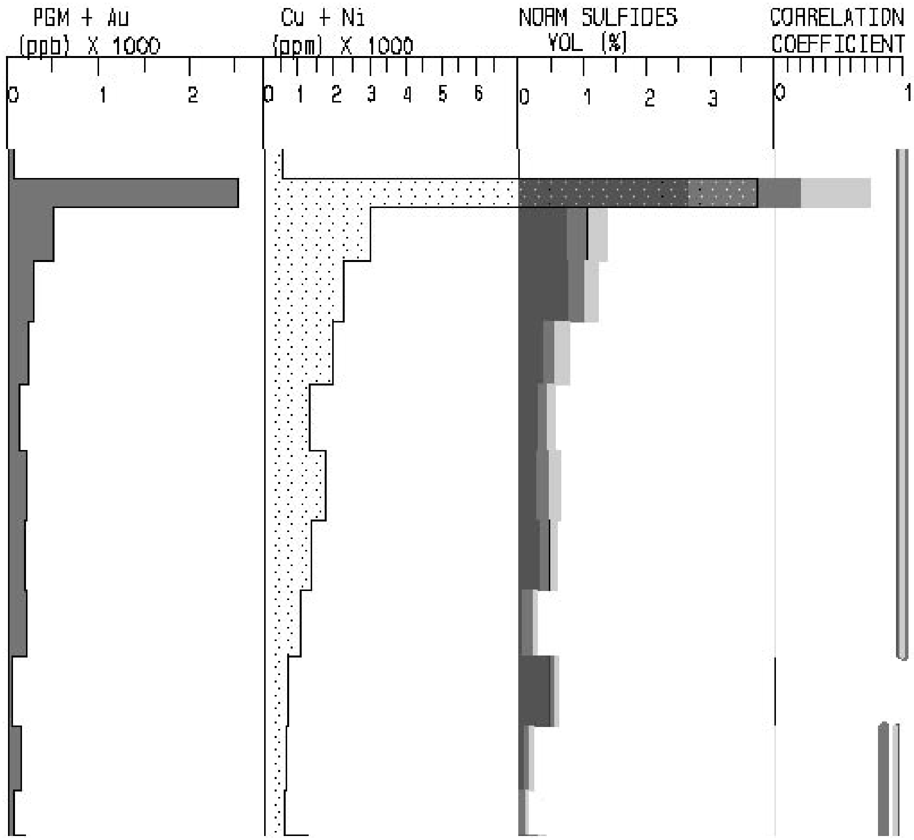


Figure 16. Example of GeoLogger log showing imported platinum group metals (PGM), copper and nickel, and normative sulfides (pyrite, pyrrhotite, pentlandite, and chalcopyrite), computed using metal and sulfur assays, and correlation coefficients of PGM with each normative sulfide.

tion of Petrololeum Geologists Bulletin, v. 69, p. 173-189.

Jackson, I., and Asch, K., 2002, The status of digital geological mapping in Europe: the results of a census of the digital mapping coverage, approaches and standards of 29 European geological survey organizations in the year 2000: Computers in Geosciences, v. 28, p. 783-788.

Kramer, J.H., 2000, Digital mapping systems for field data collection, *in* Soller, D.R., ed., Digital Mapping Techniques '00—Workshop Proceedings: U.S. Geological Survey Open-File Report 00-325, p. 13-19; < http://pubs.usgs.gov/openfile/of00-325/kramer.html>.

Montero-Sanchez, I.C., and Brimhall, G.H., 1998, Novel application of digitally integrated mapping systems for the mineralogical characterization of abandoned mines: Geological Society of America Abstracts with Programs, v. 30, no. 7, p. A-358.

Montero Sanchez, I.C., and Brimhall, G. H, in press, FST-SpecID: fast spectral analysis program for the identification of dominant secondary iron mineralogy using integrated digital mapping of abandoned mines: Remote Sensing of the Environment.

Montero-Sanchez, I.C., Brimhall, G.H., and Alpers, C.N., 1999, Use of UV/VIS/IS spectroscopy to characterize mine waste dumps in the Penn Mine, Calaveras County, California: Geological Society of America Abstracts with Programs, v. 31, no. 6, p. A-91.

Montero Sanchez, I. C., Brimhall, G.H., Alpers, C.N., and Swayze, G.A., in press, Characterization of waste rock associated with acid drainage at the Penn Mine, California by ground-based visible to short-wave infrared reflectance spectroscopy assisted by digital mapping: Chemical Geology.

Muehlberger, W. R., compiler, 1996, Tectonic map of North America: American Association of Petroleum Geologists, scale 1:5,000,000.

Soller, D. R., and Berg, T. M., 2001, The National Geological Map Database: a progress report, *in* Soller, D.R., ed., Digital Mapping Techniques '01—Workshop Proceedings: U.S. Geological Survey Open-File Report 01-223, p. 51-58; < http://pubs.usgs.gov/of/2001/of01-223/soller1.html>.

Soller, D. R., Wahl, R., Weisenfluh, J., Brodaric, R., Hastings, J., Laudati, R., and Fredericks, R., 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; <http://pubs.usgs.gov/of/2001/of01-223/soller2.html>.

U.S. Geological Survey, 1995, Draft cartographic and digital standard for geologic map information: U.S. Geological Survey Open-File Report 95-525.

Winchester, Simon, 2001, The map that changed the world: New York, Harper Collins, 329 p.

Overcoming Institutional Barriers to GIS Coordination: Building a State GIS Council—The Alabama Experience

by Berry H. (Nick) Tew, Jr.

Geological Survey of Alabama
Box 869999
Tuscaloosa, AL 35486-6999
Telephone: (205) 349-2852
Fax: (205) 349-2861
e-mail: ntew@gsa.state.al.us

Some of the primary barriers to effective development of comprehensive, coordinated state government programs to acquire, maintain, and distribute digital geographic information are institutional in nature. These barriers are often related to a general lack of communication, coordination, and, in some cases, cooperation, among state governmental agencies and others that have a stake in the development of a current, accurate geospatial-data infrastructure. Thus, many of the impediments to realization of such an infrastructure often result directly from (1) the absence of an officially sanctioned entity charged with providing a forum for discussion, (2) a mechanism for addressing pertinent issues and concerns, and (3) a process for collective decision-making. Prior to March of 2002, no such entity existed in the State of Alabama.

Over the past decade or so, governmental agencies at all levels, educational institutions, and private industry in Alabama invested heavily in technology for digital mapping and processing of geographic information. As this occurred, there was a progressive increase in the awareness of the potential that this technology offered for long-term economic and societal benefits to the State and its citizens. It was realized that a significant increase in the effectiveness and efficiency of information acquisition, maintenance, access, and delivery could be achieved through a coordinated, statewide GIS effort. It was noted that GIS technology could be used as a strategic decision-making tool in such areas as economic planning and development; water, agricultural, energy, cultural, land, and mineral resources; environmental management; forestry; geology; public health; local government services; land-use planning; public safety; social services; transportation; utilities; waste management; and wildlife conservation and management.

In 1999, a small group of primarily managerial-level Alabama GIS practitioners recognized the benefits to be gained for coordination, cooperation, and sharing in the acquisition, maintenance, and dissemination of geospatial data for the State and began holding informal meetings to begin the process of establishment of a state-level GIS coordinating body. This ad hoc committee, over the course of several months, worked to identify important issues and concerns, gain an understanding of the GIS community in Alabama and the State's data and technology needs, conduct research on GIS activities and organizations in other states, and develop a strategy to create a GIS coordinating entity in Alabama. The work of this ad hoc committee resulted in a draft legislative bill to establish a GIS council for the State, which was introduced in the 2001 Regular Session of the Alabama legislature. Although the bill was favorably received and had no substantive opposition, it failed to gain final passage.

Subsequent to this initial failure to pass the GIS council bill in the legislature, members of the ad hoc committee initiated meetings with Governor Don Siegelman's staff to discuss the issues related to geospatial data and GIS technology and the potential benefits to be gained from a coordinated, cooperative approach to development of Alabama Spatial Data Infrastructure. These discussions led to Executive Order No. 68, which established the Alabama Geographic Information Council (AGIC) and provided the mandate for the council. AGIC included members representing various agencies of the State of Alabama, regional and local government, the education community at all levels, State boards of professional licensure, and the private sector.

The duties and mandates of AGIC as established in Executive Order No. 68 included the following:

- Recommend goals, objectives, and a strategic management plan to guide the development and implementation of GIS technology for the best value and benefit of the citizens of Alabama.
- Assess all current geographic information available for the State of Alabama and make recommendations to reduce inefficiency and redundancy in geographic information collection.
- Recommend policies and strategies which emphasize cooperation and coordination among state agencies, federal agencies, environmental agencies,

regional planning agencies, municipalities, counties, academic institutions, nonprofit organizations, utilities, private companies, individuals, and other states in order to maximize the cost efficiency.

- Recommend a strategy for funding and development of an accurate, current digital base map for the state that includes commonly needed geographic information themes, including the seven framework data layers as defined by the Federal Geographic Data Committee—roads, streams, orthorectified digital aerial photography, elevation, political boundaries, cadastre (parcel ownership), and geodetic control.
- Meet at least monthly until submission of a final report on or about September 21, 2002.
- Prepare a final report, to include the following:
  a) a needs assessment by each state agency represented on AGIC regarding GIS technology,
  b) recommendations with respect to the future organizational structure of the AGIC, and
  c) recommendations for implementing a comprehensive GIS strategy for Alabama.

Major activities conducted by the AGIC membership included seven meetings of the full council, seven additional meetings by a designated Technical Subcommittee, compilation of GIS needs and uses reports for state agencies, a GIS data survey of the identified stakeholder community in Alabama, and a statewide GIS symposium entitled "Governor's 1st Annual GIS Symposium—Mapping Alabama's Future." All data collected as part of the AGIC process indicated a need and desire for a permanent executive-level policy organization such as AGIC, as well as a state-government-sanctioned operational entity ("Office of GIS") to conduct GIS activities in Alabama. It was concluded that establishment of these entities on a permanent basis can result in significant benefit to the State of Alabama. These needs were forwarded as recommendations to the Governor in the final report of AGIC and are presently under review and consideration.

Vendor Presentations and Contact Information

The Digital Mapping Techniques ‘02 workshop was attended by technical experts from selected software and hardware companies. These individuals provided technical troubleshooting and general information to the geological survey workshop attendees, and the workshop organizers offer sincere thanks for their significant contributions to the meeting. The DMT workshop series is designed as a collegial event, where information is freely shared in recognition of a common set of goals. Our colleagues in the vendor community certainly contributed to the workshop’s success.

Mike Price of Environmental Systems Research Institute, Inc., (ESRI) provided technical guidance and support for the half-day field demonstration of portable data-capture systems, which was the first event of the meeting. Mike also gave an oral presentation that described the various ESRI software products. In this presentation, he noted that ESRI has engaged the geologic community in a formal process to develop a geologic data model compatible with their Geodatabase structure. In early 2002 a planning meeting was held at ESRI headquarters, attended by representatives from the British, Canadian, Dutch, and United States geological surveys. Progress on this initiative will be posted at <http://arconline.esri.com/arconline/datamodels.cfm>. For further information, please contact Andrew Zolnai (ESRI Petroleum Manager, <azolnai@esri.com>) or Steve Grise (sgrise@esri.com). ESRI also provided operating funds for the meeting, and we sincerely thank ESRI and Mike for their generosity and for their interest in this meeting. For information regarding ESRI products and/or the geologic data model development, please contact:

Andrew Zolnai, Petroleum and Pipeline Manager
Environmental Systems Research Institute, Inc.
380 New York St.,
Redlands, CA 92373-8100
Telephone: (909) 793-2853
e-mail: azolnai@esri.com
Corporate Web site: <http://www.esri.com>

Dean Derhak of Onyx Graphics gave a presentation entitled “Using Onyx PosterShop products in GIS/Mapping—A Brief Introduction.” Onyx Graphics offers RIP software for large-format printing devices including Hewlett-Packard, Epson, Encad, and Canon, and they support a wide range of image file formats. Onyx RIP Products allow you to print directly from your Mac or PC mapping software, drive multiple printers from the same computer system, process and print on the fly, scan your maps directly into the RIP software, scale maps to almost any size, use light inks to preserve map detail and smoothness in light areas, process and print files of almost any size, and quickly open large image files and do real-time color adjustments on-screen. Contact information:

Dean Derhak
Onyx Graphics
6915 Hightech Dr.
Midvale, UT 84047
Telephone: (801) 984-5318
e-mail: Dean.Derhak@onyxgfx.com
Corporate Web site: <http://www.onyxgfx.com>

Todd Packebush of LizardTech, Inc. gave a presentation entitled “Create, Distribute, Archive and View Complex Content Efficiently, with LizardTech Image Compression Software.” LizardTech makes documents, imagery, and photographs “network-ready” for rapid distribution and access with efficient storage, resulting in improvements in productivity, measurable cost savings and dramatic increases in the value of the information by making it more accessible and useful. Contact information:

Todd H. Packebush
SW Territory Manager
LizardTech, Inc.
Telephone: (206) 652-5211
e-mail: todd@lizardtech.com

We also warmly thank Mike Price (ESRI), Chris Wayne (ESRI), and Gary Edmundo (MinGIS, Reno, NV) for their technical expertise and assistance in conducting the pre-meeting field demonstration of portable geologic mapping gear (i.e., ArcPad software on PDA’s, supported by GPS).

Alaska Division of Geological & Geophysical Surveys Geologic Database Development—Logical Model

By Larry Freeman,[1] Kathryn Engle,[2] and Carrie Browne[1]

[1]Alaska Division of Geological & Geophysical Surveys
794 University Avenue, Suite 200
Fairbanks, AK 99709
Telephone: (907) 451-5027
Fax: (907) 451-5050
e-mail: {Larry_Freeman, Carrie_Browne}@dnr.state.ak.us

[2]GeoNorth, LLC
1000 O'Malley Road, Suite 200
Anchorage, AK 99515
Telephone: (907) 677-1500
Fax: (907) 677-1502
e-mail: kengle@geonorth.com

INTRODUCTION

The Database Project at the Alaska Division of Geological & Geophysical Surveys (DGGS) was conceived in the late 1990s to stop the loss of critical geologic data that are used in the compilation of geologic maps and to modernize the way that DGGS delivers geologic data to the public. To assist in this task, in late 2000 the DGGS hired a contractor to help with the initial database design and implementation. A logical database model is now complete, and development of the physical model and subsequent building and testing of the initial database will occur this summer (2002).

PROJECT OBJECTIVE

Work began on the DGGS database project in late 2000. The objective was to create a new, comprehensive database to meet the data storage needs of the Survey, provide the basis for delivering geologic information to the public, and provide a stable data structure to interact with databases at other agencies such as the National Geologic Map Database, <http://ncgmp.usgs.gov/ngmdbproject/home.html>, Interagency Minerals Coordinating Group (IMCG), <http://imcg.wr.usgs.gov/index.html>, and the Alaska State Geo-Spatial Data Clearinghouse, <http://www.asgdc.state.ak.us/> (Freeman, 2001).

The database project faces such challenges as serving the needs of multiple users, functioning as part of multiple distributed data networks, anticipating future needs of the database and DGGS, and allowing for flexibility in the design of the database to adapt to changing technology. These challenges will be met using a fully normalized relational database model, developed in a spatial-data-capable, relational database management system. The logical model presented in our poster will be the basis for the development of the DGGS geologic database.

METHODOLOGY

During the past year, DGGS contracted with GeoNorth, LLC, of Anchorage, to design a relational database and implement it in Oracle 8i and ArcSDE on a UNIX platform. GeoNorth has developed the following models toward the completion of their contract with DGGS:

- *A business-process model*—a graphic model showing the flow of business and scientific work at DGGS from conception of a project through publication of the project's data, delivery of publications, and archiving of all project data. This model identifies and describes data entities and positions responsible for the work and the business (science) rules that define the relationships between the entities.
- A *conceptual database model*—a model that identifies the data entities, their attributes, and the relationships between entities.
- A *logical database model*—a comprehensive model of data entities, attributes including their logical data types, and the relationships between the entities. GeoNorth used Computer Associates

ERwin database modeling software to construct the model.

GeoNorth involved all members of the DGGS staff in an iterative and collaborative effort during the design period for the business-process and logical models. This is a critical factor in database design for two reasons: First, the geologists at DGGS are considered experts in their field and are most familiar with their data and how they will be using the data available to them; and second, involving the staff encouraged participation and interest in the database project. The logical model is based on data entities identified in the business-process and conceptual models. It has taken nearly seven months to design the models.

GeoNorth's next step is to convert the logical model to a physical model. Database (structured query language data definition language) code will be generated from the physical database model and a prototype database will be built. We expect that much of this design work and code generation will be performed using ERwin. An initial internal release of the database on the DGGS network is anticipated by summer 2002.

DGGS staff will test the database structure, using command-line interface (Oracle SQL+), ArcGIS 8, Microsoft Access, Map Info, and other third-party products. After the initial testing is complete and DGGS has accepted the database from GeoNorth, we will start building applications to facilitate frequently repeated data entry functions and queries and to allow for public access to the data via the DGGS Web site and through a publicly accessible workstation that will be available at DGGS's office.

THE LOGICAL MODEL

In the logical database model, <http://www.dggs.dnr.state.ak.us/Logmod_0205_web/dggs_logmod_020519.htm>, DGGS data are broken into multiple main categories connected by conceptual and logical relationships that represent the way that data are collected, edited, and analyzed by the specific work groups. Although the logical model appears complex, encompassing more than 200 entities (tables), many of these entities are subtypes, data validation lists, and descendants of six primary entities (Figure 1). The relationships between the primary entities are a distillation of the business rules identified in the business-process model and serve to link the main data categories, which are described below.

Field station data include descriptive, measured, and instrumental data that are collected in the field to support geologic mapping, resource investigations, and hazard evaluations. At DGGS, all field data are identified and located by a field station. Each field station has a point location and may have an associated geometry, such as a polygon that surrounds an outcrop or geothermal occurrence or the surface trace of a borehole.

Sample analyses consist of instrumental analyses and descriptions of geologic samples. Sample analysis data include summary and secondary analysis information as well as original laboratory data. A large number of the entities in the DGGS logical model are subtypes and descendants of a sample analysis. A sample analysis is identified by the sample number and the analysis batch (who, what, and where).

Publication information is recorded for DGGS publications and for external publications that are cited as sources in the DGGS data. Information for DGGS publications includes information about distribution, electronic files used to make the publications, and the archive location of those publications.

A geospatial data set is a set of spatially referenced, interrelated features. They are constrained by a geographic domain and exist as a separately addressable file that can be linked to the relational database that contains the map's descriptive information. All geologic map objects in the DGGS database will be derived from a geospatial data set. Some data sets are distributed by DGGS and others are distributed from other sources and used in the DGGS database. Regardless, metadata for the source data set can be queried from the database and displayed in a format compatible with FGDC-compliant metadata. Metadata for each data set will be required before it is loaded into the database.

A geologic map feature in the DGGS data model (Entity Data Spatial Classification, Figure 2) is defined by geometry and classification attributes. Geologic map features have topologic and geologic relationships that are internally consistent within a geospatial data set, but are related to geologic map features in other geospatial data sets only by their classification attributes. In other words, we are intending to preserve "geologic map edge faults" until they can be resolved. Classification attributes of geologic map features include feature type, composition, geologic age, or map unit (lithostratigraphic) name, proper name (e.g., Denali Fault), and derivative classification themes. Classification attributes are related to cartographic symbols to provide visualization of the features using multiple symbol sets. Although this model deviates from the North American Geologic Map Data Model (Johnson, and others, 1999), it gives us flexibility to create views of geologic map data in multiple configurations. The remaining entities in the database include thematic information such as mineral occurrence information and reference tables for data validation and indexing. The reference tables will contain the standard nomenclature, classification, and keywords that DGGS uses to conduct geologic work.

FUTURE WORK

The DGGS database logical model will be converted to a physical model and then a database prototype during

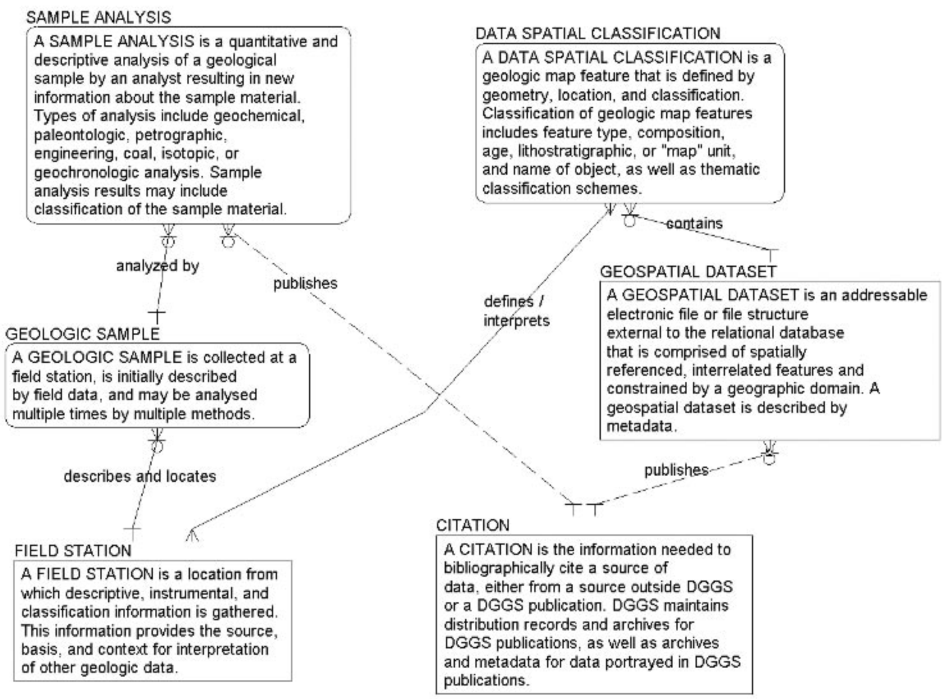


Figure 1. Selected primary entities of the DGGS logical database model. The many-to-many relationship between field data and geologic-map features is defined manually by a geologist and by geographic data manipulation. Because of the diversity and inconsistency in the way that geologic-map features are defined at DGGS, we were not able to logically resolve this relationship. Information Engineering (IE) notation (Halpin, 2000) is used in this diagram and in Figure 2; verbs describing the relationships are read from bottom to top.

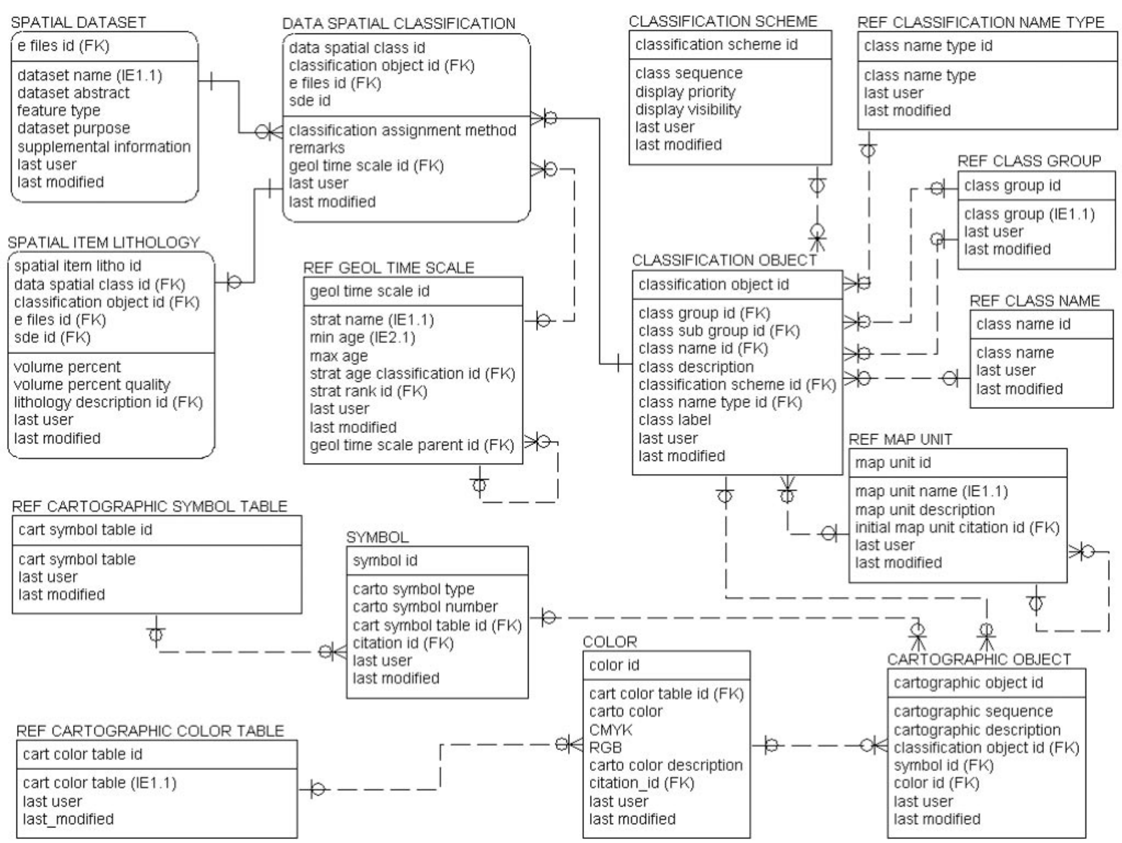


Figure 2. Entity-relationship diagram for geologic-map features in the DGGS logical database model.

summer 2002. The database will be implemented in Oracle version 8.1.7, and ArcSDE 8.2. DGGS then will begin loading and manipulating data through SQL scripts, procedure language, and Oracle SQL+ command-line interface, and viewing the data through ODBC (open database connectivity) client software and ArcGIS. This will allow us to discover any flaws in the design and implementation before we start to build custom delivery applications.

In the next year, DGGS will begin using the database to compile geologic data collected during the 2003 field season. Other DGGS projects with their own data sets will begin connecting to the DGGS database. The database project staff will continue loading data from multiple data sources, and will begin creating custom applications using the database to produce output files for Web pages, publications, and distribution to external databases.

ACKNOWLEDGMENT

The DGGS database project is funded by a U.S. Geological Survey grant through the interagency Minerals Data and Information at Risk in Alaska (MDIRA) program.

REFERENCES

Freeman, L.K., 2001, A case study in database design: the Alaska geologic database, *in* Soller, D.R., ed., Digital Mapping Techniques ‘01—Workshop Proceedings, U.S. Geological Survey Open-File Report 01-223, p. 31–34, <http://pubs.usgs.gov/of/2001/of01-223/freeman.html>.

Halpin, T., 2000, Entity relationship modeling from an ORM perspective: Part 3: Journal of Conceptual Modeling, Issue 13, <http://inconcept.com/jcm/April2000/halpin.html>.

Johnson, B.R., Brodaric, B., Raines, G.L., Hastings, J.T., and Wahl, R., 1999, Digital geologic map model; Version 4.3: Association of American State Geologists/U.S. Geological Survey draft document, 69 p., <http://geology.usgs.gov/dm/model/Model43a.pdf>.

Using NADM in a Distributed Framework

By Eric Boisvert, Annie Morin, and Martin Anctil

Geological Survey of Canada, Québec Division
880 Chemin Ste-Foy
Québec City, Québec, Canada G1V 4C7
Telephone: (418) 654-3705
Fax: (418) 654-2615
e-mail: {eboisver,amorin,manctil}@nrcan.gc.ca

INTRODUCTION

The goal of the Canadian Geoscience Knowledge Network (CGKN, <http://www.cgkn.net>) is to establish a framework "which would link all of the government geological surveys and could potentially include knowledge held within academic institutions and the private sector. The resulting 'single window' access will facilitate national and international access to Canadian geoscience knowledge and incorporate Canadian geoscience data into the Canadian Geospatial Data Infrastructure." The approach chosen by CGKN is to establish links between data sources without imposing formal structure and yet dynamically link the information seamlessly. GIS interoperability is well served by several functional tools that allow the display of maps from various sources without prior conversion. The emergence of Web mapping technology now allows a user to dynamically merge map data from remote locations through the Web. Although these technologies are functional for the geometric aspect of GIS, the content of the maps is far from easy to integrate. Various organizations use particular classification frameworks, often conflicting, based on historical usage, particular institutional interest or mandate, or they do not impose any rules at all, resulting into a gigantic patchwork of classifications that must be integrated into a single coherent database. In a previous paper (Boisvert and others, 2001) we discussed how we experimented with a small system that extracts information from several databases and presents the results into a single report without having to physically link the underlying databases. In this paper we present new developments and how some of our earlier ideas have been implemented.

DO YOU SPEAK NADM?

GEOMDB Prototype

GEOMDB (GEOscience Multiple DataBase) presented in Boisvert and others (2001) was our first attempt to connect various databases into a single interface. Its architecture was relatively simple and used the NADM-Cord[1] (Brodaric and others, 1999a) Conceptual Object Archive (COA) structure to index the information from one database to the other. (The NADM-Cord COA has evolved from the original meaning of COA, as documented in the v4.3 of the model, Compound Object Archive, but the differences are purely academic for our usage. In this paper we use COA and concept interchangeably.) The only shared part between the various databases was a unique identifier related to a concept (the COA) in the local database. The list of globally known concepts was kept in a central registry, and each local database had the responsibility to maintain a correlation between these global concepts and their local counterparts. When a query was issued at the central registry, the demand was cascaded to all local databases, which then converted the global identifier into the local identifier to search for the information. When found in the local database, the result was sent back to the central registry as a small part of a Web page that was combined with positive results from all other local databases, reassembled into a single page, and sent back to the client. The whole system depends on the fact that almost everything in NADM-Cord is tied in one way or another to a COA. This feature was the backbone for the interoperability of various NADM-Cord implementations that we tested. With a COA reference, it is possible to extract information from any related map, such as a single map legend element or descriptive attributes (including images and text), all of which can be attached to a COA. The COA really acts as feature-level metadata and can be used as a point of contact between databases.

Improvements

Two major changes, one to the data model and the

[1]In this document, we use NADM-Cord for the NADM (North American Data Model) 5.x variant used by the CORDLink (and other) digital libraries (Brodaric, 1999a). See Boisvert (1999) for a comparison of NADM-Cord and NADM 4.3.

other one to the GEOMDB framework, have put the above-discussed idea in a new perspective. The first change was to adopt an idea put forward by Brodaric and others (2001) in the U.S. National Geologic Map Database's Kentucky prototype (an object oriented version of NADM; see documentation at the Web site, <http://geology.usgs.gov/dm/steering/teams/design/>. The idea is to eliminate what are called lookup tables, or independent lists of terms that are used to populate various areas of the data model, and instead concentrate all terms in the concept domain of the data model (as COA's). This means that all keywords or terms become concepts in their own right. The immediate impact is to simplify the management of terms in the database, and the long-term impact is to establish a complex network of interrelationships between concepts. For example, a rock unit concept is related to a stratigraphic age concept to define the unit's age; this age concept can in turn be linked to another piece of information (for example, a piece of text, or another concept, like a geochron age) that was not foreseen by the person loading the rock unit in the database.

This technique has been used by Davenport (this volume) to build an emerging encoded science language. He established that certain geological concepts are best described by the union of several concepts. For example, a rock type can be described by a conjunction of material, genesis, and texture/fabric. A map unit can in turn be described by a collection of rock types and an age. The cascading effects of linking one concept to other concepts allows the linking of a map unit with genesis (and genesis to environment, and so forth). This technique seems intuitively closer to the way geological information is structured. The approach used in the GEOMDB prototype seems profitable because this small improvement in the data model opens a realm of possibilities, such as dynamic reclassification (for example, unit into ages: because the unit concepts are related to age concepts, it's possible to reclassify units as ages) and the possibility to query the database about the possible relationships between concepts even if the relationship is not encoded in a single database. The map unit-age relationship could lie in one database, and the map unit-genesis relationship could reside in another. Joining the results of both databases, a user could find where a particular genesis is found at a specific time, even if this information is not explicitly coded in one database instance.

Boisvert and others (2001) noted "we are of course toying with the idea of using XML as an exchange mechanism," and, in this year's work, we did indeed. This first prototype exchanged specially formatted HTML pages (actually, snippets of pages), and converting this information to XML added a new dimension to the project: the possibility to process rather than simply to display the result. The original approach used only a single mediator software (the piece of software that translates the database content into an HTML page) and the resulting set of pages was merely reassembled and displayed. Using XML we can now ask another mediator to receive the series of XML responses and process them, and do something useful with the result. Using XML also allows software other than browsers to use the server response.

This opens another set of possibilities in database interoperability. This approach relies on a translation mechanism that brings information stored in structure A to a portable format that can be translated back to structure B with another translator. This "lingua franca" method is already used by software like FME <http://www.safe.com> where a common format (based on SAIF) is used as a launch point toward another format. Our goal for geological information is to use NADM-Cord as a lingua franca between database structures. Because the goal for CGKN is to share database content, the goal for a specific agency participating in this exchange is to provide a mechanism to translate its local structure into NADM-Cord concepts and constructs (Figure 1). This is done usually through the creation of mediators, which are pieces of software that translate back and forth between NADM-Cord and local structure and content. Developments made during the past year advanced further the concept of connecting distributed databases using an emerging concept called Web services.

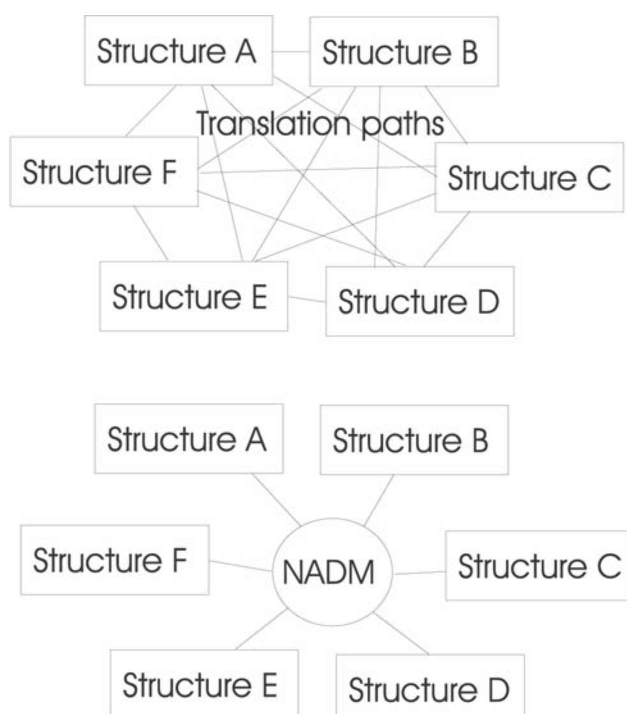


Figure 1. NADM-Cord as an interchange language between various database implementation. The top part of the figure shows a brute-force solution for interoperability; the bottom part shows our vision.

WEB SERVICES

People working in the information technology domain are aware of the Web service revolution. Simply stated, a Web service works similarly to the standard Web page server we all are familiar with, except that the Web page is formatted in such a way that another machine can read it and process it. It is called a service simply because it offers a small piece of information of processing logic to whatever client that might want to use it; it is not exclusive to a given platform or software. Many emerging standards address this concept to make it work: A software application somewhere uses a URL (Universal Resource Locator) to locate a machine where a piece of information is stored, requests the information (like a browser would request a Web page), and the server generates a specially formatted page (using XML) that is parsed back to be used by the calling software. The calling client can itself be a server for another application. The technique has many benefits, principally that it is relatively easy to implement because you need only a Web server and a scripting language to generate pages dynamically. (More sophisticated approaches are available.) The new .NET platform from Microsoft makes Web services even easier to deploy. This technology is ideal to implement our NADM-Cord as a lingua franca idea; we are experimenting with this.

PROTOTYPE SERVICES

The COA Service

The first problem we had to resolve in the distributed database project was how to store a single copy of the COA tree (see Boisvert and others, 2001, for the rationale) and give access to several users at once. In other words, we needed to determine how to remotely manipulate a list of COA's (which is the database incarnation of a concept). A service has been created on our Web server to permit manipulation of a COA tree; operations such as "Create a new COA," "Move a branch," and "Get a copy of a branch" are all possible through a series of ColdFusion pages specially formatted to be parsed by Geomatter (Boisvert and others, 2000; Brodaric and others, 1999b).

Figure 2 shows schematically how this service can be used to synchronize information between a central database and a local database. The call to the service is done through a regular URL[2] to a specific page, which activates a ColdFusion script. These scripts operate on the database and generate responses formatted in XML, which Geomatter parses back, making the necessary adjustment to the user interface (that is, creating a visual representation of the information). This is a good example in which the client of the service is not a Web browser. The user working on Geomatter never sees the XML and is not aware that a conversation is being held between Geomatter and the Web server; the user is shielded by a user interface that displays familiar Windows controls.

Map Availability Service

This service is a direct improvement on the HTML-

[2]This is only one of the legal methods to access a Web service. Other methods such as HTTP-POST and SOAP (Box and others, 2000) are other means to communicate with a Web service.

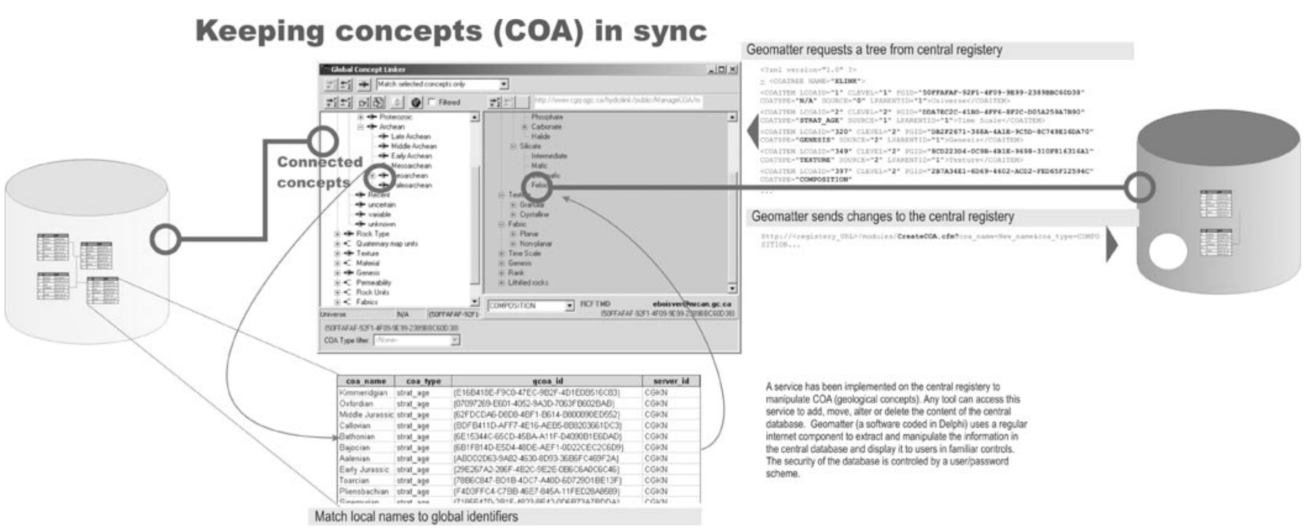


Figure 2. Geomatter as a client of a service. Geomatter calls the concept "manipulation service" and can manipulate the content of the central registry through a series of operations. The server side is encoded using ColdFusion scripts that interact with the database and return results to Geomatter using XML formatted pages.

only solution presented in Boisvert and others (2001) because XML is now used. Each server is asked to return a list of sources (maps) that contain or display a specific geological concept. This list of sources is formatted in XML (see Example 1) so it can be further processed. This is then displayed by the central portal, which shows the results from a set of local databases as if it were one seamless database.

In Figure 3, the central portal receives a request to find maps that contain a certain geological concept. This request is passed to all local databases, which return an XML segment as a response. The central database can then easily merge XML segments and present a single response to the issuer of the request. The XML (Example 1) contains all necessary information to locate the map through a standard Web Mapping Service (WMS) call. The WMS is an Open GIS Consortium (OGC) standard to request maps over the Internet (OGC-WMS, 2000); this technology is gaining wide acceptance among GIS vendors.

Example 1. *Sample of a XML response from the server.*

```
<?xml version="1.0"?>
<MAPLIST CLIENT="SIMPLE_MAPSERVICE">
<!—this is where on the web the map is➔
<MAP SRC="http://www.cgq-qgc.ca/cgi-bin/mapserv_35s.exe?map=d:/webcgq/hydrolink/data/maps/production/
english/surf_wms.map">
    Surficial geology of Canada
<!—This is its full name➔
<NAME>Surficial geology of Canada</NAME>
    <!—this tell us that the map can be accessed using WMS protocol➔
    <TYPE>WMS</TYPE>
    <!—the projection of the source map, using EPSG[3] codes➔
    <PROJ SRS="EPSG:4269" />
    <!—the limits of the map➔
    <BBOX XMIN="-140.0" YMIN="40.0" XMAX="-40.0" YMAX="85.0" />
        <!—and a list of layers composing the map➔
        <COVERS>
        <COVER NAME="10002">Surficial Geology of Canada</COVER>
        </COVERS>
</MAP>
<!—and we continue with the next map➔
<MAP SRC="http://www.cgq-qgc.ca/cgi-bin/mapserv_35s.exe?map=d:/webcgq/hydrolink/data/maps/production/
english/piedmond_wms.map">
    Carte des formations de surface de Portneuf
    <NAME>Carte des formations de surface de Portneuf</NAME>
    <TYPE>WMS</TYPE>
    <PROJ SRS="EPSG:4269" />
    <BBOX XMIN="45.5" YMIN="-71.5" XMAX="45.0" YMAX="-70.5" />
        <COVERS>
        <COVER NAME="10006">Unit&eacute; geologie de surface</COVER>
        <COVER NAME="10009">Forages</COVER>
        <COVER NAME="10031">Station GIMS</COVER>
        </COVERS>
… continues…
```

[3]EPSG = European Petroleum Survey Group. This group developed a standard set of code for commonly used projections (<http://www.epsg.org/>). This standard has been adopted by OGC to represent projections.

This XML segment can be consumed by any application that can parse XML tags. We coded a simple consumer (a *consumer* is a service user, to extend the business metaphor; this term is widely used in Web service literature) that can merge a selected list of maps into a single view (Figure 3), but because this has been established as a service, other applications can use it. For instance, someone writing software in Visual Basic might want to use this service and call it from within its code; the service is totally independent of which client is using it. The map is not trapped in this Web page. Other applications can use the service and extract the map. For instance, we wrote a

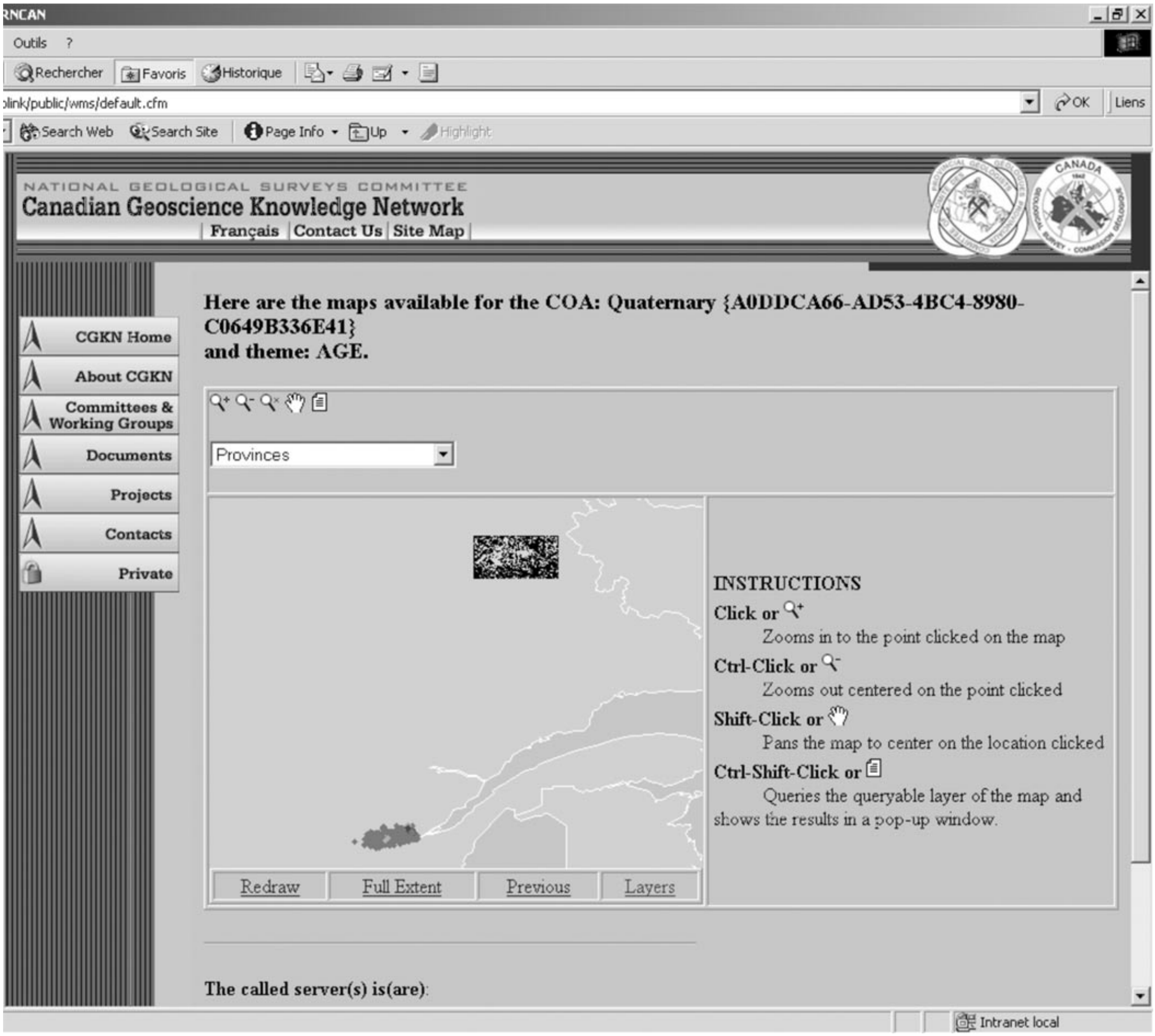


Figure 3. Web site as a client of a service. The content of this page is the result of a series of calls made to various services. The application is interacting with a server that merges information from various servers using Web services.

small application that can read this service back and build a composite view of maps aggregated from a single query to the central registry.

OGC Related Services

Because OGC is becoming an important part of Canadian spatial infrastructure, we developed, with the help of Compusult Inc., in Newfoundland, Canada, the first step of WMS compliance for NADM-Cord. This service delivers a standard WMS GetCapability and GetFeatureInfo with information extracted from the NADM-Cord framework, achieving for NADM-Cord a primary aspect of WMS interoperability. The GetCapability is a standard WMS call to identify what is available from a server (which map, layers, projections, metadata, etc.). The GetFeatureInfo allows users to query a specific feature from a map and extract its attributes. The response format is not specified in the WMS specification, and we had to create one that would fit the NADM-Cord requirements. A typical result is presented in Example 2.

Example 2. *Response for a GetFeatureInfo*.


```
<xml version='1.0' encoding="UTF-8" standalone="no" ?>
<NADM VERSION="5.2" xmlns="http://www.cgkn.net/NADM">
<!—A feature_block is created for every spatial object in the list—>

<FEATURE_INFO>
<DBSOURCE xmlns:xlink="http://www.w3.org/1999/xlink" xlink:type="simple" xlink:href="http://
www.cordlink1.org">

<REQUEST MODE="SELECT">
 <SOURCE SOURCE_ID="24">Geological Map Of Canada</SOURCE>
 <NADM_DATASET DATASET_ID="56">Geology
 <SPATIAL_OBJECT_ID ID="100">
 </NADM_DATASET>
 <NADM_SERVICE URL="http://www.cordlinkg.org"/>
</REQUEST>

<CLASSIFICATION SCHEME_ID ="125" CLASS_OBJ_ID="1265">
 <CLASS_LABEL>Dst</CLASS_LABEL>
 <CLASS_NAME>Talwar Formation</CLASS_NAME>
 <COA COA_ID ="55">Talwar Fmt
 <COA_ATT DESC_TYPE="IMAGE" DESC_ID="25"/>
 <COA_ATT DESC_TYPE="IMAGE" DESC_ID="56"/>
 <COA_ATT DESC_TYPE="TEXT" DESC_ID="1123"/>
 </COA>
 <COA_REL COA_REL_TYPE="ROCK COMPOSITION" COA_ID ="225">Calcarous limestone</COA_REL>
 interbedded with <COA_REL COA_REL_TYPE="ROCK COMPOSITION" COA_ID ="123">minor
 shales</COA_REL> of <COA_REL COA_REL_TYPE="AGE" COA_ID ="1234">devonian age</COA_REL>
```

…continues…

CONCLUSION

This work has given us an opportunity to experiment and crystallize our vision about how distributed databases can work. Since we started this project, the Web service paradigm has flourished, and large companies (e.g., Microsoft) are adopting it. Now there are more solid standards, such as SOAP (Box and others, 2000) and WSDL (Christensen and others, 2002), we can use to implement the ideas we tested. The new .NET platform makes service creation and consumption extremely easy to implement—it's a matter of adding a keyword to a function. We are now looking at these tools to redesign our interoperability platform and publish these services for the outside world to use.

ACKNOWLEDGMENTS

As always, Peter Davenport and Andrew Moore of the Geological Survey of Canada and David Soller of the U.S. Geological Survey were kind enough to review the manuscript.

REFERENCES

Boisvert, E., 1999, NADM variant fact sheet: U.S. Geological Survey, <http://geology.usgs.gov/dm/steering/teams/design/variant/CORDLink_Variant_Description.htm>.

Boisvert, E., Desjardins, V., Brodaric, B., Berdusco, B., Johnson, B., and Lauzière, K., 2000, Geomatter II, a progress report, *in* Soller, D.R., ed., Digital Mapping Techniques '00—Workshop Proceedings: U.S. Geological Survey Open-File Report 00-325, p 87-95, <http://pubs.usgs.gov/of/2000/of00-325/boisvert.html>.

Boisvert, E., Morin, A., Lauzière, K., and Lebel, D., 2001, Using the proposed North American Data Model in a distributed database environment, *in* Soller, D.R., ed., Digital Mapping Techniques '01—Workshop Proceedings: U.S. Geological Survey Open-File Report 01-223, p. 35-43, <http://pubs.usgs.gov/openfile/2001/of01-223/boisvert.html>.

Box, D., and others, 2000, Simple Object Access Protocol (SOAP): W3C Note 08 May 2000: World Wide Web Consortium, < http://www.w3.org/TR/SOAP/>.

Brodaric, B., and Hastings, J., 2001, Evolution of an object-oriented, NADM-based Data Model Prototype for the USGS National Geologic Map Database Project [abs.]: Annual Conference of the International Association for

Mathematical Geology, IAMG2001, Cancun, Mexico, <http://www.kgs.ku.edu/ Conferences/IAMG/Sessions/I/ brodaric.html>.

Brodaric, B., Journeay, M., Talwar, S., Boisvert, E., and others, 1999a, CordLink Digital Library Geologic Map Data Model Version 5.2, <http://cordlink.gsc.nrcan.gc.ca/cordlink1/ info_pages/ English/dm52.pdf>.

Brodaric, B., Boisvert, E., and Lauzière, K., 1999b, Geomatter: A map-oriented software tool for attributing geologic map information according to the proposed U.S. National Digital Geologic Map Data Model, *in* Soller, D.R., ed., Digital Mapping Techniques '99—Workshop Proceedings: U.S. Geological Survey Open-File Report 99-386, p. 101-106, <http://pubs.usgs.gov/openfile/of99-386/brodaric2.html>.

Christensen, E., Curbera, F., Meredith, G., Weerawarana, S., 2001, Web Services Description Language 1.1 (WSDL): W3C Note 15 March 2001: World Wide Web Consortium, <http://www.w3.org/TR/wsdl.html>.

Open GIS Consortium-Web Mapping Service (OGC-WMS), 2000, Web Mapping Service specification documentation, <http://www.opengis.org/techno/specs/00-028.pdf>.

Got Digital Map Data. Now What? How the Idaho Geological Survey Distributes Digital Geologic Map Data

By B. Benjamin E. Studer and Loudon R. Stanford

Idaho Geological Survey
Third Floor, Morrill Hall
University of Idaho
Moscow, ID 83844-3014
Telephone: (208) 885-7479
Fax: (208) 885-5826
e-mail: bens@uidaho.edu

INTRODUCTION

The Idaho Geological Survey has been collecting and disseminating digital geologic data for 10 years. Recently, the Idaho Geological Survey released the first publication in a new Digital Data Series: *Digital geologic map of the St. Maries 30 x 60 minute quadrangle, Idaho*.

The Digital Geologic Mapping Lab at the Idaho Geological Survey currently distributes its compiled geologic maps, when possible, as three products: a color, print-on-demand, paper map; a Portable Document Format (PDF) version of the paper map; and a digital geologic map complete with FGDC-compliant metadata and a structured data model for storing map data such as contact, fault, or symbol attributes.

Key to the success of Idaho digital geologic map data is the adoption of a digital geologic map data model. The Idaho Survey's model is a variant of version 4.3 of the North American Digital Geologic Map Data Model (NADM <http://geology.usgs.gov/dm/>). Unlike the NADM, which is designed to be a catalog of separate, digitized geologic maps, the Idaho model is designed to be a statewide database of the best available geologic map data collected in edge-matched tiles. More information is available at the Idaho Geological Survey Web site: <http://www.idahogeology.org/Lab/datamodel.htm>

TILES OF EDGE-MATCHED DIGITAL GEOLOGIC MAPPING

The Idaho Geological Survey is compiling digital geologic map data in 30 x 60 minute tiles. The map data are designed for merging with adjoining, edge-matched data sets.

- When possible, new field work is melded with existing mapping to create a more complete data set. Updated geologic data will then be released with a new version number.
- Original geologic mapping, if possible, is compiled at 1:24,000 or the most detailed scale possible.
- The original publication source and other geologic information about each feature (contacts, faults, symbols) is tracked with geologic object metadata and data-model attributes.
- When possible, both a paper map product and a digital data set are released. Both publications have the same authorship. The data sets are released as Digital Geologic Maps in the Idaho Survey's Digital Data Series. A digital compiler credit also is included with the digital publication.

BACKBONE OF THE IDAHO SURVEY DIGITAL GEOLOGIC MAP

A data model for digital geologic maps provides a framework to support the many different components that create a geologic map. The Idaho Survey variant (some might call it a deviant) of the NADM was designed to meet the following objectives:

- To provide a framework on which to store geologic map data and corollary legend information created by the Idaho Survey.
- To work with in-house systems and procedures for collecting and attributing geologic map data.
- To use the structure and design developed for the NADM v.4.3, where possible.
- To be expandable to handle new data types as the need arises.
- To permit data updates and transfers to future formats and structures.
- To develop tools to access the stored data.

SUMMARY

The Idaho Geological Survey has released the first publication in its new Digital Data Series: *Digital geologic map of the St. Maries 30 x 60 minute quadrangle, Idaho* (Lewis, and others, 2001). The format includes a digital spatial map, a PDF version of the paper map, geochemistry, and more. Five more data sets are to follow in early summer 2002. A Print-On-Demand paper map also will be available for many of the digital maps. Digital data is available online on our Web site <http://www.idahogeology.org> and may be purchased as a CD.

By creating the digital geologic map tiles and their associated data sets, a group of highly useful statewide databases are produced as a by-product: map units, geochemistry, geologic map sources, formal stratigraphic units, colors, symbols, and metadata.

This paper is the condensation of a poster presented at the Digital Mapping Techniques 2002 conference in Salt Lake City. The poster is available for viewing, in PDF format, on the Idaho Geological Survey's Web page: <http://www.idahogeology.org/Lab/DMT/>

WHAT'S TO COME

The Idaho Survey will continue to release more digital geologic-map data sets. Currently, the power of the digital map and data model is limited to users who have a working knowledge of geology and GIS techniques. IGS tools planned for the near future are: an ArcIMS-WWW Interface, ArcView software query tools, geochemistry management and distribution tools, and color and symbol selection tools.

REFERENCE

Lewis, R.L., Burmester, R.F., Kauffman, J.D., and Frost, T.P., 2001, Digital geologic map of the St. Maries 30 x 60 minute quadrangle, Idaho: Idaho Geological Survey Digital Geologic Map 1.

Digital Archives and Metadata as Mechanisms to Preserve Institutional Memory

By John C. Steinmetz, Richard T. Hill, and Kimberly H. Sowder

Indiana Geological Survey
Indiana University
611 North Walnut Grove
Bloomington, IN 47405-2208
Telephone: (812) 855-7636
Fax: (812) 855-2862
e-mail: {jsteinm, hill2, sowderk}@indiana.edu

ABSTRACT

Metadata are essential for any reliable geographic information system (GIS). Their utility extends beyond GIS, however. The Indiana Geological Survey (IGS) is employing metadata as a means to help preserve institutional memory. Due to its size and diversity, the IGS generates a multitude of data types in any given period of a few months. Unless properly documented, field and analytical data, samples, unpublished maps, well and mine records, and other similar data and information can be lost within the organization. Since much of this data is commonly used by various staff for different projects and for different reasons, the opportunity to misplace or lose it is great. Moreover, the permanent departure of any employee, whether through retirement or normal attrition, is an occasion for an organization to lose a large amount of knowledge about how things were done, the stage at which projects were abandoned, and even the physical location of important data. Once lost, and only if resources permit, can precious staff time be used to search for or even reconstruct or re-collect the data.

Virtually all data collected or processed by the IGS are geospatially oriented. Hence, the application of metadata to their cataloging within an organization is easy to envision. By constructing metadata incrementally throughout the life of a project, a means is provided to ensure that adequate documentation is captured upon publication of products, and also prior to an investigator's departure.

INTRODUCTION

Congratulations! You have just won the $10 million PowerBall Lottery! After carefully reconsidering your life's priorities, you promptly turn in your resignation before retiring to the Bahamas, leaving no forwarding address. But what about those data sets you have been working on for the last five years? Does your organization lose your knowledge about the data? Will the person who replaces you and inherits your data know enough about them to use them? Even though this will not be of concern to you as you bask on the beach, it should be of great concern to the organization.

INSTITUTIONAL MEMORY AND DISCOVERY

Data are expensive. Whether they are gathered by an investigator in the field or generated by an analyst in the laboratory, the cost of their acquisition is great. As expensive as data are, the replacement cost is even greater: Scientific staff must be paid again, and equipment maintained and refurbished, assuming the original collection site still exists.

Beyond the cost of collection, the value of scientific data lies in their use. Whether used by the original investigator or years later by someone else, scientific data only maintain value if utilized. To be of utility, they must be accessible; to be accessible, they must be discoverable. Many organizations rely heavily on institutional memory—the collective knowledge and history of an organization held by employees, especially those who have been there for a number of years (National Research Council, 2002)—to aid in the discovery and accessibility of data. All too often when a long-term and productive employee retires or leaves an organization, an institutional memory of inestimable value is lost. With him or her commonly go such simple information as the physical whereabouts of a data set. Once lost within the organization—namely, once its location is no longer known within the building—the data set is essentially useless.

Finding (discovery of) data within an organization involves identifying the existence and location of desired data sets and collections. Ancillary considerations include ascertaining data availability, quality, and format. Within the petroleum industry, it has been estimated that between 60 and 80 percent of a geoscientist's time is spent searching for data; the balance is spent organizing and analyzing it (S. Natali, Barrett Resources, personal commun., 2001). One internal goal of most public or private organizations is to shorten the discovery time so the investigator can invest more valuable time in using the data. In many instances, however, potential users gain knowledge of, and access to, data by traditional means: through personal acquaintances, letters, on-site visits, or by telephone, fax, or e-mail. Too often, knowledge of the mere existence of geoscience data is reliant on personal relations, that is, institutional memory.

Discovery of data from outside an organization requires a certain degree of public relations by the organization that is archiving data. The IGS, like many state and federal Earth science institutions, promotes its holdings not only via e-mail and the Internet, but also through mass mailings, professional meetings, posters, and CD-ROM's. Digital data catalogs and access to them over the Internet are increasingly common, but many investigators are surprised to learn that digital access to a catalog's collection is not yet available. In many instances, funds to build an electronic catalog and provide Internet access are available only when garnered from existing operational funds; new money for these efforts rarely is afforded.

Adequate cataloging of data may seem time consuming and not terribly exciting, yet the costs involved in data acquisition generally far outweigh all other costs combined; reacquisition of data, if even possible, is more costly than initial acquisition and retention. In the current economy, organizations simply cannot afford to lose the usefulness of valuable data by neglecting documentation procedures.

The IGS has undertaken the task of inventorying, cataloging, and creating metadata for new and historical data sets. To promote this initiative, the IGS has formed the Data at the Indiana Geological Survey Committee (DIGS Committee). The chair of the committee is the head of the Technology Transfer Section, and the members include staff across all IGS disciplines. They are charged with examining the factors involved in conducting a Survey-wide inventory of files and records, samples, archives, and publications. The committee's immediate goals are to capture IGS data, to develop the means of data retrieval, to develop a database to organize and access records of all IGS data, and to provide Internet access to an inventory of selected records. Some considerations the committee is taking into account are the lumping and splitting of items into various categories, design of inventory forms, necessary resources (personnel, equipment, supplies), metadata and quality assurance, public vs. proprietary data, barcoding, prioritization of data capture, efficiency and ease of the inventory process, staff training, and Internet deliverability.

Geospatial Data

With the exception of administrative records, virtually all of the data at the IGS have a geospatial component to them. ("Geospatial" refers to information that identifies the geographic location and characteristics of natural or constructed features and boundaries on the Earth.) Some of the types of data at the IGS include maps, publications, open-file studies, CD-ROM's, rock and mineral specimens, thin sections and rock analyses, fossil specimens and paleontological data, outcrop descriptions and photographs, cores and core descriptions, ground penetrating radar, downhole sample data and interpretations, shallow and deep geophysical data, field measurements, field and laboratory chemical and physical analyses, aerial photographs, digital information of an increasing variety, card files, lithologic strips, project files, reports, drawings, transparencies, X-ray data, seismic data, field notes, log picks, photographic negatives, cross sections, test procedures, gravity and magnetic data, and more.

Metadata

The phenomenal development of telecommunications in the late 1990's has been accompanied by fundamental shifts in how scientific data are gathered, accessed, and used. Until recently, most data and information were published in paper format (for example, books or maps), and access was provided through card catalogs (paper or electronic) with limited search capabilities. Now, data can be posted on the Internet, searched in a multitude of ways, and accessed through clearinghouses and numerous other portals. To facilitate discovery and access, metadata are used.

Metadata are descriptive information about data and information resources. Typically, metadata describe, point to, or otherwise complement the information content of the data to which they are related. Metadata provide a concise aid in locating desired information and help make such information easily accessible. It is particularly useful for geospatial information because federal standards have been written.

On April 11, 1994, President Clinton signed Executive Order 12906. This order, among other things, established "a coordinated National Spatial Data Infrastructure (NSDI) to support public and private sector applications of geospatial data in such areas as transportation, community development, agriculture, emergency response, environmental management, and information technology." Additionally, Executive Order 12906 mandated "the Standardized Documentation of Data . . . each agency shall document all new geospatial data it collects or produces, either directly or indirectly, using the standard under development by the FGDC [Federal Geographic Data Committee], and make that standardized documentation electronically accessible to the Clearinghouse network."

The FGDC standard describes what information is to be provided by the metadata and in what format the data should be provided. For example, the FGDC standard directs that the producer of a data set must describe the data's quality. Metadata help ensure that data remain usable in perpetuity. Moreover, metadata provide assurance that the data are of sufficient quality and validity, and eliminate one of the greatest barriers to the use of scientific data: discovery.

Training of IGS staff in metadata guidelines as specified by the FGDC Content Standard for Digital Geospatial Metadata has been provided through a series of in-house workshops. A policy has been prepared by the IGS administration (see Appendix) to require metadata creation for all new data sets and the creation of project metadata for final products. The ultimate goal is to increase the value of already valuable data and make it easier to access and retrieve. The IGS Intranet Web site currently provides easy access to staff for metadata keyword and category searches, and the IGS Internet site will ultimately provide users with data that can be downloaded.

Access

Balanced against the cost of acquisition, the cost of retaining geoscience data is a mere fraction, and data may acquire an increased value through time (Montgomery, 1999), (Figure 1), yet unless those data are accessible, they are useless. Before the electronic age, lists of data in collections were kept in (serial) logbooks or on (alphabetic) file cards. An individual familiar with the order of the record-keeping system was essential, to look up the data listing and to locate the physical whereabouts of the desired data. Access depended on a high degree of institutional memory, and on individuals who cared about the system and its organization. Archives that rely on institutional memory are prone to degrade when staff transfer, retire, or otherwise leave the institution. Today, computer databases that catalog a collection's holdings can be searched and queried by any number of descriptive parameters, even remotely over the Internet, utilizing much of the same technology developed by libraries.

The American Association of Petroleum Geologists (AAPG) has been promoting geoscience preservation and access for over 50 years. It has had a standing committee for core and sample preservation since 1948, and supports the American Geological Institute (AGI) proposal to create a centralized repository, the National Geoscience Data Repository System (NGDRS)—in effect, a Library of Congress for samples in the public domain (American Geological Institute, 1994, 1997; Montgomery, 1999). To initiate the formation of the NGDRS, AGI secured support from the U.S. Department of Energy and some petroleum companies, developed a repository data model, facilitated the transfer of some data (cores, cuttings, paleontological samples, seismic data, logs, and scout tickets) from the private to the public sector, and implemented and is currently operating GeoTrek, a software data catalog and access system available on the Internet <http://www.agiweb.org/NGDRS/>. The location of the centralized facility has not been determined, and petroleum industry support is mixed. Companies are not willing to donate materials until a repository is located (Montgomery, 1999), and many individuals feel that a network of distributed repositories at key locations in the country would foster a greater degree of use. Finally, many state geological surveys would not contribute their materials, since they already have a statutory obligation to archive state-derived data.

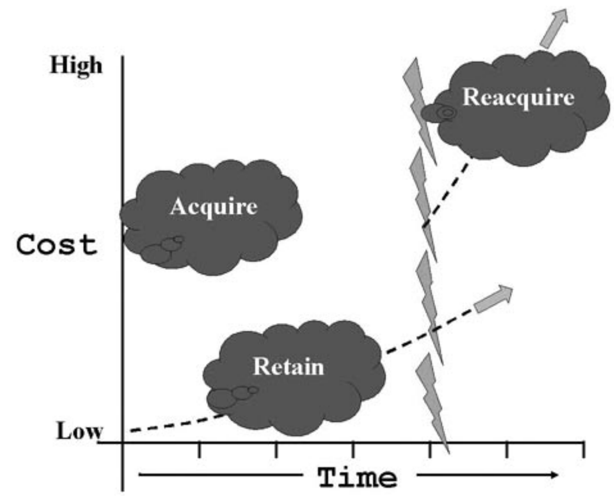


Figure 1. The short-term cost and value of data, either gathered in the field or generated in the laboratory, differ from their long-term cost and value. While the initial cost to acquire data may be quite high (left), the annual and ongoing costs for retention can be low. The costs of reacquiring the data at some time in the future (to the right of the jagged lines), if reacquisition is even possible, are typically much higher than the original acquisition costs. (Figure modified from National Research Council, 2002.)

Computerization

Computerization involves the digitizing of paper records, copying from one electronic medium to another, and/or re-formatting existing digital data. Increasingly, collections are cataloged in digital databases. Nevertheless, paper serves as an important medium of storage, if only as a visible backup. At most institutions, few specimens are accompanied by digital data when they arrive. In nearly all cases, specimen data arrive as collector-generated labels, scientific publications, and maps that accompany the samples. Specimen data are usually prepared for computer entry by initially organizing them on handwritten forms. Although this seems cumbersome, the two-step process cuts down on errors and leaves a tangible trail. The goal is an error-free inventory database.

Once in digital form, data are not guaranteed immortality. Data loss can result from physical degradation of the magnetic medium (particularly tape, which should be re-written about every 5 years), obsolete formats (and obsolete equipment to access them), the migration from one format to another, or the lack of complete auxiliary data (such as header information, recording parameters, calibration data, metadata).

CD-ROM storage currently is one of the more popular forms of digital data storage. Benefits include a simple and low-cost replication process, ability to store multiple data sets (e.g., text, images, video, and audio), and random access to the information. CD-ROM's are also expected to have a shelf life projected to exceed 25 years under standard office conditions.

Digital data also require periodic refreshing. Accessibility and retrievability can be guaranteed only if data are migrated to protect against media deterioration and technology evolution.

EARLY LESSONS

Some of the early lessons the IGS DIGS Committee learned in undertaking an institutionwide inventory effort are:

1. There are five stages of data preservation:
   a. Data acquisition or assimilation;
   b. Storage and maintenance;
   c. Awareness;
   d. Accessibility;
   e. Usefulness (sufficient quality and validity to be believable).

   Failure of any stage results in all stages being repeated. Metadata are necessary for each stage (National Research Council, 2002).
2. Standardization of data structure is essential. It contributes to a consistent vocabulary of keywords and it facilitates metadata creation and ease of use. Similarly, templates for metadata enhance the ease of data capture and help ensure compliance with FGDC standards.
3. Bar coding serves numerous purposes and is a popular means of controlling inventory. Barcoding of items in a collection not only enhances sample identity by connecting the user immediately to more complete metadata than can be recorded on a small label or boxtop, but it achieves another important and simple task: it easily signifies whether or not an item has been inventoried and is already part of the collection's catalog.
4. Each piece should only be handled once. The magnitude of inventorying an institution with many different data types necessitates, if the inventory is to be successful, a certain efficiency. Physically handling each piece once and only once is an important step in that process.
5. Resident staff participation is essential. Individual memories are an important part of the data-capturing process. Familiarity with, and personal investment in, the data would be lost if an "outsider" (i.e., contract laborer) were brought in to merely inventory physical objects.
6. Staff participation must be sought, but only after the entire process has been thoroughly designed and rigorously tested. Since staff buy-in is critical, their participation can be assured only if they understand that the inventory will be taken only once. Everyone knows that staff time is expensive. Enthusiastic and determined participation of the staff can be won and sustained only if the inventory procedure is a tested and efficient process instead of a time-wasting experiment.

CONCLUSION

Properly cataloged geoscience data (geospatial data) are a unique and unconventional resource library of increasing value. Metadata provide a means to efficiently catalog and readily access those data. Data documentation can be a long and time-consuming process, but the value of knowing the details about data far outweighs the trouble of documentation. More information is created and shared today than at any time in the past. Users of data want easy access and quick results, as well as information guaranteeing the accuracy of the data they wish to use. Organizations should make the commitment to provide data with proper metadata, and to garner information from the individuals who have created the data before they hit the lucky numbers on that big lottery ticket pay-off.

FUTURE GOALS

Numerous state geological surveys are in the process of digitizing data and providing wider access by publishing catalogs on the Internet. Financial resources for staff and equipment are, in many cases, the only impediments to digitizing and providing Internet access to data.

The application of informatics may be an important goal in geoscience data discovery and access. In such a scenario, all data would be in digital form and accessible over the Internet. Each sample could be located by its spatial coordinates, and attendant metadata would record the circumstances under which the sample was collected and would provide quality control. Such a system requires standardized formats for data archiving, software support, data mining tools, and a knowledgeable end-user community (see, for example, the Kansas Geological Survey's Geoinformatics efforts at <http://www.kgs.ukans.edu/Geoinfo2/index.html>).

The Smithsonian's National Museum of Natural History (NMNH) is creating a "Research and Collections Information System" that approaches an informatics-based system. The intention is to accomplish three main goals: (1) better collections management to track the disposition of specimens acquired, loaned, borrowed, or disposed, and their locations; (2) online access to all digital specimen data for the benefit of museum research, collections, public program's staff, scientists worldwide, and the general public worldwide; and (3) participation in national and international informatics initiatives. Using a suite of software applications that are used internationally, NMNH staff have begun to slowly implement the system in a number of science departments. The software was chosen for its stability, ability to scale, flexibility for diverse NMNH disciplines, and ability for customization. Museum officials estimate that between 40 and 50 million records will adequately represent NMNH specimens at a cost of $55 to $75 million. Presently, there are no funds for data entry, and the collections care and informatics initiatives are stalled for lack of funds.

REFERENCES

American Geological Institute (AGI), 1994, National Geoscience Data Repository System feasibility and assessment study, submitted to the Office of Fossil Energy, U.S. Department of Energy: Alexandria, Virginia, American Geological Institute, 68 p.

American Geological Institute (AGI), 1997, National directory of geoscience data repositories. Alexandria, Virginia, American Geological Institute, 91 p.

Montgomery, S.L., 1999, Core values: the growing need for repositories: Oil & Gas Journal, v. 97, no. 46, November 15, 1999, p. 84-87.

National Research Council (NRC), 2002, Committee on the Preservation of Geoscience Data and Collections, Committee on Earth Resources, National Research Council. Geoscience Collections and Data: National resources in peril: Washington, D.C., National Academy Press, 205 p.

APPENDIX

The Indiana Geological Survey Metadata Policy

The Administration and Staff of the Indiana Geological Survey recognize the inherent value of the work they undertake and the data they generate. Additionally, they recognize the geospatial nature of virtually all of these data and information. Further, they recognize that data without proper documentation are worthless, vulnerable to being lost, support questionable and tentative decisions, and may never be used again.

As stewards of public information, the IGS has an obligation to provide high-quality, well-documented data sets and information through readily searchable and easily accessible means. This policy is established and designed to ensure and facilitate full and open access to quality data for research and education.

Metadata serve as a means to efficiently collect, preserve, manage, access, and disseminate these data and information.

The intentions of establishing a Metadata Policy are to:

Preserve data for future use;
Save time, resources, and duplicated effort;
Contribute toward building the National Spatial Data Infrastructure;
Support sound science and decision-making;
Serve as a basis for an inventory of IGS holdings.

IGS Metadata Policy

- All staff will be trained (or will learn) to write and read Federal Geographic Data Committee (FGDC) compliant metadata using the ArcCatalog metadata editor or similar software application.
- All finalized data sets will be accompanied by a compliant metadata record.
- Project leaders will be responsible for ensuring the quality and consistency of metadata throughout their projects.
- Before developing new data sets for a project, project staff will search existing data (both internal and external) to ensure that duplication of effort does not occur.
- The Technology Transfer Section has created a clearinghouse node to host all our metadata on the IGS Intranet. Staff can search for metadata by filename, originator name (e.g., EPA), area of interest (e.g., specific county), or keyword.

The Metadata Working Group (of the DIGS Committee) serves as an internal resource to provide guidance by answering questions, establishing metadata templates, and in helping to assure ease of use in the metadata creation process. They will not, however, write metadata for the general staff.

After the 6-month implementation period, comments about the metadata system will be solicited. The Metadata Working Group will review all comments and determine what, if any, changes need to be made to the system.

The Metadata Process at IGS

What are required to have metadata?

- Any completed data product
- IGS publications
- IGS Open-File Studies
- Final reports on projects, both internal and external
- Maps and GIS products
- Digital data images
- Databases
- Collections of samples or data

Who needs to create metadata?

Project director and/or project staff, that is, those closest to the actual generation of the data.

Procedure to create metadata

- Follow the file-naming conventions established by the Metadata Work Group
- Utilize the IGS metadata template to create metadata
- Follow authorship/citation guidelines
- Insert all necessary keywords
- Categories: theme, place, stratum, temporal

How should the metadata be submitted?

Completed metadata should be submitted to the publications review coordinator for internal technical and editorial review.

Following approval by the director, the metadata will be available for public release, and they will also be archived by Technology Transfer in the Document Archive Database and included in the metadata search engine on the IGS Intranet.

Handling of proprietary data

Proprietary data will be kept physically separate from those that are publicly available, and they may be used only by IGS staff or publicly released with permission of the director or his designate.

How are incomplete databases to be handled?

By way of example, as large and comprehensive as the IGS Petroleum Well Data Base (PWDB) is, its data are by no means perfect, nor is it a completed database, yet it is necessary to document data quality as precisely (and candidly) as possible for the end-user to evaluate whether or not to use the data set.

Mapping Aquifer Sensitivity in Tazewell County, Illinois

By Patrick D. Johnstone

Illinois State Geological Survey
615 East Peabody Drive
Champaign, IL 61820
Telephone: (217) 244-0365
Fax: (217)333-2830
e-mail: johnston@isgs.uiuc.edu

BACKGROUND

The Illinois State Geological Survey (ISGS), with partial support from the Tazewell County Board, conducted a study to map the near-surface geology of Tazewell County. A primary purpose of the study was to classify the near-surface materials with respect to the potential for aquifer contamination within the county. The Tazewell County Board recognized their need for information to assist with land-use planning within their communities, particularly with regard to current and future landfill needs. The ISGS, having completed similar projects for other counties in Illinois, including Carroll (McGarry and Grimley, 1997) and Jo Daviess (Riggs and others, 2000), was contracted to create a series of maps at 1:62,500 scale for Tazewell County which could be interpreted and used by people with varying levels of geologic or scientific training. The basic topographic and geologic maps, and selected derivative maps, were created during the period from December 2000 to January 2002 from existing data in the files of the ISGS.

Tazewell County covers an area of approximately 600 square miles in central Illinois. The county is bordered on the west by the Illinois River, and is bisected by the east-west-trending Mackinaw River. The surficial geology of the county is varied; the eastern half is covered by thick Wisconsin-age glacial tills, and the western half is covered by Wisconsin-age and younger outwash and fluvial deposits. Tazewell County is a predominantly rural county, but its western areas are experiencing urban encroachment from the city of Peoria.

TOPOGRAPHY

Accurate mapping of the surface topography of the county was essential as a basis for detailed modeling of the subsurface, since much of the data used for modeling was assembled from well-log information, which is based on reported depth from the surface. Topographic data were collected from the 20 U.S. Geological Survey (USGS) 7.5-minute quadrangle maps (1:24,000 scale) that fall completely or partially within the county's borders. Digital line graph (DLG) versions of 15 of the quadrangles were available; contour intervals range from 5 to 20 feet. For the remaining five quads, members of the ISGS staff digitized the hypsography on screen from USGS digital raster graphic (DRG) files, using ArcEdit software.

The arc and point coverages from the 20 quadrangles were reprojected to a common projection and datum and merged. TopoGrid was used to create a 30-foot grid, with stream and lake coverages (collected from the same sources as the hypsography) used to control valley shapes. ArcMap was used for inspection of the grid, and editing of the original coverages was performed several times to correct misplaced or mislabeled features. Contours were created digitally using the CONTOUR command in ArcGrid. To reduce pixelizing, or "grid-tracing," of contours, the contour interval was set to 10.001 feet. The resultant line coverage was corrected for topology using ArcEdit, and some manual and digital generalization was performed to make contours appropriate for display at a scale of 1:62,500. The HILLSHADE and SLOPE functions in ArcGrid were used to create shaded-relief and surface-slope grids.

Three maps were created directly from these data (Surface Topography, Shaded Relief, and Surface Slopes), and all layout and presentation were performed using ArcMap. An interesting problem that emerged during layout was the inability of ArcMap 8.1 to properly label contours. The traditional technique of blanking out a section of contour line under the label can only be performed in ArcMap by using a "halo" around the label. However, this "halo" effectively blanks out any underlying themes (e.g., elevation color bands, thematic polygons). The only practical solution found for this problem was to export the contour layer into Adobe Illustrator, insert labels, then return the layer to ArcMap—a cumbersome procedure for what should be a simple step in the map-making process.

SUBSURFACE DATA

Field collection of new data was very limited due to time and budget constraints; therefore, existing data archives of the ISGS were used extensively. The ISGS maintains immense digital and paper databases of well logs, geologists' field notes, maps, and other data going back almost 100 years.

The data source for 3-D modeling was the ISGS Well Log Database. This database stores location information and well-drillers' lithologic descriptions for most water wells, engineering borings, oil and gas exploration wells, and research borings drilled in the state. For Tazewell County, more than 5,000 well records exist.

The well records for Tazewell County were first checked for locational accuracy. For those not located within a quarter-quarter-section (approx. 330 feet), the driller's location description was compared to topographic maps, aerial photography, and plat maps to refine the location. If the location information was inconclusive, the data from that well were discarded. As a second location-accuracy test, reported elevations of all remaining wells were compared to spot elevations from the 30-foot DEM produced for the project. If the reported and spot elevations differed by more than 10 feet, the location and elevation data were scrutinized, and the well location was moved or the elevation was adjusted. If there was no satisfactory remedy, the well data were discarded.

For the accurately located wells, the lithologic descriptions were classified into five grain-size categories (fine, mixed-fine, mixed, mixed-coarse, coarse) or as bedrock. The reclassified lithologic records were then viewed in ArcView using the 3D-Analyst extension. The top of each unit in the well was assigned an elevation and the thickness data were used to compute the unit's 3-D geometry (to be shown with 30X vertical exaggeration). One-mile-wide "strips" of the wells were viewed in 3-D, first east-to-west, then north-to-south. These "virtual cross sections" were useful in helping to locate spurious data, and in reconciling records that were inconsistent with nearby data.

The result of these editing/quality control steps was a reliable data set that provided accurate data input for the modeling and made post-model editing of the data set less necessary. Although more than 20% of the available data was discarded, approximately 23,000 data points from over 4,000 wells were used in the final modeling process.

BEDROCK TOPOGRAPHY

The bedrock of Tazewell County is covered by a blanket of Quaternary and Holocene sediments (collectively called "drift") that ranges from less than 5 feet to more than 400 feet thick. Most of the ground-water resources in the county are found in these sediments. As such, mapping of the topography of the underlying bedrock surface and the thickness of this unlithified cover was important.

Data on depth-to-bedrock was collected from the well records used for the 3-D modeling. Just over 400 wells reported reaching bedrock, and these were used to create an elevation grid. The grid was further enhanced by comparison in 3D-Analyst to 400 wells that did not report bedrock but pierced the surface of the preliminary bedrock-surface grid. These wells were used to "push" the surface downward where appropriate to a depth equal to the well depth plus 5 feet. Additional depth-to-bedrock data came form an earlier study covering the southeastern corner of the county (Herzog and others, 1995).

Once the final bedrock-surface grid was created using TopoGrid in ArcInfo, ArcGrid was used to subtract this grid from the land-surface grid previously produced, creating a drift-thickness grid. These grids were contoured using the CONTOUR command in ArcGrid, and the line work was manually generalized to make contours that were appropriate for display at the 1:62,500 scale.

SURFACE GEOLOGY

A map showing the Surficial Deposits of Tazewell County was produced. Preliminary data on the characteristics of the surficial material were collected from the U.S. Department of Agriculture (USDA) soil survey of the county (Teater, 1996). Because the soil survey line work is superimposed on nonorthorectified aerial photographs, the images were projected onto USGS topographic maps, and lines were adjusted to fit topographic, hydrographic, and terrain features. Once these initial polygons were drawn, map units were delineated and classified based on: information from the well logs, published reports by the ISGS and other organizations (Follmer and others, 1979; Lineback, 1979; Herzog and others, 1995; Wilson and others, 1994; Hansel and Johnson, 1996), unpublished maps and reports on file at the ISGS, and consultations with ISGS Quaternary Section geologists.

AQUIFER SENSITIVITY

One of the most important derivative maps produced for the project was an aquifer sensitivity map, to be used by county staff as a tool to assist with long-term land-use planning. The classification system used to assess aquifer sensitivity was previously developed by the ISGS and used in other areas of the state (Berg, 2001). The system is based on two criteria: depth to the top of the uppermost aquifer, and the thickness of that aquifer. In this geologic classification scheme, any lithologic unit composed of sand and/or gravel is considered to be capable of yielding economically significant amounts of water, and is called an aquifer. No hydrologic analyses were performed. Thicker aquifers represent more important ground-water resources, and a higher priority is placed on protecting them. Nonaquifer materials (such as tills, clays, or shales) covering an aquifer represent the only natural protection to keep contaminants applied at the surface from migrating to the aquifer. Where the top of an aquifer is at or near the surface, this natural protection is minimal or absent, and the sensitivity of the aquifer is very high.

Several factors were considered when sensitivity classes were assigned to certain depth/thickness categories. The classes represent sensitivity of the aquifer to application of contaminants (1) at the surface (such as agricultural chemicals), (2) in shallow trenches (septic fields, etc.), or (3) in deeper trenches (landfills). The classifications deal with aquifers when the top of the aquifer is within 100 feet of the land surface, representing a maximum depth for infiltration of materials applied at the surface, while also recognizing that modern landfill processes commonly include trenching more than 30 feet into the subsurface.

Although the classification scheme for Tazewell County did not include aquifers with tops below 100 feet, an inset map was included to show where aquifer materials were present below this depth. In the map area, there are two major preglacial valley systems that intersect, and both are partially filled with thick proglacial sand and gravel deposits. These deep aquifers represent an important local and regional ground-water resource.

The Aquifer Sensitivity Map was created from a 3-D solid model that was generated with EarthVision software. The reviewed and edited well-log data were used to generate a 3-D solid model, and regions of coarse (sand and gravel) and fine (loam, silt, and clay) sediments were delineated. The coarse material was designated "Aquifer Material" and the fine and mixed units were designated "Non-Aquifer Material." This model was sliced into layers, each representing depths found in the classification scheme (0-5 ft, 5-20 ft, 20-50 ft, 50-100 ft, 100+ ft below ground surface). For each layer, an isopach map was generated which displayed the thickness of aquifer materials present in that layer. These coverages were displayed over images that showed the materials present at the top of each layer. By analyzing these displays, the top and thickness of each aquifer unit could be determined, regardless of whether the unit was wholly contained within one layer, or extended through several layers.

An ESRI polygon coverage was created from these displays. Polygons were drawn on-screen by overlaying the isopach maps and images. Areas were designated with the units from the classification scheme described above. Polygons were drawn by their hierarchical order in the classification—A1 first, A2 second, etc. These initial polygons were edited on screen, with consideration taken of geologic setting and model limitations. Where the modeling process extended geologic units past their mapped extent, the sensitivity class associated with these units was trimmed to better reflect the geology. Likewise, where high surface relief caused model incongruity, the polygons were adjusted. Finally, areas of low data density were checked for units which may have resulted only from interpolation by the modeling software, and polylines were smoothed for display at the 1:62,500 scale and adjusted for topology (Figure 1).

PROJECT COMPLETION

In a little more than one year, using existing data, nine preliminary maps were produced and presented to the Tazewell County Board and the Tazewell County Health Services Committee. The maps included Surface Topography, Surface Slopes, Shaded Relief, Bedrock Topography, Drift Thickness, Surficial Geology, and Aquifer Sensitivity. Also included were a map showing locations of data points and a map displaying the locations of historic landfills and resource-extraction activities, such as coal mining and sand and gravel pits. These preliminary maps were well received, and will serve as informational resources to be used by county officials and others as they weigh various decisions concerning land use. The maps are currently in scientific review and in 2002 will be published in the ISGS Open File Series.

REFERENCES

Berg, R.C., 2001, Aquifer sensitivity classification for Illinois using depth to uppermost aquifer material and aquifer thickness: Illinois State Geological Survey Circular 560, 14 p.

Follmer, L.R., McKay, E.D., Lineback, J.A., and Gross, D.L., 1979, Wisconsin, Sangamonian, and Illinoian stratigraphy in central Illinois: Illinois State Geological Survey Guidebook 13, 139 p.

Hansel, A.K., and Johnson, W.H, 1996, Wedron and Mason Groups: lithostratigraphic reclassification of deposits of the Wisconsin Episode, Lake Michigan Lobe area: Illinois State Geological Survey Bulletin 104, 116 p.

Herzog, B.L., Wilson, S.D., Larson, D.R., Smith, E.C., Larson, T.H., and Greenslate, M.L., 1995, Hydrogeology and groundwater availability in southwest McLean and southwest Tazewell Counties, Part 1: Aquifer characterization: Illinois State Geological Survey and Illinois State Water Survey Cooperative Groundwater Report 17, 70 p.

Lineback, J.A., 1979, Quaternary deposits of Illinois: Illinois State Geological Survey, map, scale 1:500,000.

McGarry, C.S., and Grimley, D.A., 1997, Maps of Carroll County, Illinois: Illinois State Geological Survey Open File Series 1997-13 a through k (11 maps), scale 1:62,500.

Riggs, M.H., McGarry, C.S., and Medlin, S., 2000, Maps of Jo Daviess County, Illinois: Illinois State Geological Survey Open File Series 2000-8 a through j (10 maps), scale 1:62,500.

Teater, W.M., 1996, Soil survey of Tazewell County, Illinois: U.S. Department of Agriculture, Natural Resources Conservation Service, 210 p. and maps.

Wilson, S.D., Kempton, J.P., and Lott, R.B., 1994, The Sankoty-Mahomet Aquifer in the confluence area of the Mackinaw and Mahomet Bedrock Valleys, central Illinois: Illinois State Water Survey and Illinois State Geological Survey Co-operative Groundwater Report 16, 64 p.

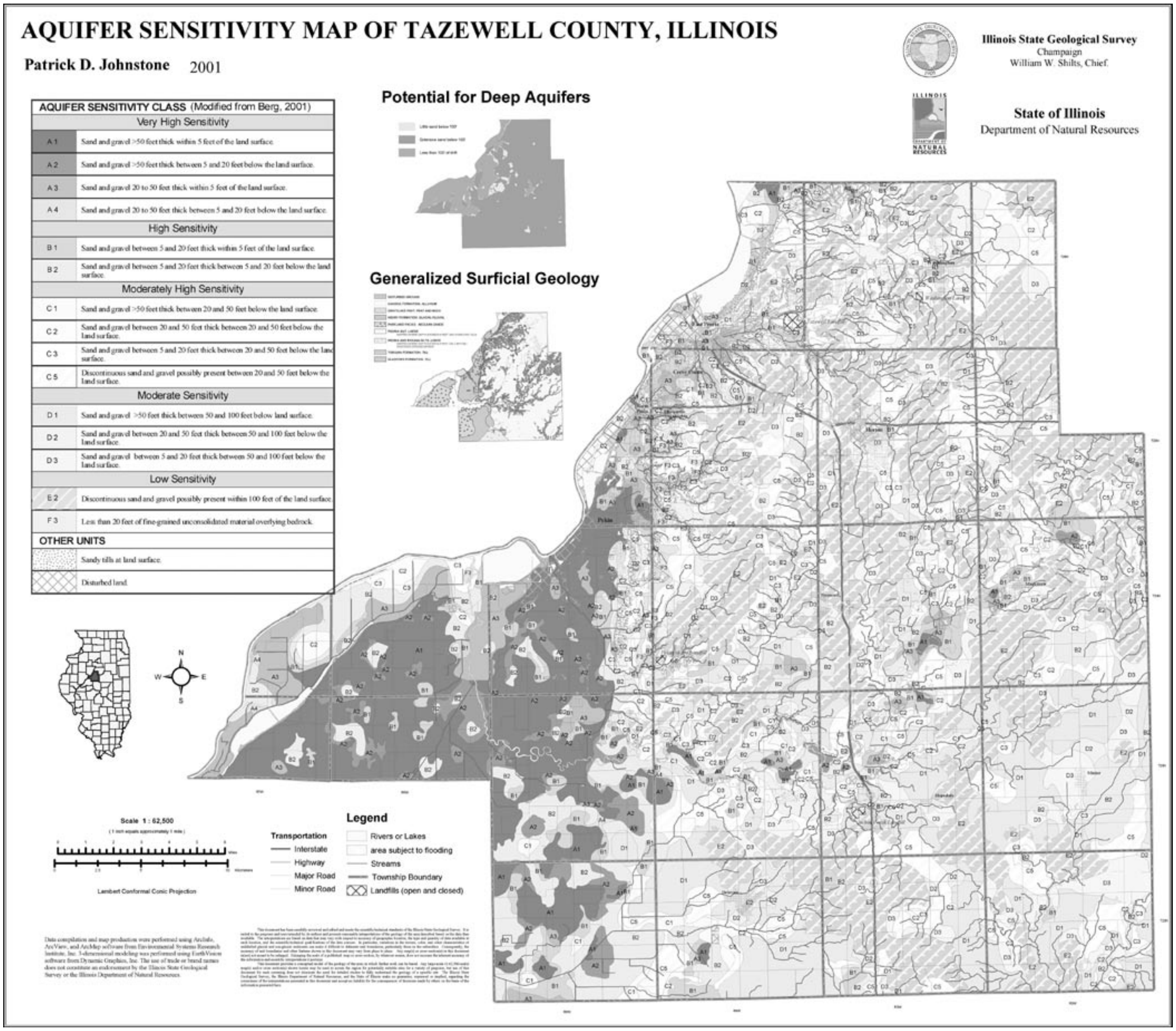


Figure 1. Aquifer Sensitivity Map of Tazewell County, Illinois, one of nine maps created for Tazewell County by the ISGS. Actual map was delivered at 1:62,500 scale in full color.

Visualizing the Uncertainty of Geologic Maps

By Patrick J. Kennelly

Montana Bureau of Mines and Geology
Montana Tech of the University of Montana
Butte, MT 59701-8997
Telephone: (406) 496-2986
Fax: (406) 496-4451
e-mail: pkennelly@mtech.edu

ABSTRACT

For this study I quantified and displayed the uncertainty of geologic contacts located from U.S. Geological Survey (USGS) topographic base maps. I used Level 1 Digital Elevation Models (DEM's), which have a root mean square error (RMSE) of 7 meters or less. Assuming a normal distribution of error, a contact traced primarily from topographic contours has a 68.3% chance of being within ±7 vertical meters of the contact's actual location, a 95.5% chance of being within ±14 vertical meters, and a 99.7% chance of being within ±21 vertical meters. For this study I used the top and base of the Tullock Member of the Fort Union Formation in the eastern half of the U.S. Geological Survey Broadus 30 x 60 minute quadrangle in southeastern Montana as an example. The geologic contact was buffered in the z direction using the underlying digital elevation model data. Resulting areas vary in horizontal width and are represented by dark gray (1 x RMSE), medium gray (2 x RMSE), and light gray (3 x RMSE). This method could serve as a reconnaissance tool, assisting field geologists in better locating geologic contacts.

INTRODUCTION

Locational error is present in and impossible to eliminate from geologic maps, but can be visualized with geographic information system (GIS) technology. The probability of a geologic contact being precisely located is a function of such diverse factors as how easily the contact can be recognized in the field, the detail of the field work, the accuracy of the manual or digital data capture, and errors inherited from the topographic base map. Although some of these uncertainties are difficult to quantify, the horizontal and vertical errors inherent in topographic base maps available from the U.S. Geological Survey (USGS) are well documented (USGS, 1998, 2000).

UNCERTAINTY OF TOPOGRAPHIC MAPS

The United States National Map Accuracy Standards define horizontal and vertical accuracy by the associated error. To meet the horizontal accuracy standard, 90% or more of well defined test points must not be in error by more than 1/30 inch for maps printed at scales less detailed than 1:20,000. Well-defined test points include benchmarks or perpendicular road intersections, but not geologic contacts. To meet the vertical accuracy standard, 90% or more of tested elevations must fall within one-half of the contour interval. The vertical standard is appropriate for geologic contacts created from contour patterns augmented with some known points.

To measure topographic uncertainty, the USGS generally uses 28 test points to determine the vertical accuracy of a 7.5-minute sheet. From these test points, a root mean square error (RMSE) can be calculated using the following equation:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} d_i^2}{n}}$$

where $d_i = Z_{estimated} - Z_{observed}$ and n is the number of sample points. Given this RMSE, the USGS has assigned all DEM's to one of three categories. Level 1 DEM's are derived using photogrammetric techniques, and contain the least vertical error. They have a RMSE of 7 meters or less (Slocum, 1999).

The RMSE is a summary statistic for a map. It gives no indication of how error is distributed, statistically or spatially. Error curves of very different shapes (e.g., normal, skewed) could have the same RMSE. Additionally, one segment of a map could account for the majority of the error in the map (Shortridge, 2001). In this study, I made the simplifying assumptions that (1) the error has a normal or Gaussian distribution (Figure 1), and (2) this error is distributed across the map.

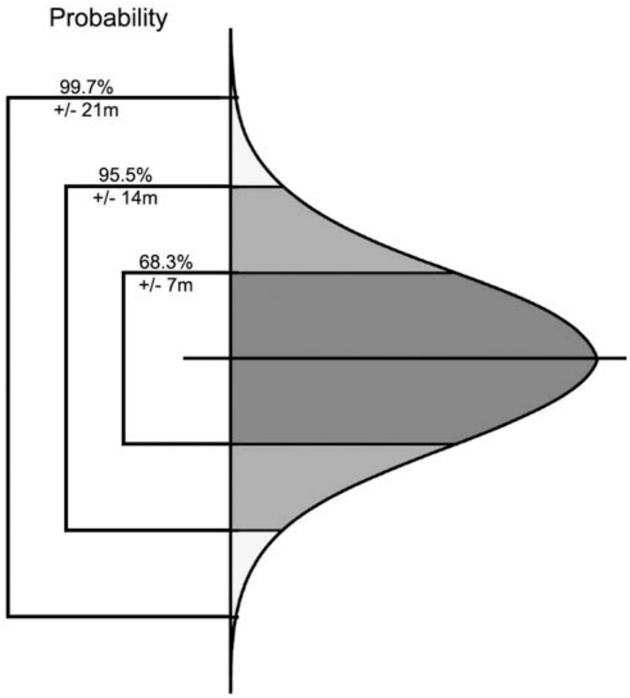


Figure 1. A normal or Gaussian distribution curve. Values within one standard deviation account for 68.3% of the values and are shown in dark gray. Values within two standard deviations account for 95.5% of the values and include the medium-gray areas. Values within three standard deviations account for 99.7% of the values and include the light-gray areas.

LOCATION AND GEOLOGY OF THE STUDY AREA

For this study I looked at the geologic contacts associated with the top and base of the Tullock Member of the Tertiary Fort Union Formation in the eastern half of the USGS Broadus 30 x 60 minute quadrangle (scale 1:100,000) in southeastern Montana (Figure 2). The Tullock is a planar-bedded light-yellow to light-brown sandstone interbedded with shale and mudstone. Its basal contact is the base of the lowest persistent coal bed. It is underlain by a cross-bedded yellowish-gray sandstone, the Cretaceous Hell Creek Formation. The Tullock is overlain by the Lebo Member of the Fort Union Formation, a gray, smectitic shale and mudstone containing lenses of gray and yellow sandstone (Figure 3). See Vuke and others (2001) for more detailed descriptions of these geologic units.

Geologic structure in this area is minimal and is characterized by low-angle bedding dipping to the northwest. The geologic contacts were constructed from a combination of field observations, regional correlations, and integration of previous mapping. In many cases contacts were interpolated between observed locations by the mapper,

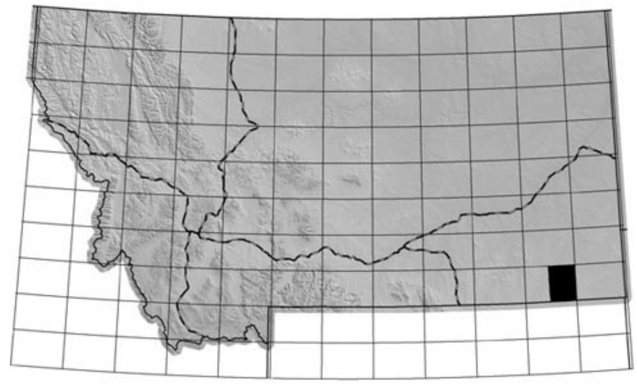


Figure 2. Location of the study area within the eastern half of the USGS Broadus, Montana, 30 x 60 minute quadrangle, southeastern Montana. The area is approximately 40 km by 55 km.

using knowledge of the structural orientation of the units and contours from the USGS topographic base maps.

The topography in the study area ranges from elevations of 900 to 1,350 meters (Figure 4). Slopes are as steep as 37°. The Little Powder River runs through the center of the study area, and associated Quaternary alluvium overlies the bedrock geology.

METHODOLOGY

Geologic contacts were converted to line themes and overlain on the 16 Level 1 DEM's (scale 1:24,000) that compose the eastern half of the Broadus quadrangle. I applied a vertical buffer using an ArcView extension developed by Damon Holzer (Department of Forest Service, Texas A & M University) and available on ESRI's Web page <http://arcscripts.esri.com>. The vertical buffer identified all grid cells within a given elevation of the contacts. Three vertical buffers were created to represent the first, second, and third standard deviations of error in the DEM's, at ±7, ±14, and ±21 vertical meters, respectively. Due to variations in slope, vertical buffers of equal interval define areas of variable horizontal width.

A horizontal distance of 200, 400, and 600 meters was defined as the maximum horizontal width associated with the vertical buffers of ±7, ±14, and ±21 meters, respectively. This definition prevented inclusion of areas that are unrealistic distances from interpreted contacts based on the geologist's assesment. The gentle dip of geologic formations was ignored in this analysis. At the maximum horizontal distance of 600 meters from the contact, dips of 5° offset horizontal locations by approximately 50 meters.

RESULTS

The resulting buffer is displayed on the geologic map in Figure 5 and the topographic map in Figure 6. Dark-

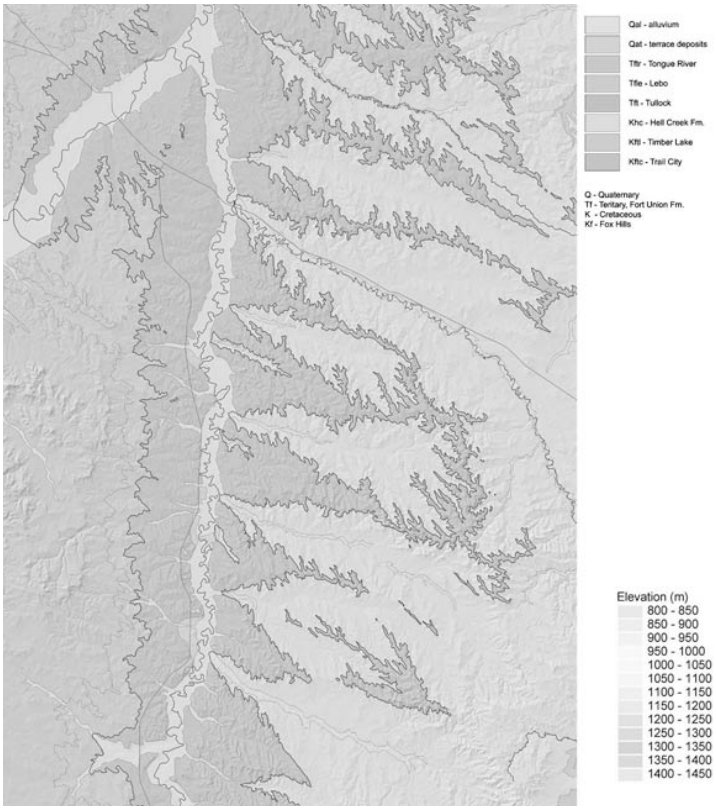


Figure 3. Geologic units of the eastern half of the USGS Broadus, Montana, 30 x 60 minute quadrangle.

Figure 4. Topography of the eastern half of the USGS Broadus 30 x 60 minute quadrangle, overlain with contacts of the geologic units from Figure 3.

gray pixels are those within 7 vertical meters; medium-gray pixels are between 7 and 14 vertical meters; and light-gray pixels are between 14 and 21 vertical meters. These shades of gray can be related to the probability, assuming a normal distribution (Figure 1). The probability of the location of the contact being within the dark-gray area is 68.3%, within the red and medium-gray area is 95.5%, and within the dark-gray, medium-gray and light-gray area is 99.7%.

The horizontal width of the resulting buffer is highly variable. For the third standard deviation (dark gray, medium gray, and light gray), the width ranges from 90 to 1,200 meters (using the arbitrary maximum horizontal distance of 600 meters on each side of the contact). The width is greater in gently sloping areas and lesser in steeply sloping areas. The effect of slope on the probability of properly locating the contact in cross-sectional and map views is illustrated in Figure 7, and shown in perspective view for a part of the geologic map in Figure 8.

DISCUSSION

The method of local, vertical buffering seems most appropriate for geologic maps because it allows interpreted contacts to be used to their fullest extent. It is similar to a method devised by Hunter and Goodchild (1995) to determine the probability of an area being above a threshold elevation. Their study, however, used only one elevation value; this study allowed the elevation of the contact to vary.

Several alternative methods could be used to create a vertical buffer for geologic contacts. If the geologic formations were horizontal, the buffers could be displayed directly from the DEM. This is not the case in this study area, where elevations of the contact range from 950 to 1,250 meters. Alternatively, elevation values could be assigned to the top and basal contacts, and then a surface could be fit to the lines in three dimensions. A best-fit planar analysis in three dimensions would be equivalent to a linear regression of point data in two dimensions; then the surface could be offset by the RMSE z-values and intersected with the DEM. This method has two disadvantages. First, all segments of contact lines need not fall directly on the surfaces. Thus, there could be local variations not interpreted by the geologist. Second, new areas could be added in regions where the geologist has interpreted the contact as not present. In fact, the latter is the case in the southeastern corner of the study area. The Tertiary Tullock Member was determined to be absent on the basis of pollen samples, although the known topogra-

Figure 5. Probability of the geologic contact being within the dark-gray, medium-gray, and light-gray areas displayed on the *geologic* map; compare with Figure 1. The areas were defined using a vertical buffer.

Figure 6. Probability of the geologic contact being within the dark-gray, medium-gray, and light-gray areas displayed on the *topographic* map; compare with Figure 1. The areas were defined using a vertical buffer.

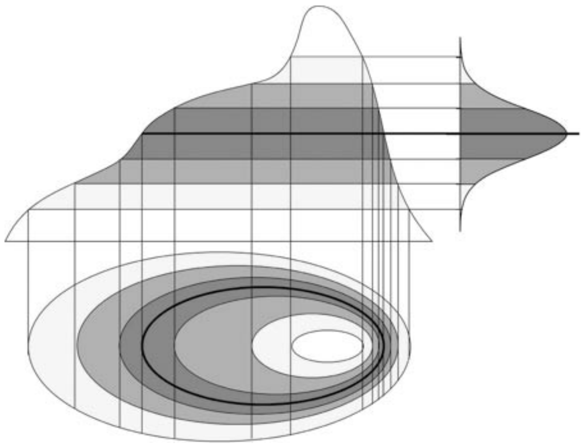


Figure 7. Normal distribution of error is shown on a cross-sectional view and map view of a sample location. Vertical widths are constant, but horizontal widths vary with the slope of the topography.

phy and geologic structure would indicate that it should be present. The geologic interpretation is that structural dip increases in the southeastern part of the study area.

Areas inappropriate for this type of analysis include those where bedrock geologic contacts are overlain by alluvium. At these locales, the geologist has interpreted the location of the contact beneath the alluvium. Thus, the surface elevation, which now represents the top of the alluvium, is no longer appropriate. In areas where the contact is overlain by alluvium, such as in the stream valley in the northwestern part of the map, an evenly spaced concentric buffer pattern is apparent in Figure 5.

The assumption that errors in elevation follow a normal curve is critical to the validity of this method. If error is random, the distribution should be normal (Wise, 1998). Systematic error, however, would not have a normal distribution. One such sampling error in the construction of DEM's is referred to as the "Firth Effect" and results in north-south or east-west lineations. This effect is caused by operators of photogrammetric equipment sampling row by row or column by column in alternating directions, and consistently underestimating elevation when moving upslope and overestimating elevation when moving downslope (Hunter and Goodchild, 1995). Visual inspection of hillshaded DEM's clearly reveal this striped pattern. These DEM's are generally not Level 1 and should be avoided for analysis.

CONCLUSIONS

The uncertainty inherited by geologic contacts from USGS topographic base maps can be quantified and visualized. The method described in this study assumes a normal or Gaussian distribution of measured error and

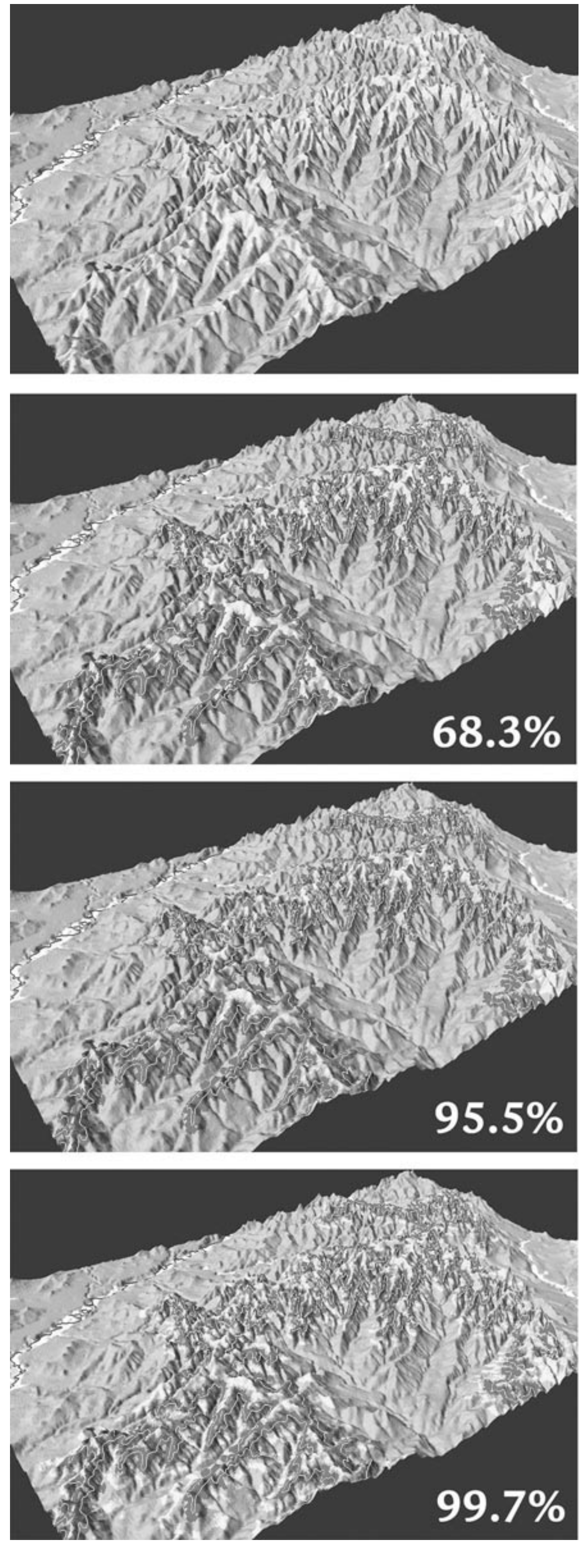


Figure 8. Perspective view of the east-central part of the geologic map from Figure 3. The Tullock Member overlies the Hell Creek Formation. Each frame adds a gray buffer to represent an additional standard deviation of locational error.

applies vertical buffers to the contacts based on these measures. These vertical buffers represent one, two, and three standard deviations based on a measured RMSE. The map width of these buffers is variable.

This GIS method could be a reconnaissance tool used by geologists to determine the areas requiring more detailed field inspection. The method would be especially useful in larger scale mapping, such as 1:24,000 scale.

REFERENCES

Brown, J.L., 1993, Sedimentology and depositional history of the lower Paleocene Tullock Member of the Fort Union Formation, Powder River Basin, Wyoming and Montana: U.S. Geological Survey Bulletin 1917-L, 42 p.

Hunter, G.J., and Goodchild, M.F., 1995, Dealing with error in spatial databases: a simple case study: Photogrammetric Engineering and Remote Sensing, v. 61, p. 529–537.

Shortridge, A., 2001, Characterizing uncertainty in digital elevation models, *in* Hunsaker, C.T., Goodchild, M.F., Friedl, M.A., and Case, T.J., eds., Spatial uncertainty in ecology: implications for remote sensing and GIS applications: New York, Springer-Verlag, p. 238-257.

Slocum, T., 1999, Thematic cartography and visualization: Upper Saddle River, N. J: Prentice-Hall, Inc., 293 p.

U.S. Geological Survey, 2000, U.S. GeoData digital elevation models: U.S. Department of the Interior Fact Sheet 040-00: Washington, D.C., <http://mac.usgs.gov/mac/isb/pubs/factsheets/fs04000.html>.

U.S. Geological Survey, 1998, National Mapping Program technical instructions: standards for digital elevation models: U.S. Department of the Interior, <http://rockyweb.cr.usgs.gov/nmpstds/demstds.html>.

Vuke, S.M., Heffern, E.L., Bergantino, R.N., and Colton, R.B., 2001, Geologic map of the Broadus 30- x 60-minute quadrangle, eastern Montana: Montana Bureau of Mines and Geology Open File Report #432, scale 1:100,000.

Wise, S.M., 1998, The effect of GIS interpolation errors on the use of digital elevation models in geomorphology, *in* Lane, S.N., Richards, K.S., and Chandler, J.H., eds., Landform monitoring, modeling and analysis: Chichester, England, John Wiley and Sons, p. 139-164.

From Paper to Digital: A Geologic Map's Odyssey

By Stephanie A. O'Meara,[1] Tim Connors,[2] Trista L. Thornberry,[1]
Joe D. Gregson,[3] Anne R. Poole,[4] and Bruce Heise[2]

[1]Colorado State University/National Park Service Cooperator
1201 Oak Ridge Drive, Suite 200
Fort Collins, CO 80525
Telephone: (970) 225-3584
e-mail: Stephanie_O'Meara@partner.nps.gov;
Tthorn@cnr.colostate.edu

[2]National Park Service–Geologic Resources Division
12795 West Alameda Parkway
Lakewood, CO 80228
Telephone: (303) 969-2093
e-mail: Tim_Connors@nps.gov;
Bruce_Heise@nps.gov

[3]National Park Service–Natural Resources Information Division
1201 Oak Ridge Parkway, Suite 200
Fort Collins, CO 80525
Telephone: (970) 225-3559
e-mail: Joe_Gregson@nps.gov

[4] National Park Service–Intermountain Region
Intermountain GIS Support
Bandelier Hall West, Room 203
Albuquerque, NM 87131
Telephone: (505) 346-2885 extension 260
e-mail: Anne_Poole@nps.gov

INTRODUCTION

Bedrock and surficial geologic maps and supporting information provide the foundation for studies of ecosystems, Earth history, ground water, geomorphology, soils, and environmental hazards such as fire history, landslide and rockfall potential, etc. Geologic maps describe the underlying physical conditions of many natural systems and are an integral component of the physical science inventories stipulated by the National Park Service (NPS) in its Natural Resources Inventory and Monitoring Guideline. The NPS has identified GIS and digital cartographic products as fundamental resource-management tools. There are few geologists employed at parks; thus, these tools are particularly important to the NPS to aid resource managers in using geologic data for park management decisions.

WHY DIGITAL?

Digital geologic maps have several advantages over paper geologic maps. Digital geologic maps can be used in a digital GIS environment where they can be integrated with other geospatial data to provide analysis of spatial relationships. A GIS provides quick, reproducible, precise analytic results. Digital data are also more easily shared and transferred between users. With digital attribute capability, a digital geologic map becomes a powerful database. The NPS is in the midst of producing other digital data sets for soils, vegetation, species distribution, and hydrology. These themes will be used in conjunction with the digital geologic themes to better promote integrated science in the national parks.

THE ODYSSEY FROM PAPER TO DIGITAL

One of the unresolved issues facing developers of digital geologic maps and data models is how to include map-unit descriptions, supplemental explanatory text (references and map notes), geologic cross sections, and the variety of other printed information that occur on published maps. The overarching development goal of the NPS Inventory and Monitoring Program is to produce digital products that are immediately useful to anyone familiar with their analog counterparts. For geologic maps, this means that the map-unit legend must be sorted and shaded appropriately by geologic age and that all textual, graphical, and other information from the published maps must be available interactively to the user. In short, the digital product must "look and feel" like its published source.

The NPS is developing most digital products in ESRI (Environmental Systems Research Institute) ArcView GIS software. ArcView interfaces effectively with other software running on the Microsoft Windows operating system. Also, integrating a variety of tools, including the NPS GIS Theme Manager, Windows Help software, a Microsoft Visual Basic graphics viewer program, and the ArcView legend editor, has allowed users to display geologic map information in a GIS.

COMPLETING THE ODYSSEY

The text below details the steps in transforming a paper geologic map to a user-friendly digital geologic map and database. First, the paper geologic map is scanned and the resulting image is georeferenced, providing a background for the digitization (capture) of geologic features. In accordance with the NPS Geology-GIS Data Model (O'Meara and others, 2002), the spatial and geologic feature types present (i.e., polygon, line, point, fault, fold, unit, etc.) are captured into appropriate GIS coverages and attributed according to the data model. These data are then incorporated into the NPS GIS Theme Manager that facilitates (in ArcView 3.3) the presentation of the various map coverages along with any Federal Geographic Data Committee (FDGC) metadata and accompanying help files that display map notes, unit descriptions, and other ancillary data from the original paper source map. Any map graphics (e.g., geologic cross sections) are scanned from the original paper map and "hot-linked" to a coverage (e.g., the cross-section line coverage) on the digital geologic map. These data are then posted on the NPS Inventory and Monitoring Program GIS FTP Web site for user access and download.

It All Starts with a Paper Map . . .

An existing traditionally printed map is the starting point in the odyssey (Figure 1).

Digitizing Geologic Map Features

Using ArcView 3.3 and a georeferenced image of the source map, linear and point geologic features are captured (Figure 2). These digitized features are attributed according to the existing NPS Geology-GIS Data Model coverages.

Attributing Geologic Coverages in ArcView 3.3 Using the NPS Geology-GIS Data Model

Geologic features are attributed according to the Data Model to preserve all components of the original source map. Figure 3 shows the geologic-units (polygon) area coverage being attributed. Geologic-unit data captured include the geologic-unit symbol (e.g., Ppc), its age relative to other map units, source map ID, and a help file link. Geologic units and contacts, faults, folds, attitude points, mine-related features, cross-section lines, and linear joints are all coverages derived from the Dinosaur Quarry quadrangle map. At present, there are 22 NPS Geology-GIS Data Model coverages.

Dinosaur Quarry Quadrangle Geologic Map and Cross-Section A-A'

A finished map is displayed in ArcView 3.3 using the NPS GIS Theme Manager (Figure 4), which organizes and presents GIS coverages complete with titles, legend files, and links to metadata, help files, and graphics. An example of such a link is shown in the cross-section graphic (Figure 5).

Dinosaur Quarry Quadrangle Metadata

Using the NPS GIS Theme Manager in ArcView 3.3, a user can display FGDC-compliant metadata related to each geologic coverage (Figure 6).

Dinosaur National Monument Geologic Map Help File

Contained within many geologic coverages are links to a geologic-map help file. Using the NPS GIS Theme Manager allows a user to display the help file in an interactive way within ArcView 3.3, complete with keyword searchability (Figure 7). The help file contains map-unit descriptions as well as other information, including reference information about the source maps used to compile the digital geologic map references, map notes, correlation of map units, and other ancillary information shown on the original paper map.

National Park Service Inventory and Monitoring GIS FTP Web Site

Digital geologic data and maps produced by the NPS are available online at <http://www3.nature.nps.gov/im/gis/ftp/ftparchive.cfm>. Data can be searched by park(s) and specified keywords. Data are available in ArcInfo 8.1 "E00" coverages or ArcView 3.3 shape (.SHP) formats and include associated FGDC-compliant metadata, help files, cross-section graphics, ArcView 3.3 (.AVL) legend files, and NPS GIS Theme Manager geology theme lists. The Web site lists the park(s) of interest, data category (e.g., geology), the data set or coverage title, the download file name, the data set or coverage file name, metadata file link, source of the data, and download file size (Figure 8).

Geologic GIS data are available at the NPS GIS FTP Web site for the following NPS units:

Arches National Park, Utah (ARCH),
Badlands National Park, South Dakota (BADL),
Bent's Old Fort National Historic Site, Colorado (BEOL),
Black Canyon National Park, Colorado (BLCA),
Bryce Canyon National Park, Utah (BRCA),
Capitol Reef National Park, Utah (CARE),
Colorado National Monument, Colorado (COLM),
Curecanti National Recreation Area, Colorado (CURE),
Dinosaur National Monument, Colorado and Utah (DINO),
Florissant Fossil Beds National Monument, Colorado (FLFO),
Great Sand Dunes National Park, Colorado (GRSA),
Hovenweep National Monument, Colorado and Utah (HOVE),
Mesa Verde National Park, Colorado (MEVE),
Natural Bridges National Monument, Utah (NABR),
Pipe Spring National Monument, Arizona (PISP),
Rocky Mountain National Park, Colorado (ROMO),
Saguaro National Park, Arizona (SAGU), and
Timpanogos Cave National Monument, Utah (TICA).

Geology of the Dinosaur Quarry Quadrangle

Figure 9 shows the geology of the Dinosaur Quarry quadrangle with hillshading added.

REFERENCE

O'Meara, S.A., Gregson, J.D., and Poole, A.R., 2002, National Park Service Geology-GIS Data Model, <http://www1.nature.nps.gov/im/gis/GeologyGISDataModel.htm>.

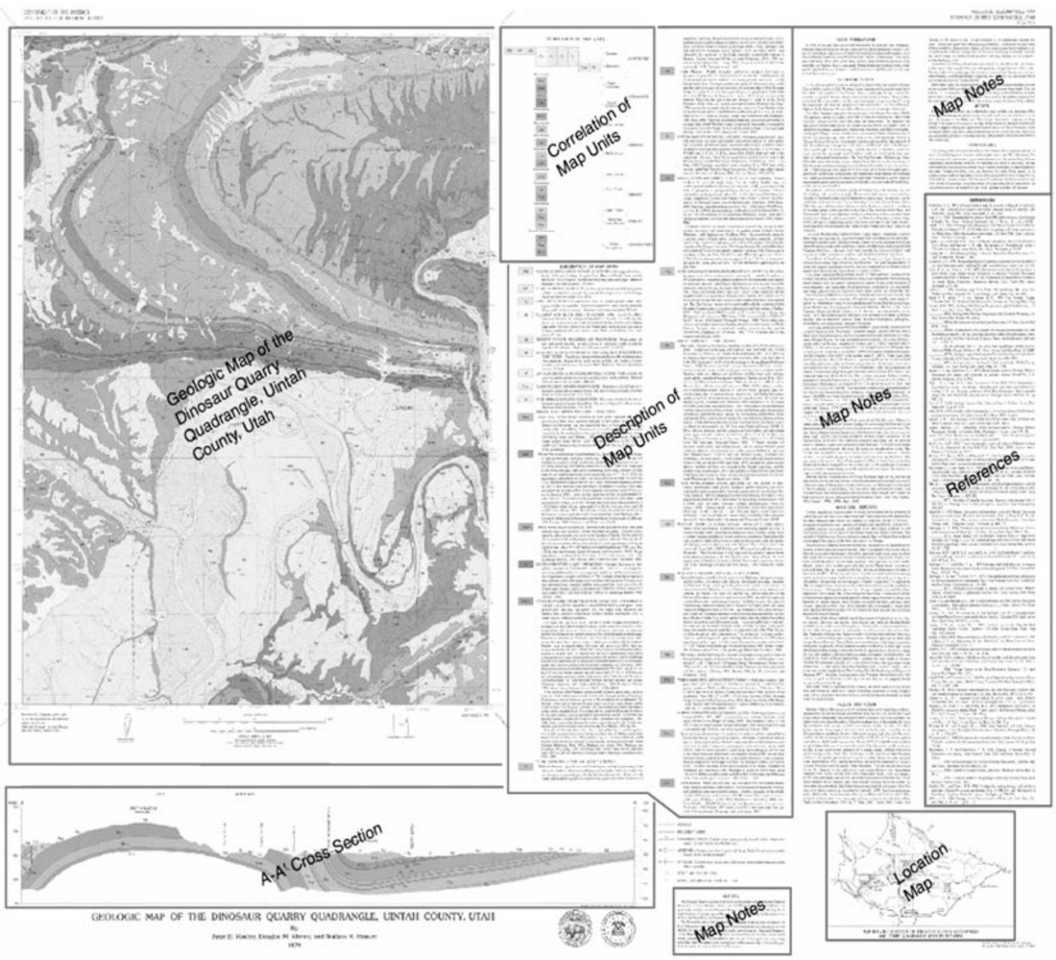


Figure 1. Image of the paper geologic map of Dinosaur Quarry quadrangle. Map components are indicated with diagonal text.

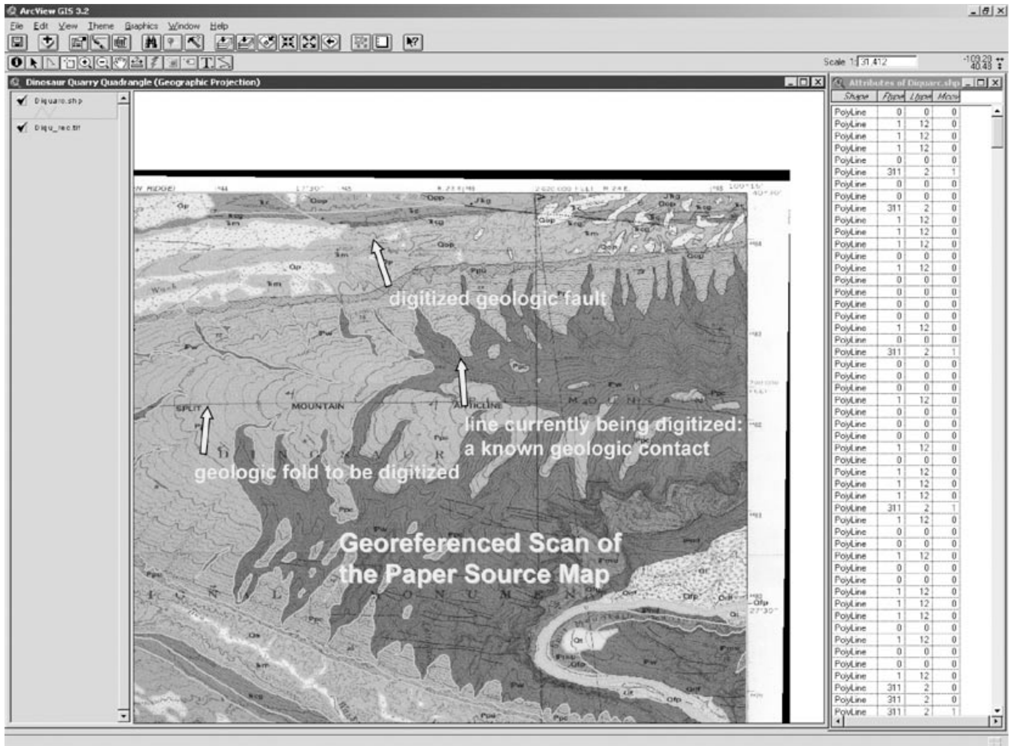


Figure 2. Image showing digitizing of a geologic map in ArcView 3.3 using a georeferenced scan of the source paper geologic map. Attribute table with codes indicates (1) type of line feature (FTYPE): geologic contact, anticlinal fold, normal fault, etc., (2) accuracy of line position/concealment (LTYPE): known or certain, approximated, concealed, etc., (3) whether the line is present in multiple coverages (MCOV; e.g., a contact between different geologic units that is also a fault).

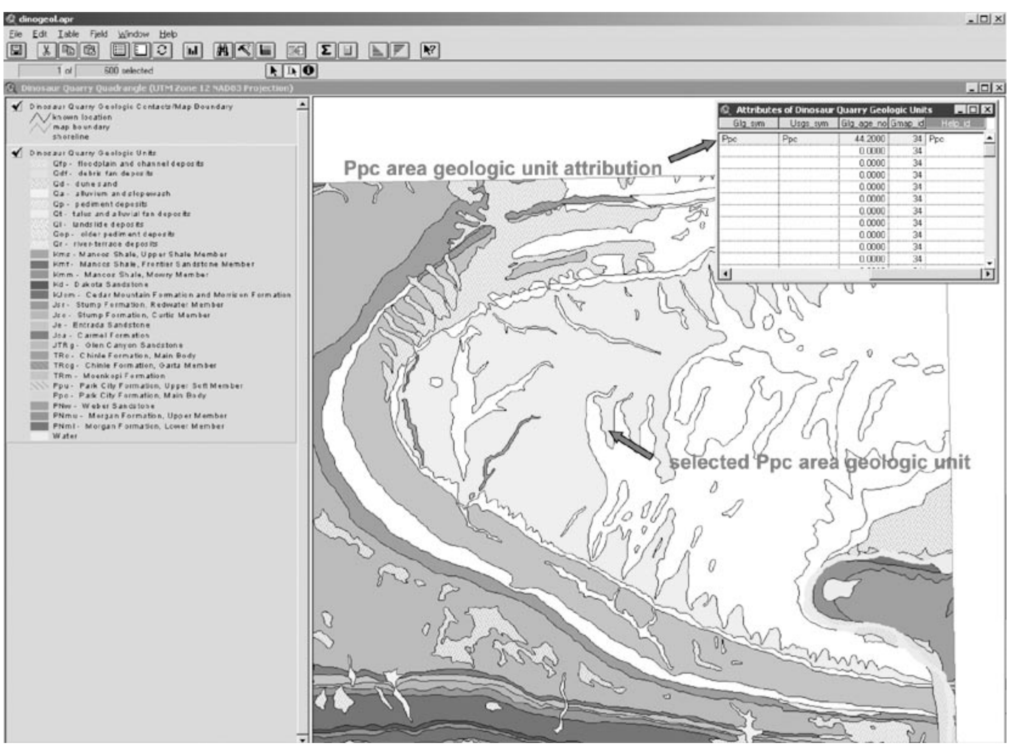


Figure 3. Image showing NPS Geology-GIS Data Model attribution of an area (polygon) geology coverage using ArcView 3.3.

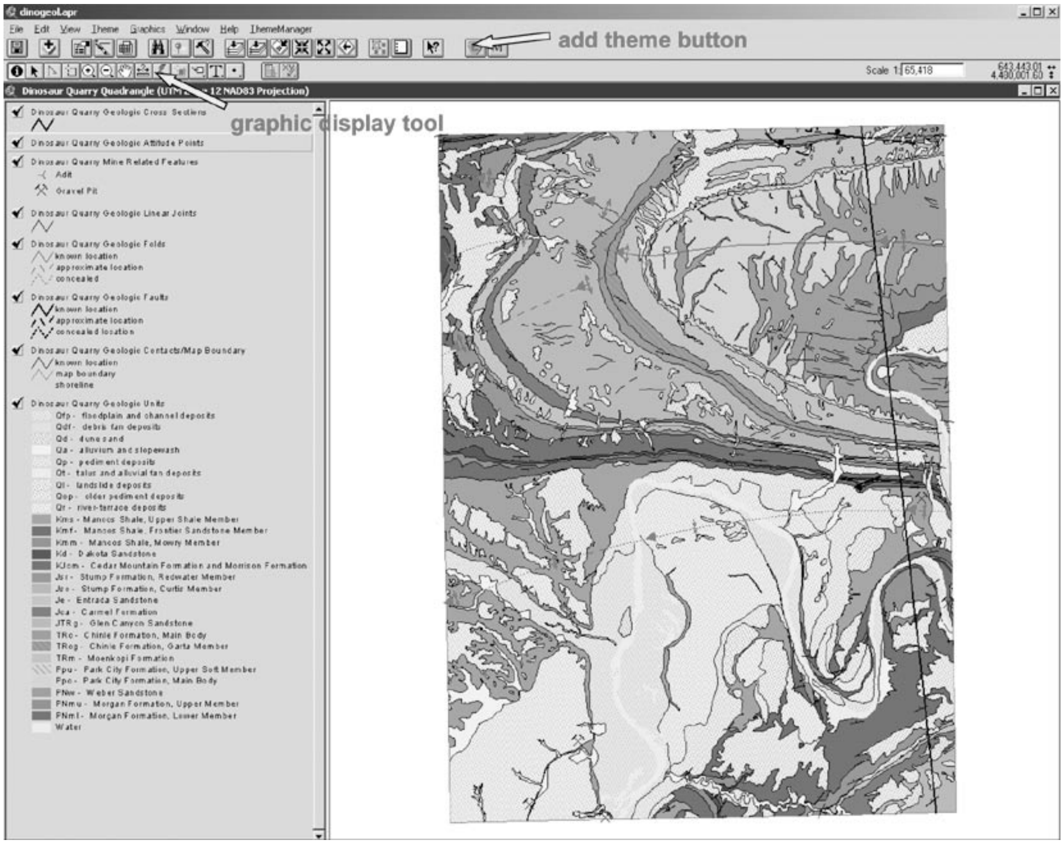


Figure 4. Image showing completed digital geologic GIS data in ArcView 3.3.

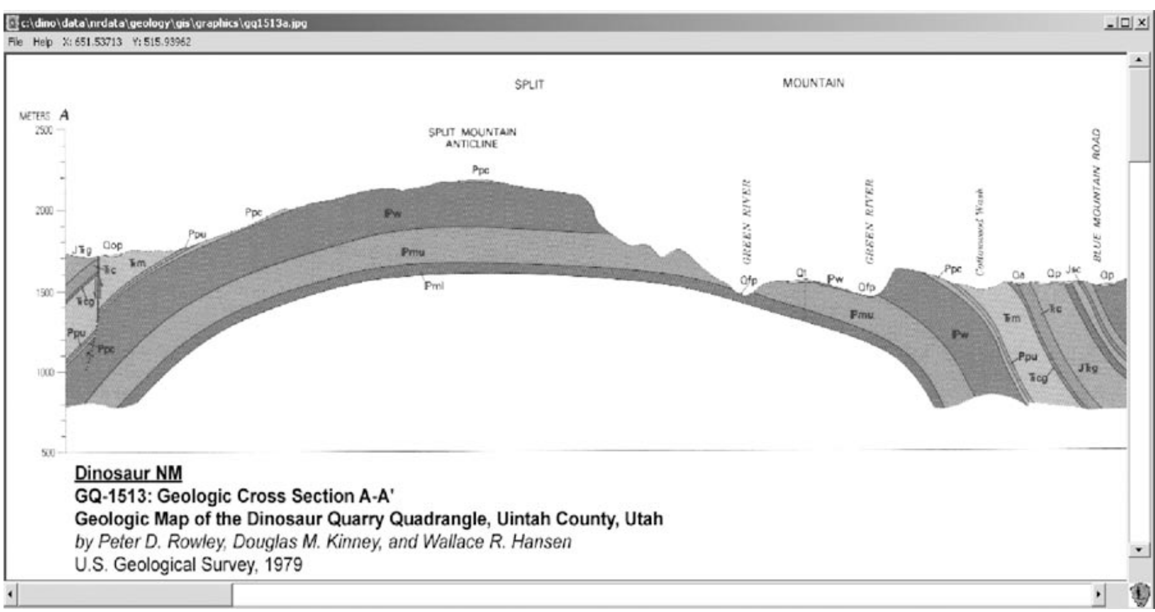


Figure 5. Image showing presentation of geologic cross section A-A' in ArcView 3.3 using the NPS GIS Theme Manager.

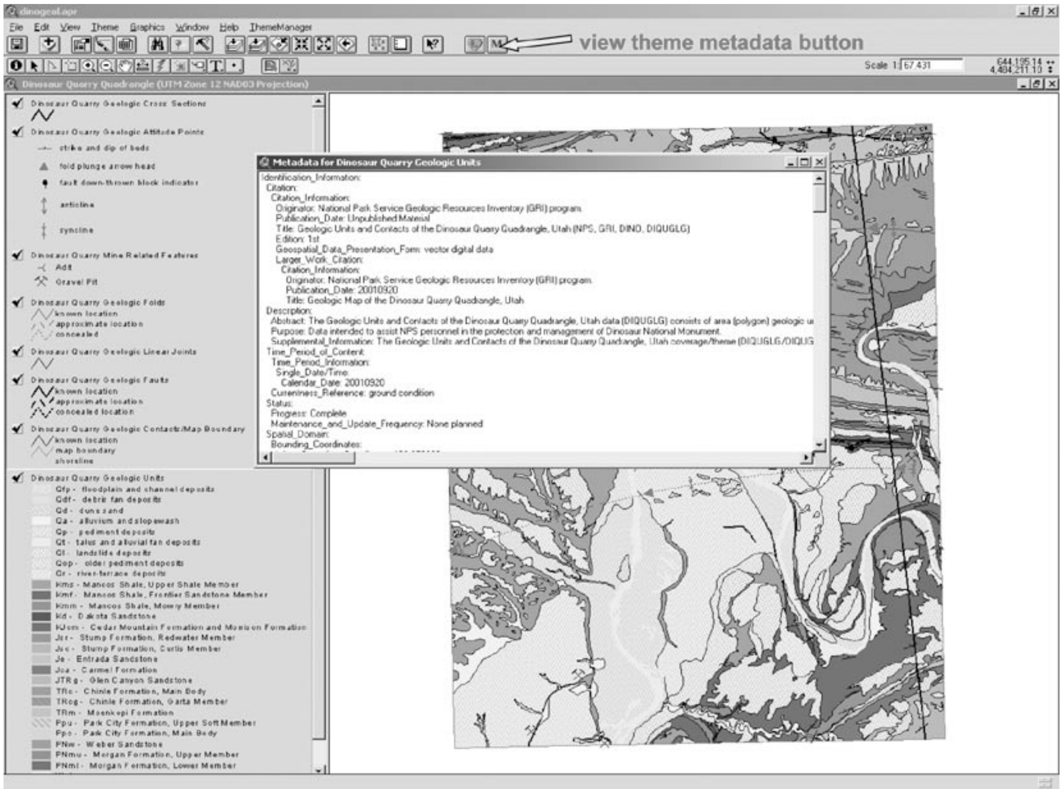


Figure 6. Image showing presentation of GIS data and FGDC metadata in ArcView 3.3 using the NPS GIS Theme Manager.

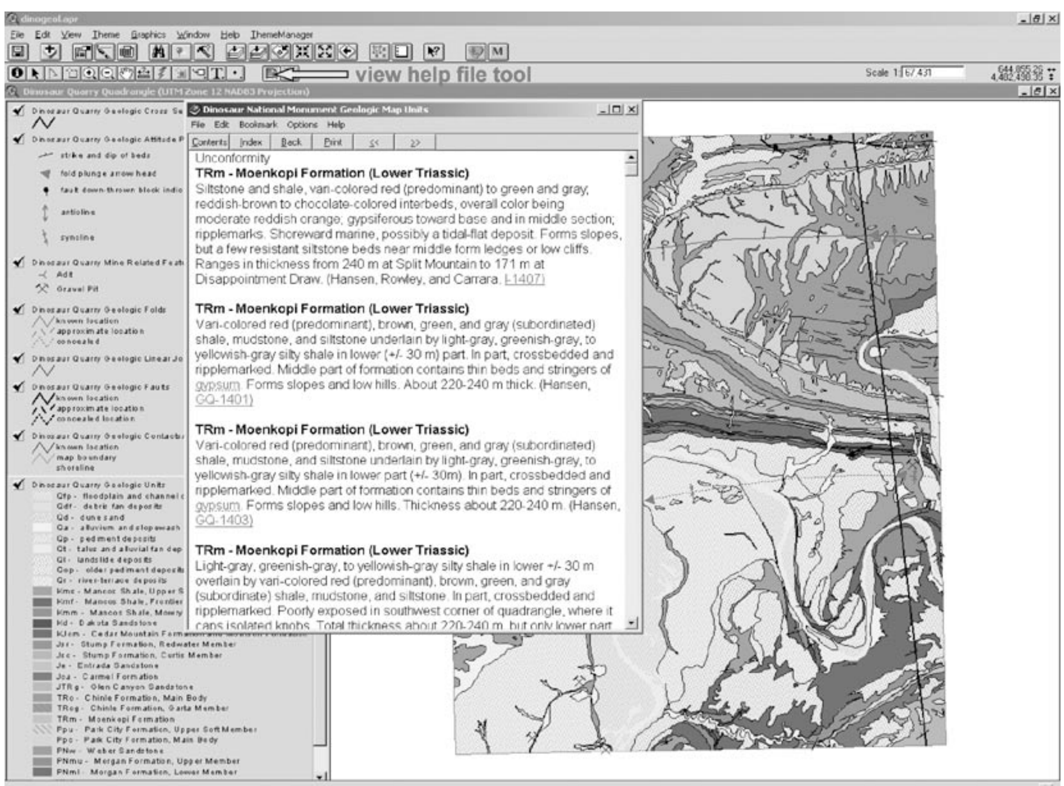


Figure 7. Image showing presentation of a geology help file in ArcView 3.3 using the NPS GIS Theme Manager.

Figure 8. Image showing available geologic GIS data available at the NPS GIS FTP Web site for the Dinosaur Quarry quadrangle map.

Figure 9. Image showing the geology of the Dinosaur Quarry (part of Dinosaur National Monument) quadrangle draped over a hillshade to depict the interrelationships of geology and geomorphology.

Cartographic Observations from Production of the Surficial Geologic Map of Northern New Jersey

By James R. Estabrook, D. Paul Mathieux, and Will R. Stettner

U.S. Geological Survey
Eastern Publications Group
National Center, MS 903
12201 Sunrise Valley Drive
Reston, VA 20192
Telephone: (703) 648-4319
Fax: (703) 648-6138
e-mail: jestabro@usgs.gov

INTRODUCTION

The *Surficial geologic map of northern New Jersey* (Stone and others, 2002) is a cooperative endeavor between the New Jersey Geological Survey and the U.S. Geological Survey (USGS). It is the final in a series of four map products that show the bedrock and surficial geology of New Jersey at a scale of 1:100,000. New Jersey is the first state to have such coverage over its entire area. Field work for the northern surficial map was begun in 1982, and compilation was done between 1989 and 1995. This product consists of three large map sheets (40 x 56 inches) and a text pamphlet. Sheet 1 shows the map at a scale of 1:100,000, along with a correlation of map units and a list of map units; sheet 2 contains 11 supporting maps and figures; and sheet 3 displays 11 cross sections. The 41-page pamphlet contains a detailed description of map and subsurface units and presents stratotype sections for eight proposed lithostratigraphic units of Quaternary age in northern New Jersey. This map adjoins the *Surficial geologic map of central and southern New Jersey* (Newell and others, 2000). The cartographic production of that map was discussed in the Digital Mapping Techniques '01 Workshop Proceedings (Stettner and Koozmin, 2001). Many of the steps described therein were also applied to the production of this map.

Our purpose in writing this paper and presenting the map as a poster at DMT '02 is to discuss (1) how we assembled the base map, (2) author compilation, and (3) how we applied Adobe Illustrator software to cartographic production.

BASE MAP

The base map on which the geology is shown is from the USGS 1:100,000-scale series and consists of a mosaic of one entire quadrangle with parts of five adjoining quadrangles. Because the mosaic was constructed in the late 1980's, mosaicking was accomplished not by digital means but entirely by hand on a large light table; film positives of each quadrangle were trimmed of their marginal information, edge-joined to create a "seamless" base image, and carefully taped onto a clear-film carrier sheet. Actually, three mosaics were constructed in order to permit the printing of the base in three colors: blue for drainage, brown for topography, and gray for culture (roads and place names). To accomplish this, we obtained film positive "separates" of the drainage, topography, and culture for all six quadrangles—18 positives in all. The culture separates were mosaicked first; then, in turn, the drainage and topography separates were mosaicked, each on their own carrier sheet registered on top of the culture to ensure exact fit. Film positives were made from each mosaic, as well as a composite negative of all three mosaics; this negative was used to produce a greenline on which the authors began to compile their geology.

At the start of cartographic production, the film positives of the three mosaics were scanned at a resolution of 300 dpi as transparent line art for graphic processing in Adobe Illustrator. Because raster images by nature have a diminished legibility, we selected for the printing of the culture a solid gray ink, rather than black ink screened 50 percent biangle, to avoid further eroding the raster image. The 50-percent biangle screen normally is used in printing to reduce the intensity of base-map information by breaking up the image. It also diminishes the clarity of the image and therefore is not recommended for use in conjunction with a raster image. Using a solid gray ink achieves a subdued image without causing any loss of clarity. For drainage, we chose Pantone 300 blue rather than the conventional cyan. Pantone 300 blue is considerably darker than cyan, so the drainage remains legible where overprinted by geologic map units composed of high percentages of cyan.

At the completion of cartographic production, we generated printing negatives and made a cromalin proof to check the registration and colors and to see how well the base image stood up against the other map features. The USGS 1:100,000-scale quadrangle maps use Souvenir (serif) and Univers (sans serif) type fonts in various sizes. We observed on the proof that most sans-serif place names, even at a small size, were pretty legible, but that serif type, even in large place names, was difficult to read. The presence of dense road networks under many feature names contributed to the problem. We rescanned the culture mosaic at several higher thresholds of optical sensitivity, but at the same resolution (300 dpi), and eventually found a threshold that yielded clearer type and linework. The improved legibility of the serif type, especially in congested urban areas, was achieved by sacrificing the legibility of very fine lines, most noticeable in some secondary roads and route numbers in rural areas. Rescanning the culture required us to carefully monitor its registration with the drainage and topography.

AUTHOR COMPILATION

Showing features that resulted from glaciation of most of the map area added much to the intricacy of the map. The complexity of the map is reflected by the number of registered overlays the authors required in order to compile their data. In addition to the greenline showing contacts and map-unit labels, seven overlays were drafted, which contained the following data: (1) patterned-area outlines (seven categories); (2) outlines of areas of thin till cover; (3) artificial fill (solid color areas and patterned areas); (4) contours showing elevation of the bedrock surface relative to sea level; (5) drainage basin divides; (6) symbols; and (7) bedrock outcrops.

To construct a mill copy for editing, we made a greenline of the base map and photographically exposed the contacts and the other information from the seven overlays to this greenline, to show in black, red, or blue. The authors then used colored pencils to shade in the map units, with the intent of actually using the same color scheme they desired to see on the final printed map. The colored mill was useful as a general color guide; however, as there are more than 175 map units, we needed some vehicle to help us readily distinguish them on the mill copy because (1) colors had to be repeated for many units, and (2) the small size of most polygons prohibited the authors from labeling more than half of them. Adobe Illustrator software proved to be most helpful with the identification problem; it is discussed in the following section.

CARTOGRAPHIC PRODUCTION USING ADOBE ILLUSTRATOR

Until recently, when geologic maps (especially simpler ones) were undergoing review, our practice was to edit the author's mill copy, ask him to make any agreed-upon changes to the drafting on his greenline, and then import his linework into Adobe Illustrator to begin cartographic production. In an effort to streamline production, we are now importing all data sets (raster and vector) into Adobe Illustrator, separating data into individual layers, and making refinements and a preliminary layout, all prior to the technical edit. The edit and subsequent reviews are made on color plots run on a large-format Hewlett Packard 3000 DesignJet. Because of the ability in Adobe Illustrator to group geologic-feature layers in different combinations, we were able to make plots showing selected geology (specific layers) for a particular review and (or) edit. This gave us a flexibility we found to be indispensable. As a result of this option, we made and reviewed approximately 30 generations of plots during the course of producing the three map sheets.

Unidentified Polygons

The contacts, patterned-area outlines, thin-till-area outlines, and artificial fill were scribed in order to get scannable linework of uniform quality. We had a positive made from each scribecoat, then contracted with Geologic Data Systems, Inc. (GDS), of Denver, Colorado, to furnish a data set of each positive by scanning the positives and tagging the polygons represented on them. These data sets were created in AutoCAD. Once imported into Adobe Illustrator, they required extensive sorting and organizing. Initially, about 5,000 polygons (one-quarter of the map's total) could not be identified by GDS owing to complexity of the mill copy. Ultimately, a review plot was sent to the authors showing all unidentified polygons in red. Using the mill copy, related compilations, and their notes, the authors resolved the unidentified polygons.

Color Selection

Of the 175 map units, 40 represent meltwater sediments deposited in major glacial lakes, and 46 represent meltwater sediments deposited in small glacial lakes and ponds. For these two categories, limited segments of the spectrum—blue to bluish green to gray for the former, and tan to violet to purple for the latter—had to represent all 40 and 46 units, respectively. To map and correlate the meltwater deposits, the authors divided the map area into five geographic regions on the basis of physiography and watershed. Within each region, for each category of meltwater sediment, the youngest deposits were given the lightest color shades, and the oldest deposits the darkest shades. For both kinds of meltwater sediment, a few regions had as many as 14 units, so some repeating of colors was necessary. Within each meltwater-sediment category it was permissible to repeat colors from one region to another. The issue was to make sure no two adjoining units on the map had the same color. By having each map unit in its own layer, we were able to selectively view a unit on screen in combination with any other units whose color and proximity we wanted to verify. Adobe Illustrator provided a very efficient method for identifying color problems. In order to perform these operations, hardware with the maximum processing speed and RAM was essential. All digital manipulations were performed on an Apple Macintosh G4 computer having 500 MHz processing speed and 1.25-gigabyte RAM.

REFERENCES

Newell, W.L., Powars, D.S., Owens, J.P., Stanford, S.D., and Stone, B.D., 2000, Surficial geologic map of central and southern New Jersey: U.S. Geological Survey Miscellaneous Investigations Series Map I-2540-D, scale 1:100,000, 3 sheets and pamphlet.

Stettner, W.R., and Koozmin, E.D., 2001, Cartographic choices for the surficial geologic map of central and southern New Jersey, *in* Soller, D.R., ed., Digital Mapping Techniques '01—Workshop Proceedings: U.S. Geological Survey Open-File Report 01-223, p. 233-239.

Stone, B.D., Stanford, S.D., and Witte, R.W., 2002, Surficial geologic map of northern New Jersey: U.S. Geological Survey Miscellaneous Investigations Series Map I-2540-C, scale 1:100,000, 3 sheets and pamphlet.

Raster to Vector Conversion of Geologic Maps: Using R2V from Able Software Corporation

By Kent D. Brown

Utah Geological Survey
P.O. Box 146100
Salt Lake City, UT 84114-6100
Phone: (801) 537-3300
Fax: (801) 537-3400
e-mail: kentbrown@utah.gov

INTRODUCTION

Raster to vector conversion of geologic maps has proven to be a valuable asset in the creation of a GIS database. However, the ability of applications to vectorize scanned (raster) maps varies widely. Most vectorizing software is optimized for the archival of old engineering drawings, architectural plans, and historical documents, where most lines meet at right angles. Geologic maps do not vectorize well with many of these programs.

The Mapping Program of the Utah Geological Survey (UGS) has attempted to scan and vectorize geologic maps of various types with mixed results. The software we used made workers frustrated with the amount of time required to vectorize a map, and the editing tools were inefficient and not user friendly.

In an effort to minimize our table digitizing and increase efficiency and positional accuracy, we evaluated five commercially available vectorizing programs. There was no real scientific method to our evaluation. We downloaded trial versions of the software and vectorized the same map using each application. Although these applications are also capable of heads-up digitizing and interactive vectorizing (line following), the focus of this test was to autovectorize the entire map. This paper summarizes the autovectorization process using R2V from Able Software Corporation.

OVERVIEW OF METHODS

Recently the UGS vectorized 32 7.5-minute geologic quadrangles to compile a 30 x 60 minute quadrangle. The source maps were paper topographic base maps on which the author had inked lines and labels and polygons were hand colored in pencil. The maps were scanned and saved as 300-dpi, 24-bit, color JPEG files.

The JPEG files were opened in R2V (Figure 1), and a process was used to eliminate the colors, preserving just the black line features from the author's original mapping. Because the base-map features were colored over with pencil, the software did not interpret them as black and they could be eliminated almost entirely, leaving the inked contacts, faults, labels, and leader lines (Figure 2). Further image editing was done (Figure 3) to eliminate the leader lines, labels, and any trace of the base map that remained.

Next, the autovectorizing tool was used. This tool traces the pixels on the raster, so the resulting vectors have a stair-step appearance with too many vertices. The "smooth lines" command was used to eliminate many of the vertices that were redundant, making the lines look better. Because R2V supports layer creation, we chose to separate geologic contacts and faults onto different layers for easier data management. After manually editing the whole map, the vectors were suitable for use in our GIS (Figure 4).

The vectors were then georeferenced by placing control points on the corners of the quadrangle and assigning correct UTM coordinates to them. Georeferencing is the final step before the lines are exported as a DXF file. We chose DXF because the layers are preserved, saving much time in GIS by importing layers from the DXF and attributing the whole layer at once. If exported as a shapefile, all layers are combined into one and we find this to be less useful.

CONCLUSIONS

R2V is designed for GIS users who are making maps and includes many useful and easy-to-use tools for that purpose. Other programs were designed with engineering drawings as the focus and are less useful for maps. Using R2V, georeferencing is easy to perform and image warping and world file creation tools are included. R2V, a Windows 98/NT/2000/XP program, is designed for exporting vectorized data directly to ArcInfo generate (.gen), ArcView shapefile (.shp), DXF (.dxf), MapInfo (.mif), IGES (.igs), and a few other common formats. At the Utah Geological Survey, we found R2V to be a very useful and efficient tool for vectorizing geologic maps in preparation for GIS map production.

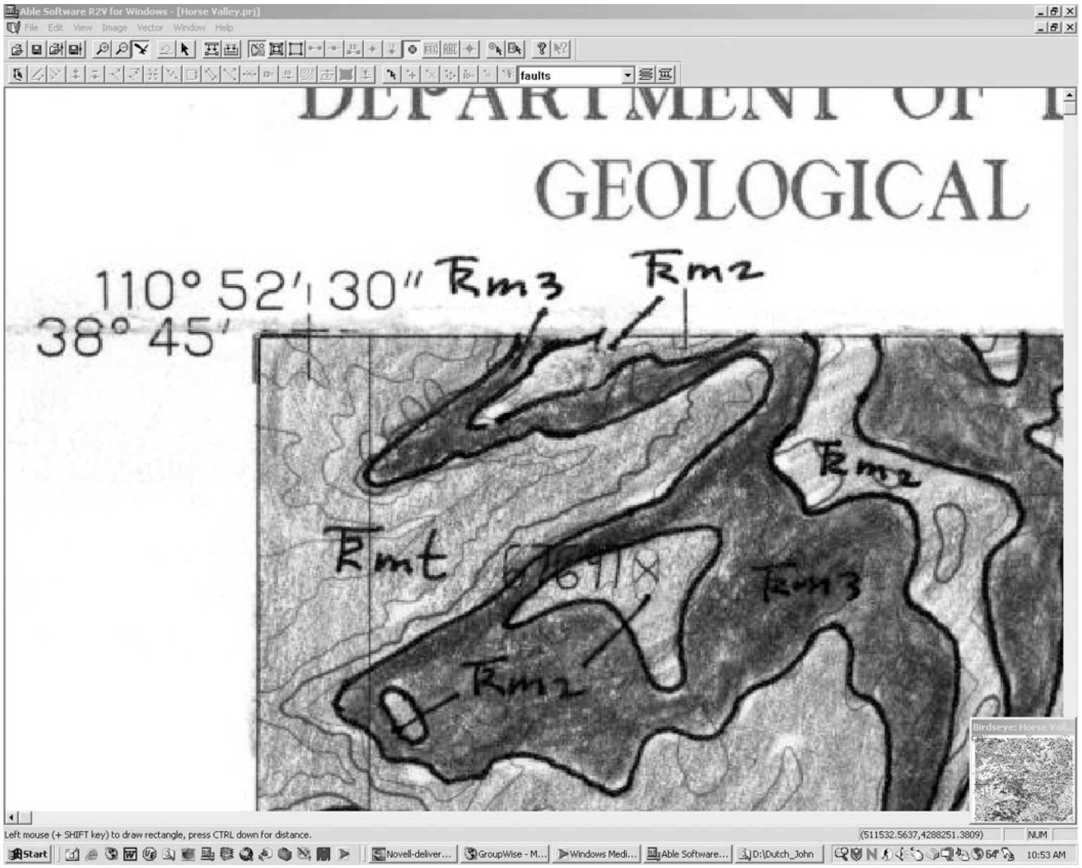


Figure 1. JPEG of original map opened in R2V.

Figure 2. Colors from the original have been removed.

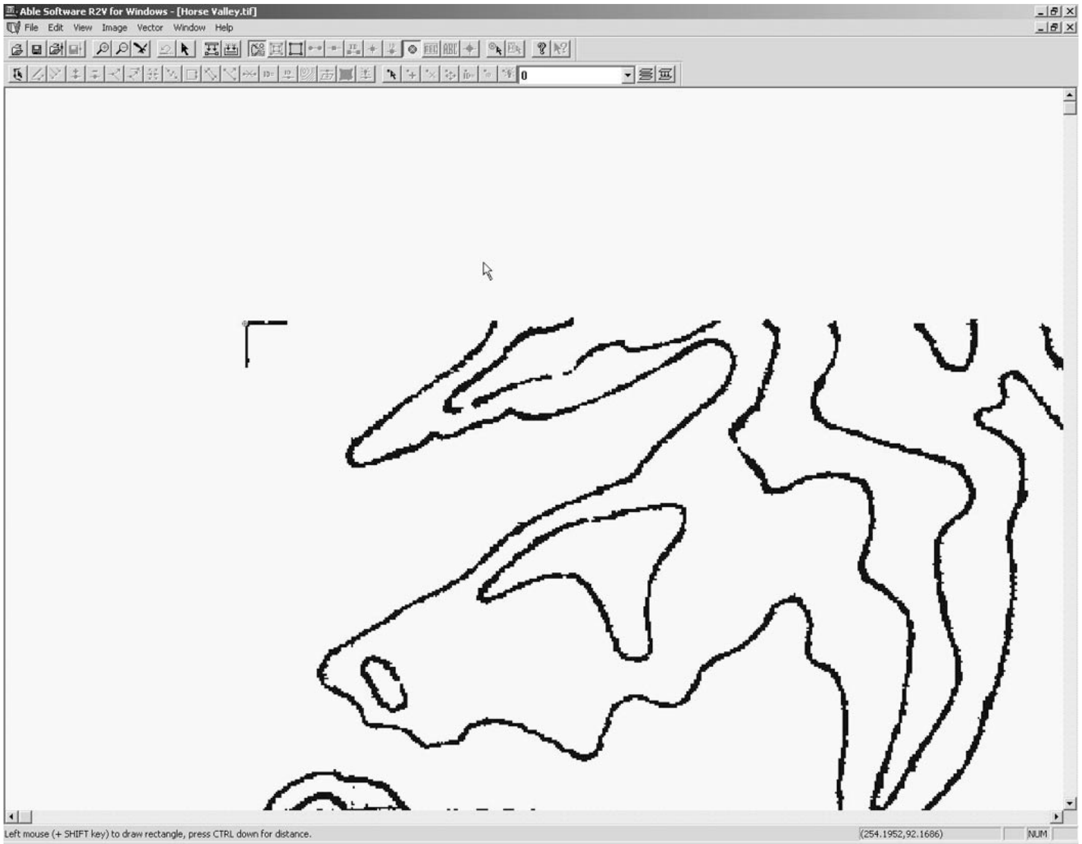


Figure 3. Only lines to be used for autovectorizing remain.

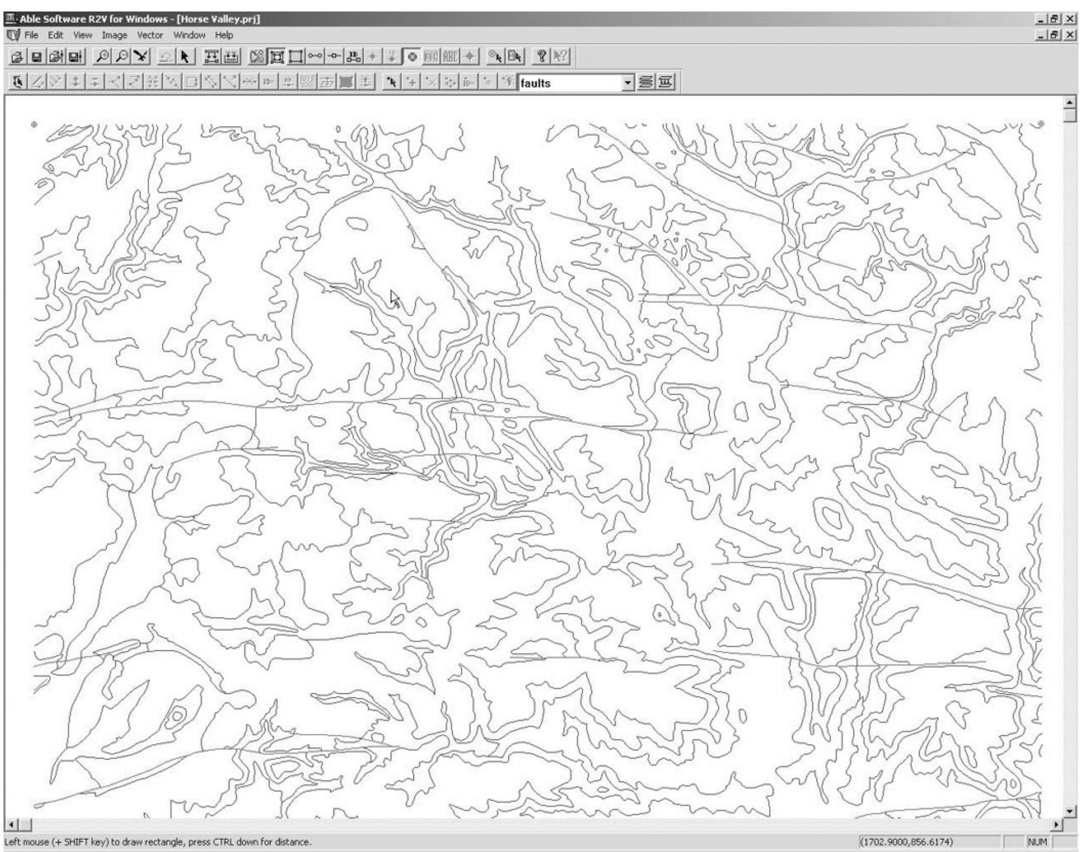


Figure 4. Vectors ready for export and use in GIS. Pixelation seen here is an artifact of the computer screen capture.

Ground-Water Quality Classification Using GIS Contouring Methods for Cedar Valley, Iron County, Utah

By Matt Butler, Janae Wallace, and Mike Lowe

Utah Geological Survey
P.O. Box 146100
Salt Lake City, UT 84114-6100
Telephone: (801)537-3391
Fax: (801)537-3400
e-mail: {mattbutler, janaewallace, mikelowe}@utah.gov

Cedar Valley, in southwestern Utah, is experiencing an increase in residential development on unconsolidated deposits of the basin-fill aquifer, the valley's primary source of drinking water. In response to local government's desire to protect the high quality of this resource, we used a geographic information system (GIS) to develop a ground-water quality classification map derived from a total-dissolved-solids (TDS) concentration map.

To delineate bedrock versus valley-fill material, we created a digital geologic coverage of rock units and unconsolidated deposits. Then we generated the TDS and the ground-water classification maps. Using GIS contouring methods within the ArcView 3x Spatial Analyst Extension, we created the TDS map using available TDS values (measured in mg/L) from sampled water wells. Total-dissolved-solids concentrations ranged from 184 to 2,190 mg/L. Contours were generated from a point shapefile, using an interval of 500 mg/L. A polygon coverage outlining the TDS class boundaries was then manually created in ArcInfo using the contoured shapefile (converted to a coverage) as a guide; this coverage was subsequently clipped to the bedrock boundary within the valley-fill areas.

A ground-water quality classification map formally identifies the beneficial use of the ground-water resource based primarily on TDS concentrations as follows: class 1A, less than 500 mg/L; class 2, 500 to less than 3,000 mg/L; class 3, 3,000 to less than 10,000 mg/L; and class 4, 10,000 mg/L and greater. Areas where individual constituents exceed drinking-water standards are class 3 water. The ground-water quality classification map was compiled from the TDS map, supplemented with a GIS-nitrate concentration coverage using these criteria. We also calculated the land-surface-area percentage of each ground-water class category by building the polygon coverage in ArcInfo and calculating the corresponding areas. On the basis of chemical analyses of water from 94 wells sampled during 1974 to 2000, Cedar Valley ground water classified as follows: class 1A, 80%, primarily in the central and western parts of the valley; class 2, 19%, primarily in the eastern part of the valley; and class 3, 1%, an area of persistent nitrate contamination northwest of Cedar City. Land-use planners can now use these maps as a basis for enacting regulations to protect water resources in this valley.

APPENDIX A

List of Workshop Attendees

[Grouped by affiliation]

Alabama Geological Survey
Nick Tew

Alaska Division of Geological & Geophysical Surveys
Gail Davidson
Larry Freeman

Albanian Geological Survey
Burbuqe Agolli
Mimiza Simixhiu

Arizona Geological Survey
Stephen Richard

California Geological Survey
George Saucedo

Colorado Geological Survey
Randal Phillips

Colorado State University/National Park Service
Victor de Wolfe III
Stephanie O'Meara
Trista Thornberry

ESRI
Mike Price
Chris Wayne

GIS Consultants
Bruce Joffe

Geological Survey of Canada
Martin Anctil
Eric Boisvert
Boyan Brodaric
Peter Davenport
Marianne Quat

Idaho Geological Survey
Jane Freed
Kurt Othberg
Loudon Stanford
Benjamin Studer

Illinois State Geological Survey
Curt Abert
Daniel Byers
Jane Domier
Patrick Johnstone
Don Keefer
Robert Krumm
Barbara Stiff

Indiana Geological Survey
Richard Hill
Kim Sowder

Kansas Geological Survey
David Collins
Jorgina Ross

Kennecott Utah Copper
Stan Nelson

Kentucky Geological Survey
James Cobb
Jerry Weisenfluh

Library of Congress
Colleen Cahill

LizardTech
Todd Packebush

Louisiana Geological Survey
R. Hampton Peele
John Snead

MinGIS
Gary Edmondo

Montana Bureau of Mines and Geology
Patrick Kennelly
Susan Smith
Paul Thale

National Park Service
Tim Connors
Anne Poole

Natural Resources Canada
Vic Dohar
Dave Everett
Terry Houlahan

Nevada Bureau of Mines & Geology
Gary Johnson

New Hampshire Geological Survey
Rick Chormann

New Mexico Bureau of Mines and Mineral Resources
Kathryn Glesener
David McCraw

North Dakota Geological Survey
Mark Gonzales
Lorraine Manz

Ohio Geological Survey
Thomas Berg
James McDonald

Oklahoma Geological Survey
James Anderson
Russell Standridge

Onyx Graphics Corp.
Dean Derhak

Oregon Department of Geology and Mineral Industries
Clark Niewendorp

Pennsylvania Geological Survey
William Kochanov
Thomas Whitfield

University of Alabama
Doug Behm
Craig Remington

University of California, Berkeley
George Brimhall

University of California, Santa Barbara
Jordan Hastings
Ada Otter

U.S. Geological Survey
Debra Block
VeeAnn Cross
James Estabrook
Bruce Johnson
Diane Lane
Peter Lyttle
Peter Schweitzer
Dave Soller
Nancy Stamm
Will Stettner
Ronald Wahl
Bruce Wardlaw
Robert Wardwell

Utah Department of Transportation
Christopher Meredith

Utah Geological Survey
Kelli Bacon
Bob Biek
Kent Brown
Matthew Butler
Bill Case
Jon King
Basia Matyjasik
James Parker
Pat Speranza
Doug Sprinkel
Neil Storey
Grant Willis

Washington Department of Natural Resources
Charles Caruthers

West Virginia Geological Survey
Jane McColloch
Scott McColloch

Wisconsin Geological and Natural History Survey
Mindy James
Kurt Zeiler

Wyoming Geological Survey
Joseph Huss

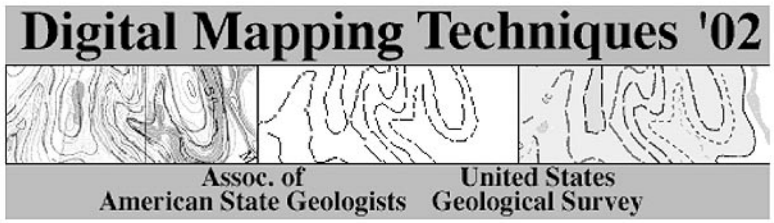
APPENDIX B

Workshop Web Site

Digital Mapping Techniques 2002 - Utah Geological Survey

Utah Geological Survey

- What's New
- Utah Geology
- Maps Online
- Publications
- Educational Resources
- About UGS
- Map & Bookstore
- UGS Home
- Natural Resources



Hosted by
Utah Geological Survey
and
University of Utah Department of Geology and Geophysics

Convened by
Association of American State Geologists
and
U.S. Geological Survey

Meeting Dates

Sunday, May 19 to Wednesday, May 22, 2002

Location

Sunday afternoon: Workshop on field mapping and data acquisition techniques hosted by ESRI. Assemble at 1:00 pm at the University Park Marriott Hotel. We will shuttle to the field site about ½ mile from the hotel.

If you arrive after 1:00 pm Sunday, drive or walk east (uphill) up Wakara Way about 3 blocks to Colorow Road. Park along the street and walk up the trail about 100 yards to the Lake Bonneville shoreline bench.

We will leave a detailed map to the field site with the hotel front desk and post a map on the official meeting web site. The field workshop ends about 5:30 pm.

Sunday 6-8 pm: Icebreaker - University Park Marriott Hotel.

Monday - Wednesday: Eccles Auditorium, 6th Floor of the Huntsman Cancer Institute, adjacent to the University of Utah Campus, Salt Lake City, Utah

APPENDIX C

List of Addresses, Telephone Numbers, and URL's for Software and Hardware Suppliers

[Information contained herein was provided by the authors of the various articles and has not been checked by the editors for accuracy]

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

Adobe Illustrator, Photoshop—Adobe Systems, Inc., 345 Park Ave., San Jose, CA 95110-2704 USA, (800) 833-6687, <http://www.adobe.com>.

ColdFusion Server—Macromedia, Inc., 600 Townsend St., San Francisco, CA 94103 USA, (415) 252-2000, <http://www.macromedia.com/>.

Colortrac 5480 scanner, Tangent scanner—ACTion Imaging Solutions US, 10499 Bradford Rd., Littleton, CO 80127 USA, (303) 973-6722, <http://www.action-imaging.com/>.

Contex scanner—CONTEX Scanning Technology Inc., 3200 Inland Empire Blvd., Suite 160, Ontario, CA 91764 USA, (877) 226-6839 <http://www.contex.com>.

EarthVision—Dynamic Graphics Inc., 1015 Atlantic Ave., Alameda, CA 94501-1154, (510) 522-0700, <http://www.dgi.com/>.

ERwin—Computer Associates International, One Computer Associates Plaza, Islandia, NY 11749 USA, (800) 225-5224, <http://www3.ca.com/Solutions/Product.asp?ID=260>.

Garmin—GARMIN International Inc., 1200 E. 151st St., Olathe, KS 66062, (913) 397-8200, <http://www.garmin.com/>.

Geolink—Michael Baker Corp., Airport Office Park, Building 3, 420 Rouser Rd., Coraopolis, PA 15108, (800) 642-2537, <http://www.mbakercorp.com/gis/>.

Geologic Data Systems, Inc.—1600 Emerson St., Denver, CO 80218 USA, (303) 837-1699, <http://www.gdata.com/>.

GeoMapper and GeoLogger—Rubicon Digital Mapping Associates, 1240 6th St., Suite D, Berkeley, CA 94710-1402 USA, <http://www.rubicondigital.com/>.

Google search engine—<http://www.google.com/>.

HP DesignJet 5000ps plotter, iPAQ Pocket PC—Hewlett-Packard, Inc., 3000 Hanover St., Palo Alto, CA, 94304-1185 USA, (800) 752-0900, <http://www.hp.com/>.

Macromedia, Inc.—600 Townsend Street, San Francisco, CA 94103 USA, (415) 252-2000, <http://www.macromedia.com/>.

Magellan—Thales Navigation, (800) 707-9971, <http://www.magellangps.com/>.

MapInfo—MapInfo Corp., One Global View, Troy, New York 12180 USA, (800) FASTMAP, <http://www.mapinfo.com/>.

Microsoft Access—Microsoft Corp., One Microsoft Way, Redmond, WA 98052-6399 USA, (425) 882-8080, http://www.microsoft.com/office/access/.

Microsoft.NET—Microsoft Corp., One Microsoft Way, Redmond, WA 98052-6399 USA, (425) 882-8080, <http://msdn.microsoft.com/net>.

Microsoft SQL Server—Microsoft Corp., One Microsoft Way, Redmond, WA 98052-6399 USA, (425) 882-8080, <http://www.microsoft.com/sql/default.asp>.

MrSID—LizardTech, The National Building, 2nd Floor, 1008 Western Ave., Seattle, WA 98104 USA, (206) 652-5211, <http://www.lizardtech.com/>.

Océ Imaging Supplies—Océ-USA, Inc., 1800 Bruning Dr. West, Itasca, IL 60143-1039 USA, (800) 714-4427, <http://www.oceusa.com/>.

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

PenMap—Condor Earth Technologies Inc., 21663 Brian Lane, Sonora, CA 95370-3905 USA, (209) 532-0361, <http://www.conndorearth.com/products>; Strata Software & Consultancy, Ltd., The Business & Innovation Centre, Angel Way, Bradford, North Yorkshire, England BD7 1BX, +44 (0)1274 841325, <http://www.penmap.com>.

PhazeOne, Consulting—3101 Broadway, Suite 320, Kansas City, MO 64111 USA, (816) 756-1300, (816) 756-1355, <http://www.phazeone.com/>.

R2V—Able Software Corp., 5 Appletree Lane, Lexington, MA 02420-2406 USA, (781) 862-2804, <http://www.ablesw.com/>.

Surfer by Golden Software—Rockware Inc., 2221 East St., Golden CO 80401, (303) 278-3534, <http://www.rockware.com/>.

Trimble Navigation, Trimble Pathfinder Power, Pro XR series—Trimble Navigation Limited, 645 N. Mary Avenue, Sunnyvale, CA, 94088-3642, (408) 481-8000, <http://www.trimble.com/>.

Widecom scanner—WIDECOM Group, Inc., 37 George St. North, Suite 103, Brampton, Ontario, Canada, L6X 1R5, (905) 712-0505, <http://www.widecom.com/>.

Xybernaut MA V—Xybernaut Corp., 12701 Fair Lakes Circle, Suite 550, Fairfax, VA 22033, (703) 631-692, <http://www.xybernaut.com/>.