Prototype GIS Database for the DNAG Geologic Map of North America

By Christopher P. Garrity and David R. Soller

U.S. Geological Survey
National Center, MS 956
12201 Sunrise Valley Drive
Reston, VA 20192
Telephone: (703) 648-6426
Fax: (703) 648-6977
e-mail: {cgarrity, drsoller}@usgs.gov

PURPOSE[1]

When plans for the Geologic Map of North America (GMNA) were being made, the notion of geologic map databases was in its infancy. At that time, and for many years thereafter, few geologists were familiar with the design and use of databases to manage geologic map information. In 1998, the Geological Society of America (GSA) and the United States Geological Survey (USGS) National Geologic Map Database project agreed to cost-share the digital preparation of this map. The plan was to digitize the hand-drawn, author-prepared geologic compilations for the four map quadrants to provide digital data for two purposes: (1) to allow GSA to print the map, and (2) to permit the National Geologic Map Database project to develop a prototype database for this map. With the map now printed, the National Geologic Map Database project has begun to design and create the prototype, based on certain assumptions regarding the anticipated content of, and uses for, the map database. The first version of this database will contain the descriptive information for geologic units shown on the map. It will serve as the fundamental entity from which products of the map can be derived. These products may be interpretive, or they may be future editions of the map.

In mid-2006, this prototype will be provided to the organizations principally responsible for map compilation (GSA, USGS, GSC, and WHOI) to initiate discussion and decisions on how the map database will be designed, managed, and distributed served to the public and cooperators. The prototype is shown here to generate technical discussion and guidance prior to formal discussion among those organizations.

To produce any future editions of the map, the database will incorporate all map revisions that are necessitated by detection of compilation errors and by new regional mapping and interpretations. Further, the geologic unit descriptions shown on the printed map can be supplemented in the database by more detailed, richly attributed information derived from the many sources that were used to compile the map. This capability to revise the printed map and include additional descriptive information for map units is one of the primary reasons for building the database; the other reason is, of course, the analytical capabilities made possible by providing the map in a digital, Geographic Information System (GIS) compatible format.

The creation of this database and its enhancement to include new mapping and more richly attributed information is a daunting task that will take a significant amount of time and effort. Recognizing that a group of dedicated and knowledgeable scientists is essential to make this database useful and to keep its content up to date, GSA will develop a consortium of geological agencies to manage the database. With prototype development of the database, the National Geologic Map Database project provides a basis for this consortium to proceed.

PROTOTYPE AREA

The GMNA prototype spans an area of about 530,206 square miles and includes both continental and seafloor geologic units in the United States and Canada (Figure 1). The area was chosen due to its relatively complex geology and abundance of both onshore and offshore map symbology displayed in the published GMNA. The prototype contains over 2,500 individual polygon features and about 5,700 line features. Polygon features are symbolized by 205 unique geologic unit values, each with custom color and pattern fills designed to mimic those used in the published GMNA (Figure 2). The prototype contains customized line symbology which nearly duplicates the symbol sets used in the published map. All custom symbol sets are stored in ArcGIS layer files. Line features are classified by

[1]Modified from Soller, in Reed et al. (2005)

Figure 1. The 2005 Geologic Map of North America (GMNA) is the first such map published in the past four decades. The map, which covers 15% of the Earth's surface (shown in yellow box), depicts the geology of the seafloor in detail never before seen on a map of this scale (1:5,000,000). It is the first geologic map of North America to be compiled since the general acceptance of plate-tectonic theory and since radiometric dates for plutonic and volcanic rocks because widely available. This map distinguishes more than 900 rock unites, 100 of which are offshore. It depicts more than seven times as many terrestrial units as are shown on the previous 1965 map, as well as detailed features of the seafloor, such as spreading centers, seamont chains, and subduction zones (Reed, et. al., 2005). The GIS prototype area (shown in red box) was chosen due to its relatively complex geology and abundance of both on- and offshore map symbology displayed in the published GMNA. The release of the prototype is meant to serve as a forum for both general comments on the overall objectives of the GIS database design, and specific comment son elements such as cartographic symbolization style.

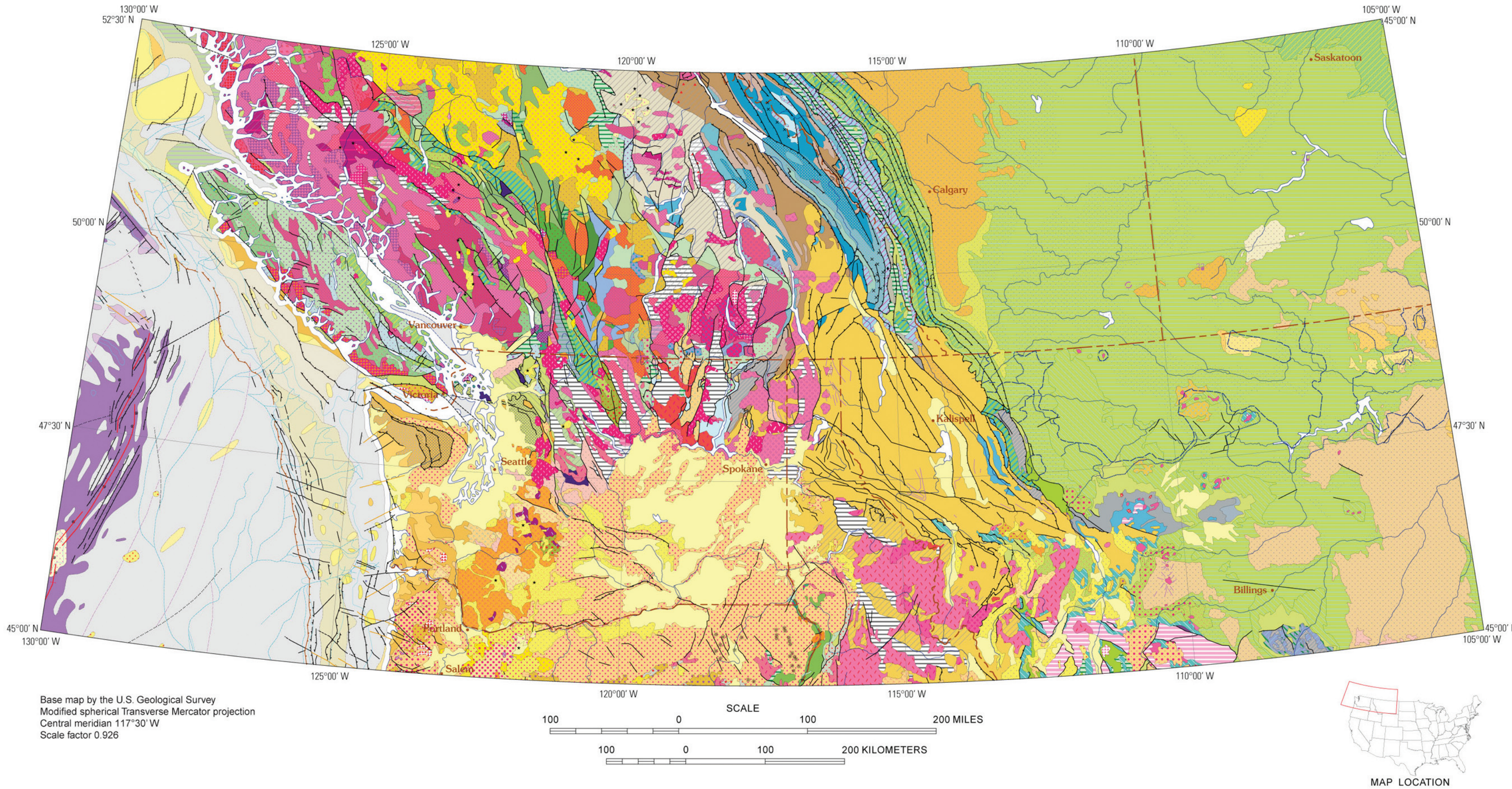

Figure 2. Layout of the prototype area exported from ArcMap. The prototype, spanning an area of about 530,206 square miles, contains over 2,500 individual polygon features and about 5,700 line features. Polygon features are symbolized by 205 unique geologic unit values, each with custom color and pattern fills designed to mimic those used in the published GMNA.

geologic line type into 26 subtypes. Geologic line types include general geologic features (contacts, faults, etc.), special submarine features (slump scars, seamount chains, spreading centers, etc.), and lithologic/age defined dikes and sills. Each subtype definition matches the explanation of the corresponding map symbol used in the GMNA.

PROCESS DESCRIPTION

A secondary objective in creation of the prototype was to determine the most efficient way to convert such an enormous digital map into a useable GIS. It was also important to determine a realistic time frame in which a project of this size could be completed. The two Adobe Illustrator files that contain the source digital data were massive, with layer counts totaling over 1,500. When analyzed at scales much greater than the intended map scale, the files showed areas where problems in topological relationships existed. Common topological problems in these areas included polygons that overlapped or had gaps between them, overlying line layers (contacts, faults, etc.) that were not coincident with polygon boundaries, and line features that self-overlapped. Although these areas were much too small to compromise the quality of the hard copy layouts, they did present problems when validating the topology in a GIS. Due to the sheer number of imported features, the errors reported after validating topology numbered in the tens of thousands. To avoid the time consuming process of correcting each error, it was decided that only the non-contact linework coincident with geologic unit boundaries would be imported directly from Adobe Illustrator. For attribution purposes, each line type (inferred thrust fault, concealed thrust fault, etc.) was imported to the GIS individually. The remaining linework (contact layer) was isolated in Adobe Illustrator and exported as a high-resolution raster image. The image was georeferenced using control points in the DNAG projection (Snyder, 1987), and auto-vectorized in ArcScan.
By setting topology rules in ArcMap, line dangles in the vectorized layer were snapped to the nearest unit-bordering line feature, which resulted in a topologically clean layer. To build unit attributes quickly, individual geologic unit layers were batch exported from Adobe Illustrator and used to overlay the newly created layer. Through spatial querying, polygons in the new layer that had their center within a specific overlay layer were attributed based on the overlay's geologic unit abbreviation. Much of the remaining attribution was completed through simple VBA field calculator scripting based on the unit abbreviation field. The final step, which proved to be the most time consuming, involved recreation of feature class symbolization in ArcGIS. 26 symbolized line types, and 205 unique color fills/patterns had to be created to mimic those in the source Adobe Illustrator file. The prototype was completed in about 10 days. GIS compatible files for the southern map sheet, which comprises the conterminous United States, are scheduled to be completed and released in FY2007. Release of GIS compatible files for the northern map sheet is scheduled to follow.

PROTOTYPE DATABASE

This prototype is intended to serve as a forum for comments on the overall objectives of the GIS database and the attribute information within. It is certain that as the database evolves, attribute information will be modified to make the database more useful. For the prototype database, a preliminary set of attributes was chosen to serve as a foundation for an eventual GMNA data model. When that data model is formalized, we anticipate that it will incorporate elements of the North American Data Model (http://nadm-geo.org/) and the International Geological Map of Europe's data model (http://www.bgr.de/karten/IGME5000/igme5000.htm). The selected attributes attempt to capture the information depicted in the GMNA explanation sheet that accompanies the published map. The attribute list includes:

- ROCKTYPE – the "top level" rock classification (sedimentary, plutonic, volcanic, metamorphic)
- LITHOLOGY – the simplified description included for each geologic unit on the explanation sheet of the GMNA
- ROCK_UNIT_NOTE – special notes associated with certain units on the explanation sheet of the GMNA. For example, selected volcanic rocks are attributed "Basalt adjacent to active spreading centers," selected metamorphic rocks are attributed "Granulate facies metamorphism," and selected sedimentary rocks are attributed "Continental deposits"
- UNIT_UNCERTAINTY – a query following the map unit code indicates uncertainty about composition, or whether the rock is in situ
- MIN_AGE – minimum geologic age for the unit. Subdivisions of time-stratigraphic units are lower, middle, and upper (lower-case), and for plutonic rocks are Early, Middle, and Late

- MAX_AGE – see comments for MIN_AGE
- MIN_AGE_CODE – code derived from the geologic age codes defined by the AAPG Committee on Standard Stratigraphic Coding (1967)
- MAX_AGE_CODE – see comments for MIN_AGE_CODE
- AGE_UNCERTAINTY – a query preceding the map unit label indicates uncertainty about the assigned age
- MAP_UNIT_CODE – the GMNA map unit code
- MIN_MAX_RELATE – the relationship (“and”, “or”, “thru”) between the MIN?MAX ages of units bounded by multiple ages

REFERENCES

American Association of Petroleum Geologists, Committee on Standard Stratigraphic Coding (George V. Cohee, Chairman), 1967, Standard stratigraphic code adopted by AAPG: American Assiciation of Petroleum Geologists Bulletin, v. 51, no. 10, p. 2146-2150.

Reed, J.C. Jr., Wheeler, J.O., and Tucholke, B.E., compilers, 2004, Geologic Map of North America: Decade of North American Geology Continental Scale Map 001, Boulder, Geological Society of America, scale 1:5,000,000.

Reed, J.C. Jr., Wheeler, J.O., and Tucholke, B.E., 2005, Geologic Map of North America—Perspectives and explanation: Boulder, Colorado, Geological Society of America, Decade of North American Geology, 28 p.

Snyder, J.P., 1987, Map projections—a working manual: U.S. Geological Survey Professional Paper 1395, 383 p.