

National Park Service Geologic Resources Inventory: Data Model Concepts and Implementation, and a Programmatic Approach to Digital Map Production

By Stephanie O'Meara and Jim Chappell

Colorado State University
Department of Geosciences
Fort Collins, CO 80523
Telephone: (970) 491-6655 and (970) 491-5147
email: stephanie_o'meara@colostate.edu and jim_chappell@partner.nps.gov

Introduction

The Geologic Resources Inventory (GRI) program is tasked with producing geologic information for 270 National Park Service (NPS) parks with significant natural resources. The program is funded by the NPS Inventory and Monitoring Division (IMD) and is administered by the NPS Geologic Resources Division (GRD). The GRI program relies heavily upon partnerships with Colorado State University (CSU), the U.S. Geological Survey (USGS), individual State surveys, and other organizations in developing its products.

In developing GRI products, CSU research associates work side-by-side with GRD staff, attending scoping meetings at parks to identify mapping needs and park-specific geologic issues, features, and processes. A scoping summary report is then produced. The geologic issues, features, and processes identified at scoping are then further explained in a geology report written for park resource managers. From the scoping plan, source geology maps, in paper, mylar, and (or) digital format, are acquired and evaluated, then processed into the GRI Geology-GIS Geodatabase Data Model (O'Meara and others, 2010), which is in ESRI's geodatabase format.

To facilitate the creation of a useful and high quality digital map product, good data model designs, as well as efficient map production processes, are needed. This paper presents prominent concepts and requirements considered in the design and implementation of the GRI data model and the approach utilized in streamlining digital map production.

Data Model Concepts and Design

In developing a geology-GIS data model, there are typically a number of design requirements that should be considered to ensure a well-developed working data model that effectively communicates geologic information to the intended data users and promotes the production of consistent quality data.

When designing the GRI Geology-GIS Geodatabase Data Model, several base design requirements, as well as factors such as geologic diversity across our national parks, variable source map scale, and map compilation considerations were addressed.

Fundamental Data Model Design Requirements

- Model is implementable in standard GIS software. The GIS software widely employed by the NPS is ESRI ArcGIS.
- Intended users of our data are park resource managers, most of whom are scientists but not geologists!
- Geologic information on source map is preserved and effectively communicated as GIS data (as features and tables) or as ancillary documents (as report text, meta-data, or graphics).

Other Design Requirements and Challenges

- Geology across the land managed by the NPS is varied and diverse, with each geologic terrain often possessing its own set of geologic features and observations. Such geologic diversity requires a data model that is flexible and can accommodate new features.
- Map scale considerations: Features may vary in their spatial representation (polygon, line, or point) depending on map scale. In this case the data model needed to accommodate changes to the spatial representation of some features, as these can vary depending on the geologic feature's spatial extent and the scale at which the feature was mapped. For example, on most maps a gravel pit is represented as a point locality; however, if the feature is of significant size and (or) the feature was mapped at a very large scale (for example, 1:12,000) the gravel pit would likely be an area (polygon) feature. This variation in spatial representation is present amongst a significant number of geologic features found on geologic maps of different scales.
- Map compilation considerations: Many GRI park maps are compiled from multiple source maps. This frequently results in the integration of geologic features not present on every source map. In these cases all features are incorporated into the compiled map (none are simply omitted), and in some cases this dictates that some features are given their own feature class (for example, the integration of two or more sets of different structure contour lines where each set is given its own feature class and is not simply merged into one structure contour line feature class).

Data Model Implementation

It was a fundamental design requirement that the GRI data model had to be implementable in ESRI ArcGIS software, which is the GIS software widely used within the NPS. The latest and most functional ESRI GIS format is the geodatabase. This format provides robust functionality that the GRI data model fully utilizes to store, attribute, and relate features. Characteristics of the GRI data model are described below.

GIS Data Format and Architecture

- Geologic-GIS data are implemented in an ESRI 9.x personal geodatabase.
- Feature class attribute tables comprise just those fields necessary to fully capture all applicable information.

- Geologic features commonly are grouped into data layers (feature classes) based upon the geologic processes that created them (for example, deformation/structural, volcanic, glacial), for ease of presentation to our intended users.
- We continue to evaluate a revision to an ESRI 9.x/10.x file-based geodatabase format.

GIS Building Blocks

The GRI data model employs much of the functionality inherent in the ESRI geodatabase format to depict, attribute, ensure spatial coincidence, and relate geologic features and observations to ancillary GIS tables.

- Geologic features are depicted as area (polygon), line, or points in discrete data layers referred to as feature classes.
- Only 25 data model attribute fields are employed for data model feature classes. Custom attribute fields also can be readily added, and both coded and ranged attribute domains are implemented (fig. 1).
- Geodatabase topology is implemented to ensure no gaps, no overlaps, and no dangles and to ensure feature coincidence between features where appropriate (fig. 2).

Field Name	Field Alias	Data Type	Allow Nulls	Implemented Domain	Precision	Scale	Length
OBJECT_ID*	NA	Object ID	—	—	—	—	—
SHAPE*	NA	Geometry	Yes	—	—	—	—
FUID	Unique Feature ID	Long Integer	No	—	0	—	—
GLG_SYM ⁽²⁾	Unit Symbol	Text	No	—	—	—	12
SRC_SYM	Source Unit Symbol	Text	No	—	—	—	12
SORT_NO	Sort Number	Float	No	—	6	3	—
NOTES	Notes	Text	No	—	—	—	254
LBL	Label	Text	Yes	—	—	—	60
GMAP_ID ⁽¹⁾	Source Map ID	Long Integer	No	—	0	—	—
SHAPE_Length*	NA	Double	Yes	—	0	—	—
SHAPE_Area*	NA	Double	Yes	—	0	—	—

* Standard ESRI 9.X geodatabase feature class attribute field.
 (1) Relationship class foreign key field to MAP table.
 (2) Relationship class foreign key field to UNIT table.

Ranged Domain Value	Definition
1	minimum value
359	maximum value
999	not applicable (NULL) value

Figure 1. Geologic Units (GLG) feature class attribute table parameters, and Strike/Trend (STRIKE_ROTATION) Ranged Domain list (lower right).

- Ancillary GIS tables consistent of a Geologic Unit Information (UNIT) and a Source Map Information (MAP) table (fig. 3).
- Feature classes are linked to ancillary tables via relationship classes using a common key field.
- Additional GIS tables, if present in the source data, can be readily added as-is, or as custom table schema created for a specific map or, if needed, implemented for other (future) maps that will contain the same table.

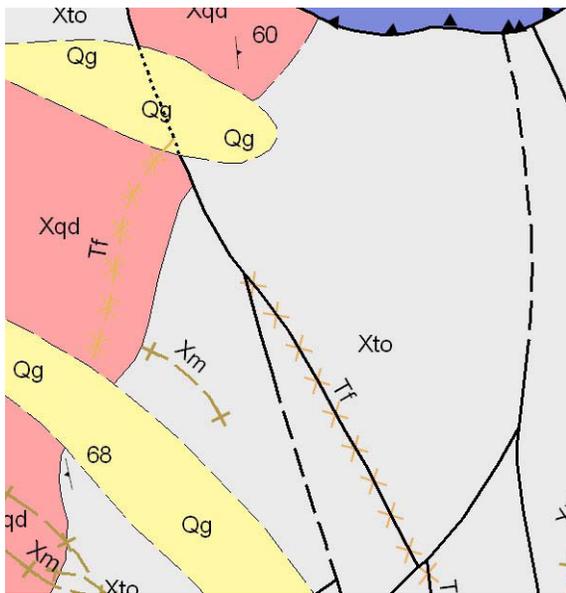


Figure 2. Dike intruded along a fault (fault is solid dark line, with dike shown diagrammatically as lighter colored X's and labeled Tf). Feature coincidence is maintained between the Linear Dikes (DKE) and Faults (FLT) feature classes via topology rules. If either the fault or dike feature is spatially edited using topology edit tools, then both features are edited.

Feature Class Implementation

Many data model feature classes can be repeated if warranted (for example, for different structure contour lines or for different area hazards). To implement many feature classes, our data model employs the use of shared schema. Feature classes share the same schema when they have the same:

- Spatial geometry (polygon, line, or point).
- Attribute fields (the minimum required to fully attribute).
- Table-to-table relationships.
- Topological rules.

Shared data model schema are referred to as a “Template Feature Class Definition” in our data model. Seven template feature class definitions are employed to represent 44 of the 56 possible GRI data model feature classes (fig. 4).

Production Workflow and Development

Capturing geologic-GIS data can be a time-consuming process. Often the steps involved in digital data production introduce a component of human error due to the repetitive and sometimes complex processes involved in digital GIS data production. A task that is seemingly simple, like adding a set of data-model-defined topology rules, can be a repetitive and time-consuming process with little control over whether the rules were added correctly. By automating certain processes like adding data model topology rules, some tasks can be significantly streamlined and errors caused by manual processes eliminated. The GRI development team has identified tasks within the GRI GIS production workflow that can be automated through custom programming (fig. 5).

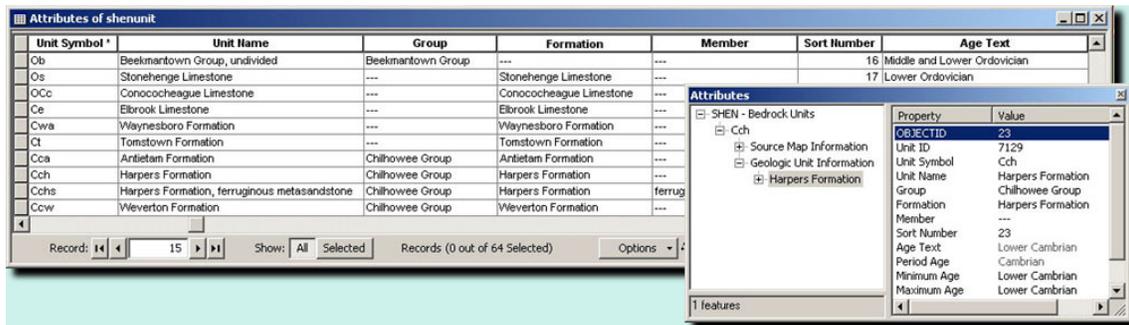


Figure 3. Shenandoah NP (SHEN) Geologic Unit Information (UNIT) Table and an ArcMap Information Window (lower right) showing UNIT table information related via a relationship class to a bedrock unit (Cch) polygon in the Geologic Units (GLG) feature class.

Figure 4. Partial extent of the GRI John Muir National Historic Site (JOMU) digital map (from Haydon, 1995) showing JOMU data model feature classes (upper left) and hazard feature classes (middle to lower left). Both area hazard susceptibility data layers, as well as the Hazard Area Features feature class, implement shared data model schema referred to as a template feature class definition in the GRI data model. The park is in the center of the figure.

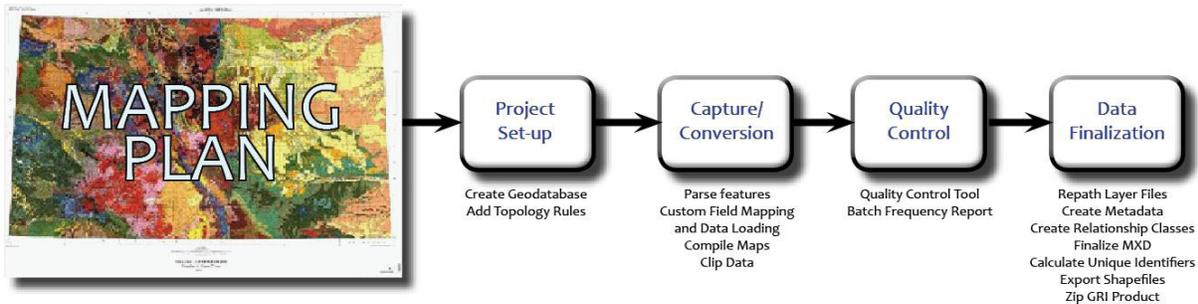
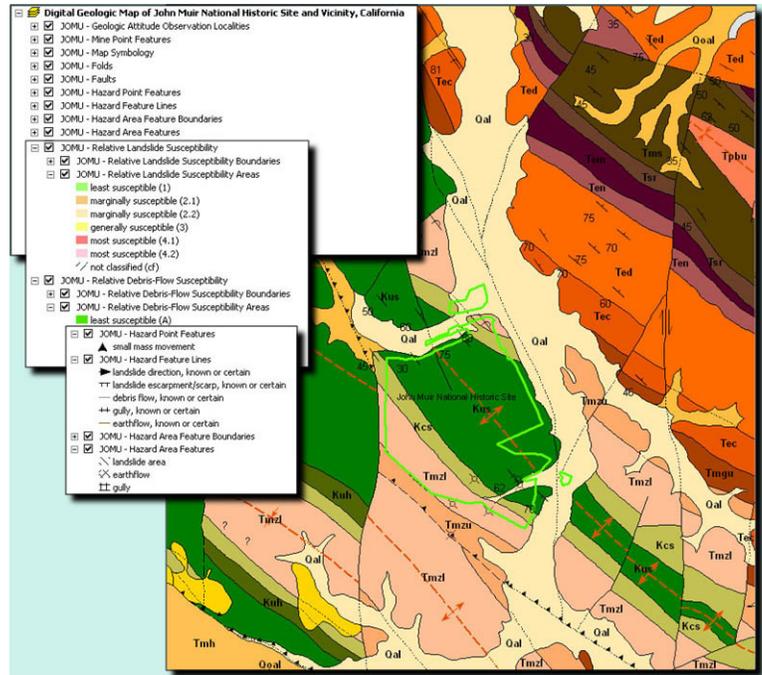


Figure 5. Overview of the GRI digital map production workflow, starting with a mapping plan for a specific park and ending with map finalization. These processes use automation tools and scripts (see captions for workflow steps) to provide efficiency and quality not possible with manual processing.

Development Approach

The GRI development team comprises project managers who have programming ability but also have annual production responsibilities. Most programming, as a result, must be accomplished when it can be fit in with production tasks. Because GRI project managers are familiar with the production workflow, they are able to identify processes that are error prone, inefficient, or could be automated. To reduce the amount of programming, developer samples, snippets, starter code, and other applicable toolsets are acquired, evaluated, and utilized whenever possible. Newly developed GRI tools are tested on real data, refined, and then deployed to the GRI production team. The resulting production tools range from simple macros run in ArcToolbox to more complex scripts and

applications utilizing ArcObjects and .NET. This simplistic approach to tool development enables the development team to get production tools into the general workflow quickly while significantly reducing development time.

Create GDB Tool

The CreateGDB tool (fig. 6) is a wizard-like tool, initially developed in VBA and later migrated to VB.NET, that enables a user to easily create a GRI data model-compliant geodatabase. It prompts the user to select applicable feature classes, create custom feature classes, and includes the option to generate ancillary data model GIS tables. Data model domains are associated with respective feature classes, and applicable topology rules are added to the final output.

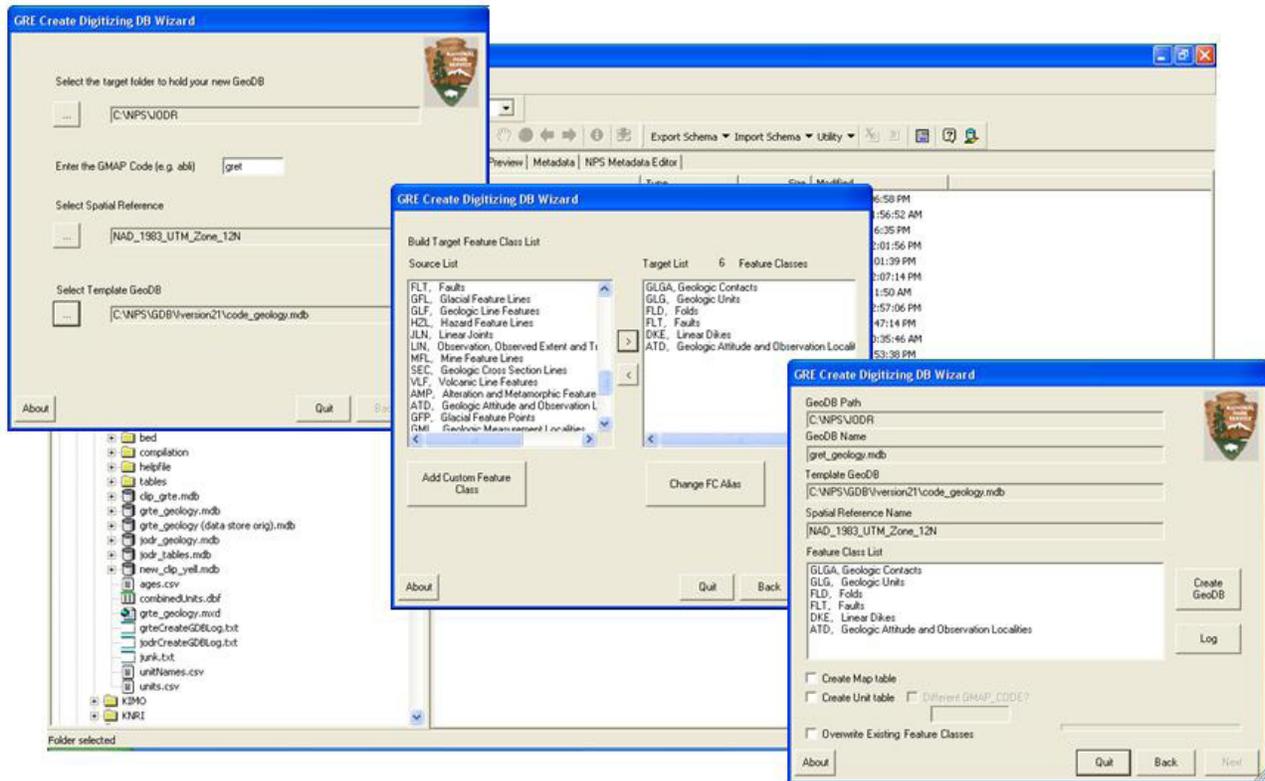


Figure 6. The CreateGDB tool. The first dialog (upper left) of the wizard prompts users for dataset name, location, spatial reference, and GRI template geodatabase. The second dialog (middle) allows users to select feature classes, create custom feature classes, and change feature class aliases, if desired. The third and last dialog (lower right) summarizes user defined parameters and provides option to create ancillary tables before generating the new geodatabase.

QC Tool

Designed in Python and later recoded in VB.NET, this tool prescribes spatial and attribute rules or tests based on feature classes present within a specified GRI data model geodatabase. For example, all water polygons must be bordered by shoreline, and contacts on the edge of the map must be attributed as map boundary. Run in ArcMap (fig. 7), the QC tool reports and graphically highlights errors while providing “zoom to” and selection options to aid in error resolution.

Summary

The GRI data model needed to be flexible and not too technical in design. Primary factors that influenced the design were (1) our anticipated data users are not geologists and their use of our data varies according to their backgrounds and the priorities established for the particular park; (2) the data model needed to preserve all source map information; (3) there is varied and diverse geology across the lands managed by the

NPS; (4) we often use large-scale source maps; and (5) we frequently produce map compilations.

The GRI geology-GIS data model is implemented in an ESRI 9.x personal geodatabase and makes use of much of the functionality (attribute domains, topology, relationship classes) that this format provides. The GRI data model preserves all source map geologic information and presents this information in data layers and attribution that can easily be understood and manipulated by our users. As a result of our design and implementation methodology, our data model can accommodate the addition of new features, as well as new data layers as these are recognized. In addition, the data model is simplified by implementing many features classes using shared schema.

The GRI production workflow has been fine-tuned through the insertion of custom-programmed tools and scripts that increase production efficiency while yielding high quality and consistent GIS data. Because programming of these custom tools and scripts is completed by project managers, who are intimate with the production workflow, the time it takes to implement is greatly reduced.

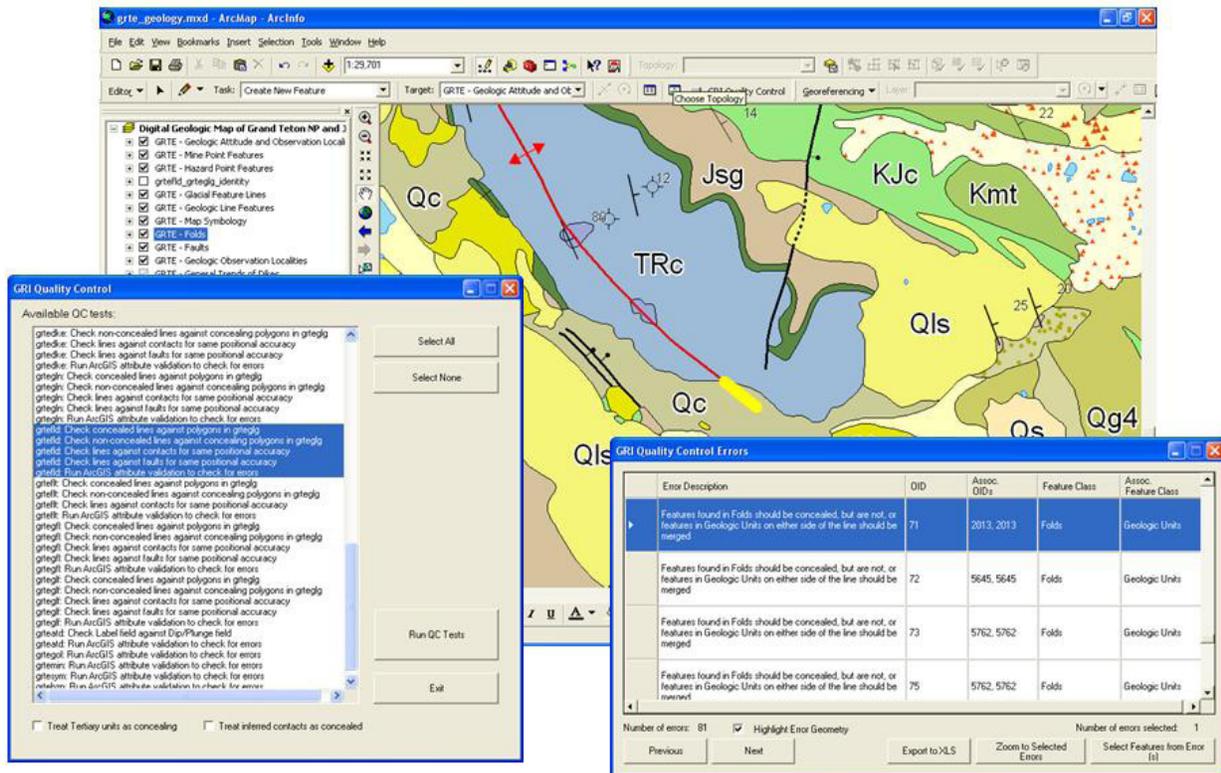


Figure 7. The QC tool. The dialog on the left shows tests prescribed for a specific collection of feature classes. The first test in the highlighted block of tests suggests checking fold axis positional accuracy with surrounding geologic unit polygons. For example, typically, most non-Quaternary linear features that have the same Quaternary unit on either side should be attributed as concealed. The dialog on the right shows results from that test. The highlighted test result shows that the feature highlighted in yellow on the map should be attributed as concealed but is currently attributed as approximate.

GRI Products

GRI GIS data and report products are available for download at the NPS Natural Resource Information Reference Search Application: <http://nriinfo.nps.gov/Reference.mvc/Search>. Enter the search word “gri” into the search text, and select the park(s) from the units listed.

Geologic Resources Inventory Products: http://www.nature.nps.gov/geology/inventory/gre_publications.cfm.

References

Haydon, W.D., 1995, Landslide Hazards in the Martinez-Orinda-Walnut Creek area, Contra Costa County, California: California Division of Mines and Geology Landslide Hazard Identification Map 32, OFR 95-12, 4 plates, 1:24,000 scale.

Johnson, B. R. and Bruce, R.M., 1991, Reconnaissance geologic map of parts of the Twin Peaks and Blanco Peak Quadrangles, Alamosa, Costilla and Huerfano Counties, Colorado: U.S. Geological Survey Miscellaneous Investigations Series Map MF-2169, 1:24,000 scale.

O’Meara, S.A., Stanton, H.I., and Chappell, J.R., with contributions from Mack, G., Hybels, G., Karpilo, R.D., and Croskrey, A., 2010, National Park Service Geologic Resources Inventory Geology-GIS Geodatabase Data Model (v. 2.1): Colorado State University and National Park Service, 95 p., <http://science.nature.nps.gov/im/inventory/geology/GeologyGISDataModel.htm>.