THE DEVELOPMENT OF A NATIONAL SMALL-SCALE DIGITAL CARTOGRAPHIC DATA BASE

Malcolm J. Stephens
U.S. Geological Survey
Topographic Division
National Center, MS 560
Reston, Virginia 22092

Michael A. Domaratz
U.S. Geological Survey
Topographic Division
National Center, MS 560
Reston, Virginia 22092

Warren E. Schmidt
U.S. Geological Survey
Topographic Division
National Center, MS 519
Reston, Virginia 22092

I. Introduction

The United States Geological Survey has identified an urgent multi-user requirement for a national small-scale digital cartographic data base. To satisfy the expressed needs of both the USGS and other user agencies, it was concluded that the data base had to be current, topologically structured, no smaller in scale than 1:2,000,000, and sufficiently flexible to permit the generation of smaller-scale maps, thematic overlays, and other graphic products. Existing data bases, software, and hardware were investigated, but all had limitations. After considerable study and planning, a series of tasks directed toward the generation of a comprehensive national data base was defined. These tasks involved primarily the modification and linking of existing soft-

ware systems, the development of some new programs, and the acquisition of additional digitizing hardware. The project was initiated in June 1979, with sample products for a selected test area planned for September. Early in 1980, a digital cartographic data base covering a major portion of the nation will be available at a scale of 1:2,000,000 and, with it, the capability to automatically display a host of new base, reference, thematic, and index maps.

In an attempt to respond to growing user demands for up-to-date cartographic data and products on a nationwide basis, the Geological Survey has been increasingly looking toward the application of numerical techniques to the cartographic and geographic sciences to help meet these needs. This approach offers the attractive features of reduced manual effort, quick response, high precision, and versatility (i.e., the means for generating products of various scales, projections, feature combinations, and output media). These advantages do not, however, come without some problems; one of the major problems is the lack of suitable data sets for the applications identified.

II. Existing Digital Cartographic Data Bases

Among the current data sets which include the United States are World Data Bank I, World Data Bank II, and the DIMECO files.* Some of the limitations related to these data sets are; the scale of the source material used for the original digitization, the limited range of features digitized, the currency of information, the accuracy of source documents, and the flexibility of merging these files with thematic data (such as socio-economic factors) available in digital form from various organizations. Using World Data Bank II as an example, source documents were digitized at scales of 1:1,000,000 to 1:4,000,000; digitizing was performed from 1973 through 1977; and only coastlines, islands, lakes, rivers, and international and State boundaries are available for the U.S. Our past experience with these data bases has shown that, while they most likely met the objectives for which they were designed, their content is not sufficiently comprehensive for many common cartographic applications.

* See Appendix for examples of WDB I, II data.

III. The National Small-Scale Data Base Project

The current effort, toward generating a nationwide small-scale digital data base, is intended to address several of these shortcomings. The source materials used for the digitizing, the General Reference Maps of The National Atlas of the United States of America, are at a scale of 1:2,000,000--the only exception being Alaska drainage (originally compiled at 1:1,000,000). The content of the final data base will include political boundaries (State and county level), Federal lands, transportation network (roads and railroads), hydrographic features (streams and water bodies), populated places, miscellaneous cultural features (airports, etc.), and hypsography. The source material is to be updated immediately before the digitizing phase, thus providing the most current information available. The data is structured to provide the adaptability required for plotting general-purpose base maps for use in portraying thematic data as well as singular-purpose base maps for scientific investigations. The coding system used is intended to allow the production of a variety of smaller scale products with varied content using minimal manual interaction.

Plans are to produce the data in two formats. The first will be a topologically-structured data set which will provide a useful geographic reference system for displaying a wide range of thematic data, allow for various types of spatial analysis, and furnish a means of automatically producing patterned or solid-fill plates for reproduction. The second is to be a graphically-oriented data set which will be compatible with the Cartographic Automatic Mapping (CAM) program. Although the final data base organization is yet to be decided, it is intended that both data sets will be as machine independent as possible.

IV. The Approach

The coding scheme used in the project is organized to allow the selection of certain classes of features appropriate for a map given its scale and intended use and/or purpose. The various features are assigned a numerical rank based on the significance of that feature relative to other features of similar type. The evaluation of a feature's significance is based on criteria established for a given set of features; the criteria varies between

different sets of features. For example, railroads are classified by annual tonnage; roads by designation (Interstate, U.S., etc.).

The work flow for the project is shown in figure 1. All steps in the flow are based on digitizing on a sheetwide, rather than state-by-state, basis in order to minimize edge join problems and other complications. The process begins with the gathering of pertinent source documents used in the compilation process. The digitizing source is then updated using traditional cartographic techniques. Following an edit of the updated compilation, the features on the source documents used for digitizing are submitted for ranking and coding. After an edit of the codes to insure uniformity across the entire data base, the features are digitized in a topological fashion indicating nodes, areas and lines. As the digitization of an overlay of features is finished, the data is processed through the USGS Topographic Division's topological data-structuring program, UCLGES (Unified Cartographic Line Graph Encoding System). Through a combination of computer printouts and plots, the attributes and alignments of various features are checked for accuracy and correctness. After the data has been captured, an automatic check is performed to assure that all indicated topological relationships among the data are correct. Corrections are made when errors in the topology are found, and the topological file is archived. The final step is the conversion of the archived file x-y coordinates to geographic coordinates in a format compatible with CAM. It is intended that topological files for the individual sheets will eventually be joined into one master data base. The target completion date for the individual sheet files is April 1980. It is intended that both the topological and the graphical files will be available through the National Cartographic Information Center.

V. Concluding Remarks

The task of creating a nationwide digital cartographic data base, as presented, is singular in its attempt at completeness, flexibility and independence of application. Moreover, a precise system of attribute codes related to basic cartographic data has been defined, and will be offered as a standard to the user community. The development of software to process, merge and transform digital cartographic information into a topologi-

cally-structured data base is underway, and the successful completion of this project will demonstrate the capability for developing a truly comprehensive digital cartographic data base.

Continuing research in the design of data base structures and the development of additional software for off-line processing are needed to establish a system that can provide an effective means for updating existing maps and digital data bases, as well as to facilitate the production of a broad range of thematic maps/graphics. A number of government and other institutions have identified their requirements for digital cartographic support for several time-critical applications related to their resources management programs. The ability to effectively administer the increasing array of projects related to the development and prudent use of our nation's natural resources appears to pivot on the availability of accurate, current, and complete base map information.

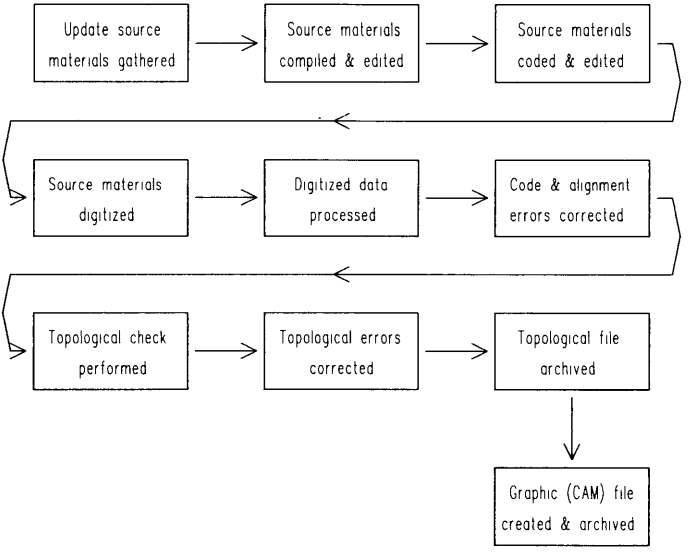
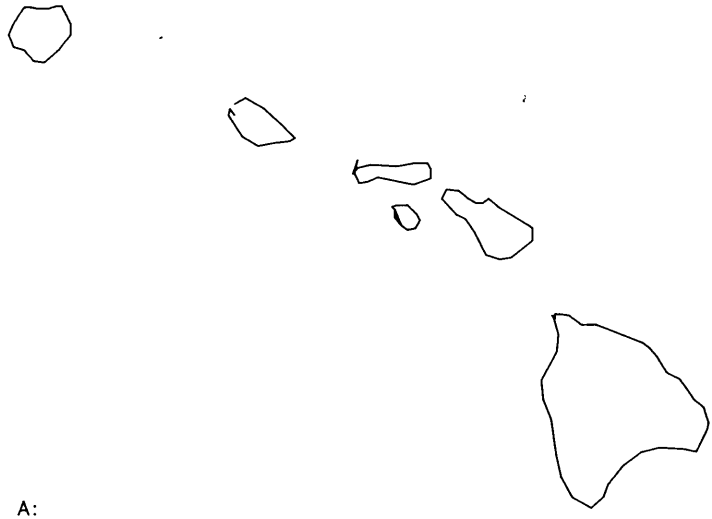
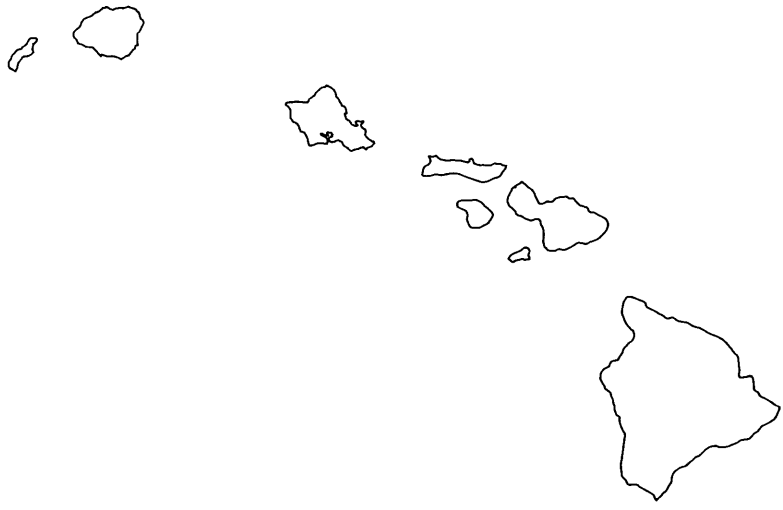


Figure 1



Appendix A:

HAWAII: World Data Bank I



Appendix A:

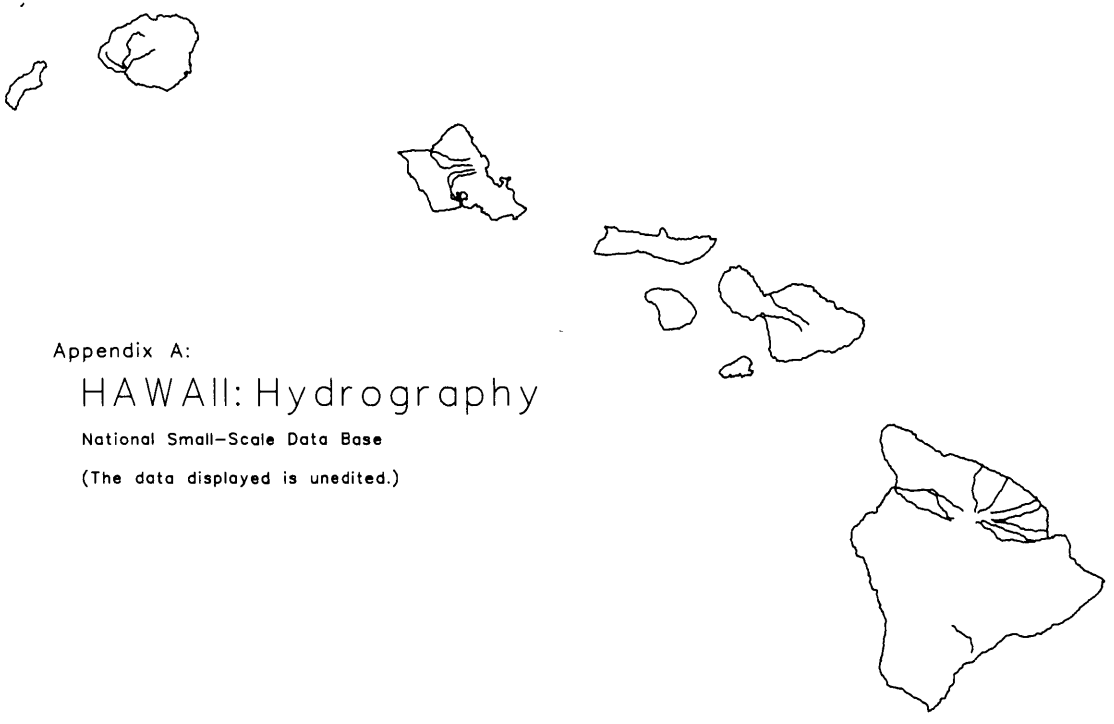
HAWAII: World Data Bank II

Appendix A:

HAWAII: Hydrography

National Small-Scale Data Base

(The data displayed is unedited.)



Appendix A:

HAWAII: Roads

National Small-Scale Data Base

(The data displayed is unedited.)