THE COLLECTION AND ANALYSIS OF NATURAL RESOURCE DATA IN THE 1980s

Doyle G. Fredrick
and
K. Eric Anderson
U.S. Geological Survey
102 National Center
Reston, Virginia 22092

ABSTRACT

In an era of increasing need for the discovery and development of our natural resources, the U.S. Geological Survey has an important role to play. The accomplishment of this role will require improved efficiencies in the storage, manipulation, and presentation of data, but that task will be complicated by current personnel and budget constraints. Increasing reliance upon technological innovations, therefore, will be essential. The Survey is already providing topographic map information in digital form and exploring means of automating more of the mapping process. Computer modeling has become a common tool of Survey geologists and hydrologists. Now these scientists are working with Survey cartographers to combine data bases in order to efficiently produce comprehensive compilations of earth science data. The use of computer graphics in all forms has also increased and is now commonly used for the analysis and presentation of data required in policy decisions. Because the Survey anticipates an increasing demand for natural resource information in the future and recognizes the many advantages of recent technical advances, increased emphasis will be placed on the use of telecommunications, automated data collection, computer-based analysis, computer graphics, and automated mapping in the remaining years of this decade.

INTRODUCTION

The United States faces many difficult resource and land management issues involving: energy and mineral exploration and prudent development, assessment of the Nation's water resources (quality and quantity), and natural hazards examination, among others. Maps and mapping have always played an important role in dealing with issues of this type and will continue to serve in a vital capacity in the future. As we near the end of this century, several problems face the mapmaker, the earth scientist, and the land manager who would meet the challenges of providing data in these vital areas:

- Location of undiscovered resources requires the application of sophisticated techniques commonly requiring the processing of large amounts of data.

Also, this will require a look at lands that are not altogether accessible--wilderness areas, OCS, etc.

- Increasing competition for, and public discussion about, the most appropriate use of land and resources require detailed consideration of the economic and environmental impacts of proposed actions, i.e., to leave open for exploration and development or set aside.

- Concern about the consequences of action or lack of action requires that analysis be done completely and quickly. This necessitates extensive modeling of natural process and proposed action. Fortunately, technology has now provided the scientists and the decisionmaker with a vast array of powerful tools to aid the process considerably.

U.S. GEOLOGICAL SURVEY ROLE

In order to provide a framework for a discussion related to management of natural resource data and the application of computer data bases and automated cartography, it might be helpful to review briefly the mission of the Geological Survey.

The Geological Survey was established in 1879 and charged with the responsibility for the "classification of public lands and examination of the geological structure, mineral resources, and products of the national domain." Over the years the evolution of the earth sciences, the need to carefully manage the Nation's nonrenewable resources and to find new sources of critical energy and mineral commodities, and mounting concern over man's impact on the environment have added numerous other duties including geographic research, hazards duties, topographic and geologic mapping, and water resources assessments. The Survey is an impartial research agency that gathers, interprets, and distributes data in order to advance scientific knowledge of the Earth so that managerial decisions related to natural resources can be based on objective information.

To meet the ever increasing needs for the more sophisticated uses of our data, scientists at the Geological Survey are embracing computer assisted cartography as an essential element of their work. Also, sophisticated computer analysis is becoming a common practice among those who use our information (inside the Geological Survey as well as outside).

As the primary Federal earth science, research, and fact-gathering agency, the Geological Survey provides information for sound planning and decisionmaking as it has for many years. Now much of that information is provided on computer tapes. The computer data bases of the Survey must meet the needs of many as a source of information and to support analysis, and the cartographic data base should also provide the connecting link to permit the integration of many disparate data sets. In addition, as a major mapping agency we must assure that data bases can be used to enhance our capability and capacity to prepare and revise maps.

THE CHALLENGE OF THE 1980's

As both a producer and a user of cartographic and earth science data, the Geological Survey, like many other agencies, faces major challenges to establish and manage these data base and to develop standards to allow easy access and convenient use. Already, there are large data bases in existence and more are on the way. In addition, we are not the only organization to collect and store data.

To assure that the cartographic, geologic, hydrologic, and other related information can be readily used to produce comprehensive insights required by today's researchers and decisionmakers, the Geological Survey and other mapping organizations are obligated to insure that high standards of accuracy and integrity are maintained in the design, development, and management of these data bases. We must provide: the basic foundation to enable the integration of many data types and sources, and truly multiple purpose data bases that will serve all.

In recognization of our obligations in these areas, we have set out to:

- Manage our internal data bases in a consistent, but not oppressive, way. We may now have as many as 250 separate data bases which require some oversight. We have had a Data Base Administrator for some time, and a great deal of progress has been made. We are now reorganizing our computer operations to insure additional oversight for the Bureau's information systems.

- Develop standards for earth science information that can be used both internally as well as externally.

  The Geological Survey is carrying out a leadership role in the development of standards for earth science data through a memorandum of understanding with the Bureau of Standards. This is especially appropriate considering the Survey's role as the Nation's mapper, which provides that core cartographic data set to which all others must tie. Committees have been established within the Survey and the Department of the Interior to examine and review current activities and develop recommendations to the Bureau of Standards for Federal standards on earth science data. The first set of standards to go through this process have just been published as Circular 878-A: <u>Codes for the Identification of Hydrologic Units in the United States and the Caribbean Outlying Areas</u>. In addition, the Survey has supported the establishment of a National Committee on Digital Cartographic Standards under the auspices of the American Congress on Surveying and Mapping. These activities represent the beginning stages of an ongoing effort to provide the necessary foundation to insure the usefulness and applicability of automated earth science data now and in the future.

None of us has the power to legislate standards--hopefully we can be convincing regarding the benefits.

- Properly design the data base to assure multi-purpose use.

- Insure user input into all decisions related to data input, etc.

DATA BASES (EXAMPLES)

Digital Cartographic Data:
Recognizing the essential role of cartographic data for resource studies and considering our responsibilities as the civilian land mapping organization, we have undertaken the development of a national digital cartographic data base. This activity encompasses all of the complexities, concerns, and obligations discussed earlier. To serve national needs, we must insure that our data base:

- Provides data that people need.
- Interfaces with almost every data bases around (proved the link).
- Contains enough information for complete analysis.
- Is easily accessible.
- Can be developed at a reasonable cost.
- Accommodates data from others.
- Support the mapping process.

The National Digital Cartographic Data Base encompassses planimetric and elevation data from the standard topographic maps of the Survey. The principal categories of planimetric data being included are: the Public Land Survey System; county, State, and Federal boundaries, highways, roads, trails, and rivers, streams, and water bodies. The primary scale of collection is 1:24,000, but pilot projects are examining the efficiency of data collection at other scales. Elevation data are being collected at 1:24,000-scale, but the Survey is also distributing elevation data at 1:250,000-scale collected by the Defense Mapping Agency. One of the most notable achievements of the Survey's digital cartography program is the completion of a national series of maps at 1:2,000,000-scale.

In response to the need for digital cartographic data suitable for applications at regional or national scales, the Survey has introduced a data base at the 1:2,000,000-scale. This data base was constructed from the general reference maps of The National Atlas of the United States of America. It includes political boundaries (State and county level), Federal lands, transportation networks (roads and railroads), and hydrographic features (streams and water bodies). The source materials were updated immediately before the digitizing phase, thus providing the most current information available.

Other forms of cartographic and geographic data are also becoming available in digital form. Land use and land cover data at 1:250,000-scale (some at 1;100,000-scale) are being produced along with digital data on county boundaries, census subdivisions, hydrologic units, and Federal lands. These will provide a baseline of information useful to land use planners and resource managers. Also, a Geographic Names Information System, incorporating over 2,000,000 names of geographic places and features in the United States, has been completed. These names have been extracted from the largest scale map covering an area and, in addition to serving Federal, State, and local requirements, are playing a role in the automation of the mapping process.

Not only are digital cartographic data important to the analytical functions of the Survey and the earth science community, but they are essential to the automation of the mapping process. Automation is playing an increasing role in the Survey's mapping activities. A major development has been the collection of digital information directly during the stereocompilation process. Research and development efforts have led to the implementation of a digital cartographic software system that has proven to be very cost-effective. The data collection subsystem consists of stereoplotter instruments fitted with digitizing units. Each station is equipped with voice data entry systems to improve operator efficiency. This system has led to improved efficiencies in the map production process in addition to collecting certain categories of digital data. Ongoing research is directed towards the automated production of map separates and the development of effective means of digital map revision.

WATER RESOURCES APPLICATIONS

The Geological Survey has accumulated extensive data bases describing the water resources of the United States. Such data bases are essential for the effective monitoring of trends in water quality and availability. Programs like the National Stream Quality Accounting Network provide the information to account for the quality and quantity of water moving within the United States, to develop a large-scale picture of how stream quality varies from place to place, and to detect changes in stream quality over time. A crucial role is played by the Water Data Storage and Retrieval System (WATSTORE), which includes data on surface and ground water supplies and quality, chemical analyses, peak flows, and water use. This data base serves numerous State, local, private, and Federal agencies in their efforts to develop and manage water resources. These hydrologic data are used not only in determining the adequacy of water supplies but also in designing dams, bridges, and flood control projects, in allocating irrigation projects, in locating sources of pollution, and in planning for energy development.

Columbia River Irrigation Project:
A recent project that illustrates the importance of such hydrologic information as well as the value of integrating different types of data is an irrigation planning project for the Columbia River. The Umatilla River basin in Oregon was analyzed in a cooperative effort with the U.S. Army Corps of Engineers to determine its potential for irrigation.

Irrigation demand results in the largest consumptive use of Columbia River water, so it must be managed both to maximize agricultural benefits and to insure the wisest use of water and power supplies. Hydropower generation, in particular, is most adversely affected by irrigation development. The development of 10,000 acres of irrigation causes an average annual loss of $310,000 of electrical energy. Future irrigation development must take place in areas where the loss of power revenues is more than compensated for by the benefits derived from the new irrigation.

To develop a solution to this problem, several types of data were required. Pumping plant data were used to describe the location and efficiency of pumps to provide the water. Detailed soil survey information was used to identify soil irrigability ratings. Elevation data were required to compute slope information needed in the irrigability potential study and to provide estimates of the potential energy consumption for irrigation. In addition, information on land cover and land ownership was utilized to conduct the irrigation potential analyses.

These data were integrated in a complex analytical model that recognized the spatial variations inherent in the problem. A composite mapping technique was applied that permitted the variables to be precisely registered to a base map and overlaid. The composite map provides ratings of development potential based on the physical characteristics of the landscape, the economic impact of energy costs, and government restrictions prohibiting irrigation. The maps produced in this study provide a starting point for planning future regional uses of land, water, and power resources. From this starting point, more cost-effective field studies will be possible, and more resource-efficient water development decisions can be made.

GEOLOGIC APPLICATIONS

Computerized Resources Information Bank:
The extensive work of the Geological Survey in the collection and analysis of data in support of energy and mineral discovery and development has led to the establishment of several major data base systems. One of the earliest of these was the Computerized Resources Information Bank,

which still provides a means for rapidly organizing, summarizing, and displaying information on mineral resources. One of the more recent developments is the National Coal Resources Data System, which provides for on-line analyses of data relating to the location, quantity, and physical and chemical characteristics of coal. These systems enable a wide range of studies to be conducted that extend our knowledge of energy and mineral resources more rapidly and more completely than would otherwise be possible. The Survey also manages the National Earthquake Information Service, which utilizes digital data recording devices all over the world to gather and analyze seismic events. The resulting data base of seismic activity provides a valuable international tool for geophysical research.

National Coal Resources Data System:
The National Coal Resources Data System incorporates resource and chemical data on an areal basis (county, township, or coal field) and point source data including field observations and drill-hole logs. Such variables as bed thickness, name, moisture, ash, sulfur, major-, minor-, and trace-element content, heat value, and characteristics of overburden, roof, and floor rocks. The system is designed to incorporate data from many sources including the Geological Survey, the Department of Energy, State geologic agencies, and private coal companies. This flexibility is a key feature of the system that will allow for data update through ongoing field and mining activities to permit reassessment of the resource estimates based on the most current information.

The system is used to calculate coal resources and quality for any geographic area, coal bed, or series of coal beds, to calculate and discriminate overburden categories, and to locate, through computer analysis, the more desirable portions of a coal deposit. The system can reproduce the basic data as maps using digital plotters. It derives and plots structure sections and structure contours, as well as ispoach maps for any analyzed chemical element and ratio maps between any desired characteristics. These capabilities mean that the National Coal Resources Data System has a crucial role in the identification and asessment of energy resources from the country.

Alaska Minerals Resource Appraisal Program:
The need for better information on mineral resources has led to the establishment of programs such as the Alaska Minerals Resource Appraisal Program. A recent project under that program involved the use of digital image analysis techniques to develop and apply a conceptual model of porphyry-type copper mineralization.

The 1° by 3° Nabesna quadrangle in east-central Alaska was selected as the study area because copper mineralization is known to occur in the area, and a large amount of data was available. A wide range of data was utilized, including geophysical, geochemical, topographic, geologic, and Landsat multispectral scanner data. All the data were registered to a map base and integrated into a raster data base.

Various processing techniques were applied to the data in order, first to identify relationships among the variables by correlation and second to develop quantitative weights for a mineral potential model. The resultant model contained 10 parameters that could be applied to the data base to quantify the likelihood of mineralization.

The applications of the model resulted in the identification of three areas most likely to contain porphyry copper mineralization. Two of the areas are known occurrences of mineralization, but the third case, which had the highest potential rating, had been unknown until this study. Field studies in August 1981 confirmed that all three areas identified by the model have significant copper potential.

CONCLUSION

We are faced with challenging scientific, data handling, and management problems if we are to fully meet the identified needs at reasonable cost. As the information revolution leads to the creation of more and larger data bases with an ever-increasing body of hardware and software tools for their use, care must be exercised. It is important to remember that the primary motivation for automation data collection and analysis was to use the tremendous power of the computer to integrate and synthesize large quantities of disparate data. If each data base is allowed to develop in isolation, the ability to integrate data will be limited. Problems are emerging that require the use of data from many different data bases. The time to address these problems is now, during the creation of these data bases. It is necessary to plan and develop systems so that the interchange of data among them is facilitated and not hindered. As part of this process, it is necessary to develop standards for computer-readable data to insure that data exchange is efficient. This is not an easy task considering the pressure to move-out smartly and reduce expenditures.