INTERACTIVE MANAGEMENT AND UPDATING OF SPATIAL DATA BASES

P. French and M. Taylor
Resources Planning Associates, Inc.
P.O. Box 2562, East Hill
Ithaca, N.Y. 14850 USA

1. INTRODUCTION

The decision making process, whether for power plant siting, load forecasting or energy resource planning, invariably involves a blend of analytical methods and judgment. Management decisions can be improved by the implementation of techniques which permit an increased comprehension of results from analytical models. Even where analytical procedures are not required, decisions can be aided by improving the methods used to examine spatially and temporally variant data.

This paper will discuss how the use of computer aided planning (CAP) programs, and the selection of a predominant data structure, can improve the decision making process.

2. SPATIAL DATA BASES - A NECESSITY

Modern society imposes a large number of constraints on the planning and management of those goods and services which affect the general welfare. One such example is the need to develop environmental impact statements for large scale construction projects. In order to assess future impacts which might result from a proposed project, a baseline condition first must be established. Data must be obtained to define the present state of the air, water, ecological and other natural resources, as well as the current economic, social and demographic characteristics. All of these types of information can be represented conveniently by spatial data bases.

In addition to presenting data as a baseline, some of these data also may be required as input to analytical models. These models will generate results (which in themselves may constitute a spatial data base) which, together with the baseline, can be used in the decision making process of project evaluation.

Defining the need for spatial data is easy. Obtaining all that is required can be difficult. Everyone collects and stores spatial data; the amount and level of sophistication vary from NASA's LANDSAT to the Department of Public Works supervisor who, after 35 years on the job, is the only one who knows the location of every storm sewer in the city. The complexity of most sources of data, however, lies somewhere between these two extremes, and invariably involves hardcopy maps on paper, Mylar, or linen.

The numerous sources of maps available create two problems: redundancy and accuracy. Since very few units of government have a Department of Cartography or other centralized spatial data base repository, it is incumbent upon each agency or department to both obtain and maintain the data it requires. Since jurisdictions (and sometimes functions) overlap, redundancy occurs. Due to differences in the frequency and level of updating, this redundancy eventually leads to discrepancies. Therefore, in addition to the problem of where to go to obtain data, one also may have to contend with the problem of which source to believe.

Problems associated with the suitability and accuracy of necessary data can be reduced through the careful development, management and updating of comprehensive data bases. Although every planner and manager may dream of the perfect data base - one that has every conceivable type of information, 100% accurate, and up to date - this "wealth" of information could prove valueless if the decision maker is unable to manipulate it, comprehend it, or assess its significance. As the volume of accumulated data increases, it becomes increasingly difficult to maintain and rely upon a system based solely upon paper maps.

One method found to be extremely useful in assisting decision makers in understanding data complexities and interactions relies upon the use of interactive color computer graphics to both help generate and display complex spatial data bases. While the ability to quickly and easily analyze and interact with spatial data is important, no less important is the display format selected for presenting the digital data. Developers of digital cartographic systems should attempt to provide a data base that, when displayed, looks as good as a high quality paper map.

3. SPATIAL DATA BASE STRUCTURES

In order to provide a data base which is easy to manipulate and comprehend, the structure selected for the data base is important. The conventional structures involve points, lines, areas or, as a simplification of a polygonal structure, grid cells. The different relationships that can exist between geographical entities, and the various searching and sorting operations which can be encountered, have been discussed by Nagy and Wagle (1979). Rather than reviewing those discussions, some of the advantages and disadvantages associated with each type of structure will be presented.

3.1 Point Data Structures

Perhaps the greatest advantage to using a point data structure is the relatively small amount of storage required. However, exclusive use of a point system limits the information portrayed to discrete locational data: utility poles, manholes, geologic core samples, well locations. For certain types of point data (such as elevation or depth to groundwater) even a sparse data set can be sufficient for determining probable values at unknown locations if appropriate techniques such as Krieging (Delfiner and Delhomme, 1975) are used. Unfortunately, if connectivity between points exists, it is not retained in a point data structure. This contributes to the difficulty of trying to geo-reference the points within the area of interest.

3.2 Line Data Structures

As with points, the use of line primitives is relatively low cost from the standpoint of storage and processing requirements. If linear representations of curvilinear segments are used, the storage and processing requirements are further reduced. Also, use of linearized primatives can provide a de facto point data structure, as long as the points are associated with the particular network (transmission towers, substations, bus stops).

While line data structures provide more information than point structures from the standpoint of visual geo-referencing, a line or line-point data structure is inadequate for representing most geographical data bases.

3.3 Area Data Structures

The use of polygonal or area data structures literally provides a new dimension in the ability to portray information, since zones of an attribute type or value now can be described and quantified. If the polygon boundaries are stored as linear line segments, along with the two attribute types they separate (thus preventing duplication of boundary definition for adjacent polygons), the storage requirements can be modest if the area/perimeter ratio is large.

From the standpoint of visual geo-referencing, polygonal data structures are less informative than linear data structures when viewed apart from everything else. Their utility is not apparent until viewed in the proper spatial context: a utility service area displayed with the transmission network, or as an overlay on an appropriately scaled map. When combined with other data structures, polygonal data can be very informative. However, while the ability to answer a question such as "Locate and quantify the miles of transmission lines within a number of service districts which provide areas not meeting certain population density requirements" becomes possible, arriving at the answer can be computationally intensive.

3.4 Grid Cell Data Structures

Given a fixed grid system, each grid cell occupies a known geographical location, has a known area, and can have multiple attributes associated with it. These characteristics facilitate the overlaying of and determination of intersecting areas of different attributes. This type of data structure can be interacted with and updated easily and, as will be shown, can provide the visual geo-referencing necessary to make the system user friendly.

Another advantage of a grid cell based data structure is the link that can be made with video technology. Grid cell data not only can be displayed on video monitors, but also can be created using video equipment.

The disadvantages of using a grid cell system primarily arise from the size of the cell. The smaller the cell size, the greater the ability to approximate a point-line-area data structure. However, if the cell size is too small, data input methods and storage requirements become critical. Indeed, manual definition of cell data may be too labor intensive and prone to error for very small cell structures. Conversely, data for large cell areas become easier to input and store, but at the expense of detail and accuracy: points and lines no longer are represented realistically.

4. VIDEO BASE MAPS: A PREDOMINANT DATA STRUCTURE

Certainly no single data base structure provides all of the advantages and none of the disadvantages associated with a geographic information system. Where man-made attributes have to be considered, point and network data become commonplace. For analyses required for site planning purposes, the use of some type of polygonal data is virtually assured. Given that different data structures are required to fully represent the requirements of geographic information systems, consideration should be given to the selection and creation of a predominant data base structure. The selection should be one which provides for the best visual quality of the displayed data, enhances the utility of the other data base structures, and is easily updated.

In our work with water and other natural resources planning, we have found that the use of computerized video base maps satisfies our criteria for a predominant data base structure. The computerized grid cell base maps, created by video digitizing aerial photography, U.S.G.S. 7 1/2 minute quad sheets, or other suitable hardcopy, contain the complex "extras" - the highschool track, the interstate's cloverleaf, parking lots, fields under cultivation - which provide for instant recognition and visual georeferencing. If additional data to describe the location of points or networks are required, they should be contained in, and extracted from, supplemental point-line data bases and registered onto the video base map. In this manner, the video base map enhances the utility of the supplemental data through improved comprehension of spatial relationships.

The use of grid cells as the predominant data structure, and color computer graphics as the display medium, enable utilization of the many private and Federal data bases which are becoming digital. Examples include worldwide LANDSAT coverage, digital terrain models which are being developed for the entire U.S., and Census data. All of these data bases, which are grid cell, point, line or polygon based, can be geo-referenced and displayed over the video base maps.

In addition to improving a person's ability to mentally geo-reference data, the use of video technology also permits data to be displayed in color. Some geographic information systems which utilize color have only destructive color displays. That is, when displaying color-coded polygonal information, any underlying data are masked by the addition of color. Since destructive displays also inhibit the ability to geo-reference data, non-destructive display techniques should be used whenever possible. These techniques have been developed and incorporated into a computer-aided resources planning package by the authors. In several test applications this package has demonstrated the value of these techniques in the planning process. Figures 1 and 2, illustrating typical displays from this package, are based upon the Rochester East and Webster, NY 7 1/2 minute quadrangle maps.

5. CREATING/UPDATING SPATIAL DATA

The ability to define and update complex, large scale spatial data bases also is facilitated by the use of video techniques. Original color (or black and white) hardcopy maps or photographs can be quickly converted to a digital grid cell structure. Since the data can be defined at up to video rates (1/30 second for a 525 line image), large amounts of information can be captured quickly in a non-labor intensive manner. Supervised classification techniques then can be used to convert three channel (red, green, blue) data to a single classified channel. With all pertinent spatial information available in a single channel, the ability to rapidly interact with the data is improved.

Interactive updating of video maps can be performed using any of several methods. Point, line or polygon data can be entered using a digitizing pen and tablet and locating the position visually, or by registering a source map on the tablet to the video image and tracing the appropriate information. Updates also can be made using the many available sources of digital information which have been suitably classified and registered to the video base map.

6. HARDWARE AND COST

There are many people who will dismiss the use of grid cell digital cartographic data bases either for the reason mentioned in section 3.4 (too large a storage requirement), or because of a lack of resolution. Concern over the latter point probably is justified with the use of commonplace raster graphics hardware, which has 480 x 512 displayable cells and provides a resolution of approximately 0.046" for a 7 1/2

Fig. 1. Results from a flood analysis package are displayed (blue) over the 1 channel (8 bit) video base map.

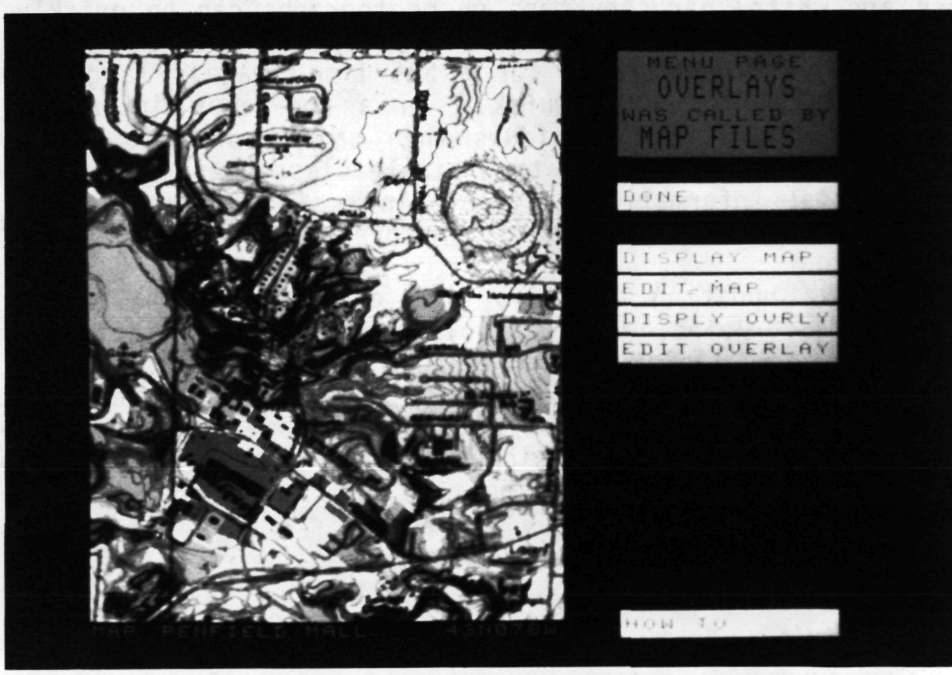

Fig. 2. Color-coded land use information, derived from polygonal data, are registered and displayed over the 1 channel video base map.

minute quad sheet. However, hardware capable of displaying 1000 x 1000 cells are readily available; units that display 2000 x 2000 cells, each with 8 or more bits of color, will be available in the next 3-5 years. A device with 2000 x 2000 resolution corresponds to a quad sheet resolution of approximately 0.01 inches/cell, considerably better than the National Map Accuracy Standards (Monmonier, 1977).

The concern over the vast amounts of data storage required for grid cell maps also will subside in the next three to five years. As computer-compatible read/write video disc equipment becomes available, digitized maps (areal photographs, terrain models, etc.) will be stored and retrieved quickly. Even using today's video standards, a 2000 x 2000 cell composited image could be displayed in approximately 1/2 second. At this resolution, digital coverage for each of the more than 40,000 7 1/2 minute quadrangles in the U.S. could be stored on 3 movie length video discs.

Equipment needed to provide the interactive capabilities mentioned can be obtained for as little as $35,000. While this price currently may be exclusionary for some of those groups which most utilize spatial data - local, county, regional and state planning agencies - the cost of computer hardware continues to decrease. Also to be considered are the cost and availability of the necessary software. Although the cost of software could approach that of hardware, we foresee a number of microcomputer-based turnkey video cartographic data systems being offered for less than $40,000 in the next 3 to 5 years.

7. CONCLUSIONS

The ability to visually geo-reference complex spatial data and spatially oriented results from mathematical models is an important characteristic required of many planning processes. The use of a predominant grid cell data structure, based upon the use of video images and color computer graphics, helps to provide the geo-referencing capability. This data structure and the computer and video hardware facilitate interactive creation, management and updating of complex spatial data. The ability to rapidly manipulate large amounts of spatial data, perform analyses with that data, and display color coded results which can be visually geo-referenced to that data, help to place in proper perspective the role of judgment in the decision making process.

8. ACKNOWLEDGEMENTS

The computer-aided planning package used to generate Figures 1 and 2 was developed at Cornell University by the authors under the direction of D.P. Loucks; funding was provided by a grant from the National Science Foundation.

REFERENCES

Delfiner, P. and Delhomme, J.P. "Optimum Interpolation by Krieging." In Display and Analysis of Spatial Data, J.C. Davis and M.J. McCullogh (Eds.), John Wiley & Sons, Inc., New York, 1975.

Monmonier, M.S. Maps, Distortion, and Meaning. RP 75-4, Association of American Geographers, Washington, D.C., 1977.

Nagy, G. and Wagle, S. "Geographic Data Processing." ACM Computing Surveys, V. 11, No. 2, June, 1979, pp. 139-181.