Reprinted from

Eighth International Symposium

Machine Processing of

Remotely Sensed Data

with special emphasis on

Crop Inventory and Monitoring

July 7-9, 1982

Proceedings

Purdue University
The Laboratory for Applications of Remote Sensing
West Lafayette, Indiana 47907 USA

CARTOGRAPHIC MODELING: COMPUTER-ASSISTED ANALYSIS OF EFFECTIVE DISTANCE

J.K. BERRY

Yale University/School of
Forestry and Environmental Studies
New Haven, Connecticut

C.D. TOMLIN

Harvard University/Graduate School of Design
Cambridge, Massachusetts

ABSTRACT

Cartographic models addressing a wide variety of applications are composed of fundamental map processing operations. These primitive operations are neither data base nor application-specific. By organizing the set of operations into a mathematical-like structure, the basis for a generalized cartographic modeling framework can be developed. Among the major classes of primitive operations are those associated with reclassifying map categories, overlaying maps, determining distance and connectivity, and characterizing cartographic neighborhoods. This paper establishes the conceptual framework of cartographic modeling and uses the techniques for distance characterization as a means of demonstrating some of the more sophisticated procedures of computer-assisted map analysis. Most of the techniques described have been implemented as part of the Map Analysis Package developed at the Yale School of Forestry and Environmental Studies.

I. INTRODUCTION

Most computer-oriented geographic information systems include processing capabilities which relate to the encoding, storage, analysis and/or display of cartographic data. The analytical operations used in many of the currently available systems[4] are embedded within application-specific contexts. By extracting and organizing primitive operations in a logical manner, the basis for a generalized cartographic modeling structure, or "map algebra" can be developed.[7,8] In this context primitive map analysis operations are analogous to traditional mathematical operations. The sequencing of map operations is thus similar to the algebraic solution of equations to find unknowns. In this case, however, the unknowns represent entire maps. The conceptual structuring of these primitive operations provides a basis for a modeling structure which accommodates a wide variety of computer analyses. This paper describes this conceptual framework and uses the techniques for assessing effective cartographic distance and connectivity to demonstrate some of the more sophisticated procedures and considerations of cartographic modeling.

II. DATA AND PROCESSING STRUCTURES

In order to use primitive operations in a modeling context a common data structure and a flexible processing structure must be used. The variety of mappable characteristics likely to be associated with any given geographic location may be organized as a series of spatially registered computer-compatible maps (Figure 1). In this way, a data base may be defined as a set of maps registered over a common geographic area; a map, or "overlay", may be defined as a set of mutually exclusive but thematically related categories; and a category, or "region", may be defined as a thematic value associated with a set of geographic locations, or "points".[9] While this is certainly not the only way to represent cartographic data[5] it is one which relates directly and intuitively to traditional graphic techniques involving conventional geographic maps. It is also one which is common to many computer-oriented geographic information systems. Differences among these systems relate to either the way in which thematic attributes are represented (i.e. numerically, literally or in binary form) or to the way in which locational attributes are coded (i.e. rectangular cells, polygons, line segments, etc.). While these differences are significant in terms of implementation strategies, they need not affect the definition of fundamental cartographic techniques.

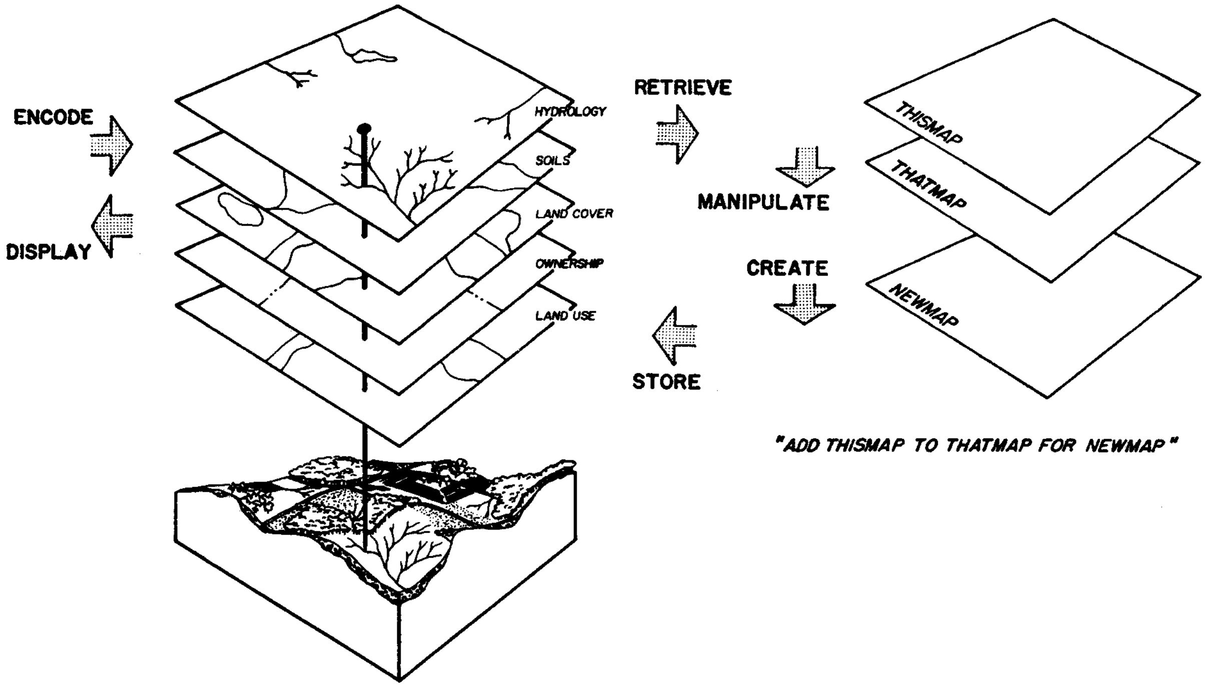


FIGURE 1. A computerized geographic information system consists of a data base of spatially registered maps and procedures for the encoding, storage, analysis and display of mapped data. The cyclical process of cartographic modeling involves retrieving one or more maps from the data base which are used to create a new map. This new map then becomes part of the data base and is available for subsequent processing.

If primitive operations are to be flexibly combined each much accept input and generate output in the same format. Using a data structure as outlined above this may be accomplished by requiring that each analytic operation involve:

* retrieval of one or more maps from the data file;
* manipulation of that data;
* creation of a new map whose categories are represented by thematic values defined as a result of that manipulation; and,
* storage of that new map for subsequent processing.

The cyclical nature of this processing structure (Figure 1) is analogous to the evaluation of "nested parentheticals" in traditional algebra. The logical sequencing of primitive operations on a set of maps forms a cartographic model of a specified application. As with traditional algebra, fundamental techniques involving several primitive operations can be identified (e.g. a "travel-time" map) that are applicable to numerous applications. The use of primitive analytical operations in a generalized modeling context accommodates a variety of analyses in a common, flexible and intuitive manner. It also provides a framework for instruction in the principles of computer-assisted map analysis that stimulates the development of new techniques and applications.[3]

III. FUNDAMENTAL OPERATIONS

Within the data and processing structures outlined above, each primitive operation may be regarded as an independent tool limited only by the general thematic and/or spatial characteristics of the data to which it is applied. From this point of view, four major classes of fundamental map analysis operations may be identified (Table 1). These involve:

* reclassifying map categories;
* overlaying maps;
* determining distance and connectivity; and,
* characterizing cartographic neighborhoods.

A brief discussion of these fundamental classes is presented below. More detailed discussions are presented in several of the references noted at the end of this paper.[1,2,7]

The first of the four major groups of cartographic modeling operations is the simplest and, in many ways, the most fundamental. Each of the operations

Table 1. Fundamental Map Analysis Operations

FUNDAMENTAL CLASSES	FUNCTIONAL BASIS	EXAMPLE OPERATIONS
RECLASSIFYING MAP CATEGORIES- operations for reclassifying map categories involve reassigning thematic values to the categories of an existing map as a function of the initial value, the position, the size or the shape of the spatial configuration associated with each category.	+ INITIAL VALUE	+ ARBITRARY SCHEME (relabeling, isolating, aggregating) + ORDERING SCHEME (ranking, weighting) + MATHEMATICAL RULE (Isolating, arithmetics with constants)
	+ POSITION	+ SPATIAL LOCATION (reference coordinates, line orientation)
	+ SIZE	+ AREAL EXTENT + VOLUME
	+ SHAPE	+ BOUNDARY CONFIGURATION (edgeness, irregularity) + SPATIAL INTEGRITY (interior holes, fragmentation)
OVERLAYING MAPS-- overlay operations result in the creation of a new map where the values assigned to every location on that map is computed as a function of independent values associated with that location on two or more existing maps.	+ LOCATION-SPECIFIC	+ PERMUTATION (category combination) + DIVERSITY COUNTING + RELATIVE PROPORTION (frequency) + ORDINAL SELECTION (maximize, minimize, median, etc) + MASKING/SIEVING + LOGICAL COMBINATION (union, intersection) + ARITHMETIC COMBINATION (add, subtract, divide, etc) + WEIGHT AVERAGING
	+ CATEGORY-WIDE	+ PERMUTATION (category combination) + DIVERSITY COUNTING + RELATIVE PROPORTION (frequency, uniqueness, overlap) + ORDINAL SELECTION (maximize, minimize, median, etc) + LOGICAL COMBINATION (union, intersection) + ARITHMETIC COMBINATION (add, multiply, etc- commutative only)
	+ MAP-WIDE	+ SOLUTION OF MATHEMATICAL/STATISTICAL RELATIONSHIPS
DETERMINING DISTANCE AND CONNECTIVITY- operations for measuring cartographic distance involve the creation of new maps in which the distance and route between points can be expressed as simple Euclidean length or as a function of absolute and/or relative barriers.	+ SIMPLE DISTANCE/PROXIMITY	+ SHORTEST STRAIGHT LINE ("as-the-crow-flies")
	+ WEIGHTED PROXIMITY	+ SHORTEST ROUTE ("as-the-crow-walks") + ABSOLUTE BARRIERS (guide movement) + RELATIVE BARRIERS (expend units of movement)
	+ CONNECTIVITY	+ STRAIGHT LINE (simple distance, intervisibility) + OPTIMAL PATH (steepest downhill path, minimal "cost" route) + OPTIMAL PATH DENSITY (networks)
CHARACTERIZING CARTOGRAPHIC NEIGHBORHOODS- these operations involve the creation of a new map based on the consideration of "roving windows" of neighboring points about selected target locations.	+ SUMMARIZING THEMATIC ATTRIBUTES	+ STATISTICS (total, mean, maximum, etc) + ANOMALLY DETECTION (deviation, proportion similar) + INTERPOLATION (weighted average, nearest neighbor) + MAP GENERALIZATION (surface fitting)
	+ 2-DIMENSIONAL FEATURE ATTRIBUTES	+ NARROWNESS (shortest cord to opposing edges) + CONTIGUITY (individual clumps)
	+ 3-DIMENSIONAL SURFACE ATTRIBUTES	+ SLOPE (topographic slope, differentiation) + ORIENTATION (topographic aspect, direction of movement) + PROFILE (patterns along sequential cross-sections)

involves the creation of a new map by reassigning thematic values to the categories of an existing map. These values may be assigned as a function of the initial value, the position, the size, or the shape of the spatial configuration associated with each category. All of the reclassification operations involve the simple "repackaging" of information on a single map and results in no new boundary delineations.

Operations for overlaying maps begin to relate to the spatial, as well as to the thematic nature of cartographic information. Included in this class of operations are those which involve the creation of a new map such that the value assigned to every location is a function of the independent values associated with that location on two or more existing maps. In simple location-specific overlaying the value assigned is a function of the spatially aligned coincidence of the existing maps. In category-wide composites values are assigned to entire thematic regions as a function of the values associated with the regions contained on the existing maps. Whereas the first overlaying approach conceptually involves "vertical spearing" of a set of maps, the latter approach uses one map to identify boundaries from which information is extracted in a "horizontal summary" fashion from the other maps. A third overlay approach treats each map as a variable; each location as a case and each value as an observation in evaluating a mathematical or statistical relationship.

The third class of operations is one which relates primarily to the locational nature of cartographic information. Operations in this group generally involve the measurement of distance and the identification of routes between points on a map surface. This class of techniques will be discussed in detail in subsequent portions of this paper.

The fourth and final group of operations includes procedures that create a new map in which the value assigned to a location is computed as a function of the independent values within a specified distance around that location (i.e, its neighborhood). The summary of information within the "roving window" of neighboring locations can be based on the configuration of the surface (i.e. slope and aspect), the characterization of contiguous features (i.e. narrowness), or the statistical summary of thematic values (i.e. average value).

IV. DETERMINING DISTANCE AND CONNECTIVITY

Most geographic information systems contain analytic capabilities for reclassifying and overlaying maps. These operations address the majority of applications that parallel conventional map analysis techniques.[6] However, to more fully integrate spatial considerations with contemporary analysis and planning, new techniques are emerging. The concept of distance has historically been associated with the "shortest distance between two points." While this measure is both easily conceptualized and implemented in either a manual or computer-oriented environment, it is frequently insufficient in a decision-making context. A straight line route may indicate the distance as-the-crow-flies, but offer little information for a walking or hitchhiking crow. Equally important to most travelers is to have the measurement of distance expressed in more relevant terms, such as time or cost.

The basis of any system for the measurement of distance requires two components -- a standard measurement unit and a measurement procedure. The standard measurement unit used in most computer-oriented systems is the "grid space" implied by the superimposing of an imaginary uniform grid over a geographic study area. The distance from any location to another is computed as the number of intervening grid spaces. A frequently employed measurement procedure involves expanding the concept of distance to one of proximity. Rather than sequentially computing the distance between pairs of locations, concentric, equidistant zones are established around a loction, or set of locations. In effect, a map of proximity to "target" location is generated that indicates the shortest straight line distance to the nearest target location for each non-target loction (Figure 2).

Within many application contexts, the "shortest" route between two locations may not always be a straight line. And even if it is straight the Euclidean length of that line may not always reflect a meaningful measure of distance. Rather, distance in these applications is best defined in terms of movement expressed as travel time, cost, or energy that may be consumed at rates which vary over time and space.

Distance-modifying effects may be expressed cartographically as "barriers" located within the space in which the distance is being measured. Note that this implies that distance is the result of some sort of movement over that space and through those barriers. Two major types of barriers can be identified as to how they affect this implied movement. Absolute barriers are those which completely restrict movement and therefore imply an infinite distance between the points they separate, unless a path around the barrier is available (Figure 3). A river might be regarded as an absolute barrier to a non-swimmer. To a swimmer or boater, however that same river might be regarded as a relative rather than an absolute barrier. Relative barriers are ones which are passible but only at a cost which may be equated with an increase in physical distance (Figure 4).

Absolute and relative barriers may be expressed as map categories represented by thematic values which indicate their ability to impede movement. For example, movement by automobile is effectively constrained to a network of roads of varying speedlimits. From a less conventional perspective, proximity can be expressed in such terms as accumulated cost of powerline construction from an existing

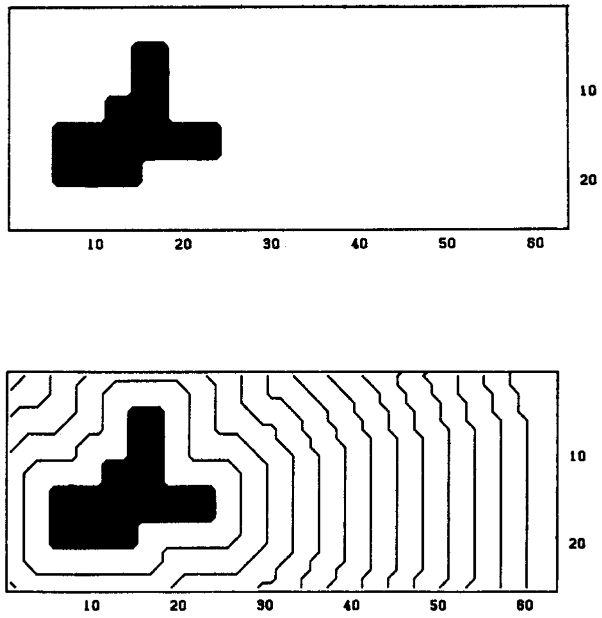


FIGURE 2. The lower map shows the concentric equidistant zones emanating from a feature that was created by assigning to each location within the mapped area a value indicating the shortest straight line distance between that location and the specified feature. The bowl-like 3-dimensional surface that is generated forms a map of proximity to the feature.

powerline trunk to all other locations in a study area. The cost surface developed can be a function of a variety of social and engineering factors, such as visual exposure and adverse topography, expressed as absolute and/or relative barriers.

The ability to move, whether physically or abstractly, may vary as a function of the implied movement. One aspect of movement which may affect the ability of a barrier to restrict that movement is speed. A highway impairment which effectively reduces traffic speeds from 55 to 40 miles per hour, for example, might have little or no effect on traffic already moving at a lower speed. Another possible modifying factor is accumulation. After hiking a certain distance, molehills tend to become mountains. A third attribute of movement which might dynamically alter the effect of a barrier is direction. A topographic incline, for example, will generally impede movement differently according to whether that movement is uphill, downhill or across slope.

A distance-related group of operations involve the delineation of paths, or connectivity, among locations. In the case of simple distance these paths locate the shortest straight line considering only two dimensions (Figure 5a). Another technique traces the steepest downhill path from a location on a three-dimensional surface (Figure 5b and c). The steepest downhill path along a topographic surface will indicate the route of surficial runoff. For a surface represented by a travel time map (or any sort of cumulative cost/distance map as described above), this technique can be used to trace the optimal (i.e. the shortest or quickest) route. If a accumulation cost surface is considered the minimum cost route will be located.

The process of determining viewsheds involves establishing intervisibility among locations. Locations forming the viewshed of an area are connected by straight lines in three dimensional space to the "viewer" location, or set of viewers (Figure 6a). Topographic relief and surface objects form absolute visual barriers that preclude connectivity. If multiple viewers are designated, locations within the viewshed may be assigned a value indicating the number or density of visual connections.

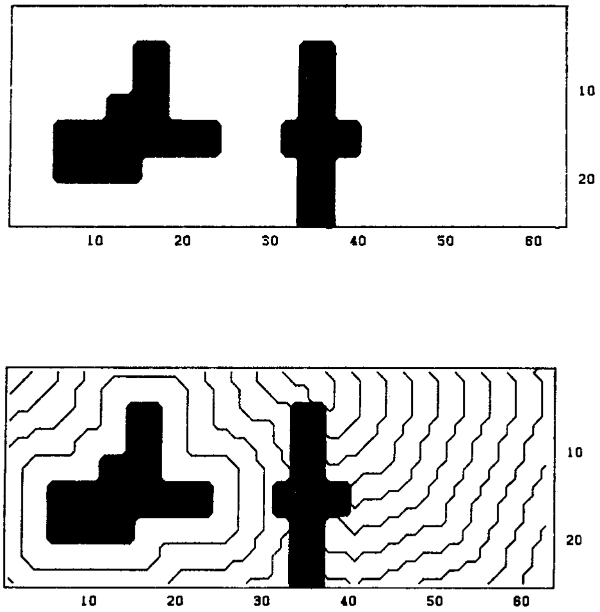


FIGURE 3. The lower map shows the equidistant zones about a feature as modified by an absolute barrier. The distance values assigned to the locations to the left of the barrier are significantly further away than the simple distance zones in Figure 2 because the measurement procedure must circumscribe the barrier.

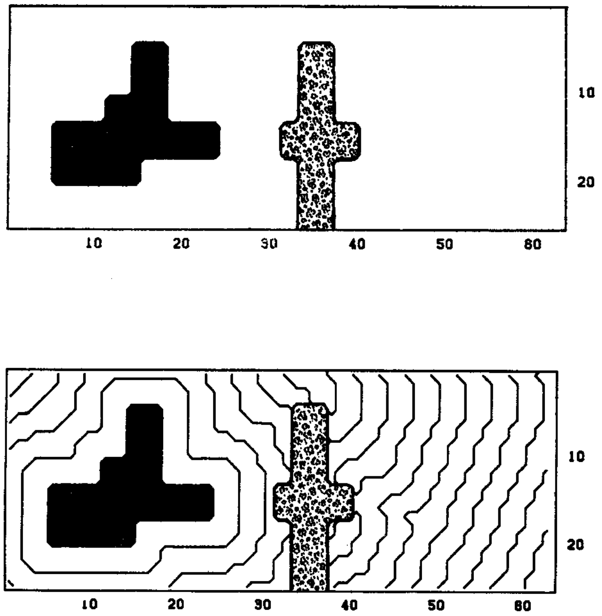


FIGURE 4. The lower map shows the equidistant zones about a feature as modified by a relative barrier. The distance values assigned to the locations to the left of the barrier are different from those in either Figure 2 or Figure 3. The distance zones represent the complex interactions between the possible shortest paths of "through" or "around" the barrier.

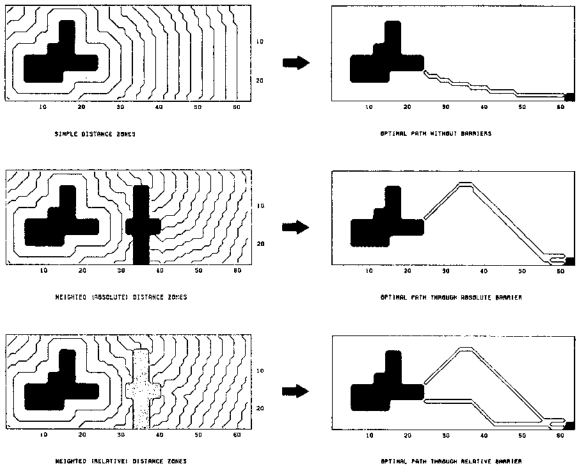


FIGURE 5. Optimal path connectivity locates the shortest route between locations in a study area. The upper pair of maps locates the optimal simple distance route. The middle pair of maps indicates the route if an absolute barrier is imposed. The lower pair of maps locates two equally optimal routes; one around the relative barrier and the other through it.

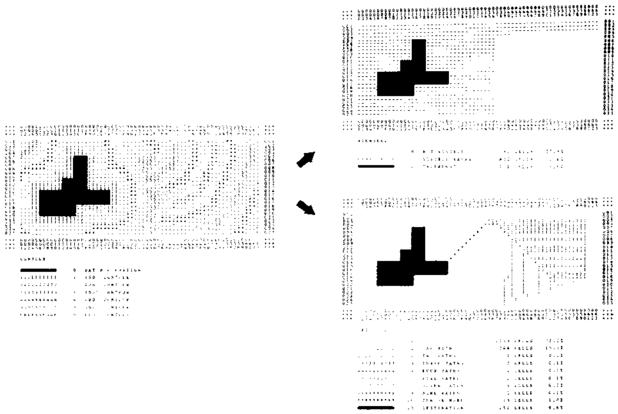


FIGURE 6. Maps of optimal path density characterize the degree of connectivity among locations in a study area. The upper right map identifies the viewshed of the feature indicated in the left map over the surface indicated in the same map. Note the effective visual barrier in the middle portion of the study area. The lower right map indicates optimal path density to the feature from a set of locations to the right of the effective barrier. The paths converge as they are forced to circumscribe the barrier.

In an analogous manner, the density of optimal paths along a three-dimensional surface may be characterized by assigning a value indicating the number of paths traversing each location in a study area (Figure 6b). An extension of this technique weighs each optimal path accroding to its importance. For example, a map could be generated that indicates the volume of timber that would ultimately pass through each location if all harvestable timber were to be accessed by way of the least cost path to a specified collection point.[10] The resulting map of path densities will help to locate major access corridors.

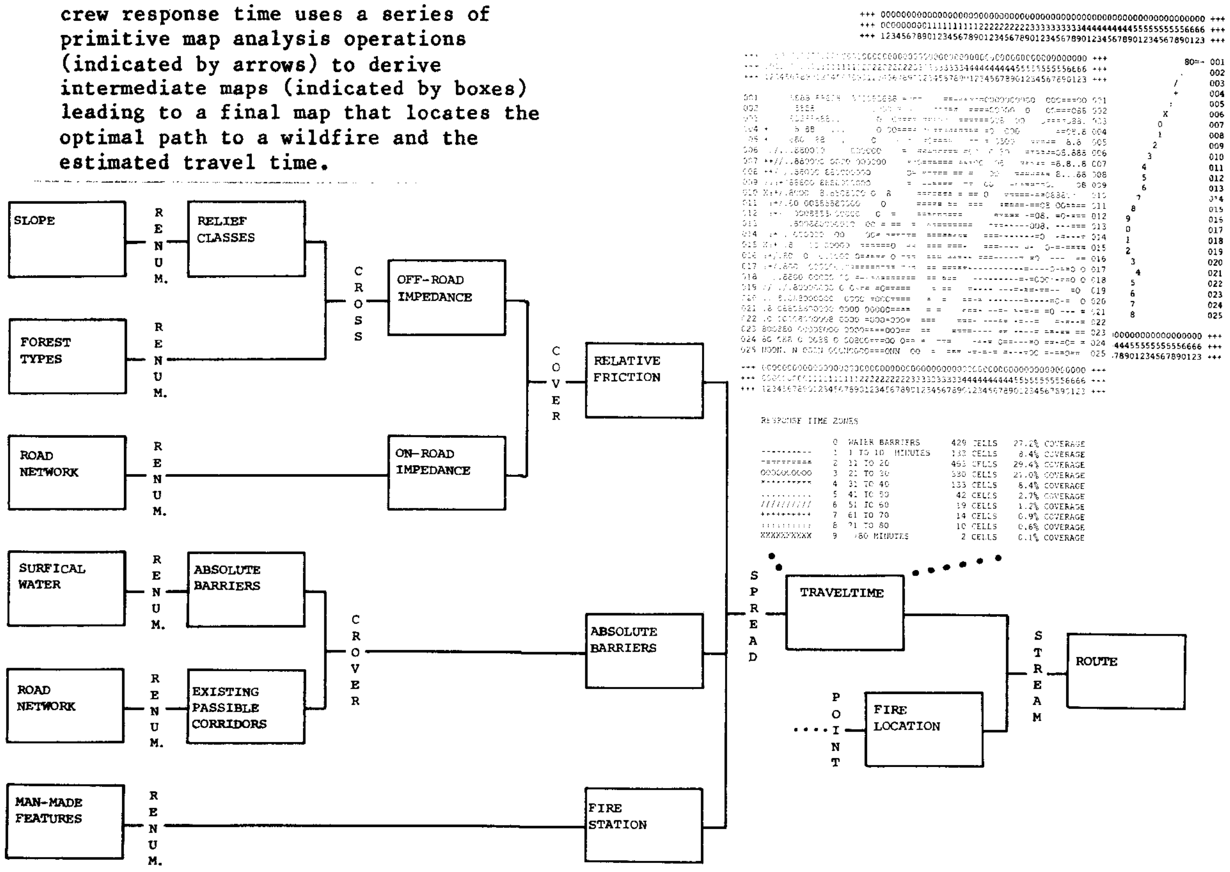
V. CARTOGRAPHIC MODEL

In order to suggest some of the ways in which primitive map processing operations might be combined to perform more complex analyses, an illustrative cartographic model is outlined below and schematically represented in Figure 7. The simplified model estimates fire crew response time and optimal path to a wildfire originating at any location within a study area. The flowchart identifies a logical sequence of Map Analysis Package[8] operations represented as lines which transform maps that are represented as boxes.

Given maps of roads, topography and forest cover the relative ease of on-road and off-road travel is used to generate a map of relative barriers representing time expended in traversing each location. A map of surficial water is used to identify off-road areas that are absolute varriers to vehicular travel. A travel-time map can then be constructed which indicates the minimum time to move from the fire station to any other location. Avoidance of absolute barriers and relative ease of highway travel favor on-road movement whenever possible. The varying degrees of ease of movement through different cover/slope conditions determines off-road travel.

The travel-time map (i.e. accumulation surface) has a bowl-like shape with the fire station as its bottom (i.e. zero time units away). The road network radiates away from this bottom with an incline (i.e increasing time zones) based on the average speed associated with the various road segements. The off-road areas form steep ridges (i.e. rapidly increasing time zones) between road segments. Locations containing absolute barriers become extremely high plateaus (i.e. infinitely far away). This topological shape is unique as it contains no saddle points, or

FIGURE 7. This cartographic model of fire crew response time uses a series of primitive map analysis operations (indicated by arrows) to derive intermediate maps (indicated by boxes) leading to a final map that locates the optimal path to a wildfire and the estimated travel time.



false bottoms. The location of a wildfire with respect to this abstract surface indicates the minimum response time. The steepest downhill path from the fire along the surface locates the optimal path to the fire.

Several natural resource related models have been developed at the Yale School of Forestry and Environmental Studies using this approach as embodied in the Map Analysis Package software. These include:

* spatial characterization of timber supply in terms of transport zones and timbersheds;
* assessing deer habitat quality as a function of weighted proximity to natural and anthtropogenic factors;
* mapping outdoor recreation opportunity as determined by an area's remoteness, size, and physical and social attributes;
* predicting storm runoff from small watersheds by spatially evaluating the standard Soil Conservation Service model;
* assessing the spatial ramifications of the comprehensive plan of a small town considering natural land use, preservation, growth and utility policies; and ,
* characterizing spatial relationships among marine ecosytems factors.

In addressing these divergent applications a common set of fundamental map analysis operations were used. The logical sequencing of these operations on different sets of mapped data form the cartographic models of these different applications.

VI CONCLUSION

The modeling approach described in this paper can be used to extend the utility of maps for a variety of applications. A broad range of fundamental map analysis operations can be identified and grouped according to generalized characteristics. This organization establishes a framework for understanding of the analytic potential of computer-assisted map analysis. By logically ordering these fundamental operations, a cartographic modeling language can be developed which accomodates a variety of applictions in a common, flexible and intuitive manner.

REFERENCES CITED

1. Berry, J.K., "Cartographic Modeling: Procedures for Extending the Utility of Remotely Sensed Data," PECORA VII Symp. on Remote Sensing: An Input to Geographic Information Systems in the 1980's, American Society of Photogrammetry, 105 N. Virginia Avenue, Falls Church, Virginia, pp. 414-424, 1981.

2. Berry, J.K., "A Practical Format for Coupling Techniques of Pattern Recognition and Cartographic Modeling," Proc. of the American Society of Photogrammetry, 46th Annual Meeting, pp. 131-142, 1980.

3. Berry, J.K, "Cartographic Modeling in the Classroom: A Fundamental Approach to Computer-Assisted Map Analysis," Proc. of the American Congress on Survey and Mapping, 40th Annual Meeting, pp. 227-241, 1980.

4. Calkins, H. and D.F. Marble, Computer Software for Spatial Data Handling, volumes 1-3, International Geographical Union, Commission on Geographic Data Sensing and Processing, Ottawa, Ontario, 1980.

5. Chrisman, N. and T. Peucker, "Cartographic Data Strutures," The American Cartographer, Volume 2, Number 1, pp. 56-69, 1975.

6. McHarg, I.L. Design with Nature, Doubleday/Natural History Press, Doubleday & Company, New York, 1969.

7. Tomlin, C.D. and J.K. Berry, "A Mathematical Structure for Cartographic Modeling in Environmental Analysis," Proc. of the American Congress on Surveying and Mapping," 39th Annual Meeting, pp. 269-283, 1979. 8. Tomlin, C.D., "Digital Cartographic Modeling Techniques in Environmental Planning," doctoral disseatation, School of Forestry and Environmental Studies, Yale University, New Haven, Connecticut, in preparation.

9. Tomlin, C.D. and S.M. Tomlin, "An Overlay Mapping Language Symposium on Regional Landscape Planning, American Society of Landscape Architects, Washington, D.C., 1981.

10. Tomlin, S.M., "Timber Harvest Scheduling and Spatial Allocation," Master's thesis, Graduate School of Arts and Sciences, Harvard University, Cambridge, Massachusettes, 1980.