Reprinted from

Ninth International Symposium

Machine Processing of

Remotely Sensed Data

with special emphasis on

Natural Resources Evaluation

June 21-23, 1983

Proceedings

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

COMPUTER-ASSISTED MAP ANALYSIS: ACADEMIC STRUCTURE AND EXPERIENCE

J.K. BERRY

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

ABSTRACT

In response to the increasing interest in the use of computer-oriented geographic information systems, academic programs in many fields have begun to recognize the need for more effective ways of presenting spatial data analysis techniques to students with a wide diversity of backgrounds. The most commonly used approach is the case study coupled with a limited exposure to computer mapping software. This paper presents an alternative approach which relies on the development of a fundamental analytical theory coupled with extensive exercises which demonstrate the practical application of that theory. This experience is further enhanced by opportunities for students to design and implement their own cartographic models. The approach used in conveying concepts of spatial data analysis develops a general framework of primitive map analysis operations that can applied to a wide variety of applications. This deductive approach appears to be effective in providing students with a basic knowledge of map analysis techniques and enables them to extend this knowledge to potenital applicaltions of their own design. This approach has been used for six terms in a graduate-level course at Yale University. The course has been implemented at several other Universities and serves as the basis of an intensive two-day workshop for professionals. Acquisition of software, course materials, and lecture notes can be arranged.

I. INTRODUCTION

Computer-assisted geographic information systems show promise for providing capabilities clearly needed for effective land planning. However, the need for such capabilities has frequently cast the developing science into operational contexts without adequate user education. The result has been a technology that often intimidates potential users and discourages what otherwise would be valuable input in the identification, clarification, and development of new applications. A major factor that contributes to this situation is the absence of a comprehensive framework that organizes the wide range of analytic techniques. Until this academic framework is established, geographic information systems will continue to be thought of as special-purpose "black boxes."

One frequently used academic approach is the case study. In this context, students are exposed to a variety of applications that demonstrate an array of analytical capabilities. Learning is by induction. When properly organized and presented, this form of instruction can be effective. Critical to this approach, however, is the selection of cases that match student interest and experience. If a mismatch occurs, confusion about the specifics of the example can cloud the fundamental principles being demonstrated. Furthermore, the selection of cases must balance the inherent complexity of interesting applications with the simplicity necessary for clearly portraying fundamental theory.

Another approach is one which relies on establishing a fundamental theory of spatial information analysis. This approach is analogous to the traditional presentation of basic mathematics in which general principles and operations are introduced and then demonstrated through practical examples. The deductive nature of this approach is particularly appropriate for audiences with diverse backgrounds and interests.

However, the development of this approach to spatial information analysis has two drawbacks. First, this developing science is continually redefining existing techniques and rapidly generating new ones. The dynamic environment makes identification of an enduring general framework difficult. Second, in order for this approach to be effective, supporting examples and exercises aimed at conversion of general theory into practical experience are required. In the case of basic mathematics exercises are easily defined and executed. However, the hardware and software resources required for "hands-on" instruction of spatial information analysis are much more sophisticated and demanding.

This paper describes the development of an academic structure and supporting materials for instruction in the fundamental principles of computer-assisted map analysis. A set of primitive operations is organized into a typology that encompasses a broad range of higher level analytical techniques. By organizing these primitive operations, the basis for a generalized cartographic modeling approach is identified that accomodates a variety of analytic procedures in a common, flexible and intuitive manner. None of the procedures are application-specific and the use of each is limited only by the general thematic and spatial nature of the data to which it is applied.

This mathematical-like structure serves as the focus of a curriculum for both a formal graduate-level course and a professional oriented workshop. Emphasis in both of these courses is given to extensive use of exercises as a means of demonstrating fundamental concepts of computer-assisted map analysis. The software, course materials and lecture notes used in these courses can be acquired from the author. The software system is written in FORTRAN IV and has been implmented in most major computer environments. A microcomputer based system is under development.

II. FUNDAMENTAL THEORY

The establishment of a fundamental theory and "hands-on" instruction requires a common data structure and a felxible processing structure. 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, or "overlays." An overlay is simply a special form of a geographic map. On an overlay, each cartographic location is characterized explicitly in terms of one and only one thematic attribute. This is in contrast to the conventional format of a USGS topographic map on which each point is explicitly characterized in terms of its elevation, its water features, its vegetation, its political designations and so on.

If primitive operations are to be flexibly combined, a processing structure must be used that accepts input and generates output in the same format. Using an "overlay" data structure 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 the new map for subsequent processing.

The cyclical nature of this processing structure 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 specified application. As with traditional algebra, fundamental techniques involving several primitive operations can be identified (e.g. generation of a "travel-time" map) that are applicable to numerous situations. The use of primitive analytical operations in a generalized modeling context accomodates 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.

Within the data and processing structures noted 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 a user's point of view, four major classes of fundamental map analysis operations may be identified. These involve:

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

These major groupings can be further classified as to whether processing is done on a point-by-point, region-by-region, or neighborhood-by-neighborhood basis.

The first group of cartographic modeling operations is the simplest and, in many ways, the most fundamental. Each of the operations involves the creation of a new map by reassigning thematic values to the categories of an existing overlay. 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 overlay and results in no new boundary delineations. Such operations can be thought of as the purposeful "recoloring" of maps.

Operations for overlaying maps begin to relate to the spatial, as well as to the thematic nature of cartographic information. The general class of overlay operations can be characterized as "light-table gymastics." Included in this class of operations are those which involve the creation of a new map such that the value assigned to each location is a function of the independent values associated with that location on two or more existing overlays. In simple location-specific overlaying, the value assigned is a function of the point-by-point aligned coincidence of the existing overlays. In category-wide compositing, values are assigned to entire thematic regions as a function of the values associated with the regions contained on the existing overlays. Whereas the first overlaying approach conceptually involves vertical spearing of a set of overlays, the latter approach uses one overlay to identify boundaries from which information is extracted in a horizontal summary fashion from another overlay. A third overlay approach treats each overlay as a variable, each location as a case, and each value as an observation in evaluating, on a point-by-point basis, 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 locations on a map. The simplest of the distance related operations involves the creation of a map in which the value assigned to each location indicates the shortest distance "as the crow flies" between that location and a specified target area. The result is an overlay of concentric, equidistant zones around the target(s). A target area need not be constrained to a single location and can be comprised of a set of dispersed points, lines or areal features.

If movement is implied in the measurement of distance the shortest route between two points may not always be a straight line. And even if it is straight, the Euclidean length may be defined in terms of factors such as travel-time, cost, or energy which, unlike miles, may be consumed at rates that vary over space and time. These modifying effects may be expressed cartographically as absolute and relative "barriers" located within the space over which distance is being measured. The resulting overlay identifies an effective proximity surface that characterizes movement from a target area over that space and through those barriers. This general class of operations can be thought of as "rubber rulers" for measurement of effective proximity.

Another distance-related set of operations determines the connectivity among specified locations. One such operation traces the steepest downhill path from a point on a three-dimensional surface. For a topographic surface, the path would indicate surficial water flow. For a surface represented by a travel-time map, this can be used to trace the minimum-time (i.e. quickest) path. Another operation determines connectivity measured only for straight rays emanating from a target area over a three-dimensional surface to identify visual exposure.

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). This general class of operations are characterized by "roving windows." The summary of information within the neighboring locations can be based on the configuration of the surface (e.g. slope and aspect) or the statistical summary of thematic values (e.g. average value).

Additional operations necessary for a complete cartographic modeling system include those associated with data encoding, display and program control. A more detailed discussion of the general modeling approach and analytical techniques used in this course is presented in other papers.[3,4,6]

III. CURRICULUM

The organization of fundamental data processing operations described above is one which is generally familar to students and one which provides for a generalized approach to instruction in spatial information analysis. The curriculum of the graduate-level course at Yale is outlined in Table 1. Note that the first half of the semester is devoted to the introduction of fundamental map processing operations. This normally involves three hours of lecture and one hour of laboratory exercises each week. Each lecture session deals with the general theory and application of a group of related functions. Coordinated laboratory sessions include discussion of forthcoming exercises associated with that group of functions and student presentations of solutions to the previous week's assignment. These exercises are designed to give students an opportunity to appply fundamental theory presented in lecture. Each of eight laboratory sessions includes approximately ten individual tasks. These are tied together as much as possible through a common, hypothetical situation in order to provide continuity and to avoid what might otherwise become a purely mechanical procedure. They are completed outside of class and normally require an average of from four to eight hours each week. Students are encouraged to work in groups as dialogue is likely to be mutually constructive. Friendly competition between working groups is also encouraged. The reading material for this phase of instruction consists primarily of specially prepared course materials.

Beginning in the ninth week, after a midterm examination students are exposed to some of the broader complexities of cartographic modeling. Here they are given several opportunities to design cartographic models in an application context and to apply their knowledge of the fundamental processing operations in a comprehensive manner. Students are then introduced to techniques of data acquisition, encoding, and display. At this point, they complete exercises calling for manual grid encoding, electronic cursor digitizing, and pen plotter display.

The remaining lectures are devoted to discussion of contemporary processing systems and applications. As much as possible, these topics are presented within the general framework developed during the first half of the course. Each student prepares short reports on an application and a software system. Reference

TABLE 1. SYLLABUS FOR INTRODUCTORY COURSE IN GEOGRAPHIC INFORMATION SYSTEMS

WEEK 1	INTRODUCTION - course outline; overview of G.I.S. and applications; fundamental approach in cartographic modeling.
WEEK 2	DATA AND PROCESSING STRUCTURES - data organization; polygon vs. grid structures; processing structures; MAP operations INFORM, LIST, DISPLAY, and STOP. Exercise: Introduction to the MAP system.
WEEK 3	RECLASSIFICATION FUNCTIONS- thematic and locational attributes of cartographic information; qualitative vs. quantitative data; reclassifying mapped data as a function of thematic value, location, size or shape; MAP program control operations INFORM, DESCRIBE, PROTECT, EXPOSE, SCALE, LABEL, RENAME, ZAP, IDENTIFY, NOTE, and PRINT; reclassifying operations RENUMBER, SLICE, SIZE, EULER, CLUMP, and SURVEY. Exercise: Reclassifying Map Categories.
WEEK 4	OVERLAY FUNCTIONS - overlaying maps to generate arithmetic, logical, or statistical combinations of mapped data; MAP operations ADD, SUBTRACT, MULTIPLY, DIVIDE, EXPONENTIATE, MAXIMIZE, MINIMIZE, AVERAGE, COVER, CROSS, and SCORE. Exercise: Overlaying Maps.
WEEK 5	DISTANCE FUNCTIONS (Proximity) - cartographic representation of movement; measuring true vs. weighted distance; generating travel-time and travel-cost maps; delineating watersheds, MAP operations SPREAD and SPAN. Exercise: Measuring Euclidean Distance.
WEEK 6	DISTANCE FUNCTIONS (Connectivity) - generating viewsheds and shortest paths; operations RADIATE, and STREAM. Exercise: Measuring non-Euclidean Distance.
WEEK 7	NEIGHBORHOOD FUNCTIONS (Surface Configuration) - characterizing points on a three-dimensional surface; topographic slope, aspect, and relief; MAP operations DIFFERENTIATE, ORIENT, and PROFILE. Exercise: Characterizing Points on a 3-D Surface.
WEEK 8	NEIGHBORHOOD FUNCTIONS (Thematic Values) - narrowness; "roving window" statistics: interpolation; contour smoothing; azimuth and distance weighting; map scale conversion; MAP operation SCAN. Exercise: Characterizing Cartographic Neighborhoods.
WEEK 9	MODELING TECHNIQUES - flowcharting and logical structuring of models using primitive operations; error handling; interactive vs. batch processing; extending MAP capabilities; MAP operations FLOWCHART, ECHO, QUIET, REMEMBER, OVERLOOK, REMIT, READ, WRITE, and INSERT. Exercise: Program Control Operations and Model Implementation. MIDETERM EXAMINATION.
WEEK 10	MODELING TECHNIQUES - continued. Exercise: Model Modifications.
WEEK 11	DATA ENCODING AND DISPLAY - encoding strategies; digitizing methods; cartographic projections and geometric registration; pageprint, raster, 2-D/3-D pen plots; computer graphics display devices; MAP operations MAP, POINT, GRID, STRIP, TRACE and RESPACE. Exercises: Data Encoding.
WEEK 12	SOFTWARE AND CURRENT PRACTICE - digital representation of information; alternative cartographic data structures, processing structures, processing structures; and command language; computational efficiency; implentation and operating costs; and available systems. TEAM PROJECT PREPARATION - problem definition; rough flowchart of proposed model. INDIVIDUAL REPORTS on an application and a software system.
WEEK 13	TEAM PROJECT PREPARATION - Model refinement and implementation
WEEK 14	TEAM PROJECT PRESENTATION - 30 minute oral presentation. FINAL EXAMINATION

material for this phase of instruction consists of selected theoretical papers, project reports, software manuals, and hardware descriptions.

The final and major component of the course involves team projects. These are intended to give students broader experience in the design, implementation, and presentation of actual cartographic analyses. Some examples of team projects include deer habitat modeling, timberlands accessibility, marine ecosystem modeling, land use planning and landfill siting. A more detailed discussion of the curriculum is presented in other papers.[1,2]

IV RESOURCES

Software support for this course was provided through use of the Map Analysis Package (MAP)[3] This package is a set of programs developed at Yale that provide for the encoding, storage, analysis, and display of cartographic information. Its development represents an attempt to accomodate many of the academic objectives outlined above. Thus, use of the package is similar in many ways to the use of traditional techniques that involve conventional geographic maps. The data processing capabilities of MAP are organized as a series of primitive operations that may be combined to perform a variety of complex map analyses. These operations are specified intuitively through a user oriented command language of English-like phrases that does not require formal knowledge of computer programming. The MAP system is written in FORTRAN IV and may be implemented for both interactive and batch processing modes. At present, the package employs a grid-cell data structure but has line segment oriented input and output capabilities.

Each of the more than 60 independent MAP operations is associated with one of five major categories (storage, analysis, display, program control, and user-defined) according to its function and its relation to the flow of information between programs, data files, and input or output media. The analytic operations are subgrouped into the previously described categories of reclassification, overlay, distance and neighborhoods.

The data base used for most of the exercises includes small grid maps of 1575 cells each, dimensioned at 25 rows by 63 columns. This configuration has proved to be extensive enough to provide adequate demonstration of processing operations while being small enough for efficient computation and display. The data base includes maps at a cell size of 250 meters square covering a small portion the tri-state area of Alabama, Tennessee, and Georgia. The base maps include land use, housing, forest cover, elevation, roads, power lines and water bodies. The team projects utilize one of three other more extensive and detailed data bases of local areas.

Hardware in support of the course inlcudes remote access to an IBM 4341 computer via normal student terminals. Hardcopy of interactive terminal sessions are directed to a high-speed line printer. Special equipment includes a Talos large bed digitizer and a Zeta pen plotter. Computer use costs are estimated to be about $175 per student.

V. EXPERIENCE

The course has been presented each spring term for the past six years. The largest course enrollment was thirty four with an average class size of approximately twelve. The periods of large enrollments resulted in the presentation of a short course (16 hours of instruction) designed for individuals who wish to develop a familarity with the field, but not an operational expertise. This condensed offering has allowed the formal course to address a smaller and more focused group of students.

Students in both courses noted that the command language and control structure of MAP were appropriate. The free-format, English-like commands proved to be particularly user friendly. Since there are no apparent differences to the user between interactive and batch processing with MAP, the transition is made easily from interactive-based exercises to batch-oriented projects.

The projects represent the culmulation of efforts for the term. Even though these reports consumed a large portion of each student's limited time, enthusiasm has continually remained high. Several student evaluations noted that the projects were particularly valuable because they provided the opportunity for experiencing the complexities of practical application. Many of the students completing the course have pursued major independent projects the following year.

VI. CONCLUDING REMARKS

This course has attempted to develop an academic framework for a deductive modeling approach to the conceptualization of geographic information analysis. This structure is useful because it allows students to develop a fundamental knowledge of techniques of spatial analysis. This general framework also enables them to extend this knowledge to applications of their own design.

Essential to this approach is the ability to provide actual experience in processing of spatial data. This opportunity gives prospective users a realistic view of the analyst's tasks. It also reinforces in a practical context, the range of fundamental assumptions that must be made in any study. The MAP system proved to be ideally suited for this practical experience.

VII. REFERENCES

1. Berry, J.K. and C.D. Tomlin. 1979. "An Academic Approach to Cartographic Modeling in Natural Resource Management", Harvard Library of Computer Graphics, Volume 11:13-19, Harvard University, Cambridge, Massachusetts.

2. Berry, J.K. and C.D. Tomlin. 1980. "Cartographic Modeling in the Classroom: A Fundamental Approach to Computer-Assisted Map Analysis," 40th Sym. Proc., Amer. Congress on Surveying and Mapping, pgs. 227-241.

3. Berry, J.K. and C.D. Tomlin. 1982a. "Cartographic Modeling: Computer-Assisted Analysis of Effective Distance," 8th Sym. on Machine Processing of Remotely Sensed Data, Purdue University, West Lafayette, Indiana. pgs. 503-510.

4. Berry, J.K. and C.D. Tomlin. 1982b. "Cartographic Modeling Computer-Assisted Analysis of Spatially Defined Neighborhoods," Conf. on Energy Resource Management, Amer. Planning Assoc., Vienna, Virginia.

5. Tomlin, C.D. 1980. The Map Analysis Package (Draft); Yale School of Forestry and Environmental Studies, New Haven, Connecticut.

6. Tomlin, C.D., J.K. Berry, and S.M. Tomlin, 1981. "Fundamental Overlay Mapping Techniques," Wkshp. on In-Place Resource Inventories, Society of Amer. Foresters, Bethesda, Maryland. pgs. 470-481.