ORNL/CSD/TM-226

MARTIN MARIETTA

Review of Geographic Processing Techniques Applicable to Regional Analysis

R. C. Durfee

OPERATED BY
MARTIN MARIETTA ENERGY SYSTEMS, INC.
FOR THE UNITED STATES
DEPARTMENT OF ENERGY

Printed in the United States of America. Available from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road, Springfield, Virginia 22161
NTIS price codes—Printed Copy: A07 Microfiche A01

This report was prepared as an account of work sponsored by an agency of the United States Government Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof

ORNL/CSD/TM-226

REVIEW OF GEOGRAPHIC PROCESSING TECHNIQUES APPLICABLE TO REGIONAL ANALYSIS

R. C. Durfee

Computing and Telecommunications Division
at Oak Ridge National Laboratory
P. O. Box X
Oak Ridge, Tennessee 37831

Major Contributors

E. A. Bright	F. E. Latham
P. R. Coleman	S. M. Margle
J. E. Dobson	G. J. Morris
R. G. Edwards+	L. E. Till
J. F. Hines*	E. P. Tinnel
P. E. Johnson	D. L. Wilson
	B. C. Zygmunt

+ Energy Division, Oak Ridge National Laboratory

* Knox County GIS, Knoxville, Tennessee

February 1988

Prepared for and partially funded by the
Environmental Protection Agency

NOTICE This document contains information of a preliminary nature. It is subject to revision or correction and therefore does not represent a final report.

MARTIN MARIETTA ENERGY SYSTEMS, INC.
operating the
Oak Ridge National Laboratory   Oak Ridge Y-12 Plant
Oak Ridge Gaseous Diffusion Plant   Paducah Gaseous Diffusion Plant
for the
U.S. DEPARTMENT OF ENERGY
under contract DE-AC05-84OR21400

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES (cont'd)

LIST OF FIGURES (cont'd)

LIST OF FIGURES (cont'd)

LIST OF DIAGRAMS

ACKNOWLEDGMENTS

This paper presents concepts, ideas, and work that have been carried out by a number of individuals at the Oak Ridge National Laboratory (ORNL), including computer scientists in the Geographic Data Systems Section, Computing and Telecommunications Division, who have worked with a variety of sponsors from other divisions and agencies. Two of the key people deserving significant recognition and appreciation are Phillip R. Coleman and Robert G. Edwards, whose research and development represented state-of-the-art work in geographic algorithms, data structures, and processing techniques since the mid 1970s. Steve M. Margle has been a key individual responsible for excellent hardware-software development and applications of the ORIDS Geographics Laboratory (Oak Ridge Interactive Digitization and Display Systems), along with Ed P. Tinnel's work with digitization systems, graphics display systems, and image processing techniques. Since the early 1970s Don L. Wilson has directed and carried out a variety of regional analysis projects whose graphic results are presented here. Paul E. Johnson has been responsible for much of the transportation analyses and mapping. Fred E. Latham has carried out many different geographic applications involving national and regional data bases. Garvin J. Morris has provided excellent support in a variety of areas including color output systems, computer digitization, data analysis, and quality assurance associated with the ORIDS graphics systems. Recognition is also due Jerome E. Dobson who has participated in and supported a variety of automated geographic applications including water resource studies and emergency management. Two key members have been Jim F. Hines and Ed A. Bright who have developed many of our geographic data bases and performed the spatial analyses and computer mapping for a wide range of applications. Lynn E. Till has been another key participant in the development of geographic workstation capabilities using high-resolution microcomputer systems. Other participants of the Geographic Data Systems Section whose geographic and systems work have been instrumental include John L. Smyre, I. Glen Harrison, Chen Liu, Greg S. Love, Beverly C. Zygmunt,Don W. Rosowitz, and Brenda H. Holt.

There have been a large number of sponsors and co-participants who have funded, directed, and participated in this work, including several ORNL Divisions, such as the Energy Division which has been involved since the beginning of this work. Robert B. Honea has been a key contributor and participant in these activities since the early 1970s when NSF sponsored a Regional Environmental Systems Analysis project (RESA) at the Laboratory. Appreciation is expressed to Richard J. Olson, David S. Shriner and others of Environmental Sciences Division who provided part of the funds for early drafts of this report as well as continuing support in geographic applications. Dave S. Joy and Lawrence B. Shappert from the Chemical Technology Division have directed and supported much of the transportation work done by Paul E. Johnson and the model development by Bruce E. Peterson from the Energy

Division. Support and encouragement has been provided on a wide range of Air Force applications by Robert B. Craig of the Hazardous Waste Remedial Actions Program. Richard M. Rush has coordinated several large projects involving the Nuclear Regulatory Commission and the Department of Interior. Cooperative work has also been carried out with many other groups including Tom W. Oakes and Karen L. Daniels in the Environmental and Occupational Safety Division, Don E. Olins and Ada Olins in the Biology Division, Ed L. Hillsman, Robert B. Braid and Robert D. Roop of the Energy Division, Robert M. Cushman with the Carbon Dioxide Information Center, and Robert M. Holmes and Jan B. Berry in Engineering and Operations Divisions. Appreciation is also due William Fulkerson, Director of the Energy Division, who has been a long-time supporter of these activities and systems. Thanks are expressed to all of these people and many others whose cooperative efforts over the last 15 years have resulted in the techniques discussed herein.

ABSTRACT

Since the early 1970s regional environmental studies have been carried out at the Oak Ridge National Laboratory using computer-assisted techniques. This paper presents an overview of some of these past experiences and the capabilities developed at the Laboratory for processing, analyzing, and displaying geographic data. A variety of technologies have resulted such as computer cartography, image processing, spatial modeling, computer graphics, data base management, and geographic information systems. These tools have been used in a wide range of spatial applications involving facility siting, transportation routing, coal resource analysis, environmental impacts, terrain modeling, inventory development, demographic studies, water resource analyses, etc. The report discusses a number of topics dealing with geographic data bases and structures, software and processing techniques, hardware systems, models and analysis tools, data acquisition techniques, and graphical display methods. Numerous results from many different applications are shown to aid the reader interested in using geographic information systems for environmental analyses.

1. INTRODUCTION

Since the early 1970s regional environmental studies have been carried out at the Oak Ridge National Laboratory using computer-assisted techniques. This paper presents an overview of some of these past experiences and capabilities developed at the Laboratory for processing geographic data. The techniques and methods described are the result of work done over the past 15 years by the Geographic Data Systems Section in cooperation with a number of agencies and key investigators from other groups, including the Energy Division. The intent in this document is not to provide an all-encompassing description of geographic information systems (GIS), but rather to introduce some concepts associated with acquiring, analyzing and displaying geographic data. A variety of examples are included to help describe these concepts, some of them utilizing various computer systems in the ORIDS Geographics Laboratory (see Diagram 1 below).

For the casual observer, maps represent a very efficient mechanism to present large amounts of information. Various spatial relationships are readily apparent in graphic form, but in order to analyze and model spatial processes quantitatively, geographic information must be transformed into digital data whose structure and resolution are sufficient for computer analysis. Geographic information is much more complex than other types of tabular data, containing both spatial (x-y-z) and temporal characteristics, as well as the thematic measurements made over time and space. Advances in the last few years in geographic information technologies have made it possible to perform complex analyses at high resolution for large geographic areas. For successful application of geographic information systems to regional environmental problems, investigators must consider and integrate five key ingredients. These include (1) the data bases and data resources, (2) the graphic and computer hardware systems, (3) the software algorithms and packages, (4) the models and analyses that need to be carried out, and (5) the resource staff and personnel who understand the real-world problems and how to use the tools at hand. This document discusses only portions of the second, third, and fourth components dealing primarily with software processing techniques and analysis methods. The primary elements discussed within these areas will include data structures and coordinate systems to represent the geographic nature of the data, various digitizing techniques for collecting the data, transformations and analyses to manipulate the data, and several techniques for graphically displaying the data in map form. Questions dealing with data access techniques, storage and retrieval problems, data management systems, source documents (e.g., maps, aerial photography, etc.), and specialized hardware equipment are mentioned but not discussed in detail in this document.

A variety of geographic data resources and analysis tools are available at ORNL for use in spatial applications such as acid rain studies, land-use modeling, terrain analysis, facility siting, transportation routing, water resource studies, remote sensing, environmental impacts, coal resource analyses, etc. Examples include river and stream locations, water quality data, political boundaries, population distributions, LANDSAT imagery, industrial and utility plant locations, highway and rail networks, airports, air monitoring data, ORNL—DWG 85-7742

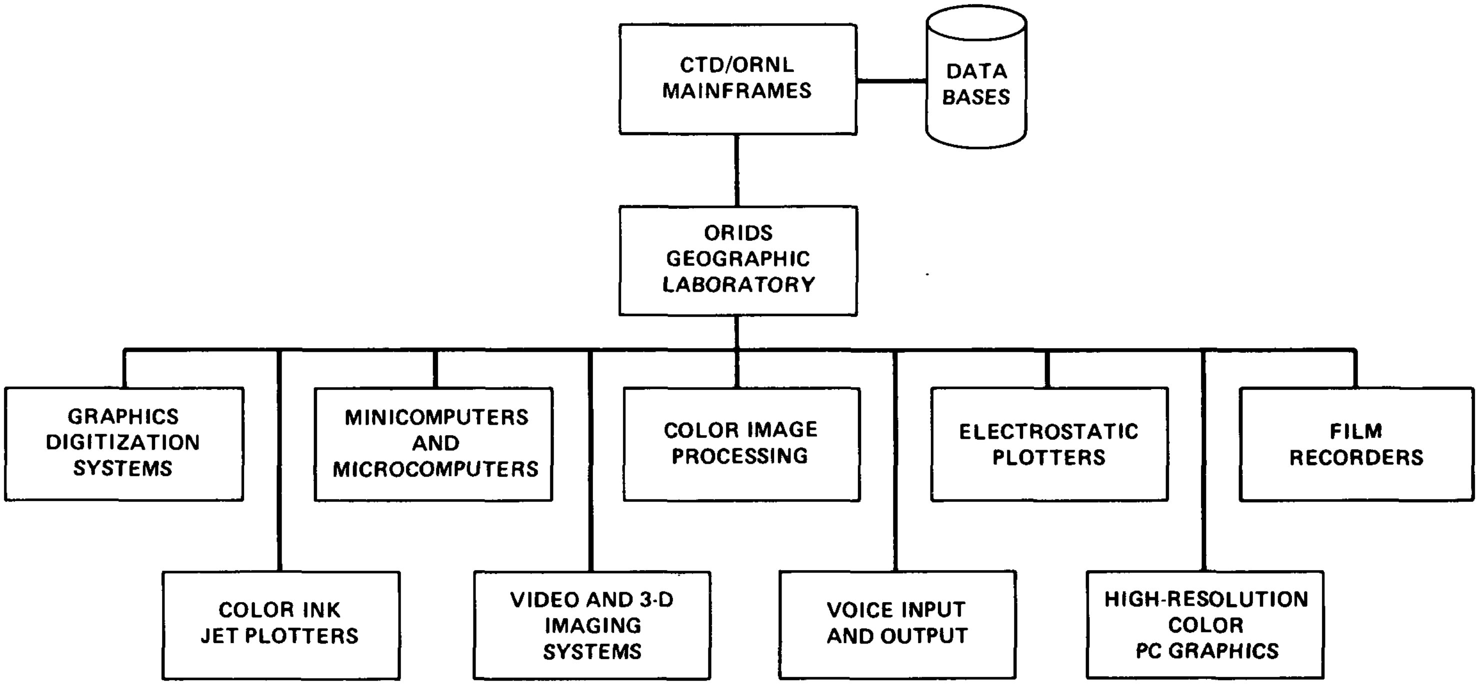


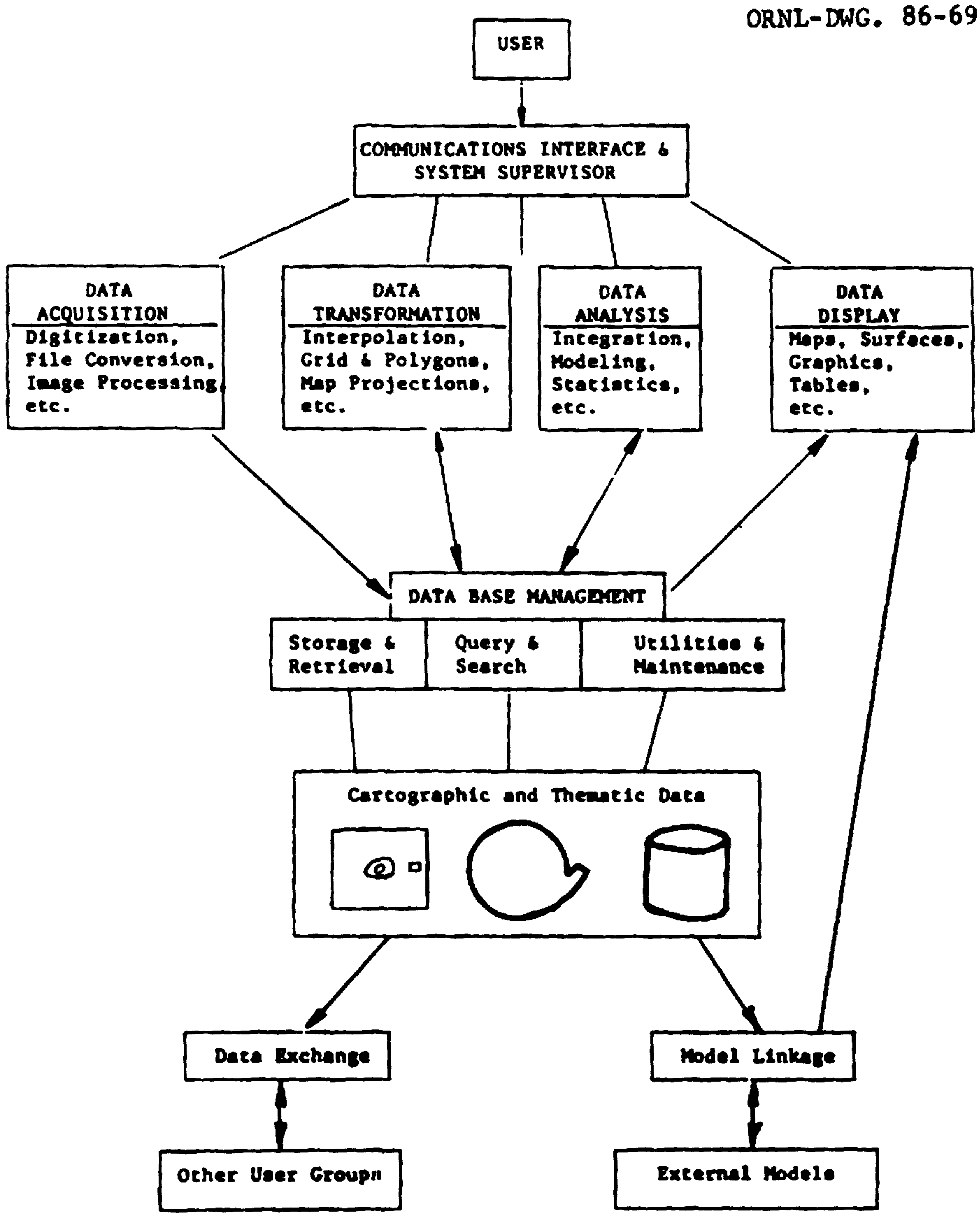
Diagram 1. Computer Graphics Systems Comprising the ORIDS Geographics Laboratory (Oak Ridge Interactive Digitizing and Display Systems).

water flow data, lake and reservoir locations, topography data, watershed basins, wind rose and meteorological data, etc. There is a wide range in resolution and scale associated with these different data bases. For example, the topographic and LANDSAT multi-spectral scanner (MSS) imagery are at a resolution of roughly one acre, whereas major watershed basins may correspond to areas that are several counties in size. The types of questions being asked, the degree of detail needed, and the methodology employed in the analysis will determine the usefulness of these different data bases. For some problems there may be other sources of geographic information which need to be digitized into machine-readable form for integration and analysis with existing data files.

In order to represent, analyze, and display the spatial relationships among these massive amounts of information efficiently, special systems and techniques are required. The geographic software comprising these systems is extremely complex due to the flexibility necessary to handle many data types, and the complexity of spatial algorithms required for efficient processing of large amounts of data. A comprehensive system can be represented functionally as a series of interlinked components (see Diagram 2 below). They include a user-friendly interface for communications between the human operator and the computer, a system supervisor for managing the programs and modules, an extensive suite of geographic functions, and a spatial data base management system. The software modules must be integrated so that data can flow easily and the integrity of the information can be preserved and presented properly to the user. The software must perform many types of conversions and transformations to assure compatibility between data structures (e.g. vectors and rasters), projections and coordinate systems, varying map scales, etc. Appendix A gives examples of typical software functions that might be included in a GIS.

Experience in solving real-world problems has demonstrated that, in many cases, the majority of the effort is spent in preparing, acquiring, and validating the geographic data bases, even before the analyses are done. An understanding of GIS technology and the availability of flexible and user-friendly software systems can ease the burden of data base creation and spatial analysis. The remaining sections discuss the concepts and methods used to manipulate geographic entities in the computer. The discussion is presented as an overview for the non-computer specialist, covering a broad range of areas and examples. The intent is to aid the regional planner who is beginning to use automated systems for geographic applications. The reader who is interested primarily in spatial analysis techniques and application models may wish to skip to Section 4.

ORNL-DWG. 86-6926



FUNCTIONAL CHARACTERISTICS OF A GEOGRAPHIC INFORMATION AND ANALYSIS SYSTEM

Diagram 2. Functional Characteristics of a Geographic Information and Analysis System.

2. DATA STRUCTURES AND COORDINATE SYSTEMS

2.1 MAP PROJECTIONS AND COORDINATE SYSTEMS

One of the first steps in planning techniques for performing regional analyses is to determine how geographic features from maps of the earth's surface are going to be represented in the computer. There are a variety of geographic data structures such as polygons, grids, networks, and point data; but the basic spatial unit for representing the location of features is an x,y coordinate location. The user must recognize that coordinate systems used for measuring x and y are based on either latitude-longitude (using an ellipsoidal model of the earth's surface) or on certain map projections whereby the three-dimensional features are transformed from the earth's surface to a two-dimensional map as shown in Fig. 1. There are a number of different map projections[1] commonly used, such as Albers' Equal Area, Lambert's Conformal, Transverse Mercator, Polyconic, State Plane, etc. Each of these map projections has its own coordinate system normally measured in meters or feet from a defined origin. It is important in capturing geographic data that either latitude-longitude or one of the standard coordinate systems be used so that other data bases can be combined with them for later analysis and display.

In the early days of geographic systems, some users captured data by placing an arbitrary grid over a base map and recording the contents of each cell without regard to any coordinate system associated with the grid. These data were then very limited in use since they could not be combined with other data files later on, nor could they be mapped at different map projections other than the original map from which they were digitized. The current philosophy of our group is to input and output data using whatever map projection is required, but convert to latitude-longitude for internal storage. Appropriate transformations are used when digitizing maps for input or when plotting output maps to meet the user's requirements. An example is shown in Fig. 2. If the internal storage is based upon a map coordinate system there are usually problems if large regions are involved. For example, in using the Universal Transverse Mercator (UTM) coordinate system there are problems at zone boundaries because the grid coordinates from one zone do not mesh with grid coordinates from an adjacent zone, thus creating a discontinuity and triangular cells at the border (see Fig. 3).

2.2 POINT STRUCTURES

Many geographic features correspond to point locations at which thematic measurements are made. Examples include water quality gauging stations, oil and gas wells, air monitoring stations (Fig. 4), factory locations, etc. Data associated with locating and identifying a geographic feature and defining its spatial relationships with its neighbors is referred to as "cartographic data". This would correspond to the coordinates of a water quality gauging station, its name, the stream on which it is located, etc. This information would allow a base map to be produced. The measured information at the geographic feature is referred to as "thematic data" and would correspond to measured pollutant levels or water flow in cubic feet-per-second. There are various ways of plotting the thematic data on top of the cartographic information as described in a later section.

ORNL DWG 75-6672

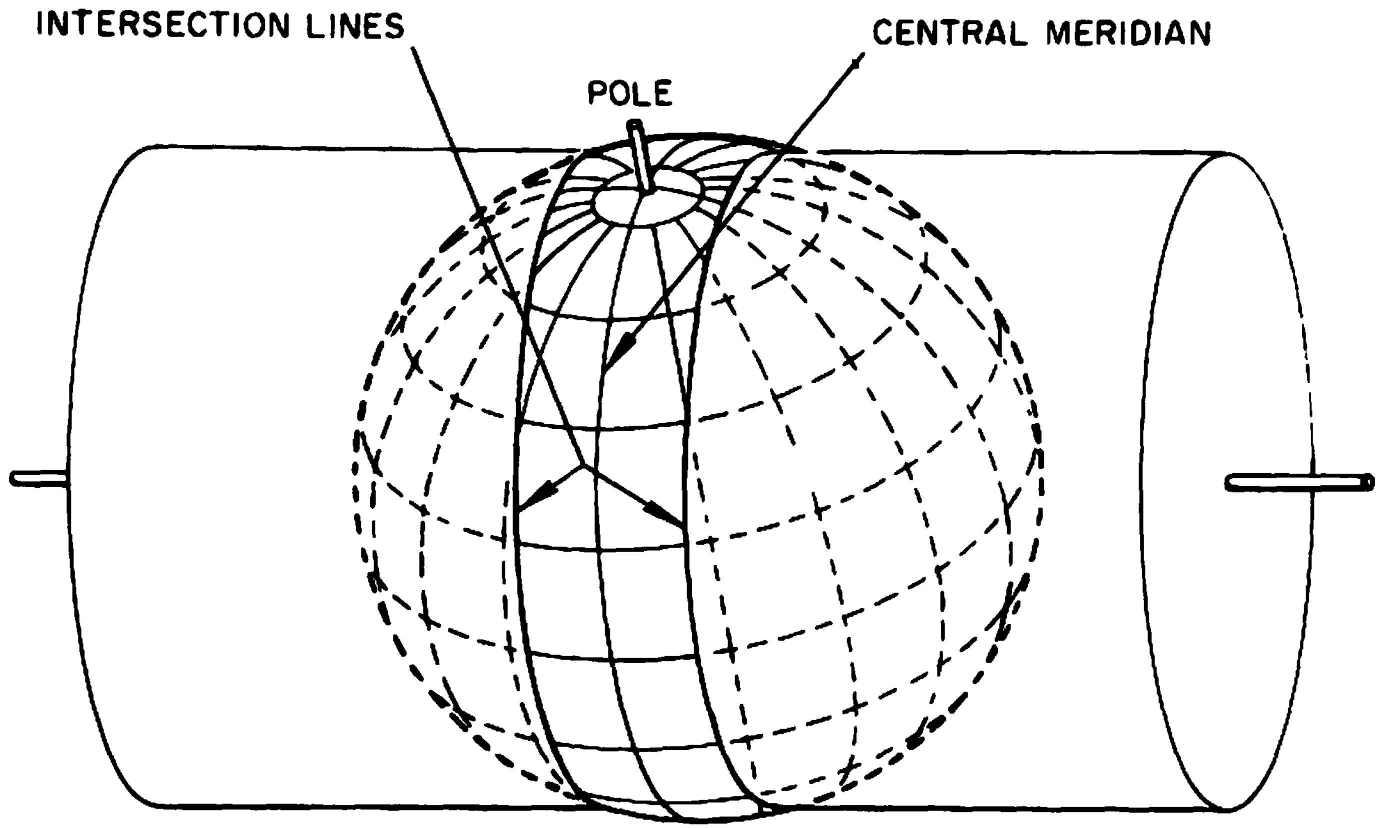
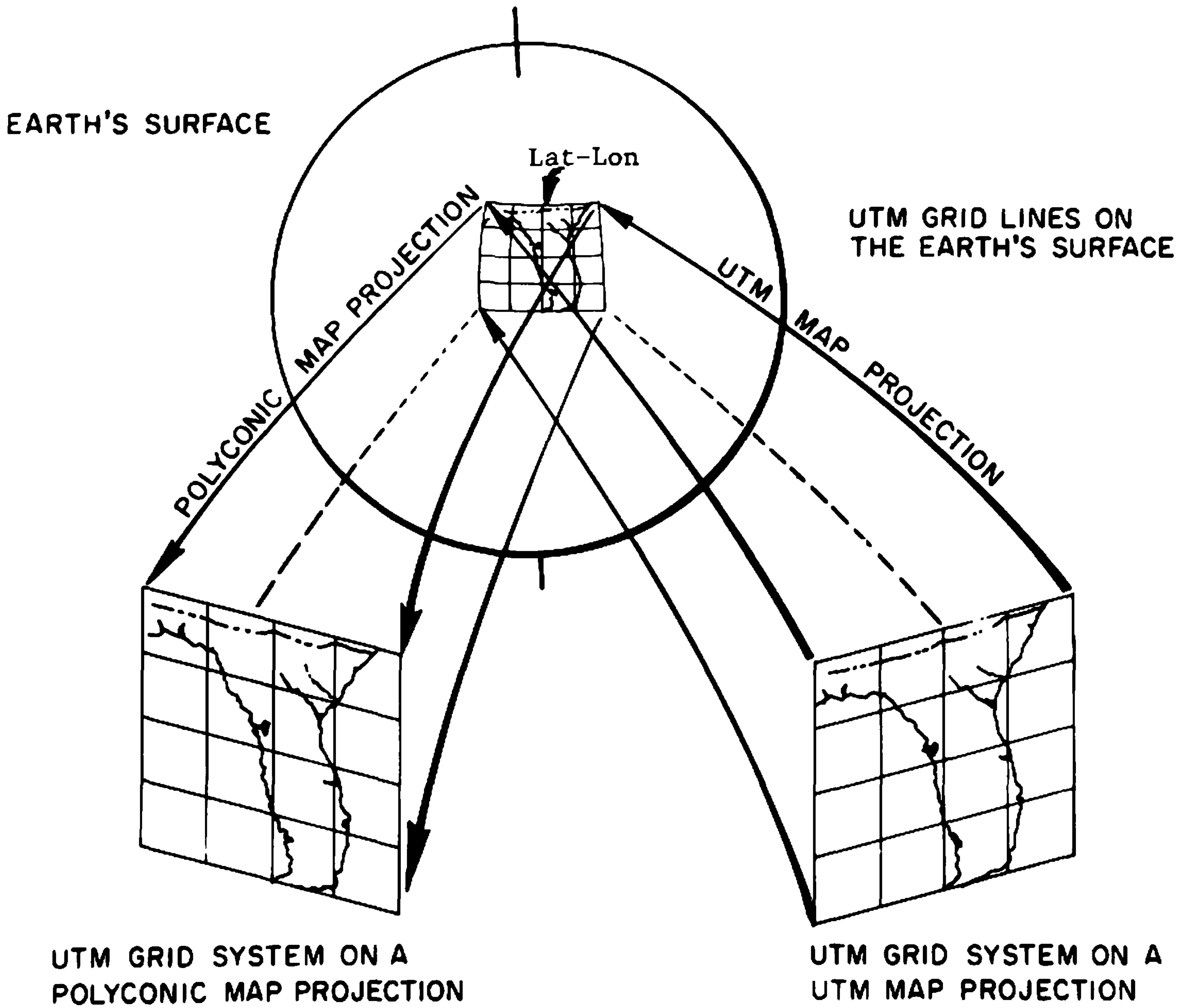


Fig. 1. Representation of Geographic Features Using The Transverse Mercator Projection.

ORNL DWG 75-6671



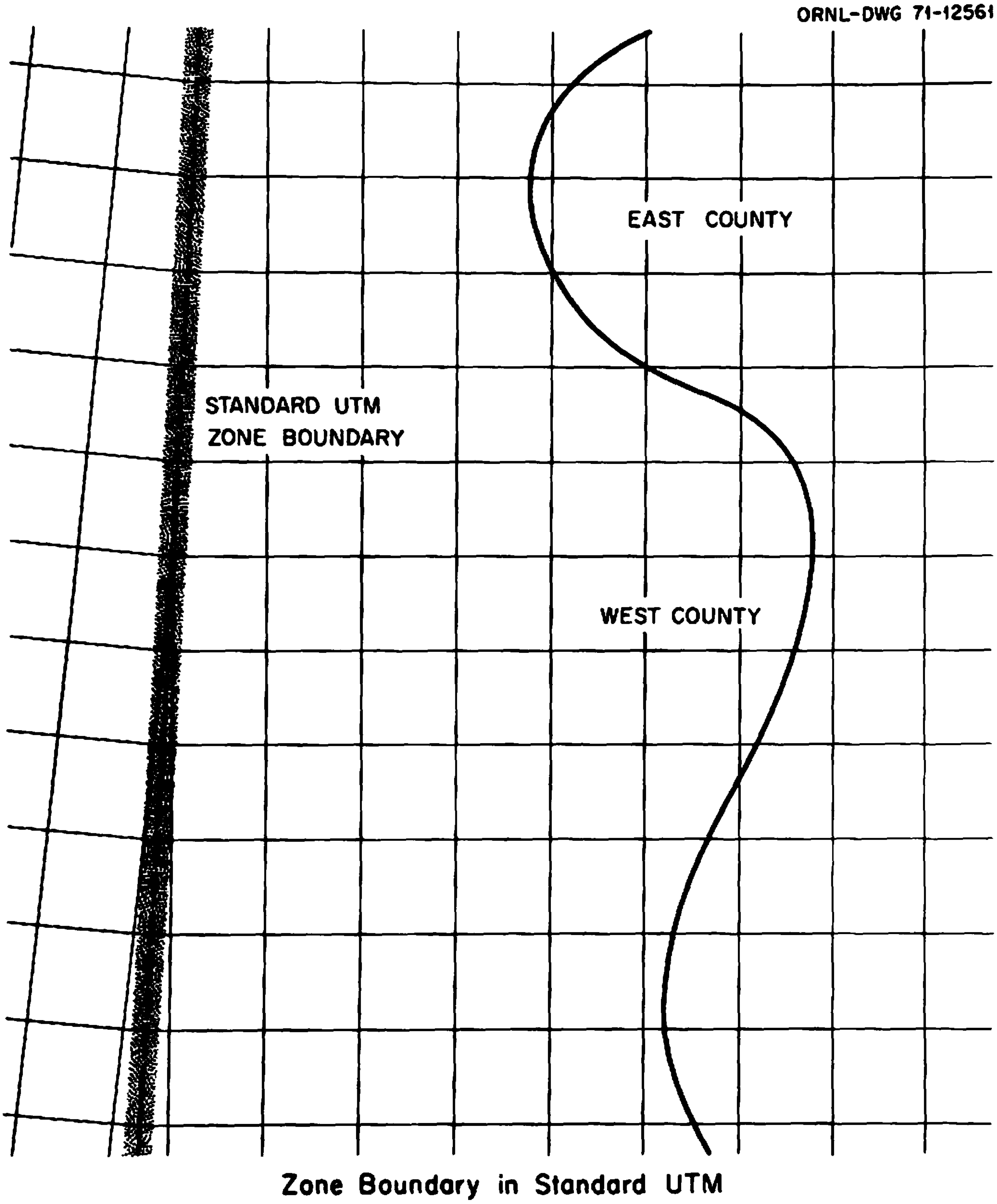


Fig. 3. Mismatch of Grid Systems Along Zone Boundary in Standard UTM.

ORNL-DWG 82-11089

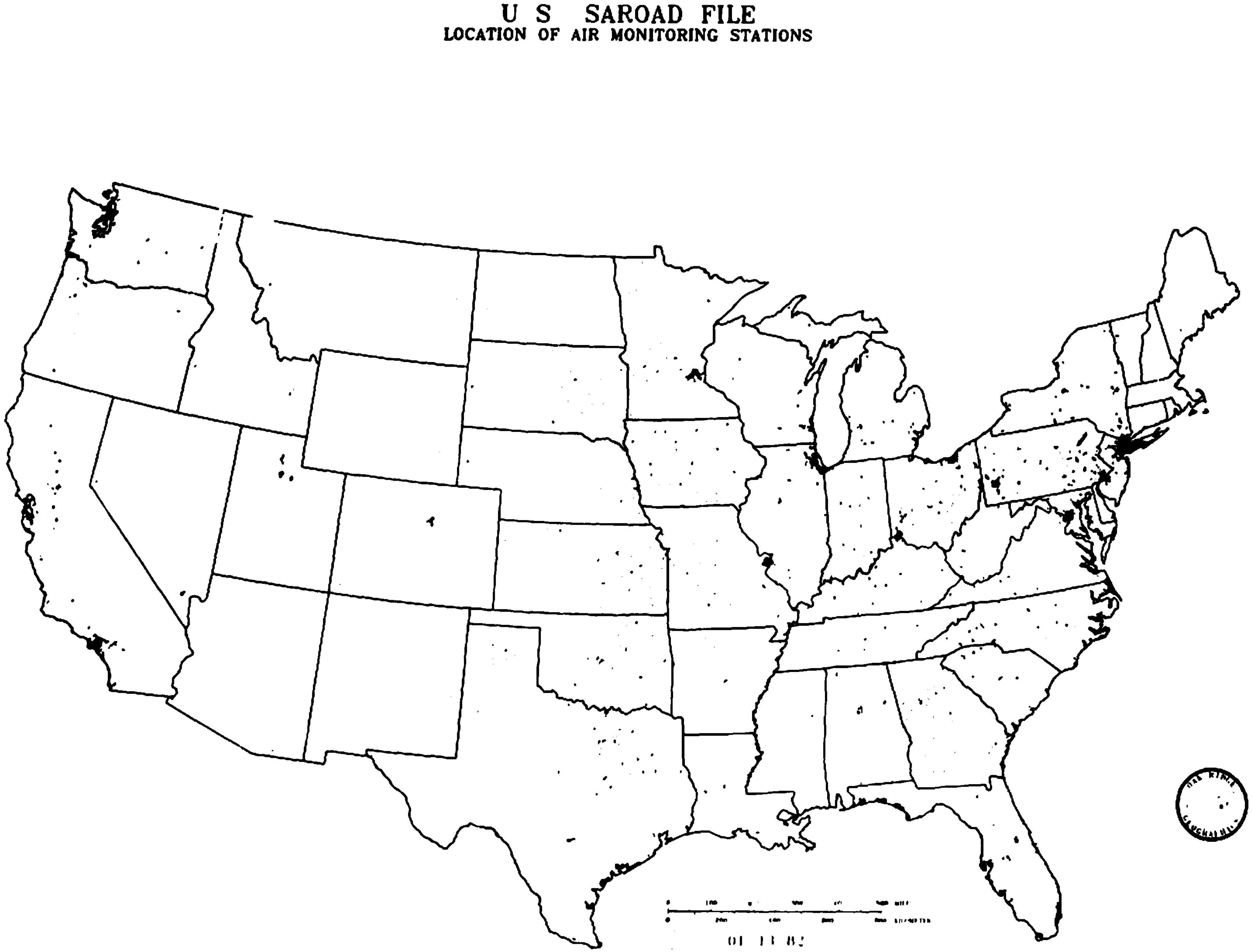


Fig. 4. Point Locations of Air Monitoring Stations Across the United States.

Point locations are used not only to represent discrete geographic entities existing at a specific location, but points may also be used to measure distributed data or continuous surface data through sampling procedures. An example might be the measurement of terrain elevations at sample points within a study area. These data may then be interpolated at a later time to create a continuous surface. The selection of appropriate data structures for representing geographic features is very important because it affects all the other uses of these data later in the application. The choice of data structures affects resolution levels, processing efficiency and costs, the types of analysis models that can be used, the types of displays that can be produced, the hardware equipment that can be used for digitizing and plotting the data, the accuracy of representing the original information, etc. A very valuable feature of any geographic information system is the capability to transform among different data structures especially when the user obtains previously digitized files from other agencies in different structures.

2.3 SEGMENTAL STRUCTURES

Other types of geographic features include linear segments (e.g., highways as in Fig. 5, rivers, railroads, transmission lines, etc.) and areas which may considered as bounded polygons (counties, landcover patterns, lakes, state parks, etc.). Although some geographic features are viewed as discrete areas bounded by well defined polygons, there are situations in which the geographic entity changes gradually over an area. For example, soils information is usually mapped as discrete units but in reality may change gradually from one soil type to another. Because it is difficult to represent such a continuously changing entity over space, discrete boundaries in the form of polygons are used. A similar situation arises with topography where the continuously changing terrain is represented as contour lineations There are several related data structures used to represent these types of features in a line segment form. An abbreviated list includes:

1. dime vectors,
2. lineations,
3. chains, and
4. polygons.

Each of these structures consists of x,y coordinates which, when connected with straight lines, depict the original feature on the map. The differences between these structures are mainly attributed to the manner in which they are to be processed and stored. Final plotter maps may look identical even though different structures were used. The remaining paragraphs describe briefly these structures. Figure 6 presents diagrams of several types including grid systems that are discussed in a later subsection.

Dime vectors are the simplest, consisting of two points connected by a straight line with a specific direction. An identifier is given for the area on the left and the right of the dime vector. This terminology comes from the DIME (Dual Independent Map Encoding) structure used by the Census Bureau for data similar to these. Groups

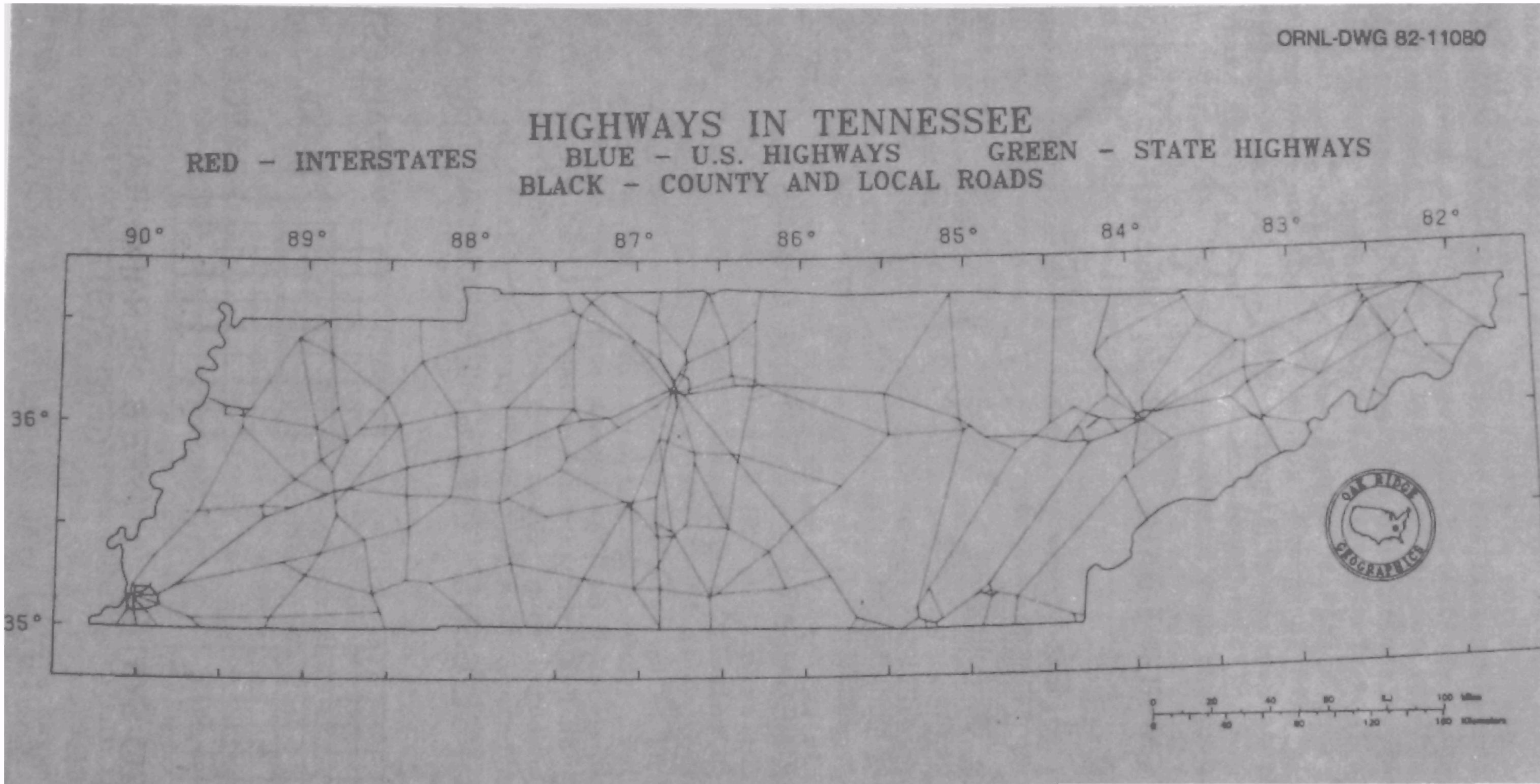


Fig. 5. Major Highway Links in Tennessee.

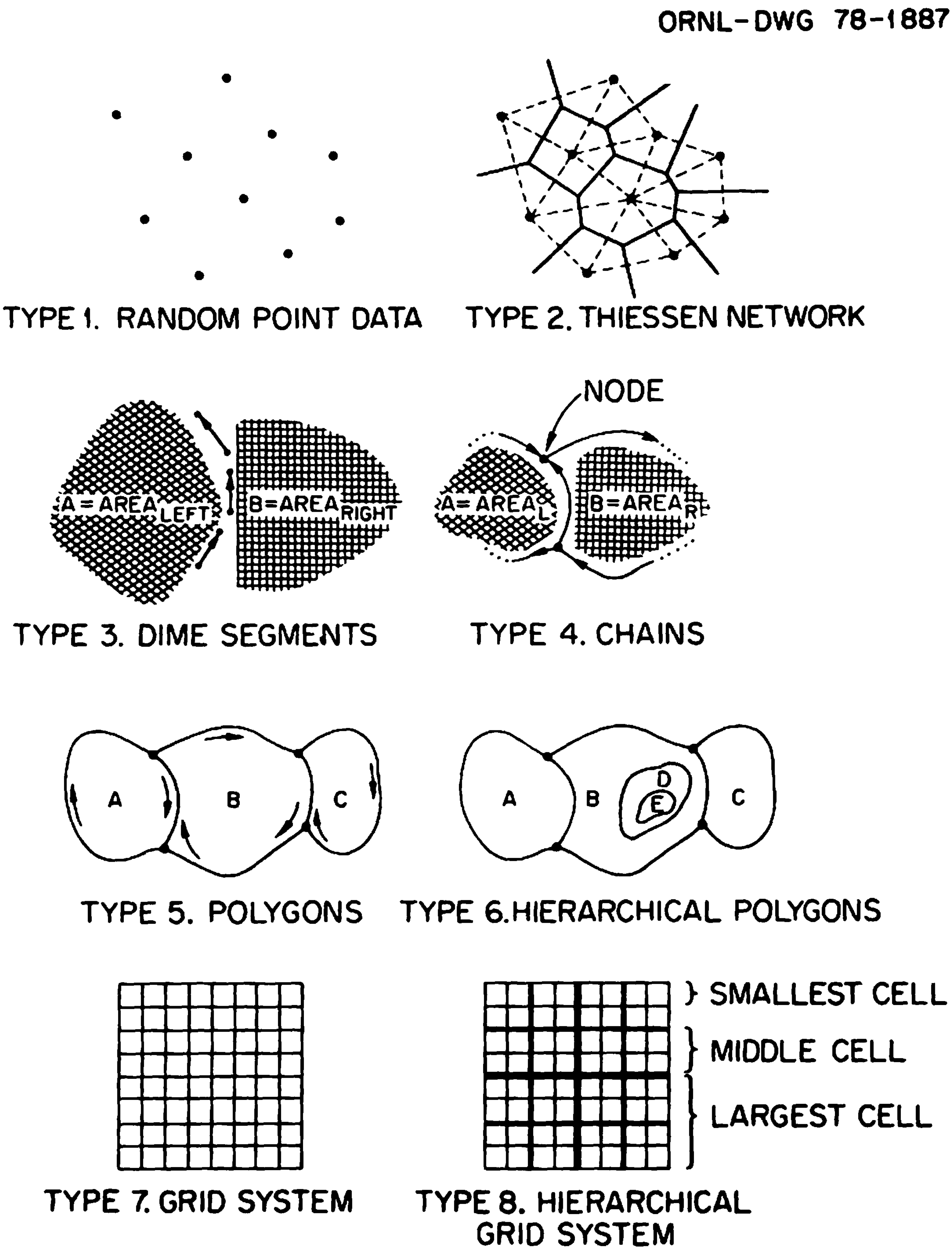


Fig. 6. Typical ORNL Geographic Data Structures.

of vectors without the area identifiers can be joined together as lineations to represent linear features such as contour lines and geologic faults. In general, lineations do not contain information to identify areas on the left or right of the linear segment. When groups of the lineations do not have topological structure describing their spatial arrangement, these data are sometimes referred to as "cartographic spaghetti". They are then used primarily for plotting purposes. When geographic features contain topological structure between the basic entities, a more complex data structure can be used to represent the spatial relationships. Groups of dime vectors may be joined together into chains with node points defined at the ends of the chains where three or more chains come together (Fig. 6, type 4). County boundaries can be easily represented with chains where the end-points of each chain would correspond to nodes at which three or more county boundaries come together. For each chain the county on the left and right would be identified.

Complete polygons can be created by simply linking together all the individual vectors around the polygon and making sure that the ending point matches the starting point (so there is no gap). One advantage in using chains over closed polygons for a data base like county boundaries is that chain coordinates are not duplicated in representing the polygons. Dime vectors are more useful for many operations because they can be sorted and processed easier than chains. Some map features have been represented as polygons inside of other polygons (Fig. 6, type 6) so that complex directories might be created to represent these relationships (depending upon the processing required). We have found that most of our spatial calculations do not require the creation of explicit hierarchical relationships among polygons but can be performed using just the dime vectors. By using sophisticated and efficient processing techniques, it is not necessary to explicitly create lots of directories and pointer lists which bog down the computations with data management and housekeeping overhead. However the topological information must be included at the dime vector level. Thus an important point that needs to be stressed again and again is that digitizing geographic data without including its topological structure (spatial relationships) severely limits the future use of data for analysis.

2.4 GRID AND RASTER STRUCTURES

One of the most commonly used data structures, especially in the early days of geographic information systems, was based upon grid cells. In this case a rectangular set of grid lines was superimposed on base maps, and data were recorded for the contents inside each grid cell. Since the grid cells corresponded to areas rather than points on the ground, different techniques were used to represent these data. One might store either a single predominant land cover or the percentage of different land cover types falling within the cell. As data bases were used at different scales (e.g., county level, state level, large regions), it became difficult to store large amounts of information for the smallest cell size in an area. Thus, hierarchical grid systems were developed whereby aggregate information was stored for large cells, with only specific variables stored at the smaller cell size for just those regions of the country requiring detailed resolution. Figure 6, type 8, shows an example. Linear features can be represented with grid cells if the cells are small enough to meet the resolution requirements. Figure 7 shows rivers, streams, and watershed basins in a grid system of one-acre cells. Notice the jagged effect from the cell edges.

The idea of using rasters or pixels (picture elements) was just an extension of the grid cell basis where the cells become very small so that the original source data are represented fairly well. As raster plotters, raster displays, and scanning digitizers became more prevalent, the use of raster images to represent geographic data was possible. Typical resolutions of 0.001 to 0.005 in. are common, although increases in computer capacities were essential to process these much larger amounts of data. Examples of raster data would include LANDSAT imagery, scanned aerial photography, cell-based topographic elevations, etc. For representing continuously changing data over a geographic region where data values are needed at a high resolution, raster arrays are the most appropriate. For sparsely spaced features of a segmental basis, one of the earlier structures (such as chains) would be more appropriate. The key to carrying out regional analyses with different types of data is to be able to combine both grid cell and segmental structures with appropriate transformations between them.

Grid systems can be defined using reference systems other than those based on rectangular axes. A more commonly used system is based on polar coordinates about a specific site with annuli or rings at different distances out from the site, and rays emanating from the site to the outer ring. A commonly used polar system contains 16 sectors corresponding to the different points of the compass around a site. Wind rose patterns are commonly described with this type of a grid system. Population counts are also accumulated in the various annuli and sectors of such a system as shown on the left half of Fig. 8. The right half shows population density contours from which the counts-per-sector are calculated.

An example application using multiple data structures is the study and modeling of sulfur dioxide pollution from industrial plants. The pollutants might be transported over large regions, followed by deposition on the earth's surface resulting in environmental problems caused by the acid rain. The surface runoff and water transport mechanisms may pollute bodies of water, thus affecting aquatic habitats. Data bases that might be used in this study and their appropriate data structures are listed.

1. Industrial locations	Point data
2. Air monitoring stations	Point data
3. Meterological stations	Point data
4. Political boundaries (such as counties)	Chains
5. Stream and river locations	Lineations or networks
6. Satellite imagery	Raster arrays
7. Landcover polygons	Chains or dime vectors
8. Reservoir and lake boundaries •	Polygons
9. Water quality stations	Point data
10. Air dispersion patterns	Contour lineations
11. Deposition patterns	Grid cells
12. Population distributions	Grid cells

ORNL-DWG 75-13578

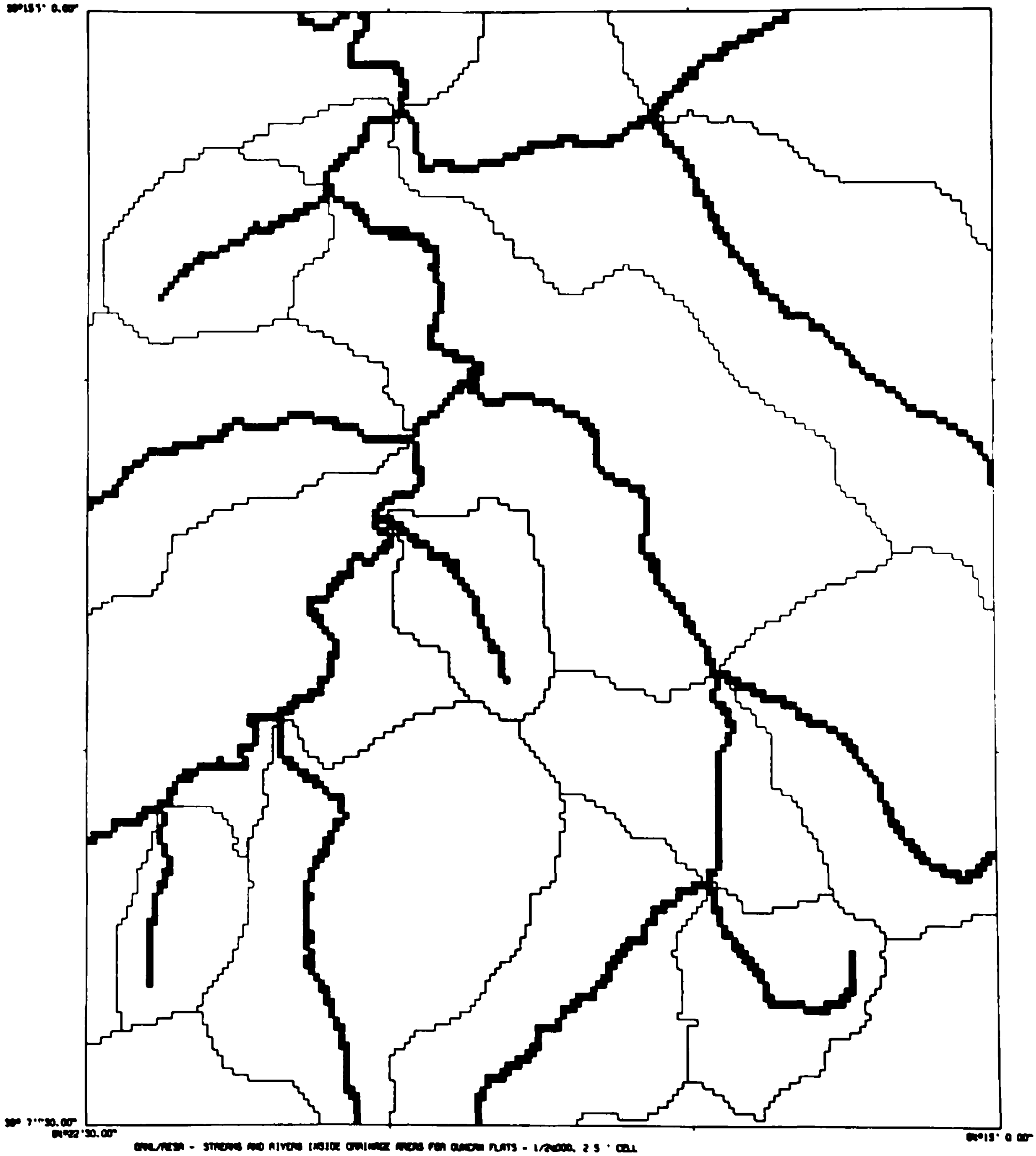


Fig. 7. Streams, Rivers, and Watershed Basins In a Grid System.

ORNL-DWG 83-7794

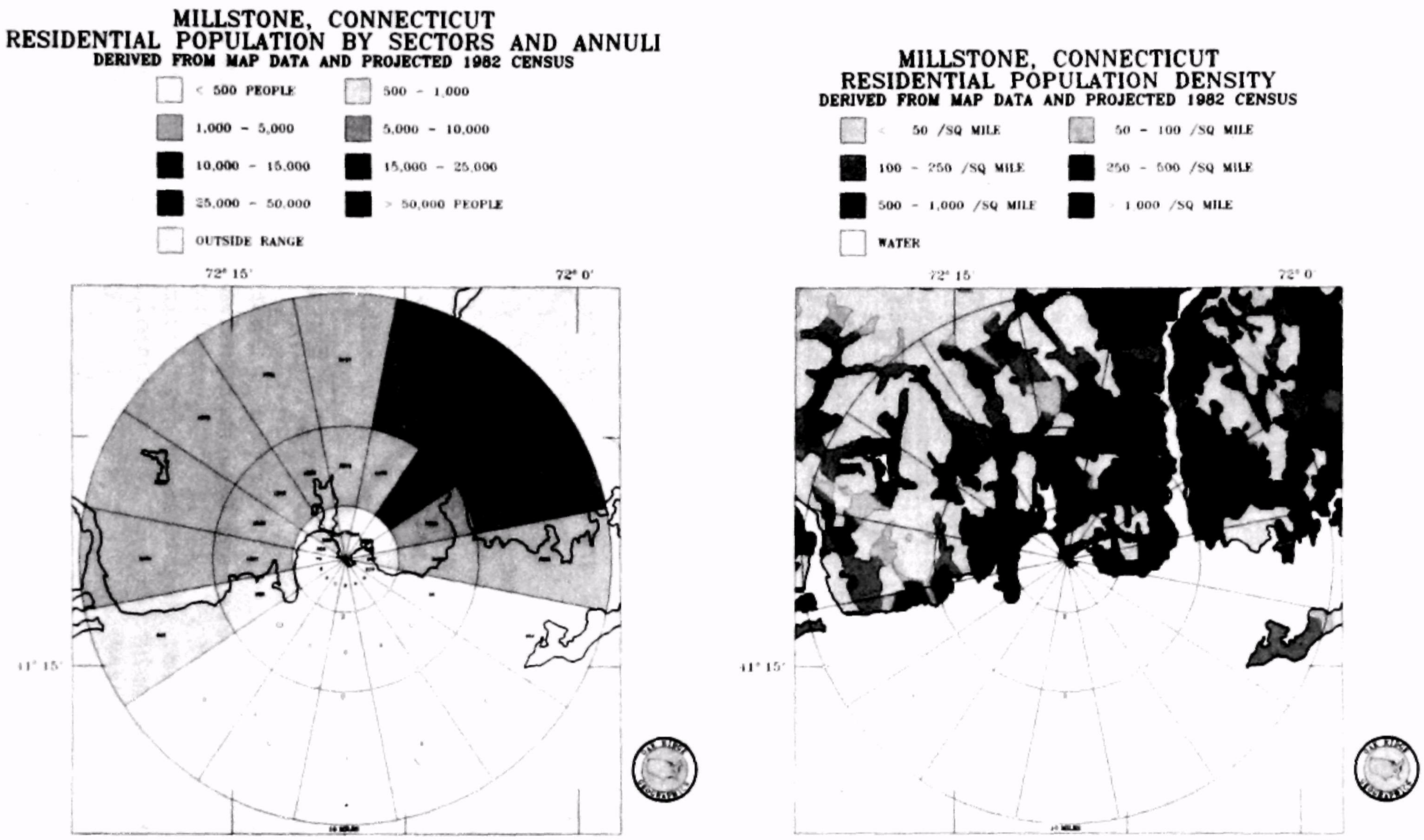


Fig. 8. A Sector-Annuli Grid System For Representing Population Distributions.

The ability to input different data structures and then transform among the structures for performing the calculations is a very important part of regional analyses studies. A later section will discuss different types of transformations and analyses.

2.5 NETWORK STRUCTURES

Brief mention will be made of network structures for representing features such as highway networks, railroad networks, stream and river systems, and networks that may be internally generated to establish spatial relationships between geographic entities. Lineation networks could be used to represent collections of highways or transmission lines which are joined together at nodes and may require a direction of flow. In order to analyze these networks for applications such as transportation routing of shipments across the United States (U.S.), it is necessary that the topological links and possible flow directions be represented in the data structure. Additional information is needed in developing a network of streams and rivers to identify the order of the stream from smallest to largest. Information is needed depicting node points where streams join together, which ones are upstream, the direction of flow, and possibly, the watershed area that each stream drains. If the flow travels in only one direction, a hierarchical tree structure as shown on the left half of Fig. 9 is sufficient. If there is no restriction on the flow direction as in an electrical transmission or highway network, any node may be connected to any other node. Thus, a tree structure is unsatisfactory. If a set of nodes are designated as sources and sinks, or are set to given potentials, the flows in the branches are established although the ordinary hierarchical tree structure may still be unsatisfactory.

Other types of networks can be internally generated in the computer to aid in spatial transformations. For example, interpolation to create a continuous surface from a set of random input points can be aided by constructing a triangulate network between the points. This network identifies the immediate neighbors around each point and the proximal or influence areas around each point to aid in the interpolation. Figure 6, type 2, shows such a network, with the dashed lines being the triangulation linking points and the solid lines being the proximal polygons. To calculate an interpolated value for a new point, the network quickly identifies the immediate neighbors, weighting values based on distance, and influence areas that surround the interpolated point.

2.6 THREE-DIMENSIONAL STRUCTURES

Geographic applications that deal with subsurface geology or air transport phenomena must consider three-dimensional data structures to represent the data and processes taking place. One technique is to use two-dimensional slices through the volume being studied and prepare profile plots or contour maps at different intervals. If the information varies continuously throughout the three-dimensional volume, another approach is to actually subdivide the volume into small units or compartments sometimes referred to as voxels (similar to pixels in the two-dimensional case). Usually the voxels are made fairly large if the data resolution will permit it, or else sample volumes are studied throughout the 3-D space with small voxels used in the samples. In a small 2-D LANDSAT application the user may be dealing with 250,000

ORNL-DWG 76-10899

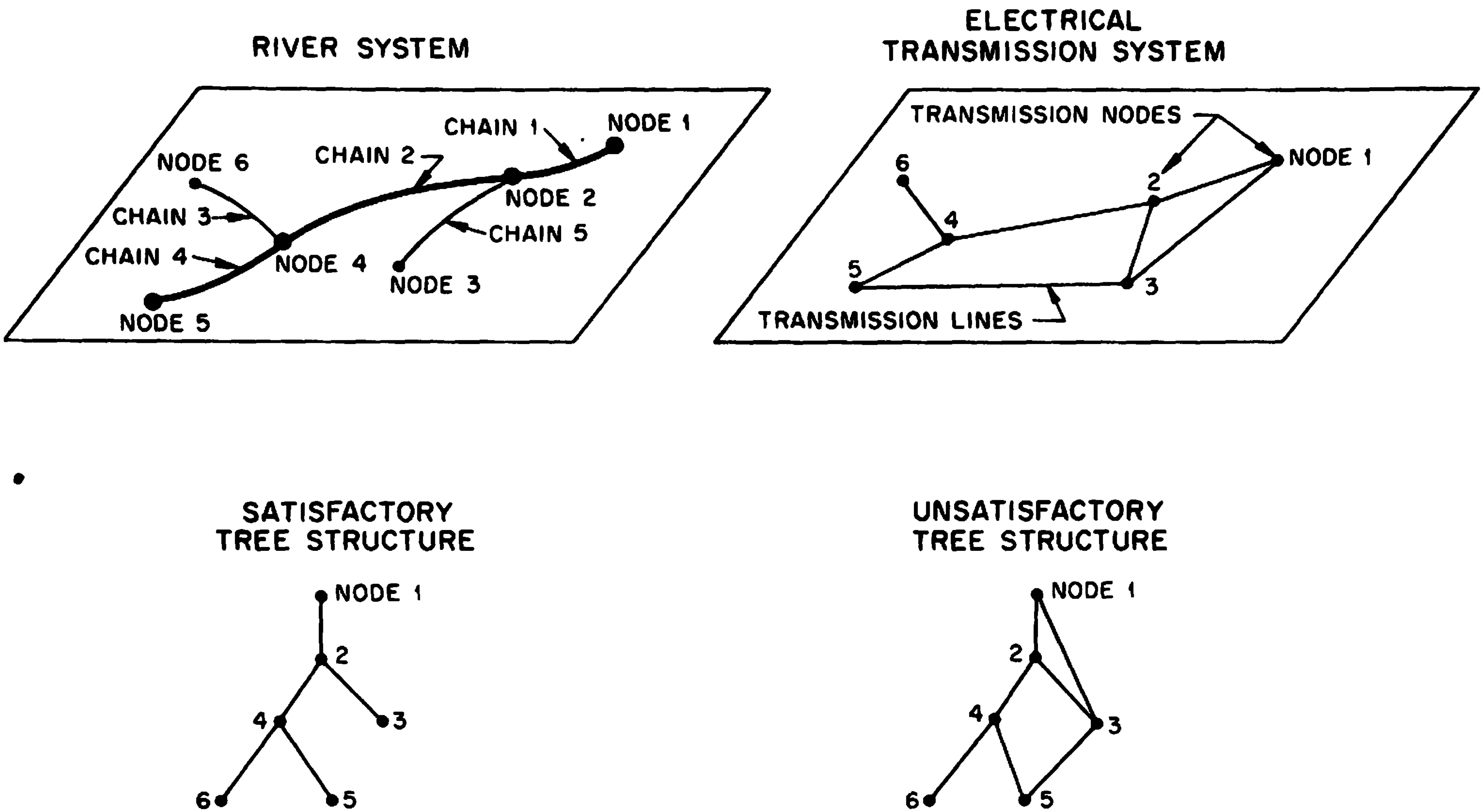


Fig. 9. Representation of Linear Network Structures.

pixels on the earth's surface to determine landcover. However, if this same resolution were used to study atmospheric phenomena to an altitude of 10 miles, then 12,000,000 voxels would be required. Because of large data requirements, these types of calculations are not normally done. Other three-dimensional geographic data structures are and will be developed in the future, perhaps based on hyperplanes or other arbitrarily shaped partitioning of three-dimensional space.

The reason for discussing all these different types of data structures is to help the user in the planning stages of his/her application. It is important to stress that the questions being asked of the data must be formulated before selecting a data structure and beginning a data collection effort. If the intent were to calculate sensitivity indices by combining variables such as landcover patterns, soil polygons, topography, weather patterns, acidity contours, etc., there are two primary choices: one using grid systems and the other using segmental (polygon type) systems. Discussion of the processing pros and cons are given in a later section.

3. DIGITIZING AND GEO-EDITING TECHNIQUES

3.1 MANUAL GRID OVERLAY

In the early days of geographic information systems, the most common method for inputting spatial data was to use a grid superimposed on the base map from which data were to be collected. In many cases this grid was hand drawn on acetate material with row and column numbers to identify the matrix of cells. For sparsely spaced data, each grid cell containing a desired feature (e.g., factory locations) was coded with a row, column number, and the identifier for that particular feature. For inputting polygonal data, a run-length encoding procedure was used whereby a starting and ending column number was given for each polygon, along with the identifier associated with all cells inside that polygon. This was done on a row-by-row basis. For example, if a given row intersected three land cover polygons, only the column numbers for six different cells had to be coded along with the land cover type between the starting and ending column numbers within each polygon. This technique was much more efficient than having to code an input number for every cell in the grid. The computer is used to fill in the intervening cells. Other manual gridding techniques have been devised such as delineating polygon borders by giving the row and column numbers around each polygon and a single identifier code inside the polygon. The computer was then used to fill out all cells falling within the polygonal border.

The use of these type techniques required very little computational power or sophisticated software but used a lot of technician manpower to code all the data. Costs were usually measured in a few cents-per-cell. In order to keep the costs to a minimum, the cells were made as large as possible while still preserving a reasonable resolution for the immediate application project. In some cases, as projects proceeded, it was determined that the cell size was not sufficient for the questions being asked or that the costs were larger than expected. In order to change the grid system, it was then necessary to recode all previous data. From our experience we have found that certain data bases created for one application may be used repeatedly for other projects that were never even anticipated at the time the data were collected. Manually gridded data are usually of less use later on because they are collected at a minimal resolution level for a particular application. However, data bases that have been digitized with a high degree of resolution using automated techniques, can be aggregated to different resolutions for different applications. In most cases it is important to maintain the original digitized data since they represent the finest resolution a user can refer back to.

If manual gridding techniques are used, it is important that some standard coordinate system be selected for the grid lines (e.g., UTM, State Plane, Lat-Lon). This will allow the data to be geographically referenced to other base maps and data files. For maps of large regions, it should be noted that the grid lines may be curved depending upon the map projection used. This can be a problem when the grid has to be drawn by hand. It would be much better to use a computer to plot the desired grid on mylar using the appropriate map projection and coordinate system.

3.2 COORDINATE DIGITIZING FROM X,Y TABLETS

Some of the most commonly used systems for digitizing geographic data from maps and aerial photographs are based on x,y tablets with a hand-held cursor. As the cursor is moved on top of the map, x,y coordinates are measured by the tablet and sent to a minicomputer or microcomputer for processing. Figure 10 shows a typical digitizing station. Tablets come in a variety of sizes ranging up to several feet on a side and measure the coordinates in various fashions. In some cases, there is a grid of wires beneath the tablet surface which, when strobed by timing pulses and detected by the cross hair cursor, allow for accuracies of 0.001 in. Tablets normally work in one of two modes, point mode or stream mode. In point mode, a coordinate is sent every time the button is pushed on the cursor; whereas in stream mode, coordinates are sent continuously as long as the button is held down.

In addition to digitizing the coordinate locations it is necessary to put in identifiers or thematic data associated with the geographic features. This may be done through the use of special menus on the side of the tablet or by typing the information on a keyboard after each feature is digitized. A spatial display of the digitized data is normally viewed on a CRT screen for error detection and correction. This is usually carried out through the use of a CRT cursor on the screen (controlled from a joy stick, mouse, or light pen, or through the use of further data digitized from the tablet). At ORNL, a special system has been built with a video projector mounted above the tablet so that the CRT image is projected directly on the source map as data are digitized (Fig. 11). In this way a green light-beam traces out on the map itself the boundaries digitized by the hand-held cursor as the operator moves it around. This allows for direct interaction with both the digitized image and the source map at the same time. A CRT graphics monitor is also used. For detailed interactive editing of large segmental data bases, an additional system has been developed using a high-resolution graphics CRT linked to a minicomputer. The operator manipulates all the spatial data directly in map form on the screen using a joy stick to locate features, move data around, window-in on interest areas, rearrange or edit chains or polygons, correct attribute information, etc.

The types of features digitized from base maps with a tablet are either point locations or line segment features such as polygonal boundaries, contour lines, highway networks, and geologic fault lines. For digitizing area-type data, polygonal boundaries are normally drawn around the different areas and an identifier assigned to represent a homogeneous feature within the polygon. If continuously varying area data (e.g., reflected color intensities of ground cover) are to be digitized, an x,y tablet is not normally used since a raster array is a more appropriate data structure for storing this information. In this case, some type of scanning device, as discussed in the next section, should be used. It is possible to digitize gridded data with an x,y tablet by positioning the cursor inside each cell on a gridded acetate overlay and typing in the cell contents. This is not a very efficient process especially if most of the grid cells contain non-zero data items for input. An efficient mechanism for creating gridded data is to digitize the raw information as linear segments or point data, and then transform the raw data into gridded information by performing a polygon-to-grid calculation or an interpolation (if the raw data were

ORNL PHOTO 1136-81

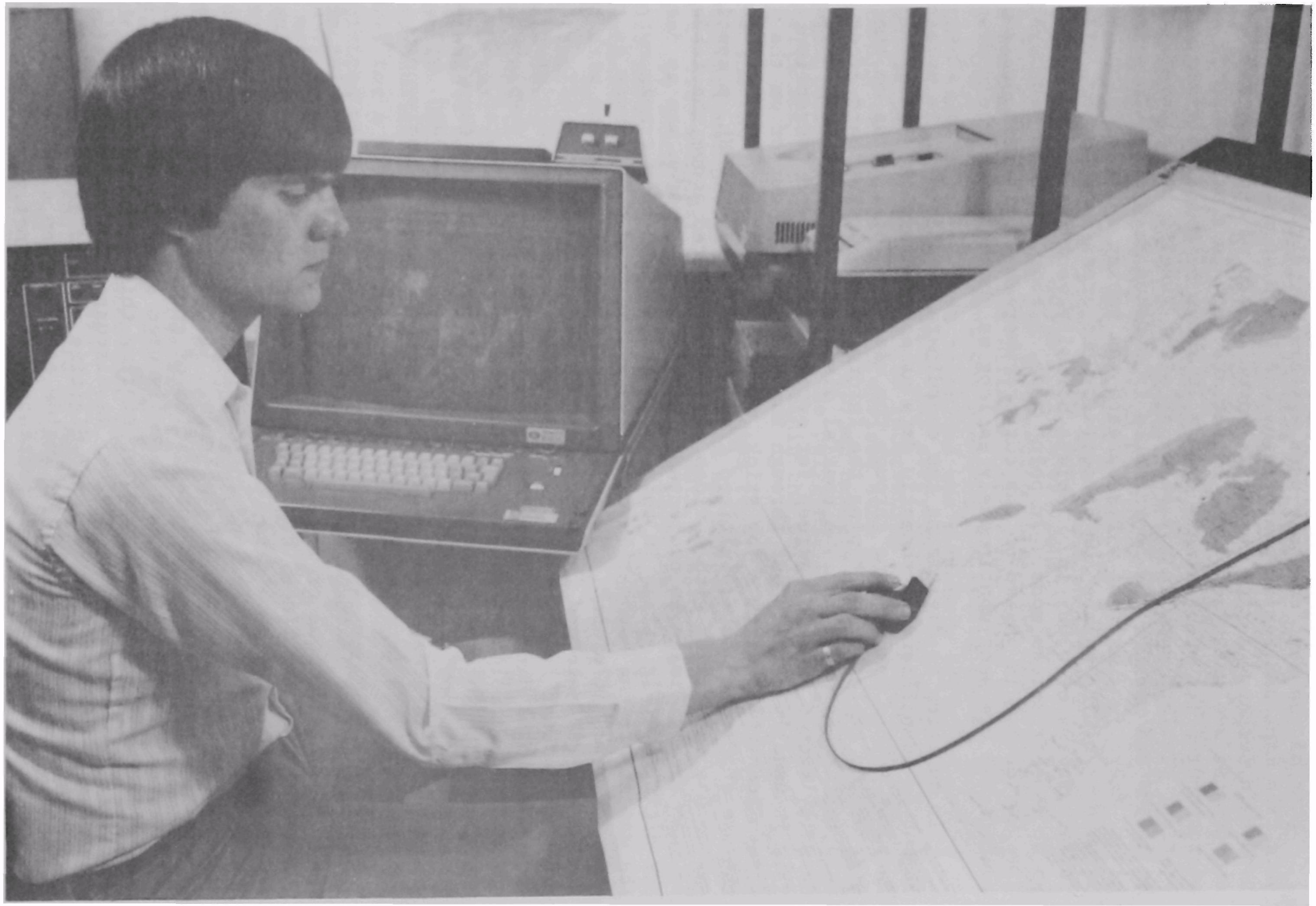


Fig. 10. Tektronix Digitizing Station Using x,y, Tablet and a CRT Graphic Display.

ORNL PHOTO 1130-81

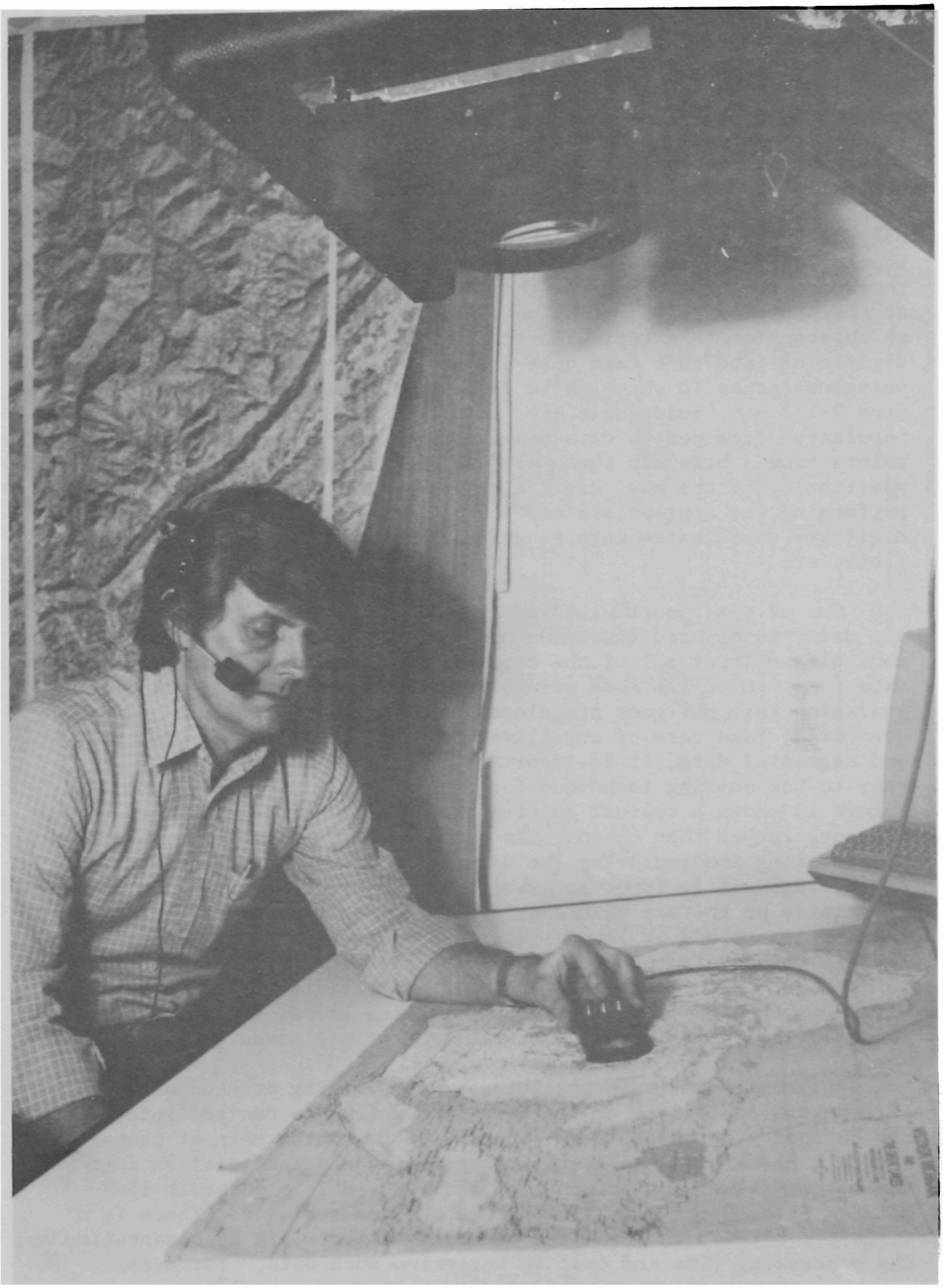


Fig. 11. Video Projection Digitization System at ORNL.

random points). This operation is more cost-effective and much less time consuming than having to input detailed gridded data by hand.

The cost of tablet digitizing is very difficult to estimate in a generic sense because the time is so dependent upon the type of data being captured. From past experience digitizing jobs can range from a couple of hours to several months in duration. A large job at ORNL required digitizing all the coal fields across the United States from a large national map which was divided into sections and photographically enlarged. These coal fields were to be delineated by type of coal and commercial feasibility of mining the coal. Because of the intricate shapes and detail in each of the polygons, the number of different polygon types merged together, and the need to preserve maximum topological structure in the data, it took a computer technician around three man-months to complete the job. Figure 12 shows an output plot at the national scale. (The detailed resolution and complexity are not as apparent at this scale.) At the other extreme, there are digitizing jobs that take only one or two hours to input exclusion polygons (areas in which there is no residential population) digitized from 7-1/2 min. quadrangle maps to aid in the distribution of population from census data bases. Even for digitizing just a few points from a base map there are certain minimal costs associated with positioning of the map, digitizing latitude-longitude control points, performing the appropriate map projections and transformation of digitizer coordinates into geographic coordinates, producing test plots, etc.

One of the important advantages in using tablets for digitizing map data, as opposed to raster scanners, is that the operator can control and input all of the topological structure needed in these data. He can define node points with the appropriate identifiers, guarantee that polygons are closed, assure that chains will be simply connected, take care of any sliver problems, etc. For complex polygons and segmental data, it is important to have an efficient and easy-to-use editing technique for handling these types of problems. Figure 13 shows a typical editing problem when digitizing independent polygons rather than chains. In some cases the operator may spend more time editing and improving the data than was spent in the original digitizing. It is important that the topological information be identified on the map before or during the initial digitization rather than trying to add it later. Digitized files which just contain strings of coordinates useful only for plotting purposes are referred to as "cartographic spaghetti". It is very difficult to do analysis with this type of data, such as calculating areas, shading polygons, intersecting polygons, and calculating proximal areas.

In some cases it is possible to use sampling procedures or surrogates (other less costly data from which the desired information can be approximated) to significantly decrease the cost of preparing new data bases. Generally the polygonal type data digitized from x-y tablets represent the original map features more accurately than gridded data, especially if large cell sizes are used. There is a trade-off between the need for accurate cartographic representation and the processing time and cost in analyzing such data. For example, it may be easier to combine gridded data sets than do polygon intersections, yet gridded output maps might not be of sufficient accuracy and resolution. These items are discussed in a later section.

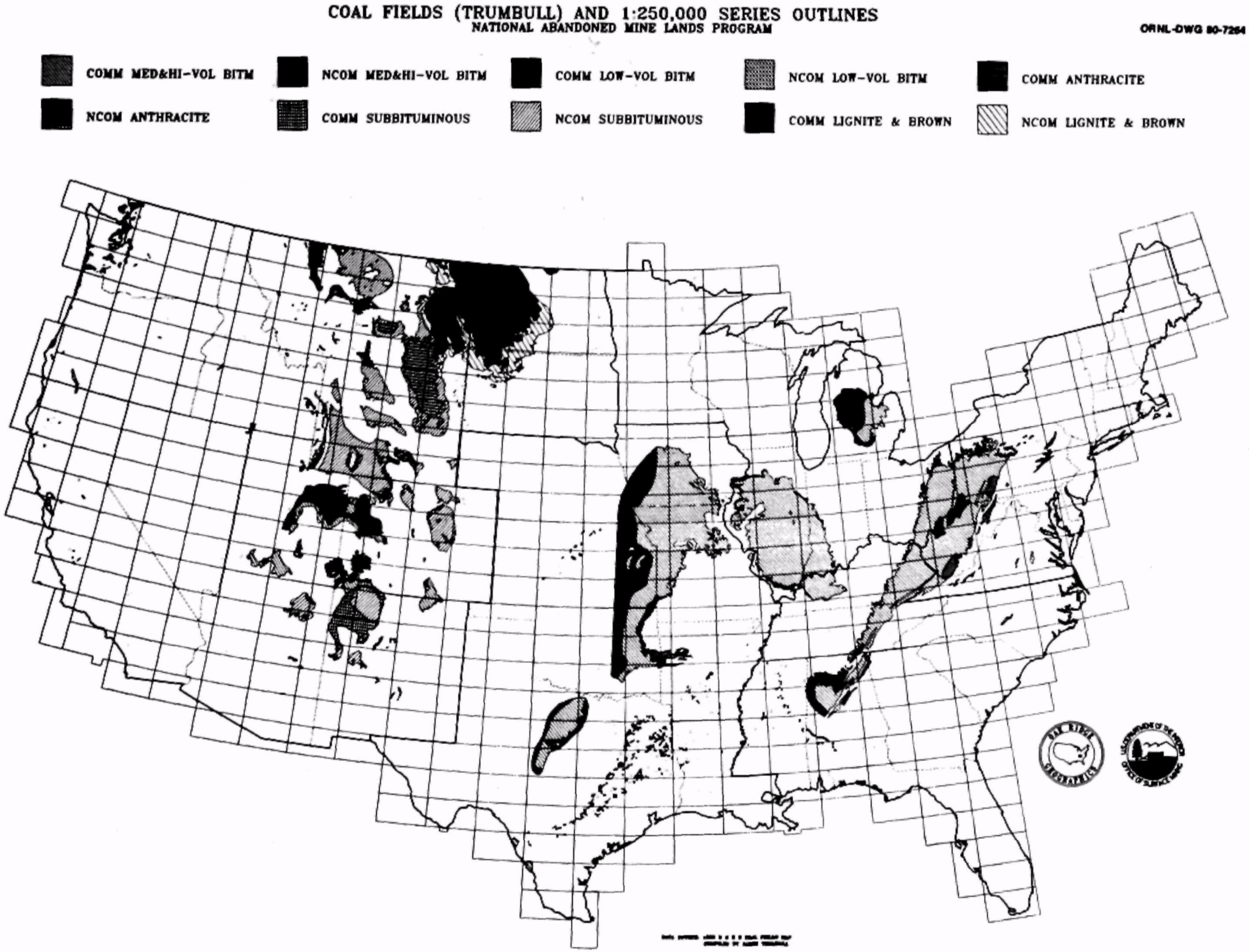


Fig. 12. Digitized Coal Fields by Type of Coal and Commercial Feasibility of Mining

ORNL-DWG 76-10893

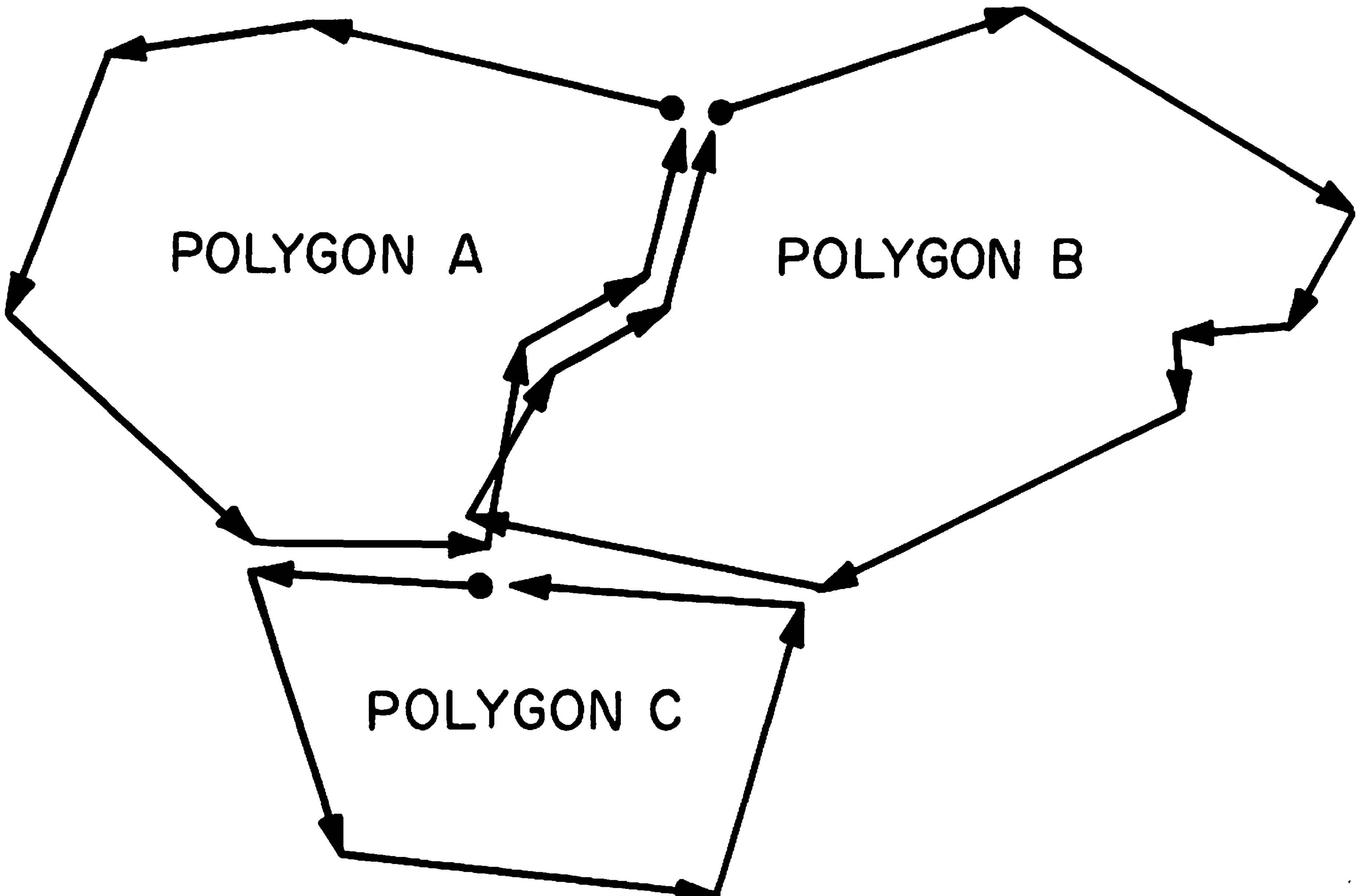


Fig. 13. Digitization Problems With the Independent Polygon Method.

3.3 RASTER SCANNING OF SOURCE IMAGES

In more recent years there have been several techniques devised for scanning maps and aerial photographs using raster-oriented devices that record a data value for each pixel element (small grid cell) on the image. One type of device consists of optical scanners which measure reflected or transmitted light intensity through the source image. These scanners convert the voltage readings into digital integers usually ranging from 0 to 255. If a transparent image is used as a base, then the light source is transmitted through the image and recorded by the scanner. If the source document is opaque, a reflective light source must be used to illuminate the document for the scanner. Scanners of this type generally work with smaller source documents, less than a foot or two in size. Optical densitometers give very high accuracy and small spot size with large amounts of data. Scanning time may take from a few minutes to a few hours. Video scanners such as the ORIDS vidicon scanning camera (Fig. 14) are generally designed to be interactive with some type of an image processing system. The scanning takes place almost instantly but does not have as high a resolution or accuracy as the optical scanning densitometers. The advantage is that the digitized data can be viewed on a CRT screen before it is actually captured so that interactive operations can be carried out quickly. A typical scanned image will range in size from 512 x 512 to 2048 x 2048 pixels. Resolutions will vary depending on the device, the spot size, and the distance of the scanner from the source document. Typical ranges might represent 100-500 rasters/in.

Newer types of scanners have been developed using laser technology. These may be either large flat bed or drum scanners and have been designed to accept source documents up to three or four feet on a side. They are usually much more expensive ranging from $50,000 and up in cost. They do produce very accurate high resolution raster data with typical scanning times of 20 to 45 min. All of these devices are recording grey-level readings associated with each raster spot on the map. Vector-oriented laser devices have also been developed to operate in a line-following mode, whereby they can locate a given black line and follow along the line, using a predetermined search procedure when junctions or the map border are hit. This technique eliminates some of the manual work required in using a tablet since an operator does not have to trace along the individual lines. However, it is not possible to input all the topological structure (e.g., polygon on the left and polygon on the right) even with the line-following laser, so further work is necessary before the data can be used for spatial analysis. The source document has to be prepared accurately and consistently (e.g., no line gaps).

One of the more common geographic uses of scanned raster data is to enhance or classify the data into categories that are meaningful in particular applications. For example, aerial photographs can be scanned to input different wave lengths or "bands" of light (by using color filters) so that landcover patterns can be classified or enhanced using normal image processing techniques. As with multi-spectral satellite data, the classifications are not completely accurate and

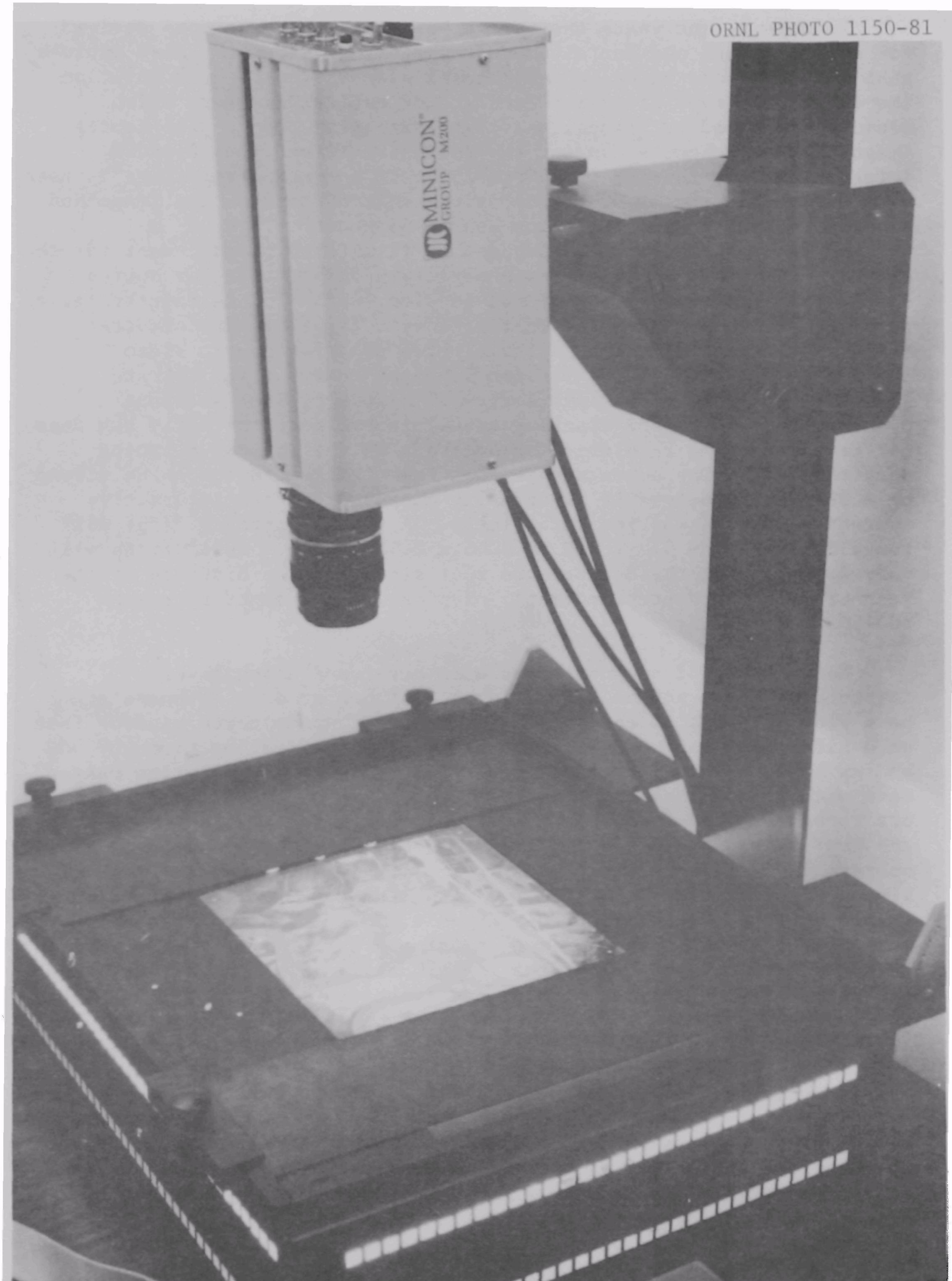


Fig. 14. Vidicon Input Scanning Camera (for Photographs, Maps, and Transparencies).

futher editing may be necessary. The data structure may be thought of as a grid system with very small cells. Because of the large quantity of data, efficient data processing techniques should be used.

Attempts have been made to transform raster data into line segment form, but they usually require significant amounts of manual editing because of limitations in the pattern recognition algorithms and poor data quality. For example, transforming a black line (captured as rows of pixels) into x,y coordinates as would be digitized from a tablet is rather difficult. The line width is usually several pixels across and the "blackness" of the line usually changes in intensity and may have gaps. Several firms have developed scanning-vectorizing systems, especailly for the CAD/CAM market. (A rather expensive system using a laser scanner has been built by Scitex American Corporation of Bedford, MA.) Extensive effort is also required to assign the attribute and topological information to the vectorized data. Experimentation has been carried out at ORNL to video digitize landcover maps in color. Attempts were then made to classify each color into a separate recognizable category with the intent of aggregating pixels inside each polygon into contiguous units whose borders could then be automatically generated. It was found that the computer was so sensitive to lighting and color variations from one part of the map to another that further work is needed to calculate accurate classifications of all pixels in each polygon.

Another function normally performed with scanned data consists of geometric correction or rectification. The digitized raster units are not associated with a geographic coordinate system and cannot be immediately combined with other geographic data. Many times the digitized data are geometrically distorted and must be corrected through "rubber-sheeting" techniques. (i.e., as if the image were made of a sheet of rubber which could be stretched to fit a geometrically correct base map.) Thus ground control points are identified in the digitized data (e.g., road intersections, building locations, etc.) for comparison with the same ground control points collected from base maps of known projections. Transformations can then be built to correct the scanned data so that the individual pixels can be referred to by latitude-longitude or some other geographic coordinate. In picking ground control points from the digitized data, it is necessary to have a computer display or plotter map to work from. These may take the form of black-and-white grey-level plots from electrostatic plotters, line printer maps, or CRT raster images with a cursor that can be positioned on top of the control point on the CRT screen.

The cost of scanning base maps or aerial photography is usually much less than digitizing similar polygonal data (e.g., land cover polygons) on a tablet digitizing system. However, the information content may not be sufficient to allow for spatial analysis of the scanned data. Also, the data processing costs are normally much higher with the scanned data because every spot on the map corresponds to a data element. An example can be used to aid the discussion. Assume an investigator wishes to combine county boundaries with vegetative patterns in a study region. Two ways might be considered for digitizing the vegetative cover. One would be to scan aerial photographs which then might be classified into different categories, hopefully similar to the vegetative patterns desired by the investigator.

The raster image would have to be rectified to match the county boundaries, or the boundaries warped to match the raster image. The second technique might delineate the actual polygons representing the different categories on the photograph or a base map, and then digitize these outlines using a tablet digitizer. If the question is to calculate acreages of the different vegetative patterns by county, either data base could be combined with the digital county boundaries to tabulate these statistics. In the first case, each pixel falling inside a county boundary would be tabulated by category. The number of pixels multiplied by the area per pixel would give the acreages desired in each county. In the second case, the vegetative polygons would be intersected with the county outlines; and areas would be computed from the resulting polygons. Assuming a proper classification could be calculated for the raster image, it would be better to use the first approach for this type of analysis question. However, if an additional requirement were to produce a computer map showing each vegetative cluster within a county, color-coded by its acreage and labeled by its type, the second case should be used. Each polygon resulting from the intersection could be colored, based on its size, and the label automatically inserted to identify the vegetation type.

Based on past experience, the appropriate mechanism for digitizing geographic data is very dependent upon the types of questions that are going to be asked of the data. Thus, the investigator should spend sufficient planning time to determine the analyses needed before choosing data structures and digitizing techniques. A combination of raster and polygonal techniques are sometimes most efficient.

3.4 CRT CURSOR DIGITIZING

Another technique used for digitizing geographic data integrates both raster and vector (or line-oriented) characteristics. This technique is only applicable where the source document can be captured in a raster image format quickly and displayed on a CRT screen. The operator moves a cursor around the image to delineate features of interest, perhaps after they are enhanced in color. This is shown on the ORIDS image processing system in Fig. 15 with a cursor in the lower right corner of the screen. The technique uses raster data as the base information with vector data created by the operator on the CRT screen. For example, if the user wished to delineate polygons of recently flooded areas from LANDSAT data, this approach might be appropriate, especially if it was difficult to automatically classify these areas with image processing techniques. The water bodies or wetland areas could be enhanced in color before vector digitizing. This is not a production oriented technique but is useful only in specific situations.

ORNL PHOTO 1148-81

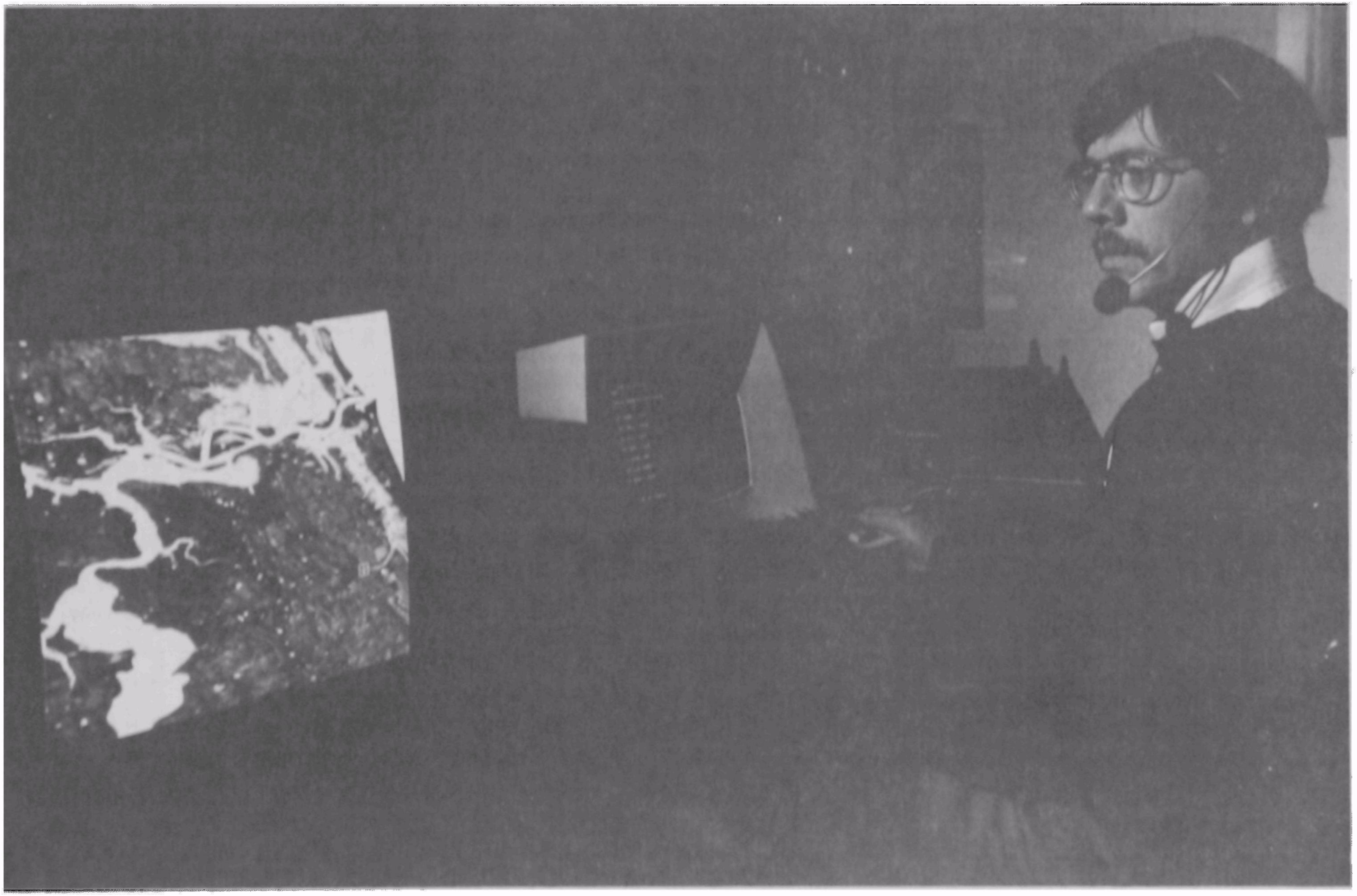


Fig. 15. Color Image CRT Display With Operator-Controlled Cursor.

4. TRANSFORMATIONS, ANALYSES, AND MODELING

4.1 SPATIAL TRANSFORMATIONS OF DATA STRUCTURES

The real power of geographic information systems is associated with the ability to perform spatial analyses and transformations on different data structures. Real-world problems are very complex and require the combination and integration of different geographic data bases to answer spatially oriented questions. This begins with the ability to transform a single point from one coordinate system to another and continues through a wide range of transformations such as interpolation, chain or polygon intersection, Thiessen polygon calculations, geometric warping, etc. Spatial transformations are required at every step of digitizing, editing, analyzing, and displaying geographic data. For example, digitizer coordinates from an x,y tablet must go through two types of transformations before these data can be stored in a geographic data base. These transformations require input of ground control information before digitizing can begin. As coordinates are captured, they must be scaled, translated, and rotated into the coordinate system corresponding to the specific map projection of the base map (e.g., polyconic, UTM, etc.). The second transformation converts from the map projection coordinates into latitude-longitude for storage on disk. The following paragraphs in this subsection will describe spatial transformations between different types of data structures. The transformations may involve either the cartographic data or the thematic information, or in some cases, both. Later subsections will discuss integration of data files and examples of different types of spatial models and analyses (e.g., terrain models, geologic models, etc.).

<u>Geometric rectification and warping</u> (i.e., rubber-sheeting) is a prerequisite for correcting and registering cartographic data to standard coordinate systems or to match other geographic data. The techniques may be fairly simple, involving rotation, scaling, stretching, and translation. In some cases, the data may contain nonlinear distortions so that either analytical functions have to be used to describe the nonlinearities, or the geographic region must be divided into sections each of which can be rectified locally and merged back together. An example of the latter case was our rectification of previously digitized drainage basin outlines for the United States. When a global rectification to latitude-longitude was made for the whole country, the individual basins were not located accurately when superimposed on 1:250,000 quadrangle maps. It was necessary to perform a local rectification, quadrangle by quadrangle, yet at the same time applying global constraints so that the drainage basin outlines merged properly at the border of every quadrangle sheet. Within each quadrangle, further localization was done by using only the nearest ground control points in an area to compute the appropriate affine transformations. An example of the first case (warping) would be the geometric correction of LANDSAT data or scanned photography based on known ground truth points. In this case, two types of transformations are required. The first is geometric and builds a polynomial fit or special projection from the LANDSAT coordinates (row-and-column numbers) to the output coordinate system (e.g., UTM, latitude-longitude, etc.). The second type of transformation is performed on the thematic data (radiometric intensities) to compute a

new intensity value at each output pixel, based on nearby neighbors in the original LANDSAT image. Different techniques are used for this calculation (e.g., nearest neighbor, bilinear interpolation, or cubic convolution). The intent is to move and stretch the data so that they geographically overlay the correct location on the earth's surface. However, there are situations where it is more efficient to distort a correct data base to fit some other mispositioned data. If the user wished to tabulate area statistics of LANDSAT data by county, it would be more efficient to distort the county boundaries to match the distorted LANDSAT data than to rectify the thousands of LANDSAT pixels to match the county outlines. Of course, to produce an accurate computer map the LANDSAT data should be corrected.

Some of the most powerful tools for manipulating geographic data involve interpolation and extrapolation procedures. These techniques not only allow data to be distributed over geographic surfaces but also allow data to be changed from one basis to another. A common approach inputs sampled data at specific points and interpolates the thematic information to cover an area or surface. Some techniques for computing population distributions are based upon interpolating from centroids of population centers (e.g., enumeration districts) to fill out a grid over the area in question. Figure 16 shows an example for a simple distribution. In this case only the total population for each district is known so the distribution within the district boundary must be estimated. This approximates the real-world situation since a distribution of people is not a continuous function. Another example would be the calculation of gridded elevations from terrain contours. Again, interpolation techniques use data points from nearby contours to estimate the elevation for the grid cells. In most interpolation procedures, the neighboring control points are weighted in some fashion (e.g., $1/r^2$) so that nearby points have a much heavier influence than those far away. If there are a large number of input control points and the output basis (e.g., grid cells) contains many elements then it is important that efficient processing techniques be used during the calculation; otherwise, large amounts of computer time may be used to search the input files for the nearest neighbors at each interpolated position. One localized technique for identifying nearest neighbors is through the use of Thiessen triangulations.[2] Given a set of input points, the computer can calculate linkages between these points as a series of triangles. Then perpendicular bisectors of these triangles will define influence polygons around each point (see Fig. 6, type 2). Once the structure is defined, any new point (e.g., the center of the grid cell) can be introduced to the network and its immediate neighbors calculated very quickly along with the polygon in which it lies. These entities can then be used to calculate an interpolated value at the new point.

Other types of spatial transformations involve smoothing of data over a geographical area, computing areas and perimeters of polygons along with their boundary direction, calculating links of chains or lineations, and performing proximity calculations. An example of data smoothing would be the averaging of gridded terrain elevations with nearby elevation points so as to round the tops of ridges and smooth out minor spikes on the surface. Another mechanism converts grid cell data into smoothly varying contours which can then be mapped with gray-level density patterns in changing from one contour to another. Area, perimeter, centroid, and direction calculations might be done to

ORNL-DWG 76-10824

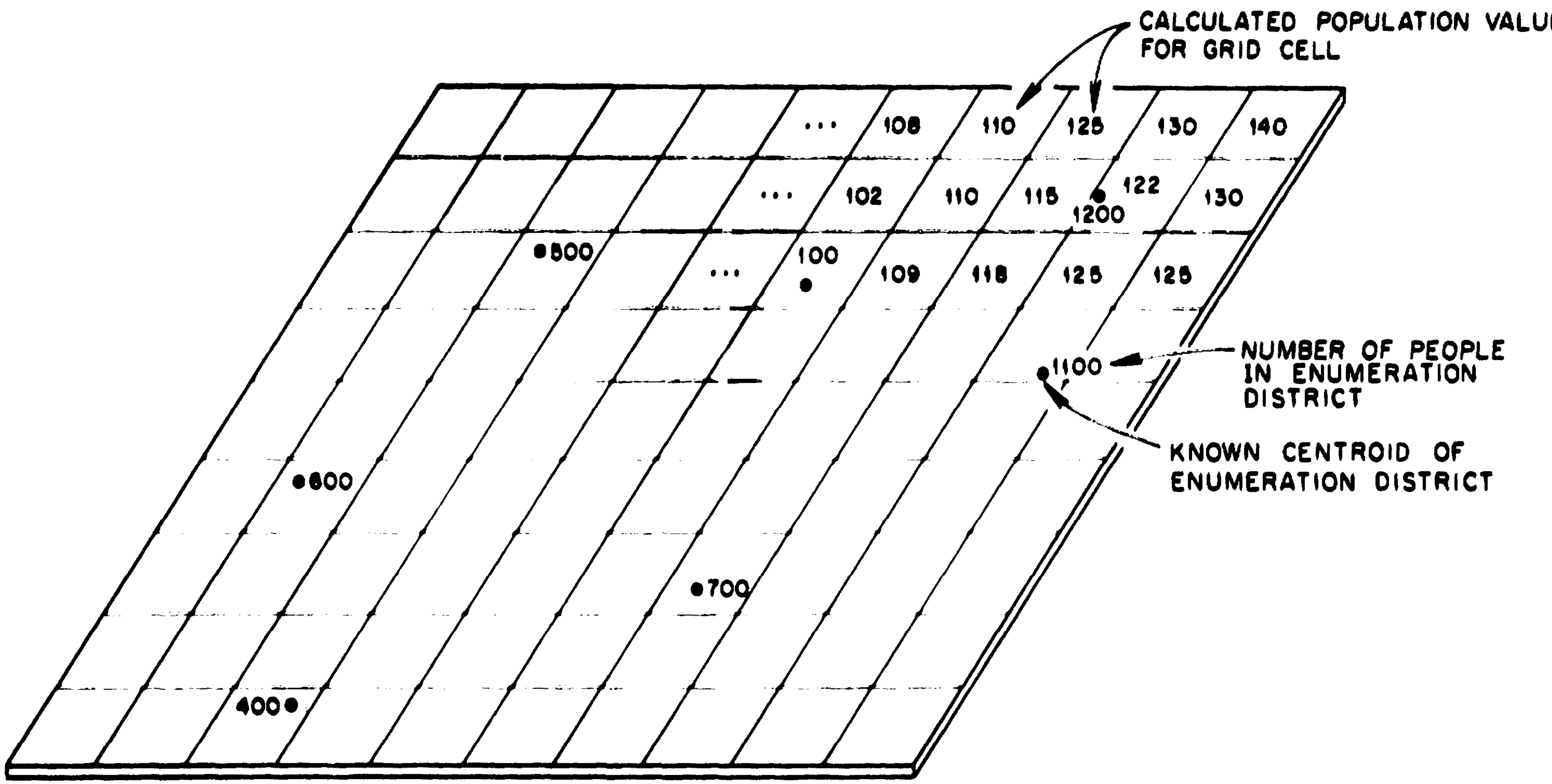


Fig. 16. Interpolation of Population Centroid Data To Estimate Counts for Each Grid Cell.

compute the acres of watershed basins, the length of streams, the center of urban areas, the direction of waterflow in a river, or the distance from a point emission source to the center of the nearest populated area. A typical proximity calculation would be determining the 10-mile buffer zone around the outer boundary of an urban area of 500 people/mile^2 or greater as shown in Fig. 17. This is a fairly sophisticated computation since the 500 people/mile^2 contour actually defines a polygonal border of arbitrary shape. Although it is easy to visualize a 10-mile ring parallel to this border, the calculation is quite complex when the computer has to determine all the points defining the outer zone. This is especially difficult when there are empty pockets or holes in the overall 10-mile zone. A simpler proximity calculation might locate factories emitting air pollutants above a certain level that are near lakes having acidity problems. In this case, the centroid of the lake might be compared with the point location of the factories. Another example would be the calculation of a transportation corridor (Fig. 18) by delineating all areas within two miles of the highway to define favorable zones for constructing a new facility.

A number of geographical applications require conversion among data structures or intersection of data structures. Figure 19 shows five different data structures used in our population distribution techniques[2]: centroid points, Thiessen polygons, county chains, grid cells, sectors, and annuli. Not shown are contour lineations computed to portray the density as the number of people/mile^2. Manipulating these structures simultaneously is necessary to carry out the spatial analyses required. Figure 20 shows conversions involving point data, gridded data, and polygonal data. Typical examples include transforming from points to grids, grids to polygons, polygons to grids, grids to grids, etc. To perform these calculations, algorithms must be used to calculate intersections or unions among the different data structures. Examples would include point-in-polygon, chain-to-chain intersections, polygon intersections,[3] dime vector-to-chain intersections, etc. The "sliver" problem is an important part of the polygon intersection problem that is difficult to solve. If the output data base contains hundreds of little slivers (that should really be removed or merged in with much larger neighbors), further processing and display is very cumbersome and may be unacceptable. Consideration must also be given to the thematic data in selecting the appropriate technique for transforming between data structures. For example, in combining gridded data and polygonal data, it might be possible to assume a grid cell was inside a particular polygon if its center point lay inside. However, if the grid cells were fairly large, it might be necessary to intersect their quadrilateral borders with the polygonal outline itself to determine how much of the grid cell lay inside each polygon. If the gridded data represented terrain slope and the desire were to calculate an average slope for soil polygons, the first approach might suffice. However, if the grid cell represented land cover categories in large cells, the second approach would be better to calculate the acreage of various land cover types in each soil polygon. Another example might be to calculate the length of rivers falling within contour bands that depict SO_2 air pollution over a large region. This would require intersection of the linear contours with the river chains. If the problem were rephrased to calculate the areas of lakes within each of the contour levels, then a polygon-chain intersection might be used. It becomes

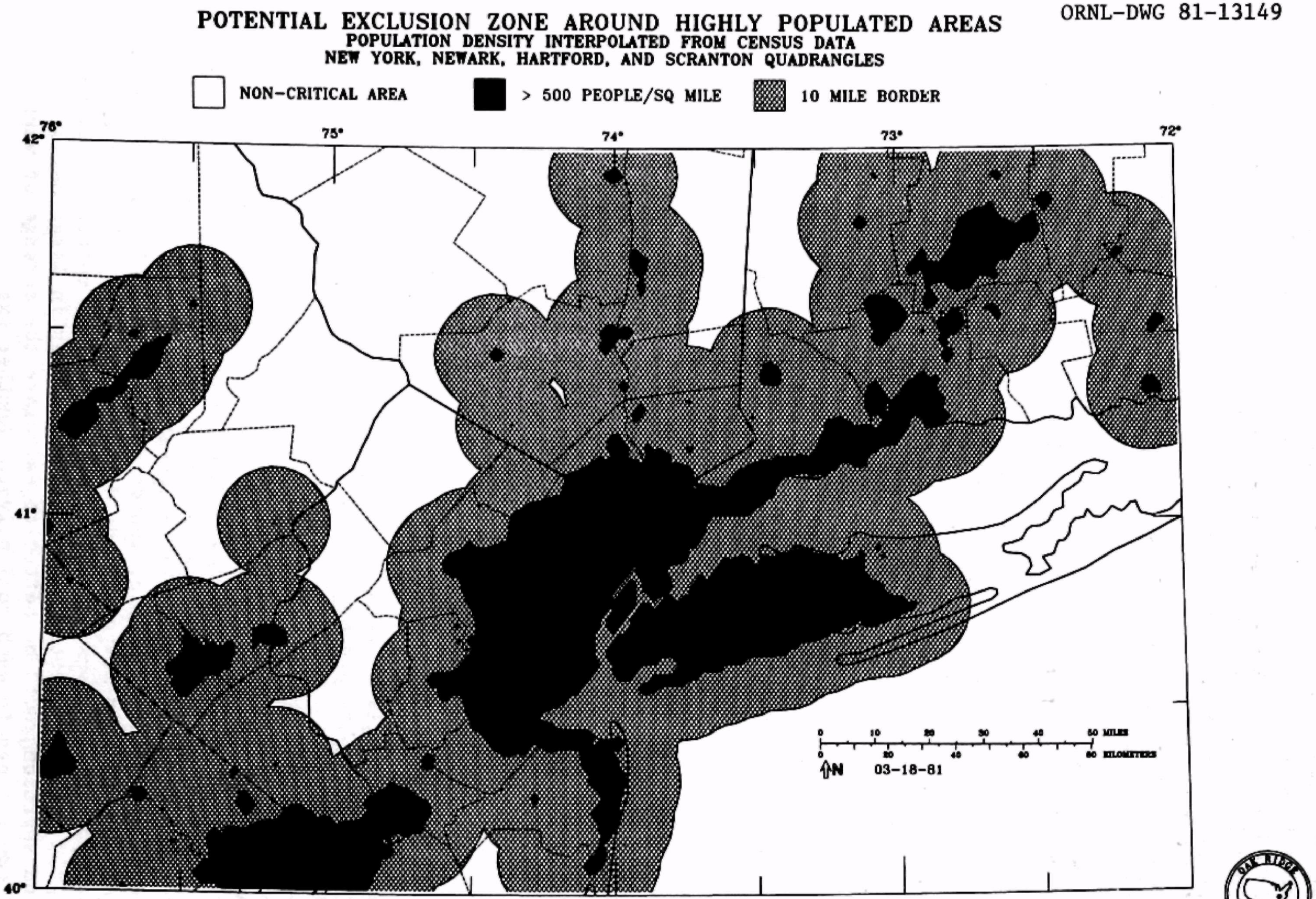


Fig. 17. Proximity Calculation of 10-Mile Zone Around Urban Areas.

ORNL-DWG 82-11063

HIGHWAY PROXIMITY MAP FOR TENNESSEE
ZONES WITHIN 2$^1/_2$ MILES OF A HIGHWAY

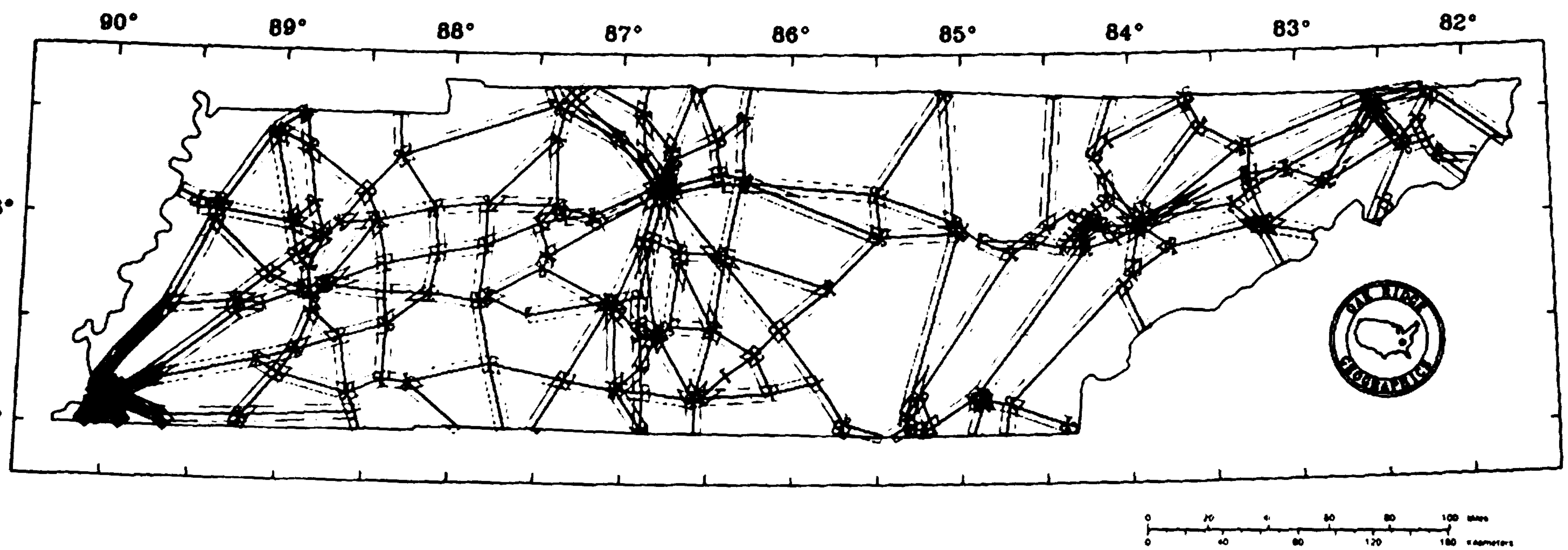


Fig. 18. Proximity Calculation of Transportation Corridors to Define Zones Within 2-1/2 Miles of Highway.

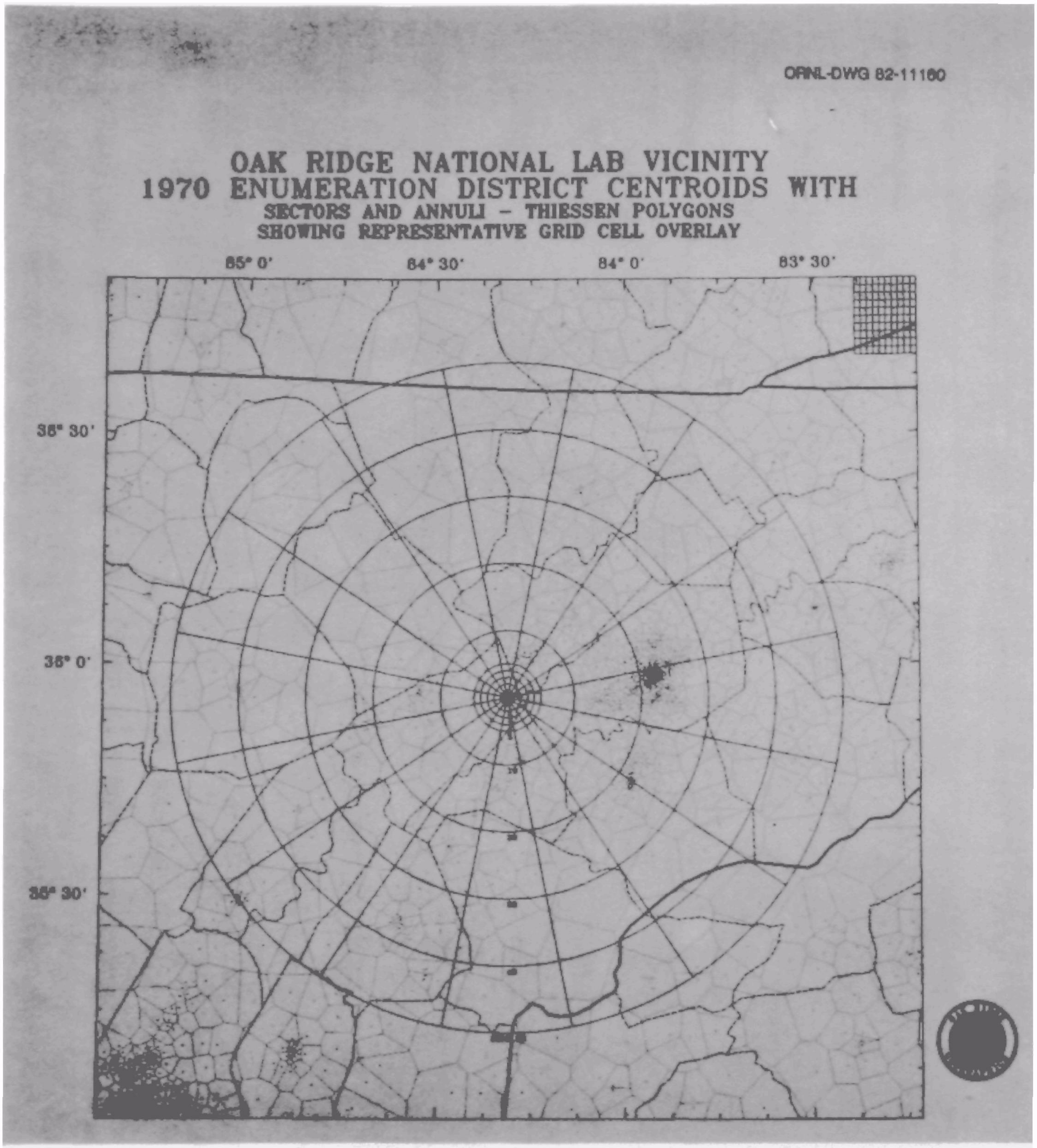


Fig. 19. Superposition of Various Data Structures Used in Geographic Transformations and Analyses.

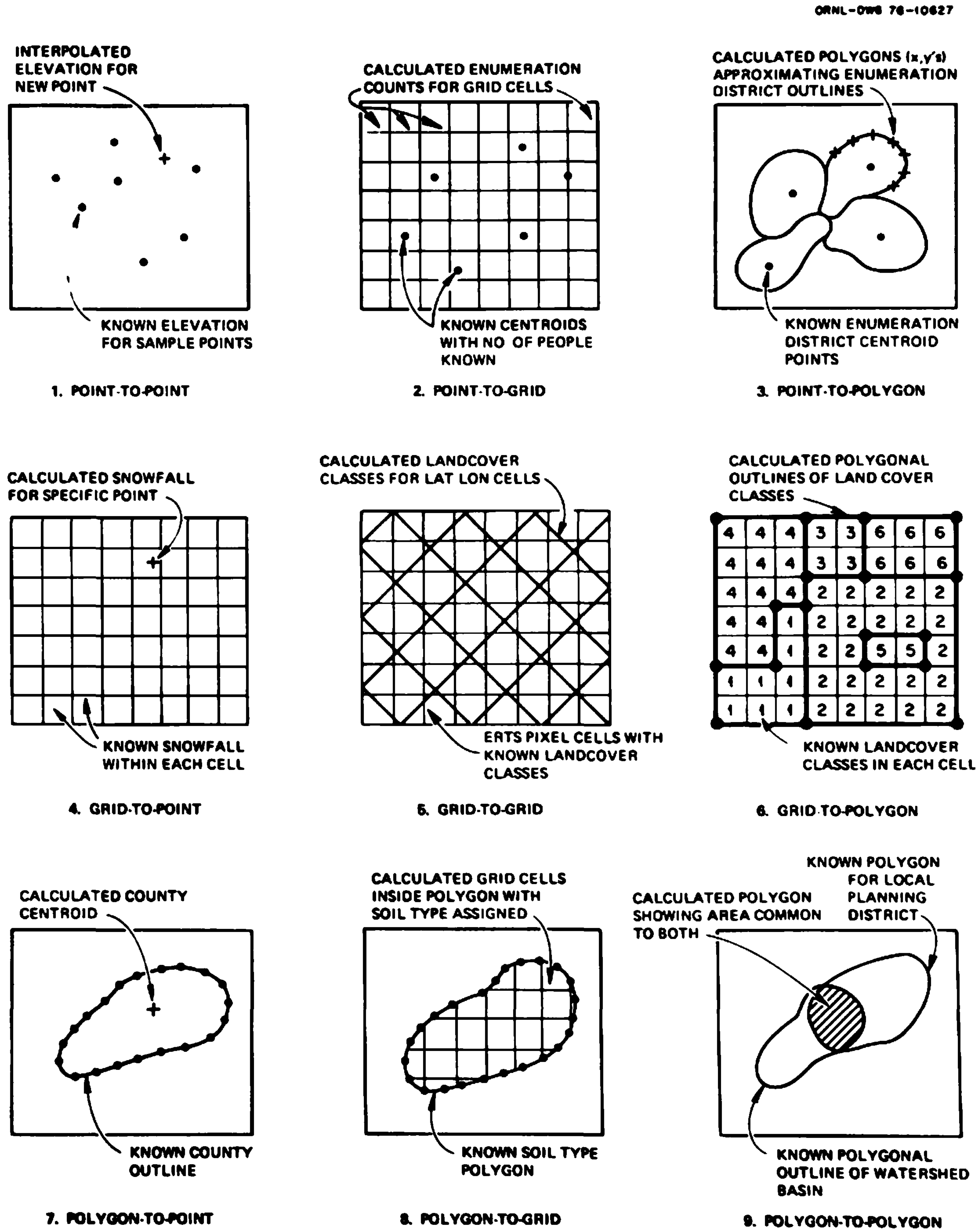


Fig. 20. ORNL Geographic Transformations Between Data Structures.

apparent that the choice of data structures and the questions to be answered determine the type of algorithms used.

Using special processing techniques, many of these spatial calculations can be performed using just the dime vector or chain representation of the geographic features rather than having to construct the actual polygons. Some of these techniques depend heavily on data sorting procedures and can increase the speed of calculations by a factor of 5. The processing algorithms have a large effect on the computing requirements, especially as the number (n) of geographic elements (e.g., vectors, polygons, etc.) increase. A system with requirements that vary as a function of n^2 or n^3 may be acceptable for test cases with a few elements but, in comparison to n or n log(n) relationships, would be totally unacceptable for real problems. With dime vector processing memory requirements also decrease significantly. All these types of spatial transformations and variations deal with two-dimensional data, primarily on the earth's surface. Some discussion of three-dimensional modeling is given in a later subsection.

4.2 DATA INTERPRETATION, INTEGRATION, AND INDEX CALCULATION.

The spatial transformations and algorithms discussed in the previous subsection make up an important part of the tools required to integrate and interpret different data bases to carry out users' applications. In addition to manipulating the cartographic data, consideration must be given to how the thematic variables are to be analyzed and combined. A wide variety of methods have been used including such functions as categorizing, ranking, index calculations, overlay analysis, decision matrices, empirical functions, statistical tabulations, regression analyses, correlation analyses, cluster analyses, etc. Later sections discuss more specific and topical types of modeling activities. The intent of this subsection is to show the use of some of these general methods in a few examples and indicate their relationship to geographic processing techniques.

Simpler methods use single thematic variables processed so that interpretations can be made from the data. Soil scientists frequently rank soil types according to erodibility and then convert the raw data into output maps where dark shadings may correspond to highly erodible soils and lighter shadings to less erodible soils. By categorizing and ranking the many different soil types into groups, it is possible to plot soil interpretation maps for different uses such as prime farm land, septic systems, commercial construction, etc. For multivariable analysis a commonly used technique is to compute suitability or sensitivity indices. This type of overlay analysis can be broken into two different processes: the first deals with the thematic values of the calculated variable, and a second deals with the spatial domain of the calculated variable. Figure 21 presents a very simplified overview of the combination of four types of data (contours, polygons, network, and points) to determine suitable areas for residential development. In this diagram the final result is shown as a simple "yes" or "no" answer for the suitability of each subarea.

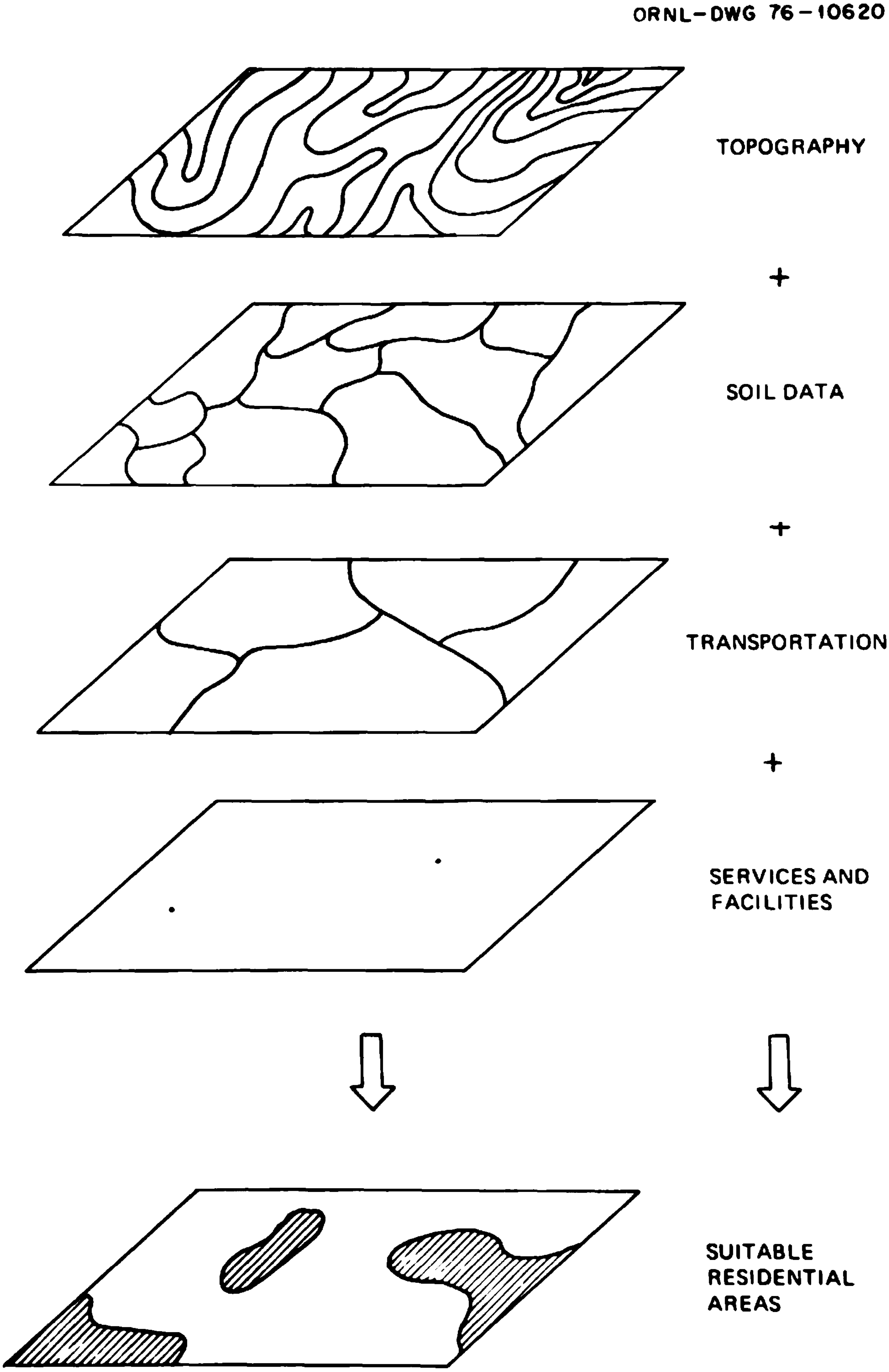


Fig. 21. Simplified Overview of Combining Four Variables to Compute Suitable Areas for Residential Development.

As another example, consider the problem of choosing a suitable site for locating a power plant based upon five different variables: seismic activity, water availability, urban demand areas, soil and geologic stability, and railroad accessibility. A site suitability index may be calculated as a linear weighted sum of the input data variables. After scaling and standardizing the input data, weights are chosen to reflect the relative importance of each variable on the final composite score to be calculated. After determining this thematic process the investigator must then tackle the problem of evaluating the function spatially for each geographic location or subarea in the study region. The complexity of this problem depends on whether a point system, grid system, segmental representation, or combination of structures is used. If a point or grid system is used, the software techniques are relatively easy, although the data processing may be prohibitive if high resolution is needed for large regions. In this case, most of the processing involves manipulating the thematic data for every cell in the grid; whereas in a polygonal system, the heavy processing is involved with intersecting the cartographic data. By using special techniques with the cartographic data in dime vector (segmental) form,[4] the processing can be made much more efficient than the traditional method of intersecting independent polygons. If a polygon overlay is performed, the "least-common-denominator" polygons define the areas for which the suitability equation is calculated using the selected weights. The polygon intersection is referred to as spatial processing, and the suitability equation represents thematic processing. If several data structures are involved, it may be necessary to determine criteria for converting the data into compatible structures. For example, the seismic data may correspond to epicenter points of known magnitude. The criteria may state that exclusion zones must be formed as circles whose radii are proportional to the magnitude of the earthquake. Railroad accessability may be defined in terms of distance to the nearest railroad. These calculations are all a part of the spatial processing.

If the input data variables are relatively homogeneous over large areas so that only a few polygons are needed to represent the data, it may be more cost-effective to use segmental techniques. However, if the input variables change quite rapidly from one location to another so that a large number of small polygons are required, it may be more efficient to use the grid cell approach. Some variables such as precipitation, wind patterns, magnetic intensity, and temperature vary continuously over the earth's surface. If polygonal techniques were used in calculating an index representing sensitivity of U.S. regions to acid rain, it would be necessary to precategorize the continuous variables for polygonal representation (since all the area inside a given polygon is assumed to have the same numerical value). If small grid cells are used to give good cartographic representation, they can also represent thematic variations of continuous variables more readily. It is difficult to compare costs on a general basis between grid cell processing and polygon intersection techniques because actual geographic variables, resolution requirements, and specific analysis needs affect the parameters used in estimating costs. Because grid techniques were easier to program in the early days of geographic information systems, they were the most commonly used structures by modelers and planners. However, cartographers and those interested in sophisticated map output used segmental systems to more accurately portray the data.

Some analyses can be defined with very specific spatial criteria rather than applying subjective weights to groups of variables. In assessing the environmental impacts of strip mining[5] (see Fig. 22), one procedure identified potential areas needing reclamation as those surface mines on slopes greater than 15% within 200 ft of a stream. Another application identified exclusion zones for future nuclear power plant sites based upon population density criteria.[2] Figure 23 presents population density contours for the Northeast. One criterion stated that any candidate site must have (1) a density below 250 people/mile^2 within 2 miles of the site, (2) a density below 750 people/mile^2 within 30 miles of the site, and (3) not more than 2250 people/mile^2 on one side of the site. Otherwise the site would be unacceptable. Figure 24 shows all possible exclusion zones based on these criteria. Some investigators may not want to specify hard-and-fast criteria for automated analysis, but would rather have their data variables combined visually so they can make manual interpretations. Figure 25 shows LANDSAT data around the Sequoyah nuclear plant in Tennessee, superimposed with population distributions and highway networks to aid in evacuation planning in the event of an emergency.

Another commonly used technique for calculating a composite interpretation from several variables involves the use of a decision matrix rather than a mathematical function. A simple example would have each variable in the study categorized and arranged along an axis of the matrix. The entries in the matrix are the calculated suitability scores corresponding to each set of data values for the input variables. The calculation of a suitability score would be a simple table lookup in which the data values for each variable specify row and column numbers in the decision matrix. This method is useful when complex relationships may have to be determined empirically or through subjective judgments rather than through analytical functions.

A number of statistical techniques are useful in studying geographical relationships among both the thematic variables and the spatial patterns. Examples of the tools include descriptive statistics, probability theory and distributions, tests of significance, analysis of variance, correlation analysis, regression analysis, discriminant analysis, factor analysis, cluster analysis, etc. Most of these techniques are oriented around similar observation units and thus have been used with grid systems or other types of compartmental structures. This brief overview is not intended to explain or even introduce the techniques of statistical analysis. However, a few brief examples can indicate the variety of geographical problems that can be analyzed through statistical techniques. Correlation analyses can be used to determine positive or negative relationships between different variables over a geographical region. When two or more variables are measuring similar characteristics, there may be an overemphasis on one aspect of a study. For example, in looking at the acid rain deposition problem, a variety of soil parameters may be used as individual variables, whereas only a single variable dealing with weather and deposition patterns may be included. This overemphasis on soil may tend to create spatial patterns which, when mapped, reflect the different soil types more than possible sensitivity to acid rain. Principal components analysis can be used to transform the input variables into a set of orthogonal components which can be treated

ORNL-DWG 77-10486

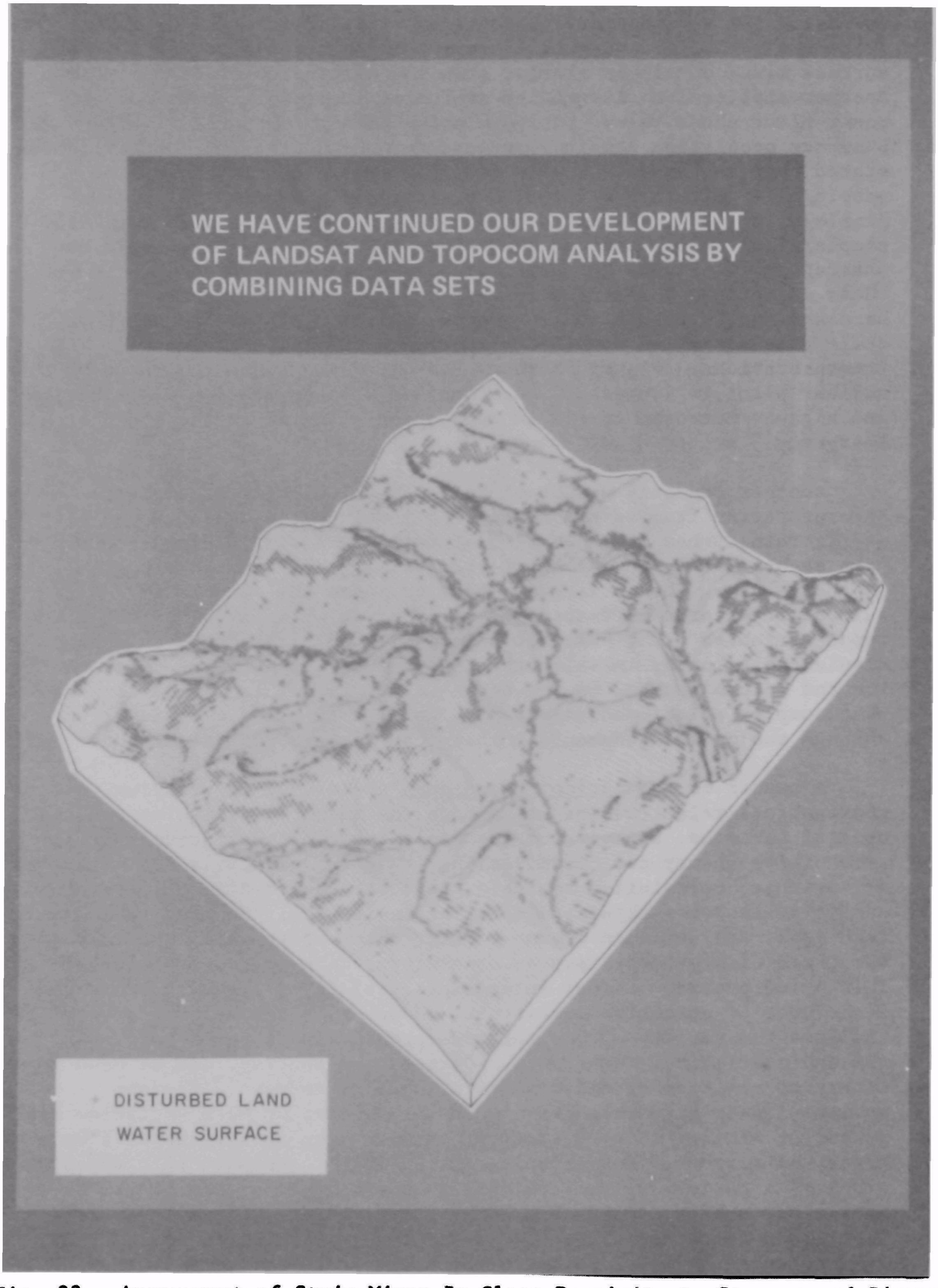


Fig. 22. Assessment of Strip Mines In Close Proximity to Streams and Rivers.

ORNL-DWG 83-7742

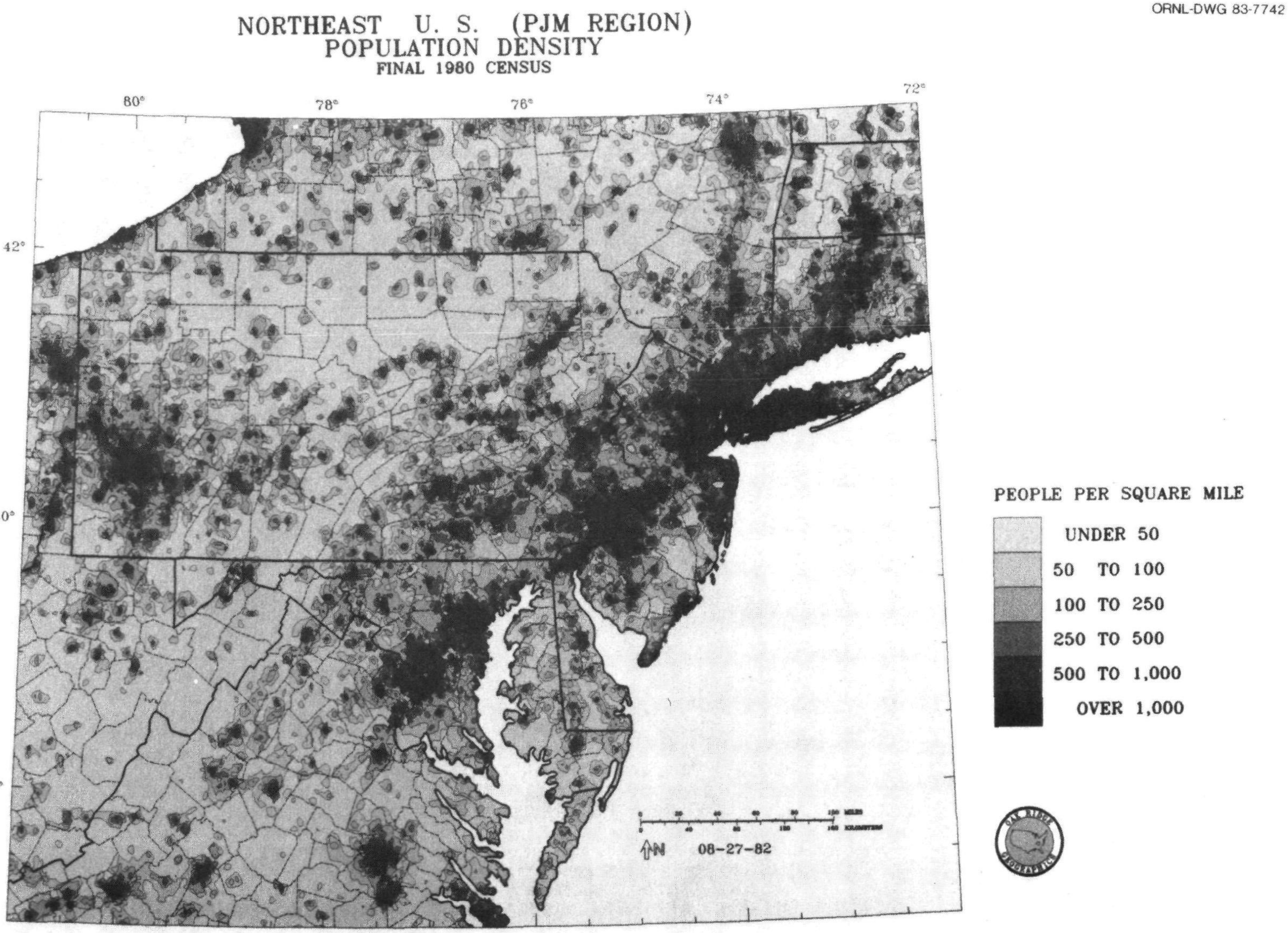


Fig. 23. Population Density Contours For the Northeast.

ORNL-DWG 81-18368

Fig. 25. Superposition of LANDSAT Data, Population Contours, Highway and Railroad Networks Around the Sequoyah Nuclear Plant.

as independent variables measuring about the same overall information content. These could then be used for further statistical analysis. Factor analysis is a further extension through which input variables can be grouped into meaningful factors that can be interpreted. An example might be the construction of composite land use indices which may be used to describe the basic spatial characteristics (e.g., residential, accessibility, agricultural, etc.) of a geographic region. Cluster analysis has been used to develop homogeneous subregions through a clustering of cells or parcels that cover an entire study area. The grouping of parcels is based upon their similar and dissimilar characeteristics. Clustering techniques have been used with remotely sensed data to calculate clusters of land cover patterns. These types of techniques have traditionally been used with gridded or raster data, although some statistical software algorithms are not suited for handling the millions of observation units (cells) needed in representing geographic data. In social or economic analyses, the number of observation units are typically less than several thousand.

There are many different approaches to analyzing geographic data, limited only by the imagination of the investigator and the ability of the computer scientist to implement these ideas. Subsection 4.1 presented a variety of basic transformation tools that are available to the analyst. This subsection presented ways in which these tools might be integrated to solve general types of geographic applications.

4.3 TOPOGRAPHIC ANALYSES, LANDFORMS AND VISIBILITY CALCULATIONS

The previous two subsections discussed general spatial transformations and analyses performed on all types of geographic data. This subsection and the ones following will present specific types of analyses and models that are associated with particular geographic data. Terrain modeling is utilized in a number of environmental applications where parameters associated with the earth's surface are important. Typical landform parameters include topographic elevations, contours, slope, aspect, watershed drainage patterns, inflow and outflow points, etc. Many computerized models use a grid system to represent the terrain, with elevations determined at intersections of the grid lines. Automated techniques may use stereo plotters to capture data directly from aerial photographs. Source data may consist of contour lineations which can be interpolated onto a gridded basis. Normally a rectangular grid system is used, although, in some cases, a network of triangular cells has been employed. During the interpolation, an analytical function may be calculated for each grid point. Typical functions include planar approximations, quadratic fits, or cubic polynominals. By evaluating the gradient of the function it is possible to determine the slope at each grid point. Figure 26 shows a simplified diagram for depicting slope from gridded elevations and contour lines. The computation of aspect provides an orientation of the hillside with respect to north. For example, an aspect of 90° has an eastern facing slope, 135° is southeasterly facing, etc. Slope information is important in determining runoff and erodibility, while aspect is important for determining sun angles and exposure. Dramatic differences may be observed in vegetation patterns depending upon whether the terrain has a north-facing or south-facing slope. Figures 27-29 present the terrain contours, slope, and aspect, respectively, in the East Tennessee area around the Oak Ridge Department of Energy reservation.

ORNL-DWG 76-10621

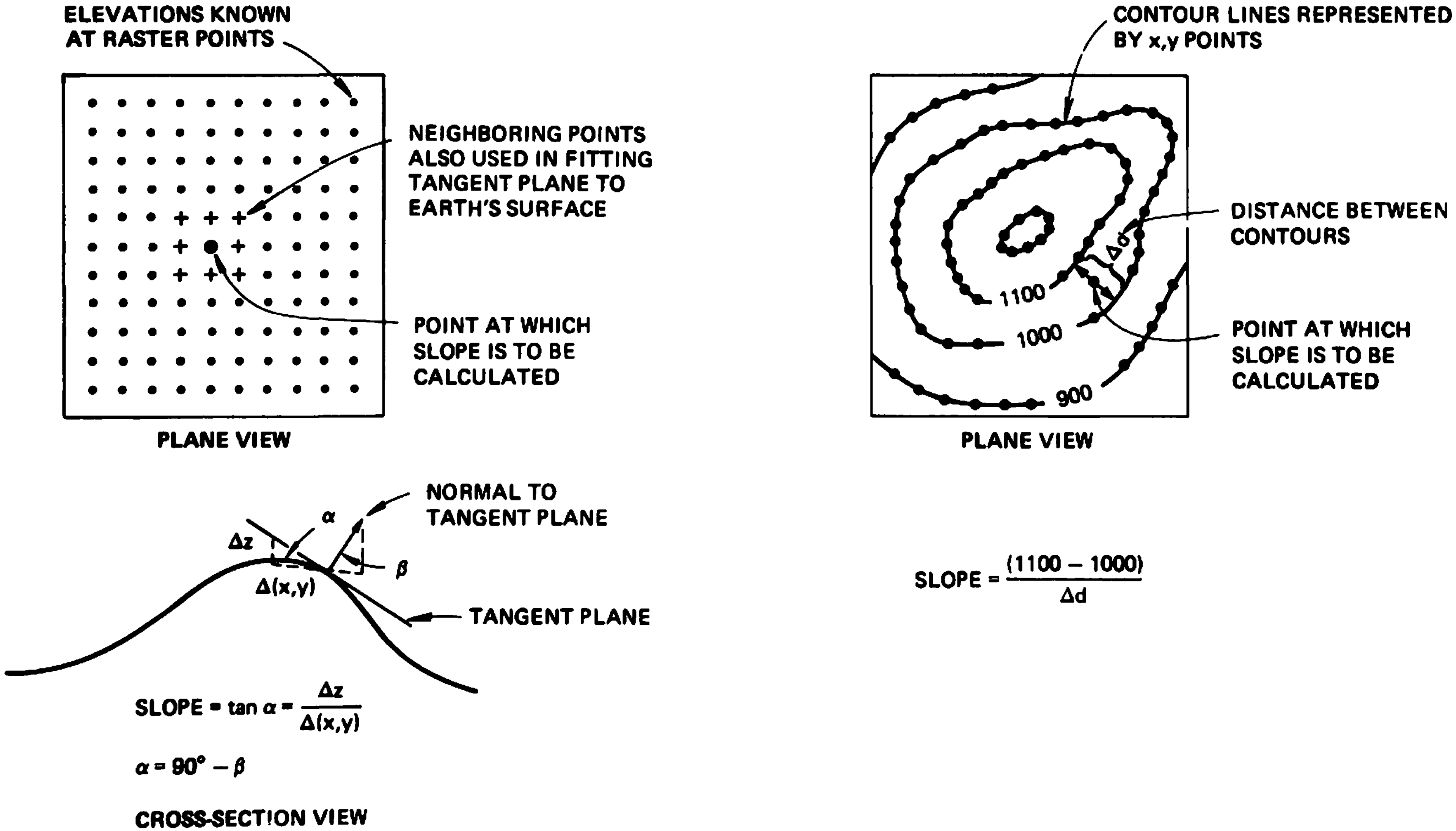


Fig. 26. Simplified Diagram Depicting Slope From Raster Data and Contour Lines.

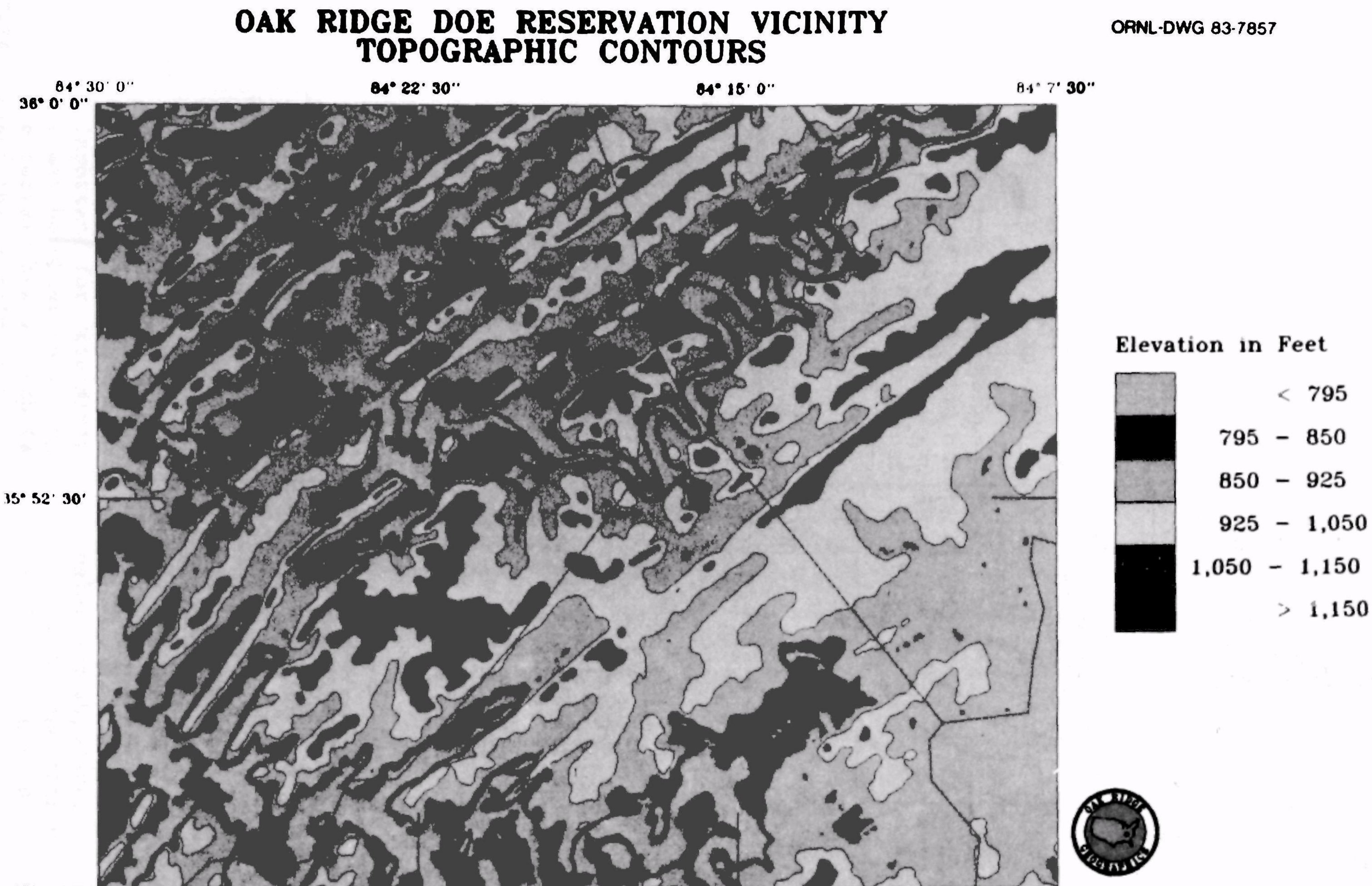


Fig. 27. Terrain Contours For Oak Ridge DOE Reservation Vicinity.

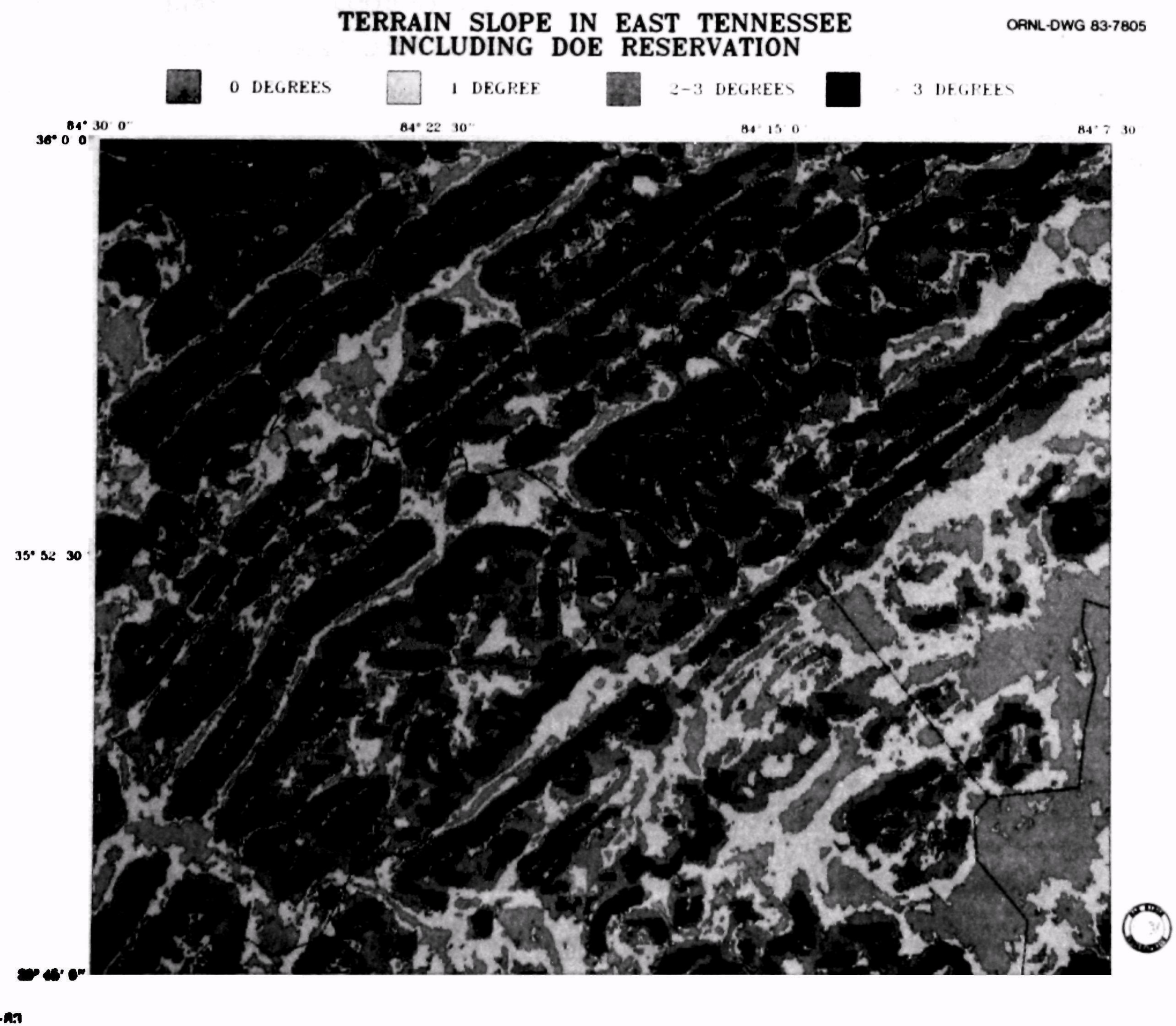


Fig. 28. Terrain Slope For Oak Ridge DOE Reservation Vicinity.

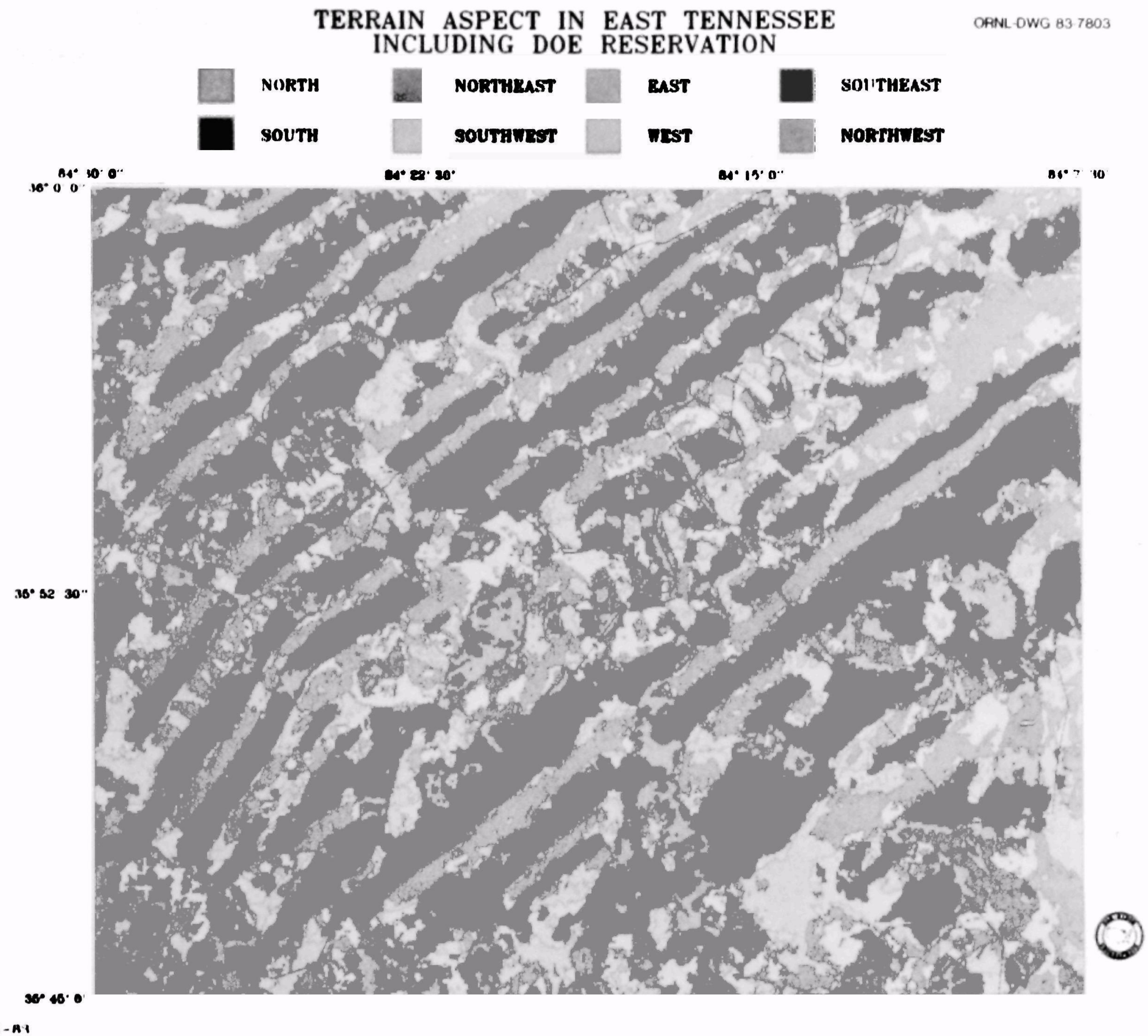


Fig. 29. Terrain Aspect for Oak Ridge DOE Reservation Vicinity.

In addition to these point calculated parameters, there are other land-forms which can be determined on an areal basis. For example, R. G. Edwards has worked on computational techniques to calculate the boundaries of watershed drainage basins by determining the direction of simulated water flow as it strikes each grid point. The computer simulation of a hydrologist's expertise in aggregating smaller watersheds into larger basins is a complex problem. Quantifying the subjective judgments that come from years of experience, as well as handling the geometric problems, is difficult. It is also useful to compute the inflow and outflow points within and among watersheds. The simulation of stream patterns can be computed within the watershed and compared with base maps or aerial photography. The identification of flood plain areas may also result from this type of calculation. Since high-resolution grid systems are normally used to represent terrain models accurately, these techniques may be more appropriate for small area studies requiring less processing. Aggregation techniques can be used for generalizing the data on a regional basis, although small variations in the earth's surface will not be represented.

Topographic analyses are used in many different types of impact and assessment problems associated with land use and environmental planning. Although topography includes structural features on the landscape, the more common usage emphasizes the terrain or relief. The following list gives examples of such problems:

1. determining the effect of topography on wind rose and air movement to model air pollution from power plants, factories, etc.;

2. determining slope suitability of an area for building construction, strip mining operations, agricultural purposes, etc.;

3. siting of transmission corridors, cooling towers, highways, rail lines, etc., with respect to visibility to surrounding areas, cost of cut-and-fill operations, etc.;

4. determining optimal location of office buildings, residential homes, and solar facilities with respect to aspect, sun angle, and sun exposure for visual and heating purposes;

5. determining land area to be covered by water from proposed dams and calculating total reservoir capacity;

6. determining water runoff during storms that may affect stream quality and cause flooding, especially if land use patterns and activities such as strip mining, forest clearing, and concrete surfaces are close by;

7. studying the patterns of vegetation growth, soil types, and geologic characteristics where spatial distribution on the landscape is a function of topographical properties; and

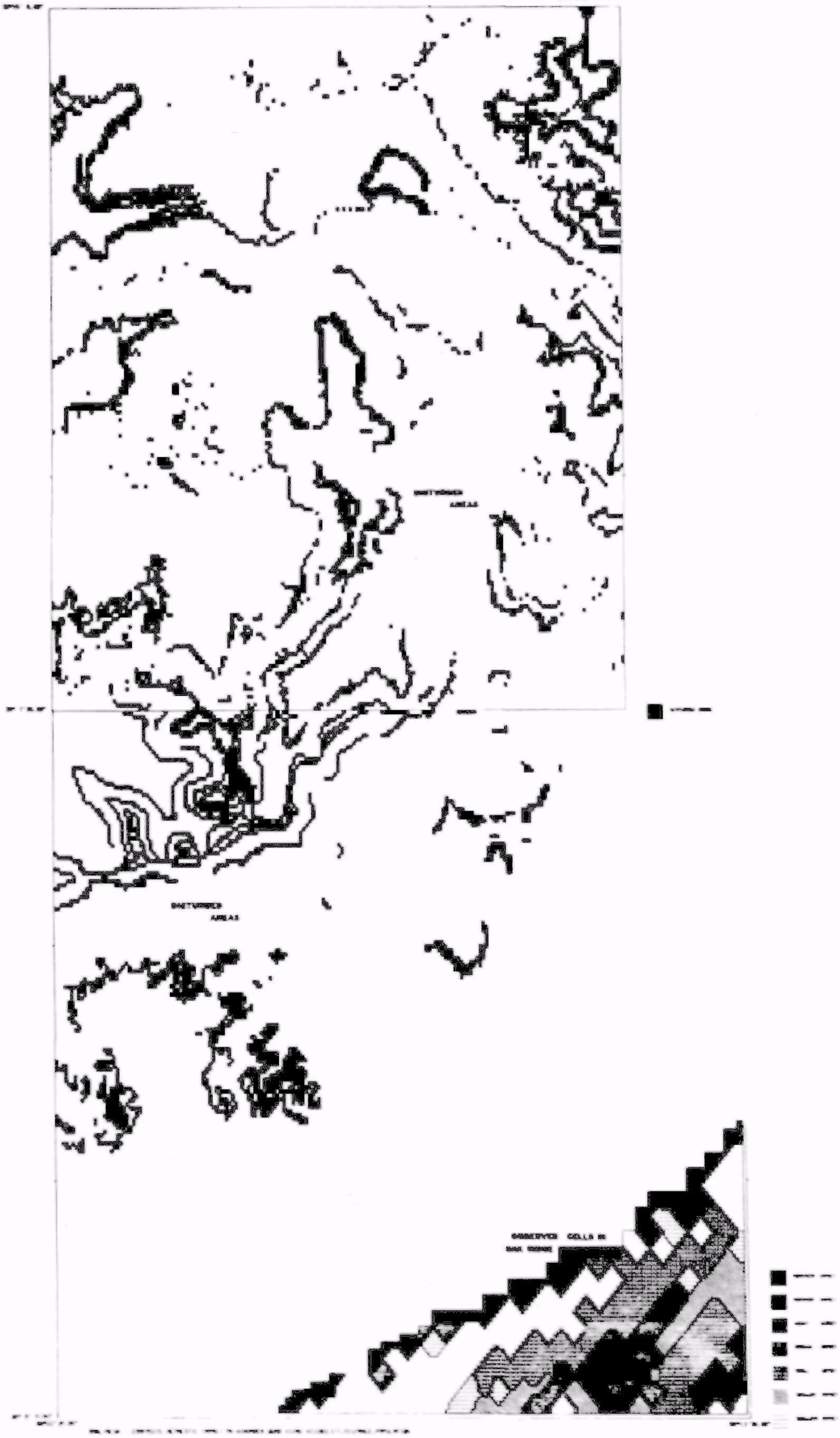
8. enhancing and correcting remotely sensed radiometric data from aircraft and satellites as a function of terrain shading to improve the classification of land cover and surface features.

One special application involving the use of terrain models is that of determing visual impacts of surrounding land features.[5] For example, one aspect of strip mining operations and reclamation practices is the visual effect on nearby population centers and transportation corridors, (e.g., interstates and major highways). Modeling these effects requires line-of-site calculations from an observer position to the affected area. In some cases, the reverse can be computed easier by determining all areas surrounding a given point which are visible from that point. The distance from the observer to the affected area, the percentage of the viewing scene impacted, and the viewing population can all be incorporated to determine an impact score. Figure 30 shows strip mines north of Oak Ridge, Tennessee, with shaded areas in the southeast corner that represent the composite impact score on the population as a function of the parameters just mentioned. To determine which strip mines are most visible, thus needing initial reclamation, the complement calculation is performed, as shown in Fig. 31 for the northernmost strip mines. The darkest shades are most visible. By assessing such visibility impact scores over a region, it is possible to determine minimum-impact corridors for siting new construction such as highways or transmission lines.

4.4 IMAGE PROCESSING AND REMOTE SENSING

With the advent of digital LANDSAT data in the early 1970s, a large number of groups around the country (both private and government) accelerated their development of computer-assisted remote-sensing analysis techniques. As a result, special hardware systems have been available for a number of years to process and display raster-oriented data from satellite imagery and aerial photographs. This subsection will discuss a few of the image processing techniques and applications associated with these types of systems. A color-image processing system (manufactured by International Imaging Systems, Milpitas, California) with capabilities for inputting and analyzing satellite data, base maps, micrographs, photographs, and other spatial data (e.g., population density, topography, etc.) has been in operation at ORNL for several years. Capabilities for producing video tapes, viewgraphs, glossies, slides, and large color plots are part of the system. Figure 32 shows a portion of the hardware. Many different analysis and display functions are available to the user, but this subsection will discuss just a few. A high-resolution color system has been installed in the Geographic Data Systems Section to handle four times the amount of data and resolution as compared to the earlier system. This newer system, initially manufactured by Ikonas Graphics Systems, Raleigh, N. C., contains high-speed hardware transformations to aid in processing vector and raster data simultaneously.

Satellite data used for computational processing (rather than manual interpretation) are normally acquired on magnetic tape. Aerial photography is normally digitized through a scanning process, using color filters if the source image is in color. The pixel data normally consist of integers ranging from 0 to 255 representing the light intensities measured during collection of the original imagery. These measurements may correspond to different wave lengths of reflected sunlight, referred to as bands. Multispectral LANDSAT data normally contain four bands with a pixel size of approximately an acre. The newer Thematic Mapper sensor collects data for seven bands with pixels of approximately one-fourth acre resolution. The French Spot satellite



COMPOSITE AESTHETIC IMPACT ON OBSERVER CELLS

Fig. 30. Composite Aesthetic Impact on Oak Ridge Observer Cells From Strip Mines North of Oak Ridge.

ORNL-DWG 75-13556

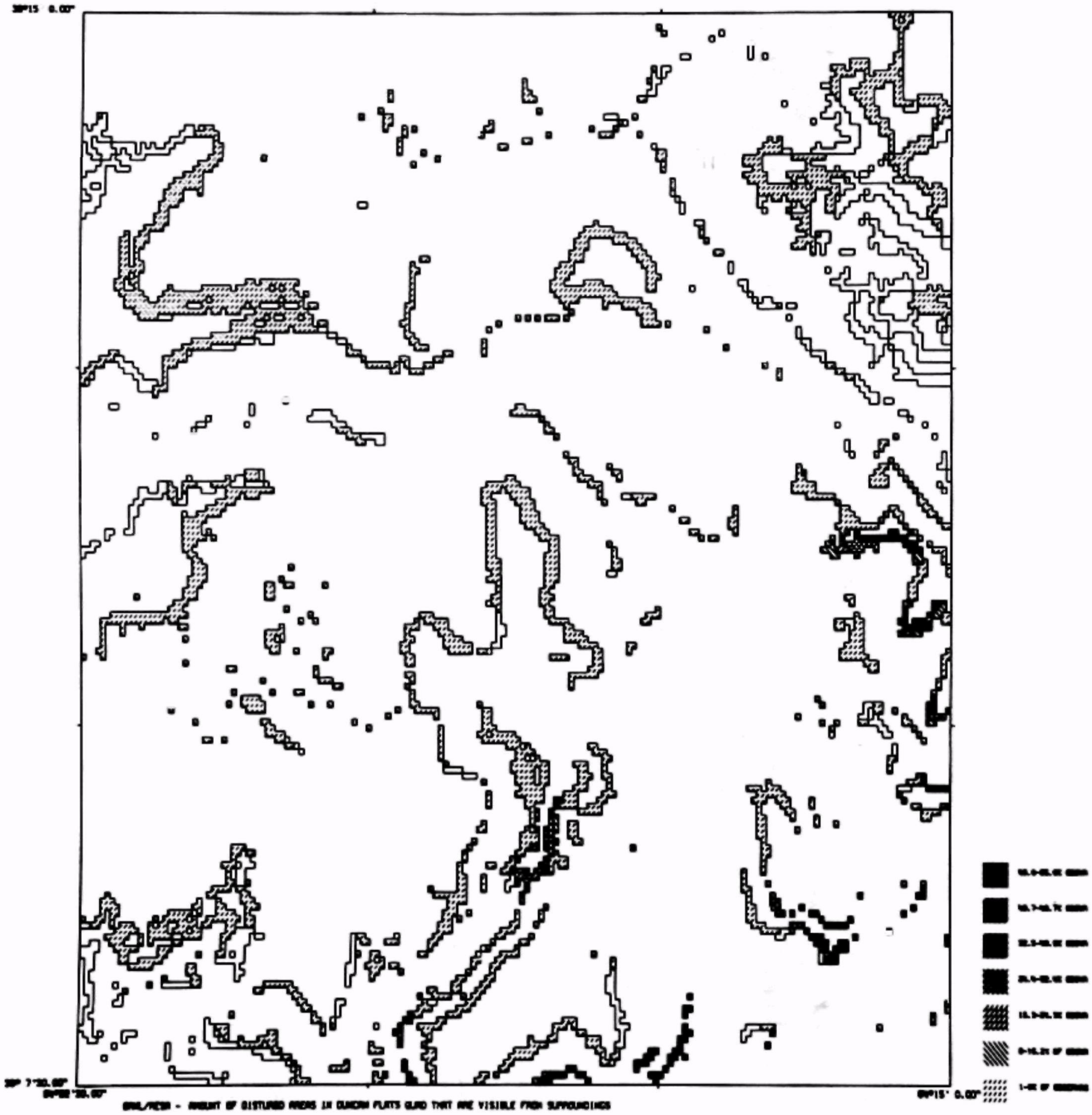


Fig. 31. Visibility Index of Disturbed Strip Mine Cells.

Fig. 32. ORNL Color Image Processing System with Computer or Video Input and Output, and Hardcopy Display.

can provide one-band data at a 10-meter resolution or three-band data at a 20-meter resolution. The intensities represent radiometric information, whereas the spatial location of the pixels represent the geometric attribute of these data. Image enhancement techniques are generally designed to transform the radiometric intensities into new values which, when displayed in color, provide more discernible information for certain features. The radiometric calculations may use statistical techniques such as cluster analysis and maximum likelihood discriminate analysis to aid in analyzing features. Geometric processing normally uses techniques of interpolation and coordinate transformations as discussed previously.

Once data are in a raster format, a variety of techniques can be used to extract information for the analyst viewing his/her results on a color CRT. Typical examples would include geometric rectification of the data, enhancement of the images to visually highlight unique features (Fig. 33), classification of data into land cover categories, edge-detection techniques to depict sharp structural changes in data, visual combination of multiple images (e.g., superimposition of other geographic features on the land cover), frequency distributions and radiometric profiles to study spectral responses across the image, pattern recognition, etc. This work is done interactively, with the analyst using a track ball to position a CRT cursor on the screen to create input information. Land cover classification may be performed using supervised or unsupervised techniques. In the later case, algorithms such as cluster analysis are used with predetermined criteria to create characteristic groups of pixels representing different land cover classes. In supervised classification polygons are drawn around specific land cover areas to create training samples as input to discriminate analysis routines for grouping the pixels into specific land cover categories. Figure 34 shows raw LANDSAT data for Minneapolis-St. Paul, and Fig. 35 shows a classification into four broad land cover categories using preselected training samples.

There are numerous applications for which image processing of remotely sensed data is useful (e.g., agricultural studies, weather patterns, demographic studies, facility siting, ecological modeling, strip mine analysis, environmental impact assessment, resource inventories, geological exploration, etc). One of the problems currently being studied at ORNL deals with acid rain impacts on water bodies, vegetation, urban structures, etc. A potential use of remotely sensed data in the acid rain problem would be to identify and classify certain features on the earth's surface that might indicate affected areas. Examples could include affected tree growth patterns, spectral differences in water bodies due to excessive acidity that has affected normal algae or plant growth, etc.

Since imagery is available several times a season, change detection studies can be performed to determine variations in landcover features over time. One study at ORNL is assessing the change in aquatic habitats at different time periods along portions of the East Coast of the United States. The integration of other types and sources of geographic data greatly enhances the use of remotely sensed imagery. For example, in the LANDSAT aquatic study, identification of tidal marshes, sand dunes, vegetative patterns, urban areas, etc. have been improved by incorporating digital data bases representing land use interpreted from aerial photography. In many geographic applications

ORNL PHOTO 3772-81

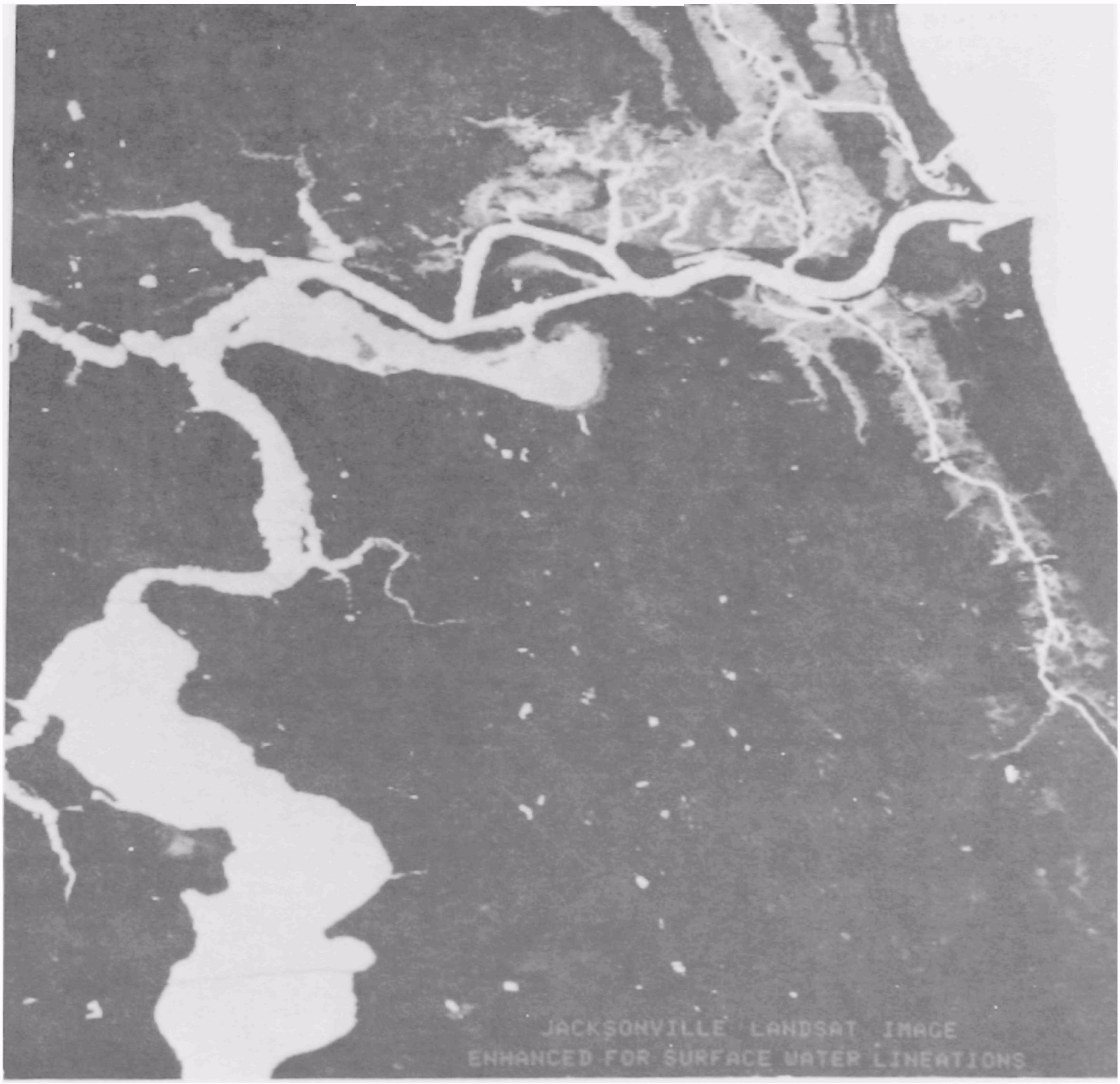


Fig. 33. Enhancement of LANSAT Imagery Around Jacksonville, Florida, to Depict Surface Water Including Streams in Swamp Areas.

ORNL-DWG-81-13289

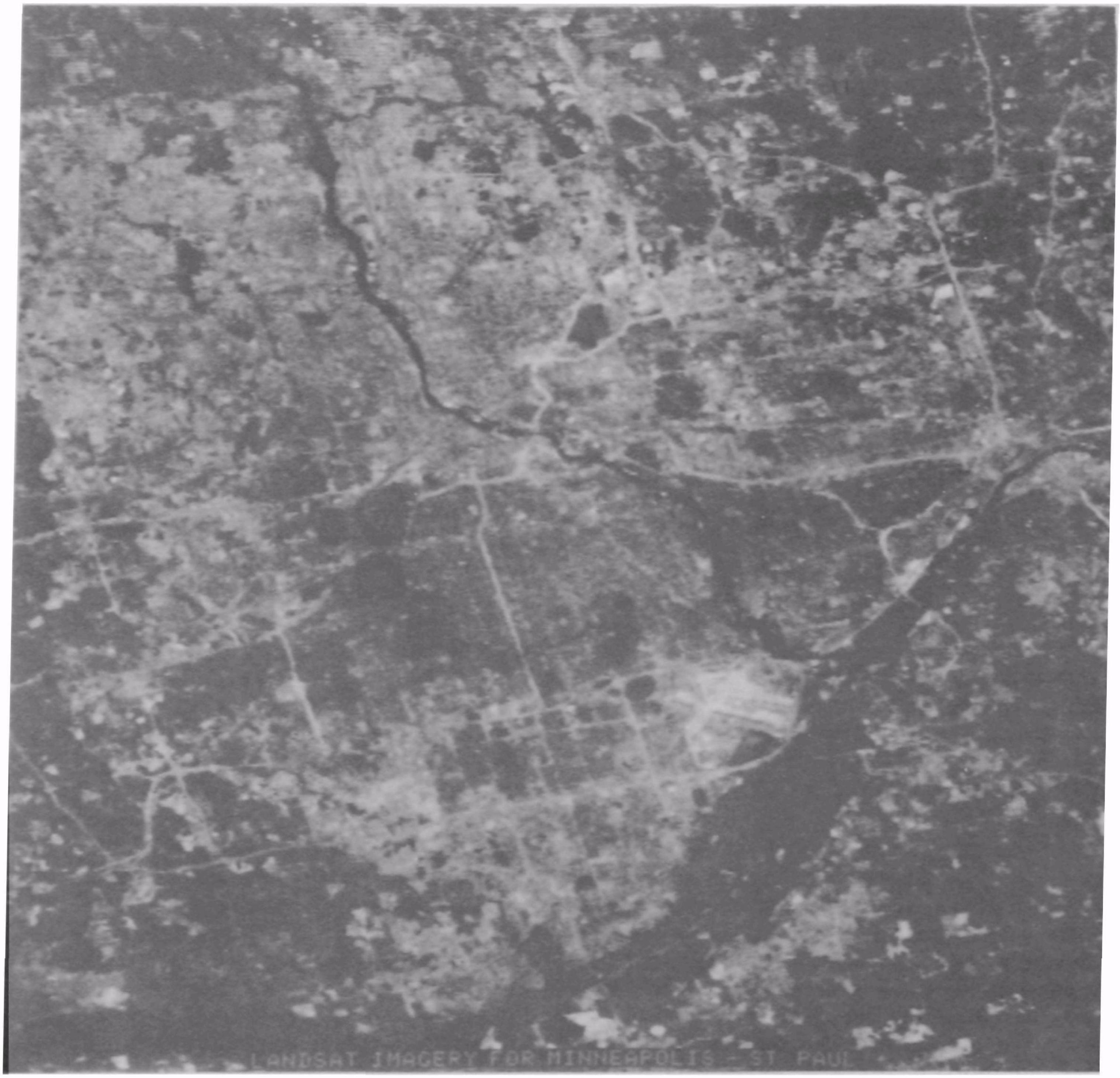


Fig. 34. Raw LANDSAT Imagery for Minneapolis-St. Paul.

ORNL-DWG 81-13279

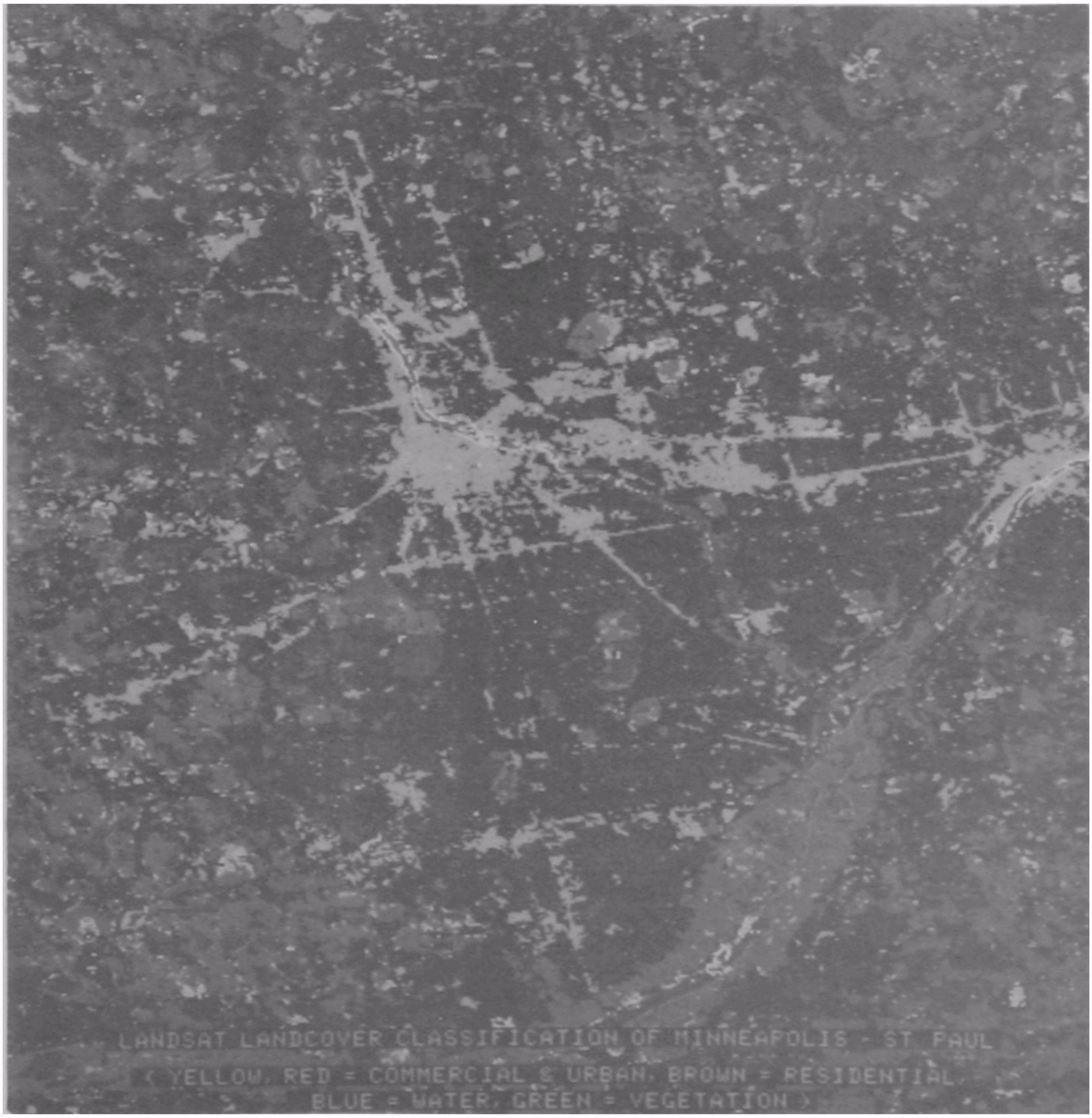


Fig. 35. Classified Landcover Patterns for Minneapolis-St. Paul.

the LANDSAT interpretations represent only one source of input for studying complex spatial problems. Thus, the capability to geometrically transform and integrate multiple data types (e.g. polygons, rasters, networks, etc.) is very important.

4.ɔ WATER RESOURCE ANALYSES

An important topical area for many environmental studies is the supply, use, and impact on water resources around the United States. In estimating the availability of water, models have been developed to predict flow (e.g., cubic feet-per-second) at gauging stations around the United States. Statistical models use historical flow data collected over a period of many years to predict water flow under different conditions. Such a model was developed at ORNL by Jeff Jalbert and Alf Shepherd.[6] The results from these models were processed geographically to interpolate along major rivers or streams in computing flow estimates between gauging stations. This is an important part of assessment studies dealing with water withdrawal or pollutant discharges into rivers. These tools are critical in estimating relative pollutant impacts on streams which have low dilution or buffering capacities during drought seasons. Combining water flow with power plant consumption for cooling purposes can also provide an estimate of seasonal impacts. Figure 36 shows those counties whose consumption for power plants in the Ohio River Basin would exceed 5% of the low flow during a drought period based upon a high demand for electricity in the year 2020. This work was done by Alf Shepherd,[7] R. B. Honea, and J. E. Dobson at ORNL.

To perform these types of computations and display the results in map form, it is very helpful to have hydrologic data bases delineating rivers, streams, reservoirs, lakes, drainage basins, etc. In the last few years a few data bases providing these types of data have become available from the United States Geological Survey[8] and the Environmental Protection Agency. Figure 37 shows Water Resource Council drainage basins for the United States, and Fig. 38a shows rivers and streams in the eastern United States. A cooperative effort with Richard J. Olson at ORNL resulted in a water quality assessment model using data from the EPA STORET system. An analysis of pollutants and water uses allowed for identification of impairments by type of use at selected gauging stations as shown in Fig. 38b (the same region as shown in Fig. 38a). Effects of individual pollutants such as iron concentration, acidity, metals, etc., could be analyzed and displayed. In studying water quality impacts for large areas, it was necessary to process pollutant data measured at many gauging stations. No digital information was available to allow easy association of these gauging stations with digitized rivers and streams. Thus, computer searching techniques were developed to locate the nearest stream within a few hundred feet of the gauging stations. The results of the water quality analysis could then be displayed geographically by river and stream as well as political jurisdiction.

Another example of geographic processing of water-related data was the estimate of water use and withdrawal as compared to water supply for major types of uses across the country (industrial, commercial, residential, agricultural, etc.). These types of analyses were done for the Water Resources Council on a hydrologic basis for major drainage basins across the United States. However, the results had to

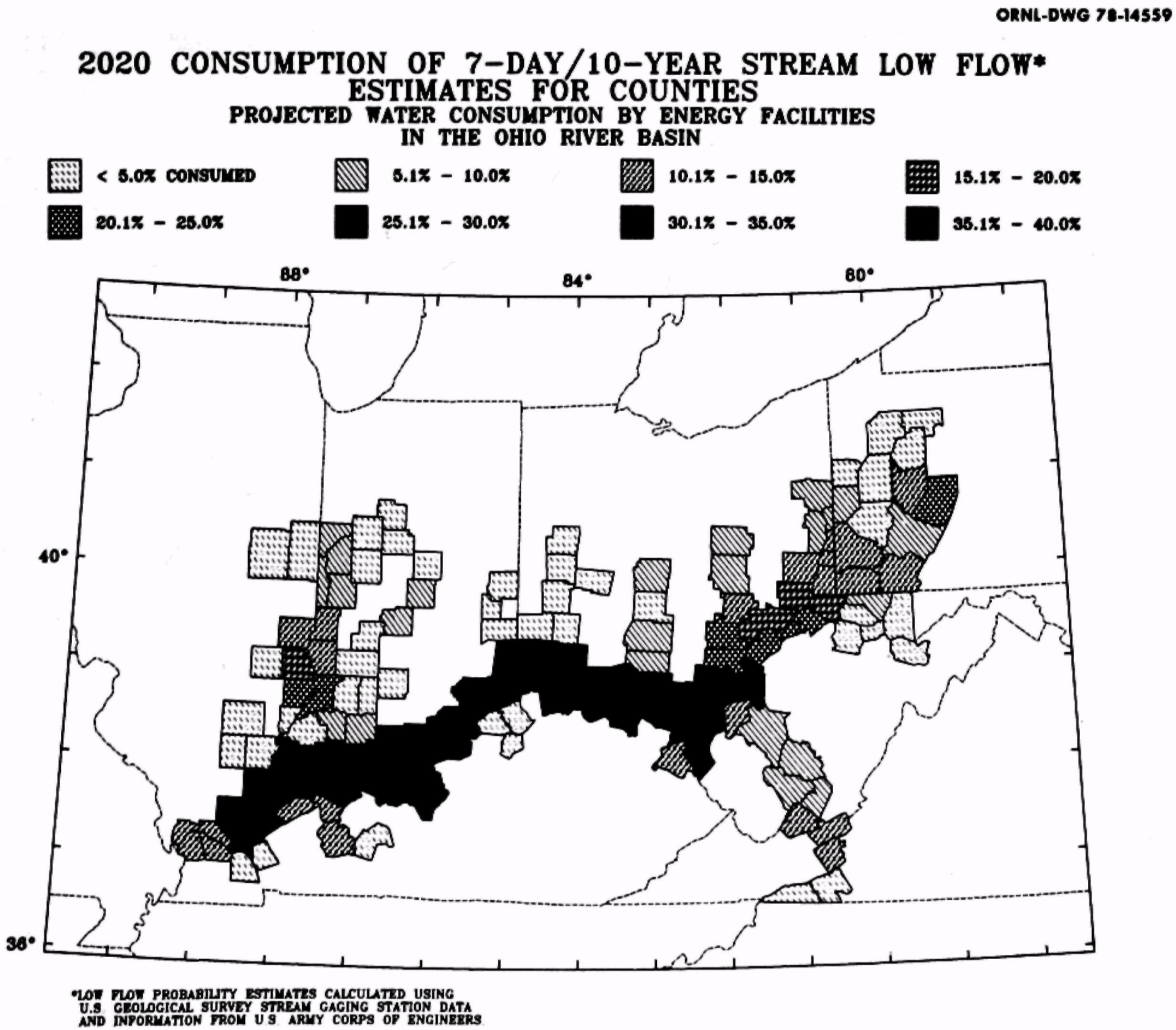


Fig. 36. Counties Whose Water Consumption for Power Plants Exceeds 5% of the Low Flow During a Drought Period (Based Upon High Demand for Electricity in 2020).

ORNL-DWG-79-21109

Fig. 37. WRC Drainage Basins for the U. S. With 1°x2° Quadrangle Borders Shown.

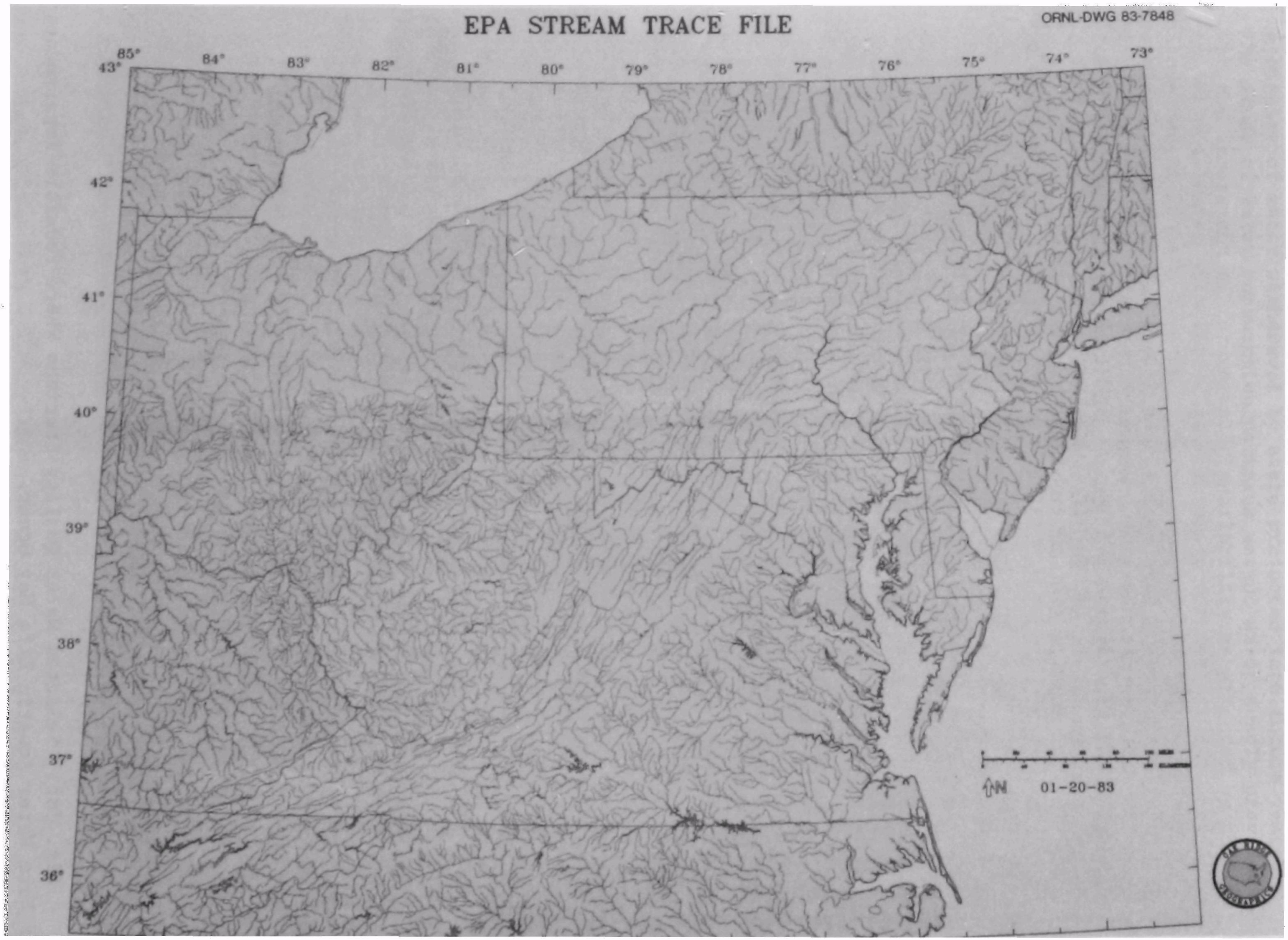


Fig. 38a. Streams and Rivers in the Eastern United States.

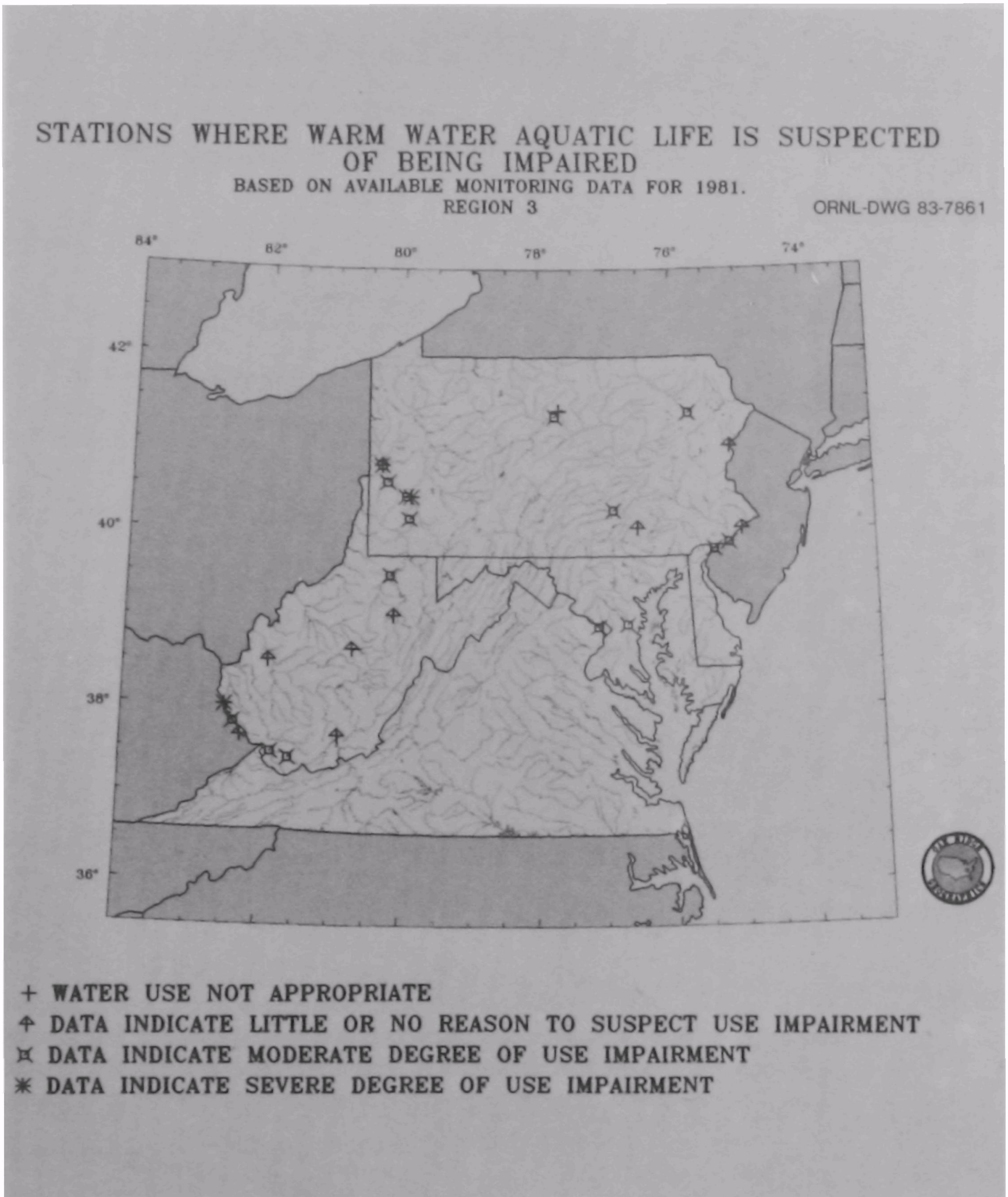


Fig. 38b. Selected Water Quality Stations Where Possible Impairments of Warm Water Aquatic Life May Occur.

also be reported by state. Since these hydrologic units crossed state boundaries, it was necessary to devise a geographic procedure for calculating state portions within each of the basins. Figure 39 shows a display of water withdrawal by state projected into future years. The study of water consumption patterns versus water supply provided for the assessment of future water availability as affected by changes in national water policies. This work was under the direction of Jerry E. Dobson at ORNL.[9] In each of these examples the initial water analyses were done by special models developed independent of any formal geographic information system. Intermediate results were then interfaced with geographic analysis routines to perform the spatial calculations which could be displayed in map form with other geographic data superimposed.

4.6 TRANSPORTATION NETWORK ROUTING AND MODELING

Another geographic application at ORNL has been the development of algorithms to compute optimal transportation routes across the United States for shipment of various commodities including radioactive waste material and coal. Three modes of transportation were developed including highways,[10] railroads, and barge channels. Figures 40 and 41 display the highway and rail networks, respectively. Figure 42 shows computed railroad routes from South Carolina to New Mexico. The initial railroad routing model was developed by Bruce E. Peterson with support from Dave S. Joy at ORNL. As with the water analyses, these primary routing models were developed as stand-alone packages which were interfaced with the geographic systems to handle the creation, processing, and mapping of the transportation networks and routes. This type of arrangement is efficient and works very well. Data are digitized and edited with the geographic systems, interfaced with specially developed models that do not require significant spatial transformations, and the results are passed back to the geographic systems for further processing and mapping.

Another transportation application has been the incorporation of airports into the systems[11] for emergency planning. For example, if an accident occurs during shipment of radioactive material that requires flying in special equipment, the system can locate the nearest feasible airport at both the source and destination and can select the optimal route. The study of commodity flows (hazardous materials, coal, etc.) has been useful to determine what parts of the country might be avoided for certain types of hazardous shipments, or to determine the capacity of the current network to handle increased flows if large coal mining areas were opened up. Figure 43 shows a hypothetical example in which routes and flows were computed from reactor sites to three possible repositories in the United States. In computing different routes it is possible to consider other parameters such as quality of the track, transfer time between trains, highway speeds, distances, special areas that block hazardous shipments, etc. The intent of these examples is not to describe the computer algorithms but to give an overview of geographic network applications that can be solved using the computer.

4.7 GEOLOGIC MODELING

Work has been done in the past on modeling geologic structures under the earth's surface. The initial application of these techniques was for estimating coal seam parameters[12] for input to external coal

ORNL-DWG 79-15164

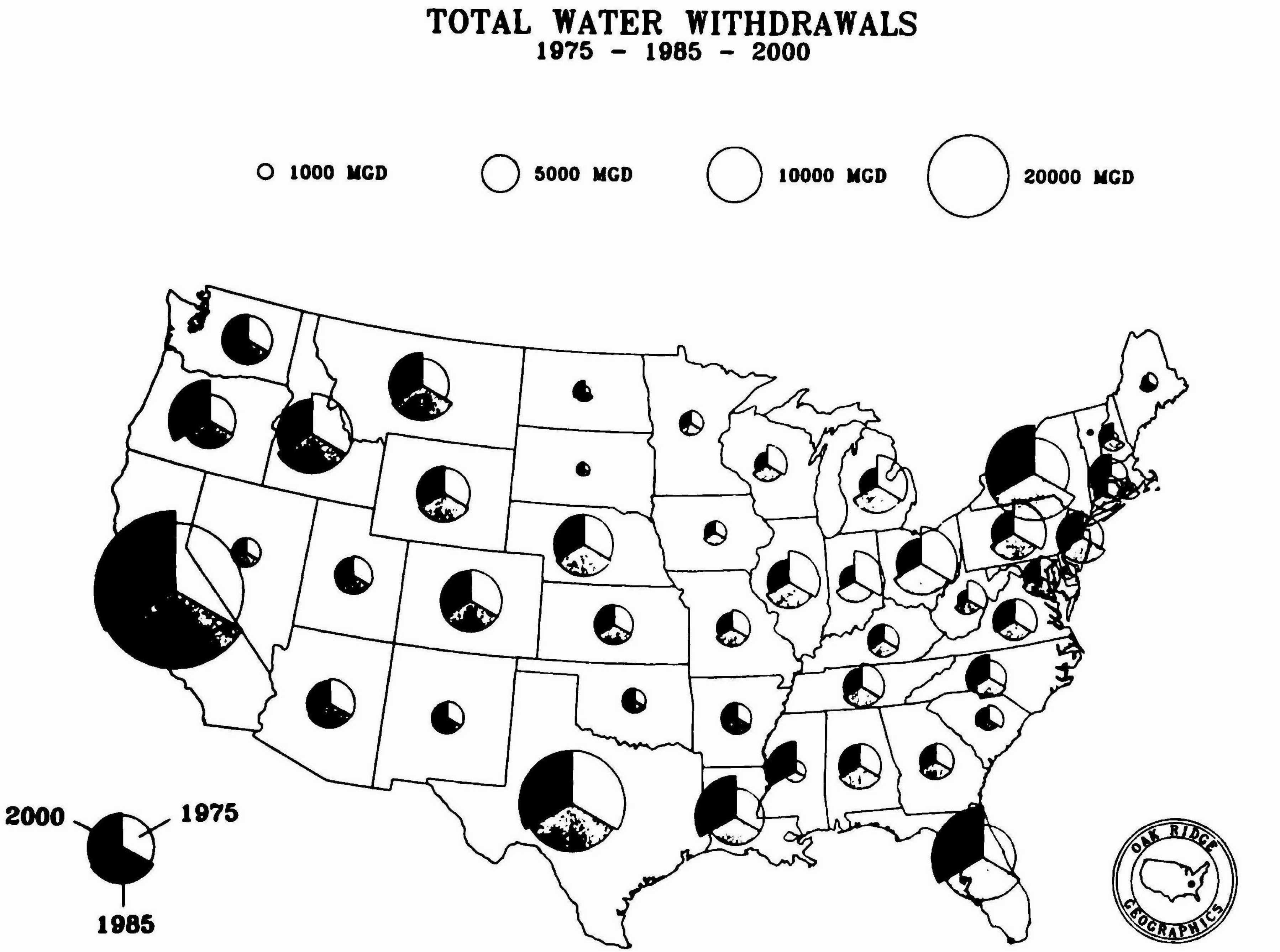


Fig. 39. Total Withdrawal of Water by State Projected Into Future Years.

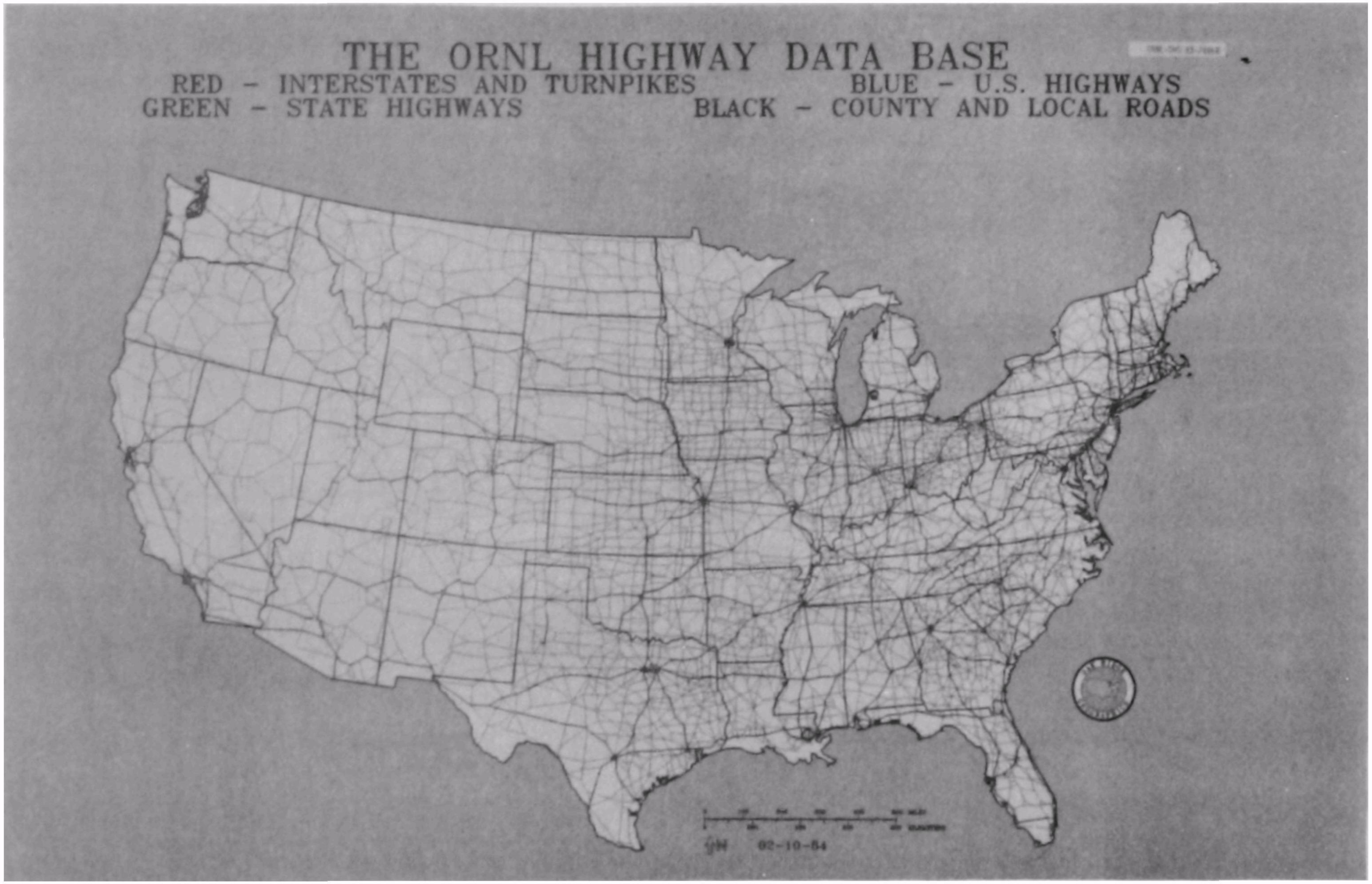


Fig. 40. The ORNL Highway Network.

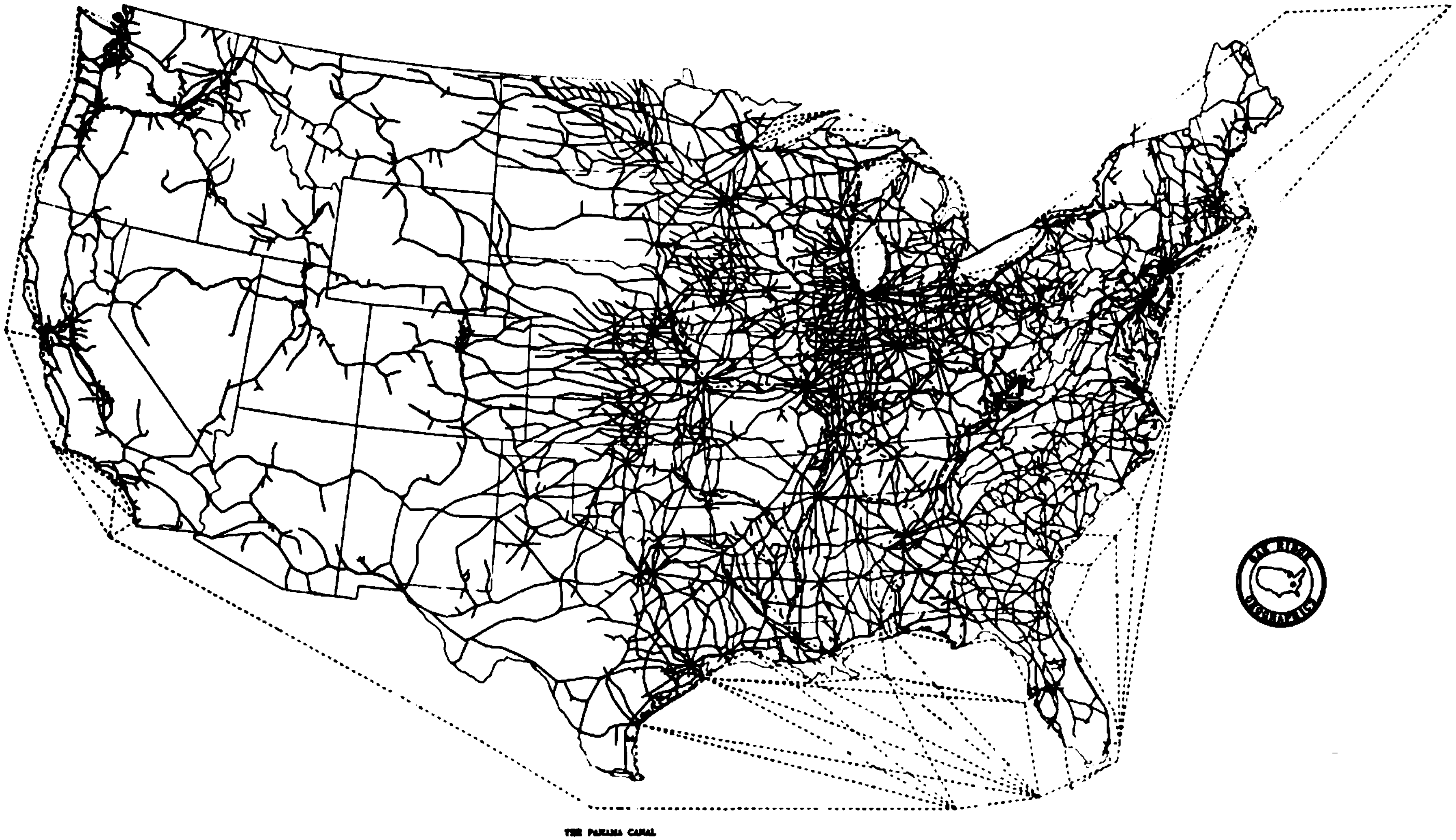


Fig. 41. The ORNL Railroad and Barge Channel Network.

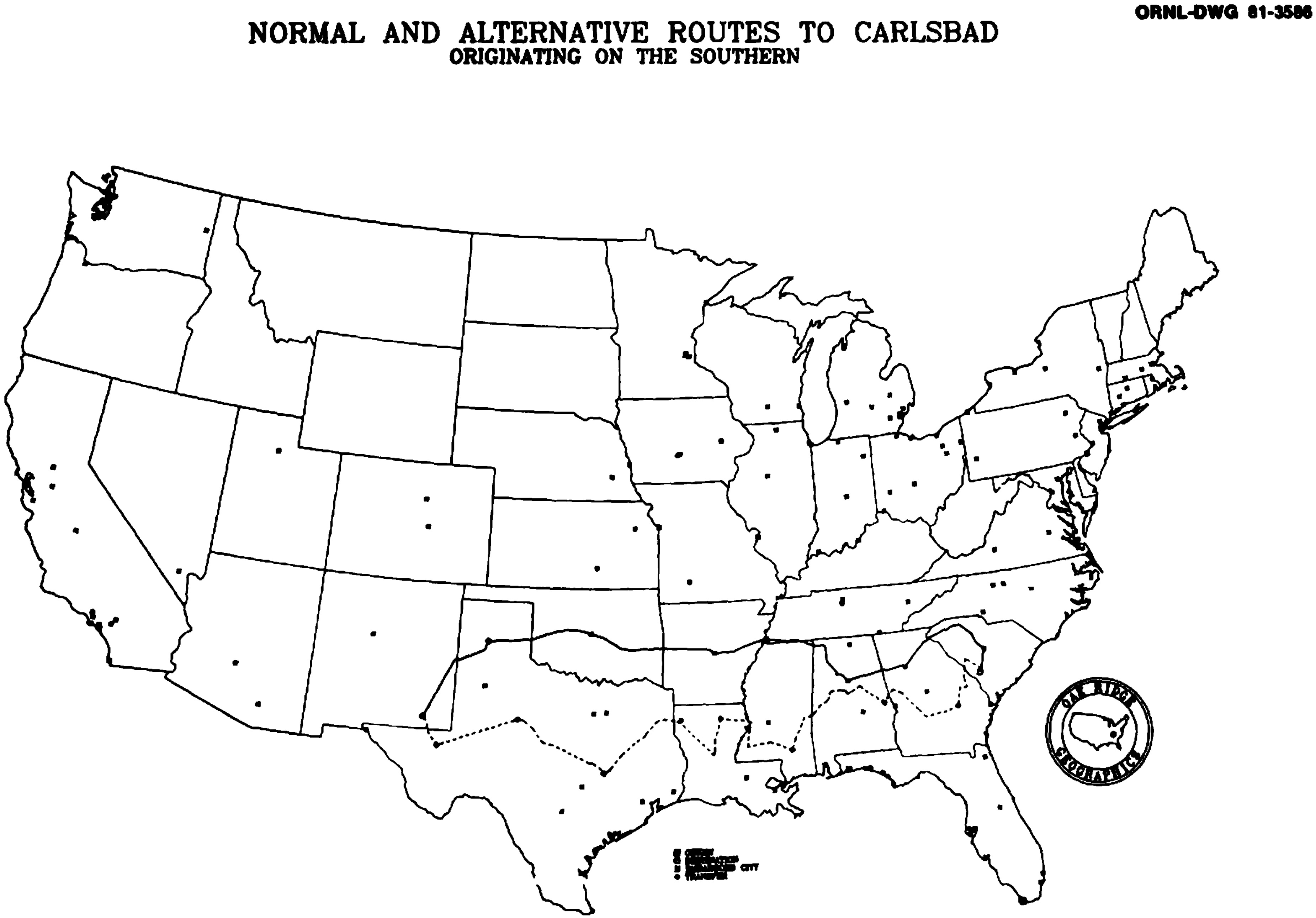


Fig. 42. Computer Railroad Routes From Barnwell, S. C., to Carlsbad, N. M.

SPENT FUEL MOVEMENT TO THREE REPOSITORIES

ORNL-DWG 80-14370A

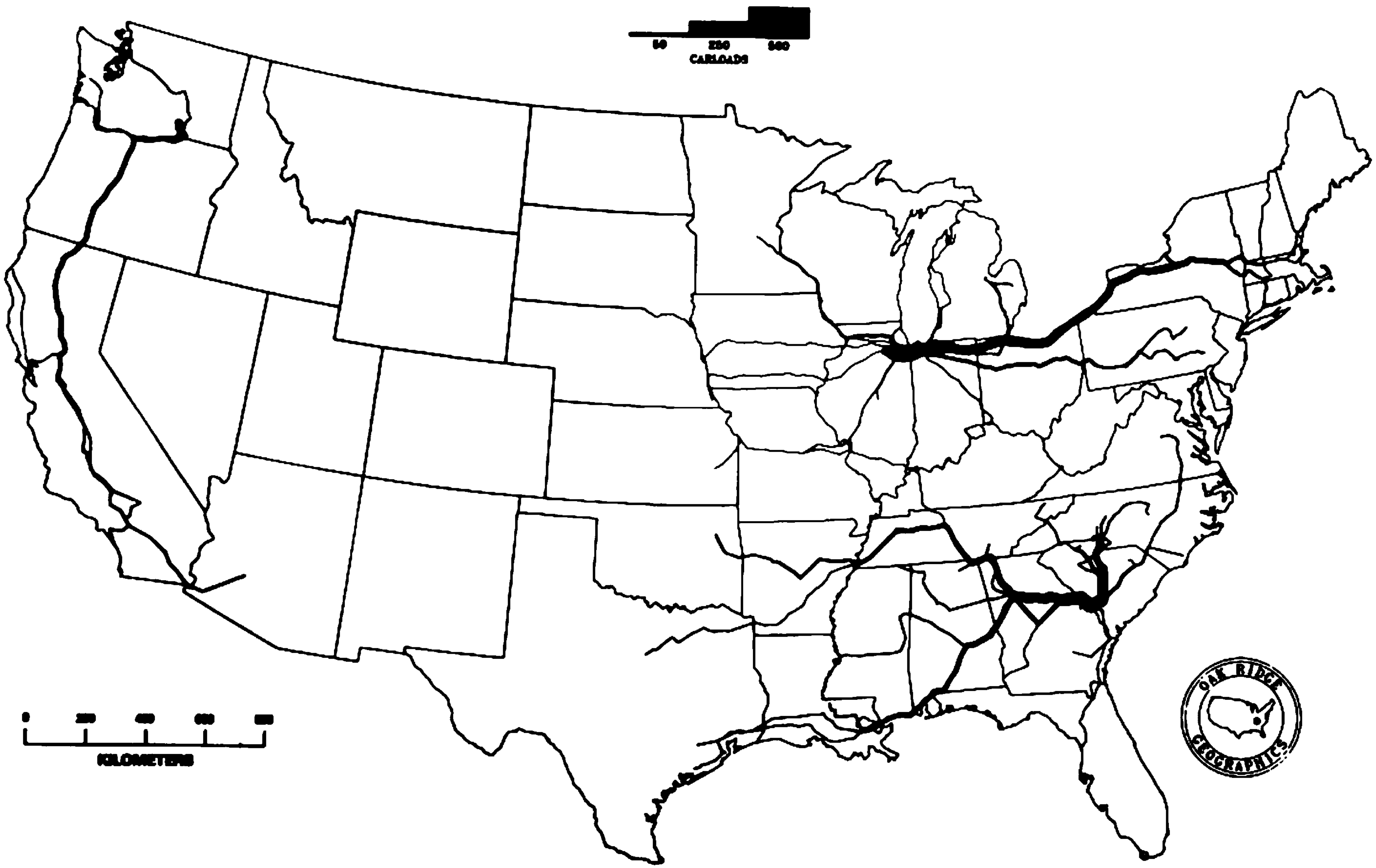


Fig. 43. Hypothetical Example of Movement of Spent Fuel From Existing or Planned Reactors With Rail Access to Possible Repositories in Washington, South Carolina, and Illinois.

cost models in estimating the type of mining and costs that would occur. The geographic systems were used to calculate parameters such as coal seam thickness, overburden, reserve estimates, stripping ratios, seam slope, etc. The information could be presented in tabular or map form as shown in Fig. 44 for coal reserve estimates in a 7-1/2 min. quadrangle. Input data were collected in three different data structures: point locations of bore holes, polygonal outlines of seam outcrops, and gridded ground surface elevations. In this model all the data were transformed into a latitude-longitude grid system after rectification to ground truth. Since only one seam at a time was being processed, it was not necessary to create voxels (see Subsec. 2.6) to represent all of the subsurface structure from ground level down to the coal seam.

Another application involved the analysis of bore hole data to determine subsurface profiles of multiple geologic formations from ground level down to the bottom of the boreholes. Figure 45 shows eight boreholes along a particular cross-section line, with their geologic layers shaded with different patterns. Interpolation procedures were used to estimate the thickness and extent of the geologic structures between bore holes. By selecting boreholes near the cross-section line, it was possible to compute seam thicknesses and locations for a grid mesh between the boreholes. Layers of each structure were created beginning at ground level and moving down one layer at a time. Problems arose at the greater depths where extrapolation had to be done, because data were not available from the more shallow boreholes. The data structure used in this analysis was a two-dimensional vertical mesh representing the vertical cross section down through the earth's surface.

ORNL-DWG 77-4195

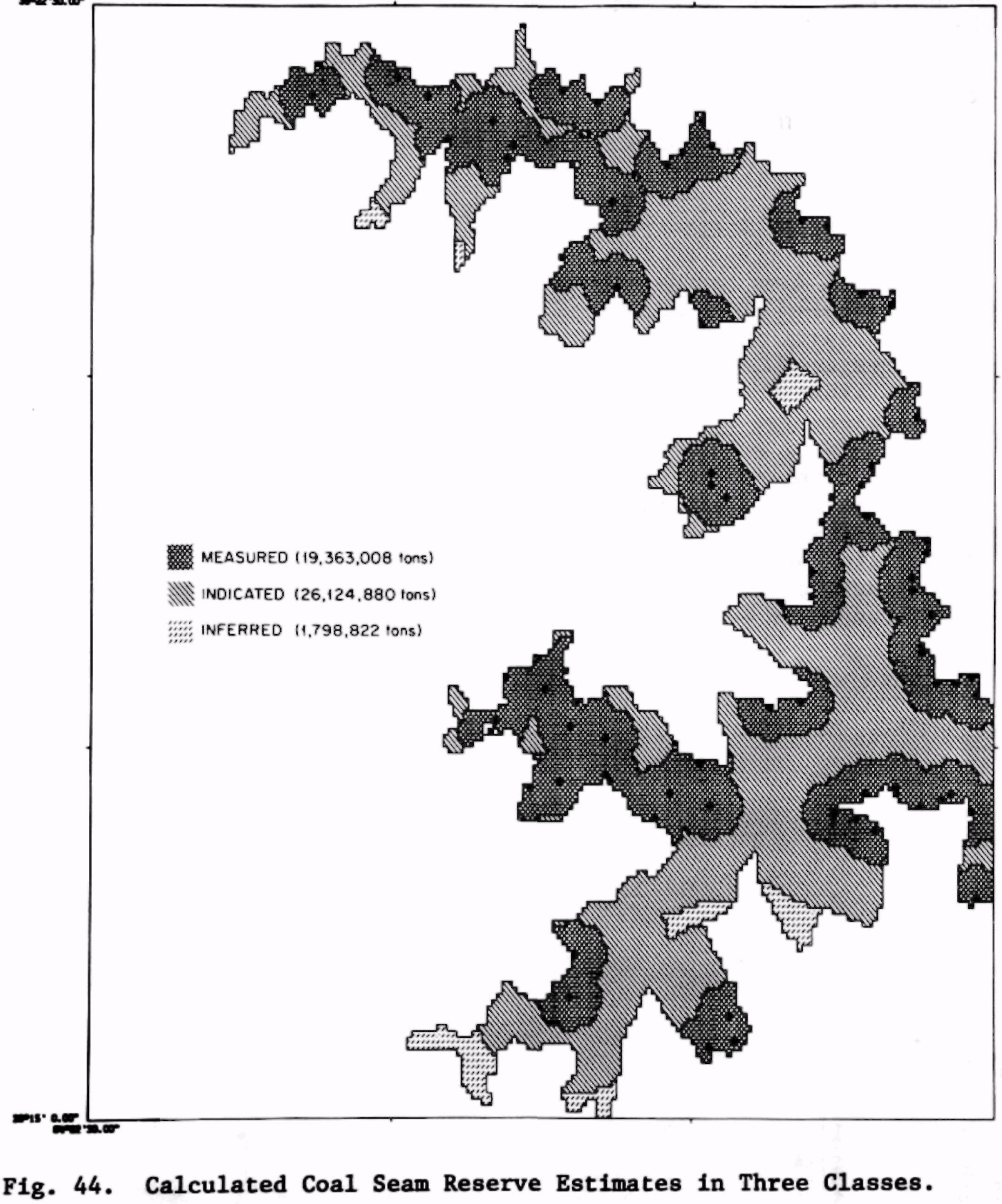


Fig. 44. Calculated Coal Seam Reserve Estimates in Three Classes.

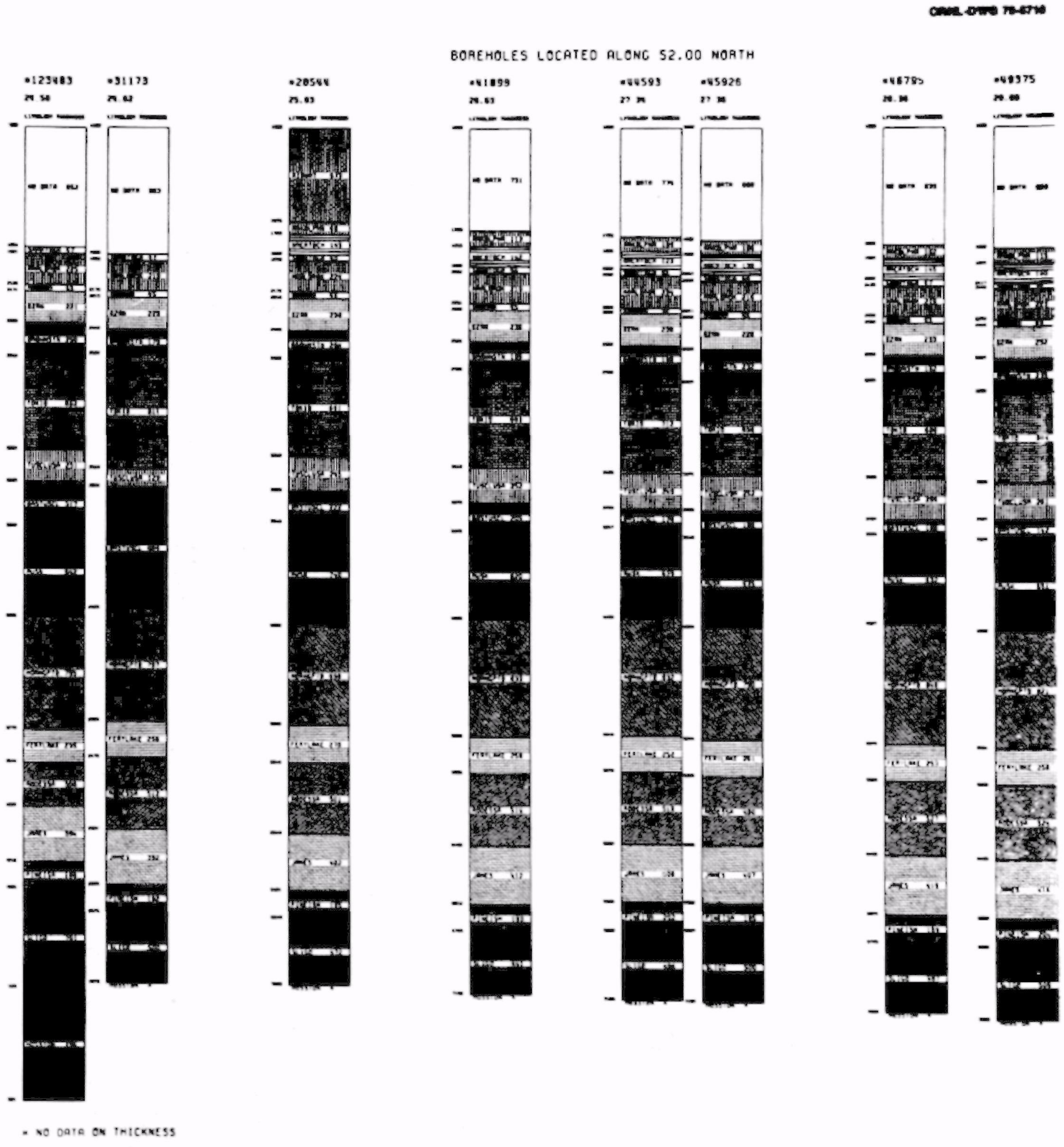


Fig. 45. Boreholes Spaced Along a Cross-Section Line with Shading Pattern Depicting the Geologic Layers.

5. GEOGRAPHIC DISPLAYS

5.1 CARTOGRAPHIC MAPPING WITH SEGMENTAL DATA

A variety of geographic displays have been presented in earlier sections to aid in explaining particular concepts or techniques. The discussion in this subsection and those following has been organized so that different mapping techniques can be discussed in an organized fashion. The intent is to give the reader an overview of the different types of displays that can be produced to best present his/her information graphically. Most mapping applications require the display of base reference information (e.g., state and county boundaries, cities, highways, rivers, etc.) to which thematic data may be referenced. For example, a map of power plants or industrial facilities across the United States would require, at a minimum, state boundaries for reference purposes. These types of geographic features are normally digitized in a segmental structure such as chains or polygons. Techniques for displaying these features are referred to here as cartographic mapping because they focus on presenting the cartographic base features traditionally mapped by cartographers to specific map standards.

The first requirement of segmental mapping is to simply plot points or linear features from one location to another. An example would be mapping roads, rivers, or oil wells. Lines may be dashed, colored, or of varied width or darkness. The specific map projection and scale may affect the plotting of line segments between coordinate points. For example, a latitude line drawn from one point on the west coast to another point on the east coast in an Albers' Equal Area projection is curved and cannot be plotted as a single straight line between the end points. It must be automatically broken into many small sections by the computer so that, when combined, a smooth curved line will be displayed at the scale chosen. Some applications require the labeling of linear segments with their appropriate names (highway names, county names, river names, etc.) as shown in Fig. 46 for the Oak Ridge DOE reservation. Locating and selecting proper sizes for the labels is a difficult procedure if attempted through automated computer techniques with complex maps. In many cases, it is better for the digitizing operator to input the position and size of label names so that they will follow the curvature of the line segments.

The display algorithms are not only dependent on map projections, data structures, and the types of graphics desired (e.g. boundary maps, shaded polygons, raster densities, perspective surfaces, etc.), but are also related to the computer systems and output devices used. Mainframe batch processing with output to a large off-line plotter may use massive sorting and subsectioning procedures, whereas an interactive graphics processor on a user workstation may integerize the data and use special hardware functions and lookup tables to produce color displays on a CRT screen. Many of the computer maps in this report have been produced on hardcopy plotters. However, an inexpensive

OAK RIDGE RESERVATION INCLUDING N.E.R.P. AREA

ORNL-DWG 83-7856

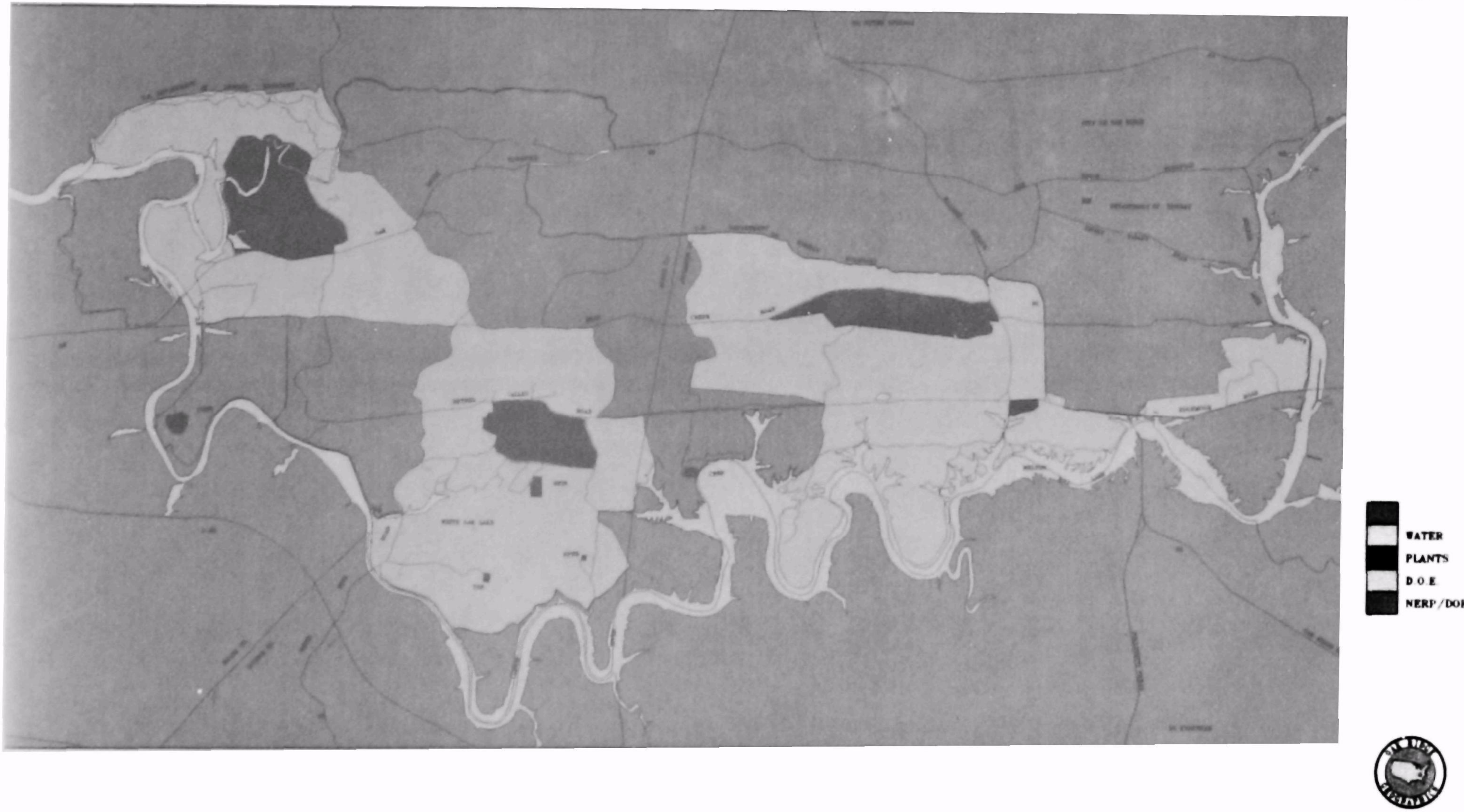


Fig. 46. The Display of Cartographic Segmental Data Including Highways, Plant Facilities, County Boundaries, Rivers and Lakes, Label Identifiers, and the DOE Boundary.

high-resolution color system has also been developed by the Geographic Data Systems Section as an interactive workstation for map display and geographic analysis. It is based upon an IBM AT personal computer linked to a graphics processor with a 1024 x 1024 color monitor and a digitizing tablet. A typical CRT display would overlay a LANDSAT Thematic Mapper image with segmental data such as landuse polygons and transportation networks. The user might interactively move a proposed "target polygon" around the screen (e.g., a shopping center or industrial complex) to find and store the optimal location. A hard copy of the result could be produced locally or off-line. The microcomputer can be linked to a host so that work can be done on either machine depending on the capabilities needed. The prototype system is shown in Fig. 47. Interactive use of this type system will allow different analysis and mapping techniques to be tested easily with multiple data bases before making a hard copy. Analysis results can be viewed quickly to discover errors, validate hypotheses, and perform the graphic editing needed to best present the results. Both vector and raster data can be simultaneously displayed and manipulated. Since the basic data are stored in latitude-longitude coordinates, any data base can be integrated with any other file at any scale or resolution. A data management system provides efficient storage, retrieval, query, selection, and updating of the files.

The mapping of polygonal features, such as county outlines, provides an opportunity for further display techniques such as shading the area inside the polygons with different patterns or colors. Figure 48 shows counties shaded to represent sulfur dioxide measurements in micrograms per cubic meter. The computer techniques will depend on whether a vector or raster device is used for plotting the information. A raster device such as an electrostatic plotter or raster driven CRT can produce continuously varying shades of gray or color within the polygon, whereas a vector device, such as a pen-and-ink plotter, must produce patterns through cross-hatching lines within the polygon. Figure 49 shows a raster ink-jet plotter that can produce multicolored displays by superimposition of three color patterns (magenta, yellow, cyan). Vector data can normally be plotted on a raster plotter through software transformations, but raster data cannot easily be plotted on a vector device. For specific polygon identification, it is useful to provide numbering or labeling at an optimal position within the polygon. Simple techniques compute a centroid for the polygon to determine a labeling position. However, in some cases, the centroid may actually fall outside the boundary of a complicated polygon, so more sophisticated techniques are needed. One approach at ORNL is to find the position for placing an ellipsoid inside the polygon which will maximize the ellipsoidal area. Then the major axis of the ellipsoid may be used to find where the label should be written, its angle of rotation, and its size. Data processing requirements for these types of complex algorithms using large data bases (thousands of polygons) are normally batch-oriented, rather than performed interactively on a microcomputer.

There are many other general graphics functions that must be performed in producing map displays such as windowing, clipping, rotation, stripping for plots that are too large, compositing, etc.

ORNL PHOTO 2013-87

Fig. 47. High-resolution Color Geographics System utilizing a Microcomputer-based Workstation.

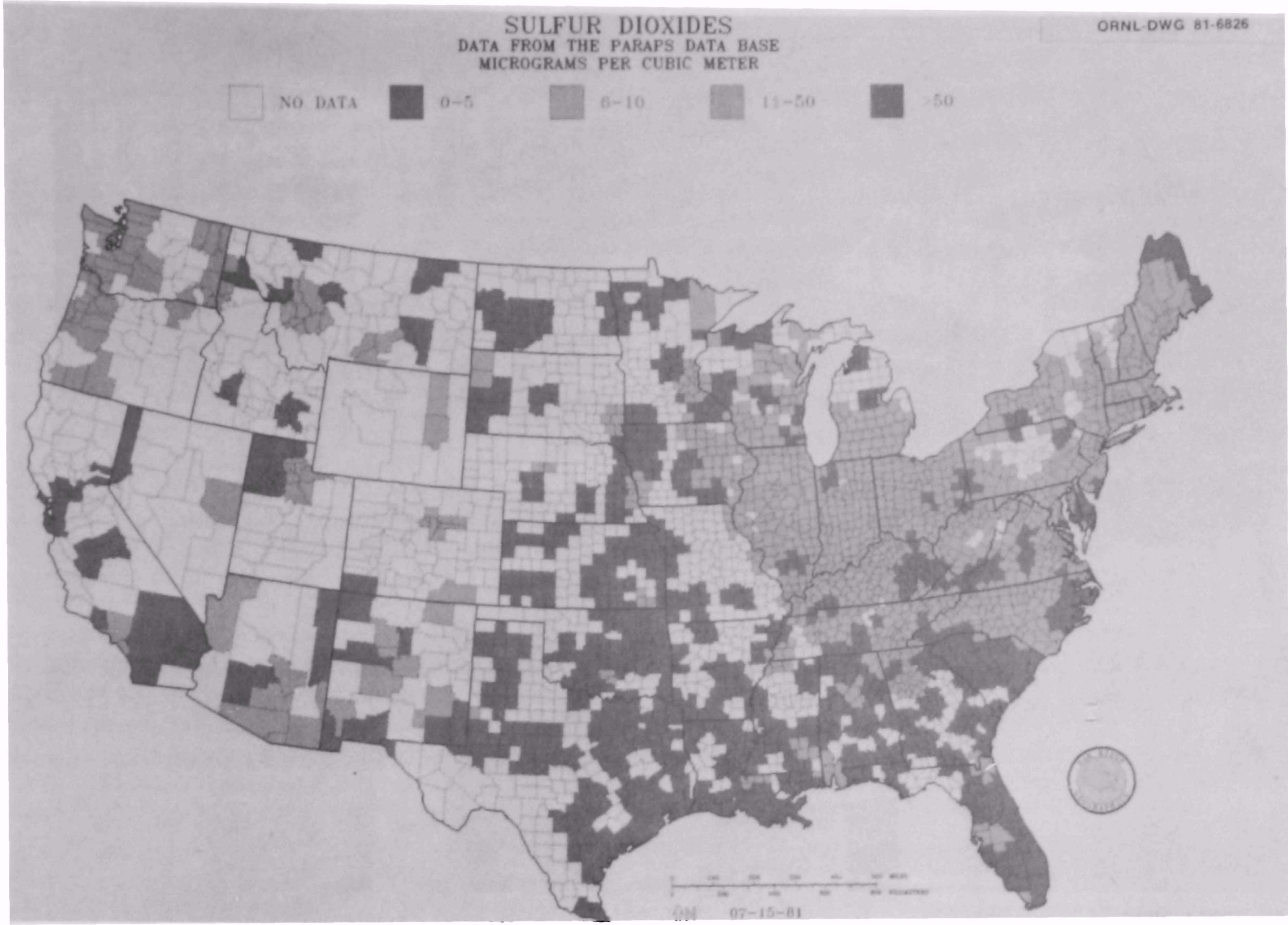


Fig. 48. Shaded Counties Corresponding to Sulfur Dioxide Measurements Averaged Over the County.

ORNL PHOTO 0284-82

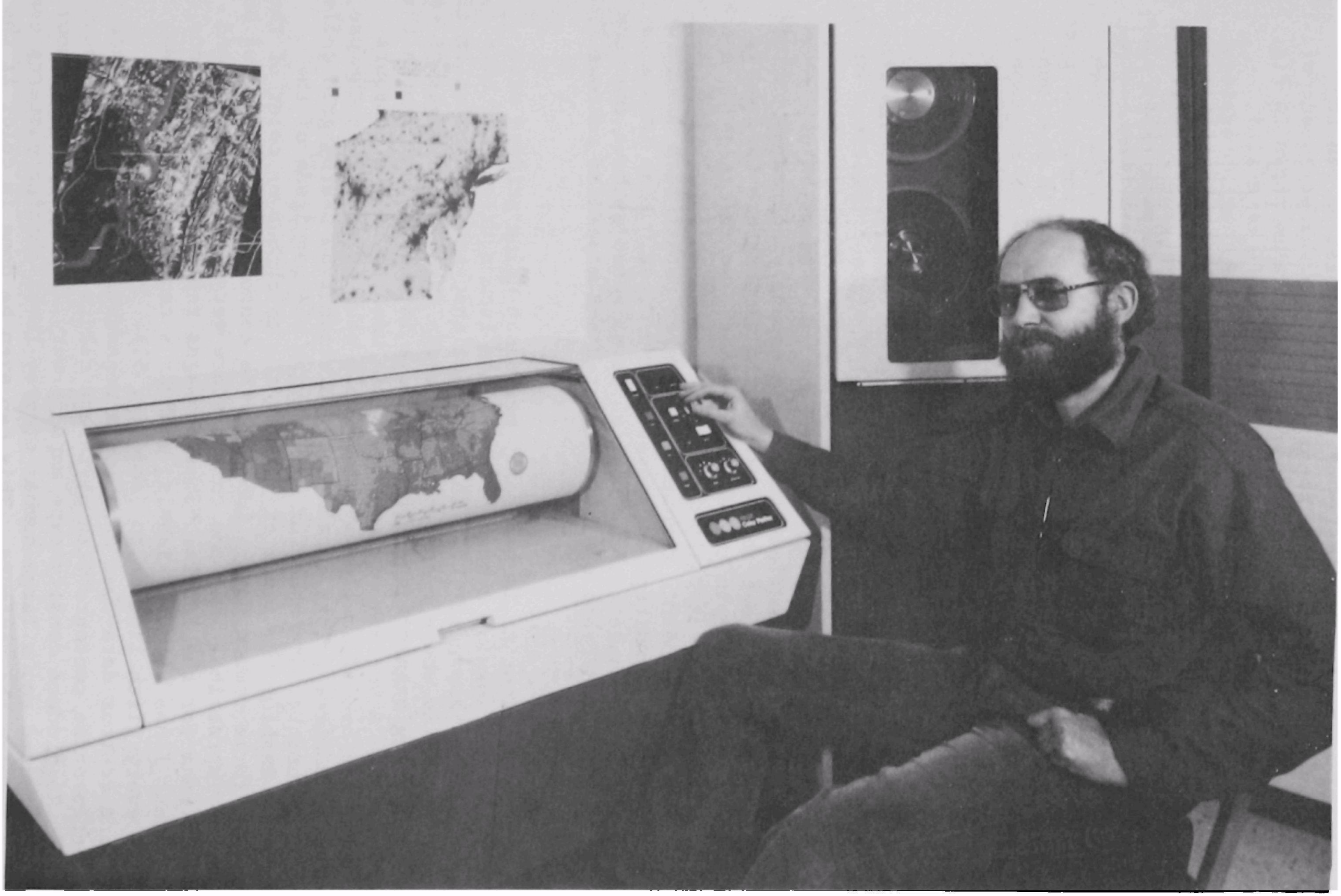


Fig. 49. Color Ink-Jet Raster Plotter.

One example of compositing is the overlaying of different polygonal files on the same map. A display of current population density patterns computed from one file superimposed with shaded Bureau of Economic Analysis (BEA) polygons from another file is shown in Fig. 50. The intent is to present current conditions in 1980 at a fine resolution level with projected growth areas by 1990 at an aggregate level. A clipping function was required to plot state river maps where the streams were to terminate at the state border. This required an intersection of the river chains with the state outlines. When maps will not fit on the plotter paper in one direction, the computer can sometimes rotate the information so that it will fit in the other direction. Plotting of titles and legends is also an important part of map display. A variety of techniques have been shown in previous examples. A number of these functions are not unique to geographic mapping but are required for all types of computer graphics.

5.2 CONTOUR MAPPING

Contour maps are normally used to present geographic data that varies continuously over a two-dimensional surface. Locations which contain data values of the same magnitude are connected to form isolines within local areas throughout the map. Examples of continuously varying functions that are frequently contoured include terrain elevations as feet above sea level, temperature in degrees, barometric pressure, depth to bedrock in feet, and thickness of a coal seam in inches. Other geographic entities, such as population, may be located at discrete points throughout a region rather than occurring at every spot on the landscape. However as the distribution increases in urban areas and a scale of presentation is selected without too much detail, density functions can be calculated and mapped as contour lines (people/mile^2).

Machine contouring is generally performed using gridded data either in a rectangular system or sometimes from a triangular base. Data values are known at the grid intersections, and interpolation is done to determine where a contour line passes through the edges of each cell. Plotting segments through these edge intersections creates the contour line. A variety of techniques are used to smooth contours including the use of smaller meshes within each grid cell or the use of additional control points in the interpolation from nearby grid cells. Several techniques are used for identifying the magnitude of the contours. Dashing the lines or varying the thickness or color of the lines are three common ways. Labeling the contour lines is useful but sometimes impractical when the contours are very close together and leave no room for the labels. Some packages replace small sections of the contour lines with the labels, much as a cartographer would do, perhaps every fifth line or so (see Fig. 51). An excellent way topresent contour data that can be comprehended quickly is to shade the area between the contours with different colors or patterns as shown in Fig. 52. To produce density type maps,[13] very closely spaced contours can be computed and a large number of grey-level shading patterns can be used between the contours. This can even be done in color if the plotter or CRT has the ability to slowly change from one hue to another.

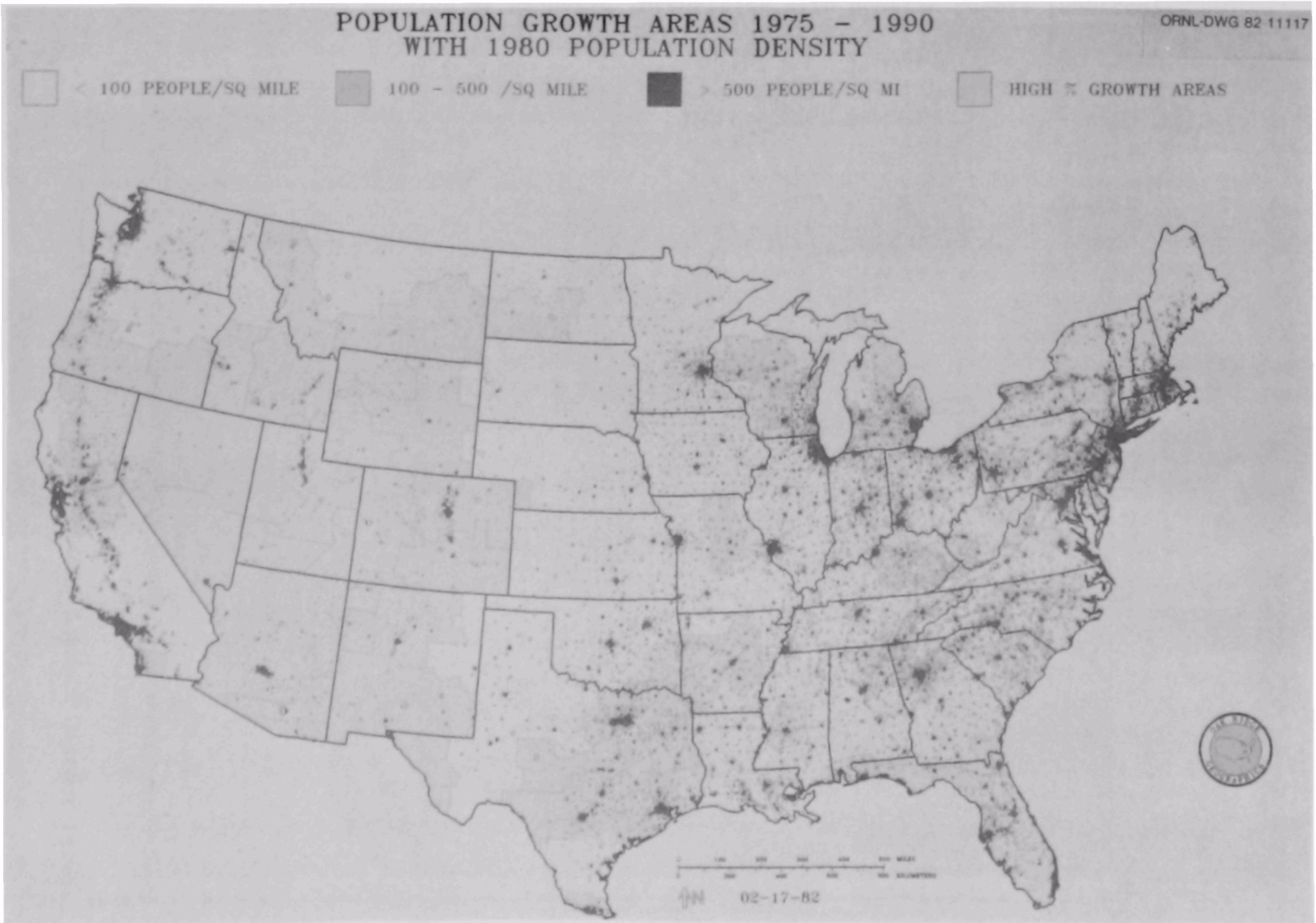


Fig. 50. Superposition of Population Density Patterns With BEA Areas of High Projected Growth.

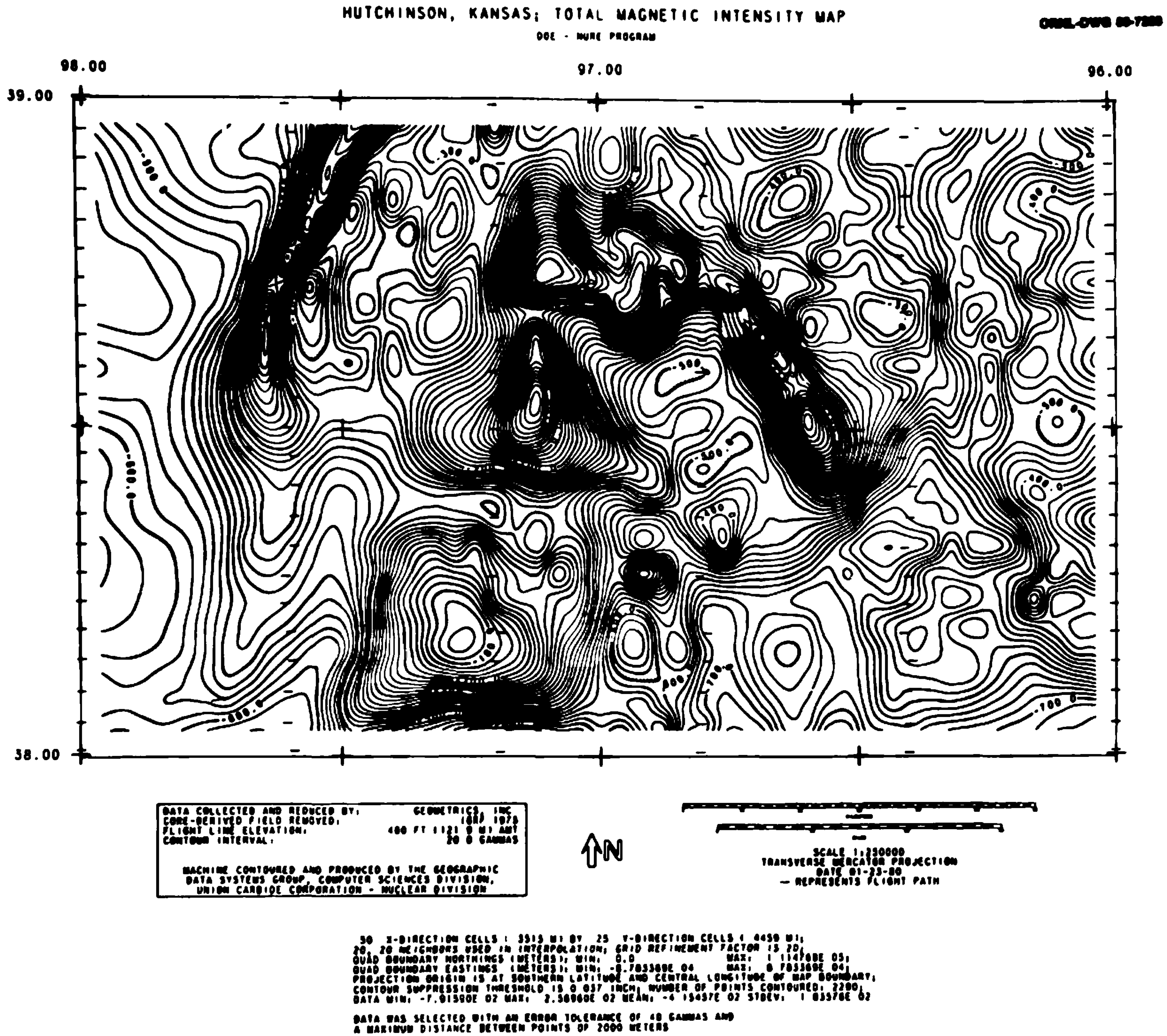


Fig. 51. Contour Mapping of Magnetic Intensities With Numeric Labels Inserted.

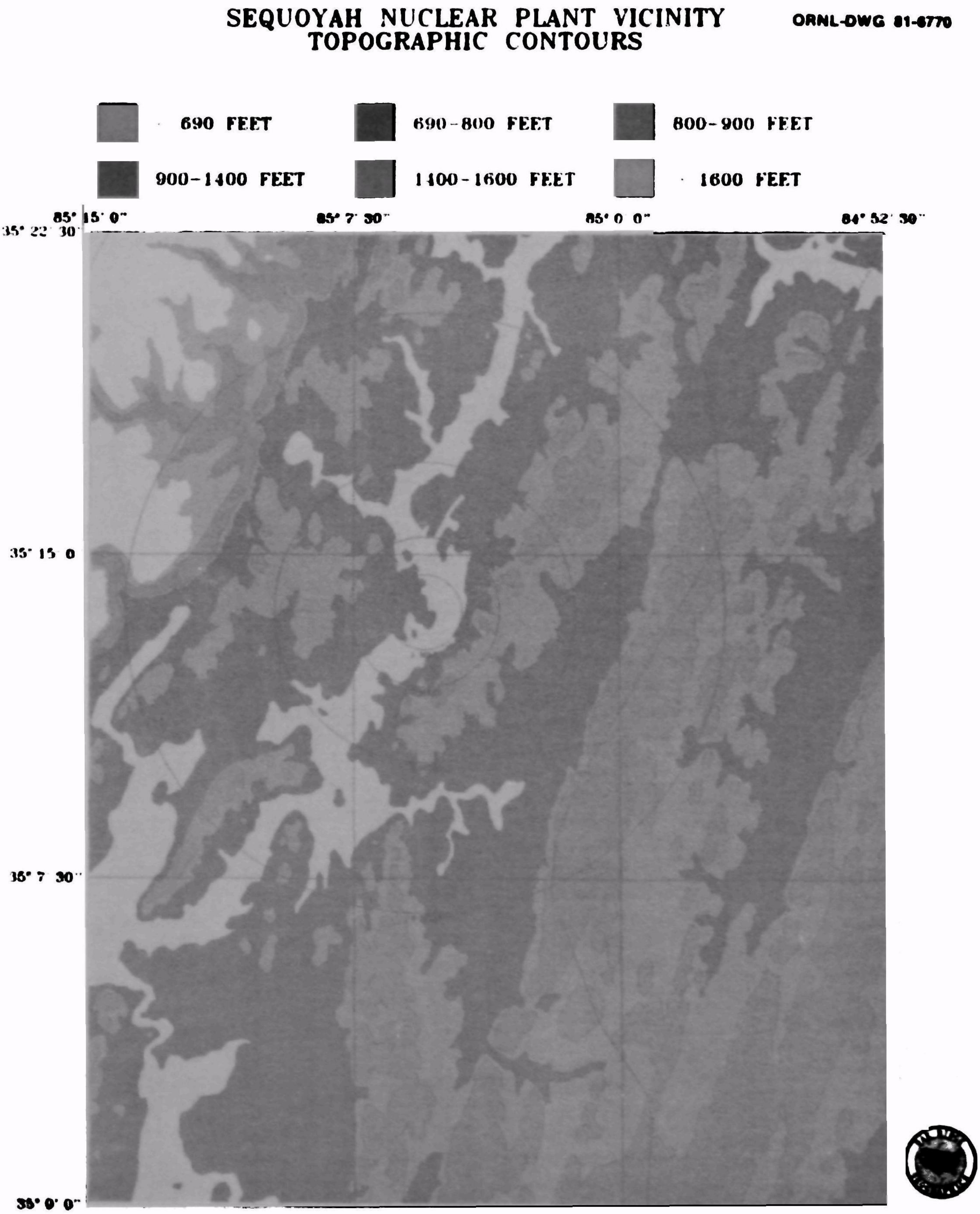


Fig. 52. Topographic Contours With Shading Between Contour Levels.

Normally the calculation and plotting of contours is considered to be a display mechanism. However, with appropriate algorithms the contours can be transformed into a chain or polygonal data structure and saved as a new data base for further computations. For example, on a regional basis, population density contours of 500 people/mile^2 have been saved as a polygonal data base representing urban areas. Then, it is possible to approximate the miles of highway within an urban area or the square miles of urban land affected by air pollution from nearby emissions. Many of the software techniques used in producing graphic displays are identical to those used in the spatial transformation procedures discussed previously.

5.3 GRID CELL AND RASTER MAPPING

A wide variety of techniques are available for producing maps of gridded data. The earliest techniques consisted of over-printing different characters on a line printer to produce grey-level maps where each printer position corresponded to a grid cell. Since the printer positions were fixed, it was not possible to produce cartographic maps as a function of scale and map projection. More recent techniques shade the individual cells with different patterns or colors, much as the polygons are shaded. Again, the techniques are somewhat dependent on the type of plotter or CRT used.If a cluster of grid cells have the same data value, their outer boundaries can be linked together to form polygons and the same shading technique used to color in the area, as shown in Fig. 53 for a soil interpretation map. The outer boundary of each cluster of cells can also be drawn rather than having to draw all the individual grid lines. Numbering or labeling of grid cells can be done for each individual cell or placed at the center of a group of cells all having the same value. This is done by computing the centroid of the polygon enclosing the cluster of cells. Again the usefulness of simultaneous grid and polygon processing is evident.

On raster plotters it is possible to create grey level patterns similar to photographic screens used in printing maps. These different grey levels or color densities can be used to depict data which change from cell to cell, such as the LANDSAT intensities shown previously. Shaded relief maps can be plotted from gridded terrain models by using the density shading techniques in combination with topographic analysis models described previously in another section. The slope and aspect of each grid cell are combined with sun angle and line-of-sight calculations to compute the proper shading. This gives an effect of bright and shadowed areas across the terrain.

The majority of display techniques discussed in these subsections are oriented toward traditional hard-copy plotters, both raster and vector. There are other hard-copy devices on the market (e.g., film recorders and thermal plotters) which can be linked directly to CRT analog signals coming from the computer so that photographs (or other hard copy products) can be created to reproduce screen images directly. In a film recorder the red, green and blue output signals from the computer can be displayed respectively on a very high-resolution CRT inside the unit. A polaroid camera, 35-mm camera, or movie camera mounted in front of the CRT exposes the film. Thus, high-quality

ORNL-DWG 83-6766

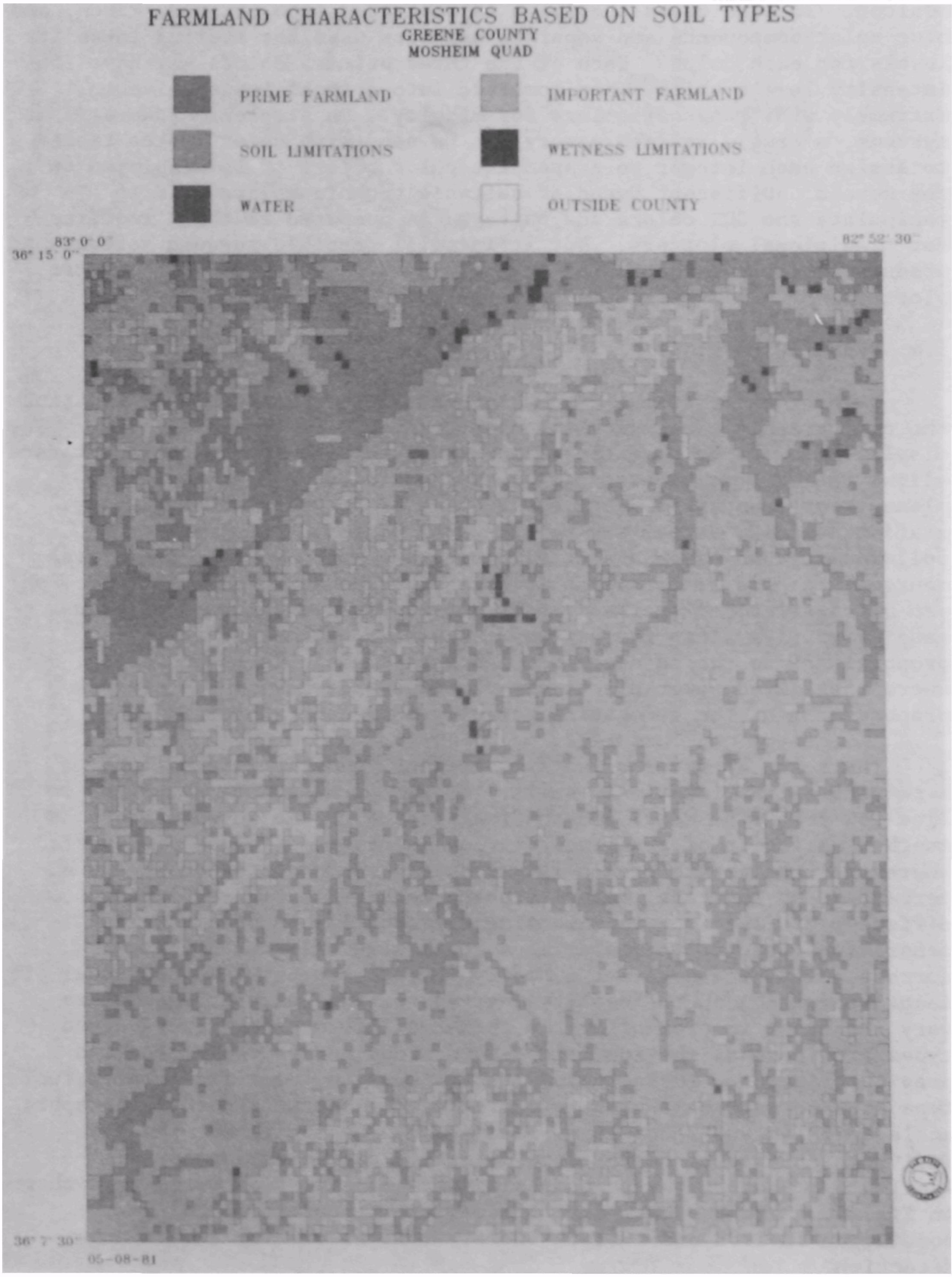


Fig. 53. Grid Cell Map Depicting Farmland Characteristics From Soil Interpretations.

photographs are made directly from the color image. High resolution raster data is displayed on a color CRT by feeding the pixel intensity values into a refresh memory (frame buffer) that drives a color monitor. In some cases the image is separated into its red, green, and blue color components and separate memories used for storing intensity levels for each color. Each of the three primary colors may have 255 intensity levels which, when combined into a final image, give an extremely wide range of colors for display. On simpler hardware systems, a single refresh memory can be used with color lookup tables to assign each integer to a specific color before it is displayed on the screen. Different types of software techniques are used to manipulate the CRT colors and patterns as compared to those required for traditional plotters. Yet it is still possible through software to produce large paper copies of CRT color images using off-line raster plotters.

5.4 THEMATIC MAPPING

Much of the previous discussion has been oriented around plotting the cartographic base information associated with geographic data. The display of measured thematic variables along with the cartographic data allows the user to understand the spatial relationships among the elements or quantities being measured. For example, water gauging stations along a stream may be measuring different parameters or pollutants, such as pH and total dissolved oxygen. After the stream course is mapped as cartographic data, the thematic variable (e.g., pH) can be superimposed (perhaps as graduated circles centered at each gauging station along the stream). The size of these circles is proportional to the pH measured. Another example is shown in Fig. 54 where coal-fired power plants are plotted as circles whose size is proportional to the generating capacity on a county basis.

There are a number of ways to present thematic information depending on the locational entity being represented (e.g., point-, line- or area-type data) and the imagination of the investigator. Another variation of graduated circles is referred to as pie charts, where the circles are broken into sections, each corresponding to a percentage of the total quantity being measured. Figure 55 shows different pie sections representing different types of electrical generating capacity in India. The total size of each circle corresponds to the megawatt capacity in each state. The individual pie sections are shaded to identify the type of fuel used. Legends are very important in presenting the thematic information. Four legend types are shown at the top, right-hand side, and bottom of the two previous figures. These include the title, the identification of fuel type, the circle scale used in depicting megawatts, and the geographic scale in miles and kilometers. Numbers can be put inside the individual pie sections for more detail if sufficient space is available. Another way of comparing data for two time periods is shown in Fig. 56. The two half-circle sizes are proportional to the population in 1971 and 1981 in India. This allows for easy change detection.

Thematic data gathered for polygons can be depicted using the same techniques discussed previously (e.g., shading or coloring the polygon proportional to the thematic data measured). Numerical values can actually be plotted inside the polygons, although they may be difficult

COAL FIRED POWER PLANTS, 1980

ORNL-DWG 82-11147

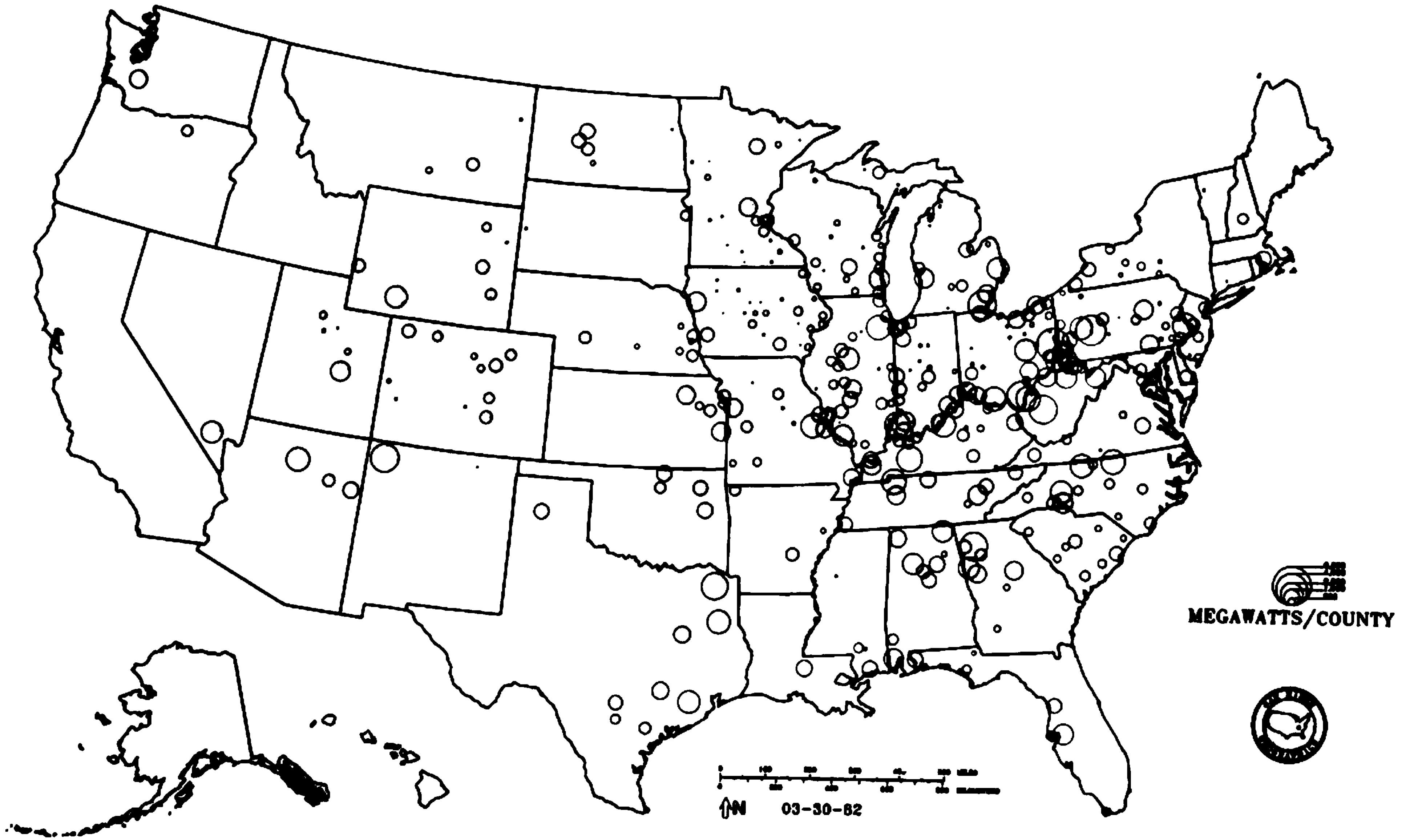


Fig. 54. Coal-Fired Power Plants by Generating Capacity at the County Level.

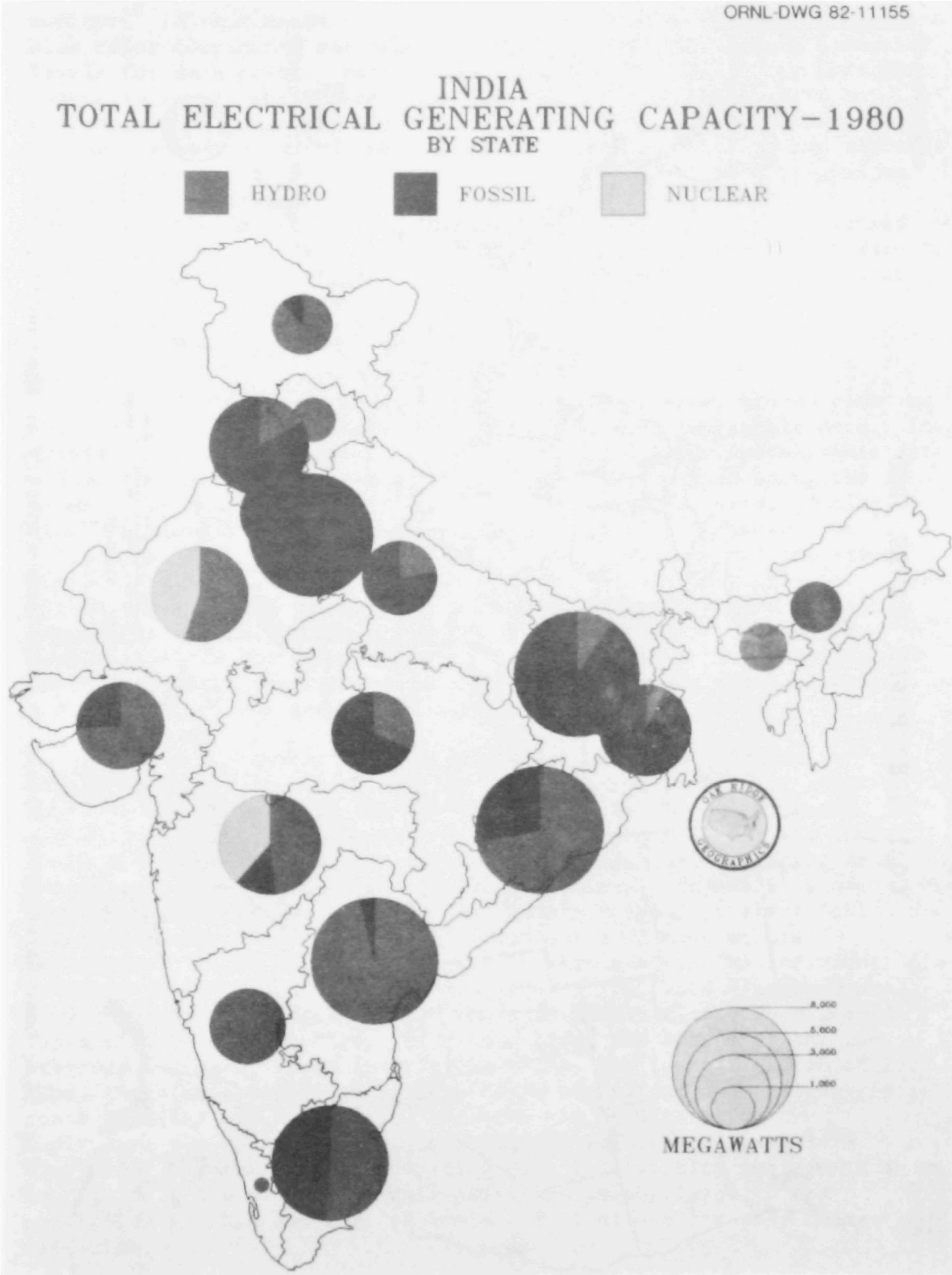


Fig. 55. Pie Section Display of Electrical Generating Capacity by Type in India.

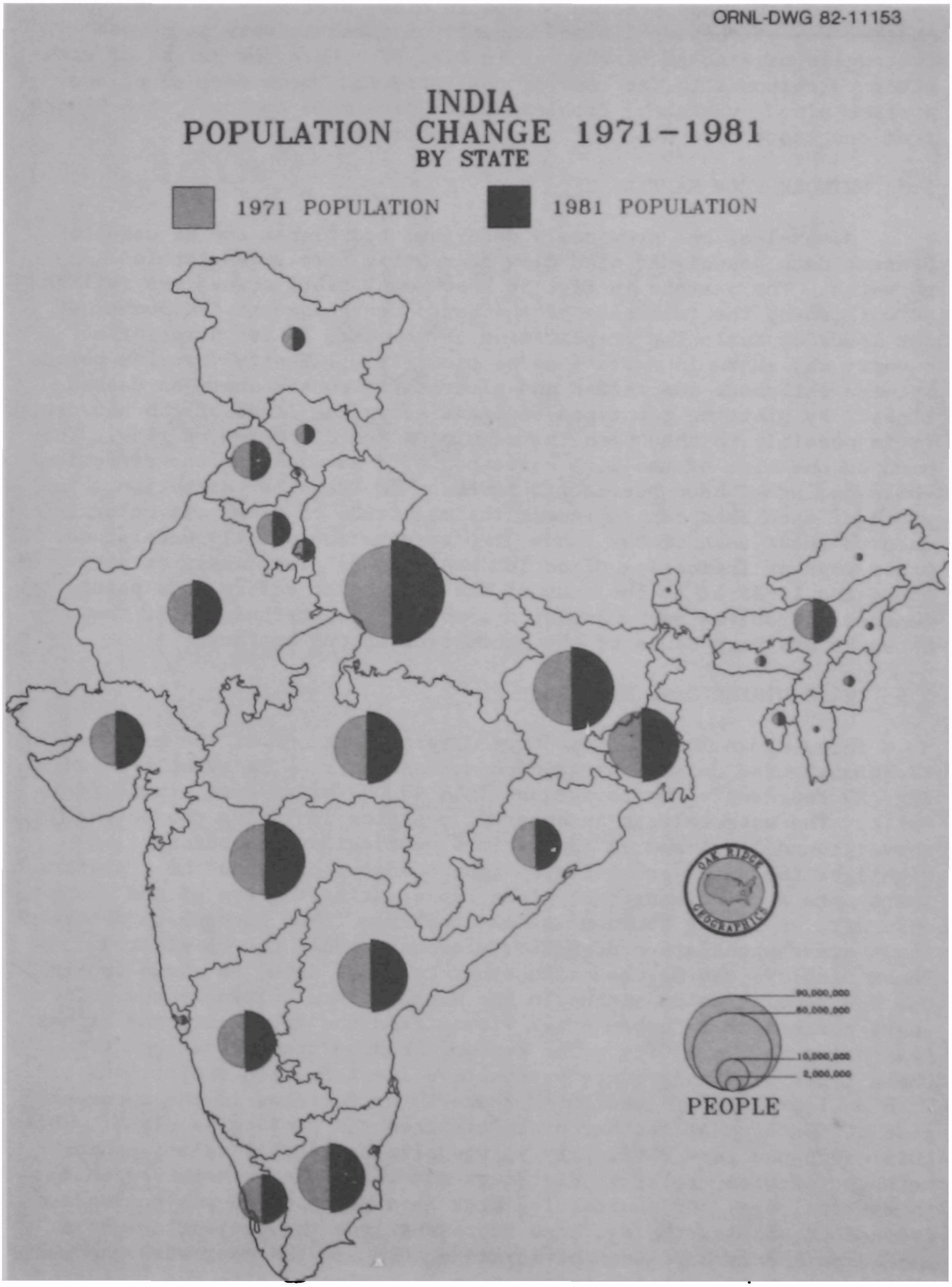


Fig. 56. Population Comparisons in India for 1971 and 1981.

to read if the polygons are small. Dot patterns can be used to present distributions of data such as national population counts. A dot might be plotted for every hundred people in close proximity to each other. Another way of showing limited amounts of thematic data is to use rectangles or stacked blocks, as in Fig. 57. Here the height of each block corresponds to the tons of coal produced, both deep mined and surface mined by state. Problems occur with this approach when blocks from one state hide blocks from a neighboring state.

5.5 NETWORK FLOW MAPPING

Several of the previously described techniques can be used to present data associated with flow or routing across geographic networks. The example in Fig. 58 shows coal flows across the railroad network where the thickness of the bands correspond to the number of car loads of coal. The presentation of computed routes across the country was shown in a previous section (Fig. 42) with transfer points between railroads identified and alternative routes shown as dashed lines. By plotting two types of bands along the links of the network, it is possible to show both the magnitude and direction of flow. The band on one side of the link corresponds to movement in one direction, while the other band represents movement in the other direction. The width of each band can represent the magnitude of flow. If color is used to shade each of the bands, the results are quickly understood. Other ways of indicating direction would be to plot labels or arrows along the links or at the ends of the links just before node points are reached. Problems may arise with some of these techniques if the networks are very dense or the bands overlap one another.

5.6 THREE-DIMENSIONAL DISPLAYS

Three-dimensional perspective displays are one of the more common techniques used to present gridded surface data. The example shown in Fig. 59 represents the topography in a 40,000-mile^2 area around Idaho Falls. The user selects an observer position including the height above ground level and an appropriate vertical exaggeration to highlight the peaks and valleys. The computer projects the transformed image onto a two-dimensional plane representing the eye of the observer. A moving horizon is computed from front to back to determine those areas which are hidden from view and should not be plotted. These displays can be used with other types of data, as shown in Fig. 60, for population densities in the Northeastern United States. The peaks correspond to urban areas viewed from the southeast, the highest peak being New York City. The two most common transformations for these types of displays are perspective and isometric projections. With an isometric projection,[14] there is no decrease in the apparent size of features at farther distances from the projection plane. While this may cause some difficulty in visualizing spatial relationships between features, relative distances and sizes can be measured on an isometric image. Of course, features in a perspective projection are reduced in size as the distance increases from the projection plane. Some experiments have been performed at ORNL in the past with panoramic displays in which the 3-D information is projected onto the inside of an upright cylinder rather than a plane. The cylinder can be unwrapped, so to speak, and laid out flat to make a two-dimensional image. This technique would normally be associated with wide-angle data where the observer is very close to the viewing scene.

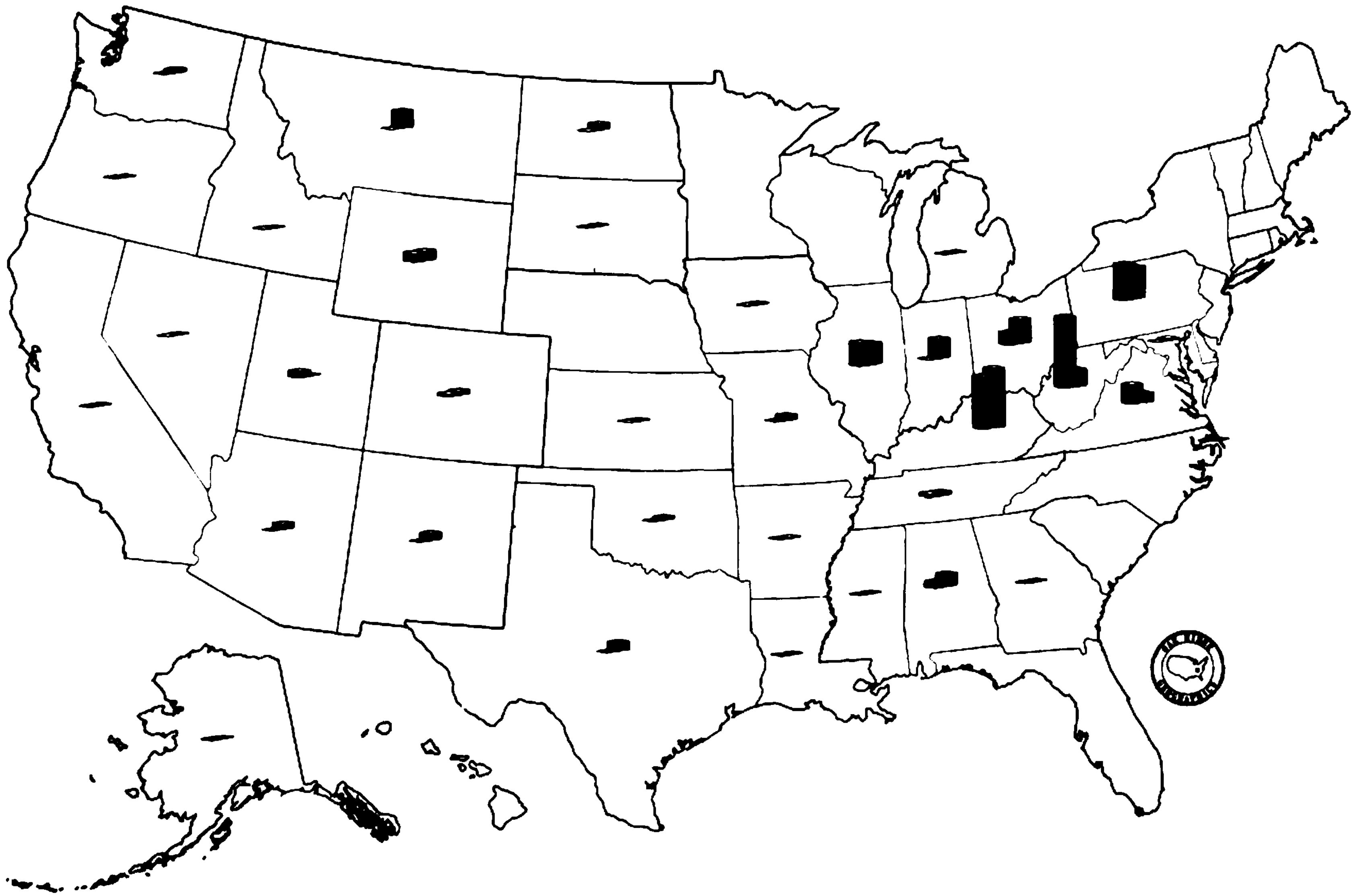
ORNL-DWG 79-12[illegible]

1975 STATE COAL PRODUCTION IN QUADS OF BTU'S

LEFT BLOCK: DEEP MINED – RIGHT BLOCK: SURFACE MINED

ONE INCH HEIGHT EQUALS 12,000,000 QUADS

NATIONAL ABANDONED MINE LANDS INVENTORY



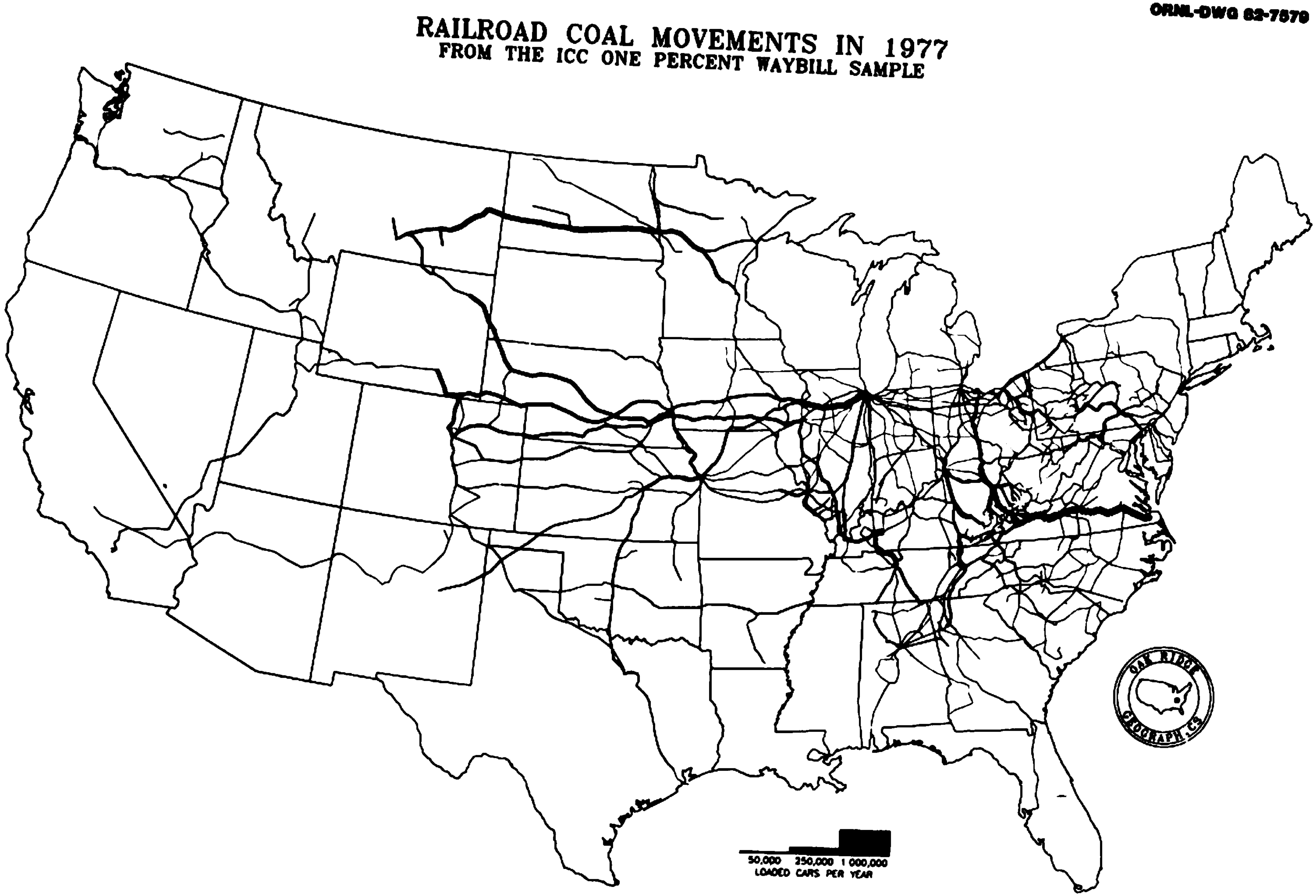


Fig. 58. Coal Flows by Railroad in 1977.

ORNL-DWG 81-3661

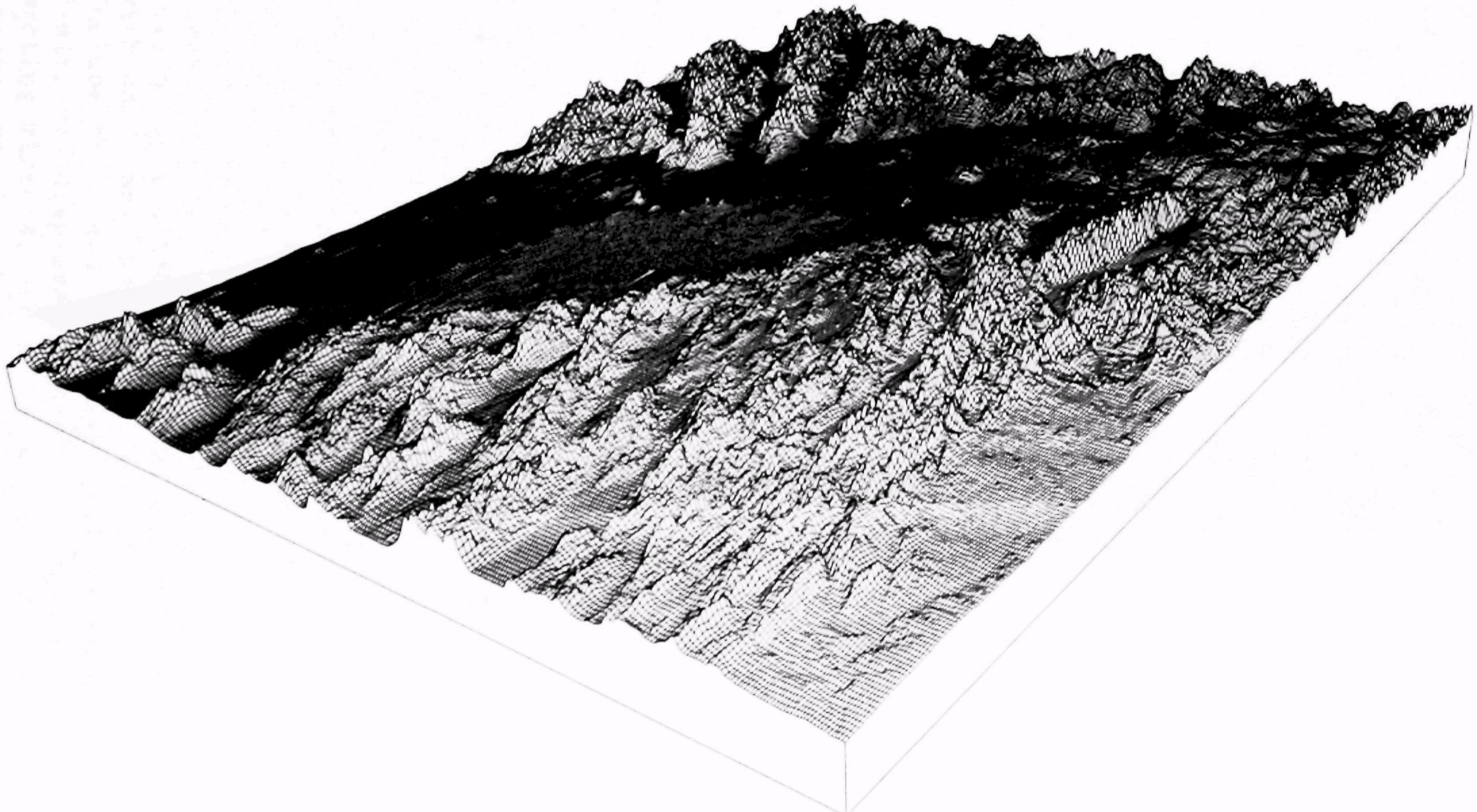


Fig. 59. Perspective Display of Topography for 200 x 200 Mile Area Around Idaho Falls.

ORNL-DWG 83-7743

POPULATION DENSITY (PEOPLE/SQ. MILE)
NORTHEAST U. S. (PJM REGION)

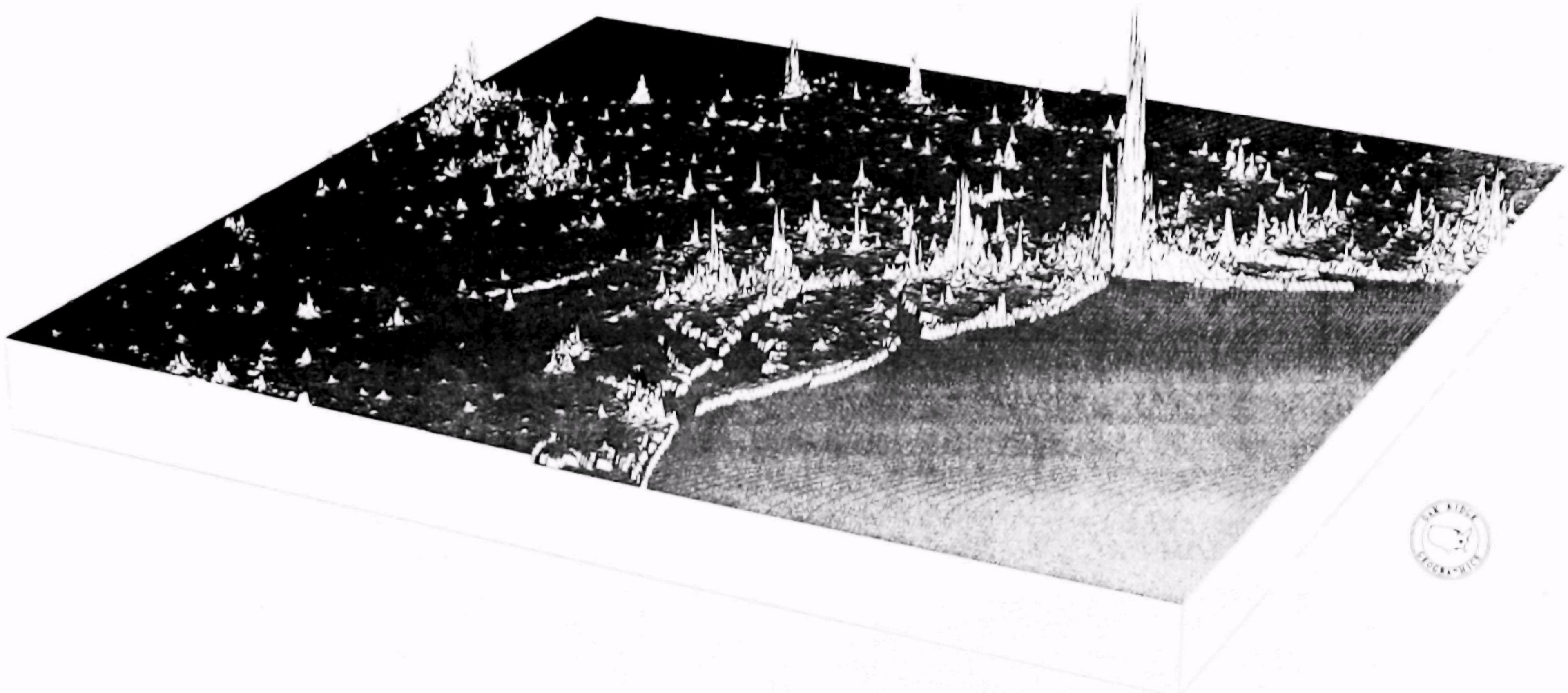


Fig. 60. Perspective Display of Population Density in the Northeastern United States Viewed From the Southeast.

To provide more information on a 3-D display, other features can be superimposed on the surface. For example, in Fig. 61 a three-dimensional surface has been created by interpolation from point monitoring data whose location is indicated by an X. The data value at each point is also superimposed on the surface. The elevation (or Z Axis) corresponds, not to terrain, but to concentrations of ^{131}I measured in milk samples around the United States following the Chernobyl nuclear accident. Density patterns proportional to the concentrations have also been overlayed on the surface to indicate regional distributions across the country. State boundaries are also included. With this graphic approach, it is easy to see the relative difference in peaks and valleys, especially in the Northwest where the prevailing winds resulted in the highest concentrations. This gradual variation in density patterns (referred to as a "fuzzy contour" approach) portrays trends without implying a high degree of resolution computed from sparse data points. The title, legend, and annotation are also plotted in 3-D to enhance the perspective orientation of the display. Color patterns portray the information dramatically. To map all the 3-D features, elevations must be computed for every line segment and density pattern plotted. The hidden-line algorithm must also take into account these additional surface features as they are plotted on top of the grid mesh. Several transformations are required among different coordinate systems (digitizer coordinates, latitude-longitude, grid mesh coordinates, and plotter raster coordinates). In this technique, the grid mesh features are eventually converted into an array of raster pixels before plotting. The development of efficient techniques for simultaneous vector-raster processing represents a newer area of computer graphics.

Another variation of superimposing ancillary data on top of a 3-D surface, is called the "melted-cheese on chicken-wire" problem. The technique is intended to build a 3-D perspective display of the terrain with a raster image (e.g., a scanned aerial photograph) superimposed on the ground surface. The raster image would contain color intensities which, if properly colored on the terrain, would represent land cover features similar to what would be seen from a normal photograph of the landscape. The "chicken-wire" corresponds to the mesh of grid lines representing the topography, and the "melted-cheese" corresponds to the raster image which has to be warped or fitted onto the terrain surface. The computer algorithms for handling this problem are different from superimposing linear features on the topography since pixel color intensities and perhaps sun angle calculations are involved. There are much larger amounts of data to process. Approaches for developing this technique have been investigated, and partial implementations have been completed at ORNL. For example, shaded land cover classes from photo-interpreted aerial photography have been superimposed on 3-D terrain models.

Another way to present digital three-dimensional information is through the use of stereo techniques. A variety of methods have been used, some with hard copy images, but more with CRT monitors. The techniques include the use of red-green glasses, polaroid filters and lenses, half-mirrored displays, flickering images with synchronized glasses, vibrating mirrors, and stereoscopes. Most of these techniques create two images, one for the left eye and one for the right eye, and

Fig. 61. Superimposition of Contours and Graphic Features on a 3-D Perspective Model of ^{131}I Milk Sample Concentrations Across the United States Following the Chernobyl Accident.

use some optical-mechanical system to present each image to the appropriate eye simultaneously. Red-green glasses have been used with a color CRT imaging system at ORNL, where the left-eye image is placed in one memory (green) and the right-eye image is placed in another memory (red). Viewing the two images superimposed simultaneously with red-green glasses presents a very nice stereo display. With true 3-D data it is possible for the information to be viewed from different locations or for these data to be rotated. The viewer perceives motion taking place in 3-D space. However, there are significant amounts of computations involved. Hard-copy output in the form of color plots or photographs can be produced and viewed in stereo with the glasses. The image flickering technique is slightly different and does not allow for hard copy representation. The CRT flickers between the two images very quickly so that, even with the naked eye, a stereo effect is perceived because the brain tends to integrate the two images. Glasses with special shutters synchronized to the flicker rate would make the perception even better. Each eye would see only one image at such a fast rate that it appears to be continuous stereo.

The vibrating mirror approach[15] projects an image from a CRT monitor onto a special mirror that dynamically changes the distance between the viewing plane and the observer's eye very quickly so that a true 3-D image is perceived. Figure 62 shows the components inside the viewing station with the monitor mounted vertically so that it projects down onto the vibrating mirror. The monitor is driven by a high-speed image processing system that continually sends display data to the screen. For explanation purposes, the three-dimensional information may be thought of as being divided up into very small slices (e.g., 30 to 512 slices, depending on the application); and each slice is presented very quickly on the CRT one after the other. A mirror (silvered mylar) is positioned in front of the CRT and vibrates in synchronization with the presentation rate of the slices. As the viewer looks at the mirror, it moves farther and farther away as slices deeper and deeper in the data are presented. The cycle is repeated continuously with such a high frequency that the multiple images are seen as one true 3-D object that can be viewed from different positions as the viewer moves his head. In effect, the mirror vibrates so as to change the focal length continuously between the CRT and the mylar. The change is synchronized with the depth of the specimen, and the brain integrates the multiple images so that a "hologram-type" effect is achieved. Three-dimensional cursors can be used to move interactively around the data. The user may also think of his/her data as x-y-z information which can be displayed and analyzed in 3-D space.

One factor that should be considered in all these methods is whether hard copy is required and whether color images are necessary. For example, in the red-green technique, color variations in the images cannot be used because the two images themselves are composed of red and green intensities. Stereo photographs can be made from the vibrating mirror image, but a stereoscope is needed for viewing the photographs. (Newer developments in 3-D display, such as laser technology and holograms, are not discussed in this paper.)

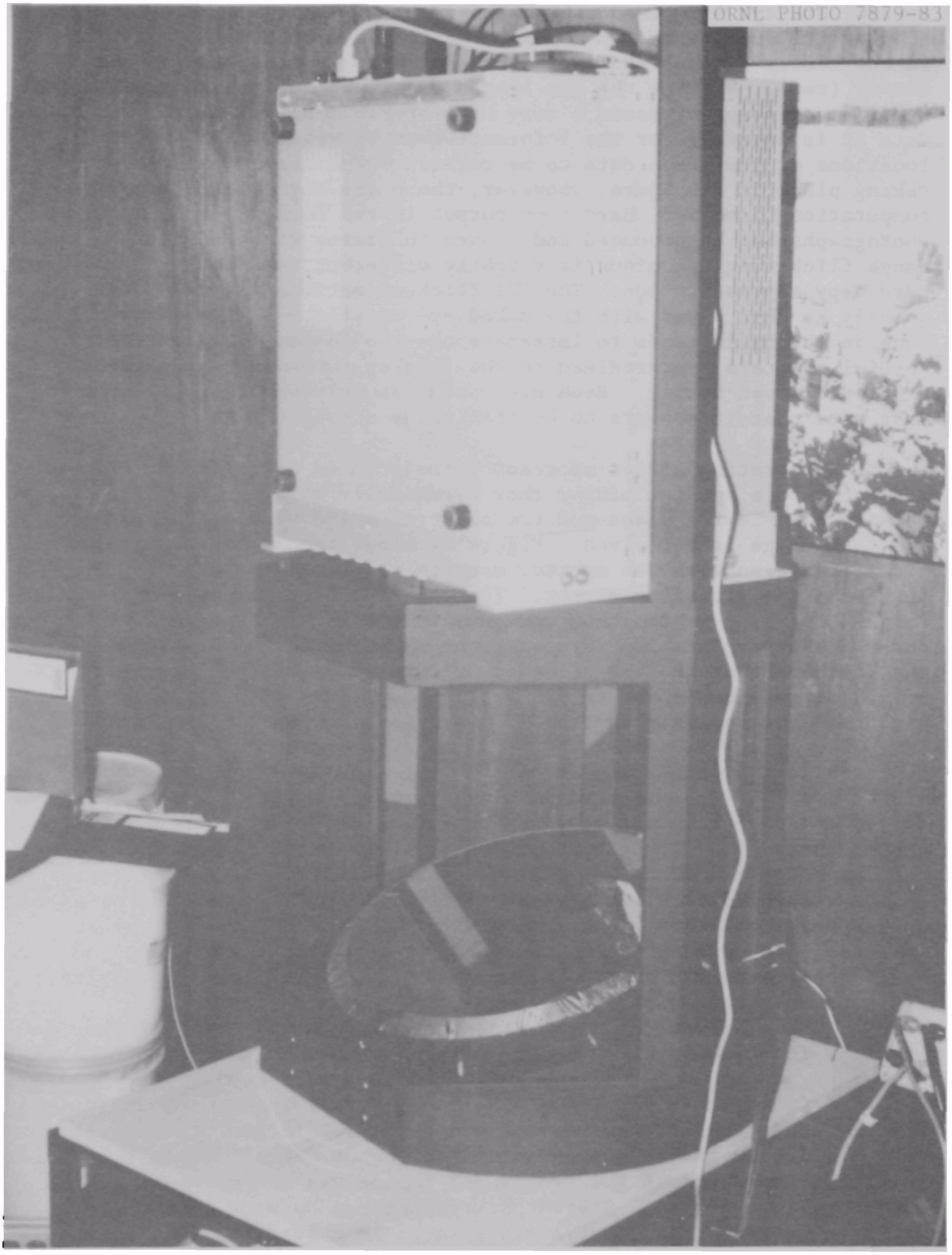


Fig. 62. Components of a Three-Dimensional Display Station Utilizing a Vibrating Mirror System.

5.7 SUBSURFACE MAPPING

In a previous subsection on geologic modeling, a brief description was given on the calculation of subsurface structures from borehole data. Figure 63 shows an example display of borehole data identifying the layers, their geologic age, and possible thickness variations. By using groups of boreholes it is possible to calculate cross-section profiles of geologic layers and display them with different shading patterns for each of the layers. This is shown in Fig. 64 with the ground surface as the top layer. The software used in these two displays were initially developed by L. W. Cobb and R. G. Mashburn, respectively, in conjunction with early waste repository studies at ORNL. By combining 3-D perspective displays of the earth's surface with cross-section cutaways, it is possible to simultaneously display information on the surface of the ground as well as the subsurface structure. Only preliminary planning steps have been carried out at ORNL for computer implementation of this display technique.

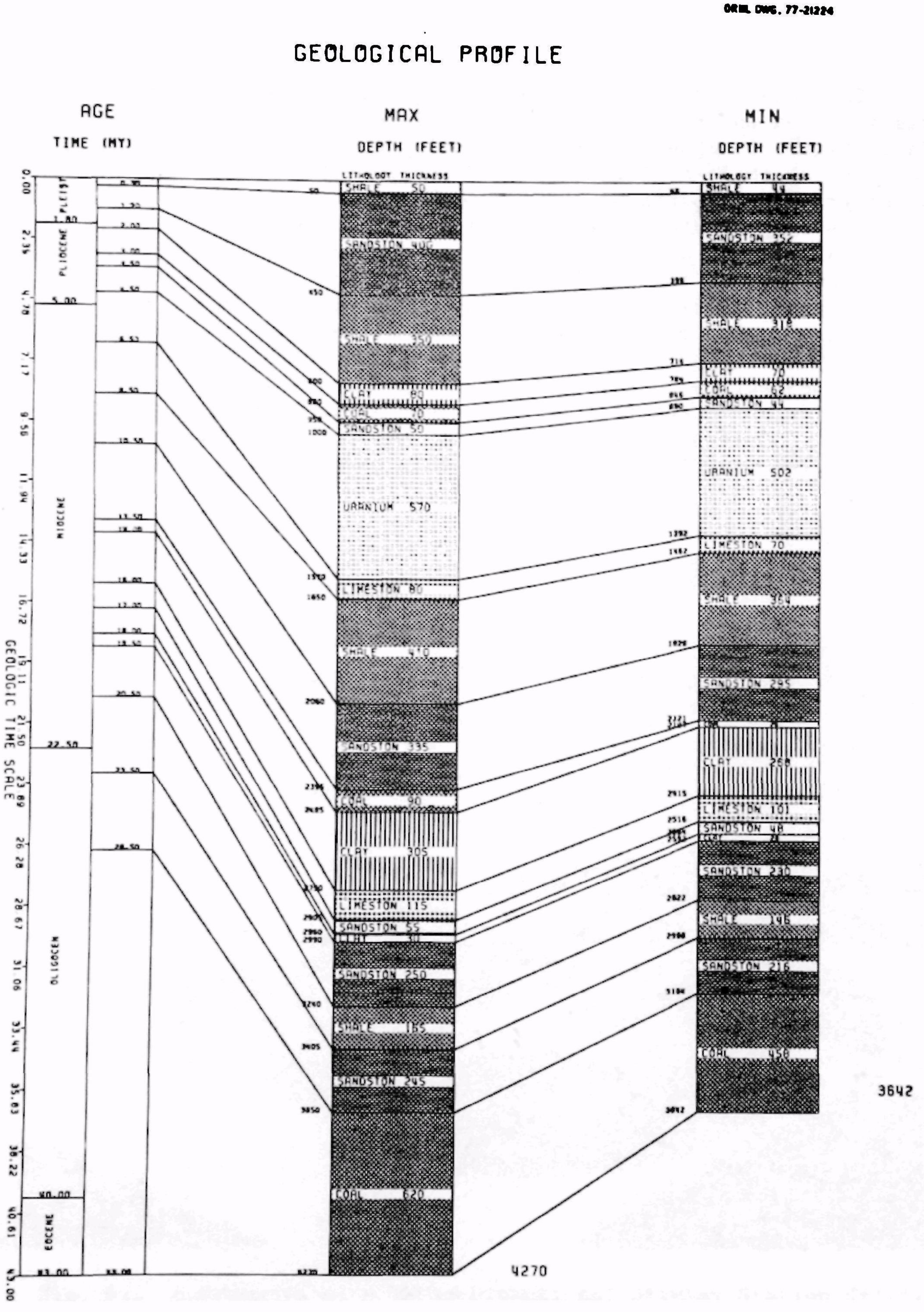


Fig. 63. Geological Profile of a Borehole Showing Minimum and Maximum Thickness Variations.

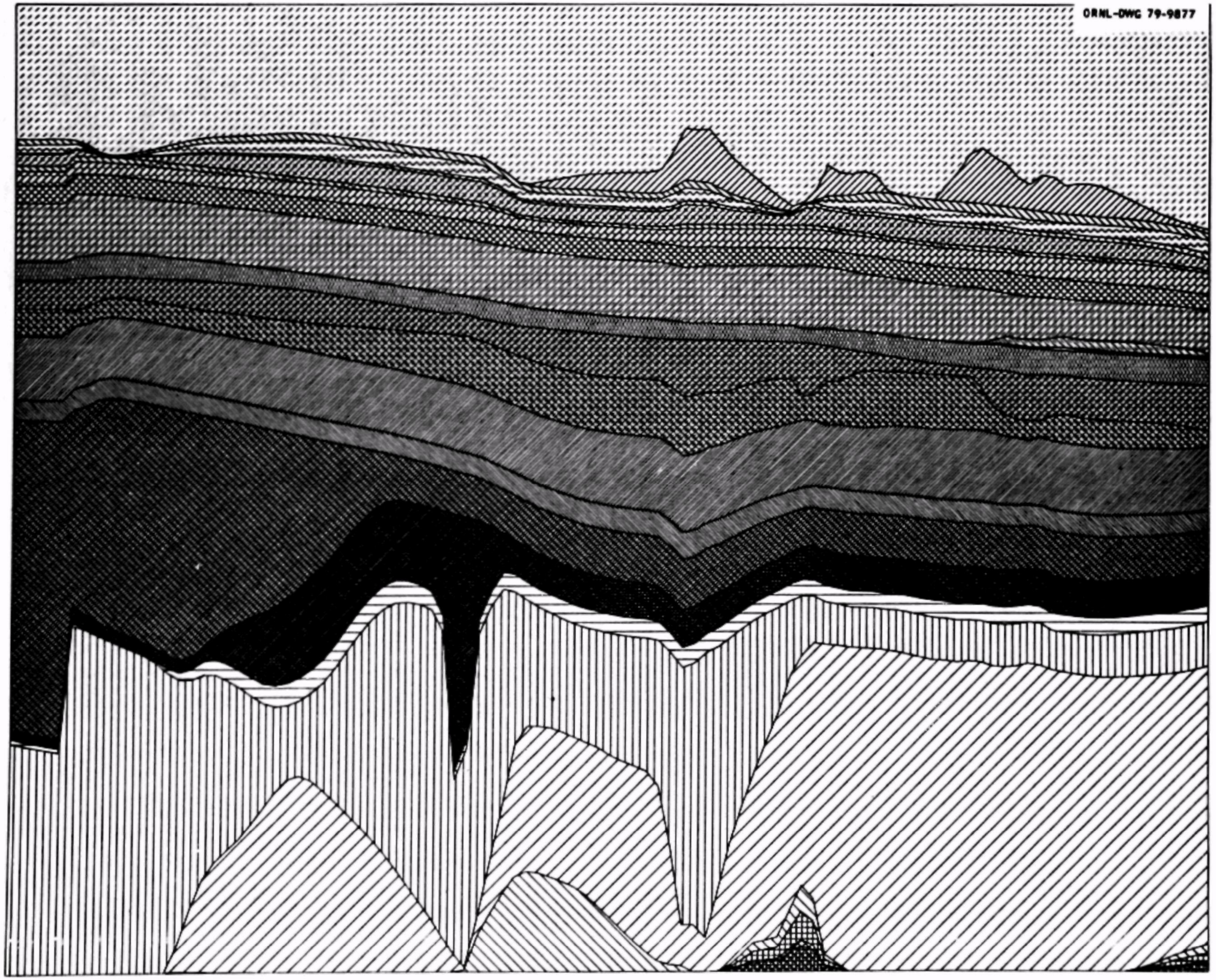


Fig. 64. Subsurface Geologic Profile Layers Calculated From Boreholes.

6. CONCLUSIONS AND OBSERVATIONS

From the analyst's standpoint, geographical information is far more complex than other forms of tabular data. In the first place, the quantity of data is proportional to the number and type of data structures used to represent the distribution of each phenomenon on a continuous earth surface. Additional data are required to locate each point in a spatial coordinate system (e.g., latitude/longitude), and in many cases temporal data are needed as well. If the data are obtained from different maps and other data sources, they cannot be integrated without careful transformation and conversion to account for differences in scale, projection, resolution, data structure, and coordinate systems. Traditionally, maps have been used as the primary mechanism for displaying geographical information because the human brain can visually assimilate these huge volumes of data more easily than it can comprehend the columns of numbers that each image represents. In the past the potential for geographical analysis has been severely constrained by the cartographer's limited capacity for manual computation of complex data conversions. Geographic information and analysis technologies have advanced rapidly and now constitute one of the most remarkable achievements of the "Information Revolution." It is now possible to perform geographic analyses of complex problems at high resolution for very large areas, to accelerate all types of geographic studies (large and small) and to address many problems in basic and applied research that could not have been addressed before.

Over the last several years, geographic systems have become readily available either commercially or from the public sector; and a variety of cartographic data bases are now provided by government agencies (e.g., the United States Geological Survey National Cartographic Information Center). There have been several national committees functioning to coordinate and establish standards for geographic data (e.g., the Federal Interagency Coordinating Committee on Digital Cartography). Graphic hardware systems are springing up at a very rapid pace (although most are not tailored to geographic analysis). These new capabilities have given rise to a variety of automated technologies ranging from computer cartography and graphics to geographic information systems, spatial modeling, digital remote sensing, and image processing. Users are becoming aware of these recent advances in spatial data processing and are attempting to acquire high-technology capabilities to meet their own environmental and geographic-oriented needs. In some cases there is a tendency for users to focus on only one or two aspects associated with the application of geographic information systems. Some may be interested primarily in the data bases themselves, rather than how they are to be used in solving real-world problems; others may focus on the sophisticated mapping output rather than the modeling techniques that calculate the results; a few are caught up in the software algorithms and efficiency of processing; and it is always easy to be enthralled with the graphic hardware used to display and interact with data. All of these components are important and experience gained over the last decade at ORNL has shown that a balanced understanding and integration of each aspect is required to successfully apply the technology.

The intent of this paper was to introduce some of the concepts and methods associated with processing geographic data for regional analyses. The primary focus has been upon describing the data structures for representing geographic features, techniques for digitizing data, various types of spatial transformations and analyses, geographic modeling, and graphic display in map form. Results from a variety of real-world applications have been shown to aid in describing the techniques. The selection of data structures and spatial transformations to solve a particular geographic application are very important. Those decisions will determine to a large extent the cost of digitizing data, performing the analyses, and the types of graphic displays available. They will also affect the resolution of the data captured and the scales at which it can be used. The investigators must not only be familiar with the data elements used in their analysis, but they must be aware of what various thematic transformations on a spatial basis can do to their data. For example, the conversion of county-level data to watersheds, based on proportionate area techniques, can sometimes distort the distributions so that erroneous conclusions are made on a hydrologic basis.

To help the planner who is beginning to use automated systems for geographic applications, a few general observations are given below. These have resulted from experiences in carrying out regional analyses over a number of years. However, they are not hard and fast conclusions, and in some cases may not even be characteristic of certain types of problems. They are just considerations to be kept in mind during the planning phases. These are given in a random fashion: It is generally better to digitize raw information rather than interpreted data whenever possible; segmental structures will better represent the original map data than grid systems; surrogates (e.g., existing or less costly data from which the desired information can be approximated) or sampling procedures can sometimes be used to reduce the cost and time of digitizing large data bases; some approaches digitize raw information in segmental form and convert to grid cells for analysis, thus avoiding the sliver problem of polygon overlay techniques; statistical techniques and models are many times associated with the thematic data, whereas spatial transformations generally involve the cartographic data; simultaneous vector-raster processing is becoming more common and efficient; although it is best to let the problem define the appropriate data and techniques to be used, many times existing data bases determine what is feasible and can be tested; any sizable amount of digital geographic data should be referenced to some standard earth coordinate system; the use of standard information systems for storing geographic data is not necessarily appropriate for the spatial or cartographic part because of the unique structures and processing requirements; no single geographic information system is suitable for handling all types of geographic problems; display techniques are becoming more dependent on the types of graphic hardware devices used; newer techniques including the use of sorting procedures and local processing, make mini-computers more feasible for geographic processing although very large data bases may still need to reside on mainframe computers with subsets transferred for mini-computer processing; small applications can actually be carried out on microcomputerswith medium-and-sometimes high-resolution displays; certain types of spatial processing, especially with three-dimensional data, can utilize the power of Class VI super-computers; the information content in a large volume of tabular output can sometimes be presented in a few graphic images; software development is generally several years behind the hardware development; it is important for new software to be programmed in machine independent form as much as possible to take advantage of new hardware systems; etc.

It should also be pointed out that large data bases of general utility require continued maintenance and updating, as well as processing improvements and the addition of new data. Simply storing data in machine readable form (tapes, disks, etc.) represents a recurring cost, especially if the data are not allowed to deteriorate over long time periods. Many sponsors are unaware of these needs and costs unless they are encouraged to become involved at a working level. Many times geographic data are digitized or collected based on just mapping considerations, especially by cartographers. However, these data bases represent important resources for spatial analyses and modeling efforts, but significant amounts of work must be done to convert the information into suitable analysis structures. This work could be incorporated during the digitizing effort and might not add much to the cost.

As was mentioned in the Introduction, there are five primary components essential to carrying out regional analyses: (1) data bases and data resources, (2) the hardware systems, (3) the software packages, (4) the analysis models, and (5) the resource staff and personnel. New users who are interested in acquiring in-house capabilities and who have had limited experience with geographic information systems may tend to focus on the hardware and equipment, especially when their planning is based solely on reviews of vendor systems. However, their planning should include appropriate reviews of their own information needs and requirements. These needs can then be evaluated along with the technologies available to develop a successful strategy for acquiring and using automated approaches in solving geographic problems. With the proliferation of microcomputers it is possible to acquire small geographic systems initially and conduct prototype applications and testing. This provides excellent hands-on experience before investing thousands of dollars in the large systems that may be required to carry-out the mission of major institutions.

The analyst who has access to these capabilities should be aware of other pitfalls in beginning geographic-oriented projects, especially with outside sponsors who may themselves lack understanding as to what is required. It is difficult to comprehend the work required "behind-the-scenes" when the final results can be presented graphically on a computer plot or a CRT screen in a matter of minutes. The optimal situation is to work closely with the sponsors providing intermediate results throughout the project, suggesting alternatives when problems arise, and continually trying to gain a better understanding of their information needs. The sponsors will depend on the geographic system specialist to select appropriate algorithms and processing techniques, but they should become familiar with the approaches and the types of results that can be obtained. Hopefully this document provides a start at gaining such familiarity.

REFERENCES

1. R. G. Edwards, MAPPROJ2 - FORTRAN Map Projection Subroutines - Version 2, ORNL/CSD/TM-10, Union Carbide Corp., Nuclear Div., Oak Ridge National Lab., August 1976.

2. R. C. Durfee and P. R. Coleman, Population Distribution Analysis for Nuclear Power Plant Siting, NUREG/CR-3056, ORNL/CSD/TM-197, Union Carbide Corp., Nuclear Div., Oak Ridge National Lab., December 1983.

3. R. G. Edwards and P. R. Coleman, IUCALC - a Fortran Subroutine for Calculating Polygon-Line Intersections, and Polygon-Polygon Intersections, Unions, and Relative Differences, ORNL/CSD/TM-12, Union Carbide Corp, Nuclear Div., Oak Ridge National Lab., August 1976.

4. R. G. Edwards, An Algorithm to Calculate Proportional Area Transformation Factors for Digital Geographic Data Bases, Spatially-Oriented Referencing Systems Association, SORSA-83 Symposium, University of Maryland, October 1983.

5. R. C. Durfee, R. G. Edwards, M. J. Ketelle, and R. B. Honea, "Assignment of ERTS and Topographical Data to Geodetic Grids for Environmental Analysis of Contour Strip Mining", International Symposium on Landuse, Phoenix, Arizona, October, 1975, American Society of Photogrammetry, Washington, D. C., February 1976.

6. J. S. Jalbert and A. D. Shepherd,, A System for Regional Analysis of Water Availability, ORNL/NUREG/TM-82, Oak Ridge National Lab., July 1977.

7. A. D. Shepherd, A Spatial Analysis Method of Assessing Water Supply and Demand Applied to Energy Development in the Ohio River Basin, ORNL/TM-6375, Union Carbide Corp., Nuclear Div., Oak Ridge National Lab., August 1979.

8. Michael A. Domaratz, Cheryl A. Hallam, Warren E. Schmidt, Hugh W. Calkins, "Digital Line Graphs From 1:2,000,000-Scale Maps", USGS Circular 895-D, U.S. Department Interior, National Cartographic Information Center, U.S. Geological Survey, Reston, Va., 1983.

9. Jerome E. Dobson and Alf D. Shepherd, Water Availability for Energy in 1985 and 1990, ORNL/TM-6777, Union Carbide Corp., Nuclear Div., Oak Ridge National Lab., October, 1979.

10. D. S. Joy and P. E. Johnson, HIGHWAY, A Transportation Routing Model: Program Description and Revised Users' Manual, ORNL/TM-8759, Union Carbide Corp., Nuclear Div., Oak Ridge National Lab., September 1983.

11. D. S. Joy and P. E. Johnson, Airport Locator Program: Description, ORNL/TM-8610, Union Carbide Corp., Nuclear Div., Oak Ridge National Lab., April 1982.

12. R. B. Honea, C. H. Petrich, D. L. Wilson, C. A. Dillard, R. C. Durfee, J. A. Faber, Computer Software to Calculate and Map Geologic Parameters Required In Estimating Coal Production Costs, EPRI EA-674, Electric Power Research Institute, April 1979.

13. Richard E. Groop and Paul Smith, "A DOT Matrix Method of Portraying Continuous Statistical Surfaces", The American Cartographer, 9, NO. 2 1982.

14. T. C. Tucker, Catch: Computer Assisted Topography Cartography, and Hyposgraphy, Part III, PERSPX: A subroutine Package for Perspective and Isometric Drawings, ORNL-TM-3790, Union Carbide Corp., Nuclear Div., Oak Ridge National Lab., February 1973.

15. Henry Fuchs, et al., "Design of an Image Editing With a Space-Filling Three-Dimensional Display Based on a Standard Raster Graphics System, Society of Photo-Optical Instrumentation Engineers, 367, (1982).

APPENDIX A - EXAMPLES OF SOFTWARE FUNCTIONS FOR A GEOGRAPHIC INFORMATION AND ANALYSIS SYSTEM

USER-FRIENDLY COMMUNICATIONS INTERFACE

- Menus
- Windows
- Command Structures
- Help Files
- Control Languages
- etc.

SYSTEM SUPERVISOR AND MANAGER

- Task Allocation and Control
- Memory Management
- Parameter Transfer
- Interrupt Processing
- File Control and Transfer
- Root Segment Control
- etc.

DATA ACQUISTION AND CONVERSION

- Reformatting External Files
- Map Projections/Coordinate Systems
- Scaling, Rotation, Translation
- Rectification
- Resolution Determination
- Filtering
- Digitizing
    - Manual Grid Overlay
    - X,Y Tablet
    - Raster Scanning
    - Add Identifiers
    - Topology Assignment
        - Area
        - Network
    - Attribute Assignment
- Editing
    - Chain Editing
    - Addition
    - Replacement
    - Modification
    - Topological Error Detection
    - Repositioning
    - Edge Matching
- Image Processing
    - Geometric Rectification
    - Radiometric Modification
    - Image Enhancement
    - Classification and Interpretation
    - Pattern Recognition
    - Standard Functions (Zoom, Pan, Blotch, Statistics, etc.)

DATA TRANSFORMATION AND INTEGRATION

- Grid-to-Polygon (or Chain)
- Polygon-to-Grid
- Point-in-Polygon
- Polygon Intersection
- Grid-to-Grid
- Interpolation
- Extrapolation
- Spline Generation
- Thiessen Polygon Construction
- Dime Vector to Chain to Polygon
- Boolean Operations
- Arithmetic Operations
- Filtering, Generalizing, and Smoothing
- Aggregation/Disaggregation

DATA BASE MANAGEMENT

- Structure
  - Flat File
  - Relational
  - Hierarchical
  - Network
- Spatial Data Processing
- Attribute Data Processing
- Data Loading
- Storage/Retrieval
- Editing and Updating
- Concatenating and Merging
- Query
- Record or Key Searching
- Sorting
- Backup and Copy
- Renaming
- Listing and Report Generation
- Record and File Summaries
- Catalogs and Directories
- Protection and Security
- Activity Logs
- Utilities and Maintenance
- Linkability to External Systems and Models

DATA ANALYSIS, STATISTICS, AND MODELING

- Mensuration
  - Linear Distance
  - Area
  - Perimeter
- Centroid
- Direction
- Proximity Calculations

- Categorization
- Class Intervals
- Ranking
- Statistics
  - Mean
  - Mode
  - Median
  - Standard Deviation
  - Correlation
  - Spatial Autocorrelation
  - Regression
  - Minimum Agtregate Travel
  - Cluster Analysis
  - Factor Analysis
  - Frequency Distribution
- Modeling
  - Spatial Index Computation
  - Screening Models
  - Terrain Models
    - Slope
    - Aspect
    - Drainage Patterns
    - Viewshed
  - Pattern Recognition
  - Network Flow
  - Corridor Analysis
  - Routing and Shortest Path
  - Linear Programming
  - Gravity Models
  - Diffusion Models

GRAPHIC OUTPUT AND DISPLAY

- Polygon/Segmental Mapping
- Contouring
- 3-D Perspective and Isometric
- 3-D Imaging
- Grid Cell Mapping
  - Cartesian
  - Raster
  - Polar Coordinates
- Graduated Circles
- Pie Charts
- Flow Charts
- Graphs
- Line Symbolism
- Graphic Overlay
  - 2-D Overlay
  - 3-D Overlay
- Mapping Vertical Data Samples and Strata
- Mosaic
- Legends
- Labels
- Titles

Text and Annotation
Scaling
Windowing
Zoom or Magnify
Pan
Rotate
Polygon Shading
Hashing
Grey Level or Density
Color Patterns
Histograms
Bar Charts

INTERNAL DISTRIBUTION

1. J. B. Berry
2. R. B. Braid
3. E. A. Bright
4. S. A. Carnes
5. C. W. Chester
6. S. M. Chin
7. P. R. Coleman
8. R. B. Craig
9. A. G. Croff
10. R. M. Cushman
11. K. L. Daniels
12. R. M. Davis
13. J. E. Dobson
14-18. R. C. Durfee
19. R. G. Edwards
20. L. D. Eyman
21. M. P. Farrell
22. W. Fulkerson
23. D. L. Greene
24. B. H. Holt
25. S. G. Hildebrand
26. E. L. Hillsman
27. R. B. Honea
28. D. B. Hunsaker
29. P. J. Johnson
30. D. S. Joy
31. P. Kanciruk
32. F. C. Kornegay
33. F. E. Latham
34. R P. Leinius/H.P. Carter
35. A. S. Loebl
36. S. M. Margle
37. R. C. Martin
38. R. A. McCord
39. G. J. Morris
40. T. W. Oakes
41. R. J. Olson
42. D. C. Parzyck
43. B. E. Peterson
44. C. H. Petrich
45. D. H. Pike
46. D. E. Reichle
47. C. R. Richmond
48. A. L. Rivera
49. P. S Rohwer
50. R. D. Roop
51. M. W. Rosenthal
52. D. W. Rosowitz
53. T. H. Row
54. R. M. Rush
55. M. J. Sale
56. L. B. Shappert
57. D. S. Shriner
58. E. J. Soderstrom
59. E. P. Tinnel
60. L. E. Till
61. R. S. Turner
62. R. I. Van Hook
63. A. H. Volker
64. L. D. Voorhees
65. D. P. Vogt
66. G. W. Westley
67. G. E. Whitesides
68. D. L. Wilson
69. T. J. Wilbanks
70. B. C. Zygmunt
71-94. Geog. Data Sys. Sec.
95. Central Research Lib.
96. Doc. Ref. Sec. Y-12
97-98. Laboratory Records
99. Lab. Records -RC
100. ORNL Patent Office

EXTERNAL DISTRIBUTION

101. Dr. Ronald F. Abler, Director, Geography and Regional Science Program, National Science Foundation, Washington, DC 20550

102. Captain W. David Alley, United States Air Force, School of Civil Engineering, Air Force Institute of Technology, Wright-Patterson AFB, OH 45433

103. Dr. K. Eric Anderson, U. S. Geological Survey, Mail Stop 521, 12201 Sunrise Valley Drive, Reston, Virginia 22092.

104. Dr. F. P. Baxter, Tennessee Valley Authority, Ridgeway Road, Norris, Tennessee 37828.

105. Dr. Thomas L. Bell, Professor and Assistant Dean for Research, Department of Geography, University of Tennessee, Knoxville, Tennessee 37916.

106. Dr. David A. Bennett, Acid Deposition Research Staff, U.S. Environmental Protection Agency, 401 M Street, SW, RD-676, Washington, DC 20460.

107. Dr. Ralph Bernstein, IBM Scientific Center, 1530 Page Mill Road, Palo Alto, California 94304.

108. Mr. Dwight Briggs, U.S. Department of Transportation, Federal Highway Administration, 400 7th Street, SW, Washington, DC 20590.

109. Mr. F. R. Broome, U.S. Bureau of Census, Federal Office Building, 4, Room 1217, Washington, DC 20233.

110. Mr. Paul E. Bryant, National Preparedness Programs, Federal Emergency Management Agency, 500 C Street, SW, Washington, DC 20472.

111. Professor Hugh Calkins, 310 Oakbrook Drive, Williamsville, New York 14221.

112. Dr. James R. Carter, Associate Professor, Department of Geography, University of Tennessee, Knoxville, Tennessee 37916.

113. Mr. Robert L. Chartrand, Senior Specialist in Information Policy and Technology, Congressional Research Service, The Library of Congress, Madison Office Building, Washington, DC 20540.

114. Ms. L. B. Cobb, 27 E. Shady Lane, Houston, Texas 77063.

115. Mr. Jack Dangermond, Environmental Systems Research Institute, 380 New York Street, Redlands, California 92373.

116. Dr. George J. Demko, The Geographer, Department or State, Washington, DC, 20520.

117. Mr. George Detrich, Bettis Atomic Power Laboratory, P. O. Box 79, West Miflin, Pennsylvania 15122.

118. Dr. K. Dueker, School of Urban Affairs, Portland State University, P. O. Box 751, Portland, Oregon 97297.

119. Ms. Dolores Eddy, Librarian, U.S. Environmental Protection Agency, Region VIII, 999-18th Street, Suite 500, Denver, CO 80202-2405

120. Dr. Robert L. Edwards, National Marine Fishery Service, National Oceanic and Atmospheric Administration, Northeast Fisheries Center, Woods Hole, Massachusetts 02543.

121. Dr. Jack Estes, Department of Geography, University of California, Santa Barbara, Santa Barbara, California 93106.

122. Mr. Robin G. Fegeas, U. S. Geological Survey, Geographic Applications Program, National Center, Mail Stop 710, Reston, Virginia 22092.

123. Mr. Timothy Foresman, U. S. Environmental Protection Agency, EMSL-LV, AMS, P. O. Box 15027, Las Vegas, Nevada 89114.

124. Captain Dennis J. Foth, United States Air Force HQ AFESC/SIE, Tyndall AFB, FL 32403-6001

125. Mr. Terry Gossard, Engineering Staff Unit, U. S. Forest Service, U. S. Department of Agriculture, P. O. Box 2417, Washington, DC 20013.

126. Mr. Martin Gottlieb, City of New York, Department of Ports, International Trade & Commerce, Battery Maritime Building, New York, NY 10004

127. Ms. Gloria Hagge, U.S. Air Force, HQ SAC/DEPV, HQ SAC/DEPV, Offutt AFB, NE 6813-50

128. Mr. David A. Henney, Chief, Air Base Planning, HQ SAC/DEPVA, Bldg. 500, Room 3F7, Offutt AFB, NE 68113

129. Mr. Jim Hines, 315 Scenic Drive, Kingston, TN 37763

130. Dr. Elizabeth R. Holbrook, 304 Lido Cove, Niceville, FL 32578

131. Mr. Thomas R. Horton, Environmental Protection Agency, Eastern Environmental Radiation Facility, P. O. Box 3009, Montgomery, Alabama 36193.

132. Dr. Paul G. Huray, Associate Dean for Research and Resources Development, College of Liberal Arts, University of Tennessee, Knoxville, Tennessee 37916.

133. Mr. Robert H. Iveson, Program Manager, Electrical Systems Division, Electrical Power Research Institute, 3412 Hillview Avenue, Palo Alto, California 94303.

134. Mr. Doug Jansing, HQSAC DEPV, Building 500, U. S. Air Force, Offitt Air Force Base, Nebraska 68113.

135. Dr. Robert T. Jaske, Technical Hazards Division, Federal Emergency Management Agency, 500 C Street, SW, Washington, DC 20472.

136. Mr. Michael Kaltman, Nuclear Regulatory Commission, Phillips Building, 7920 Norfolk Avenue, Bethesda, Maryland 20016.

137. Mr. Jim Kelly, Department of Energy, Office of Environmental Assessment, V-22, Forrestal Building, 1000 Independence Avenue, Washington, DC 20585

138. Mr. Don Kestyn, U.S. Department of Transportation, Federal Highway Administration, Highway Statistics Division, HHP-42, 400 7th Street, SW, Washington, DC 20590.

139. Dr. Richard F. Kott, Suite 890, General Electric Company, 1331 Pennsylvania Avenue, N.W., Washington, DC 20004.

140. Dr. Tom Mace, Lockheed Engineering and Management Services, P. O. Box 15027, Las Vegas, Nevada 89114.

141. Dr. John L. Malanchuck, Office of ADIS, U. S. Environmental Protection Agency (RD-676), 401 M Street, SW, Washington, DC 20460.

142. Professor Duane Marble, Department of Geography, State University of New York, Buffalo, New York 14226.

143. Mr. C. R. Meyers, Office of Surface Mining, U. S. Department of Interior, South Interior Building, 1951 Constitution Avenue, Washington, DC 20240.

144. Dr. Harold Moellering, Numerical Cartography Laboratory, 158 Derby Hall, Ohio State University, 154 North Oval Mall, Columbus, Ohio 43210.

145. Dr. Richard L. Morrill, Professor, Department of Geography, University of Washington, Seattle, Washington 98195.

146. Dr. Joel Morrison, Senior Scientific Advisor, Geography, National Mapping Division, U. S. Geological Survey, Mail Stop 516, Reston, Virginia 22092.

147. Mr. W. Allen Nixon, U. S. Air Force, HQ AFESC/DEMB, Tyndall Air Force Base, Florida 32403.

148. Dr. Jim Omernick, Environmental Research Laboratory, Environmental Protection Agency, 200 Southwest 35th Street, Corvallis, Oregon 97333.

149. Dr. W. R. Ott, Siting and Environmental Branch, U. S. Nuclear Regulatory Commission, Willste Building, 7915 Eastern Avenue, Silver Springs, Maryland 20906.

150. Dr. Risa I. Palm, Associate Professor, Department of Geography, University of Coloado, Boulder, Colorado 80309.

151. Dr. Robert W. Peplies, Geography Department, East Tennessee State University, Johnson City, Tennessee 37601.

152. Mr. Frank R. Perchalski, Photogrammetry and Remote Sensing Section, Mapping Services Branch, 216 Haney Building, Tennessee Valley Authority, Chattanooga, Tennessee 37401.

153. Dr. Wes R. Pfarner, Censor Applications Division 0323, Sandia National Laboratories, P. O. Box 5800, Albuquerque, New Mexico 87185.

154. Dr. William Pillinger, Savannah River Laboratory, Building 773-42A, Aiken, South Carolina 29808.

155. Mr. Bill Regan, U. S. Nuclear Regulatory Commission, Phillips Building, 7920 Norfolk Avenue, Bethesda, Maryland 20016.

156. Dr. Craig N. Robertson, National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Northeast Fishery Center, Sandy Hook Laboratory, Highlands, New Jersey 07732.

157. Mr. Bruce Rowland, Division of Forestry, Fisheries and Wildlife, Tennessee Valley Authority, Norris, Tennessee 37828.

158. Mr. Richard H. Schultze, Trinity Consultants, Inc., 100 North Central Expressway, Suite 606, Richardson, Texas, 75080.

159. Dr. Ralph L. Scott, Chief, Environmental and Systems Branch, U.S. Department of Energy, Morgantown Environmental Technology Center, P. O. Box 880, Morgantown, West Virginia 26507.

160. Mr. Gary A. Shelton, National Aeronautics and Space Administration, Mail Stop 240-6, Ames Research Center, Moffett Field, California 94035.

161. Ms. Susan Sherwood, National Park Service - 424, Washington, DC 20240.

162. Dr. Charles W. Smart, Division of Forestry, Fisheries and Wildlife, Tennessee Valley Authority, Norris, Tennessee 37828.

163. Mr. J. L. Smyre, Route 1, Box 292, Catskill, New York 12414.

164. Dr. David Straight, Department of Computer Science, Ayres Hall, University of Tennessee, Knoxville, Tennessee 37916.

165. Mr. Paul Svercl, Highway Statistics, HHP-44, Federal Highway Administration, 400 7th Street, SW, Washington, DC 20590.

166. Mr. Gale W. TeSelle, Director, Cartographic and Geographic Information Systems Division, Soil Conservation Service, U.S. Department of Agriculture, P. O. Box 2890, Washington, DC 20013.

167. Dr. Waldo R. Tobler, Professor, Department of Geography, University of California, Santa Barbara, California 93106.

168. Dr. James P. Thomas, U. S. Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fishery Service, Washington, DC 20235.

169. Dr. Michael Thomason, Department of Computer Science, Ayres Hall, University of Tennessee, Knoxville, Tennessee 37916.

170. Mr. Al Voss, Photogrammetry and Remote Sensing Section, Mapping Services Branch, 200 Haney Building, Tennessee Valley Authority, Chattanooga, Tennessee 37401.

171. Dr. Richard E. Whitmer, Department of Interior, U. S. Geological Survey, Mail Stop 710, Reston, Virginia, 22092.

172. Mr. R. T. Williams, Department of Energy, Office of Environmental Policy, EV-22, Forrestal Building, 1000 Independence Avenue, Washington, DC 20585.

173. Division of Engineering Mathematics and Geosciences, Department of Energy, Washington, D. C. 20545.

174. Office of Assistant Manager for Energy Research and Development, DOE/ORO, Oak Ridge, Tennessee 37831.

175-184. Technical Information Center, Oak Ridge Operations, P.O. Box 62, Oak Ridge, Tennessee 37831

☆U.S. GOVERNMENT PRINTING OFFICE 1988—548-118/80006