centro de educación continua
división de estudios superiores
facultad de ingeniería, unam

CARTOGRAFIA AUTOMATIZADA APLICADA A LA PLANEACION

Tacuba 5, primer piso. México 1, D. F.
Teléfonos: 521-30-95 y 513-27-95

centro de educación continua
división de estudios superiores
facultad de ingeniería, unam

CURSO INTENSIVO SOBRE:

" CARTOGRAFIA AUTOMATIZADA APLICADA A LA PLANEACION"

PROGRAMA

LUNES 18

9 a 10 h:	Introducción al Curso. Presentación de profesores y participantes. Provisión del material necesario.
10 a 13 h:	La producción de gráficas mediante computadora y sus aplicaciones en planeación.
PROFESOR	ERIC TEICHOLZ.
13 a 14 h:	Comida.
14 a 17 h:	Cartografía automatizada. Procesos necesarios y consideraciones técnicas para su implementación.
PROFESOR	THOMAS K. PEUCKER.
17 a 18 h:	Coctel.

MARTES 19

9 a 10 h:	Estructuras de datos cartográficos.
10 a 11 h:	Digitalización.
11 a 13 h:	Ejemplos de sistemas de información geográficos.
PROFESOR	THOMAS K. PEUCKER.

MARTES 19

13 a 14 h: Comida.

14 a 18 h: Visita a Cetenal.
Demistración de equipo y trabajos realizados.

P R O F E S O R E S: ING. ENRIQUE SOTO
ACT. JOSE OLIVERES.

MIERCOLES 20

9 a 11 h: Mapas interactivos.

P R O F E S O R E S: ERIC TEICHOLZ
THOMAS K. PEUKER.

11 a 13 h: Modelos digitales de Terrenos.

P R O F E S O R THOMAS K. PEUKER.

13 a 14 h: Comida.

14 a 18 h: Presentación de los programas del
Laboratory for Computer Graphics and
Spatial Analysis de la Universidad de Harvard.

P R O F E S O R ERIC TEICHOLZ.

JUEVES 21

9 a 13 h: El Sistema de Cartografía Automatizada
de Cetenal.

P R O F E S O R Ing. Alberto Torfer Martel.

13 a 14 h: Comida.

14 a 16 h: Trabajos realizados en el Centro de Computación
de la Dirección General de Planeación. S.E.P.

16 a 18 h: Trabajos realizados en el Laboratorio de Planeación
Urbana de la Sección de Planeación.
D.E.S.F.I.
UNAM.

VIERNES 22

9 a 13 h:	Mesa Redonda sobre aplicaciones a casos de interés de los participantes al curso.
PROFESORES:	THOMAS K. PEUCKER. ERIC TEICHOLZ. CARLOS ESTRASSBURGER. ALEJANDRO VILLANUEVA. MARCIAL PORTILLA.
13 a 14 h:	Comida.
14 a 18 h:	Sesión de Práctica con los programas SYMAP, GRID y SYMVU.
PROFESORES	del Laboratorio de Planeación Urbana.

SABADO 23

9 a 11 h:	Evaluación de los resultados obtenidos en la Sesión Práctica.
11 a 12 h:	Reunión para establecer mecanismos de cooperación profesional para la aplicación y avance de este tecnología en México.
	CONACYT
12 a 13 h:	CLAUSURA.

COMPUTER CARTOGRAPHY

Thomas K. Peucker

COMMISSION ON COLLEGE GEOGRAPHY

RESOURCE PAPER No. 17

ASSOCIATION OF AMERICAN GEOGRAPHERS

WASHINGTON, D.C.

1972

Supported by a grant from the National Science Foundation

ASSOCIATION OF AMERICAN GEOGRAPHERS

Commission on College Geography Publications

General Series

*1 Geography in Undergraduate Liberal Education, 1965
2 A Basic Geographical Library—A Selected and Annotated Book List for American Colleges, 1966
*3 Geographic Manpower: A Report on Manpower in American Geography, 1966
4 New Approaches in Introductory College Geography Courses, 1967
*5 Introductory Geography: Viewpoints and Themes, 1967
6 Undergraduate Major Programs in American Geography, 1968
7 A Survey Course: The Energy and Mass Budget at the Surface of the Earth, 1968
8 A Systems Analytic Approach to Economic Geography, 1968
9 A Geographical Bibliography for American College Libraries, 1970
10 Geography in the Two-Year Colleges, 1970

Technical Paper Series

1 Field Training in Geography, 1968
2 Computer Assisted Instruction in Geography, 1969
3 Evaluating Geography Courses: A Model with Illustrative Applications, 1970
4 Living Maps of the Field Plotter, 1971
5 Simulation of the Urban Environment, 1972
6 Computerized Instruction in Undergraduate Geography, 1972
7 The Interface as a Working Environment: A Purpose for Physical Geography, 1972
8 Land Use: A Computer Program for Laboratory Use in Economic Geography Courses, 1972
9 Bibliography of Statistical Applications in Geography, 1972

*Out of print

Resource Paper Series

1 Theories of Urban Location, 1968
2 Air Pollution, 1968
3 Perspectives on Geomorphic Processes, 1969
4 Spatial Diffusion, 1969
5 Perception of Environment, 1969
6 Social Processes in the City: Race and Urban Residential Choice, 1969
7 The Spatial Expression of Urban Growth, 1969
8 The Political Organization of Space, 1971
9 An Introduction to Spatial Allocation Analysis, 1971
10 Man and Nature, 1971
11 Tropospheric Waves, Jet Streams, and United States Weather Patterns, 1971
12 The Spatial Structure of Administrative Systems, 1972
13 Residential Mobility in the City, 1972
14 The Periglacial Environment, Permafrost, and Man, 1972
15 Conservation, Equilibrium, and Feedback Applied to Atmospheric and Fluvial Processes, 1972
16. Metropolitan Neighborhoods: Participation and Conflict Over Change, 1972

COMPUTER CARTOGRAPHY

Thomas K. Peucker
Simon Fraser University

ASSOCIATION OF AMERICAN GEOGRAPHERS

Commission on College Geography
Washington, D. C. 20009

RESOURCE PAPER NO. 17

Library of Congress Catalog Card Number 72-75261

Supported by a grant from the National Science Foundation

FOREWORD

The Resource Papers have been developed as expository documents for the use of both the student and instructor They are experimental in that they are designed to supplement existing texts and to fill a gap between significant research in geography and readily accessible materials The papers are concerned with important concepts in modern geography and focus on three general themes geographic theory; policy implications, and contemporary social relevance. They are designed as supplements to a variety of undergraduate college geography courses at the introductory and advanced level These Resource Papers are developed, printed, and distributed by the Commission on College Geography under the auspices of the Association of American Geographers with National Science Foundation support. The ideas presented in these papers do not imply endorsement by the AAG. Single copies are mailed free of charge to all AAG members.

John F. Lounsbury
Project Director
Commission on College Geography

Paul W English
Chairman, Panel on Resource
and Technical Papers

Panel Members

Ian Burton, University of Toronto
Leslie J King, McMaster University
Harold M Rose, University of Wisconsin-Milwaukee
Robert H T Smith, Queen's University
Edward W Soja, Northwestern University
Philip L Wagner, Simon Fraser University
Julian Wolpert, University of Pennsylvania

TABLE OF CONTENTS

TABLE OF FIGURES

Many of the figures were produced as student projects. Since the understanding of the concept had priority over the accuracy of the data, some graphs contain slight errors. Therefore, scale and projection frequently are not indicated.

TABLE OF DIAGRAMS

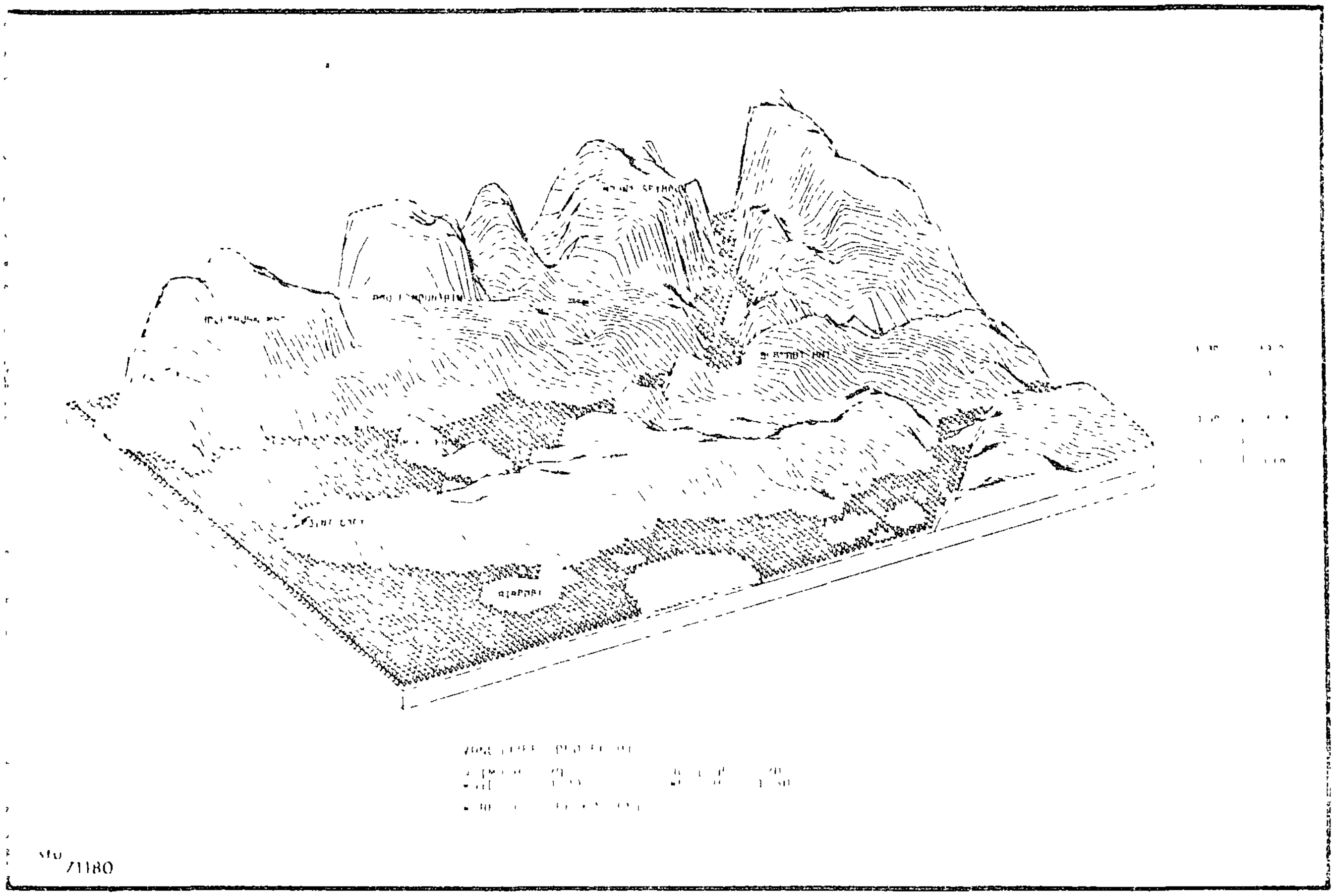


Figure 1.1 Perspective View of Greater Vancouver, View from S.W. A typical example of the advantages of computer cartography: the repetitive computation of the picture coordinates of a point in three-dimensional space (approximately 10,000 times in this case), followed by controlled plotting. Program: SYMVU, data: D. Mark, University of British Columbia, production: W.D. Rase, Bonn-Bad Godesberg

I. INTRODUCTION

Geographers use maps for the analysis, communication and storage of their findings And it seems that maps are recently more and more in demand In addition, there has been an increase in the availability of information with different kinds of "location identifiers" and an increase in the demand for thematic maps showing population densities, social class distribution in cities, land use maps, and perspective views of terrain and the like

Parallel to this growing interest in maps, the computer has played an increasing role in most areas of human life. Universities are major users of the computer and many scientific disciplines including geography are making increasing uses of the computer in research and teaching The development of computer cartography was initially hampered by the lack of satisfactory graphic instrumentation and by the misconception of the academic community that it was very difficult to handle non-numerical subjects by computer

Computers possess three different communication abilities literacy, numeracy and graphicacy [1] Like man, they are not equally well versed in all of them Among people, unevenness of performance is largely a result of uneven education and training For example, in the educational system the emphasis is upon literacy, whereas numeric and especially graphic training is relatively neglected

An analogy can be made between human communication and education and computer programs what is education for the child has its equivalent in the programs for the computer The computer-hardware is ready to do everything but able to do nothing without efficient programming

> n b There is a difference between hardware and software Hardware is the set of electronic and mechanical equipment used for processing data Software is the set of instructions (programs), developed by people to operate the computer The better the software, the more usable the computer

Until quite recently, the emphasis in software development has been on the solution of numeric problems most likely because scientists and engineers were the first to use the computer for their problems perhaps because they had more funds than others The initially slow progress in graphics was caused mainly by the need for further development of special equipment such as plotters and cathode ray tubes (See glossary) (Figure 1 2)

One of the earliest disciplines of those traditionally classed among the humanities to contribute to the development of computer science was linguistics In doing so, linguistics has expanded its own intellectual content considerably in connection with this contribution Linguists were able to use their experience with natural languages to investigate the problem of communicating with machines since a program as a set of commands is structurally very similar to a set of sentences in natural languages

In other words, a program has to be written in a language with a clear cut syntax (grammar) and a limited number of primitive symbols Many computer languages have been developed, some of which are highly *machine-oriented* (assembly languages), while others are machine independent Machine-oriented languages make programming difficult but speed up the resulting task (execution), *procedure-oriented* languages prescribe how the process of solving the problem is to be carried out, and finally *problem-oriented* languages necessitate only to state the problem and not the technique of solution

Although linguists have been working for centuries on problems of grammar, they have not had to think about languages with such an explicit syntax as those needed by computer languages Since numeric computations had to be put into a strict and logical form long before the computer was developed, the first computer languages were developed exclusively for numerical computations Only recently have we been given programs which handle strings of characters as well as numbers, and we should be better able to manipulate verbal data the more we learn about the basic theory

The work on graphic (two-dimensional) languages is still very slight and no definite language has been designed at this point Researchers in computer graphics make distinctions among diagrams (e g , flow-charts and circuits) sketches (e g , line drawings) and gray-scale pictures (which can be photographs, but also surfaces since gray-scales are only one type of z values, as is height, etc) Theoretical studies and language developments are in progress for all groups, but the emphasis is on diagrams and sketches rather than on pictures

---

[1] See Balchin, W G V and A M Coleman 1967 "Cartography and Computers *The Cartographer* Vol 4, 120-27

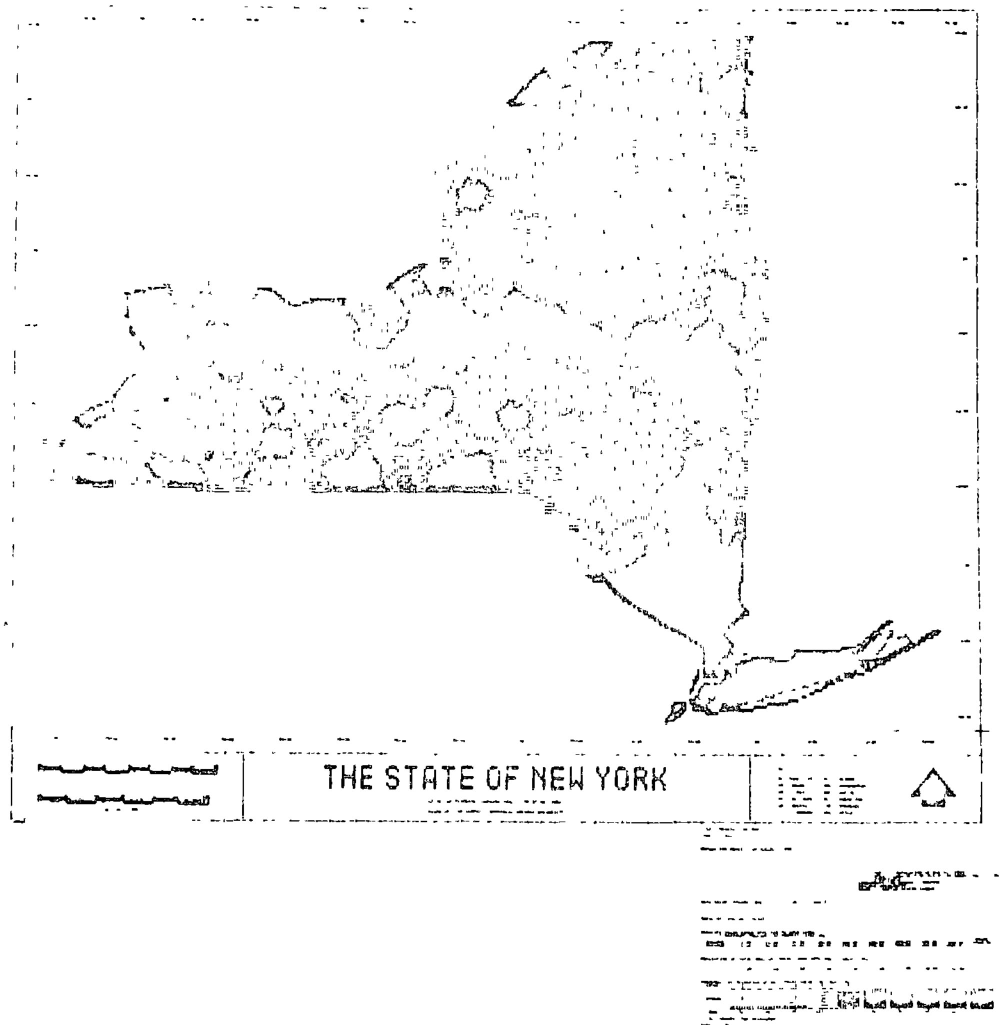


Figure 1.2 New York State Population Density, 1990. A good example of a line-printer map. Program SYMAP, data Office of Planning Coordination, State of New York, production C. Steinitz, Harvard University

This resource paper's emphasis will be on theory, rather than on actual production processes. Only little space will be used for the discussion of hardware, since the lifetime of any particular technology is short and the student ought to learn subjects that have a reasonably long "lifetime."[2] Our intent is to bring the student to a level of understanding from which he can set out on his own. No general theory is available, but an attempt has been made to bring some of the varied ideas together within a single framework. Neither is there one package of techniques to guide one in a cookbook fashion through the subject matter. Finally, no textbook exists and the author will sometimes refer the reader to sources not normally cited in introductory texts. (Figure 1.3)

[2] Tobler, in Tomlinson, 1970.

Information Theory and Cartography

Computer cartography must be seen in a wider framework to take advantage of developments in other fields in order to transpose one's findings to a higher level of generality. We will therefore present the subject here in the framework of information theory.

In the general use of the word, *information* means *instruction*. A transmitter relates news to a receiver. The news is embedded in familiar items. For a clearer definition, the novel elements in the news will be called information and the familiar items *redundancy*.[3] In a statistical sense, terms such

[3] The following paragraphs are based on Alsleben, K., 1962, *Aesthetische Redundanz*, Hamburg.

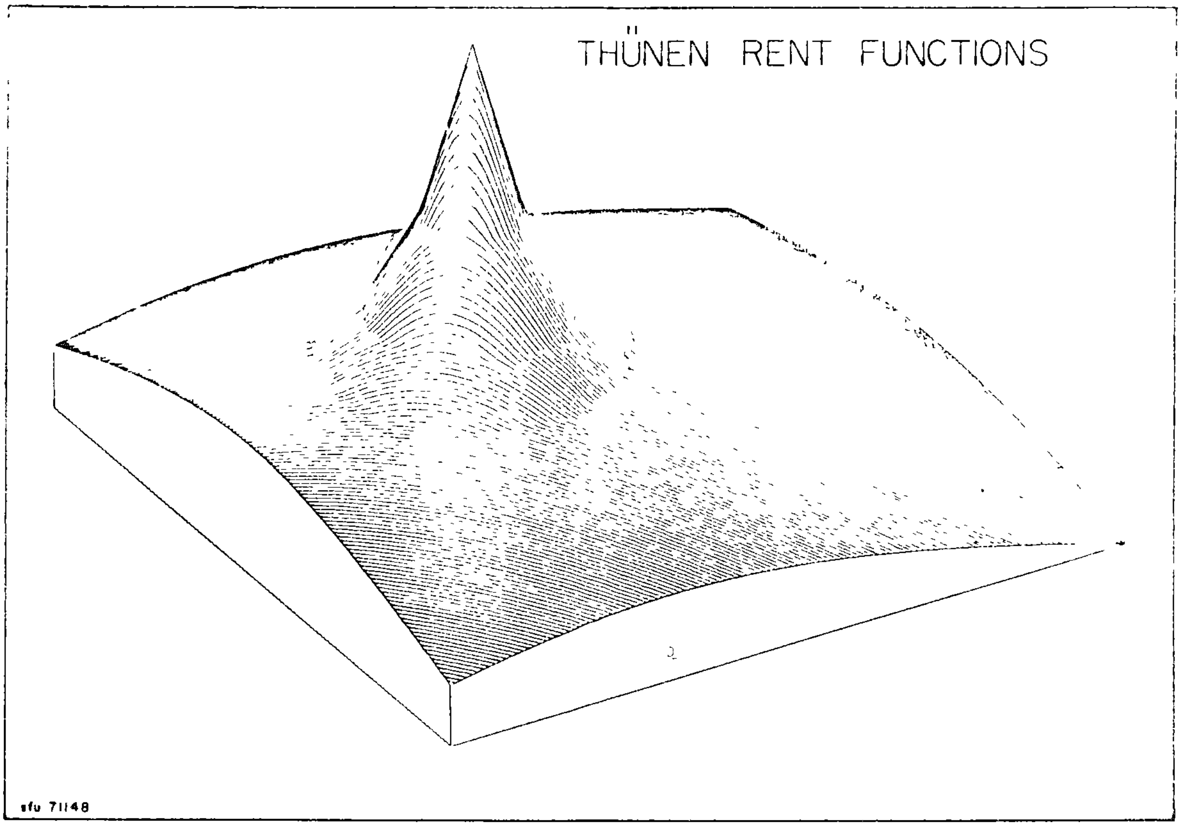


Figure 1.3 Three-Dimensional Rent Functions The graphic representation of geographical models can aid our understanding of them Programs SPATFUN, SYMVU, production T K Peucker

as "information," "news," "surprise," and "improbability" correspond to the term "rareness" A highly unexpected news item is a great surprise, it has a high information content In information theory, a framework has been developed to determine this content quantitatively The measurement unit is one bit, denoting one two-way (binary) decision The number of bits (i e , the information content of a news item) is therefore determined by the number of binary decisions one has to make to sort the item out from the sum of all expected items[4]

The problem of measuring information is usually more complicated since information items vary in their probability of occurrence[5] However we will not go beyond this definition since it is not pertinent to our understanding in this paper

If information relates to instruction, then the reception of information relates to learning No news can be taught totally independent from known knowledge Redundancy is the complement of information, it links information to the framework of knowledge. With any graphic display one therefore has to combine news with the familiar Those elements of a map which inform about the variation of the object (the earth's surface) relate to information, those elements which account for the visual effect of the map relate to redundancy (Figure 1 4)

There are as many different types of information and respective types of redundancy as there are map elements We talk about color, brightness, pragmatic information, effective information, typographic information, etc One classification would have two groups of information according to the mode of the viewer's attention One group would cover the intellectual and rational, the other, the sensory area (Figure 1 5)

When examining a map, one observes in three phases

1. Selective phase The viewer selects consciously or unconsciously, from the large number of signs (e g , parallel lines)
2. Synthetic phase The viewer recognizes composite signs (e g , a street)
3. Analytical phase The viewer studies the structure of the sign (e g , the street has intersections)

[4] As an example the checker-board, with its 64 fields, needs six binary decisions to reach any of the 64 (2^6) fields on the board

[5] For example, in the English alphabet the letter "e" has a much higher frequency and thus a lower information content in prose than, for example, the letter "k"

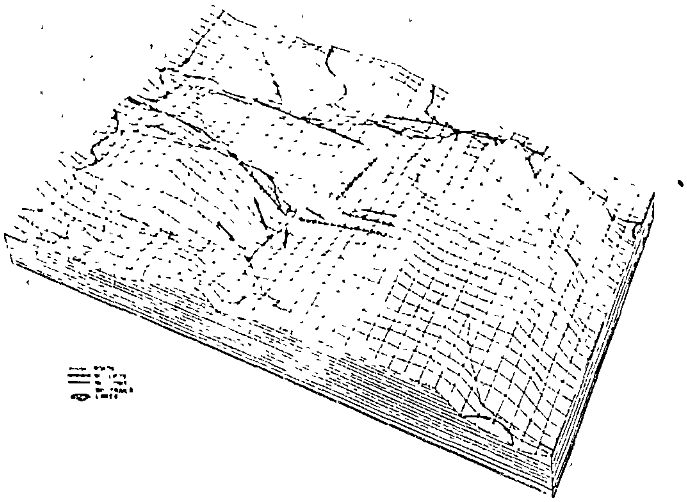


Figure 1.4 Ski Area. Perspective plots can aid planning and advertising. Program: VIEWBLOK; data: D. Douglas, University of Ottawa

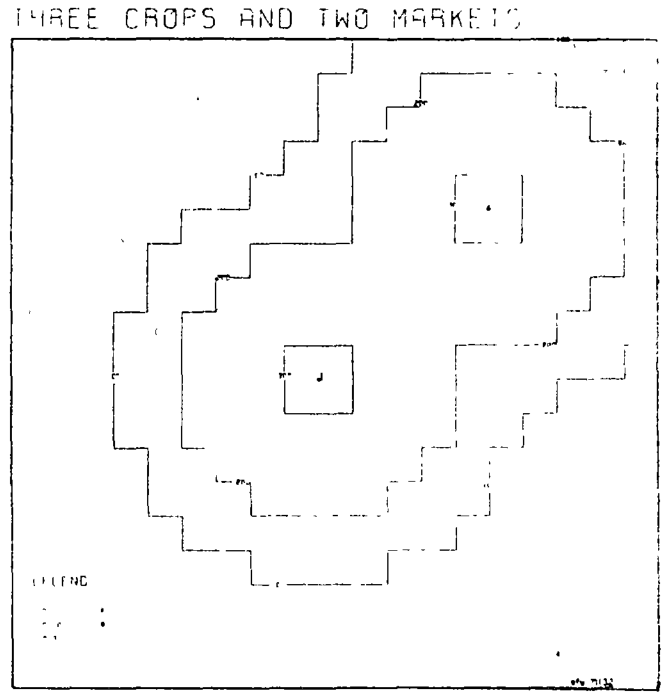
During the process of map-examination, the viewer will go through these phases several times, moving on to higher levels of observation. Composite signs become simple signs on the next level building an inverted tree of signs of increasing composition.

Two types of classes of signs, similarity and hierarchy, can be considered. The human brain classifies a large number of signs as similar, e.g., II = ll = 2. It also sums up signs to more complex signs, e.g., a number of houses with streets become a settlement.

This ability to summarize content is largely due to the propensity of human perception to reduce a large supply of information to a manageable size. It has been found that the eye alone can receive three million bits per second. However, only sixteen bits per second can be consciously present in the brain.[6] The reduction of information can be accounted for by 1) a shift of information into unconscious areas (where it can release reflexes), 2) concentration into classes, 3) the linkage to familiar knowledge and 4) a conscious limitation to a single field of content.

A cartographer must keep these points in mind when designing a map. To be effective, a map must be based on the perceptual ability of men. For example, in producing a tourist map for drivers which exhibits terrain besides traffic information, the relief shading is preferable to isarithms, since it leads directly to the desired composite sign--relief--whereas one has to go through several steps of map-examination to understand relief from an isarithmic map. (We can be confident that the driver will not want to have exact heights on each point of the map.) (Figure 1.6)

[6] See Alsleben, K., 1962, *Aesthetische Redundanz*. Hamburg



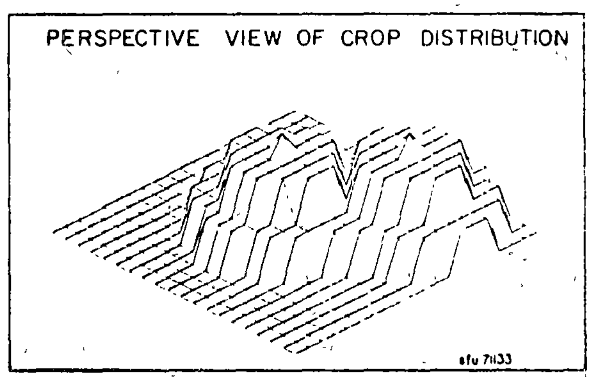


Figure 1.5 Crop Distribution Simulation. The simulation of land use with two markets based on the Thünen theory. Programs: SIM, PERS, CNTOUR; production: S. Witiuk, Simon Fraser University

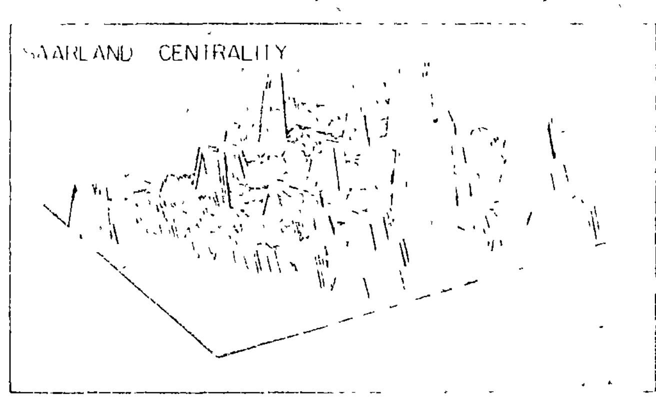


Figure 1.6 Saarland, Centrality. The results of a factor analysis of 103 variables and 356 communities. Factor 1: centrality. Programs: FAKAN, SYMVU; data and production: W.D. Rase

The Features of Numeric Cartography and Their Structures

Dacey (1970) formulates the concept of language as follows

> A *sign* is the smallest unit that designates, and that which is pointed out is called the *designation* of the sign. A *sign process* is an arrangement of signs that designates something. *Language* is an institutionalized collection of signs that have common designations to members of the community using these signs. The signs are producible by members of this community and they may be combined in some ways, but not in other ways to obtain sign processes which also have a common designation to the users

One basic feature of formal and natural languages is concatenation (i.e., linkage of signs in one dimension). The only ordering concepts in these "linear languages" are the serial orders of "before" and "after." In map-languages which are intrinsically two-dimensional, the concepts of neighborhood and juxtaposition in two dimensions are basic features. These features can be expressed by properties such as proximity, connection, superimposition, above, below, left, right, close-by, clustered, dispersed, distance decay, spatial autocorrelation function (see chapters III, IV, V) spatial overlap, etc.

Another point must be considered. Not only is the concatenation in two dimensions important for map languages but the structure of the map (i.e., the interrelationships of signs), also has to be studied. Classes of signs cannot only be organized horizontally between the elements of a single level, but also must be hierarchical (i.e., represent connections and associations between the various levels) (Narasimham, 1969).

An example would be a large-scale map showing every building by a symbol. Although every building is different in reality, a map will show a limited number of building types. Thus, the very large number of types of buildings which can be found on, say, an aerial photograph is already reduced to a small number of building classes or signs. To create structure, one can build a set of super-signs such as the following

> a farmstead is a very dense clustering of more than one but less than seven signs of the classes "house" and "barn."
>
> a hamlet is a dense clustering of more than four but less than twenty signs of the classes "house" and "barn."
>
> a village is a dense clustering of more than fifteen but less than two hundred signs of the classes "house," "barn," and one sign of the class "church."

This very simple - and probably in most cases insufficient—description of the classes farmstead, hamlet and village gives an idea of what might be pertinent in a hierarchical map language. Of course, in an operational language the properties "dense" and "very dense" would have to have some quantitative meaning. The classes overlap, as in reality. For automated map analysis, one must add other property dimensions, or make the process interactive (i.e., in each case of ambiguity, the machine reports, leaving the decision to man) (Figure 1.7).

In a learning machine, one starts with a very small vocabulary of signs (symbols) and a very simple structure (syntax, grammar) and builds both up during the process of map analysis. As soon as the machine gets to a sign which is ambiguous or unknown to the machine, man has to make the decision by either allocating the sign to a known class or by creating a new one. The machine would store the decision or the new class and act accordingly when the case occurred again. At the start of the analysis, man has to act very often, but the frequency decreases quickly.

Syntax is not the only aspect of cartographic analysis and representation. There are other factors playing a role which should become clear when reading the rest of this resource paper and other publications in cartography.[7]

The methods of cartographic presentation can be classified into five groups: semantic methods, syntactic methods, psychological methods, physiological methods and signal methods.[8]

1. Semantic methods. These methods cover all meanings which can be attached to signs, including the object of representation, whether a perspective view of a tree or a surface function. Semantic methods take effect through the object.

2. Syntactic methods, as discussed before, are the

---

[7] See especially Commission on College Geography Resource Paper No. 19, *Thematic Cartography*, by Philip Muehrcke.

[8] After Alsleben, K., 1962, *Aesthetische Redundanz*, Hamburg.

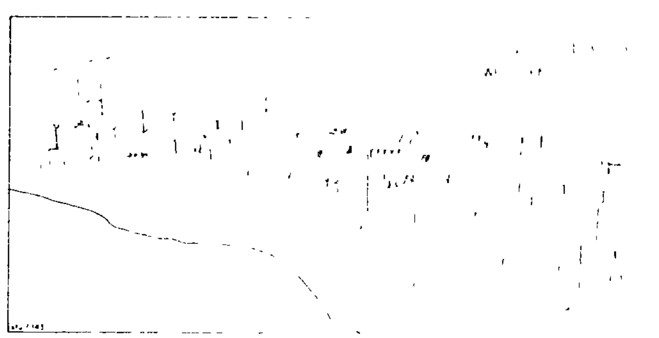


Figure 1.7 Vancouver, West End. Vancouver's major high-rise apartment area. All houses are represented by blocks. The data bank of houses can easily be changed to show residential development. The coastline is drawn in manually. Program BLOCKS, data A. Turguson, coastline L. Nelson, both from Simon Fraser University.

geometric and statistical arrangement and distribution rules They include structural elements such as neighborhood, periodicity inner connection, etc

3 Psychological methods These methods cover phenomena considered by gestalt psychology, such as similarity classes, perceptive process, etc and terms such as rhythm and dynamism (in an artistic sense), intensity, harmony, etc. Psychological methods take effect through perception

4 Physiological methods are the results of the conditions of the optical organ such as visual field, depth of focus, contrast sensitivity, color vibration, etc

5. Signal methods represent perception elements which take effect by nothing but themselves Background and base-color are examples in cartography (Figure 1 8)

All of these methods must be used in cartography However, their application is highly intuitive at present because tests for the analysis of the methods' effects are virtually non-existent Such tests, under laboratory conditions that change only one factor at a time are very difficult to achieve It is hardly possible to produce manually a series of maps which differ only in selected aspects As the cartographer draws one map after the other, he improves the overall impression of the map unconsciously Tests have only been run on abstract distributions of symbols and only rarely on very simple maps (Figure 1 9)

We are now able, however, to automatically produce series of maps under strictly controlled conditions and should therefore be able to advance our knowledge of the means of map representation at considerable speed [9]

Points, Lines and Surfaces

In the following chapters, three different types of map features will be discussed–points, lines and surfaces All three are spatial signs of different dimensions In addition, a map contains a variety of symbol-types which are signs without spatial dimensions but usually with specific locations

Points are zero-dimensional signs which either denote the location of content (name, value, etc) or constitute a part of the other two features A line consists of one or more line-segments which are straight lines (vectors) bound by a point on each end Theoretically, any curved line has an infinite number of line-segments of infinitesimally short lengths, but in computer cartography, continuously curved lines can be constructed only in rare cases (when analog plotters are available), in practice, a line is composed of a number of straight line segments with discrete lengths Lines can have a variety of contents (functions) (i e , they can

[9] Psychopictorics is a field very close to this attempt See Lipkin and Rosenfeld, (1970)

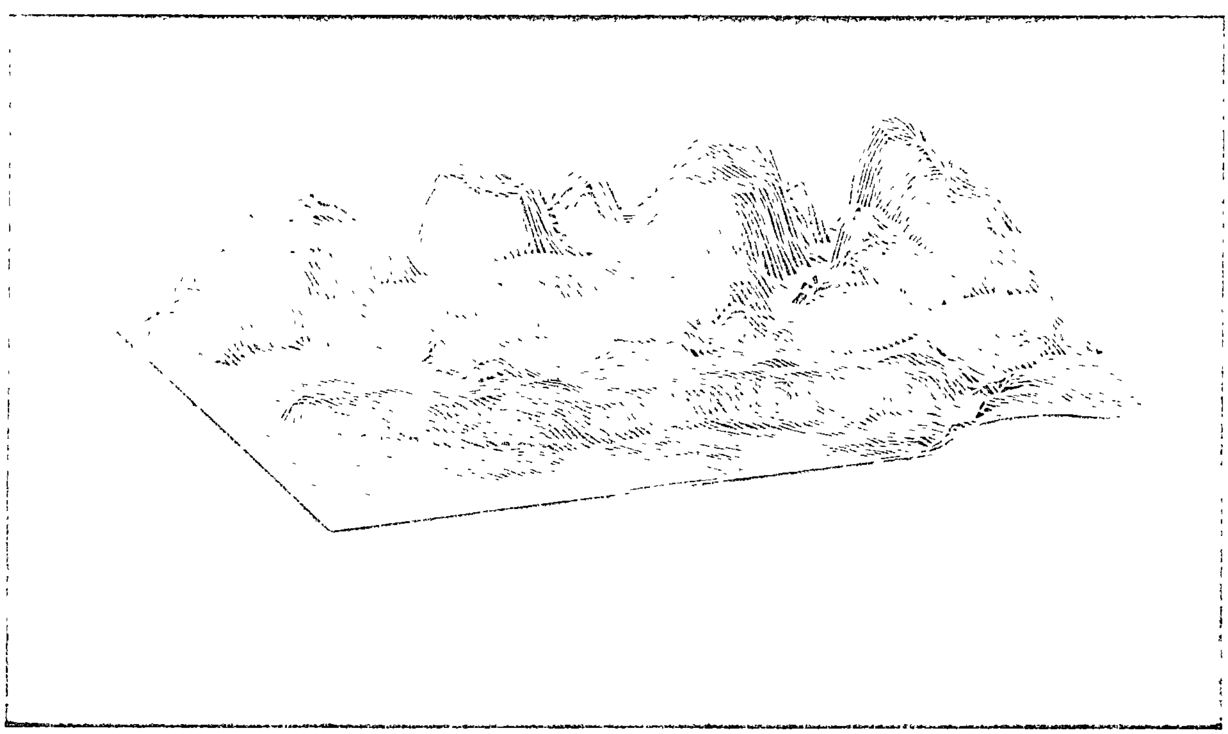


Figure 1 8 Vancouver This negative of a plot is an example of the signal means of cartographic representation Program SYMVU data D. Mark, production W D Rase

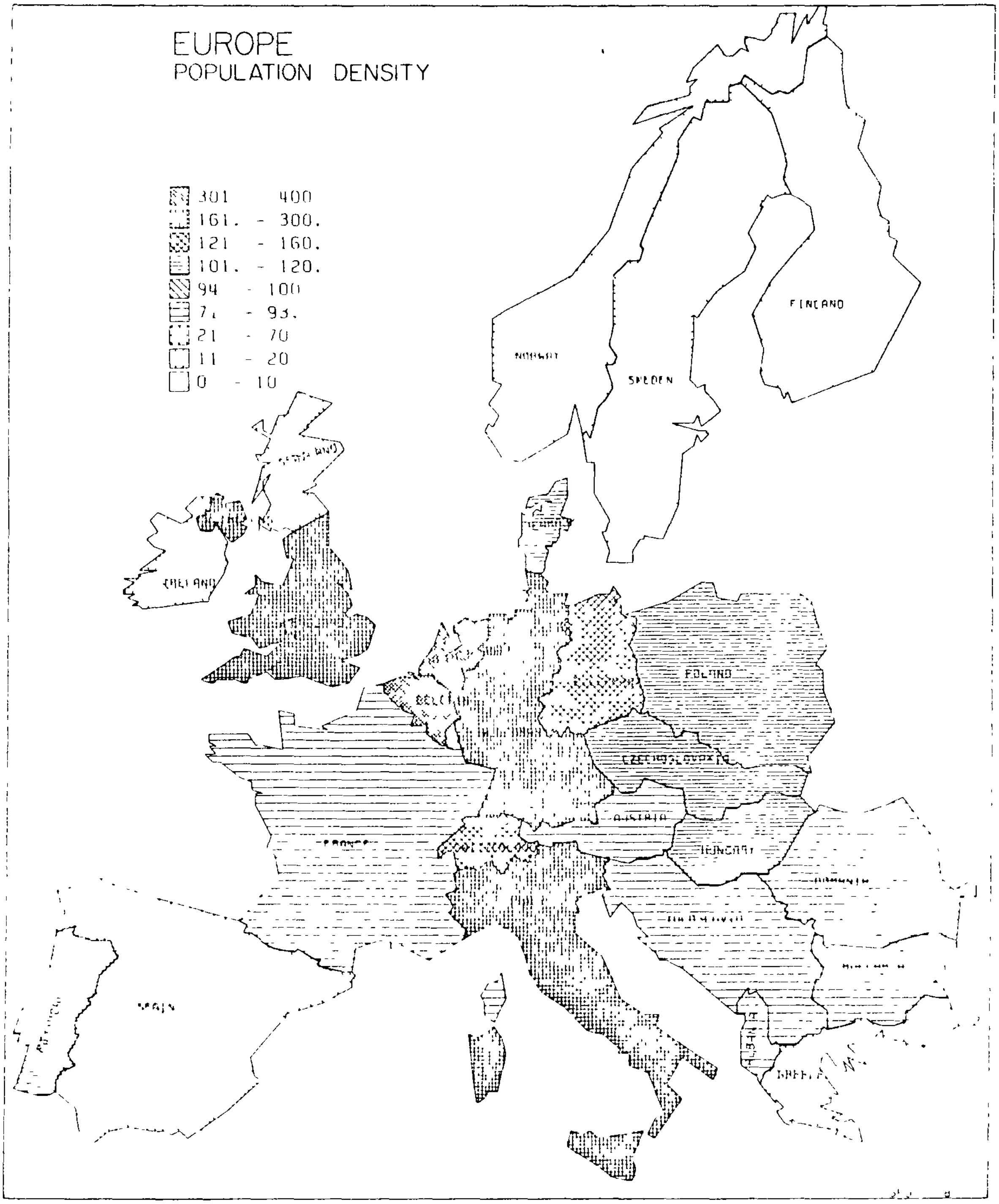


Figure 1.9 Europe Population Density. The plotter produces very fine shading but the computing and plotting time is prohibitive if used on a day to day basis. Program CALFORM, data G. Hayward, production C. Johannson

represent rivers, roads, boundaries, flows, contours and so forth.)

Surfaces can be determined in a numeric fashion by a grid of points, a series of lines, or by one of several mathematical functions. The content (value) of a surface can represent not only altitude, but also density, pressure, proportion, intensity, potential etc. (Figure 1.10)

Traditionally, the emphasis of computer cartography at the universities has been on surfaces, especially surface analysis and then only representation, and less on points and lines. The reason lies mainly in the kind of equipment which a researcher can use at a university. A large computer as graphic machinery generally satisfies only very modest demands in terms of quality and accuracy. Therefore, the development of the point, line and symbol production has been left almost entirely to governmental agencies which have developed impressive hardware and software systems.

The opportunities for cooperation between the academic and governmental groups are enormous. Theory and concepts from the former, and practical experience and available equipment from the latter could complement each other tremendously. Unfortunately, it must be admitted that this cooperation hardly exists. Neither group can particularly be blamed for this, since communication even within each group is poor. It might be the newness of the subject or the diversity of the interested people and institutions, or simply the lack of any common organization. However, especially with regard to organization, very promising efforts have been undertaken recently (see Tomlinson, 1970).

Hardware

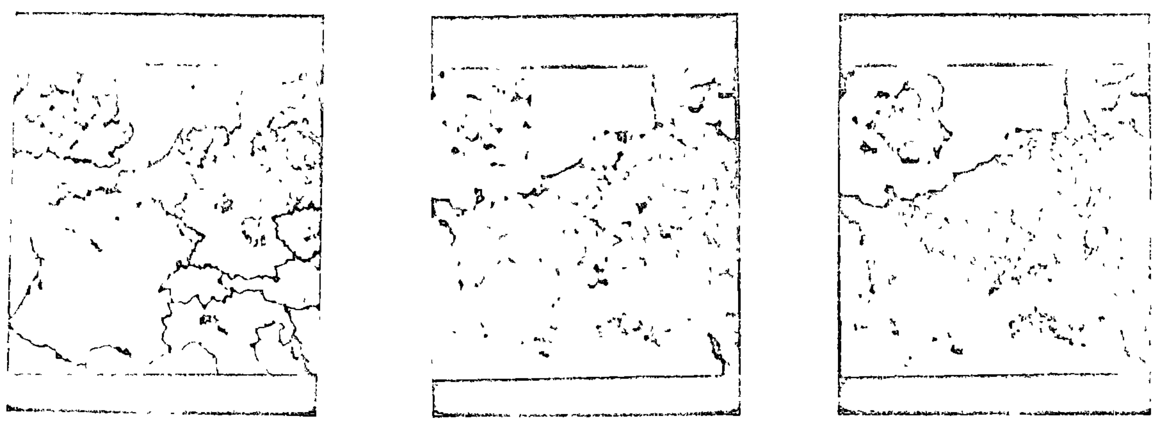
It must be said again and again that a successful study of computer cartography is not dependent on the availability of special graphic instruments. Some of the most exciting concepts have been developed without any adequate hardware. Some of the more interesting developments of graphic input-output devices have been published with photographs of paperboard models and simulated results. The construction of the machinery is dependent on the commercial demand for it; ingenuity cannot and must not wait for commercial needs. Two examples should clarify the problem, one from the area of projections, the other from analytical hill-shading.

The construction of map projections always involved time consuming calculations. It was natural for those involved in this job to use the computer, at a time when even line-printers were too messy and slow to produce printer-maps. But even the computation of a terrestrial grid in Cartesian coordinates was a great savings. However, those interested in studying projection systems in a more theoretical framework[10] were very well aware that they were developing concepts to be applied on the proper hardware, often available incidentally, by the time they had completed their work. (Figure 1.11)

Another example is that of hill-shading. The manual approach is developed to an impressive level. However, its use is restricted by cost and the scarcity of qualified cartographers. If one accepts the general sense of Turing's theorem (i.e., that any procedure which can be specified in a finite number of steps can be automated), the automation of the

---

[10] See for example: Tobler, W. R. (1962), "A Classification of Map Projections," *Annals, AAG*, Vol. 52, 167-175, and Kao, R. C. (1967), "Geometric Projections in Systems Studies," in W. L. Garrison and D. F. Marble, eds. (1967) *Quantitative Geography, Part II*, Evanston, 243-323.



a. original map — b. density map (PAX-cluster analysis) — c. a and b

Figure 1.10 Agricultural Density: Acreage in Hay, 1935. Some local averaging would have improved the image, but the primary purpose was the elimination of country boundary lines. Shows also the advantages of half tone shading. Program: PAX II, data and production: E. E. Pfaltz, University of Virginia.

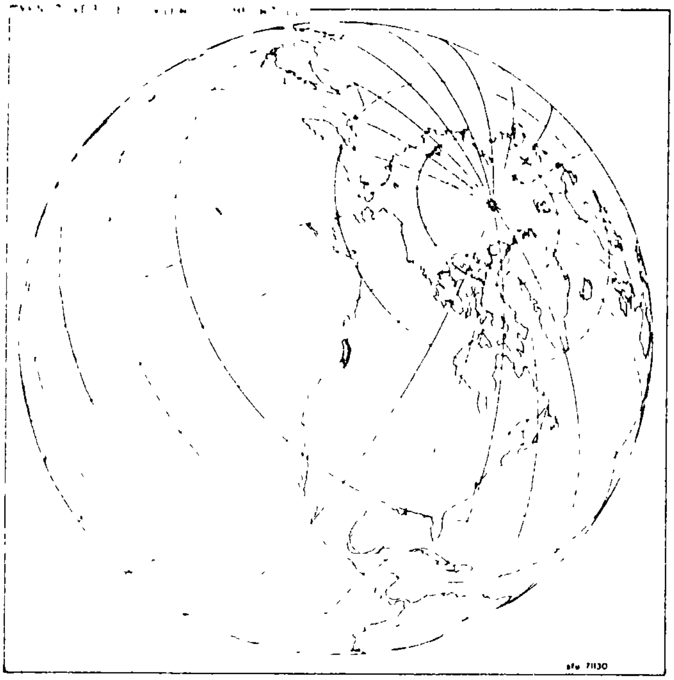


Figure 1 11 Azimuthal Projection First, the grid is drawn and then over 5,000 points from a data bank are recomputed and plotted Program SUPERMAP, production W D Rase

method is no surprise[11] However, Yoeli did this[12] with graphic equipment which gave him no chance even to come near the quality of manual shading he produced a dense grid of gray values, generated on the line-printer different gray tones through overprinting, cut them out, and pasted them onto a large sheet in the order indicated by the computer[13] The result was then photographed This author reprogrammed the method for a raster-recorder (with a 64 level gray scale and a grid of up to 4,000 x 4,000) located 3,500 miles away Of course the usage of such a large grid involves considerable numerical development but the credit for the method clearly stays with Yoeli (See Figure 5 8)

The discussion of the hardware will therefore be restricted to a very general description and a very limited outlook The basis for all work in computer cartography is, of course, the computer, a device capable of accepting information, manipulating it by a sequence of stored commands, and supplying the results of the manipulation It usually consists of arithmetic and logical units and a control unit, storage devices, and input-output devices The cost can vary from little over a hundred dollars to close to a quarter of a million dollars rental time per month The typical computer for student use at the university is medium to large, with 32,000 to 128,000 words of usable fast memory, several magnetic disc and tape drives for secondary storage, and a high-speed printer Many systems include a drum plotter and some of the larger ones also include an interactive cathode-ray tube The market is rapidly expanding for minicomputers, but cartography is not affected as yet, despite many valuable potential applications

We will first discuss different graphic output and input devices The line printer is certainly the most available device Just as one can produce pictures with the typewriter by taking advantage of the varying gray level of the different characters, so can the line printer produce maps Several mapping programs with examples in this paper are adequate proof With up to 2,000 lines per minute it is also a very fast instrument. The drawbacks are its low resolution, with 10 characters per inch across and 6, 8 or 10 per inch down, and the relatively quick fading of the carbon-ribbon which often results in a different gray-tone from one sheet to the next

Both disadvantages have been eliminated by electrostatic line plotters which plot lines of points, 100 per inch at high speed However, the paper is about ten times more expensive than printer paper, which makes many computing centers hesitant to install such a unit They prefer the very versatile plotters which draw lines defined by a sequence of coordinate pairs There are two different types of line plotters on the market, either the cheaper drum plotter which uses the rotation of the drum as the movement along the x axis and the horizontal shift of the pen as the movement along the y axis, or the flatbed plotter where the pen moves in both x and y direction over a table They have either low or high resolution (100 to 10,000 steps per inch), low or high speed (1 to 40 inches per second), one exchangeable pen or a multi-pen head, etc, depending on the price Most of the plots in this paper are produced on a thirty inch drum plotter, 200 steps per inch, 1 5 inches per second and one pen (Figure 1 12)

The cathode ray tube is familiar to everybody by its use in television Unfortunately for cartography, the CRTs used in connection with the computer are often quite different They usually operate under vector mode and not raster mode (i e , they draw straight lines of any length with a maximum of 3,000 to 6,000 lines), and do not permit selective turning on or off of individual points in a matrix of between 256 and 4,096 points square They also usually do not offer varying gray-tones, but only dark lines on light background or vice-versa But they are very fast in response so fast that one can even produce computer movies They are also often interactive, that is, they allow the user to point to different

[11] Tobler p 128 in Tomlinson 1970

[12] Yoeli P (1965) "Analytical Hill Shading" *Surveying and Mapping*, Vol 25 573-579 and Yoeli P (1967) "Mechanization in Analytical Hill Shading," *The Cartographic Journal* Vol 4 82-88 An example can be found in Robinson and Sale, 1969, p 171 and 185

[13] The method is indicated in Chapter V

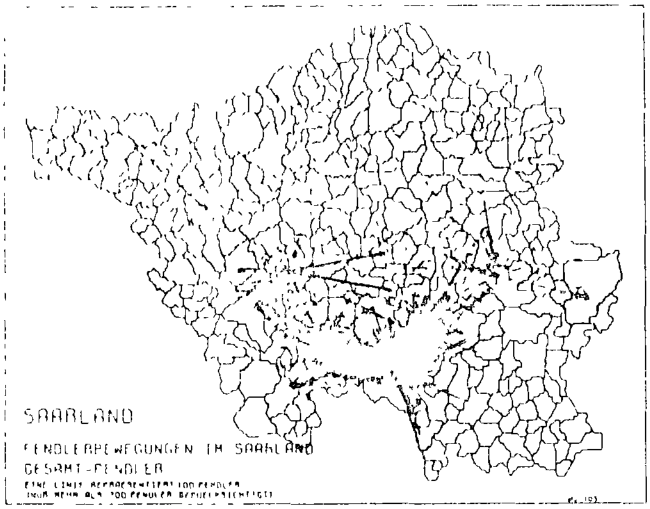


Figure 1.12 Saarland, Commuting. The commuting pattern in this German province. Program: PFEIL, data: W.D. Rase

places on the screen and make the computer react in a programmed way. Examples of the usage of such a device can be found in chapters five, six and seven, and at the end of this chapter (Figure 1.13).

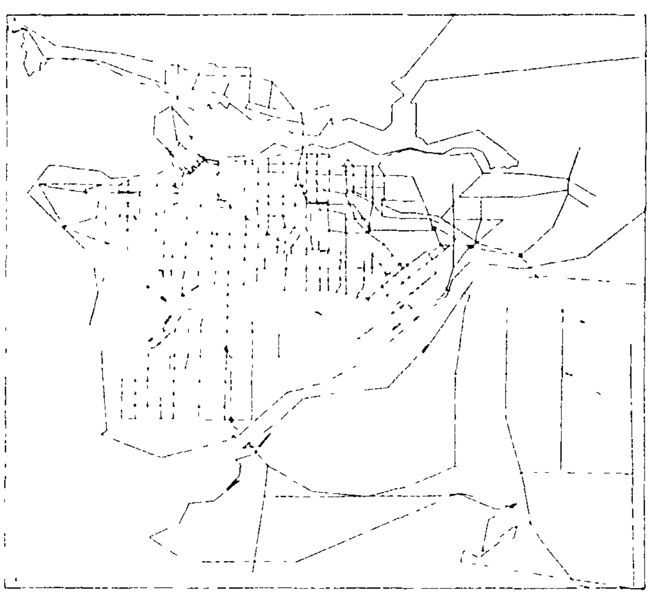


Figure 1.13 Vancouver Street Network With Coastline. Data: S. Witiuk

The vector-mode CRT is too coarse for the production of most final maps. Lines in cartography are usually too complex to be represented by a relatively small number of vectors. In addition, hachuring is almost impossible. However, as an intermediate device, any CRT can be of enormous value because it allows the user to work with the map. Half-tone CRTs might be more useful in computer cartography, especially for final copies. Figure 1.10 shows the results from an input-output scanner on film with a point matrix of up to 4,000 by 4,000, and 64 halftones, developed at the University of Maryland.

Many input-devices such as picture and map-scanners, height scanners on aerial photographs, automatic line-followers and special pattern-recognition computers have been developed. However, they are seldom available to students of computer cartography. The only device which is frequently available in universities is an xy-digitizer similar to a drafting table with any one of several mechanisms to record the x, y position of a courser or magnetic pen. The digitizer is very accurate and relatively fast, but does not relieve the user of the need for careful preparation (Figure 1.14).

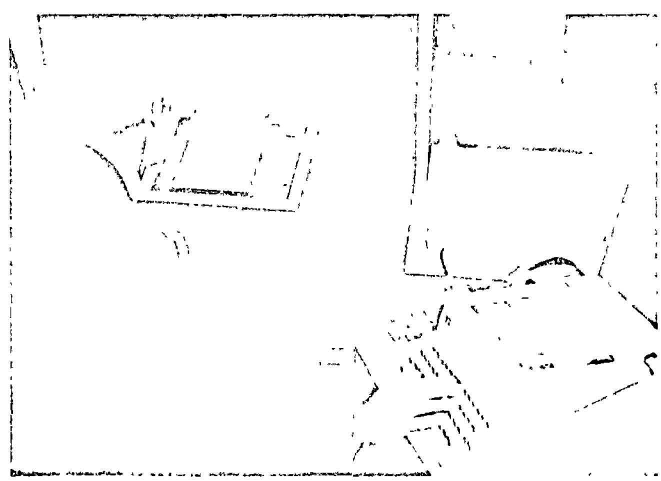


Figure 1.14 Interactive Storage Tube. The tube is in an almost horizontal position for working convenience. The operator has his right hand on a teletype console for alphanumeric interaction with the computer. Between the teletype and the tube is a "mouse" for two dimensional interactions with the computer. Courtesy A.R. Boyle, University of Saskatchewan

II. A SHORT INTRODUCTION TO ONE MAPPING SYSTEM

Many mapping systems have been developed, but only a few to the stage at which they can be used by people other than their programmers. The mapping system to be described is one of the most widely distributed, SYMAP, developed by the Laboratory for Computer Graphics and Spatial Analysis at Harvard University. Its widespread use is not necessarily due to its quality—although it is a very good program—but its good documentation and its availability through purchase.

The usability of SYMAP is based on two instruments available in virtually all computing centers: a medium-sized computer and a line printer. Just as the typewriter can produce pictures by taking advantage of the different gray-scales of letters, numbers, and signs, so can the line printer, but at a speed of 600 to 2,000 lines per minute, usually 132 symbols per line. The resolution is ten symbols per inch horizontally and six or eight symbols per inch vertically. The width of the output sheet (132 symbols) is no restriction since one can print a map in strips and then fasten the parts together (Figure 2.1).

Other positive features of the program are flexibility of input and the great variety of output. The standard form of input is on cards (punch-cards, I.B.M. cards), but some or all the information can also be read in from magnetic tape or drum. This process would normally be the result of some previous computation such as data standardization, regression analysis, factor analysis and the like. There is a standard card format, that is, the computer expects certain data in certain "fields" (groups of columns), but even this can be changed.

The main input for the production of a map consists of coordinates for points which represent either data-points or locations along lines. The different types are grouped in so-called packages which consist of a card with the package title, the point-cards (one card per point) and package-end card. The different packages are:

Outline	Barriers
Conformolines	Values
Data-Points	Map
Otolegend	

Only the map-package is mandatory.

The output can be in three types of maps. The contour map represents the z values of a set of (usually) irregularly distributed data-points by contour lines with up to twelve class intervals. The choropleth map portrays statistical areas with given boundaries, again by up to twelve class intervals. The proximal map is similar to the choropleth map, but the boundaries of the statistical areas are not given. It creates these boundaries between pairs of data-points (Figures 2.2, 2.3).

Input is governed by the choice between a fixed built-in data-structure or a flexible user-supplied one. In addition, the program can be operated by using only a partial selection of the packages and electives of the map-package. For example, one might produce a map without legends and choose the standard size (13 inches). In this case, one can leave out the Otolegend-package and does not have to specify the size-elective (command), but can rely on the default option of 13 inches.

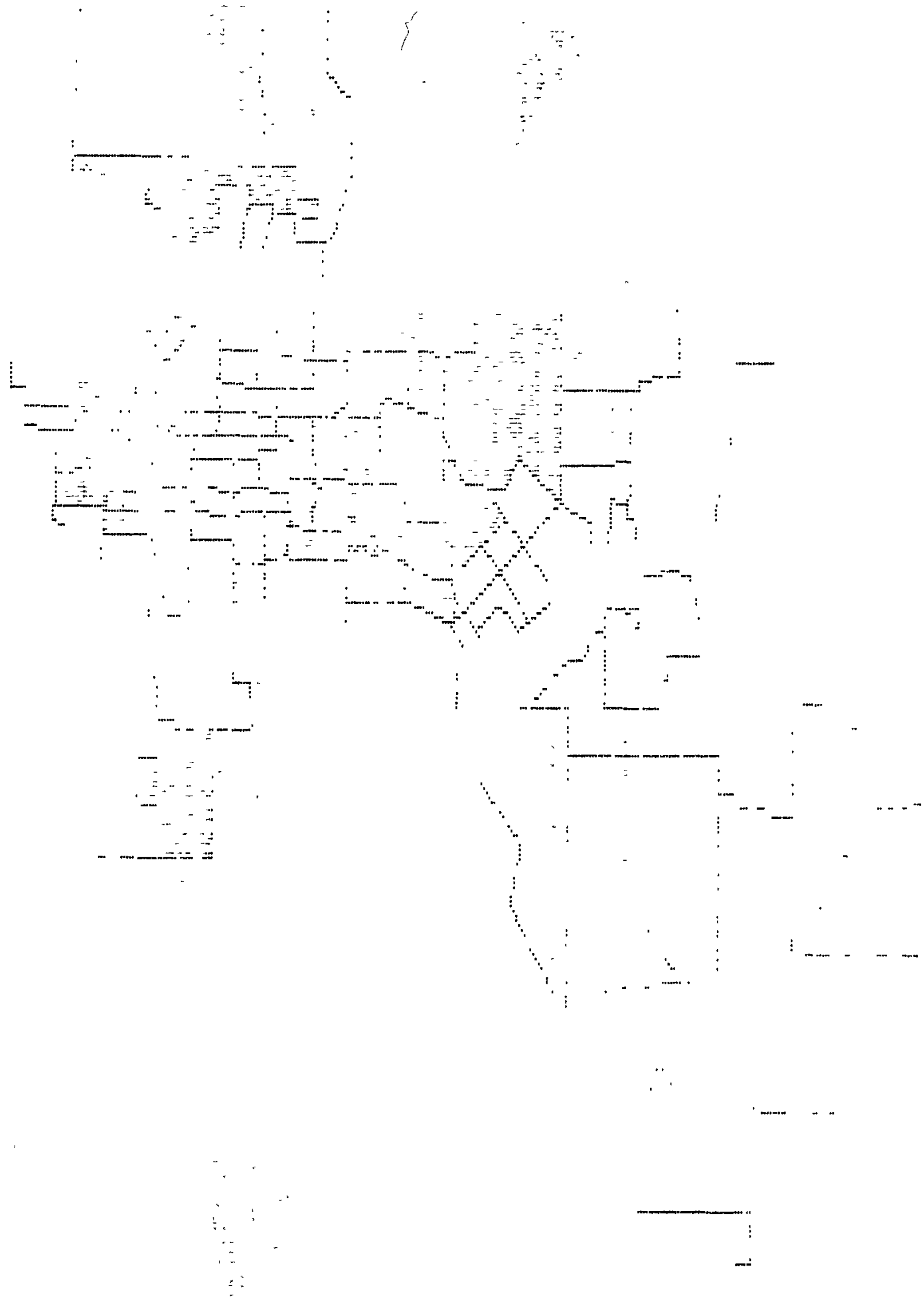
The possibility of operating the program with a smaller set of commands, however, can turn the program into a burden to both the user and his computing center. This is the fault of the user not the program. Even in its simplest form (i.e., with everything left to default), SYMAP has to reserve core storage and use up computer time for the unused packages and options. Many of the maps produced with SYMAP could be run faster and more cheaply with a smaller program. For some routine work, the development of such a program would pay off.

We should be encouraged, however, to look into SYMAP in more detail so that we do not use a restricted set of commands simply out of ignorance. We will start with the simplest image and then add other features.

The only mandatory package in SYMAP is the map-package which controls the production of each map to be printed. It has approximately thirty options, none of which have to be used for the first case. Only the package name, three cards of title and two termination cards are necessary to produce a 13 x 13 inch printout with a frame only.

By adding a set of data-points, one obtains a contour map. The saddle-point picture, which represents an often discussed matter in cartography, is created by the four corner points where the points at the ends of one diagonal have high values

Figure 2.1 Grouse Mountain, North Vancouver. Ten class intervals have been used for the contouring. The map shows the different combinations of letters used. The numbers show the location of data-points and their level. Program SYMAP, data D. Mark



sfu 71174

Figure 2 2 Vancouver A choropleth map showing the results of a factor analysis of the census of population factor (79 variables, 118 census tracts) Data T.K Peucker, production W D Rase

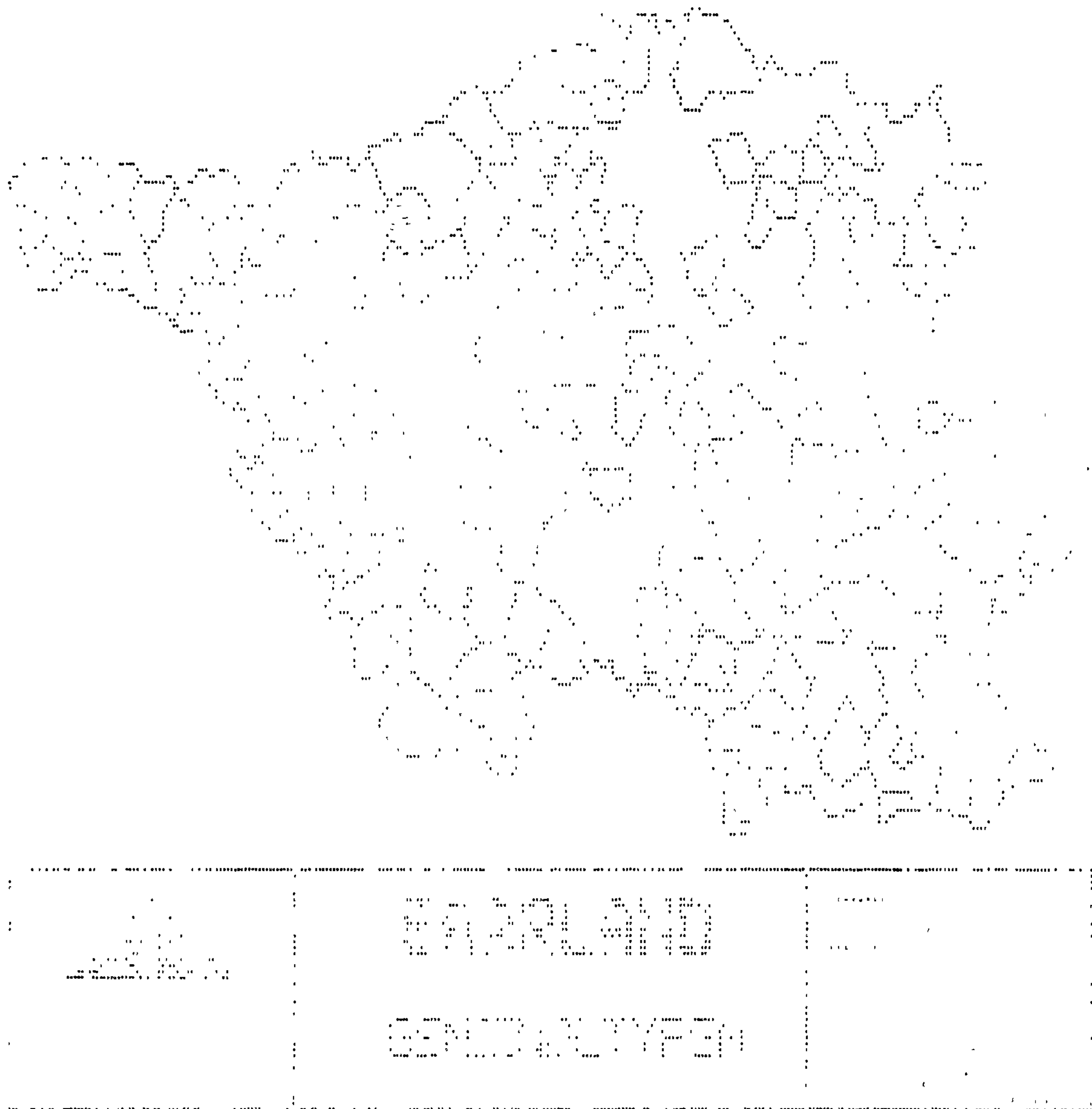


Figure 2.3 Saarland, Centrality. This example shows that the proximal-type map can be very suitable for classification mapping. The areas of little and high variation are well delineated by the density of borders. Program SYMAP, data W. D. Rase.

and the points on the other diagonal have low ones. The points are represented in the data-point package with their coordinates and in the values package with their z-values (height, etc.). As distinguished from the standard mathematical coordinate system, SYMAP assumes the origin in the northwest corner of the map and measures first down and then across. The reason for this sometimes very inconvenient regulation lies in the program's history. In earlier versions of SYMAP, the coding had to be done by lines and columns of the future output. In today's more user-oriented version, this is done by the computer after shift and scale parameters have been taken into consideration (Figure 2.4).

The cartographer's area to be mapped is usually not rectangular.[14] Cities have irregular shapes; countries are limited by boundaries and water beyond which data are either unknown or non-existent. The outline is defined by the outline-package. The area outside the outline is usually left blank, but can be filled with any desired symbol through one of the electives in the map-packages (Figure 2.5).

[14] We are discussing thematic, not topographic maps.

Figure 2 4 Saddle Point The data-points are at the four corners of the square No electives of the map package are used Program SYMAP, production G Hayward, Simon Fraser University

In addition to the outline, the base-map may also have internal boundaries such as those of states The conformo-line-package contains the boundary-coordinates of the statistical areas for the choropleth map However, the State boundaries could also be produced by line-legends in the legend-package The legend-package gives the user the opportunity to add to the map information which should be repeated every time the packages are run with different values and/or electives Examples are an asterisk at the place of a city with the first letter or the full name on the side, strings of a symbol as streets, boundaries, rivers, etc, areas filled with one symbol for lakes, unpopulated areas, parks, north-arrows, large titles, etc A careful preparation of the legend-package can add considerably to the map (Figures 2 6, 2 7)

One can organize the SYMAP electives into four groups

1 Size, section, and scale electives
2 Gray-scale levels and symbolism electives
3 Interpolation electives
4. Miscellaneous electives (text information, restrictions, output onto tape, etc)

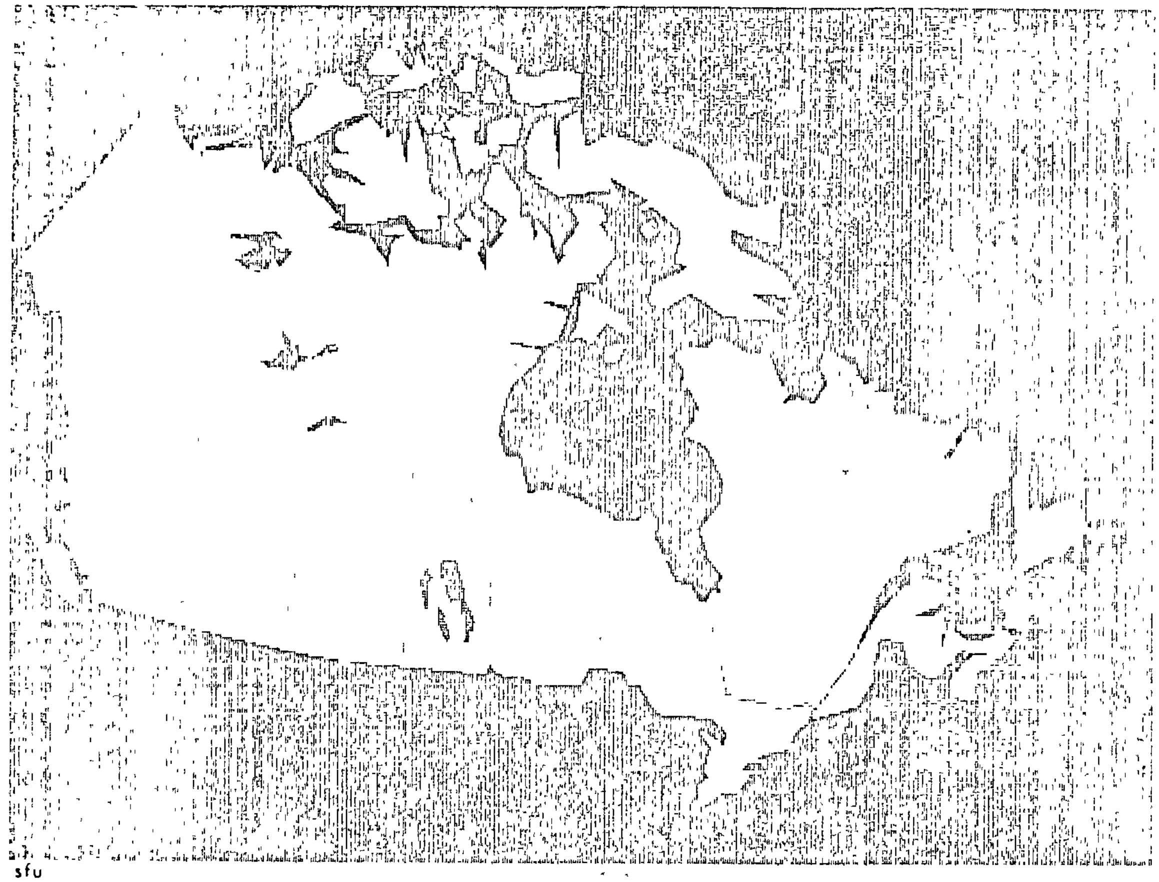


Figure 2 5 Canada Outline The use of the outline package and the size elective. Program SYMAP, data D Hatfield, Simon Fraser University

1 SYMAP is totally scale-independent. In other words, the measurement unit of the input does not in any way restrict the output One can define the coordinates in inches, miles, centimeters, etc. but also in non-distance units such as annual precipitation versus mean annual temperature The program assumes that everything is in inches merely to calculate its internal scale (printed out with the map), and produces either a map of thirteen inches (the default option) or of the size specified in an elective

The user of SYMAP is not bound by the area he wants to map by the area he inputs into the system By defining one elective he can take a section of the map and show it at any scale The section has to be rectangular, however, with a horizontal base

2 When a cartographer maps a thematic surface or terrain by choropleth or contour methods, he must employ a system of generalizing the array of data Class intervals are the numerical categories of such a system and should be thought of as being bounded by class limits A geographer knows that equidistant class-intervals are not always the ideal choice For example, in a rural area with one city, the population density is at least five times the average rural population density With the SYMAP default option of five class intervals, the rural population density would fall entirely into class one whereas the city would fall into class five with nothing for classes two to four Therefore, one might benefit from taking the lowest and the highest rural density and establishing four classes for rural population and the fifth for urban population Another possibility would be to produce more class intervals, for example ten, but then one would again get eight urban classes (for only one city) and two rural ones (Figure 2 8)

Both alternatives and any variation of them are possible One can also substitute the standard symbolism with one of his own choice

3 The data-points used to create a surface defined by a regular grid of z values do not have to be distributed regularly. SYMAP interpolates for a regular grid using a sophisticated algorithm, which has four main components [15] (Figure 2 9)

[15] See Shepard, D (1968) "A Two-Dimensional Interpolation Function for Irregularly Spaced Data," *Proceedings*, 23rd Nat Conf, Association for Computing Machinery, 517-524

Figure 2.6 New Zealand. The use of the otolegend Program SYMAP, data: T. K. Peucker

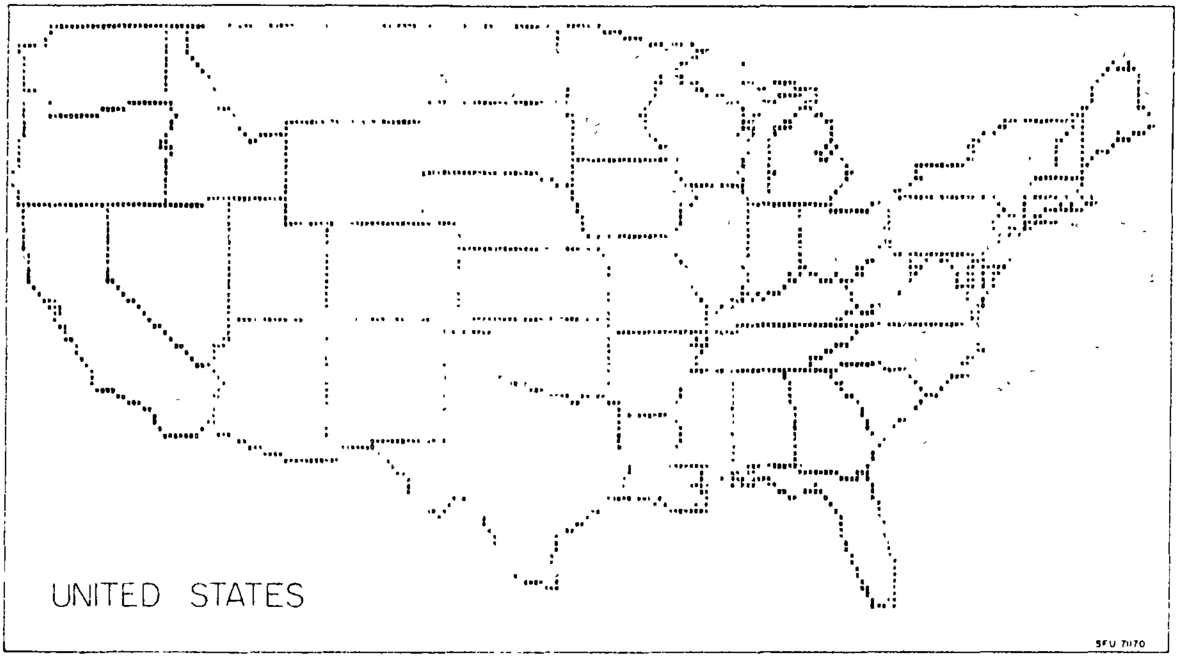


Figure 2.7 U S A All the boundaries are produced using the otolegend package Program SYMAP, data T K Peucker and R Mercready, Vienna, Virginia; production G Hayward

a) The point selection The attempt is to find the six data-points closest to the grid-point in question This is done by computing an "initial" search radius which encloses six data-points "as the map average." If the search results in between four and eleven points, the number is considered satisfactory, otherwise the search radius becomes shorter or longer respectively

b) Distance weighting The data-points chosen contribute with their z-values to the value of the grid-point inversely proportional to the square of their distance to the grid-point

c) Directional bias The distance-corrected weight of a point is reduced according to the location of a point behind another point with respect to the grid-point.

d) Additions to the algorithm account for computation errors and discontinuities of slope at the data-points

4 Finally, some general electives can nullify some built-in safeguards, allow for additional information beyond the title, and permit repetition of electives from previous maps, special output, etc One elective of special interest allows the copying of the map information onto tape for the purpose of producing perspective views and for allowing other operations based on a regular grid of z values. (Figure 2.10)

SYMAP is a highly flexible program which produces very acceptable results within the limits of a line printer. Eventually it may be replaced by an updated version which allows easier manipulation of the program itself, such as the inclusion of some special options for interpolation, the production of analytical maps, the approximation and filtering of data, and so forth. It should also be possible to reduce the program to a size which would make it faster and usable for smaller computers. However, little needs to be added from a user's point of view, which is quite a compliment for a system conceptualized almost ten years ago

An alternative to SYMAP is GEOMAP. GEOMAP was developed over a somewhat parallel time period as SYMAP but basically by only one person [16]. GEOMAP produces the same three types of maps, based on the quite different concept of neighborhood-smoothing, but is restricted in map width (to a page) and does not come close to the flexibility of the legend package and the electives of SYMAP However, it needs less core-memory. (Figure 2.11)

[16] D Steiner, University of Waterloo, some assistance to the operationalization of the program was given by Otto F Matt, University of Zurich

Figure 2 11 Switzerland–Three Maps by GEOMAP A Identification map, B. unproductive and uncultivated land as percentage of total district area, C forest as percentage of total district area. Since the program can print only a page width, the data have been rotated 90° to allow for a larger map Data D. Steiner and O F. Matt, University of Waterloo

III. SOME THEORY OF THE SURFACE

In topographic and thematic cartography, every point on a map can be identified by its x and y coordinates relative to an origin. In most cases it can also be related to one single quantitative or qualitative value. If the value is qualitative, the location gets a certain symbol or color according to the class of the value. In the quantitative case the set of all points can be considered as building a surface with the value as the "height" of the surface at that particular point. Such values can be population density, percentage of the population belonging to a certain ethnic group, atmospheric pressure, etc. Although these values often occur as classes (ordinal scale in the case of choropleth maps), the discussion will be restricted largely to values measurable by a continuous (interval) scale.

The Information Content of a Surface-Point

Some properties of surfaces with respect to sampling (i.e., taking values off the surface at specific points) will be discussed. One can record a surface in at least three different ways.

1. By keeping one dimension constant and going along the others in equidistant steps or following the surface along the constant dimension. The first case would produce a rectangular grid of z values, the second a horizontal contour line (isarithm).

2. By randomly sampling the surface (i.e., with random x and y jumps). In practice, this represents most of the statistical surfaces.

3. By coding only "surface-specific" points, that is points which furnish more information about the surface than only their coordinates such as peaks and passes.

Without delving into all the problems of sampling a surface, it can be said that a given regular grid of sampling points can depict only those variations of the data with "wave lengths of twice the sampling distance or more."[17] In other words, if we measure the elevation of a topographic surface at an interval of 50 feet, we will lose all the humps and bumps of 100 feet width and less. If we collect census data per census tract with three thousand inhabitants we will be able to find information on variations of groups with an average size of six thousand or more.[18] In this case we will not be able to acquire any knowledge about neighborhoods, for example, since neighborhoods are much smaller.

If one chooses a regular grid, one has to adjust the mesh width to the smallest desired variation in the map, even if this variation happens only in a portion of the map area, and one could get along with a much wider net for most of the area. In order to be representative in areas of rough terrain, the grid must be much finer than is necessary to represent regions of smoother terrain. These difficulties have led to the formulation of variable grid methods, in which the mesh size is small in rough regions and large in smooth regions. However, there is no application in geography using such methods as yet.

Points in a uniform grid are very definite with respect to their coordinates, they have a specific location relative to longitude and latitude or to any other orthogonal or non-orthogonal reference system. However, they are random with respect to the surface. In other words, if the surface is unknown, a regular grid is the best way to produce a sample of values (elevations, etc.)

The case is rare in which a regular grid of values is obtained with no previous knowledge of the surface available or without acquiring this knowledge during the process of digitizing. It would therefore be foolish not to use such knowledge for a more economical digitization. Several approaches to digitization using previous knowledge of the surface will be considered.

If a person looks at a map representing a topographic or thematic surface, he does not sample the surface randomly, that is, let his eyes jump over the map to find indications of the surface's "height" at random points, neither does he scan the map in equidistant steps, but rather he searches along the steepest slopes for local maxima or minima, very similar to

[17] W. R. Tobler, 1969, explains this in "Geographical Filters and their Inverses," *Geographical Analysis*, Vol. 1, 234-253, esp. 243.

[18] The relationship between sampling distance and depicted variation is stated in the sampling theorem which Tobler (1969b) explains as follows: "The theorem states that if a function has no spectral components of frequency higher than W, then the value of the function is completely determined by a knowledge of its values at points spaced ½ W apart. The lack of freedom of the function for variation from the prescribed path between sample points is a consequence of its lack of spectral components of frequency higher than W."

the hill-climbing technique in statistics and numerical analysis. In other words, he most probably looks for hills and valleys to get an idea of the surface-structure.

Any continuous surface can be approximated by a finite number of planes. In fact, this is how surveyors and photogrammeters usually perceive the topographic surface when collecting data. The surveyors measure only the corner-points of polygons (usually triangles) which cover the surface entirely. The photogrammeter traces along one constant axis and only records the "breaks" of the surface (i.e., the points where the surface's slope suddenly changes).

It is evident that this method can be of advantage for surfaces with little periodicity and large differences in the density of variation. A regular grid of values can, in such cases, be interpolated by finding the polygon inside of which each grid-point lies and then computing the "height" the plane has at the location of this point.[19]

In many cases one can acquire a higher level of preliminary information about a surface in the form of its "surface-function" or "autocorrelation-function." The autocorrelation function which will be discussed in a later part of this chapter in more detail, represents the mean "surface behavior."

This leads to the discussion of a set of features which are an expression of the surface itself, called "surface specific" points and lines. Three different pairs can be distinguished which have different relationships to the surface. These pairs are peaks and pits as maxima and minima of the surface, passes which are at the same time maxima in one direction and minima in the other and ridge-lines and course-lines as local maxima and minima connecting pairs of peaks and pits, respectively, and intersecting at passes (Figure 3.1).

Peaks are summits of mountains and hills, and as such are maximum points or positions of unstable equilibrium. Pits, on the other hand, are bottoms of holes and as such are minimum points or positions of stable equilibrium. Two neighboring peaks are linked by a ridge line, two neighboring pits are linked by a course line, both types of perpendicular lines being lines of local maxima and minima respectively. Passes, the crossing points of the two types of lines, are points of mixed equilibrium.

The type of equilibriums here is related to the type of vergency (convergency, divergency, mixed vergency) of flows along the surface. Thus a hill is an area of divergent flows bounded by course-lines which are lines of convergent flows. On the other hand, dales are areas of convergent flows bounded by ridge-lines which are lines of divergent flows. Peaks and pits have the same vergency of flow as their respective areas, whereas passes are points of mixed vergency.

[19] The "linear" interpolation is fastest since it only involves the solving of one linear equation per grid-point.

Table 3.1 gives a summary of the different surface-specific features with their vergency.[20]

Surface Behavior

Given a set of regularly or irregularly spread data points with their x, y and z values, the task is to map these points, that is, to create a surface continuously determined and single-valued. For practical purposes, it is sufficient to postulate that this surface is fully represented by a dense regular grid of points. If no other information than the position of the data points in three dimensional space is available, some assumptions have to be made about the "behavior" of the surface between the data-points. Again, for operational purposes, it may be assumed that these behavioral assumptions have to be made for the dense grid of points.

1 The discontinuous surface

The first assumption shall be that the value of any one data-point is representative for the values in its neighborhood. In other words, the surface is a series of level subsurfaces with different heights. Neighborhood in this context can be defined in two different ways.

First, a grid-point can accept the value of the closest

[20] For a detailed discussion of surface-specific features see Warntz, 1966.

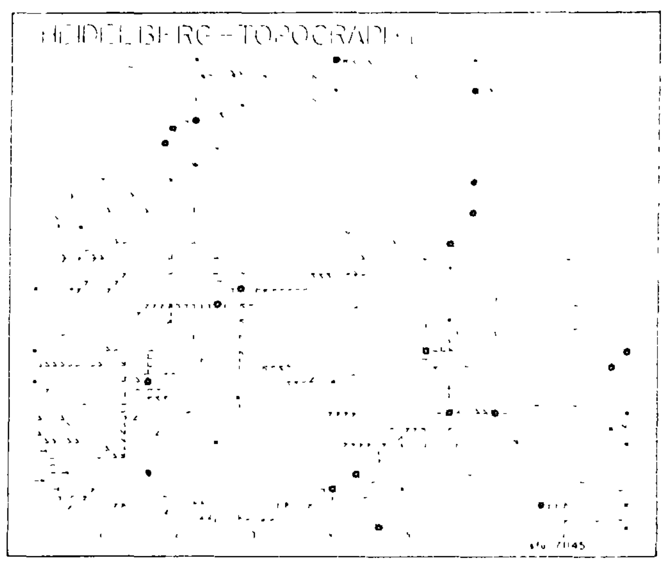


Figure 3.1 Heidelberg, Topography. A study of the surface of Heidelberg by a slope-analysis program. The specially marked points are peaks. Program: WATERSHED, data: T.K. Peucker, production: W.D. Rase.

TABLE 3 1

Dimension \ Vergency	Divergency	Convergency	Mixed Vergency
Point	Peak	Pit	Pass
Line	Ridge	Course	
Area	Hill	Dale	Territory

data-point. The type of map resulting is often called a "proximal" map and used in several mapping programs since the procedure to find boundaries between two areas of different values is very simple.[21] (Figure 3 2)

Second, neighborhood can be defined by areal extent, that is, each datum has a given subarea with known boundaries associated. The resulting map, the choropleth map, has been well studied and used in both traditional and computer cartography. It is highly suited for a manual presentation of statistical data. However, for a presentation by the computer, it is not the most efficient approach. Time and costs for coding the boundaries, storing the areal extent of the sub-regions and shading the map often make it inferior to the method of producing a continuous surface.

[21] In SYMAP, the closest data-point is found for each grid-point and its value is assigned to the grid-point In GEOMAP, a field is created around every data-point which is expanded parallel to the other fields until it hits other fields at every direction of expansion The results are very similar

2. The continuous surface through data-points

The second assumption is that each data-point represents a sample of a single valued continuous surface. The surface can be continuous in the first derivative or not (i.e., have "breaks" or not). However, the question of continuity becomes irrelevant with an increasing number of data-points in a discrete pattern of grid-points.

Discontinuity in the first derivative can mean that the surface is approximated by a set of planes as discussed in an earlier part of this chapter This represents the simplest, fastest and often the least misleading interpolation method at

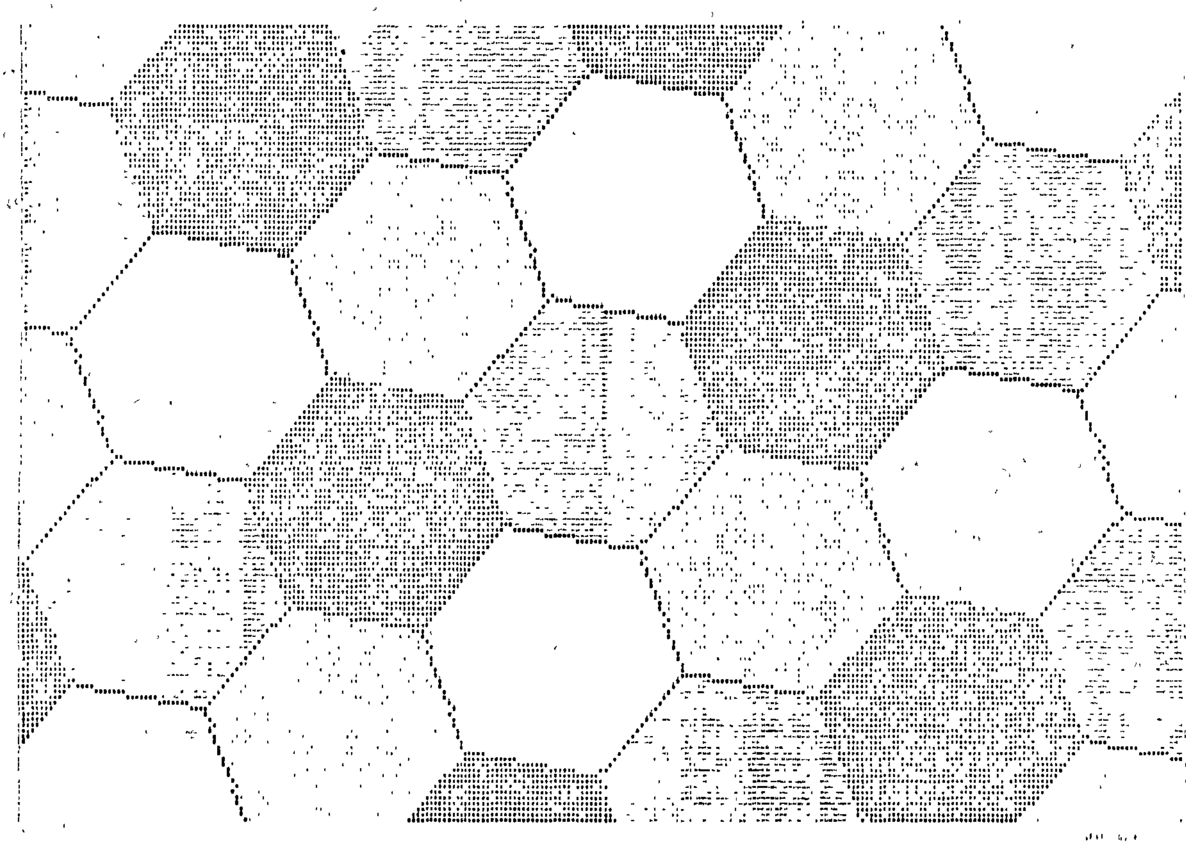


Figure 3 2 Hexagons Produced by the proximal-map electives The data-points are arranged in regular triangles Program SYMAP, data W D Rase hand if the planes are represented by triangles [22] most of the contouring algorithms in use today are based on a rectangular grid divided into four triangles by computing the center point of the rectangle

Continuity in the first and higher derivatives with the surface passing through all data-points at the same time is usually accomplished by distance-weighting functions The z values of a number of closest points are weighted by the inverse of the distance with an exponent which ranges from one to four.[23]

3 Trend surface

The third assumption is that the sampled values include measurement errors, that is, that the surface does not necessarily pass through all data-points It is also assumed that the surface is smoother than or at least as smooth as the variation of the data-values indicate, and that the errors sum up to a minimum under one of several conditions discussed later and that these errors are random (i e , no clusterings of negative or positive errors)

The first and most frequent condition is that the surface can be expressed by a polynomial or trigonometric function of x and y The minimum condition is that the sum of the squared differences between the actual and the computer surface at the data points be at its lowest possible value The resulting surface is usually called "trend surface" in geology and geography and "response surface" in statistics [24] (Figure 3 3)

Another minimum condition is that the sum of the absolute differences between the actual and the computed surface at the data points be at its lowest possible value This condition is accomplished by the method of linear programming The method has some advantages over the trend surfaces in those cases in which the distribution of data points is highly irregular However, it is relatively expensive for computation [25]

Another approach to the creation of a surface which is smoother than the original surface is by filtering out the high frequencies This will be discussed in more detail in Chapter IV

4 Autocorrelation function

Up to this point, no knowledge or only little knowledge of the general shape of the surface was used for its explanation It has been mentioned earlier, however, that very often a fair amount of information is available either from other similar surfaces or from the surface itself as one goes along coding it.

First, the areas of different roughness of terrain might be known. This will influence the distribution of observation points since more points are needed to reach a certain accuracy of surface representation in rougher terrain than in shallow areas This then leads to sampling grids with varying point density Second, the so-called breaks might be detected easily and thus enable a more surface-oriented sampling

The most powerful knowledge of a surface, however, is the knowledge of its autocorrelation This function gives the average degree of similarity between all pairs of values as a function of the distance between the respective points For the computation of the autocorrelation function the mean plane (i e., the average of all heights) has to be computed first and then subtracted from the surface, yielding positive and negative values The autocorrelation function then indicates the expected similarity of two values depending on the distance between them Diagrams 3 1a and b show two typical autocorrelation functions in one dimension (distance) and Diagram 3 1c gives an example of a two-dimensional function They start with 1 at distance zero indicating that the values of two points are expected to be identical if their location is identical which is tautological if a continuous surface is assumed and go to zero at an infinite distance In other words, there is no similarity between the values of two points if they are separated by large distances

The application of the autocorrelation function for the estimation of the value of a point (interpolation) at a certain distance from a data-point, is based on the idea that the value is most likely between the data value and the mean plan value (i e , that the surface always "converges" to the mean plane) The relative decrease in the value from the data-point with increasing distance is given by the autocorrelation function which takes the data value as one and the mean plane as zero (standardized data) and then multiplies the autocorrelation function with the difference If more than one data point is used for the computation of a value, the procedure becomes more complicated But the basic idea of the use of knowledge of the surface-behavior for a better definition is maintained (Figure 3 4)

[22] See Bengtson B E and S Nordbeck (1964) "Construction of Isarithms and Isarithmic Maps by Computer " *B I T Nordisk Tidskrift for Informations Behandling* Vol 4 87-105

[23] For example there exist versions of SYMAP which allow the exponent to be altered from 2

[24] See Chorley and Haggett (1965), Harbough and Merriam 1968 and Box, G E P (1960) "Fitting Empirical Data," *Annals, New York Academy of Science*, Vol 86 792-816

[25] See Dougherty E L and Smith (1966) "The Use of Linear Programming of Filter Digitized Map Data," *Geophysics* Vol 31 253-259

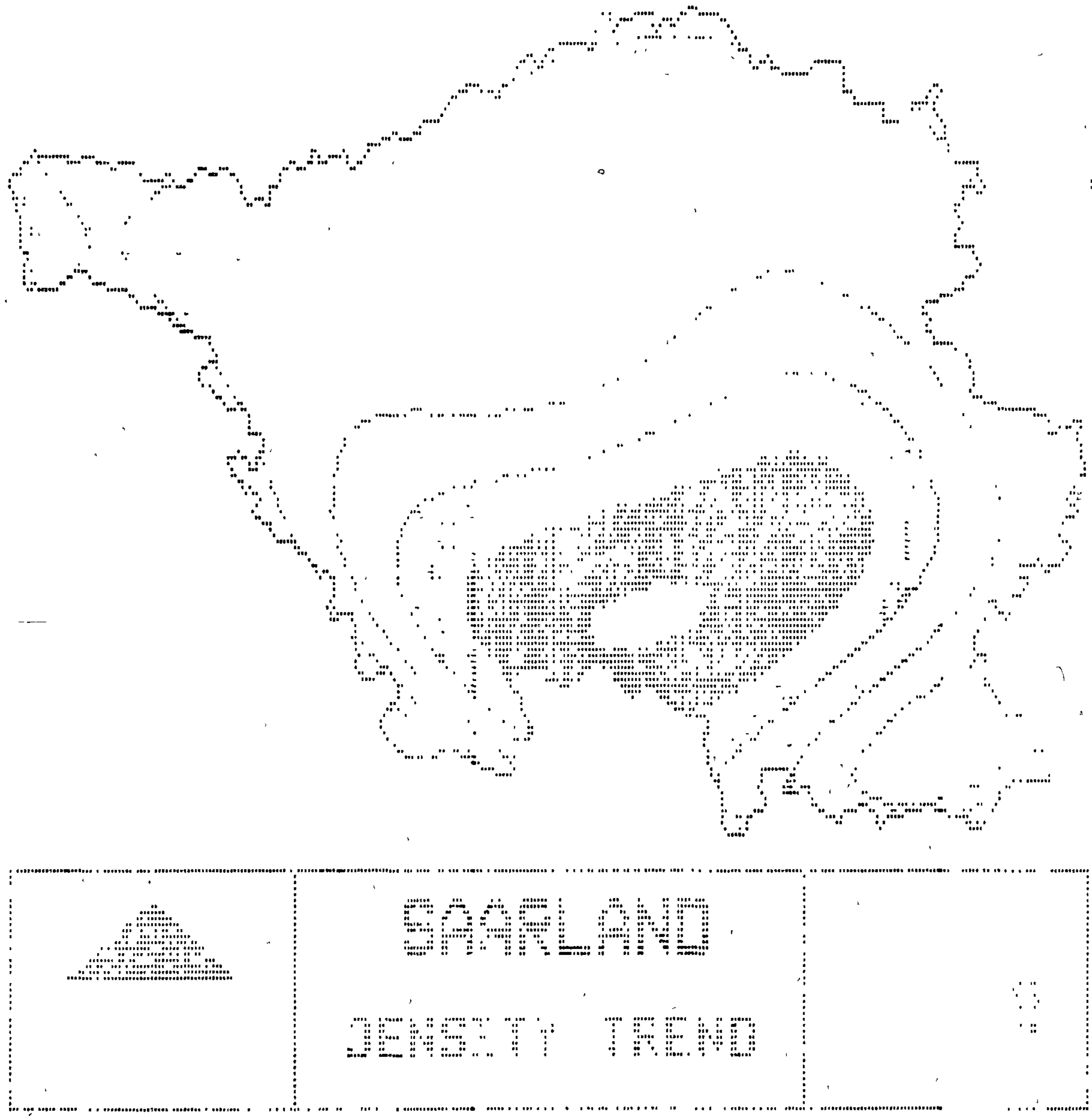


Figure 3.3 Population Density, Saarland A fourth-order trend surface.
Program SYMAP with trend-option, data W D Rase

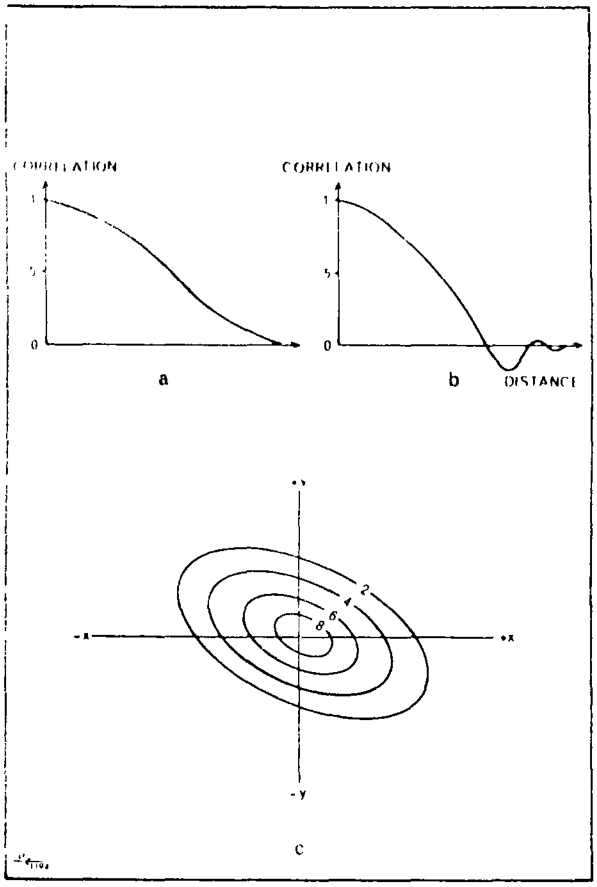


Diagram 3 1 Autocorrelation Function

Surface Features and Information Hierarchies

It may be assumed again that a surface is determined by a dense but discrete grid [26] The object may be taken as that of describing a surface "satisfactorily " Obviously, when going through the process of describing a surface, some people will be satisfied earlier than others. If one defines description as the reduction of uncertainty then one could say that some people have a higher error tolerance than others, usually because the purpose of the surface study varies.

There are several points on a surface which represent surface features with a higher information content than the average surface point For optimal surface description, these features should be sorted out with decreasing information content so that the rate of error reduction becomes smaller with added features The importance of the individual features might vary from case to case, but for many applications the following hierarchy might prevail

[26] Discrete in this context, relates to the non-continuity of the sampling process

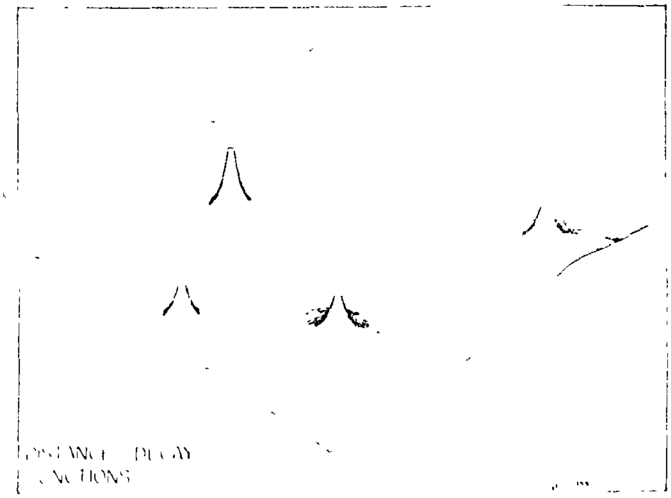


Figure 3 4 Distance Decay Function With Four Centers Programs SPATI-UN, SYMVU, production T.K Peucker

1. Surface behavior

In some cases, the pure description of the nature of an area's terrain is enough to satisfy one's curiosity or to define a relatively simple surface completely If, for example, one is told that a certain area is flat, has a rolling terrain or a mountainous structure, this might be all he wants to know These terms represent autocorrelation functions of a very summary type, not directly quantifiable, but distinct quantitative conceptions The mathematical types of the autocorrelation function are one-dimensional functions or two-dimensional functions in which the correlation can drop with varying speed in different directions

The functions can be expressed in singular equations or in vectors and matrices showing the function at discrete points. The autocorrelation function usually plays an important part in the explanation of a surface.

Other examples of surface behavior functions are distance decay functions, polynomial and harmonic trend surfaces (without the constant element), etc.

2 Surface height

Although the surface behavior might offer some indication of the height of the surface, this might be misleading For example, flat terrain suggests low altitude, but level areas also can be found in high altitudes In many cases, a height indicator must be added to the surface behavior The absolute element in the trend-surface functions, the value of the highest point on a surface, or the mean plane in the autocorrelation function are examples

3 Reference surface

The mean plane can be extended to another surface feature of some more geographic significance It could also be viewed as a reference plane to which the surface converges if any other information is lacking. But then, one might ask, does it have to be a horizontal plane as requested in engineering applications? In topographic terrain, one could think of a plane which follows the general drainage system. it could be the linear trend-surface of the terrain which would indicate the general water flow if the area is limited to one drainage basin.[27]

Other reference planes can be imagined, such as a zero plane to which distance decay functions in density, gravity and potential models converge and infinity planes to which transportation functions in cost-surfaces or isochrone maps converge

[27] The reference "plane" can also be of higher order than the first, but it is logical that it has to be simpler than the actual surface

4 Surface-specific points and lines

Surface-specific points have a higher information content than surface-random points Surface-specific points, however, exhibit a hierarchy among themselves These points are, in decreasing order, peaks and pits, passes, and ridge and course lines These features not only tell their values, but also give some idea about their surroundings In the environs of a peak, we can say that everything else will be lower, and with the autocorrelation function we can say to what degree.

5 Grid-points

A surface-random point on the other hand, gives only its own height and nothing more In other words, if a surface were represented by a grid of 100 by 100 points, knowing one point would bring us 1/10,000th closer to a full knowledge of the surface. Unfortunately, no studies yet tell us how many points one surface-specific point of a certain kind stands for.

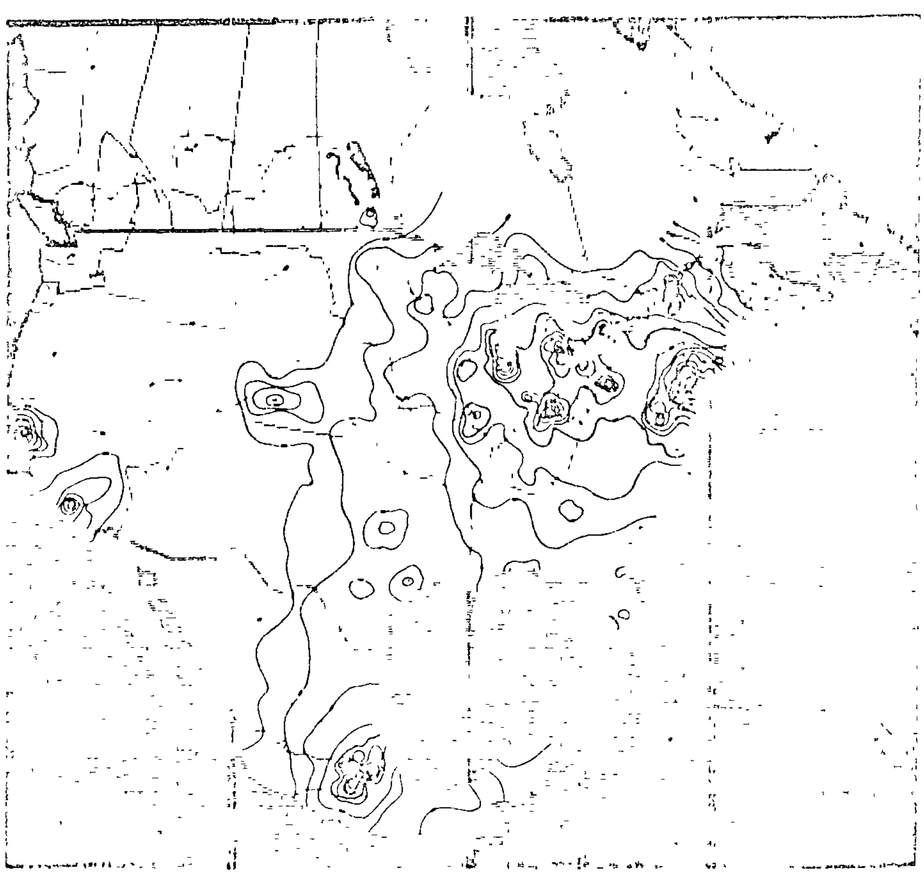


Figure 3 5 North America, Population Potential The concept of surface-specific points and lines can be exemplified by the population-potential map Source Douglas, David H., *Illustration of Regional Disparities in North America by the Use of the Gravity Model* M A Thesis, Carleton University, September 1969

6 Hierarchy

Such a hierarchy of surface features can supply a framework for a spatial linguistic system of surface description The procedure of information compaction– the reverse of what had been done above - should exemplify how the process of surface description could be organized on a linguistic basis Of course, there are many ways to do this, only one of which will be shown here

By processing every grid point in conjunction with its neighborhood, surfaces of slope gradients[28] and directions can be computed By using the direction of steepest descent as the criteria for vergency (convergency, divergency, mixed vergency), one can find the different surface-specific points and lines Surface-specific areas (hills and dales) can then be defined with different detail peaks and pits alone will give the relative location of hills and dales, respectively, passes will indicate the "points of indifference" (i e , the points shared by two areas of the same type), and ridge and course lines will determine the exact boundaries of dales and hills, respectively. (Figure 3 5)

Having acquired information about individual surface-specific areas, one can go on to structure the surface as a whole. The reference surface gives the general trend of the surface, whereas the surface-behavior function (the auto-correlation function) supplies the average local variation

This framework has to be elaborated upon in order to be operational. However, with some procedures for feature analysis and the knowledge of some typical surfaces, it should be possible to develop a surface recognition and classification system similar to the procedures of picture recognition and classification. (Figure 3.6)

[28] = the first derivative of the surface

Figure 3 6 North America, Sales and Population Potential The map shows the weighted difference of the population potential and the retail sales potential (Harris' market potential). The result is a map of purchasing power disparities (darkest area = negative) Source as in Figure 3 5

IV. SURFACE PROCESSING

In most cases spatially collected data are not, or should not be, the end-product of an inquiry. For example, a map of population density does not tell us very much apart from the mere distribution of densities. Too many irritating undulations of the surface in small areas occur (high variation), and we need some additional information to make sense of the distribution of densities.

The first defect is related to the problems of smoothing, filtering, approximation, surface-generalization, etc, whereas the second is related to problems like map comparison, spatial correlation, overlay-analysis, and generally spatial analysis It is not advisable to go into much technical detail since the student typically will be able to use program-packages and does not have to worry about aspects of computation, rather he has to study the theory and the implications of the different methods.

Geographers are not the only persons who have worked intensively in the area of surface processing. Geologists, ecologists, computer scientists, and others have developed a variety of concepts which the geographer can use for his own purpose These disciplines often refer to surfaces with different names, the most frequent being "picture" The term picture suggests a flat square object whose appearance varies from point to point. The variation can have one component (black and white) or several (color), but picture analysis and processing is largely restricted to the black and white case.[29]

The affinity between surfaces and pictures is obvious. Both have one finite value at each point (the gray level, in the case of the picture). Both are assumed to be continuous, single-valued, etc.[30] It is therefore possible to use all the development in digital and optical picture-processing for surface analysis which is usually summarized under the term of filtering. However, geographers sometimes use the term filtering for a wider aspect of surface processing.[31] It is extended to all those methods which give a more general idea of a sometimes very confusing empirical surface In the context of information theory, all the methods aim at reducing the information content of a surface while simultaneously retaining the "important" characteristics and discarding the unimportant ones. (Figure 4.1) It is therefore crucial that the researcher

1. know what is important and know it in a fashion which can be implemented.
2. use procedures which are adequate to the expected results.
3. receive the result in a form which he is able to interpret.

Three types of surface processing are those which strive to

- create a well-defined surface (as a function or a regular grid of values).
- eliminate or emphasize certain variations of the surface for a better interpretation
- and compare two or more surfaces.

These types can occur in combination, and several can often be satisfied with one procedure. They are therefore not independently discussed but will be mentioned during the presentation of the different procedural models.

Prediction

The first point relates to the problem of interpolation and extrapolation or prediction as the two procedures are sometimes called[32] The general problem of two dimensional interpolation is to estimate the values of a surface at every intersection of a regular grid (the grid-point) from the values of a number of surrounding given points (the data-points). The type of interpolation depends highly on the assumptions the researcher has made about the surface, and the purpose of the interpolation. Although these two aspects interact highly, they have been treated independently in this paper. The underlying assumptions have been discussed in the previous chapter and there is little to add

It is, however, important to keep in mind that the purpose of the interpolation often determines the underlying assumption or at least implies a restricted set of

[29] This discussion follows the more detailed definition of pictures in Rosenfeld, 1969, Chapter 1

[30] Rosenfeld, 1969, shows that any picture is "indistinguishable from an analytically well-behaved function," and Nordbeck and Rystedt show the same for maps as long as they can be transformed into density maps See Nordbeck, S and B Rystedt (1970), "Isarithmic Maps and the Continuity of Reference Interval Functions," *Geografiska Annaler*, Vol 52, Ser B, 92-123

[31] See Tobler, (1969b)

[32] See Heiskanen, W A and H Moritz (1967), *Physical Geodesy*, San Francisco, Chapter 7

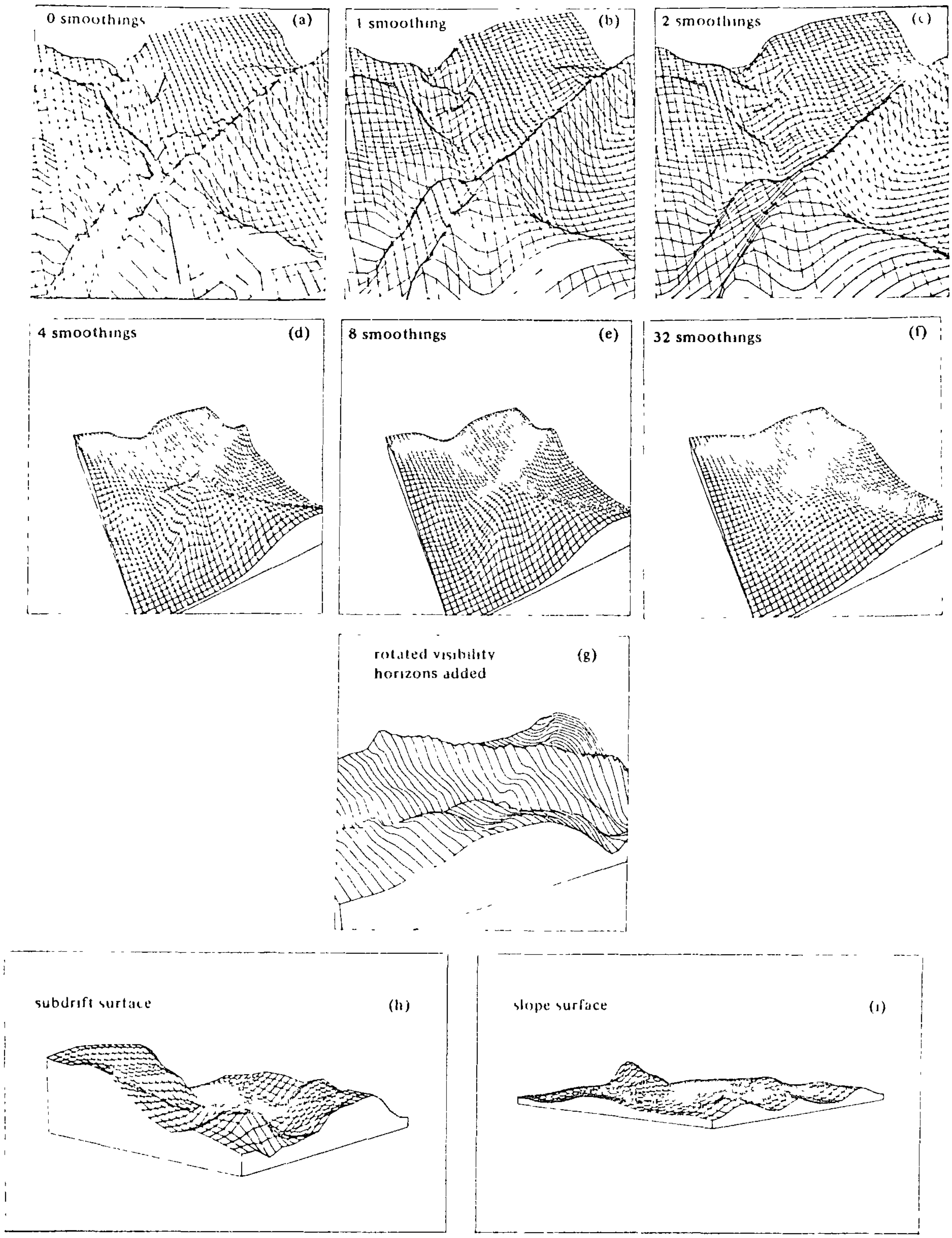


Figure 4.1 Surface Processing. Graphs a, b, c, show Leamy Creek in a close-up with 0, 1, and 2 smoothings; d, e, f show the same surface with 4, 8, and 32 smoothings. Leamy Creek is rotated in graph g and visibility horizons are added. Graphs h and i show a subdrift surface and its slope surface, respectively. Program: VIEWBLOK; data: graphs h and i, P. Johnson, University of Ottawa; all others D. Douglas, University of Ottawa.

assumptions. It is therefore important to keep the latter in mind when discussing the former.

The choropleth and proximal map approach is often used for a first look at data since it offers a quick and easy overview especially when used manually. However, for an evaluation of data by machine, linear or even other types of interpolations might furnish a more acceptable and less expensive result. A surface without "steps"--although "breaks" might occur–represented by contours frequently gives better possibilities for interpretation It must be admitted that it also opens the door for misinterpretations but the advantages usually prevail (Figure 4.2)

In the case of the linear interpolation, the three closest data-points are searched for a grid-point. Some procedures provide for the special situation that all the three points are on a straight line or close to it in which case a fourth point is searched for. The result of this type of interpolation is a surface which consists of a set of triangles with the data-points as corners A type of linear interpolation which produces somewhat smoother results is based on convex quadrilaterals and interpolates between opposite sides [33] Another linear method uses two overlapping triangles and averages the two resulting values for a grid-point It is clear that the result differs fundamentally from those of the previous methods The obtained surface does not pass through all the original data-points and the result will be smoother than the variation of the data-points would suggest.

The first group of interpolation methods--which included the already extensively discussed distance-weighting method–repeats every hill and dale given by the data. The variation created by these methods is equal to, or even higher than, the variation of the set of data-values The second group of methods which we might call approximations, smooths the surface (i e , produces a surface with less variation than the original one)

The most frequently applied model of surface approximation is that of the polynomial "trend surface",[34] as it is called in geology and geography (or "response surface" in statistics) The objective for the approximation is usually the least square criterion, that is, the condition that the sum of the squared deviations of the data points from the computed surface be minimized In other words, to take the actual value of a point, compute its estimated value by entering its x and y position into the function, subtract one from the other and square the difference, the condition is that the sum of all the squared differences has to be minimum. (See Figure 3.3).

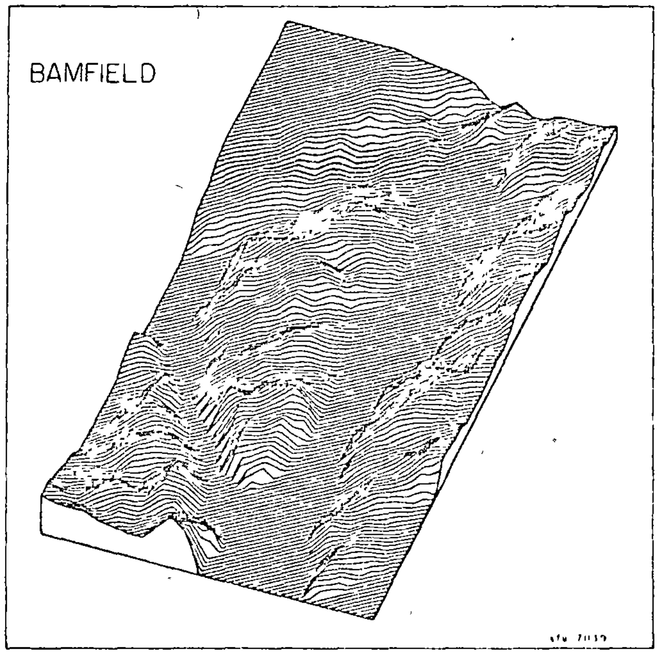


Figure 4.2 Bamfield, Linear Interpolation The elaborate SYMAP-interpolation algorithm is used only for every second point down and third point across as default This example shows that this can create unpleasant patch-surfaces Program SYMAP, data G Neilly, McMaster University, production T K Peucker

Another objective for the approximation is the condition that the sum of the absolute deviations be a minimum, again another is based on the minimum curvature-criterion. This method, called spline approximation, is usually applied in a series of surface patches[35] and therefore has a similar effect to that of the location-invariant smoothing methods. The size of the patches is proportional to the size of the moving matrix in its smoothing results

Polynomial trend surface has been introduced into spatial analysis by geologists whose studies of sedimentary layers are expected to be simple, and easily and meaningfully described by polynomials of the second or third order. In geography, however, especially in human geography, the frequent use of polynomial trend surface often seems to be based on the lack of better programs. Tobler (1969) writes

> There appears to be a temptation to apply this model rather indiscriminately to all sorts of geographical situations, for no better reason than the fact that a computer program is available It seems rather foolish when applied to social data in an urban area, for example, since there is no theoretical reason for believing such a model to be useful This does not

---

[33] Kaplan, M A and R A Papetti (1971) "A Note on Quadrilateral Interpolation," *Journal of the Association for Computer Machinery*, Vol 18, 1971, 576-585

[34] See Chorley and Haggett, 1965

[35] See Holroyd, M T and B K Bhattacharyya (1970) "Automatic Contouring of Geophysical Data Using Bicubic Spline Interpolation," *Geological Survey of Canada*, Paper, No 70-55

mean that there are no geographical situations for which this model might serve, but only that it is not always appropriate

Another least-square approximation is that by trigonometric functions (Fourier series).[36]

One reason for the frequent use of the two least square approximations is their ease of computation However, their results can be very misleading if one does not provide for a relatively even distribution of data points, otherwise, the surface is determined too much by clusters of data-points and deviates extremely in areas of a sparse distribution of points The danger is somewhat reduced by Fourier series because of their periodicity.

The periodic character of harmonic function also gives more explanatory power Wherever one wants to test theories with periodic features, Fourier series fit the purpose very well Many spatial theories imply periodic features such as settlements, diffusion and the like. (Figure 4 3)

Least-square approximations require a minimum number of data-points to produce non-trivial fits (i e , fits with a less than perfect correlation between the actual and the computer surface) This minimum number is equal to the number of elements in the function. A plane—a polynomial of the first order—therefore, needs three points for definition (a table can stand on three legs, for example). The number of points beyond the minimum number are the "degrees of freedom" of the surface, a basis for different confidence measures for the surface.

Most spatial theories have a distance-decay component.[37] It would therefore be appropriate to develop methods which allow testing of these theories A procedure for this purpose has been developed recently [38] It first fits a linear least square cone for every data point and selects the one with the highest explanatory power (with the lowest sum of residuals) It then repeats the process with the residuals, etc. An expansion of the procedure[39] allows

[36] See Harbaugh and Merriam (1968)

[37] See Commission on College Geography Resource Papers No 1, *Theories of Urban Location* by B J L Berry, No 4, *Spatial Diffusion* by P Gould, and No 8 *The Political Organization of Space* by E Soja

[38] See Cassetti E , and R Semple (1968) "A Method for the Stepwise Separation of Spatial Trends," *Discussion Paper No 11*, Michigan Inter-University Community of Mathematical Geographers

[39] Tobler, 1970

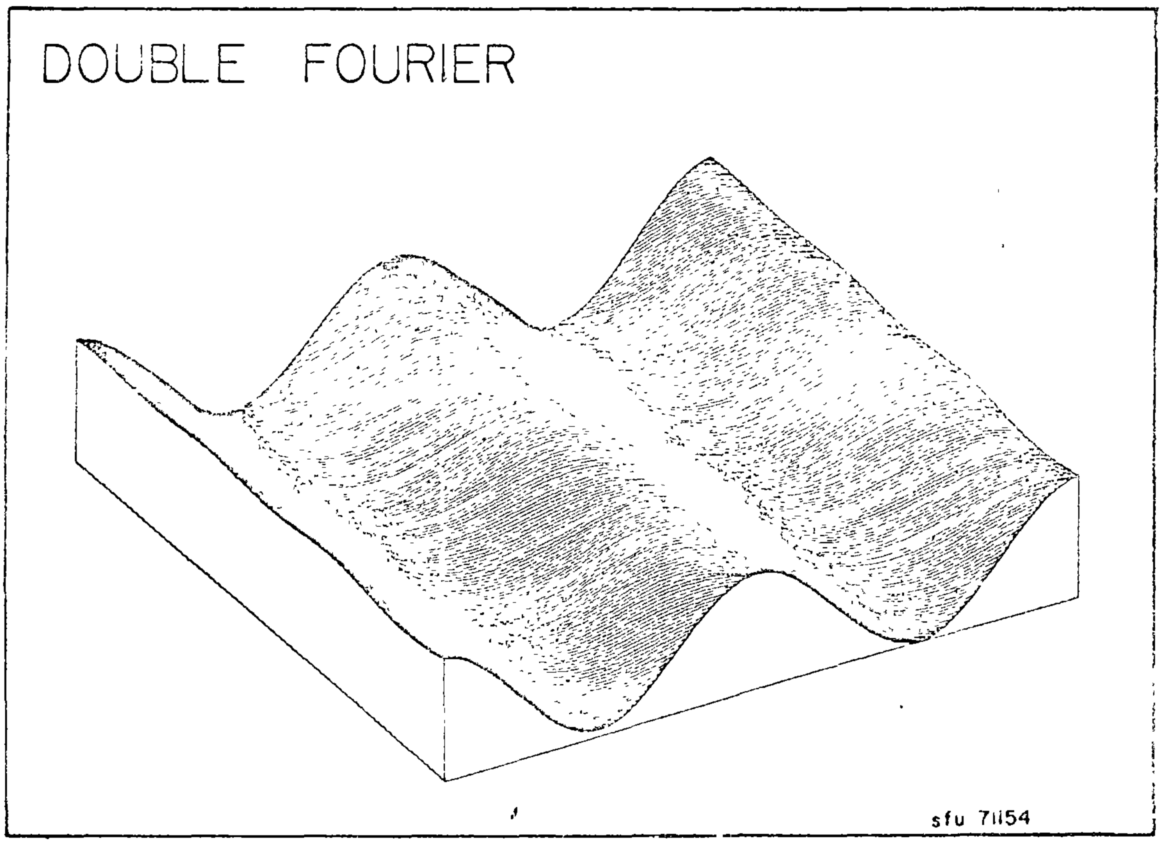


Figure 4 3 Double Fourier Series Program SYMVU, production. W.D. Rase for transformation of the data for non-linear cones (Figure 4 4)

Spectral Analysis

Closely related to harmonic trend analysis is spectral analysis. Both are based on Fourier series and fit a surface by a series of trigonometric functions with increasing frequency; the first to create a smoother object and the second to obtain a frequency-breakdown of the surface. But the spectrum can serve as a basis for many different purposes. Spectral analysis has long been used in the study of time-series but the applications of two-dimensional spectra are relatively rare because neither were there any good descriptions of the method nor programs available.[40] The spectrum displays the frequencies and their amplitudes which contribute to a surface. Therefore, spectral analysis is a way to disaggregate a surface into all its wave components. (Diagram 4.1) If the surface is periodic, then the number of contributing frequencies is finite and can be very small The spectrum is then called discrete and has the form of Diagram 4.2a.

In a non-periodic surface, theoretically an infinite number of frequencies contribute to the spectrum which has the form of Diagram 4.2b, and thus is continuous. In practice, the number of frequencies is still finite since the data-base is discrete, but it is usually very large. Of course, the one-dimensional case can be expanded to two (and more) dimensions. With two dimensions the discrete spectrum would become a map of points and the continuous spectrum would be a surface The computational methods involved in both types are very similar, but their statistical confidence is very different, as the figure might suggest

It should be mentioned that one data-set can produce many different spectra depending upon the various filters and windows used in the computation. Therefore, the student cannot be satisfied with one single analysis but must experiment to obtain the best frequency analysis

[40] This has changed with Rayner's (1971) very thorough introduction into the different aspects of spectral analysis

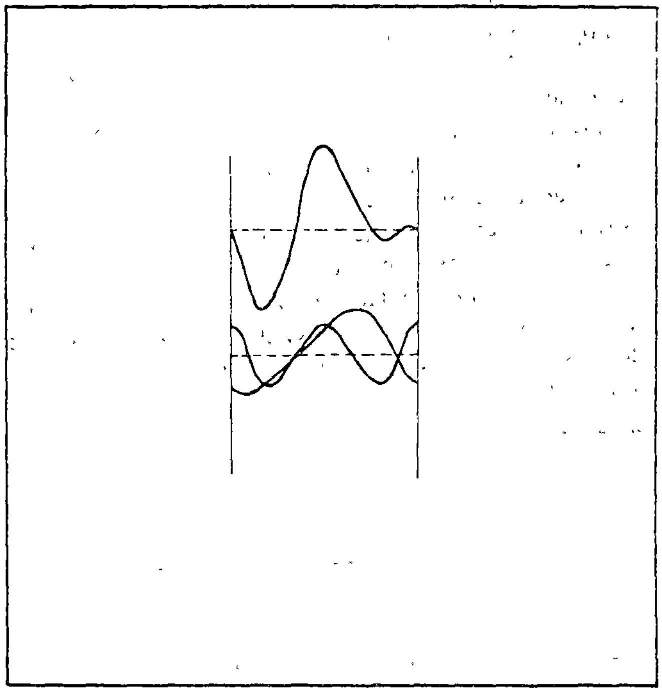


Diagram 4 1 Wave Components

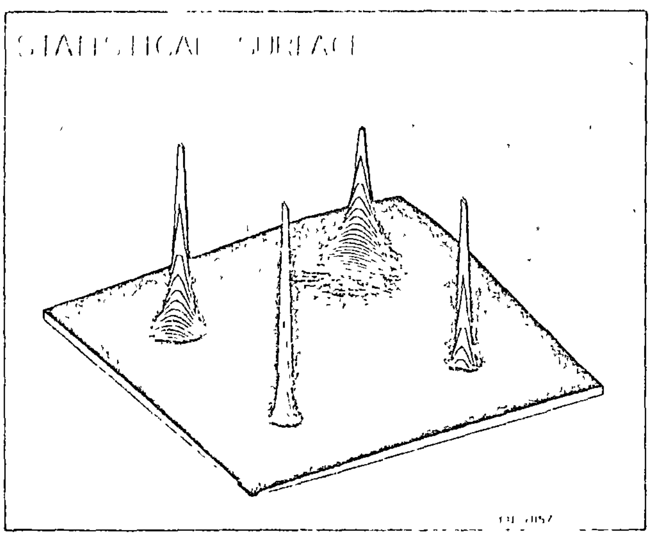


Figure 4 4 Statistical Surface, Distance Decay Function Programs SPATFUN, SYMVU, production T K Peucker

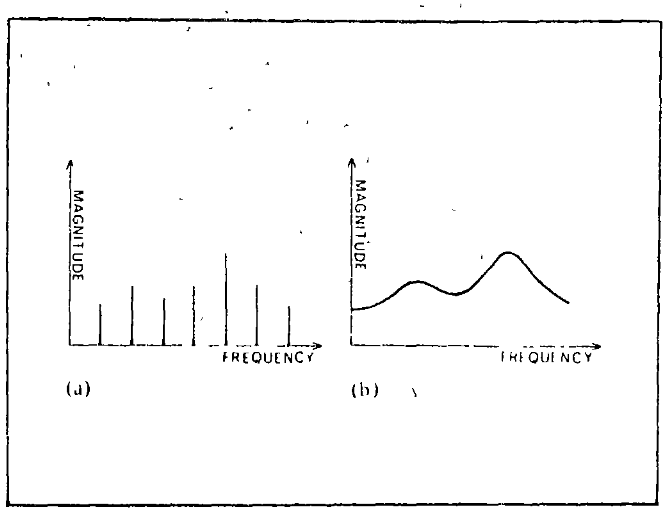


Diagram 4 2 One-Dimensional Spectrum

possible. One area of the analysis centers on the spectrum itself. For example, if one had to find out whether there are any periodicities inherent in the data, the analysis would indicate abrupt peaks in the spectrum. On the other hand, a fairly even distribution of variance along the frequency axis indicates that the data were generated by a random process. The variance spectrum (i.e., the spectrum of standardized data) is preferred by many researchers.

The study of the relationship between two spatially dependent variables is another area of interest. The cross spectrum gives the scale breakdown of regression parameters such as covariance and correlation coefficient.[41] If two variables are correlated at one scale and not at another, a regression analysis would average the two out and give a low overall correlation which would be geographically meaningless.

[41] Rayner, 1971, p. 8.

Other applications of spectral analysis use the spectrum as an aid to computation. Rayner mentions the areas of interpolation, filtering, differentiation, integration and pattern recognition. Interpolation might be done directly by applying the inverse Fourier transform to an augmented matrix or by using the autocovariance or autocorrelation function as a weighting function for irregularly spread points. In this case, the autocorrelation function is the surface-behavior function which indicates the surface-development between the points. (Figure 4.5)

Filtering with the spectrum is a very elegant way of extracting data from a surface. By discarding all those frequencies undesirable for the purpose of the study, one can extract the pertinent information. Smoothing would be performed by multiplying all high frequencies by zero; removing trends for the study of residuals would mean multiplying high frequencies by one and low frequencies by zero. The detection of high rates of change and the

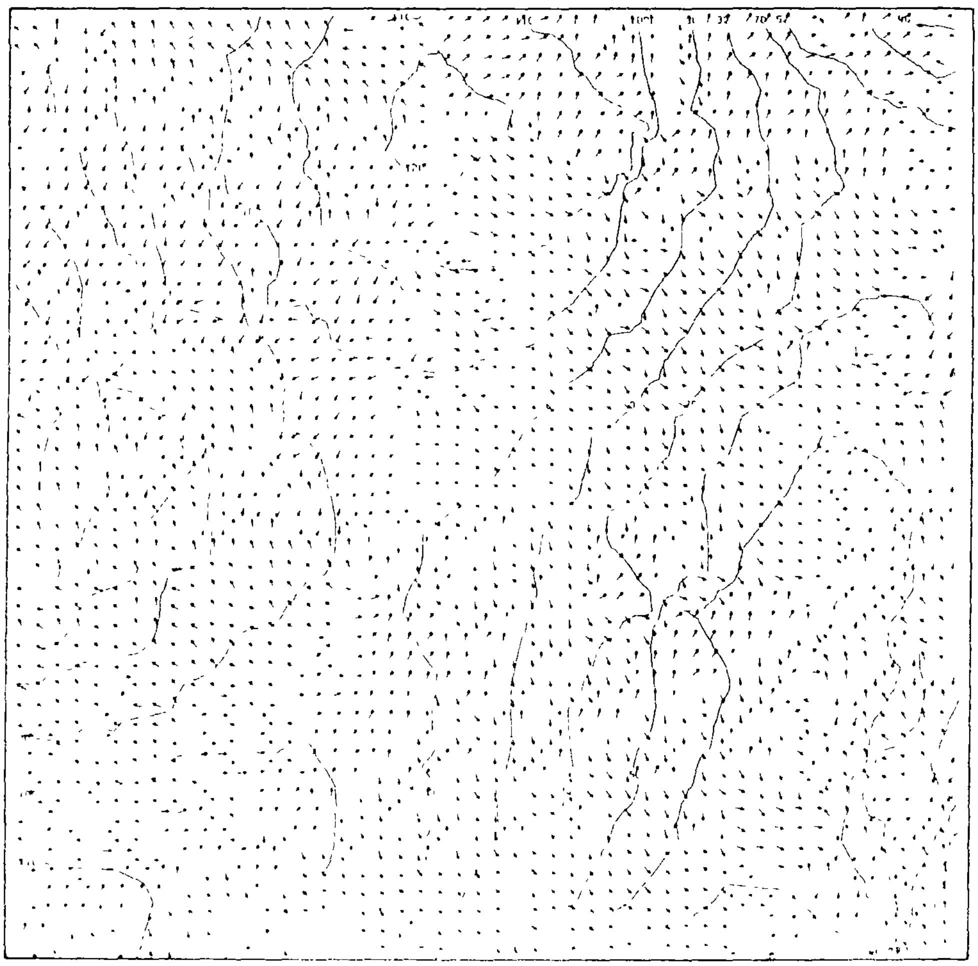


Figure 4.5 Grouse Mountain Slope Program SLOPE, data D Mark, production W D Rase

enhancement of boundaries requires a differentiation of the surface. Slopes (the first derivative), changes of slope (the second) etc., can be obtained by the filter process in the spectral context. (Figure 4 6)

Pattern recognition could also be considered as a filtering process in the spectral domain. The comparison of a known pattern with an unknown to extract correlations between the two involves a multiplication of the two spectra with an inverse transform.

Picture Processing

Most of the digital operations of spectral analysis can also be performed optically and the optical approach has clear advantages in problems where the original data are already in picture form (i.e., where gray tones correspond

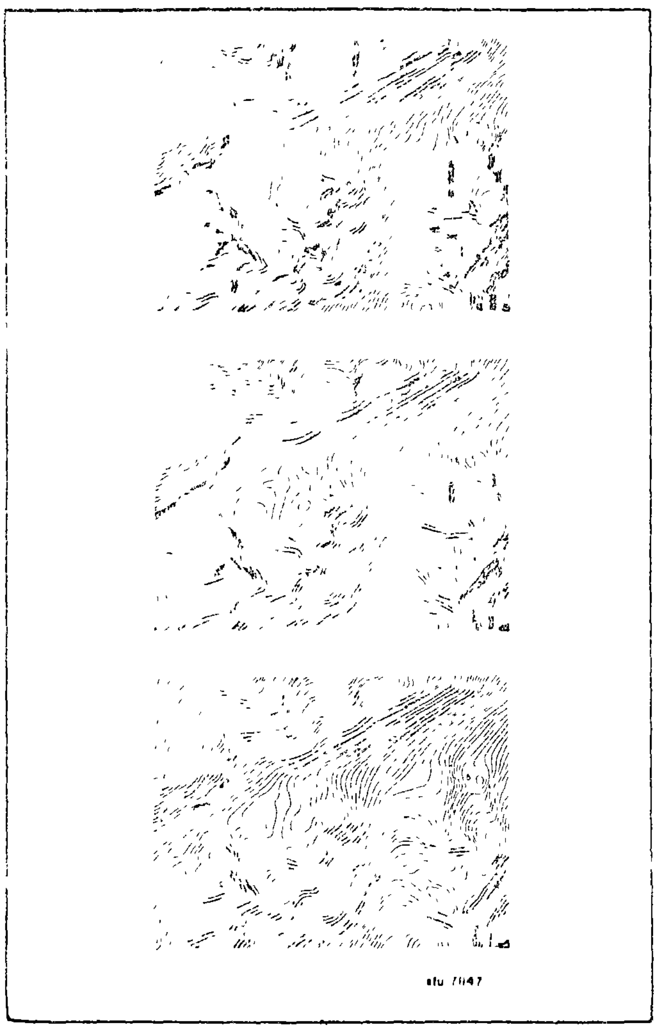


Figure 4 6 Lake Louise West, Smoothings. This example shows that filtering has a scale-reduction effect. Program INCLIN, data T.K. Peucker

to heights), because of the relative ease with which the researcher can modify filters in response to the observed output.

Optical operations are position-invariant operations, that is, their effect on one point does not depend on the position of the point in the picture. Digital position-invariant operations are frequently used in the area of picture processing.[42] Therefore, it is not surprising that spectral analysis and picture processing largely share their theoretical basis, although the procedural implementations are very different. Some of these operations are explained below.

Among *geometrical operations*, shifting is position invariant, whereas rotation and scale change are not, since they change one point's position with respect to its neighbors. Shifting operations, very frequent in pattern matching problems, can also be of use in cartography, for example in a study of spread-functions through time.

Point operations change the gray level at each point in a manner that does not depend on the rest of the picture. Of course, neither does it affect the geometry of the picture. Examples of point operations are intensification and quantization. When intensifying a picture, one multiplies the gray level by a constant, whereas the quantization of a picture involves the reduction of the number of gray levels, often by adding random noise (i.e., adding or subtracting one gray level in a random sequence) to "break up" false contours. The selection of class intervals in contour and choropleth maps is a type of quantization.

Another type of picture processing is *local operations*, in which a point is operated upon with respect to its neighborhood. In other words, the value of every point is changed and the change is a function of the neighboring values. All filtering can be considered as a local operation. The neighboring values contribute to the value of the point by the computation of a weighted average (smoothing) or differential (edging) of the sub-surface. Of course, "neighborhood" can have a directional bias, that is, be along a line (line-spread-function) or only on one side of a line (edge-spread-function).

The digital implementation of these operations poses computational problems insofar as the number of multiplications and additions increases rapidly with the increasing size of the picture. Conventional computers are basically sequential, that is, they can perform only one arithmetic operation at a time. If many identical operations could be performed simultaneously (i.e., "parallel"), great savings could be realized. Two attempts in this direction can be observed. One is the construction of special "parallel" computers with many processing units. Such computers are

[42] See Rosenfeld, 1969, and Lipkin and Rosenfeld, 1970

very complex and costly but the computational savings justify their development

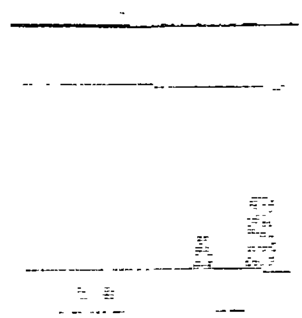
The other approach is to take advantage of the fact that even a conventional computer can perform limited logical and shifting operations on each binary digit of one memory "word" in a parallel fashion In this case, a surface is represented by one or a stack of "bit-planes" (i e , grid values may be 0 or 1 at a point) Values larger than that can be created by using several planes, one on top of the other [43] With this type of data arrangement, a parallel processor can be simulated rather effectively [44] (Figure 4 7)

Despite their efficiency, parallel processing languages are not very frequently used, since they involve extensive programming efforts Popular procedures use much simpler programming Still, the approach to first create a regular grid of cells is used very often for the treatment of land-use vegetation coverage and geologic data, etc For every cell, it is recorded whether a certain coverage is present or not, or the percentage of coverage if the unit cell is relatively large. The grid method allows very efficient computation of the area of a certain coverage type and its centroid. It also lends itself to some very fast overlay operations of two or more types of distributions as the computation of the intersection and union of several distributions and their areas

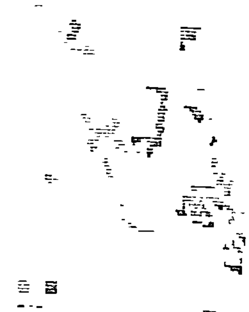
Some procedures allow for variable grid-size to guarantee a balance between detailed polygon representation and low storage needs [45] If the polygons of a map vary considerably, the grid-method might need excessive storage space since the unit cell has to be very small. Overlay procedures based on boundary intersections are very useful especially if the boundary lines can be simplified to very few edges [46] Another method is to represent polygons by skeleton as described in Chapter VI, which facilitates certain operations like intersection and shading, but makes others (e g , area computation) very difficult

The case of two-dimensional surfaces (i e , with a z value of 1 or 0), occurs very often in geography and planning The collection of land-use and land-occupation data is massive As a result, the largest geographical data banks have as their main storage requirements boundary coordinates of land-use areas and possess impressive manipulation facilities in software and hardware [47]

The comparison of three dimensional maps is somewhat



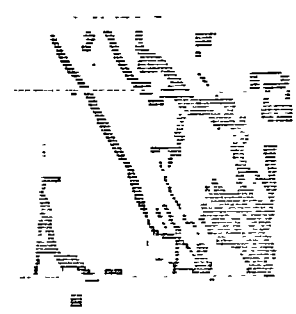
a existing development



b wood lot quality

c. soil values

d. 5 and 15% slope



e land not recommended for development

Figure 4 7 Erin Mills South Program RAPP, data T Stanhope University of Waterloo

more complex One method, cross correlation, has already been mentioned in an earlier part of this chapter. The two dimensional cross spectrum gives the similarity between two surfaces as a function of distance and direction from any point of one surface

Another much simpler approach mathematically is to compare two surfaces point by point This can be done by subtracting one surface from the other if they are based on the same measurement unit or their standardized values if they are not. It can also be done by taking the value of one surface as the dependent and the value of the other as the independent variable The two varibles then can be

---

[43] Six bit planes, for example, give a range from 0 to 63 (i e , 64 = 2^6 different values)

[44] A procedure simulating a parallel processor is described in several parts of Lipkin and Rosenfeld, 1970, but especially on pp 427-512

[45] Dangermond, J J , PIOS, San Diego County Comprehensive Planning Organization, San Diego, 1971

[46] See Chapter VII

[47] Tomlinson, 1970 gives a good account of these data banks

compared by correlation and regression analysis. By using regression analysis, the parameters of the regression equation would furnish new estimates for the dependent variable which could be subtracted from the original values. The estimates and the residuals could be mapped to show the corresponding variation and its deviation.

Problems can arise if the two data sets are collected from two different regional units. This occurs very often with social statistics where the statistical units vary from one collecting agency to the other. Interpolation to a common regular grid based on any of the previously mentioned assumptions can be the solution here.

V. REPRESENTATION OF SURFACES, LINES AND POINTS

It should not be too difficult to realize that surfaces present a concept with wide theoretical ramifications and practical applications. The question is how to represent different types of surfaces cartographically, especially by computer. A large number of manual methods exist. However, which of these methods should be automated? When should they be used? And can any methods be developed which are entirely different from manual methods?

Clearly the answer to the last question has to be negative if time and costs are restrictions However, the answer could be reversed by saying that there is no problem which will not eventually be handled by computer. However, costs will confine the number of tasks to be dealt with to those which have wide enough relevance to justify the development of software and possibly hardware

In Chapter III it was pointed out that surfaces do not have to be represented in pictorial form, especially not in the usual form of contour-lines (isarithms) In this chapter we will discuss different ways to represent surfaces, starting with the most "plastic" ones and becoming more abstract as the chapter goes along. Some problems which do not entirely fall into the area of surface representation but are needed in places for its understanding will also be handled here.

Three-Dimensional Information Elements

The discussion of the process of viewing three-dimensional objects will be inserted here One type of perspective view in two dimensions employs two pictures, one solely for the right eye, the other for the left This effect is relatively easy to produce, but the presentation of such images demands special equipment which people might find awkward to use

However, psychologists found out a long time ago that one does not need two eyes to see depth. Depth can be simulated by different processes

movement
one and two-point perspective
differences in size
shades and shading
structure in the representation
visibility

By moving the object with respect to the observer or the observer with respect to the object, the impression of depth can be created. The visual system responds very sharply to angular and size changes The ability to compare several aspects of an object at different angles or distances adds very much to three-dimensional visibility This comparison can be made through time (movement) or by varying distance (perspective and change of size)

If one displays the shadow of an object, one in fact produces two views of the object from different angles. The human eye is so used to the recognition of shades that it does not mind the transformation If the object (for example the surface) casts shadows onto itself, the display of these shadows greatly enhances the perspective impression.

If an object is given structure by any of the number of different types of structure-lines (i.e , a large number of lines arranged in some regular manner such as parallel profiles on a surface), the eye can easily recognize depth. An object with a well-chosen structure is usually also aesthetically pleasing However, this is only the case if those structural components hidden by closer components of the surface, are deleted from the picture The hidden surface problem is important in computer graphics and will therefore be discussed in more detail, after a treatment of the picture plane. (Figures 5.1 and 5 2)

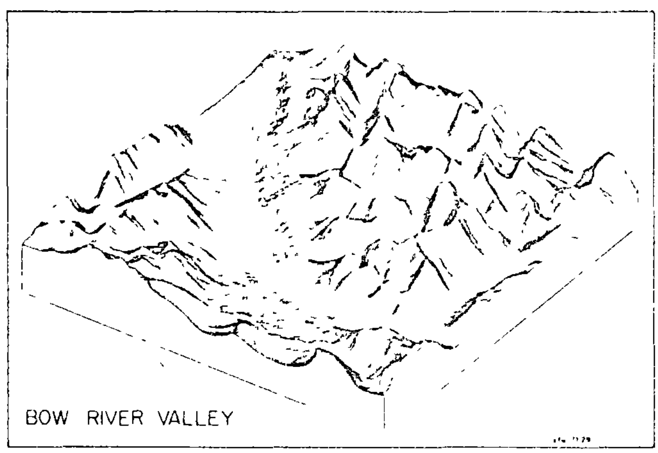


Figure 5.1 Bow River Valley, View from N W Shows very clearly the glacial features of the area Visual effect given by 298 lines running parallel over the surface Program SYMVU, data G Neilly, McMaster University

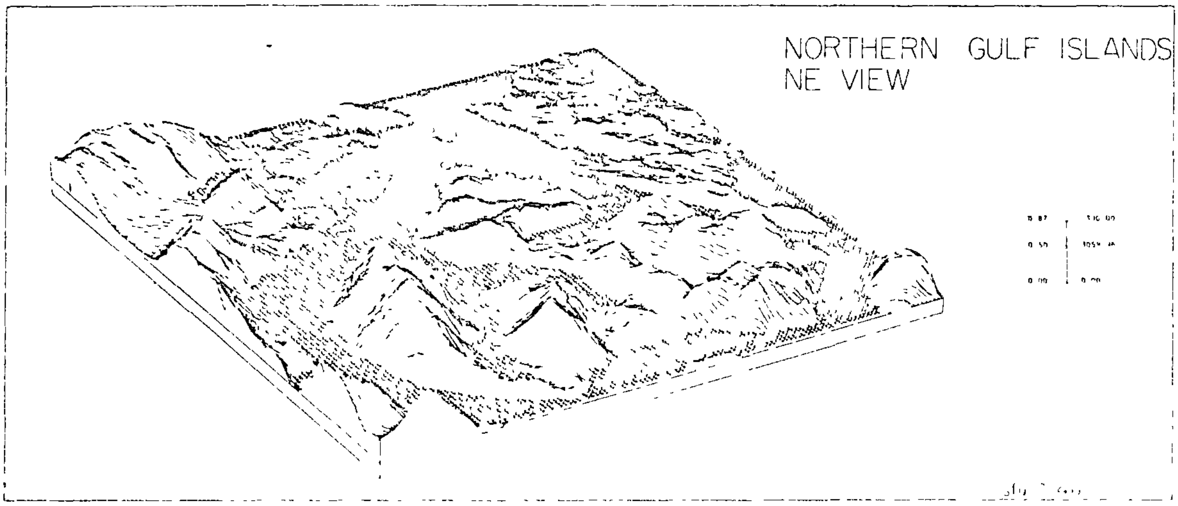


Figure 5 2 Northern Gulf Islands, Water Given Background Symbol Program SYMVU, data W Wolferstan, production W D Rase

Picture Plane

The easiest way to convert a three-dimensional object into a two-dimensional image is to place a glass plate between the object and the observer, who then traces what he can see. This does not require any artistic talent or experience, just a steady hand The picture will not be beautiful, but it will be correct.

Diagram 5 1a shows a hypothetical object at a given point (P for point) and its projection on the picture plane (I for image), with the location of the observer (O) at the distance (d) from the picture plane Diagram 5.1b shows a cross-section of Diagram 5.1a along the x axis. O is taken as the origin of the object space Both show the special case of a one-point perspective For the isometric projection, d becomes infinitely long, for the two-point perspective, the plane has to be tilted It is clear from the figures that the computation of the position of any point from the object-space to the picture plane does not pose conceptual problems.[48]

Visibility

For the general problem of visibility, Diagram 5 1 can be of service again. The outstanding property of a visible point, P, of the object is that the line segment $\overline{OP}$ contains only one point of the object, P Conversely, if P were hidden, $\overline{OP}$ would contain at least one point of the object other than P.

The easiest conceptual method for proceeding would be to test every point on an object for visibility A surface defined by a matrix could therefore be subdivided into a large number of sub-surfaces–each rectangle forms two triangles–and one could then test for every point whether the segment $\overline{OP}$ intersects with any of the subsurfaces

Although the algorithm for this test is straightforward, it consumes a considerable amount of computer time. Therefore, several "short-cuts" are used to reduce the computer time, taking advantage of the special characteristics of the data basis and the representation algorithm. Those pertaining to the representation of single-valued surfaces will be discussed, neglecting those applying to the representation of solids (e g , buildings) which usually need some additional considerations

A simplified version of the above algorithm is used for the computation of the visible area around a point of observation located on or above the surface[49] On the rim of the matrix or along a circle around the observation point, grid points are connected to the observation point along an imaginary line All grid points along the line are tested for their view-angle from the observation point starting with the closest point and going outwards As long as the view angle continues to rise, the points are visible As soon as one angle drops down to a lower level than the previous one, the point is not visible The highest angle is stored and the procedure continues If at any point the

[48] Methods for shifting the origin to the observation point and rotating it so that the picture plane is parallel to the y axis, as well as other geometric operations in three-dimensional space can be found in Ahuja, D V and S A Coons (1968), "Geometry for Construction and Display," *IBM Systems Journal* Vol 7, 188, 205

[49] The basic algorithm was developed by Amidon, E L and G H Elsner (1968) *Delineating Landscape View Areas–A Computer Approach*, USDA Forest Research Note, PSW 180, Pacific S W Forest and Range Experiment Station, Berkeley, California

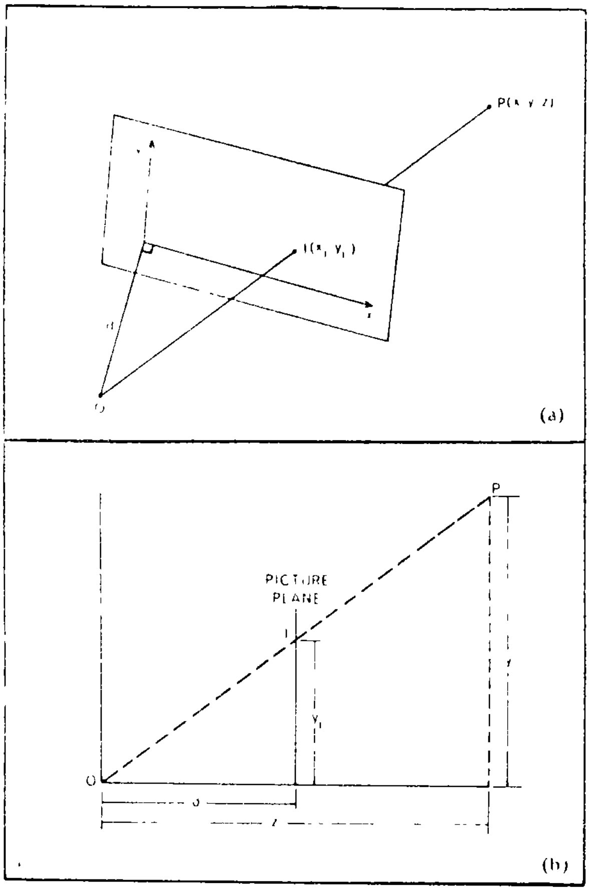


Diagram 5.1 Picture Plane

angle rises again above the stored value the area becomes visible again. This process has to be repeated as often as there are points on the rim (Figure 5.3).

If the surface is given by irregular points and an interpolation to a regular grid proves not feasible, a method similar to the first must be utilized. One may first test which of the triangles of the surface are facing towards the observation point and which are not. This test can be done by computing the equation of the plane of each triangle and then testing whether the observation point lies above or below the plane. In the case of triangles not facing the observation point, the plane will be invisible in every case and can therefore be deleted for the rest of the visibility tests.

If a surface structure (i.e., a series of parallel lines) has to be developed, the advantage of sequence can be used for an efficient algorithm. The lines are arranged according to their situation with respect to the viewer's position. One starts with the closest line and draws it. Drawing the next line reveals whether it crosses the first one at any time (i.e., goes below the first line). If this is the case, the section below is not drawn and the visible parts are stored as a "horizon" with those parts of the first line which hide portions of the second. This procedure is used for all subsequent lines, only those parts are drawn which are above the "horizon" and the visible parts of the line contribute to the horizon of the next line.

Since these parallel lines run across a matrix of z values one could flag any point which is visible. This information would thus be stored for later drawings on the same surface, such as lines perpendicular to the previous set of lines, roads, rivers, settlements, etc. (Figure 5.4).

We use another approach when a perspective view of a surface represented by its isarithms is constructed, and if the observer is situated above the surface he starts with the highest contour line and works his way downward. Inside the area bounded by the highest contour, no other contour can be seen. The second highest contour will only be drawn when it does not lie inside the first one. The union of the two contours (i.e., the area which is occupied by either of the contour areas or by both together) will be the "covered area" for the next contour, etc.

Block Diagrams

The block diagram or perspective view of a surface is gaining in popularity as are other mapping methods which provide an easier and quicker comprehension of the surface displayed. Different kinds of perspectives are used: the isometric view (which maintains parallel lines) is the most popular, but one and two point, and even cylindric[50] perspectives are also used.

It has to be admitted that the block diagrams produced by the computer are very simple compared to what has been produced by hand. They are usually only lines along the surface with a few relatively simple symbols.

Most of the procedures for making block diagrams start with a matrix of z values (i.e., with a regular grid of x, y and z coordinates) in which the first two coordinates do not have to be specified since they are implicitly given by the point's situation in the matrix. The procedures compute the location of each point on the view plane and then link all points along rows, columns or diagonals or combinations of these. Most users, however, employ diagonals if they are available since they give a higher density of lines. It is easy to show that the number of diagonals is equal to the sum of columns and rows minus two (Figure 5.5).

[50] The viewer is considered as standing inside a cylinder looking 360° around through the cylinder; the image is the unfolded cylinder.

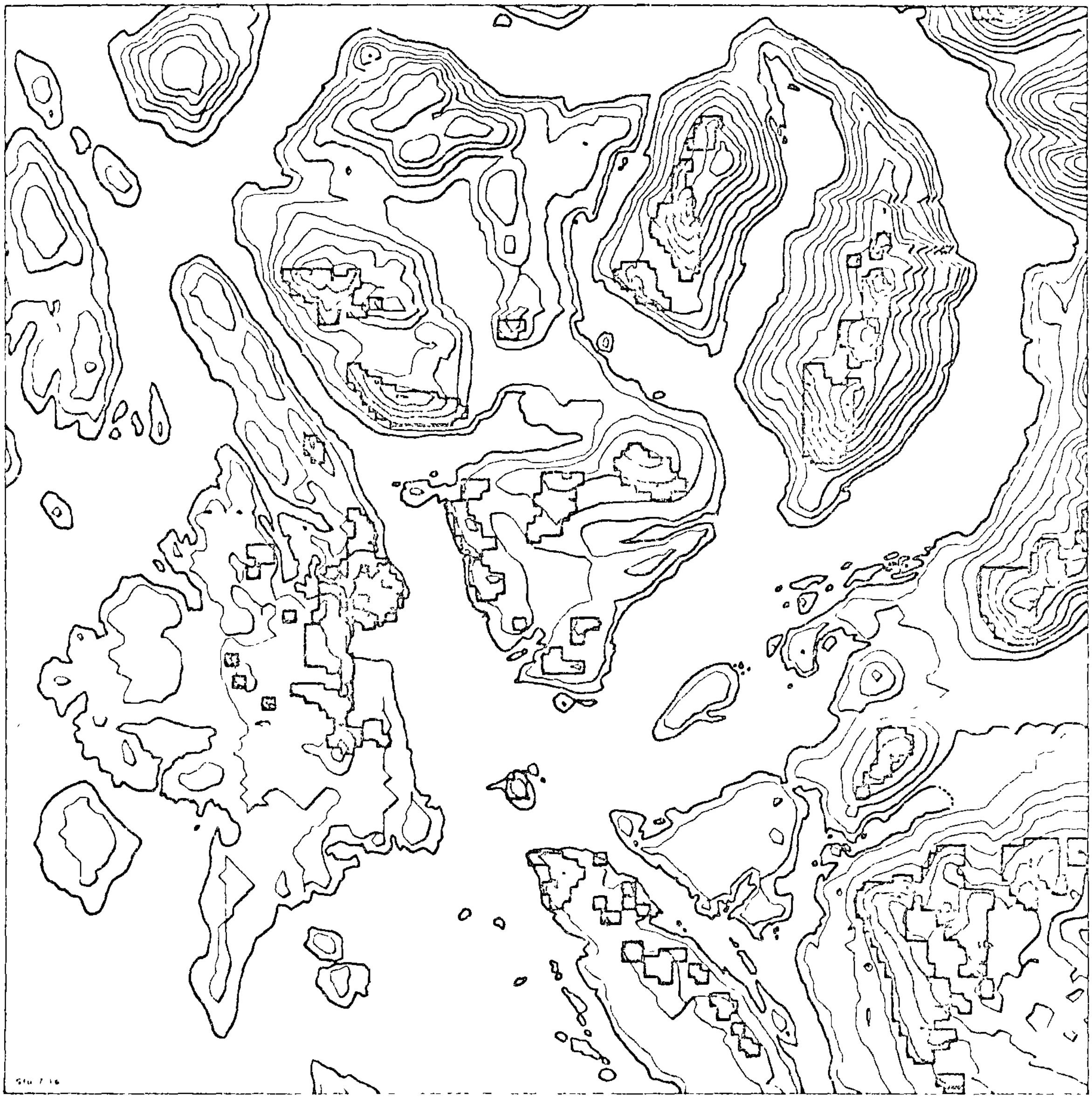


Figure 5.3 Northern Gulf Islands, Visible Area From One Point Programs VIEW, SHDCTR, data W Wolferstan, production W D Rase and G Brady

Block diagrams have one major defect Whereas horizontal space on the map usually represents planimetric position, part of it has to be reserved for horizontal position In other words, high features can hide low ones behind them Block diagrams therefore cannot be used for any general purpose information since the type of information that can be seen on a surface depends very much on the azimuth and angle of view However, block diagrams can be very good to show certain aspects of a surface, especially for teaching purposes In Figure 5.6, a perspective view of the Northern Gulf Islands of British Columbia has been used to show ice covering and flow during the last ice age.[51]

Planimetrically Correct Surface Representations

It is sometimes important to have a planimetrically correct map which displays terrain at a quick glance In fast

---

[51] Nieuwenhuizen, W., *The Late Pleistocene and Recent Glacial Morphology and Chronology of S.W. British Columbia and N.W. Washington*, M.A. thesis in progress, Simon Fraser University

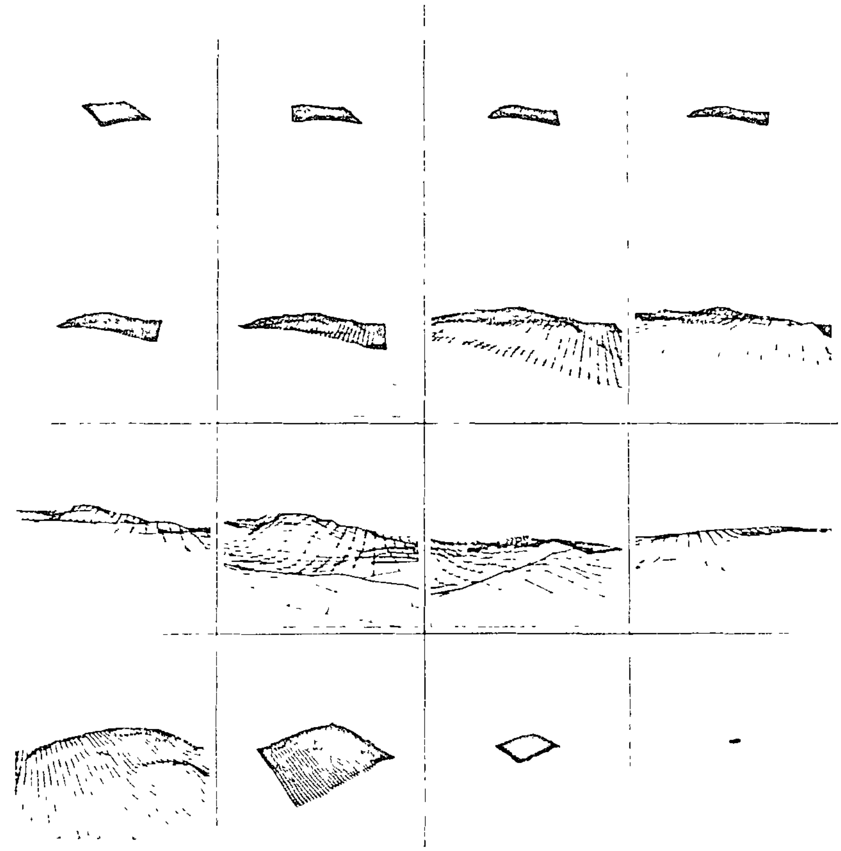


Figure 5 4 Camp Fortune, Views Seen by Approaching and Passing Aircraft The turn to look back is at the 12th graph VII WBI OK computes the drawing of one set of parallel lines and flags the points (1-visible, 0-not visible) for the perpendicular lines or other features Program VII.WBLOK, data D Douglas

moving transportation such as automobiles and airplanes, it is not necessary to know terrain variations in the degree of detail displayed by a contour map. In fact it can be extremely dangerous to have to study contour-lines on the map carefully in order to understand the relief, because it keeps the eyes off the road for too long a period

Automatically-produced planimetrically correct surface representations are all based on the idea of relief shading Assuming a perfectly white, matte body and a light source at a certain angle, then the amount of light falling on each subsurface is proportional to the cosine of the angle between the direction of the light and the normal vector perpendicular to the subsurface [52] (Figure 5 7)

[52] In one study of relief shading it is hypothesized that the tangent of the angle is more appropriate to make ridges come out white or, in other words, to mix oblique and vertical illumination See Marzig, Z, 1971, "Automatic Relief Shading" *Photogrammetria*, Vol 27, pp 57-70

Several approaches to this theory of hill-shading have been developed The most direct one is that of analytical hill shading in which the relief is simulated by gray tones [53] Until very recently, instruments were a serious bottleneck, since it was difficult to produce a satisfactory gray-scale, but this problem seems to be disappearing

The other two methods were developed in the late twenties and early forties respectively by Kichiro Tanaka [54]

[53] See Yoeli, P, 1965, "Analytical Hill Shading" *Surveying and Mapping*, Vol 25, 573-579, and Yoeli, P, 1967 "Mechanization in Analytical Hill-Shading," *The Cartographic Journal* Vol 4, pp 82-88

[54] See for the first Tanaka, K, (1932) "The Orthographical Relief Method of Representing Hill Features on a Topographical Map," *Geographical Journal*, Vol 79 213-219 and Robinson, A H and N J W Thrower, 1957 "A New Method for Terrain Representation," *Geographical Review*, Vol 47 507-520, and for the other Tanaka, K, 1950 "The Relief Contour Method of Representing Hill Features on a Topographical Map," *Geographical Review*, Vol 40, pp 444-456

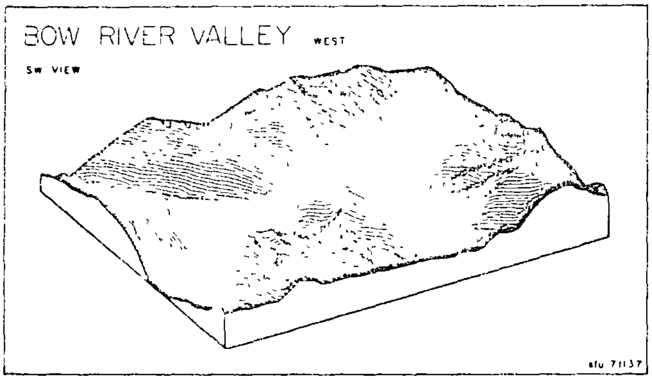


Figure 5 5 Bow River Valley, West Shows the amount of area which can be hidden in a perspective plot Program SYMVU, data T K. Peucker, production W.D Rase

The first method represents the intersections of the surface with parallel inclined planes. In other words, the procedure draws inclined contour lines The degree of inclination and the azimuth of the planes should be orthogonal to the direction of the illumination source, usually from West, Northwest, or North. The other method draws contour lines in black or white on a gray background with their width varying to approximate brightness (Figure 5.8)

Isarithms

If a surface is intersected by a plane parallel to the basis and at a distance z, we call the trace of the plane with the surface an isarithm The isarithm consists of one or more closed loops In case of a saddlepoint of exactly the height z, the isarithm of height z will cross itself at the pass[55]

The loops are only continuous if the surface is continuous and thus single-valued, that is, if no vertical or over-hanging cliffs exist Piece-wise continuous, single-valued functions allowing for vertical cliffs have been worked on[56] but hardly any procedures allow for this[57] However, to simplify the explanation, only contour-line algorithms on continuous surfaces will be discussed

A frequent misconception is that isarithmic maps are the most accurate representation of a surface The only two types of maps which are as accurate as the original data-gathering (terrain-survey, etc) are tables giving x, y and z coordinates, and maps showing the surveyed points and their values

A very important factor for the accuracy of a contour map is the density of points. Figure 5 9a - d shows a one square kilometer section of the Allatoona Dam, Georgia area in the original grid of 101 x 101 points[58] and a subsequent reduction of points by a factor of two (51 x 51), three (34 x 34), and four (26 x 26) On the basis of the slight difference between the first and the second map one might say that the first surface is overrepresented But, again, too little research has been done to give a clear notion of the optimal number of points and the governing conditions.

Depending on the distribution of the data-points, different types of preprocessing are necessary before proceeding to the contouring. If the data-points are distributed irregularly, either a set of non-overlapping triangles has to be created to build the basis of a rather complicated contouring procedure, or the surface is interpolated to a regular grid It has been shown[59] that a grid of regular triangles offers the most consistent way of contouring, but since a rectangular grid allows a more efficient storing of grid-point values, the latter is used most frequently.

[55] This can be shown easily by two hyperbolas perpendicular to each other and with the same axes, with the foci moving towards the intersection of the axes When the foci converge in the intersection, the hyperbolae will be identical with the axes

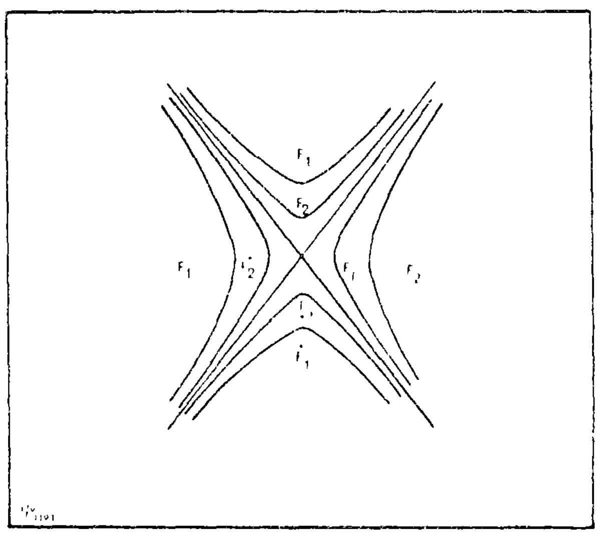


Diagram 5 2 Contour on a Saddle-Point

[56] Morse, S P , 1968, "A Mathematical Model for the Analysis of Contour Line Data," *Journal of the Association for Computer Machinery*, Vol 15, 205-220

[57] One program that does allow for it is SYMAP, which has a package for interpolation barriers

[58] The data have been generously provided by Dr Dean Edson, U S Geological Survey, Topographic Division, Washington, D C

[59] Bengtson B I and S Nordbeck (1964) "Construction of Isarithms and Isarithmic Maps by Computer," *B I T , Nordisk Tidskrift for Informations-Behandling*, Vol 4, 87-105

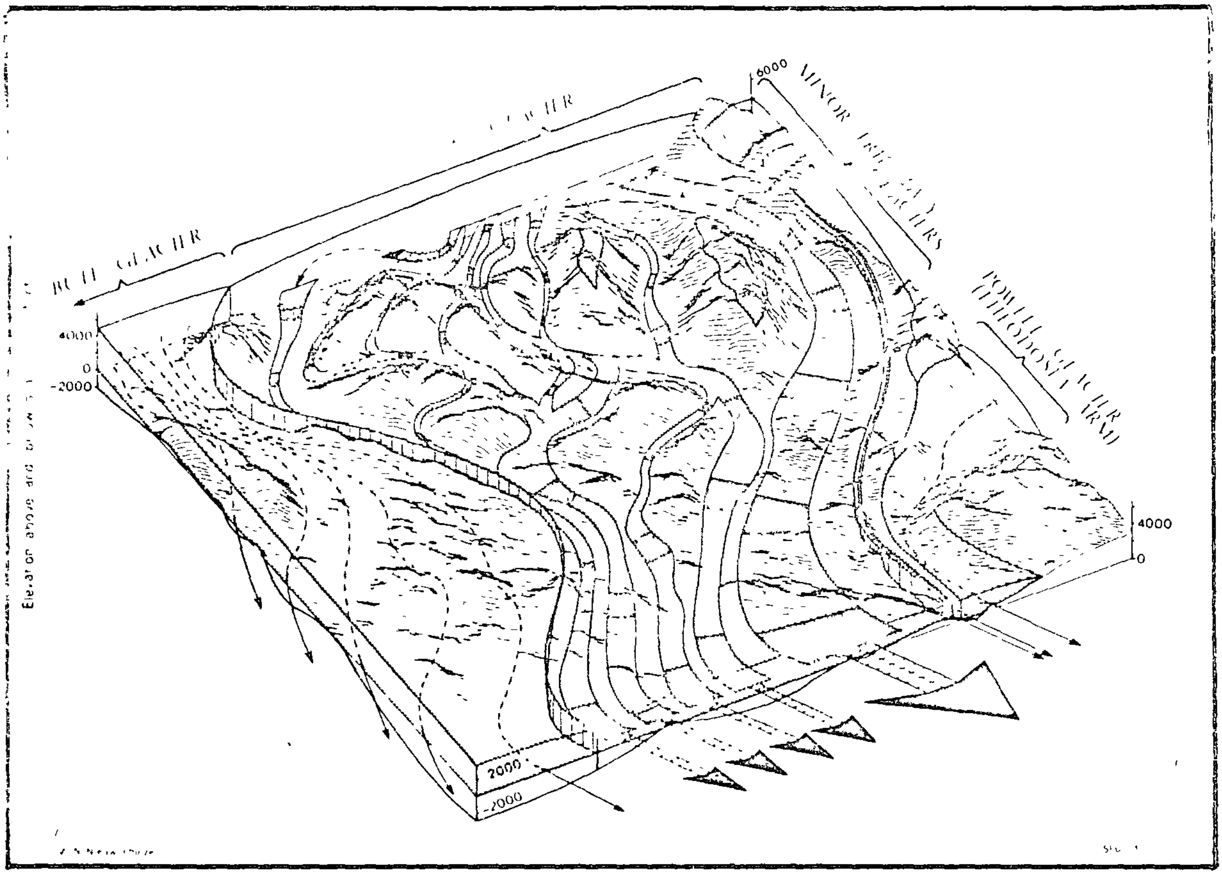


Figure 5 6 Northern Gulf Islands, Ice-Flow Map. A computer-drawn view used as the basis for the analysis and description of ice-flow patterns Analysis and ice flow W. Nieuwenhuizen, Simon Fraser University, program SYMVU, data W. Wolferstan

The problem of contouring a rectangular grid can be broken down into three parts manipulation of one cell, creation of a string of lines, and smoothing

One cell is defined by the four corner-points of a rectangle Since the procedure of testing a grid for a contour-segment has to be repeated very often (for example, for a map with a 100 x 100 grid and 20 contour-levels–200,000 times), the procedure has to be as fast as possible. The basic idea is to test whether any one of the four edges of the rectangle intersects or touches a contour level Some algorithms do this and then produce one straight line through the square. Others account for the fact that a square has one point too many for an unambiguous definition of a plane, and therefore either create two triangles with one diagonal as the common base for the two, or compute the center of the cell as the mean of the four cell points, and allow it to be the common point for four triangles. Others again fit a polynomial surface through the cell-points and the neighboring points to get a smooth curve as contour-line It is clear that the first method is the fastest but crudest and the last the slowest but smoothest. The other two methods lie in between with respect to computing speed and smoothness

The simplest method of creating a set of contour lines is to scan the matrix row by row or column by column. This can also be done with the least computer memory occupied, but makes smoothing very difficult. (Diagram 5.3) The other method is to test the rim of the matrix and search along each column for the start of a contour line and then follow the line through the matrix until it reaches the rim again or closes the loop This method needs much more fast memory, since the whole matrix has to be kept in core (computer memory) for the whole time of the job run

However, it makes smoothing the lines very simple since it creates a sequence of segments which only has to be "ironed out" by different line smoothing techniques. With

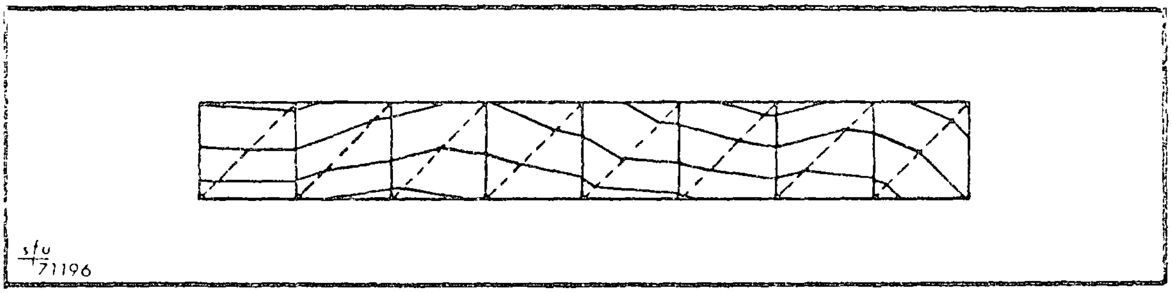


Diagram 5.3 Contour Scanning

the equipment presently available, it also saves plotting time since it allows a continuous plot of a full line.

Many specialists consider the smoothness of contour lines the most critical requirement for the general user in automated isarithmic mapping. It is therefore almost essential to apply smoothing methods to the scanning procedures. The least-square fitting through a patch of points with the cell as a center has already been mentioned, another approach would be to store the results of several rows and only plot segments when they can be smoothed with their neighbors on both sides. It has to be mentioned, however, that a smooth contour-line depends highly on a smooth surface. (Figure 5.10)

As can be seen from a comparison of the isarithms of smoothed surfaces, the very high and very low ones shrink with continuous smoothing. In other words, smoothing flattens the surface with the mean remaining constant. This is unfortunate, because one usually would like to retain the extrema since they are often measured with higher accuracy than other points on the surface. It has therefore been suggested to restore the surface by post-multiplying it with a correction factor which reestablishes the extrema.[60]

The production of a very rough contour map is quite cheap. Improvements mean extra development and computer time, with decreasing returns. Somewhere along the continuum of increasing costs there is a point where the resulting improvement is too small to justify the additional costs. At this point, the editing abilities of an experienced cartographer are more efficient. It must be admitted that the special skills of the cartographer have not yet been used enough for the production of contour maps. A large amount of computer time and program development has been wasted and good talents have been unused in the attempt to manufacture final results.

Figure 5.7 Theoretical Surface, Relief Contour Method. Program SHDCTR, data D. Wolfe, Simon Fraser University, production T. K. Peucker

[60] Tobler, W. R. (1966), *Numerical Map Generalization*, Discussion Paper No. 8, Michigan Inter-University Community of Mathematical Geographers. The formula is

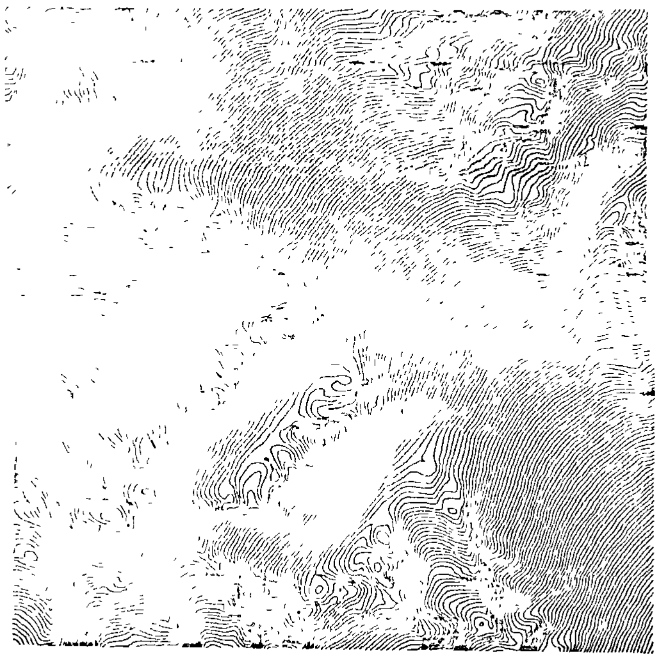
$$Z^r_{ij} = A + B\,(Z^s_{ij} - A) \text{ where}$$

Z^r_{ij} = the restored value of the point ij

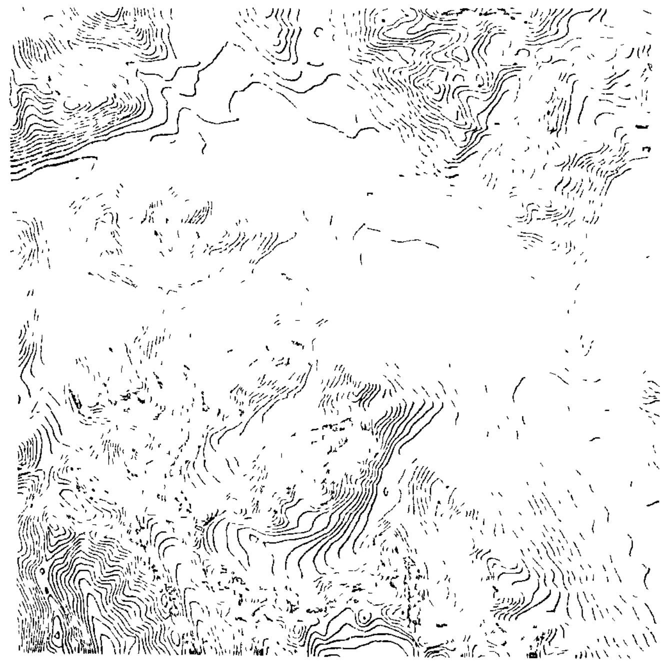
Z^s_{ij} = the smoothed value of the point ij

A = the mean value

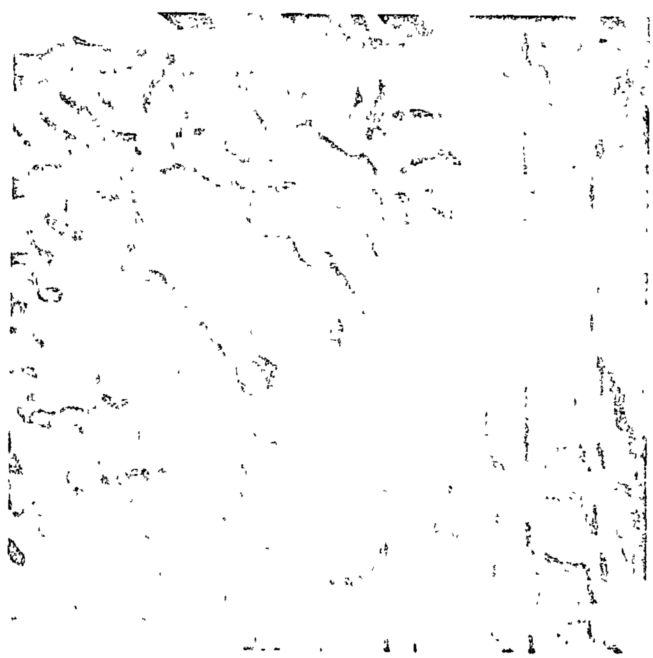
B = the correction factor = $\dfrac{Z_{max} - Z_{min}}{Z^s_{max} - Z^s_{min}}$



a. orthographical relief method, INCLIN



b. relief contour method, SHDCTR



c. analytical hillshading, YOLLI

Figure 5.8 Bow River Valley, Three Methods of Hillshading. Data G. Neilly, production a and b, W.D. Rase, c, A Pilipchuck, University of Maryland

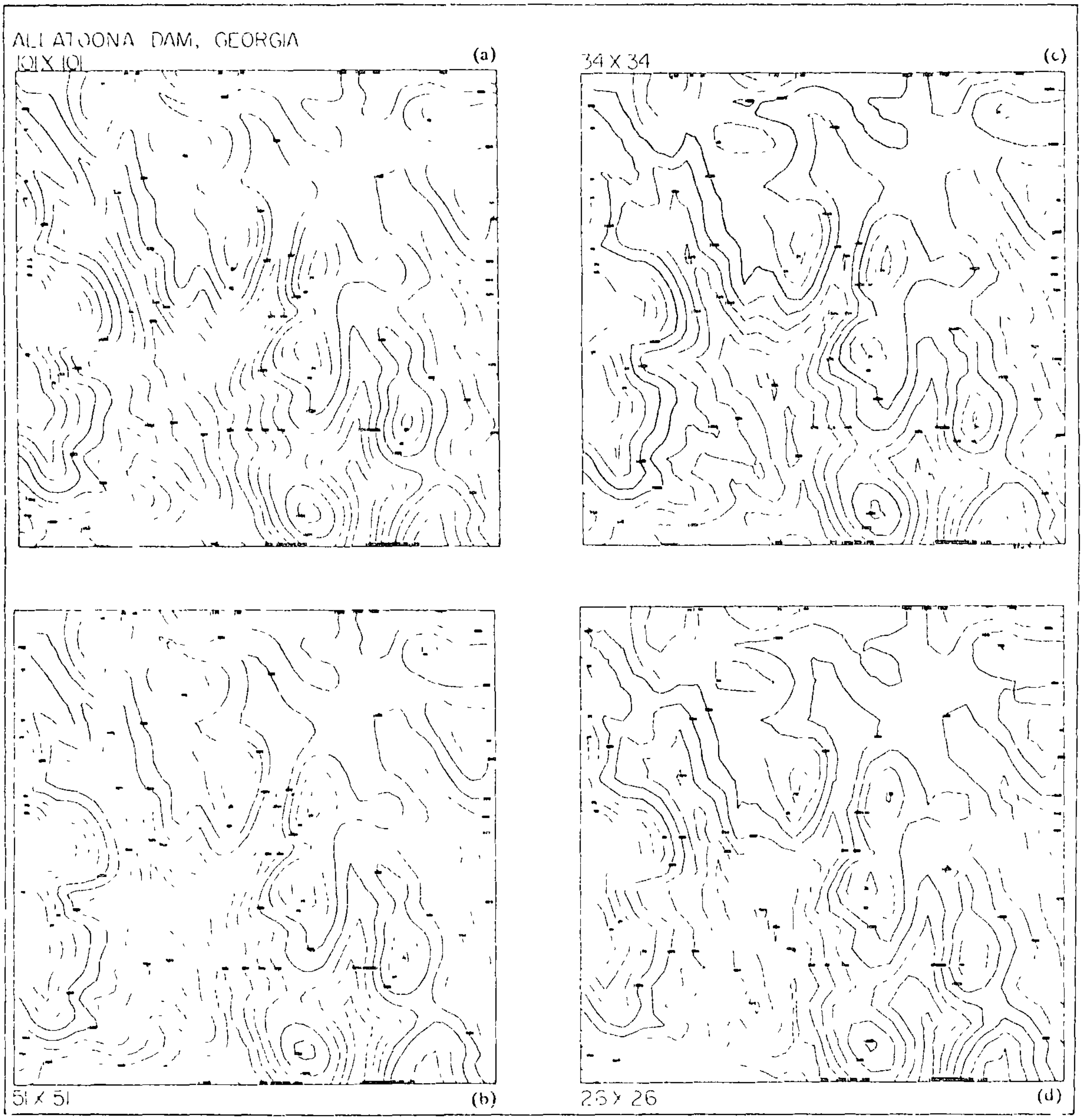


Figure 5 9 Allatoona Dam. Contouring With Varying Grid Sizes Program KOPPE, CNTOUR, data courtesy D Edson, U S Geological Survey, production T K. Peucker

Data Grids and Functions

Surfaces do not need to be displayed in an image to qualify as maps In many studies, for example, in most of the work on response surfaces, one needs only the location of the maxima and minima of a surface, and for any surface manipulations, a regular grid of values has great advantages over display or data in the form of a traditional map. Again, the operation of map overlay in many cases can be performed without any graphic output, for example when areal extent is asked for or when two types of land occupancy (e g., residential and a certain range of slope) are used to create a new land-type by their intersections (e.g., a specific residential zone) for later manipulation.

In many cases the optimal surface type for manipulation is different from the optimal type for storage, as it is also different from the optimal type for interpretation by man. For example, an irregularly distributed set of data-points, or sometimes a set of functions, will use the least storage space, but the best form to manipulate is upon a regular grid, and the ideal form for visual interpretation is a relief map Also the most economical storage of land-use data usually designates the boundaries, the fastest manipulation can be performed on bit planes, and many people would find only a colored map acceptable.

Points and Lines

Although points and lines could be considered special cases of binary surfaces, they will be treated separately since their computational and graphic manipulations turn out to be quite different from those of the surface.

In picture processing, the relationship between surfaces and lines is very strong In a dense point-array, a line is initially a long-stretched surface, often interrupted if the picture of the line includes much noise Before description, such a surface has to be thinned and connected. Only then can it be classified as a certain type of line However, the cartographer is rarely concerned with this kind of problem since his task is not the recognition but the production of points and lines on the map. Therefore, a definition within the framework of Euclidean geometry is more appropriate

A point divides a line into two segments (break point, node, corner, etc.), a line divides a surface into two areas (boundary, frontier, etc.), and a surface divides space into two positions (front, wall, stratum, etc.) To use these terms for the boundary definition of spatial items, one can say that two points on a line isolate a segment, that a closed line on a surface isolates a region, and that a closed surface in space isolates a volume.

Points as zero-dimensional surfaces are used for those items which have little or no extent relative to the scale. The symbol usually occupies more space on the map than its extension in reality. It is coded by its coordinates and a name, number, or any other content-sign (see Chapter VI).

A line can be coded in several ways as described in Chapter VI. The principle is to represent a line by a vector or a set of vectors (straight lines) For computational, but rarely for storage purposes, curves are sometimes represented by functions.

The content of points and lines can be represented in a variety of ways Different types of characters and symbols, changing size and color in two or three dimensions, can be used The computer does not add any new type of symbols to those drawn manually, but creates several problems easily handled by a trained cartographer. Two problems will be discussed here.

The first problem arises from the fact that the standard computer is fundamentally sequential, whereas the eye operates in a parallel way The eye therefore can perceive spatial relationships which the computer cannot, unless so programmed. This calls for programs which simulate parallel computers, for highly complicated procedures for hidden-line elimination, etc.

An example should clarify the problem before some of the solutions are pointed out. A careful look at the edges of the isarithmic maps will show that many of the height labels are unreadable because of overlapping A draftsman can easily avoid this, by not writing where he had previously written However, the computer would have to keep every stroke of previous drawings in memory and at each drawing-step search through its memory for possible clashes, thereby imposing tremendous computing demands in the case of a complicated graph For an isarithmic map, one could eliminate labeling at the rim and only label contours in areas of small slope (i e , where isarithms are far apart) This procedure would avoid most overlapping

Methods do exist, however, in thematic cartography, for organizing the data and the computational sequence to avoid overlapping, without having to keep everything in memory. If the map is displayed by the line printer or on a television-like storage-tube, it is produced point-by-point and line-by-line. Since the program has to accept information from different display categories for each point (e g , contours, streets, labels, etc.), one can give these categories different priorities. The procedure is to store different aspects of the map on secondary storage (disc or tape) and run the different aspects off memory point by point on three different tapes. Once the computations are finished, the tapes are rewound, and for every point the information is read from all three tapes. If there is something on all three tapes for one point, the label, if given first priority, will be drawn. If there is no label, the street should be next and then the contour. One program operating in such a

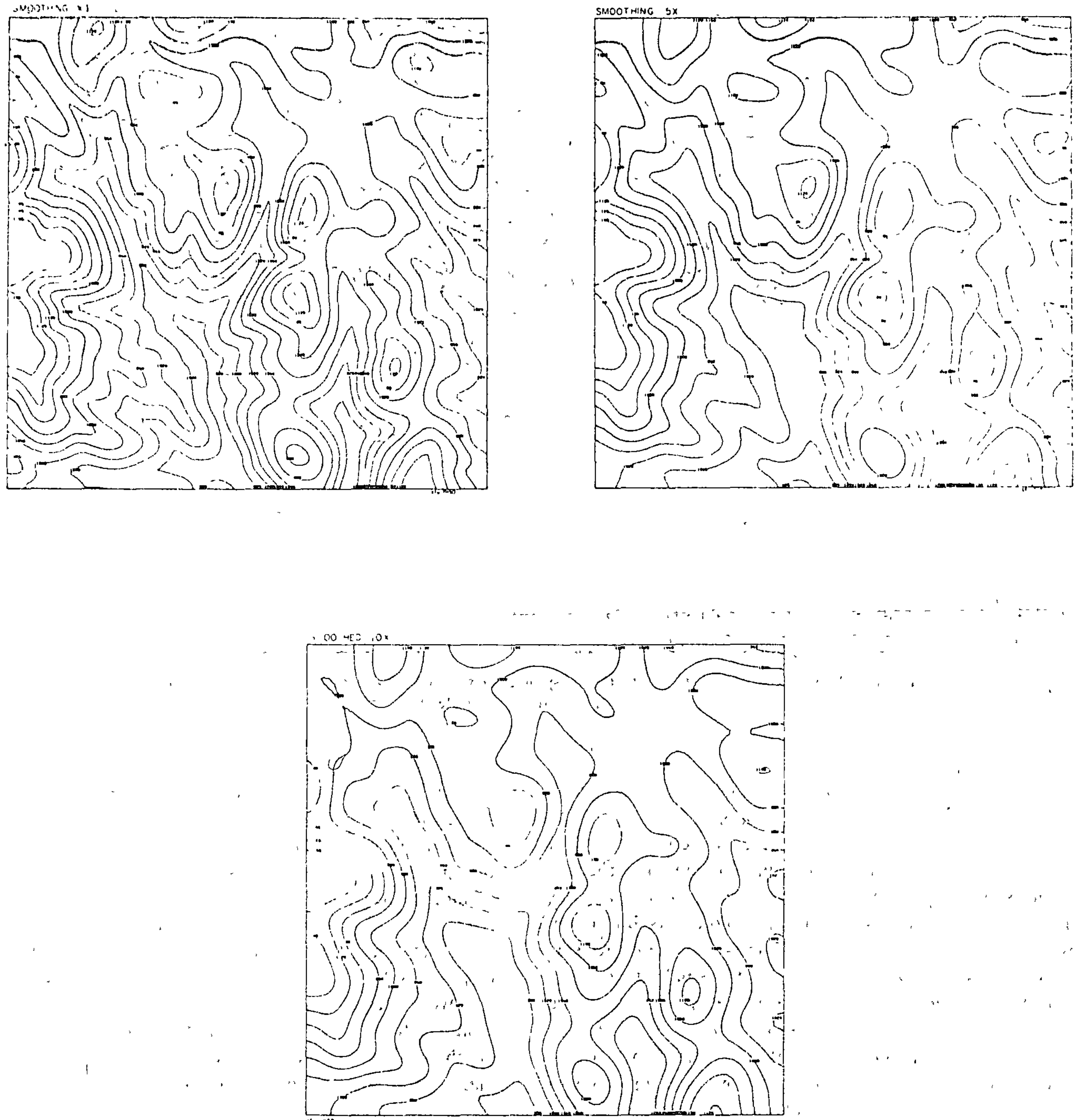


Figure 5 10 Allatoona Dam, Contouring With Smoothings. Programs SMOOTH, CNTOUR, data courtesy D. Edson, U.S Geological Survey; production. T K. Peucker

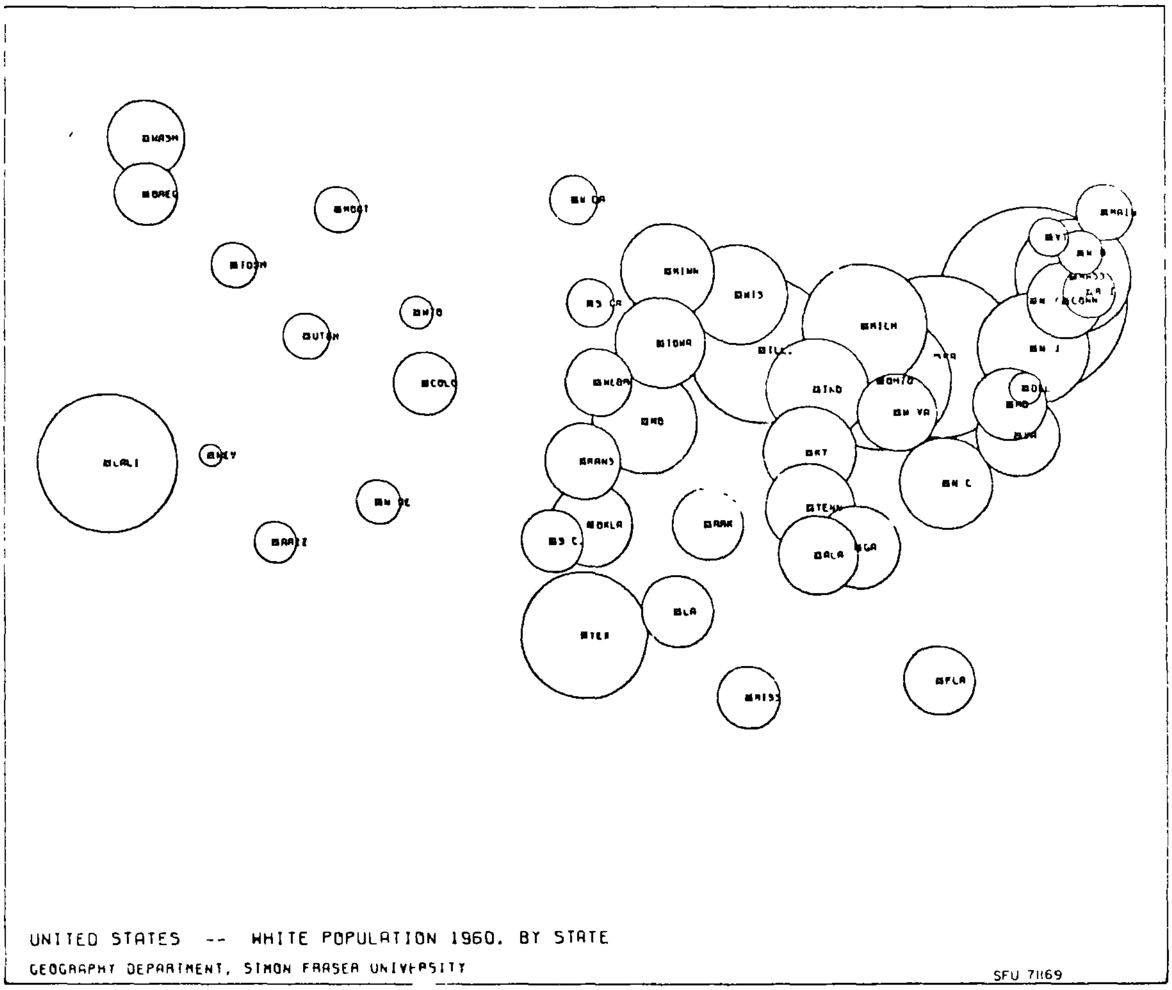


Figure 5 11 Display of Absolute Data Overlapping circles with removal of overlapping parts Program SIRKLL, data D Hatfield

fashion is SYMAP However, for this purpose, it uses discs which are faster than tapes and have less restrictions

Another method is to sort the items according to their probability of overlapping or being hidden, and then to suppress overlapping parts In the case of point symbols, circles, squares, etc , are good size indicators but many overlaps would offend the eye. One could, therefore, sort the data first by size, starting with the smallest item and with each subsequent item test for the intersection to be deleted Similar things could be done with lines, although stretched lines that cross do not have the same adverse visual effect as point symbols Three-dimensional histograms, on the other hand, would look confusing without the removal of hidden lines (Figures 5 11, 5 12)

A third method of avoiding overlap is interactive First, the program plots out a rough map onto a screen, an operator clears up lines, shifts labels etc , and then produces a hard-copy on another device The updating of the North American aeronautical charts is done in this manner [61]

As mentioned above, smoothing a surface (i e , filtering out high frequencies), works like a generalization of the surface Unfortunately, filtering is rarely used for scale reduction, and generalization has been applied to isarithms, not always with the best results because the generalization of point and line symbols poses some additional problems.

Several "laws of generalization" have been developed all pointing to the general rule that the ratio between the numbers of symbols in the original and in the generalized map should be the inverse of the square root of the

[61] Luetje, I H and R I Gard (1968) "Computer-Assisted Cartography, A Graphic System for Chart Composition and Revision," *Information Systems Symposium*, Sept 4 6, 1968, Washington, D.C

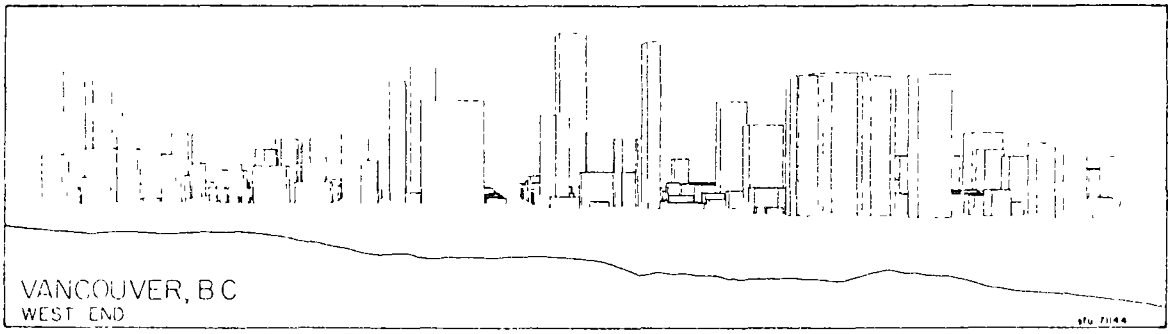


Figure 5 12 West End, Vancouver, Advantages of Hidden Line Removal. Program BLOCKS, data A Furguson

map-scale ratios[62] At first sight, this seems a straightforward rule without ambiguities If the scale is reduced by a factor of four, the number of settlements, the total length of streets, isarithms, etc., has to be reduced by two But what types of settlements should be chosen, which streets dropped? Simply to sort all settlements (e g., on a U.S. map), by size, take the median, and drop everything below it, would leave large areas without any city and create overcrowding in other areas. It is clear that the "law" has to be specified to give enough information for a suitable selection process

Some additional problems have to be answered in automated line generalization. Not only do lines have to be selected, but for those retained, the amount of undulation (the length) has to be reduced. This can be done in at least two ways.

[62] i e the "Radical Law" in Toepfer, F , and W Pillewitzer (1966), "The Principles of Selection," *The Cartographic Journal*, Vol 3, 10-16

The first and simplest approach is to select every n-th stored point for plotting, where n is dependent on the scale change and the density of points. This reduces computation time, but does not reduce the complexity of the line. In other words, it does not reduce the number of features of the line. Some corner-cutting could occur if the original density of points were not high enough, and this could only have adverse effects on the legibility of the map.

The next approach is line smoothing. The number of features is reduced but in a highly uncontrolled way. Furthermore, certain line-features on a map would look very unrealistic if smoothed out For example, the coastline in a river delta is sharply indented landwards, and smoothing this out would be unacceptable. The answer is to maintain some of the larger river branches and suppress the others. This demands constant decisions or a very detailed catalogue of conditional rules. The main difficulty for an automation would be in recognizing feature classes to be maintained or deleted.

VI. DATA STRUCTURES

Most quantitative research in the social sciences involves large data sets This is especially true for geography We try to disaggregate our regional units as far down as possible, even if accompanied by a loss in precision of the data When it comes to geographical data banks, some people would like to have the coordinates of every tree in the world (and a few could even use this information)

A well-organized data structure is therefore crucial for any successful operation in computer cartography The points to consider range from labeling a card deck to the preparation of very sophisticated procedures for list-processing of data For many of these, packaged solutions are available, however the prospective numerical cartographer should always spend some time on the methods of bookkeeping and processing *before* he starts to code his information. Otherwise he will not be able to identify the content of his decks later on, or will have to duplicate decks with shifted columns to fit them to his programs In this chapter, some guidelines are given and basic problems are discussed Using this knowledge as a basis the reader should be able to understand and use existing data banks, he will, however, have to do some more reading to develop his own data structures

Throughout this chapter, we will assume that we have to develop a data base for our state or province For the present time, the bank might be very small (i e , it might include only a few items), but the base should be flexible enough to handle a host of other information To be more specific, we might only have to code county data with county names, population size, county boundaries and centroids, but we might be requested to allow for later inclusion of the population census, boundaries of overlapping sets of statistical and administrative areas such as school and hospital districts, and also to facilitate retrieval of the information for computational and display purposes

Coding

We have basically three kinds of devices to store our data in machine-readable form cards, tapes and discs. Most of the primary coding is done on cards, but one has to deal with data sets on tapes and discs relatively early in his career as a numerical cartographer. It is therefore quite important to have a good notion of the problems of coding ahead of time.

If one has to code a point, at least its coordinates and a uniquely identified "name" must be registered At least one value is usually attached to the point, such as the elevation on a topographic map, the population of a city, or the air pressure at a meteorological station. However the second example especially shows that there can be quite a few more items, such as several hundreds of variables for census tracts (identified by their centroid, in this case), and several thousands in time series of weather data (i e , every hour for ten years).

The coordinates of a point can be "x" and "y" coordinates, or a system's geographical coordinates (longitude and latitude), or scales in a two-dimensional diagrammetric system (such as temperature and precipitation in the x and y direction with crop yield in the third direction representing the individual value)

If one attempts to compare or join data-sets from two different coordinate systems, he should think about some organization in advance to avoid loss of data when he can no longer remember the projection or the origin in a matrix-like grid, and to prevent large computation errors that could occur because of incompatible projections If geographical projections are involved, experts therefore suggest keeping two geographical data files, one with the preferred projection and another with the geographical coordinates (longitude and latitude) In this case it is easy to digitize data in any projection and convert them to longitude and latitude, since this usually poses less problems than conversion from one projection to another It also allows users to trade data with other agencies if they also have a double system. The experts consider this better than trying to have everybody agree on one projection, an impossibility because the various systems are already quite advanced in the use of very different projections

Even when dealing with a regular grid system, however, problems of coordination can occur The mathematical coordinate system has its origin in the S W corner of the area, the first coordinate shows the location of a point to the east of the origin (x direction) and the second coordinate indicates the distance of the point to the north (y direction). To many people, this is the only possible coordinate system. As we have seen with SYMAP, such is not the case. SYMAP has its origin in the N.W. corner and measures south first (DOWN) and then east (ACROSS). Other coordinate systems can be envisioned.

In fact, eight possible systems can be identifed by three binary choices;

1. x or y coordinate first coded
2. origin west or east
3. origin south or north

The three choices can be presented in a system of binary numbers and translated into the decimal system so that every one of the 8 possible coordinate systems gets a number between 0 and 7. Table 6.1 gives the eight coordinate systems with their binary and the decimal numbers. Of course, some of these eight systems occur very rarely, others very often. System 0 is the traditional mathematical coordinate system, 5 is the SYMAP coordinate system.

Can one avoid misunderstandings and make sure that one never produces a map which is upside-down or a mirror image of the original? Two approaches are possible.

1. One agrees to a system and operates only with that one, or
2. One indicates on every data set and every program which coordinate-system it is based on and converts with the proper conversion routines.

The author would advise the reader to opt for the second alternative. The first approach would mean that any data coming from outside has to be converted immediately, and every program from outside has to be reprogrammed, a process far more difficult than reading the data set in with the proper routine.

We are now ready to digitize the counties of our state or province. The lucky student can use an automatic digitizer on campus, others should not despair, however, as long as the number of points to be coded does not exceed a thousand or so. In any case, the actual coding of points is only a part of the job, and the rest has to be done manually in any case.

TABLE 6.1

First Coord	Origin		Binary Number	Decimal Number
X	West	South	000	0
		North	001	1
	East	South	010	2
		North	011	3
Y	West	South	100	4
		North	101	5
	East	South	110	6
		North	111	7

Line-Storage

With rare exceptions, the point is the elementary coding unit for cartographic information. Lines are coded as a set of points, as are surfaces. The digitizing of surfaces is relatively straightforward as far as the data structure is concerned, since the result is an array of x, y, z coordinates or a regular grid of z values. It demands considerable reflection upon the sampling of points, however. This operation has been discussed in the chapters on coding of surfaces and surface behavior.

The digitization of lines, on the other hand, is straightforward with respect to the selection of points--we will play down the problems occurring in the case of generalization for the moment by assuming that we will keep scale and emphasis constant, however it involves quite complicated questions concerning the development of a line-data-bank. The digitization of county boundaries is a good example. One could digitize these boundaries by starting with the first county, coding its borders point-wise, around the clock, for example, and then proceeding with the second county, etc., until one has coded all the counties. In this case, most of the boundaries would have to be digitized twice, and the points where three counties come together three times. Furthermore, one is bound to produce errors even with the most accurate digitizer, when a point has slightly different coordinates for two counties, the discrepancy can disappear or be magnified by the rounding process inherent in any measurement. Apart from that, several program systems are unable to recognize a boundary between two areas if this boundary is not identified by exactly the same set of points for both areas.

One solution is to number all the points to be coded, digitize them in sequence, and then prepare lists of "pointers" as the boundary-points for each county (Diagram 6.1). This would also save considerable storage space by the reduction of x, y coordinates to be stored.[6.3]

A more complex data preparation system might involve a program which does the job of creating a pointer list for each county. In such a case one has simply to digitize boundaries without much care about the density of points, the accurate definition of start and end of a boundary, etc. Such a program might first find adjacent areas, merge lines

[6.3] With most computer-languages, one can pack two or three pointers into one word, thereby leading to a further reduction of storage space.

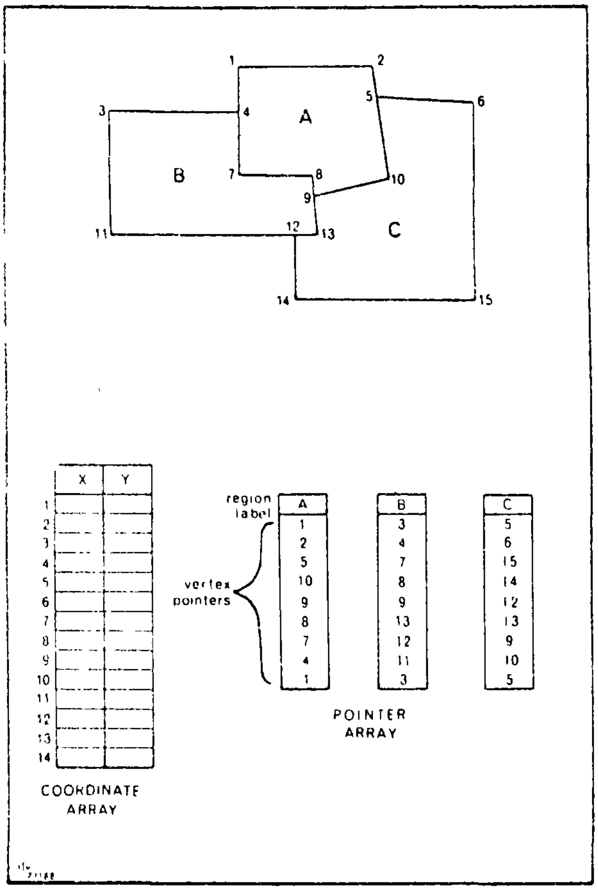


Diagram 6.1 Boundary Encoding

which are common borders, and find the optimal number of points according to some alignment criteria With such a program, most of the data preparation work would be performed by the computer

The highest storage requirement in geographical systems will most likely occur with line data It is therefore very important to make a good choice of a line-storage system. Four types can be distinguished [64]

1. In the x, y coordinate method, each line segment representing an arbitrary line may be encoded by storing the x and y coordinates of its endpoints The method is straightforward in its logic and digitizing It is very suitable when scaling and rotation are required often, and when area and circumference have to be computed. It is relatively poorly suited, for example, to cases in which the shading of areas with straight line segments is needed. Shading requires the repeated calculation of the points of intersection of a line with the area given, a process which requires much computer time for data encoded in this form.

2. Incremental encoding is very similar to the absolute x, y coordinate method. The difference is that it stores every point as if the previous point were the origin The method is as fast as the previous one for scaling, not much slower for rotation and area computation, and much faster for line-length. Its major advantage, however, lies in the storage savings by a factor of at least two.[65]

3 Chain encoding Each line can be approximated by a series of eight possible steps numbered from 0 to 7 as in diagram 6 2a. The line in diagram 6 2b would then be identified by the string 001020765. The choice of the basic unit's length (the grid size) according to the exactness required for the reconstructed drawings is important The method is very well suited for the representation of fine outlines of areas, for the computation of areas and for

[64] This section is largely based on Decker, G. F., (1970) *Interactive Graphics and a Planning Problem*, MSC thesis, Dept of Computing Science, University of Alberta, Edmonton

[65] If no increment is larger than 127 units (e g , 1 27 inches with a resolution of 1/100 inch) two x,y pairs can be stored in one word, whereas one needs at least one word for an absolute x,y pair (IBM-360 system)

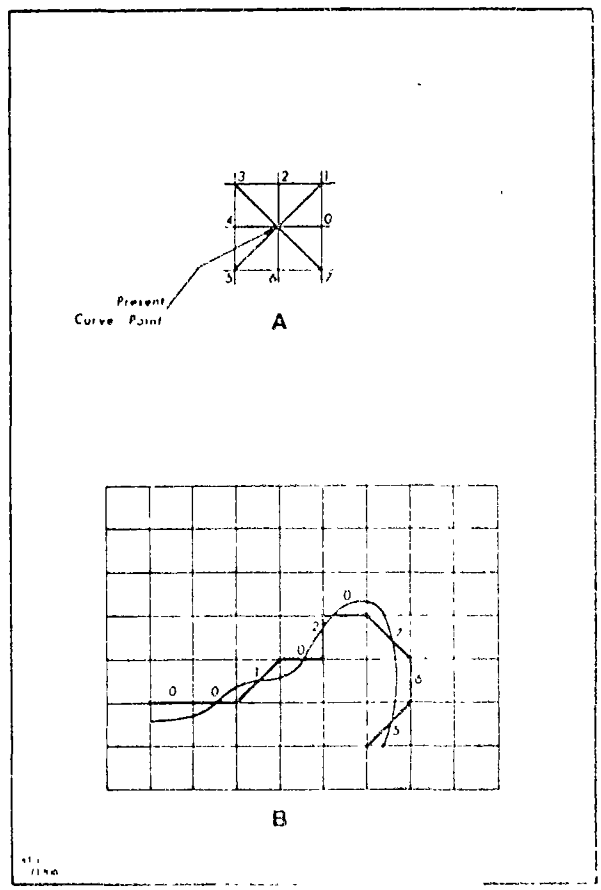


Diagram 6 2 Chain Encoding

scaling, but is poor for finding the intersection of two lines, rotation, shading, etc.[66]

4 Skeleton encoding Any area can be defined by a set of rhombi. A rhombus is the locus of constant block-face distance (i.e., with a constant sum of distance in the x and y direction) Diagram 6.3 shows a skeleton encoding of a part of an area Algorithms for the computation of "maximal neighborhoods" are developed.[67]

This method is the best of the four for the determination of intersections, unions of areas and shading of areas, but it is relatively slow for obtaining the area and the perimeter of a region and is poor for the encoding of open curves

It might be appropriate here to go into some more detail about data structures before proceeding with the example

Geographical data are of two types, *content data* and *place data.*[68] Content data relate to recorded observations upon phenomena and place data relate to geometric characteristics of segments of space. Content data refer to observations on phenomena with a particular place and time-dimension Properties of size, shape, area, and connectivity in space are considered place data

Data have three dimensions *phenomena, time* and *location. Phenomena* can be in the form of variables, in which case they are entities measurable with an assumed origin (temperature) or a fixed origin (weight, age, etc) Phenomena can also be identified by their membership in classes which can consist of non-ordered groups denominated by codes (sex, race, etc), of ordered groups denominated by codes (social rank, occupation, etc.), or can be assigned to class-intervals with known distances between classes (age and income groups, etc)

Time can be a period (the Civil War, the nineteenth century, 1955, a month, etc) or a point in time (end of the year, 1st of July 1955, etc.)

Location identifiers assign places, lines, or areas to content and place data The simplest type is the external index which is a descriptive property or an address whose location has to be determined by master index, external to the system Examples are street addresses, (identifying points), county-names (pointing to a set of boundary coordinates), the name of a typhoon (indicating a list of coordinates and points in time showing the typhoon's movement through space and time), etc

Location identifiers in an ordinal scale are relatively rare but seem to be gaining in use. In environmental psychology, for example, people seem to be quite willing to rank perceived distances but unable to give any measurable notation.

The most accurate form of location identifiers in computer graphics is the coordinate form, whether it be an x,y system or any projection The conversion from one system to another has already been discussed

Data-Organization

Any geographical system must develop an efficient data-organization. "Efficient" means that the amount of memory for the data system is low, but that the data may also be arranged in such a way that they can be found without much searching and computation. Data organization can be built up from three basic structures *sequential, random,* and *list* Computer memories are sequential in access, that is memory locations are numbered serially If,

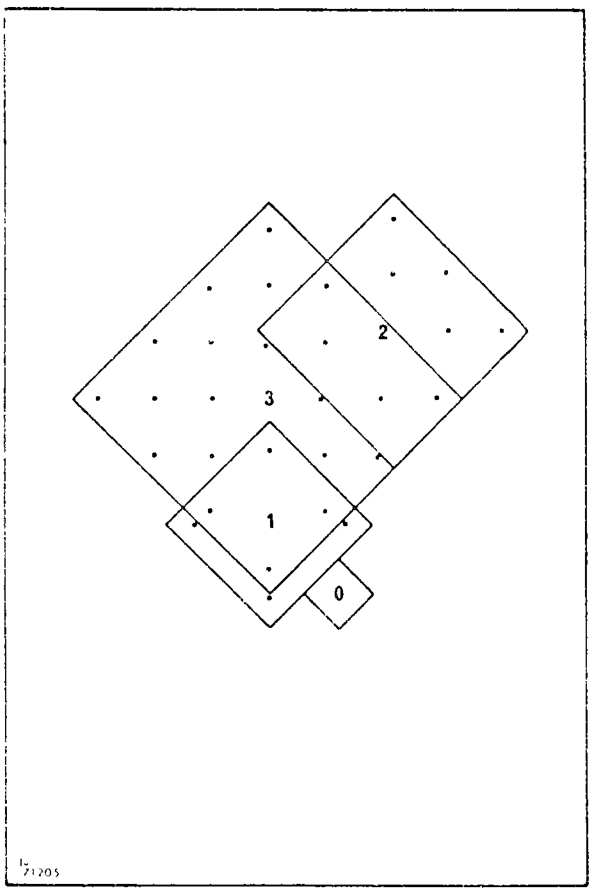


Diagram 6 3 Skeleton Encoding

[66] See Freeman, H (1961) "Techniques for Digital Computer Analysis of Chain Encoded Arbitrary Plane Curves," *Proceedings of the National Electronics Conference*, Vol 17, 421-432

[67] Pfaltz J L. and Rosenfeld, A , (1967), "Computer Representation of Planar Regions by their Skeletons," *Communications of the Association for Computer Machinery*, Vol 10, 119-125

[68] This terminology has been adopted from the Urban and Transportation Information Systems project at Northwestern University as described in Dueker, K (1966)

therefore, one is able to arrange one's data sequentially, one can call them ("fetch" them) without having to allocate storage space to the addressing procedure. A typical sequential data organization is a matrix (i.e., a rectangular two-dimensional array of data). Most of the census statistics would fall into this category.

Any data retrieval and manipulation which does not change the dimensions of the data set and the order of the data element is best performed with a sequential data organization. Records can be retrieved quickly, but insertion of a record and deletion of an old one is a difficult, slow process, because the whole file of records must be updated each time a change occurs. Since insertions or deletions of records happen very often in computer cartography, the sequential organization by itself is of little use for geographic and cartographic data files.

In the random data file, each record has a name or an "address" of its location in memory as a key for retrieval. The county boundary file could be arranged as a random file. On one hand, one would have a file containing pairs of x, y coordinates of all digitized points which would have an address implicit through their location in the file, on the other hand, he would have a series of "pointer" files, that is, for each county a file of addresses (numbers) which refers to the points contained in the boundary.

Both types of files are sequential in structure. The file of x and y coordinates contains all the digitized points and the position of a point becomes its address, the first point has the address "1", the ith point has the address "i". One boundary file contains an ordered sequence of addresses, that is, the numbers of the points forming the boundary in a clockwise direction. Although the structure of both types of files is sequential, their interconnection is random, that is by going down a boundary file the address can indicate any part of the point list, totally independent of the address of the previous or the subsequent pointer.

This method saves space on cards, tape, and disc when stored, and core space when operated upon, but it is not very convenient for several types of processings. We might, for example, add another set of regions to our county set in which the sets overlap in their units. This problem occurs frequently in geography, since different sets of statistical data are collected by different agencies in different areal units. We might want to find the set of intersections of the two sets of regions.

If we digitize two sets or regions at the same time, we can create a point file, and then three sets of boundary files (or two boundary files and the third computed from the two), the same way we created one before. However, if we digitize the two sets at different times, we will create new points when we intersect two regions. The coordinates of the points of intersection would be added to the bottom of the point file without difficulties, but it would take some manipulation to insert the addresses into the boundary fields. The insertion would require pushing all the addresses after the address of the inserted point, down one place, which is relatively time consuming.

Facilitating or speeding up the process is connected with an expansion of the storage area needed. The list structure avoids resorting of address files, but expands the files by adding pointers. In short, a list organization is one in which records are chained together by pointers. In other words, each record contains a pointer to the next record in a sequence where the pointer of the last record in the sequence refers back to the start of the sequence so that one can also follow through the sequence backwards. This case is also called ring structure, since the last pointer closes a ring pointing back to the first record. To insert records here means that the pointer of the previous record has to be chained to the address of the inserted record with the pointer of the inserted record referring to the address of the next record. Deletion is undertaken in a similar manner.[69] In the following, the usability of the different data structures in cartography will be shown by developing our example of a data bank for a state or province.

County and other boundaries are only rarely straight lines. They can be approximated with straight lines, however. This is adequate or even necessary for some mapping systems (as, for example, SYMAP) but renders very disappointing results with others (especially those based on the plotter as an output device). The question is how to create a data base which allows for both cases with a minimum of storage space.

One could consider certain points along a boundary as feature points and give them a special function. One could, for example, define all those points which one would use for a rough image on the line printer and have a set of points between them for the detailed representation of the line. In fact, storage space could be reduced even further by defining the "vectors" in incremental rather than absolute coordinates (i.e. by only indicating x and y distance of a point from the previous point rather than from an origin).

In order to guarantee flexibility for retrieval, we have to arrange our segment point in a list structure, however, we do not need a pointer for every vector, but can group a number of vectors together in one record with a pointer to the next record. Seven vectors would be an example.[70] (Diagram 6.4)

For many applications, only the feature points would be necessary. For example, a first test of intersection of two

[69] The interested reader can find more on data structures in Williams, (1971) and Dueker, 1966, Chapters 8 & 9.

[70] In the IBM/360 systems, four words would be needed for this, fourteen bytes for the point and two bytes for the pointer.

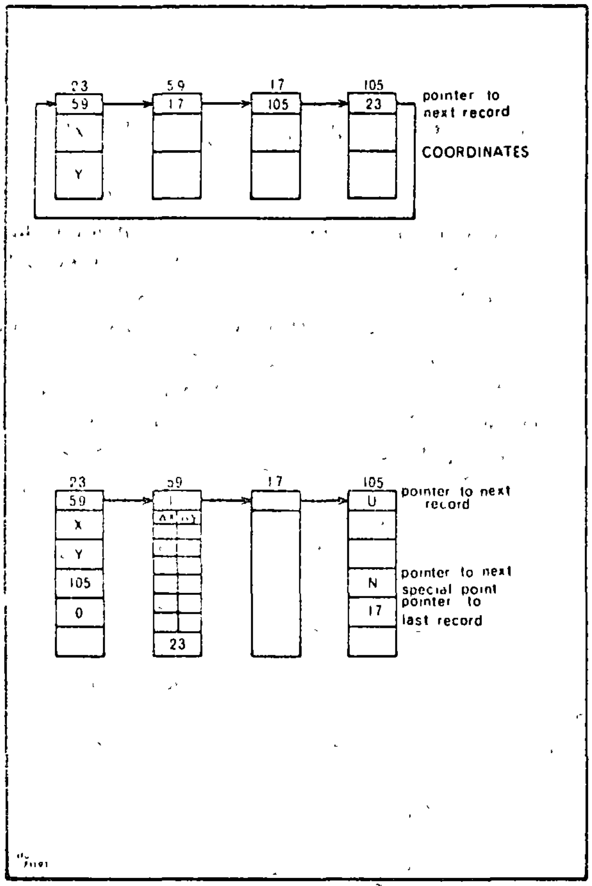


Diagram 6.4 List Structure of Line Encoding

areas could be done with them, or the computation of the area of a county if the result did not have to be extremely accurate. These points should therefore not only have pointers to the next record of segment points, but also one to the next end-point of a segment.

A number of segments represents a region by its boundaries. At this level, statistical tables come into the picture. Therefore, we not only need a file containing the addresses of the starting points of the segment which comprise the region but also pointers to some statistical tables. In other words, the statistical tables require addresses which can be connected to the addresses of the region.

A set of one sort of regions (counties, etc.) comprises another type of regions (states, etc.). Algorithms can be developed which compute the boundary for the "hyper-region"[71], the area of the hyper-region, etc.

Line-information such as streets and power-lines, and point information such as settlements or surface-heights can be stored similarly. They belong, however, to different "classes" or data. They can be included, however, in the data bank, described and recalled together with boundary and statistical information.

---

[71] The concept is very simple each segment not shared by two regions must be part of the hyper-region's boundary Finding the right sequence of segments is somewhat more difficult

VII. COMPUTER CARTOGRAPHY FOR WHOM

Most students of computer cartography did not become interested in the subject because they saw its great potential for geography, planning and/or other sectors of society, nor because they realize the many theoretical problems inherent in the development of graphic procedures, but simply because the newness of the subject and the unusual appearance of computer maps promised "a lot of fun" Later on, other things became more important, although students might still feel excitement when producing a new map.

One might distinguish between two types of maps intermediate and final ones Intermediate maps serve as an aid for further analysis during a research project, whereas final maps are part of the result of the project Theoretically, every map is intermediate since it serves as a "data-bank" and can always be used for another study, but it makes a decisive difference for the researcher's attempts at attaining high map quality to know that he will use the map only for his own information, or largely for the information of others.

The intermediate map can play a very important role in a recursive research cycle of hypothesis formulation, data gathering, data analysis, and reformulation of hypotheses,[72] since the map is the only device which stores data in a two-dimensional context The maps do not have to be of high quality for this purpose, but they do have to be produced quickly Computer maps fulfill exactly these requirements and many programs (i.e , SYMAP) were initially developed just for this purpose

Speed is not the only factor. The cost component has already been considered in a very early stage of the development Tobler gives the following cost comparison[73]

> To draw a 17 x 22 inch map of the world on some obscure projection costs about $4 00 for the plotting, $6 00 for the computing and perhaps $10 00 for the setup time I do not know any place you can get a draftsman to do a 17 x 22 inch map of the world on some strange projection for less than $100 00

The cost-comparison would not be as favorable in other mapping areas, but one has to expect a considerable reduction of computer costs in the next decade, so that what might seem expensive in the student's freshman year will be cheap by the time he graduates.

Another criterion for the advancement of computer cartography is the need for automation because of lack of personnel for the manual production It is estimated that at the most, 40% to 50% of the earth's land surface has been mapped at a scale of 1 100,000 or larger and only 3 to 4% at the scale of 1 25,000 or larger.[74] The reason is not so much the lack of survey data which could be gathered through aerial photography in a relatively short time, as it is the mapping methods which cannot keep pace with the provision of data Therefore, new types of maps have been proposed and developed such as the data-matrix[75] and the orthophoto map, a map type based on aerial photographs At the same time, the automation of manual methods is progressing rapidly.

High-Accuracy Systems

Another factor in the promotion of computer cartography is the increasing amount of cartographic data collected in machine-readable form A good example is the system of automatic cartography developed for the Canadian Hydrographic Service [76] The input of this system is survey data recorded at sea in digital form However, existing maps are also digitized to take advantage of new and old data in a cohesive manner The aim of the system is to draw chart overlay sheets automatically which should have the quality and accuracy acceptable to the user and to the Service

Eventually, most of the data will be soundings recorded automatically on sounding vessels in machine-readable form (i e , on magnetic tape) In the meantime, however, large

---

[72] See Haggett, P (1965) *Spatial Analysis in Human Geography*, London

[73] In Tomlinson, 1970

[74] After Yoeli P (1967) "Topographische Karten und Rechenautomaten," *Kartographisch Nachrichten* Vol 17, 116-118

[75] Yoeli, P (1967), *ibid* and Aumen W C (1970) "A New Map Form Numbers," *International Yearbook of Cartography* Vol 10, 80-84

[76] Boyle A R , (1970), "Automation in Hydrographic Charting," *The Canadian Surveyor* Vol 24, 519-537

volumes of data have to be converted from hydrographical charts. A digitizer is interfaced with a minicomputer for reliability control. For example, the computer stops the process if the operator moves the tracer too fast, it joins the end of a contour-line to its start, etc. All data are stored on magnetic tape.

Two output-units are provided One is a high-accuracy flatbed plotter with a lightspot pen Ink pens and scribers are of little use to automatic systems since they dry up, blob, etc. The plotter is driven by another minicomputer for accuracy control The other output-unit is a storage tube CRT for compilation purposes. A man-machine arrangement is used to employ the opinion and decision-making capability of a cartographer One group of interactions, the "interactive data base manipulation," mainly involves the comparison of old and new data, deciding which are better, and shifting and reorienting them to fit the two sources to exactly the same reference base The second group, computer aided compilation, involves checking the output data arranged by the computer for chart drawing and modifying the selection of data or details of line or symbol work, as considered most appropriate by the operator Two dimensional interaction is performed with the aid of a "mouse," a manual interactive unit which runs on two wheels in perpendicular position one to another By pushing the "mouse" around on a plane surface, one moves a point on the screen to initiate line additions, deletions, etc. (Figure 7.1)

A typical manipulation of the CRT goes as follows The computer reads line and point information from tape onto disc to allow faster operation. The operator then displays the picture on the screen, enlarges a section (command zoom), corrects certain parts with the mouse and then shifts (command pan) to another subsection All the changes are recorded on disc (together with the initial data) and read onto tape at the end of the job

In this manner, input is provided through a minicomputer onto tape and output is from tape to plotter or CRT via minicomputer. Between those two steps, several computations are performed which have to be based on a large computer. This can be done easily since all the data are stored on compatible tape. Some of the functions are to join separately digitized sections of a chart, correcting for shift and orientation, taking data from the data-bank and making it suitable for machine drawing with lines in a smoothed form Diagram 7 1 shows a chart of the whole system This system is of special interest for the geographer for it shows the successful application of interactive devices for spatial problems Presented below are some ideas of an interactive cartographic unit and a few problems geographers and cartographers could tackle with it

Small Interactive Systems

The basic units of such a system would be a digitizer maybe operating with a magnetic pen, a storage tube CRT, a disc-drive for intermediate data storage, a tape-drive for input and output, a teletype for text-interaction and programming, all interfaced to a minicomputer. The digitizer would not only function as a digitizer, but also as pointer and tracer in an interactive mode with every move being recorded on the screen The storage tube is a special CRT which does not need refreshing of the picture (e g , 30 times a second) This cuts down costs tremendously but is also slower However, a few seconds to fill the whole screen is still very fast for the geographer's needs

A frequent problem is the display of data by statistical area (choropleth map) The determination of class-intervals needs experimentation, since equal intervals often lead to serious suppressions of valid information, and even their determination on the basis of frequency distributions does not always supply the gray-tone distribution desired. The system therefore should include the possibility of modifying class intervals through the teletype and representing them in different gray-levels The procedure could be accompanied by different statistical computations, histograms, etc on the small computer, whereas the actual creation of the value-grid would be performed on the large processor

The minicomputer could play different control functions during the digitization process, such as maximum movement control during the line-digitizing, point and shape control by displaying coded points and lines concurrently on the tube, etc. An interactive sampling procedure could also be developed First, one would have to code a relatively wide-meshed net (regular or irregular) of a surface and then compute an estimate of each point's value, using different interpolation algorithms If the difference between actual and estimated values is more than a given threshold for a point, a narrower net would have to be coded around the point.

All map computation procedures discussed earlier could also be implemented (i e , line correction, addition and deletion, interactive name and symbol placement, etc). One could actually produce several overlays for one map interactively, high resolution gray-tone sheets on a half-tone recorder and combine the sheets to one map by using different color filters Such a map is already possible, although its quality would not yet reach that of a manually fabricated map

Another possible use of the cartographic unit could be its development as a test unit for map usage At this time when everybody is talking about new map needs and many new map types are being produced, the knowledge about

Figure 7 1 Computer Aided Compilation Operator can change lines, add data sets (b), zoom the picture to see a section in more detail, (c, d, e,), etc. Courtesy A R Boyle, University of Saskatchewan

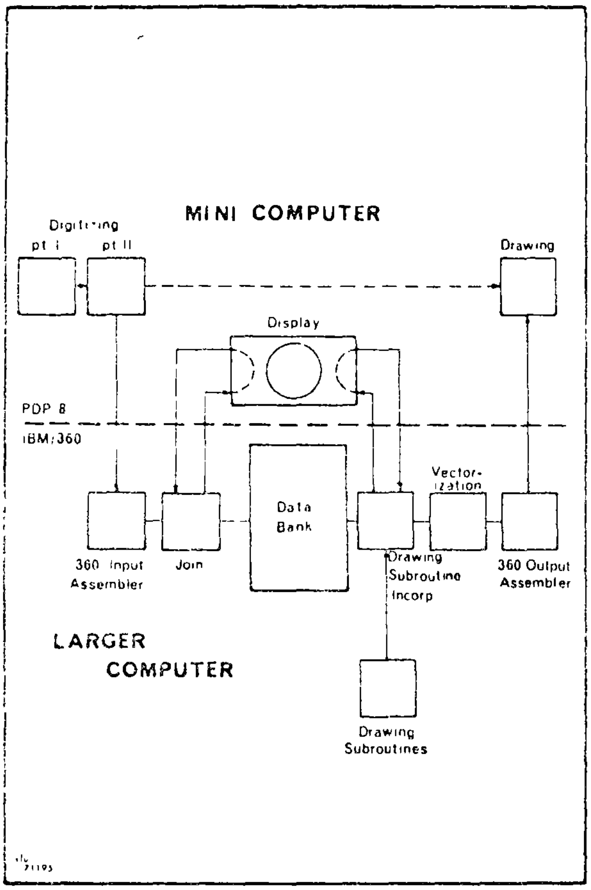


Diagram 7 1 The Automated Cartography System for the Hydrographic Service of Canada

the usefulness of certain map types for certain purposes is very small Psychophysical tests are rarely made, and if so, they are usually on abstract patterns and not on actual maps With the cartographic unit, highly sophisticated tests could be performed For example, suppose that several methods of relief representation are to be tested for their suitability for different map uses One map type would be mounted onto the data grid and several control points would be digitized to bring the map in concordance with the computer's coordinate system The first question would then appear on the screen, for example "Please draw with the magnetic pen all the valleys you can find on the map " The computer would not only record all the valleys found and test for the accuracy of the guess, number of valleys of a certain type found, etc , but also measure the time it takes the test object to find different features Other answers could be typed into the teletype, etc

Finally, it might be mentioned that the cartographic unit could of course also perform all jobs commonly dedicated to small computers. Some of these are computer-assisted instruction in this instance improved for geographers by the availability of the CRT - interactive text editing, key-word retrieval of bibliographies, etc The usage is unlimited The hardware for such a system is available at costs which are manageable now by large departments and in five to ten years, by medium-sized departments However, the software for such a system is to a large degree nonexistent, and might be for years.

Applications

Most of those who have a firm experience in computer cartography—at this time mainly planners rather than geographers—are not interested in working on theories and developing procedures, but rather want to use whatever programs are available. And one cannot blame them The demand for people to use their knowledge is quite high, as is the financial incentive Furthermore, the problems they are commissioned with are usually so interesting that they can only pity us theoreticians. Some of these problems shall be presented here.

"Northstar" is a proposed ski resort in the vicinity of Lake Tahoe Several experts had given conflicting evaluations of the "best" places for the ski runs, the developers were looking for a study based on a quantitative and cartographic evaluation of areally-collected data.[77] The study involved collecting data on a grid basis for the variables of topographic elevation, vegetation type, vegetation density, vegetation age and hydrology Previously collected data included sun angles and locations at different times during the day and spot samples of snow depth for four years From this data, several other variables were derived such as slope gradient and direction and sun intensity in heat transfer units (Figure 7 2)

A multiple regression analysis was performed with thirty snow sampling stations through four years with elevation, slope, vegetation, sun intensity and water areas as independent variables The result was a map of snow depths which gave some preliminary results on snow quality To get more information on snow quality, a questionnaire was sent out to ski consultants and people responsible for ski slope grooming at existing ski resorts

The combination of all the variables provided the basis for the final ski slope suitability model Each cell on the map represented the "Goodness" of the spot for skiing in ten levels. Areas either too steep for skiing or having other constraints, such as water areas or marsh lands, were not considered Areas which were excellent for skiing in all variables except vegetation were defined by the symbol "V" indicating that these areas could be cleared and would then become acceptable skiing areas. (Figure 7 3)

[77] The following has been described in a letter to the author by J Dangermond, Redlands, California

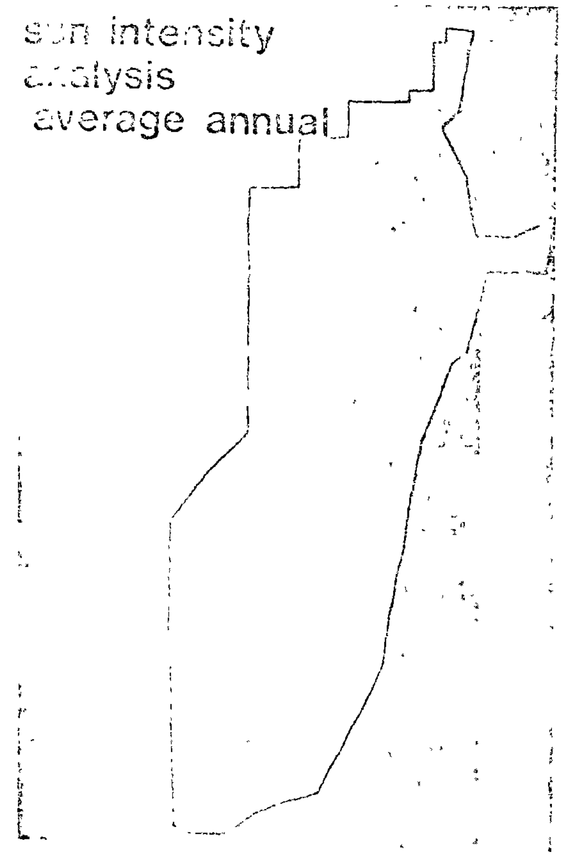


Figure 7 2 Northstar Development, Sun Intensity Analysis Courtesy J Dangermond, Redlands, California

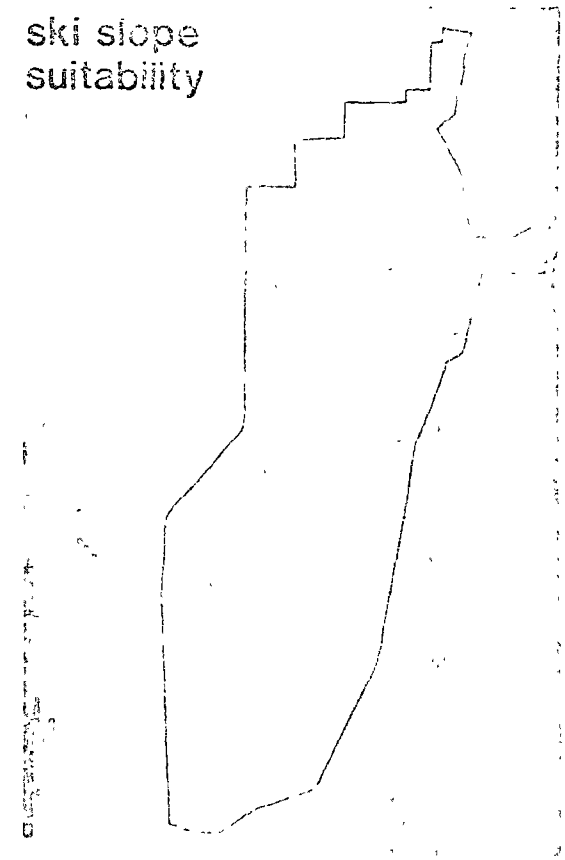


Figure 7 3 Northstar Development, Ski Slope Suitability Courtesy J Dangermond, Redlands, California

The combination of various cells to ski-hills was done manually. One could think of an extension of this study by a downhill "climbing" procedure with gradient constraints to find ski slopes for the different experience-classes, etc.

The other study to be discussed here is a systems analysis for planning the multiple use of the Honey Hill area of Swanzey, New Hampshire [78] The study gives a good example of different resource evaluation methods in a practical application The data for the study have been derived from aerial photographs and field surveys at a regular grid at the scale of 1/100 sq km (2.4 acres). The second aspect of the analysis was the development of a series of quality indices, such as visual quality, ecological damage, wildlife habitat quality and others. These were developed as models. Site attractions or constraints were measured in these terms for a variety of recreation types and other activities. These use-quality evaluations were rank-ordered, thus leading directly into a planning evaluation process for site development The environmental quality models were used as an integral part of the planning and evaluation system which constituted the third aspect of the study (Figure 7 4)

These models were to investigate the implications for environmental quality of various potential uses of a particular study area This step involved the application of the formal mathematical models of systems analysis Three different approaches to the development of plans for the site area were investigated

The first approach used a linear programming model Given certain physical, ecological and economic constraints, the objective function of the model was the maximization of net benefits for the development of the Honey Hill reservoir. The result of the model provided "optimal" development proposals within some fifty predefined zones. It also gave parametric functions of the system's response to changes in the demand parameters.

The second approach involved "best professional judg-

[78] Rogers P, C Steinitz, F Murray, D Sinton, R Toth, D Way *Honey Hill A Systems Analysis for Planning the Multiple Use of Controlled Water Areas*, U S Army Corps of Engineers Research Contract, Graduate School of Design, Harvard University, 1971 The description in the text is extracted from Steinitz, Rogers Associates Inc (1971) Selected Projects, Cambridge, Mass

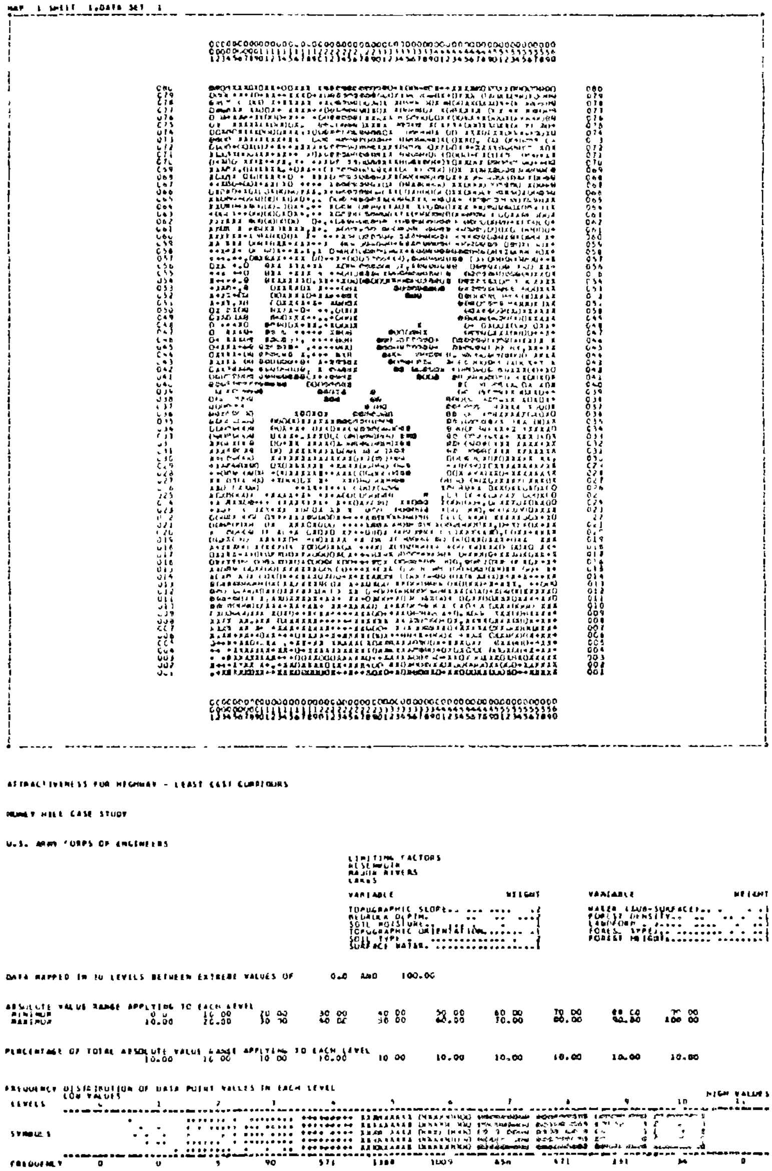


Figure 7 4 Honey Hill, Attractiveness for Highway Courtesy C Steinitz, Harvard University

ments" to develop the various plans by the members of the research team based partly on the analysis produced in the first two stages of the research The third approach was the development of a simulation model. Total demand was given, but activity preferences were keyed to nine combinations of income and travel time-distance to the site. The model was used to evaluate the alternatives arising from the "best professional judgment" and to test the plans implied by the linear programming model (Figure 7 5)

The value of such studies is evident They are based on a solid data base This data base is brought to its optimal use by the combination of analytical techniques and immediate mapping The problem with conventional cartography in planning is that the mapped data are out of date before the map is even finished and the cartographer is incapable of reacting to the fast-changing perspectives and needs of the planner and politician The result is obvious Planning atlases suitable for anything but planning, a delay in the adjustment process because these atlases become law by indolence, others are encouraged to produce other planning atlases because they look so good

The modern regional and planning atlas should not be bound. Actually, the ideal atlas would be a well-organized data bank which could be called upon when a certain

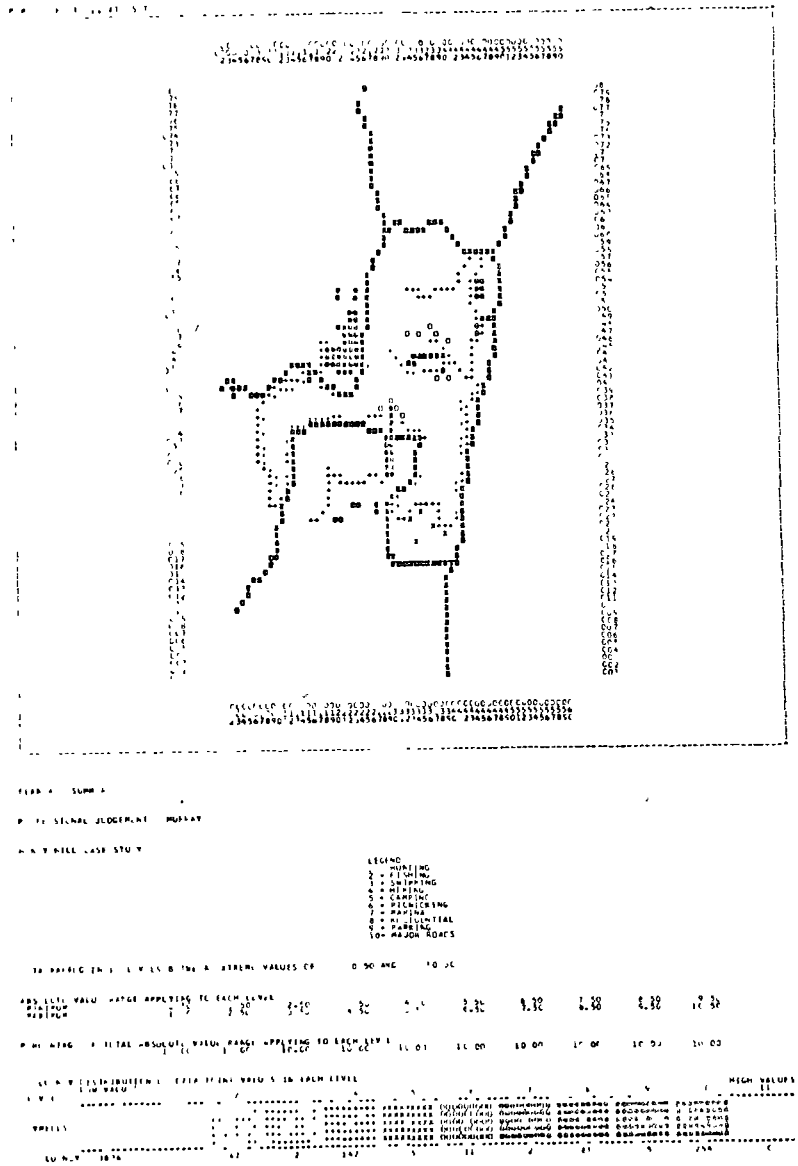


Figure 7 5 Honey Hill, Professional Judgment, Murray Courtesy C Steinitz, Harvard University

question comes up This would guarantee quick responses with the latest data Some questions could be answered immediately by calling a particular variable, some would involve little or very complex computations Some maps would consist only of tables, others of quick printer maps, a few could be prepared with great care and special hardware to attract the interest of a large public (Figure 7 6)

Computer cartography is increasingly gaining attention in many disciplines and public activities Geography is not the hardest pusher in the field, but geographers are needed to provide the link between theory and practice Much of the opposition against computer cartography comes from cartographers who are afraid of being replaced by the computer This attitude may be ridiculous, for the new possibilities may create more work for cartographers.

Anybody in the field will agree that one will find ten good programmers before one can find one good cartographer.

In the next decades, the work of many cartographers will change from drafting to cartography as a science with its own theory, techniques and applications. This event is not necessarily caused by the computer but by a considerable change in the requirements of the public. The implication, however, is that this trend will involve a new type of cartographer with expanded interests and expertise, in which computer cartography will play an important role

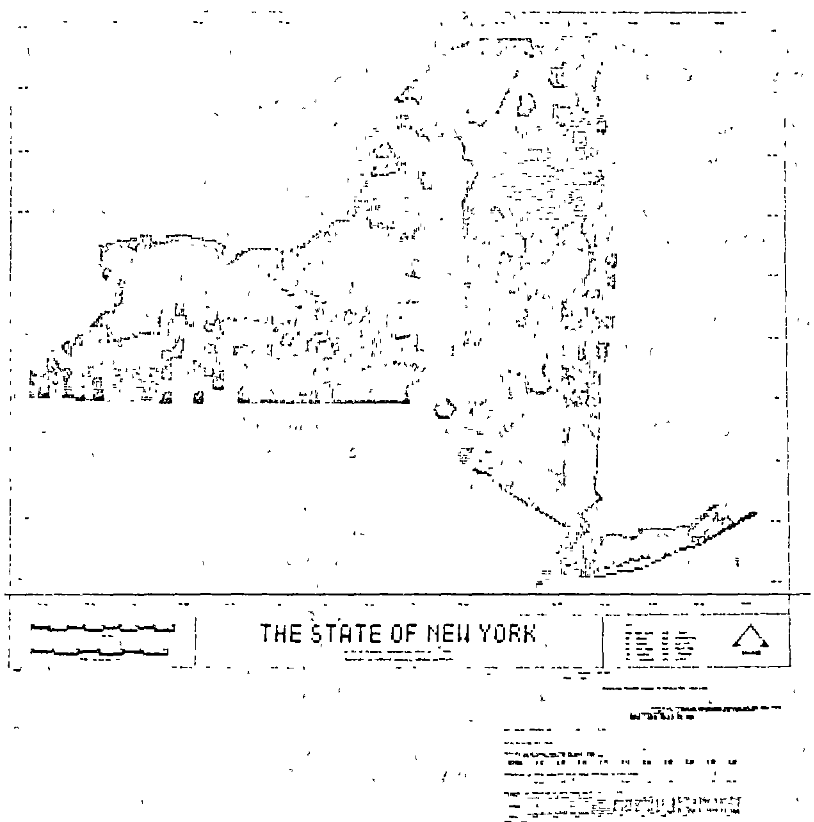


Figure 7 6 New York, Predicted Population Change 1960-1990 Courtesy C Steinitz

FURTHER READING

In the following some literature for the interested student is cited. Since the subject is relatively new, the items appear in some unusual publications and are sometimes difficult to obtain.

Balchin, W. G. V. and A. M. Coleman, 1967, "Cartography and Computers," *The Cartographer*, Vol. 4, 120–127.
A short and lively introduction.

Board, C., 1967, "Maps as Models," in Chorley, R. J., and P. Haggett, 1970, *Models in Geography*, London, pp. 671-725. A good and readable introduction into many conceptual ideas in cartography.

Boyle, A. R., 1970, "Automation in Hydrographic Charting," *The Canadian Surveyor*, Vol. 24, 519–537.
The description of a system of high accuracy with some interactive components.

Chorley, R. J. and P. Haggett, 1965, "Trend-Surface Mapping in Geographical Research," *Transactions, British Institute of Geographers*, No. 37, 47–67.

Dacey, M. F., 1970, "Linguistic Aspects of Maps and Geographic Information," *Ontario Geography*, No. 5, 71–80.
An overview of problems of map languages.

Dodd, G. S., 1969, "Elements of Data Management Systems," *Computing Review*, Vol. 1, 117–133.
A detailed tutorial on data organizations.

Dueker, K., 1966, Spatial Data Systems, I, II and III. *Technical Report*, Department of Geography, Northwestern University, AD652005/6/7, Clearing House for Federal Documents, Springfield, Va. 22151.
A study of geographical data systems, but at the same time a good tutorial on problems connected with this area. Very readable since it was intended as a tutorial.

Gould, P., 1970, "Computers and Spatial Analysis: Extensions of Geographic Research," *Geoforum*, Vol. 1, 53–69.
A general introduction.

Hagerstrand, T., 1967, "The Computer and the Geographer," *Transactions British Institute of Geographers*, No. 42, 1–19.
A more detailed overview with some remarks on the Swedish coordinate grid system and some examples of Nordbeck's work in Computer Cartography.

Harbaugh, J. W. and D. F. Merriam, 1968, *Computer Applications in Stratigraphic Analysis*, New York.
A "standard" book with chapters on mapping techniques trend analysis (time, polynomial, harmonic), classification, and simulation.

International Business Machines, Data Processing Application, *Numerical Surface Techniques and Contour Map Plotting*, IBM E20-0117-0.
A good discussion of the basic problems of automatic contour plotting.

Interactive Graphics in Data Processing, IBM Systems Journal, Vol. 7, No. 3 and 4, 1968.
A special issue on computer graphics with many fundamental articles on graphics concepts, geometry, languages, and applications. Generally not too technical.

Lipkin, B. S. and A. Rosenfeld (eds.), 1970, *Picture Processing and Psychopictorics*, New York.
Many problems pertinent for computer cartography are discussed here from the viewpoint of picture processing. Therefore very important but demanding readings.

Marble, D. F., 1967, *Some Computer Programs for Geographic Research*, Department of Geography, Northwestern University, Evanston.
A collection of geographical programs in FORTRAN IV.

Narasimham, R., 1969, "On the Description, Generation, and Recognition of Classes of Pictures," pp. 1–42, in Graselli, A. (ed.) *Automatic Interpretation and Classification of Images*, New York, 1969.
An authoritative discourse on picture languages.

Negroponte, N., 1970, *The Architecture Machine*, Cambridge, Mass.
A lively excursion into the questions of artificial intelligence and man-machine interaction in architecture and planning.

Nordbeck, S., 1968, "Coordinate Mapping Techniques," *Plan, Tidskrift for planering aw landsbygd och tatortor*, Special Issue, "Urban and Regional Research in Sweden," 1968, 101–117.
Nordbeck has built a complete system of computer cartography on the basis of the square grid "coordinate real estate register" in Sweden. Isarithmic maps, point in polygon programs, range, fraction, carriage, correlation, traffic and potential maps are some of the developments. Some of the procedures are published in Nos. 7, 8, 9 of the Lund Studies in Geography, Ser. C.

Pfaltz, J. L. and D. L. Milgram, 1970, *An Experimental Map Description System*, Technical Report 70-130 GJ 754, Computer Science Center, University of Maryland, College Park, Md.
A very readable description of a mapping system based on bitplanes, with examples. Very good as an introduction.

Rayner, J. N., 1971, *An Introduction to Spectral Analysis*, No. 2, Monographs in Spatial and Environmental Systems Analysis, London.
Does not demand any special mathematical knowledge. Guides the student to a competent use of this flexible method. Unique in the discussion of two-dimensional spectral analysis.

Robinson, A. H. and R. D. Sale, 1969, *Elements of Cartography*, New York.
The student of computer cartography needs to study general cartography thoroughly. There are several good texts, this one shows a few computer maps.

Rosenfeld, A., 1969, *Picture Processing by Computer*, New York.
A mathematically demanding treatment of this subject.

Tobler, W. R., 1967, "Of Maps and Matrices," *Journal of Regional Science*, Vol. 7, 275–280.
An elegant treatment of numerical filtering and surface generalization.

Tobler, W. R., 1969a, "An Analysis of a Digitized Surface," pp. 59–86, in C. M. Davis *A Study of the Land Type*, Department of Geography, University of Michigan, ORA Project 08055.
Ten different filters were applied to a surface. The results were correlated and three groups extracted, (elevation, slope, roughness)

Tobler, W. R., 1969b, "Geographical Filters and their Inverses," *Geographical Analysis*, Vol 1, 1969, 234–253
Puts the discussion of geographical filters on a theoretical level.

Tobler, W.R., (ed.), 1970, *Selected Computer Programs*, Department of Geography, University of Michigan, Ann Arbor, 1970. 18 programs of geographical interest with short descriptions.

Tomlinson, R. F. (ed.), 1970, *Environment Information Systems*, The Proceedings of the UNESCO/IGU, First Symposium on Geographical Information Systems, Ottawa, September 1970 A publication of the IGU Commission on Geographical Data Sensing and Processing, Ottawa
A good discussion of the state of the art in large geographical data systems.

Walters, R. F., 1969, "Contouring by Machine A User's Guide," *The American Association of Petroleum Geologists, Bulletin*, Vol. 53, 2325–2340.
All one should know when using contouring programs.

Warntz, W., 1966, "The Topology of a Socio-Economic Terrain and Spatial Flows," *Papers, Regional Science Association*, Vol 17, 47–61.
Discussion of surface-specific features and applications.

Williams, R., 1971, "A Survey of Data Structures for Computer Graphics Systems," *Computing Surveys*, Vol. 3, 1–21.
Types of data structures, languages for creating and manipulating data structures, examples of implemented data structures.

GLOSSARY

Many terms are based on Sippl, C J (1966), *Computer Dictionary*, Indianapolis, ACM/SIGRAPH (1971), Computer Graphics Glossary, *Computer Graphics*, vol 5, No 1, 1–14, and *Definition of Terms in General Use for Automated Cartography*, International Cartographic Association (1971), Commission III on Automation in Cartography

Many terms have different meanings. The one used in this research paper has been defined

access – concerns the process of obtaining data from, or placing data in, storage.

accuracy – freedom from error

address – a label, name or number identifying a location or unit where information is stored

algorithm – a defined process or set of rules that leads and assures development of a desired output from a given input at sequence of formulas and/or algebraic/logical steps to calculate or determine a given task.

analog – the representation of numerical quantities by means of physical variables (e g., translation, rotation, voltage, or power), contrasted with digital.

binary – a numbering system based on twos rather than tens which uses only the digits 0 and 1 when written

bit – abbreviation of binary digit May be equivalent to an on or off condition, a yes or no, etc A unit of information capacity of storage device. The capacity in bits is the logarithm to the base two of the number of possible states of the device.

break – the point of sudden change of slope along a scanning line.

byte – a group of binary digits usually operated upon as a unit Usually 6 bits or 8 bits

CAD – Computer Aided Design

CAI – Computer Aided or Assisted Instruction

card – a machine-processable information-storage medium of special quality paper stock, generally 7⅜ by 3¼-inches, containing 80 columns and 12 punch positions.

cathode ray tube – CRT – An electron tube whose face is covered with a phosphor that emits light when energized by its electron beam.

centroid – the mean position of a polygon Determined by the x and y means of the polygon points

command – the portion of an instruction word which specifies the operation to be performed

compaction – reduction of storage space using an efficient storage system.

compatible – executable on more than one computer

compression – reduction of storage space by selecting data from a large data-body

contiguous – adjacent or adjoining.

contrast – the relationship of the brightest to the darkest portions of a display image.

coordinate – an ordered set of data values, either absolute or relative, which specifies a location

CRT display – display using cathode ray tubes as the viewing element (raster scan, storage tube, directed beam)

cursor – a movable marker visible on a CRT display used to indicate the position at which the next operation (insertion, replacement, erasure) is to take place

data-point – point on a surface, given by its coordinates before computations (see grid point).

data structure – the arrangement and interrelation of records in a file.

default option – programmed command which takes effect if the respective elective is not specified

digital data – information represented by a code consisting of a sequence of discrete elements.

digitizer – a device that codes images into digital computer-usable form

digitizing – convert analog measures (e g , length) into digital form.

directed beam – also called vector mode In the CRT method these are the elements of a display image where the beam motion is in straight lines from point to point (see raster scan)

file – collection of related records treated as a unit A collection of informational items similar to one another in purpose, form, and content

flag – indicator used to tell some later part of a program that some condition occurred earlier

graphic language – software interface between the programmer and the display device.

grid point – point on a surface, defined in a regular net of grid points, usually computed from data points by interpolation

halftone device – graphic instrument which can produce several graytones besides black and white

hardware – the mechanical, magnetic, electrical and electronic devices or components of a computer

hidden lines – line segments obscured from view in a projected image of a three-dimensional object

input – information or data transferred or to be transferred from an external storage medium into the internal storage of the computer

intersection the area that two overlapping regions have in common

isarithm – surface-line of constant z value Also called contour-line

juxtaposition – the positioning or placing of items adjacent to each other or side by side

label – an identification device for introducing a record, groups of records, or an address

light pen – a stylus which detects within a limited area (the aiming circle) light generated on a CRT to determine either positional or display element identifying information

line printer a printer in which an entire line of characters is composed and printed at a time

list – data structure which divorces the logical organization of a file from its physical organization by employing pointers to indicate record-sequences

memory – an organization of storage units, primarily for the retrieval of information

mouse – a hand-held device, with two perpendicular wheels, which is rolled around on a flat surface to provide coordinate input to the display device

noise – errors introduced into data during measurement or display

off-line – a system where the peripheral equipment is not under the control of the central processing unit, as it is in the on-line case

origin – a reference point whose coordinates are all zero.

output – computer results such as answers to mathematical and statistical problems, plots, etc

plotter – a graphic device for making permanent copies Usually with a moving pen Plotters can be digital incremental or analog (continuous), drum or flat-bed plotters, etc

pointer – an address in a record which refers to a related record

precision – the degree of exactness with which a quantity is stated, contrast with accuracy, which refers to the absence of error regardless of precision

raster scan – a technique for generating or recording an image with an intensity (z value) controlled line-by-line sweep across the entire display surface Also called point mode display

record – a set of one or more consecutive fields of related data items

redundancy – the fraction of the gross-information content of a message that can be eliminated without a loss of essential information

refresh rate – the rate at which a display is regenerated in order to remain visible

residual – the difference between the estimated and the actual value of generated line or surface at a data point

resolution a measure of the ability of a device to differentiate value, e g plotter step size or raster unit retrieval – the act of finding stored information.

secondary storage – storage whose primary function is to augment the capacity of internal storage for handling data and instructions. Must be transferred to internal storage to become operable. Usually on disc, drum or tape.

software – the internal programs prepared to simplify programming and computer operation. More generally every type of computer program.

storage tube – a CRT which retains an image for an extended period without refreshing.

union – the area covered by either or both overlapping regions.

MAJOR PROGRAMS USED FOR THE PRODUCTION OF THE FIGURES

BLOCKS – Plots planar projections of three-dimensional block models. Author: R. M. Ray, University of North Carolina.

CALFORM – Produces choropleth maps on the line plotter. Author: J. Cartright, Harvard University.

CNTOUR – Subroutine to compute one contour in a regular grid of points. Author L. Coulthard, University of British Columbia.

FAKAN – Factor analysis (principle axis method) Author: W. D. Rase, Bonn-Bad Godesberg.

GEOMAP – Line-printer program for choropleth, proximal, and contour maps. Author: D. Steiner, University of Waterloo.

INCLIN – Program for the Orthographical Relief Method Authors: T. K. Peucker, M. Tichenor; Ottawa; and W. D. Rase.

KOPPE – Selects smaller grid from surface grid, computes absolute and mean errors of interpolation and plots residuals Author T. K Peucker.

PAX – Parallel picture processing Universities of Illinois and Maryland.

PERS – Plots perspective views. Author L. Coulthard.

PFEIL – Subroutine to plot flow-patterns. Author. W. D. Rase

RAPP – Resource Analysis and Presentation Programs to overlay distribution data Author T. Stanhope, University of Waterloo.

SHDCTR – Produces Relief Contour Method Author T. K. Peucker and M. Tichenor.

SIM – Simulates land-use patterns and plots results. Author S. Witiuk, Simon Fraser University.

SIRKEL – Produces circles for absolute data plotting with deletion of overlapping parts. Author: D. Hatleiid, Simon Fraser University.

SLOPE – Subroutine to compute slopes in a surface. Author: W. D. Rase.

SMOOTH – Subroutine to smooth a surface on the basis of a moving average. Author W D Rase.

SPATFUN – Outputs rectangular array of z-values from a choice of 13 spatial functions with user-supplied parameters Author T. K. Peucker

SUPERMAP – Draws maps and other data in a number of projections. Author R. L Parker, University of California, San Diego.

SYMAP – Flexible program to produce different types of maps on a line-printer. Author Laboratory for Computer Graphics and Spatial Analysis, Harvard University. Extensions (e g. inclusion of polynomial trend-surface, point distribution coefficient, entry for user-supplied subroutine), W. D Rase

SYMVU – Plots perspective views of regular grids. Authors F. Rens, State University of New York, Buffalo and Laboratory for Computer Graphics and Spatial Analysis, Harvard University

VIEW – Computes visible areas on a grid of height from an observation point and plots results Author. T K. Peucker, after Amidon & Elsner.

VIEWBLOK – Plots perspective views of regular grids. Author D Douglas, University of Ottawa.

WATERSHED – Surface analysis based on slope analysis. Author D Shepard, Harvard University

YOELI – Analytical Hillshading Authors: M. Tichenor, T. K. Peucker and W. D. Rase.

ASSOCIATION OF AMERICAN GEOGRAPHERS

COMMISSION ON COLLEGE GEOGRAPHY
Membership-January 1972

Dr Robert H T. Smith, Commission Chairman, Queen's University
Dr. John S Adams, University of Minnesota
Dr Ian Burton, University of Toronto
Mr. Richard D Dastyck, Fullerton Junior College
Dr. Paul W. English, University of Texas
Dr. Gordon J. Fielding, University of California, Irvine
Dr J Thomas Hastings (Education), University of Illinois
Dr Leslie J King, McMaster University
Dr. Richard E Lonsdale, University of Nebraska
Dr. Ian R. Manners, Columbia University
Dr. Marvin W. Mikesell, University of Chicago
Dr. Harold M. Rose, University of Wisconsin–Milwaukee
Dr. Harold A. Winters, Michigan State University

Ex-Officio Members, President, Edward J Taaffe; Vice President, Wilber Zelinsky, Secretary, Wesley C. Calef; Treasurer, Donald A. Wise, Executive Director, J. Warren Nystrom

Dr. John F. Lounsbury, Project Director, Arizona State University, Tempe, Arizona 85281
Dr. Salvatore J. Natoli, Associate Director, Association of American Geographers

THEMATIC CARTOGRAPHY

Phillip Muehrcke

COMMISSION ON COLLEGE GEOGRAPHY
RESOURCE PAPER No. 19

ASSOCIATION OF AMERICAN GEOGRAPHERS
WASHINGTON, D.C.
1972

Supported by a grant from the National Science Foundation

ASSOCIATION OF AMERICAN GEOGRAPHERS

Commission on College Geography Publications

General Series

1. [illegible] in Undergraduate Liberal Education, 1965
2. A Basic Geographical Library A Selected and Annotated Book List for American Colleges, 1966
3. [illegible] Manpower—A Report on Manpower in [illegible] Geography, 1966
4. New Approaches in Introductory College Geography Courses, 1967
5. Introductory Geography—Viewpoints and Themes, 1967
6. Undergraduate Major Programs in American Geography, 1968
7. [illegible] Survey Course The Energy and Mass Budget at the Surface of the Earth, 1968
8. A Systems Analytic Approach to Economic Geography, 1968
9. [illegible] Geographical Bibliography for American College [illegible], 1970
10. Geography in the Two-Year Colleges, 1970

Technical Paper Series

1. [illegible] Training in Geography, 1968
2. [illegible] Assisted Instruction in Geography, 1969
3. [illegible] Geography Courses A Model with Illustrative Applications, 1970
4. [illegible] Maps of the Field Plotter, 1971
5. [illegible] of the Urban Environment, 1972
6. [illegible] Instruction in Undergraduate Geography, 1972
7. The [illegible] as a Working Environment A Purpose for Physical Geography, 1972
8. LAND USE A Computer Program for Laboratory Use in Economic Geography Courses, 1972
9. [illegible] of Statistical Applications in Geography [illegible]

Resource Paper Series

1. Theories of Urban Location, 1968
2. Air Pollution, 1968
3. Perspectives on Geomorphic Processes, [illegible]
4. Spatial Diffusion, 1969
5. Perception of Environment, [illegible]
6. Social Processes in the City [illegible] and [illegible], 1969
7. The Spatial Expression of [illegible]
8. The [illegible] Organization of [illegible]
9. An Introduction to Spatial [illegible]
10. Man and Nature, 1971
11. Tropospheric Waves, Jet Streams, and [illegible] Weather Patterns, 1971
12. The Spatial Structure of Administrative Systems [illegible]
13. Residential Mobility in the City, 1972
14. The Periglacial Environment, [illegible] 1972
15. Conservation, Equilibrium, and [illegible] Atmospheric and Fluvial Processes, [illegible]
16. Metropolitan Neighborhoods Participation and Conflict Over Change, 1972
17. Computer Cartography, 1972
18. Society, The City and The Space [illegible] of Urbanism, 1972

*Out of print

THEMATIC CARTOGRAPHY

Phillip Muehrcke

COMMISSION ON COLLEGE GEOGRAPHY
RESOURCE PAPER No. 19

ASSOCIATION OF AMERICAN GEOGRAPHERS
WASHINGTON, D.C.
1972

Supported by a grant from the National Science Foundation

ASSOCIATION OF AMERICAN GEOGRAPHERS

Commission on College Geography Publications

General Series

1. Geography in Undergraduate Liberal Education, 1965
2. A Basic Geographical Library: A Selected and Annotated Book List for American Colleges, 1966
3. Geographic Manpower: A Report on Manpower in American Geography, 1966
4. New Approaches in Introductory College Geography Courses, 1967
5. Introductory Geography: Viewpoints and Themes, 1967
6. Undergraduate Major Programs in American Geography, 1968
7. A Survey Course: The Energy and Mass Budget at the Surface of the Earth, 1968
8. A Systems Analytic Approach to Economic Geography, 1968
9. A Geographical Bibliography for American College Libraries, 1970
10. Geography in the Two-Year Colleges, 1970

Technical Paper Series

1. Field Training in Geography, 1968
2. Computer Assisted Instruction in Geography, 1969
3. Evaluating Geography Courses: A Model with Illustrative Applications, 1970
4. Living Maps of the Field Plotter, 1971
5. Simulation of the Urban Environment, 1972
6. Computerized Instruction in Undergraduate Geography, 1972
7. The Interface as a Working Environment: A Purpose for Physical Geography, 1972
8. LAND USE: A Computer Program for Laboratory Use in Economic Geography Courses, 1972
9. Bibliography of Statistical Applications in Geography, 1972

Resource Paper Series

1. Theories of Urban Location, 1968
2. Air Pollution, 1968
3. Perspectives on Geomorphic Processes, 1969
4. Spatial Diffusion, 1968
5. Perception of Environment, 1969
6. Social Processes in the City: Race and Urban Residential Choice, 1969
7. The Spatial Expression of Urban Growth, 1969
8. The Political Organization of Space, 1971
9. An Introduction to Spatial Allocation Analysis, 1971
10. Man and Nature, 1971
11. Tropospheric Waves, Jet Streams, and United States Weather Patterns, 1971
12. The Spatial Structure of Administrative Systems, 1972
13. Residential Mobility in the City, 1972
14. The Periglacial Environment, Permafrost, and Man, 1972
15. Conservation, Equilibrium, and Feedback Applied to Atmospheric and Fluvial Processes, 1972
16. Metropolitan Neighborhoods: Participation and Conflict Over Change, 1972
17. Computer Cartography, 1972
18. Society, the City, and the Space-Economy of Urbanism, 1972

Out of print.

THEMATIC CARTOGRAPHY

Phillip Muehrcke
University of Washington

ASSOCIATION OF AMERICAN GEOGRAPHERS

Commission on College Geography
Washington, D.C. 20009

RESOURCE PAPER NO. 19

Library of Congress Catalog Card Number 72-77214

Supported by a grant from the National Science Foundation

TABLE OF CONTENTS

LIST OF FIGURES

FOREWORD

The Resource Papers have been developed as expository documents for the use of both the student and instructor. They are experimental in that they are designed to supplement existing texts and to fill a gap between significant research in geography and readily accessible materials. The papers are concerned with important concepts in modern geography and focus on three general themes: geographic theory, policy implications, and contemporary social relevance. They are designed as supplements to a variety of undergraduate college geography courses at the introductory and advanced level. Another recent Resource Paper relating to the subject matter of this paper is No. 17 *Computer Cartography*.

These Resource Papers are developed, printed, and distributed by the Commission on College Geography under the auspices of the Association of American Geographers with National Science Foundation support. The ideas presented in these papers do not imply endorsement by the AAG. Single copies are mailed free of charge to all AAG members.

John F. Lounsbury
Project Director
Commission on College Geography

Paul W. English
Chairman, Panel on Resource and Technical Papers

Panel Members

Ian Burton, University of Toronto
Leslie J. King, McMaster University
Harold M. Rose, University of Wisconsin-Milwaukee
Robert H. T. Smith, Queen's University
Edward W. Soja, Northwestern University
Philip L. Wagner, Simon Fraser University
Julian Wolpert, University of Pennsylvania

THEMATIC CARTOGRAPHY

I. INTRODUCTION

Cartographers are continually under pressure to reduce mapping time and costs while still producing more and better maps. Satisfying these demands historically has meant that cartographers must quickly take advantage of general technological and informational advances in society. Major cartographic developments are, in fact, easily traced to such inventions as the compass, the printing press, the chronometer, lithography, aerial photography, scribing, computer graphics, and remote sensing from space vehicles. Obviously many of the developments alluded to here have been in primary (i.e., original) mapping activity and have often been accomplished through expensive equipment acquisition and a fragmentation/specialization of labor inputs. The scale economies of primary mapping and the nature of the demand for the resulting maps have actually encouraged centralization to the point that most activity is now carried out in a few large governmental agencies. Fortunately, all of this has been to the benefit of the so-called thematic cartographer, who generally operates on a limited budget and utilizes information from established data collection agencies to compile specialized maps on a limited production basis

Both the environmental sciences and data processing have developed extensively over the past few years By and large, thematic cartographers have kept up with these developments and continued to produce a useful product. Sometimes vast changes in data inputs and mapping methodologies were involved. Recent responses in thematic mapping to general technological and informational developments have actually produced major modifications in cartographic emphasis and technique The conceptual, methodological and technical developments underlying these changes have revealed or created a wide array of cartographical problems as well as attracted attention to the importance of basic cartographic research Unfortunately, developments have often been so diverse or of such specialized nature that their significance in a broader sense has not always been clearly recognized or understood Too often for the individual the result has been to obscure general research trends and goals The purpose of this monograph is to review recent developments in the field of thematic cartography and, further, to identify interesting problem areas which warrant future attention Although it is impossible to treat exhaustively the extensive material available in a single paper of this nature, the selection of topics presented should illustrate adequately the number of unresolved problems facing thematic cartographers and those who utilize their specialized maps. Keep in mind of course that this writer's priorities in topic selection and emphasis do not necessarily reflect the opinions one might solicit from a committee of noted cartographers.

The problems and research activities of concern here are best associated with the continuing trend from *map making* to *cartography* This distinction is important Map making has received most attention and is defined as the aggregate of those individual and largely technical processes of data collection, cartographic design and construction (drafting, scribing, display), reproduction, et cetera, normally associated with the actual production of maps A large body of technical information and intuitive procedures has accumulated over the years which has permitted map makers with, in many cases, little formal scientific training in the various aspects of cartographic representation to compile quite acceptable maps for many purposes Cartography, on the other hand, refers to the little studied philosophical and theoretical bases principles, and rules for maps and mapping procedures The current inability of thematic cartographers to synthesize effectively the amassed wealth of individual mapping techniques reflects their lack of attention to these conceptual foundations of mapping. The result too often is inarticulate practitioners of an intuitive discipline Cartographic activity explicitly introduces the graphic model nature of maps and the resulting implications for map making methodology Since this includes map use in addition to map making, cartography explicitly concerns the relative effectiveness of various mapping techniques and requires a clear understanding of the uses to which map information will be put In simplest terms, cartography is characterized by conceptual, problem-oriented research directed at formalizing the *science* underlying the *art* of cartography.

One may deduce from the introductory remarks that numerous *external influences* combine to set real constraints on the mapping process. Developments relating to these outside influences obviously will have a direct effect on mapping as well. It is interesting, therefore, to identify the more important influences that underlie the significant changes that have taken place in cartographic emphasis and technique in the past decade. It appears that cartographic trends during this period can be attributed to at least four broad, interrelated factors: a scientific revolution in map-related disciplines, the emergence of a general theory of communication, the advent of high-speed digital computers and peripheral devices, and improvements in drafting materials, equipment and techniques.

First, in recent decades there was a rapid "scientification" of disciplines most associated with the use of maps. The whole transformation process has been described as the *quantitative revolution*, in spite of the fact that the result was much more than mere quantification of disciplines. The "real" revolution was one of intense new interest in theoretical and methodological development The high degree of largely uncritical acceptance enjoyed by maps historically, due largely to the fundamental role they played in geography, came under long overdue scrutiny. Emphasis quickly changed from a qualitative to a quantitative approach with the assumption that subjective map analysis is not sufficiently rigorous to provide the foundation for a scientific discipline never far out of mind Maps had to take their proper place with other methods of information processing on which environmental policy and planning decisions depend. So-called *quantifiers'* major response was to borrow a host of statistical procedures from other fields for *objective* (i.e., repeatable) analysis of spatial distributions. Because of uncertainties surrounding the potential and role of both graphics and statistics, it is still not clear when the use of one or the other technique is more appropriate, nor what the most judicious mix of the two approaches should be. Probably the most important cartographic development to arise from this scientific revolution is an increased concern for the role of maps, or, more generally, graphics, in scientific argument. The literature of cartography, graphic communication, and scientific model building all become significant Cartographers have been forced to evaluate their methodology and adjust to a more sophisticated approach to research.

Secondly, in the last two decades *communication* and *information* captured the attention of scientists Ramifications from the communication and information theoretic approaches now permeate most scientific endeavor The development of a general theory of communication was bound to have important consequences for cartography, since maps are essentially information transmitting devices involving the basic rules of graphic communication. Although cartographers have not been able to couch their subject matter in formal communication theory terms, there has been an increased emphasis on the information aspects of cartography as evidenced by our jargonized literature (e.g., note the use of such words as *signal, noise, bit, channel capacity,* and others) The context of information theory provides clear-cut decision criteria for retaining or eliminating map elements on the basis of their communication effectiveness. But it is still necessary to determine what information is irrelevant, or only marginally relevant, to the map user and what combinations of symbol forms are confusing and distracting from overall map legibility. Probably the most imporant consequence of these new scientific perspectives on communication has been the explicit recognition that cartographic processing is, in fact, an active feedback system involving data collection, data processing, information display, and image processing. This "systems" approach, with its strong emphasis on the interdependence between the steps in cartographic information processing, is slowly replacing an historical tendency to consider cartography as essentially a "mapping" problem. The new approach has also caused a major rethinking of cartographic methodology and opened vast areas of potential research and development. It particularly emphasized the general nature of information and led cartographers to a consideration of more abstract forms of data.

Thirdly, recent decades have witnessed the coming of *computers* and *automation* The range of computer applications is essentially unlimited. Any problem that is *computable*, in the language of automata theory, can be solved theoretically by a general purpose computer. The condition of computability is defined by the so-called *Turing theorem*, i.e., any procedure which can be reduced to a finite sequence of explicit steps can be automated Certainly most cartographic applications can be reduced to a finite number of explicit operations and thus are potentially capable of being automated It is only a question of understanding cartographic practice, technological achievement, required speed, development and implementation costs, and so forth. It is hardly surprising, then, that the development of computers and peripheral equipment has had an immeasurable effect on contemporary cartography and the future outlook Cartographic technique has been radically transformed in two ways. On one hand, computers made feasible much more rapid and sophisticated statistical manipulation of data sets This, in turn, led to higher levels of scientific model building. In order to continue producing a useful product, cartographers have had to map these more abstract and often complex notions On the other hand, remote sensors, automated stereo-compilation machines, optical scanners, computer-directed line-printers and plotters, cathode ray tube (CRT) displays, and other devices have begun to replace manual map compilation techniques One result has been a phenomenal increase in compilation speed, accuracy, and the flexibility of data manipulation and display Although increased magnitudinal and positional (i e planimetric) accuracy is not especially critical and possibly should even be de-emphasized in thematic mapping, increased production speed, along with greater mapping flexibility, means that larger numbers of up-to-date maps can be produced to serve, particularly, urban and regional planning and policy decision needs A second consequence of automation derives from the fact that computer-assisted cartography must be based upon a logical scheme of unambiguous rules or instructions Cartographers have been forced to answer questions heretofore considered only implicitly by intuitive procedures Thus, while the automation of established cartographic practice was entirely predictable and therefore not of major theoretical consequence, the actual discipline involved in reducing the "art of cartography" to a sequence of explicit steps that form the individual computer programs could well prove to have a lasting beneficial effect on cartographic development The thought processes involved in making vague mapping ideas concrete cannot help but to facilitate the development and implementation of entirely new cartographic practices.

Finally, a wide range of drafting and reproduction media have been developed in recent years which require or encourage entirely new approaches to traditional problems Plastics, sensitized papers and films, and new nonprinting processes have had a particularly great impact on cartography. Preprinted materials, which require only map positioning, are now ubiquitous aids in cartographic laboratories. These materials permit the preparation of higher quality maps with less manual labor and in less time than was possible with previous methods.

Taken together, the recent changes in cartographic emphasis and technique have revitalized and reoriented a discipline founded upon convention In combination, they particularly emphasize that understanding and communication between data collector, cartographic processer, and user must be excellent if the cartographic information system is to function optimally To illustrate these interdependencies, the basic organization of cartographic procedures typifying cartographic information processing has been schematically outlined in Figure 1 The present survey of cartographic and related research closely follows the successive (but interrelated) steps in this system In particular, the cartographic process is viewed as a series of transformations involving the selection of data from the

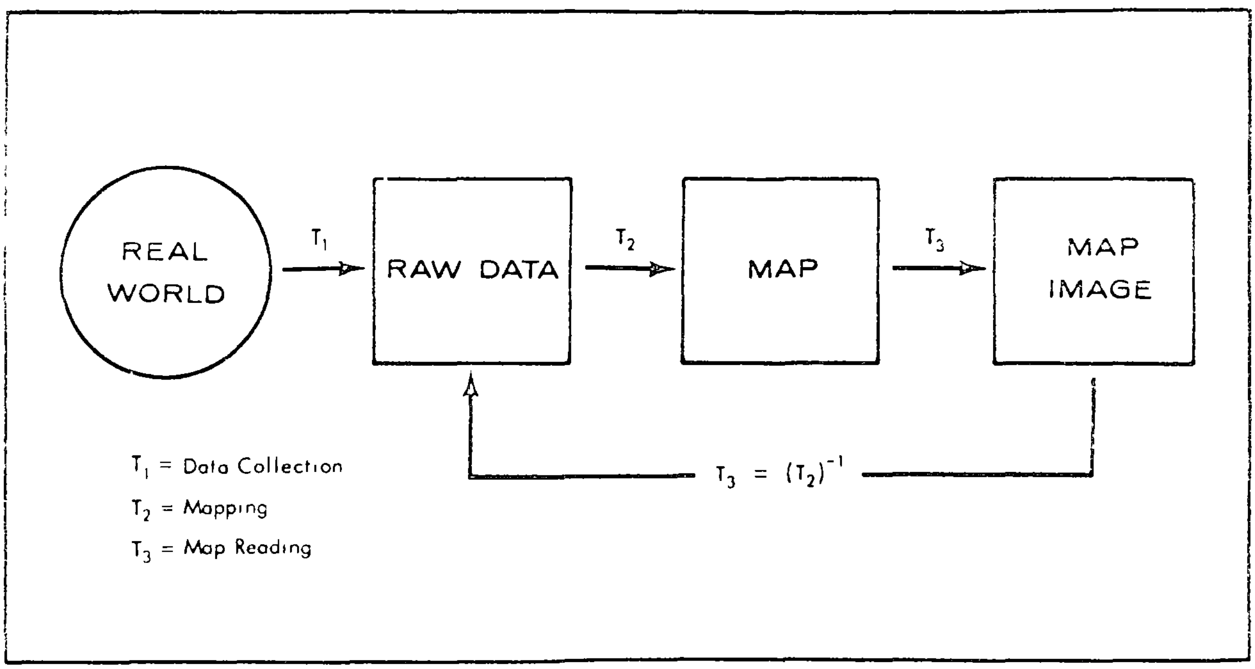


Figure 1 The cartographic processing system

real world (T_1), the transformation of these data into a graphic map (T_2), and the retrieval of information through an interpretative map reading process (T_3). A measure of the communication efficiency of the cartographic system is related to the amount of transmitted information, which is simply a measure of the correlation between input and output information. The cartographer's task is to devise better and better approximations to a transformation, T_2, such that output from T_3 is equal to input T_2, i.e., $T_3 = T_2^{-1}$. This conceptual simplification is not meant to imply, of course, that the transformations themselves are either simple or unrelated. Our ability to measure the amount of information that survives the various cartographic transformations is also quite limited.

II. DATA COLLECTION

Data capture actually is a better term than data collection, since it more expressively conveys the idea that data must be actively sought, as opposed to merely acquired in a passive way. The continuing information explosion in the environmental sciences long ago overwhelmed our limited data handling capabilities. Our ability to "capture" environmental data far exceeds our ability to interpret and manage it cartographically. The result is a growing demand for (a) lower cost maps produced in less time, (b) maps utilizing new forms and types of data, and (c) maps based on increasingly more sophisticated measurement and data sampling schemes. Satisfying these and other needs entails a wide spectrum of cartographic activity. More theory is needed, new techniques must be devised and, above all, more cartographers must be trained.

(1) Automation

Thematic mapping has been affected directly by the advent of versatile, high-speed digital computers which provide the basis for modern electronic data processing (EDP) The demand for numerical information spatially formatted with unambiguous locational identifiers brought about by the application of data processing techniques has caused major changes in established non-terrain data collection procedures ADP has provided the basis in thematic cartography for the development of automated spatial analysis and mapping systems, both of which are creating a need for ever larger amounts of spatially ordered digital information. It must also be remembered that for computer processing it is wise to record data at the largest scale that is likely ever to be desired, since theoretically there is no real internal storage problem while the task of augmenting existing data files with more detail is significant In practice, the unlimited nature of potentially mappable information and pragmatic limits on the physical size and costs of establishing and maintaining data banks make "recording data at the largest scale that is ever to be desired" only the ideal in theory We should therefore expect specialized, map specific data files to remain in common use, although these files will increasingly be in a form acceptable for EDP Whenever practical, future data collection schemes must be designed or modified to output forms of information which will be generally accessible with a minimum of subsequent data handling to the widest possible variety of cartographic processing techniques This means that observed measurements must be reported directly in machine usable forms Although today "machine-usable" essentially implies digital storage, we should expect future increased use of analog storage techniques such as microfilm, microfiche, and so forth, as pictorial computer input becomes a reality (see later discussion).

Continued modification of conventional data collection procedures along with the development of entirely new techniques will make ever larger volumes of data accessible to automated processing. The most efficient approach is to obtain *digitized pictures* or *numerical maps* directly from remote sensors, although, strictly speaking, this usually entails automatic conversion between some combination of graphical, electrical, optical and digital modes (Figure 2). Portable digitizing equipment which can be used to record observations directly as digits or convert them immediately into spatially ordered numerical forms will become essential to efficient field work. The manufacture of compact computers which are rugged enough to withstand installation in ships, airplanes, satellites, and other mobile vehicles means that input (data) may often enter the cartographic processing system directly from the environment with little, if any, human intervention. Computer interpolation and extrapolation algorithms can be used to "fill" data matrices from observations at scattered control points or sparse grid networks (Figure 3) In any case, the output from modern data acquisition systems will consist of coded information stored on punched cards, punched paper tape, magnetic tape, or in some other computer compatible form.

In cases where information has already been collected but exists only in a tabular, textual, or graphic format, these data have to be converted into a digital mode prior to most computer processing Tabular and textual information sources (e g., most existing census data) are easily but laboriously transferred directly to computer compatible cards or tapes using a typewriter-like device Numerical data are usually extracted manually from maps by superimposing one of the common areal sampling designs on the

DATA MATRIX (Digital map)

0.0	.8	1.6	2.1	2.6	2.7	2.7	2.4	1.9	1.3	.7	.0	-.5	-1.0	-1.3	-1.3	-1.2	-.9	-.5	-.0
1.6	2.4	3.1	3.7	4.1	4.3	4.2	3.9	3.4	2.8	2.1	1.4	.8	.3	.1	-.0	.1	.3	.8	1.2
2.8	3.6	4.3	4.9	5.3	5.4	5.3	5.0	4.5	3.8	3.1	2.4	1.8	1.3	1.0	.9	1.0	1.3	1.7	2.2
3.5	4.3	4.9	5.5	5.8	6.0	5.8	5.5	4.9	4.2	3.5	2.8	2.1	1.6	1.3	1.2	1.3	1.6	2.0	2.5
3.4	4.1	4.8	5.3	5.7	5.7	5.6	5.2	4.6	3.9	3.2	2.4	1.8	1.3	.9	.8	.9	1.2	1.6	2.1
2.6	3.3	4.0	4.5	4.9	4.8	4.7	4.3	3.7	2.9	2.2	1.4	.7	.2	-.1	-.2	-.1	.1	.6	1.1
1.3	2.1	2.7	3.2	3.4	3.5	3.3	2.8	2.2	1.5	.7	-.1	-.8	-1.3	-1.6	-1.7	-1.6	-1.3	-.9	-.3
-.0	.7	1.3	1.7	2.0	2.0	1.7	1.3	.6	-.1	-.9	-1.7	-2.4	-2.9	-3.3	-3.4	-3.2	-2.9	-2.5	-1.9
-1.2	-.5	.0	.5	.7	.6	.4	-.1	-.7	-1.5	-2.3	-3.1	-3.8	-4.3	-4.7	-4.7	-4.6	-4.3	-3.8	-3.2
-1.9	-1.3	-.7	-.3	-.1	-.2	-.4	-.9	-1.6	-2.4	-3.2	-4.0	-4.7	-5.2	-5.5	-5.6	-5.4	-5.1	-4.5	-3.9
-1.9	-1.3	-.8	-.4	-.2	-.3	-.6	-1.1	-1.8	-2.6	-3.4	-4.2	-4.8	-5.3	-5.6	-5.7	-5.5	-5.1	-4.6	-3.9
-1.3	-.7	-.2	.2	.3	.2	-.1	-.6	-1.3	-2.0	-2.9	-3.6	-4.3	-4.8	-5.1	-5.1	-4.9	-4.5	-3.9	-3.2
-.1	.5	1.0	1.3	1.4	1.3	1.0	.5	-.2	-1.0	-1.8	-2.6	-3.2	-3.7	-3.9	-3.9	-3.7	-3.3	-2.7	-2.0
1.3	1.8	2.3	2.6	2.7	2.6	2.3	1.8	1.1	.3	-.5	-1.2	-1.8	-2.3	-2.5	-2.5	-2.2	-1.8	-1.1	-.4
2.5	3.1	3.5	3.8	3.9	3.8	3.5	3.0	2.3	1.5	.8	.0	-.6	-1.0	-1.2	-1.1	-.9	-.4	.3	1.0
3.3	3.9	4.3	4.6	4.7	4.6	4.2	3.7	3.1	2.3	1.6	.9	.3	-.1	-.3	-.2	.1	.6	1.3	2.1
3.5	4.0	4.4	4.7	4.8	4.7	4.4	3.9	3.2	2.5	1.8	1.1	.5	.2	.0	.1	.5	1.0	1.7	2.5
2.9	3.4	3.8	4.1	4.2	4.1	3.8	3.3	2.7	2.0	1.3	.6	.1	-.3	-.4	-.2	.1	.7	1.4	2.2
1.6	2.1	2.5	2.8	2.9	2.9	2.6	2.1	1.5	.9	.2	-.5	-1.0	-1.3	-1.4	-1.2	-.8	-.3	.5	1.3
.1	.5	1.0	1.3	1.4	1.3	1.1	.6	.1	-.6	-1.3	-1.9	-2.3	-2.6	-2.7	-2.5	-2.1	-1.5	-.8	.0

NUMERICAL MAP
(function)

$$Z_{(X,Y)} = \mathrm{SIN}(X+Y) + 2\,\mathrm{SIN}(2Y) + 3\,\mathrm{SIN}(3X)$$

Figure 2. Digital and numerical maps (after Jones and Gallet, 1962 and 1962b)

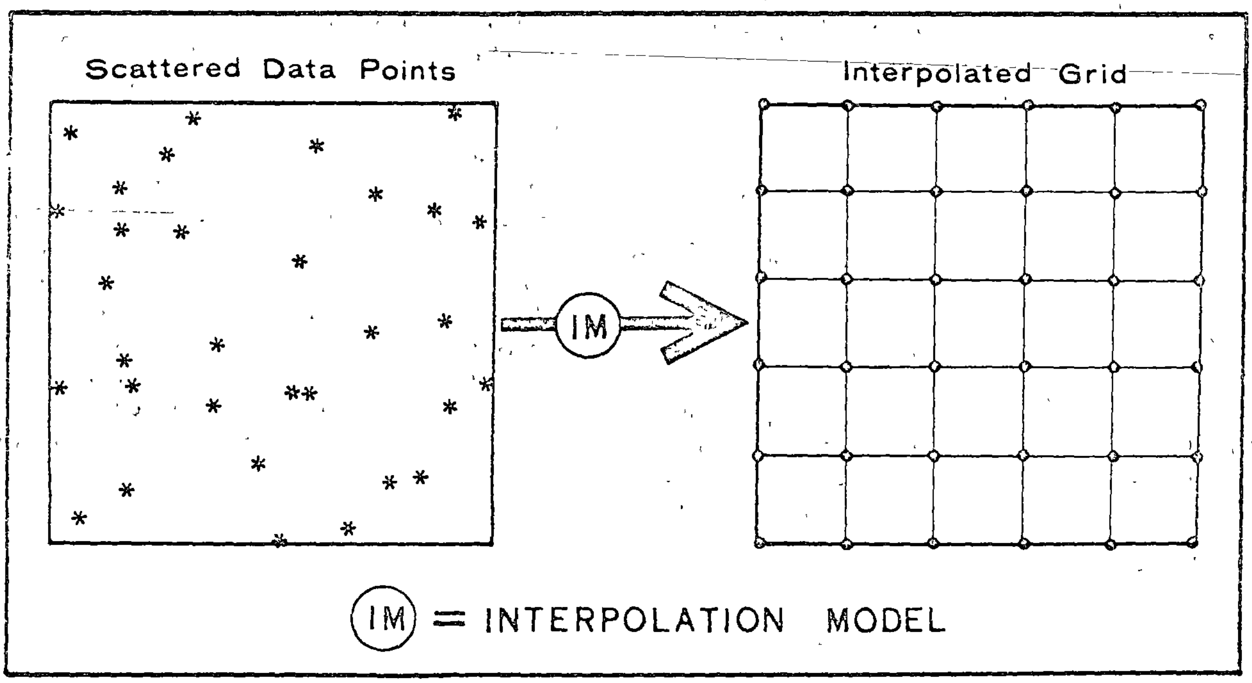


Figure 3 Predicting missing data using interpolation model

map surface, a topic to be discussed again later A large number of digitizing instruments and associated techniques have been developed for the task of converting geometric data from charts, maps, photographs, etc , into numerical form These instruments range in complexity (and cost) from relatively crude, manually operated tracing or pointing devices to partially automatic line tracing and fully automatic optical-mechanical scanning and digitizing machines Manually operated devices linked with coordinate digitizers and alphanumeric keyboards have already proved successful and should become standard cartographic laboratory equipment in this decade. The more expensive, volume-oriented automatic equipment still must undergo substantial improvement before it is generally practical Even then it will undoubtedly be restricted to only the largest mapping organizations for some time Research associated with the actual development of the types of mode changing equipment mentioned above is technical and not commonly pursued by thematic cartographers, therefore, the topic will not be considered further On the other hand, the degree and nature of data error (noise) introduced by system distortion is an important and relevant topic for study by thematic cartographers Any gain or loss of information must, of course, be weighed against system convenience and final map accuracy standards

Since development of computing systems that could handle graphical input would eliminate the numerical input bottleneck to increased automation, the literature of pictorial data-compression techniques, automatic pattern recognition and picture processing by computer is of interest (cf. Rosenfeld, 1969) Unfortunately, computer pattern analysis is still little more than crude processing In spite of its theoretical promise, most techniques developed to date in the picture processing area involve simple character recognition and apparently show limited direct cartographic applicability It will probably be some time before complex graphic data similar to maps can be processed automatically and thereby provide satisfactory solutions to the graphical input problem In the meantime, spatially ordered numerical data will remain a high priority item in data acquistion systems

(2) Type of Data

A trend toward mapping more conceptual, elusive (ephemeral) and abstract forms of information is apparent. Sometimes this has involved the development of revolutionary new methodologies, materials and equipment For example, in the case of using remotely sensed data, cartographers are increasingly looking beyond conventional black and white photography and the visible-light portion of the electromagnetic spectrum In fact, simultaneous sensing throughout the spectrum is already at least theoretically possible Researchers are currently involved with comparative analysis of such diverse multiband sensing techniques as black and white and true-color photography (in various film-filter combinations), false-color and black and white infrared photography, thermal infrared imagery, side-looking radar, electro-optical (television) imagery, imaging microwave imagery, and others Once the relative capabilities of these sensing methods taken individually and in combination are understood, the major remaining tasks will be (i) to learn how to match this impressive array of materials, equipment, and techniques to individual survey requirements, and (ii) to learn how to interpret and use the results (output) in the most efficient and effective way.

The growing acceptance of the concept of quantitative *geographical volumes* as described by *statistical surfaces* has given impetus to the mapping of many intangible distributions (Figure 4) The practically unlimited potential of topical mapping has been further exposed by a related trend toward general recognition of the fact that such abstract notions as *amount, value, density, significance, pressure*, etc , can be usefully substituted in place of the conventional *altitude* in studies of spatial variation (Figure 5) More widespread use of the statistical surface notion in conjunction with non-terrain phenomena in the future should result in much greater analytical sophistication, since, for purposes of analysis, inherently discrete data (e g , population) can be treated as continuous and differentiable throughout their domain As cartographers learn to utilize this powerful technique, we can expect to see an increased mapping of percentage, ratio, density, cost potential, demand, derivative, and other surfaces Since depiction of these non-terrain surfaces generally involves considerable data manipulation prior to mapping further discussion of this topic is deferred to a later section on data pre processing

Common practice is to map single characteristics which are considered diagnostic of their respective spatial phenomena. But it is not always possible to find a single diagnostic character for a phenomenon, and even if one is found it carries the risk of serious mismappings due to random aberrations inherent in the single feature used or errors in its measurement An alternative and often more useful technique is to map measures that summarize many characteristics, each of which is considered partially diagnostic of its respective spatial phenomenon Utilizing a number of common (or shared) features diminishes the "error" associated with any single attribute and increases the information content of the subsequent map symbols

More and more often cartographers will be confronted with distributional information which does not originate

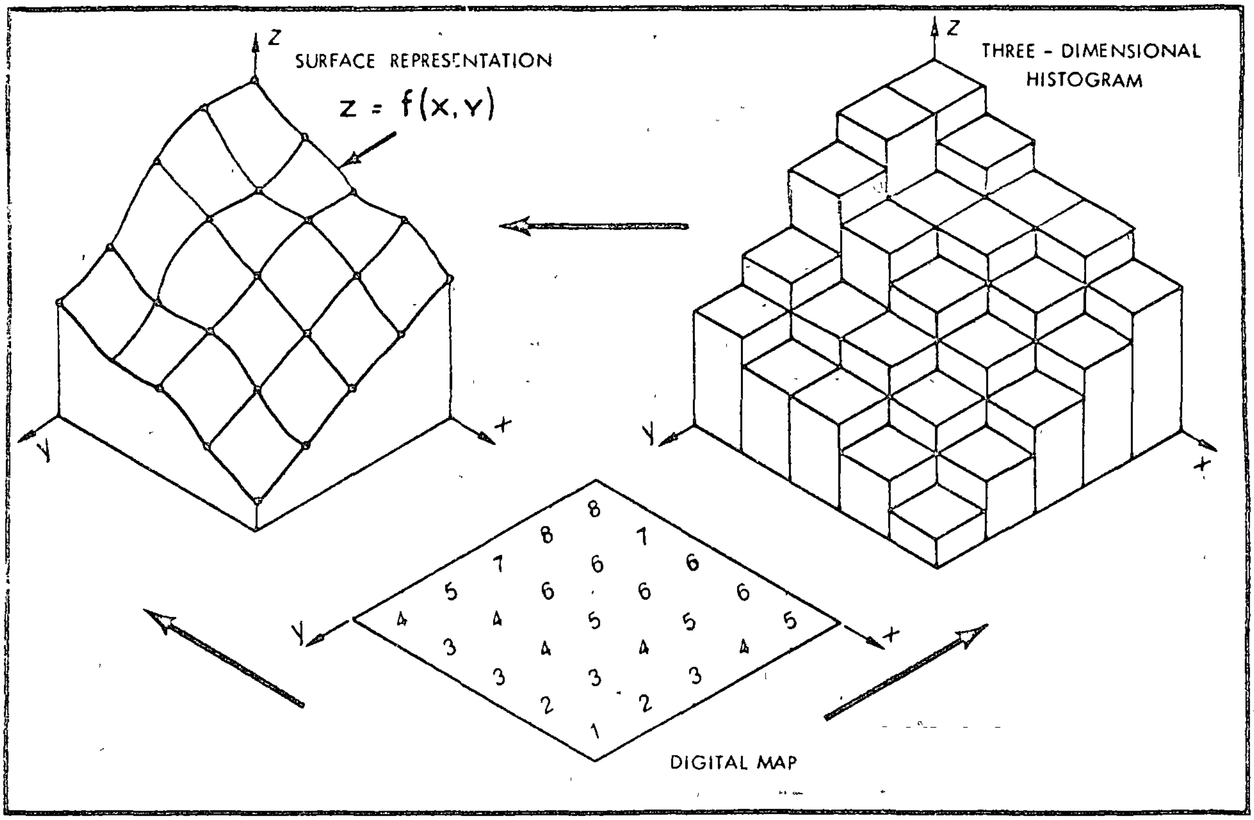


Figure 4 The statistical surface concept in mapping (after Robinson, 1961)

directly with existing phenomena. One example is the task of mapping spatial images held in the human mind in an attempt to delimit man's functional space (Figure 6). Mapping functional relationships deduced from a set of spatial process assumptions is another example. These data may represent (i) analog or digital computer simulations of abstract population and innovation diffusion models (Figure 7), or (ii) simply abstract functional descriptions Although the mapping task remains unchanged, we can expect certain conceptual and methodological problems to arise in the process of displaying these somewhat non-conventional data forms. The cartographer's responsibility will be to learn to cope with these less familiar map inputs as effectively as possible.

The temporal aspect of thematic map data is also undergoing change. Emphasis has always been on relatively static, easily managed information even though these data may not represent the most crucial environmental variables Maps of urban populations based on night-time residential distribution are good examples Potentially more interesting and revealing maps of daytime, rush hour, or non-residential night-time populations seldom exist! Color schemes of physical maps are often based on the natural landform and vegetation colors. But why do cartographers usually choose the wet season over the dry season, summer over winter, fall or spring? Why are monthly averages so often mapped even when they have no theoretical bases with respect to the phenomena depicted? Fortunately, some of these and other time conventions are breaking down and we can look forward to more temporal freedom in thematic mapping.

In the past the cartographer has commonly made do with information provided by fundamental data collection agencies such as census bureaus, government surveys, etc. Such information, however useful it seemed to the original data collection agencies, often was little suited to serve as the basis for environmental decision making. Now, as cartographic systems become formalized, cartographers are having much more chance to influence data collection More than ever before, the cartographer has the opportunity, and the obligation, to be responsive to the needs of those people responsible for the use of maps, rather than concern himself, as so often in the past, with presenting familiar or readily mappable phenomena. Cartographers

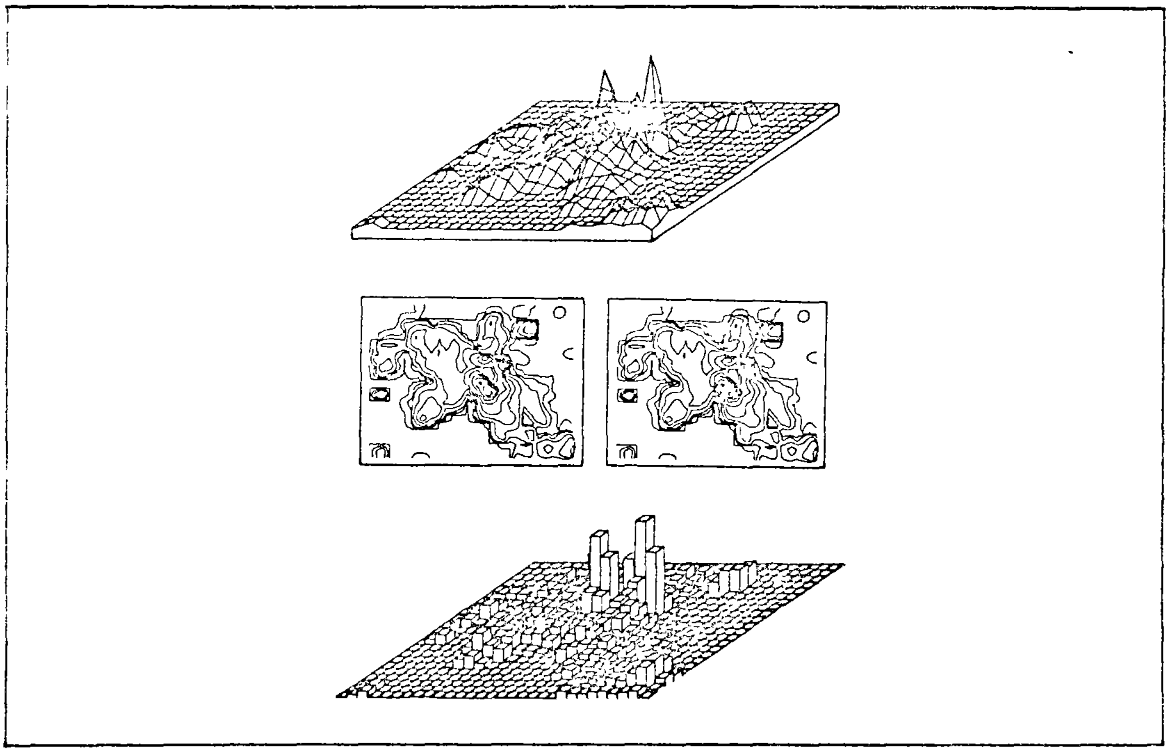


Figure 5 Population maps of Ann Arbor, Michigan (from Rens, 1967, Tobler, 1970)

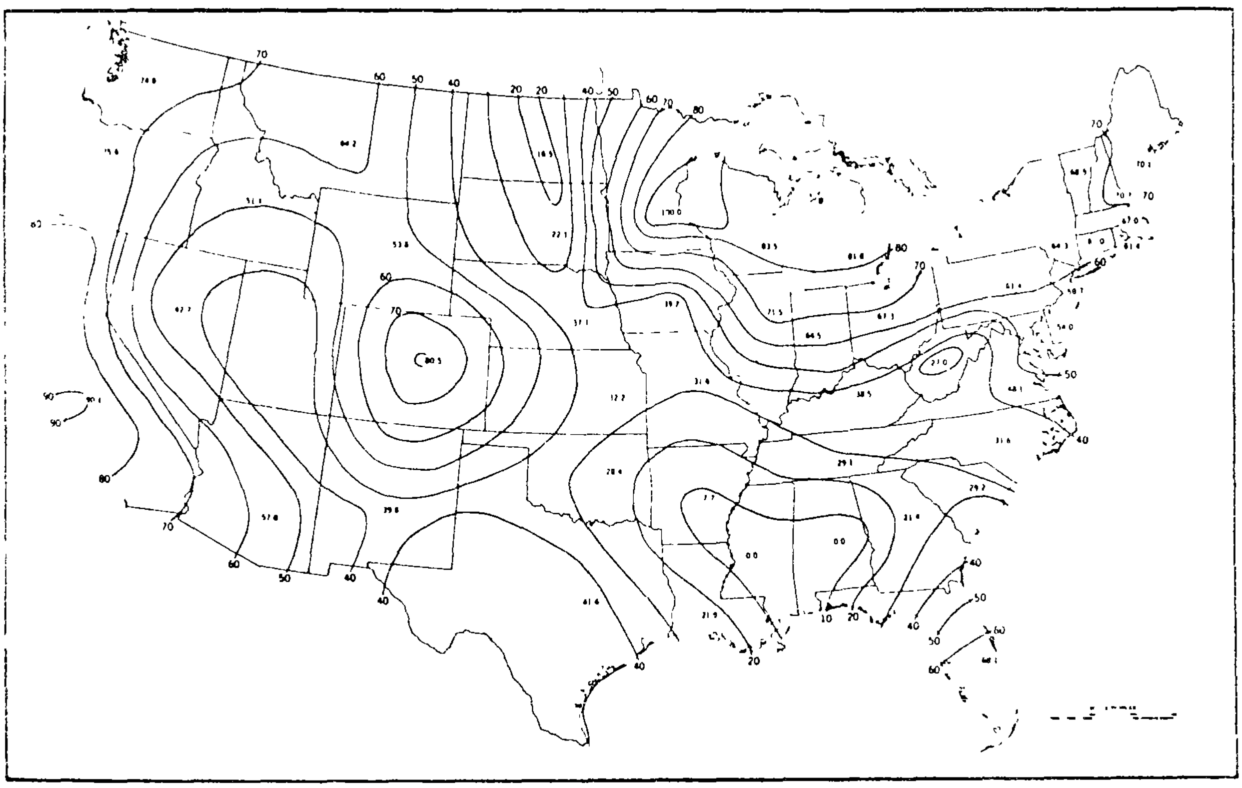


Figure 6 Overall view of the residential desirability of the U S as seen from Minnesota (from Gould, 1966)

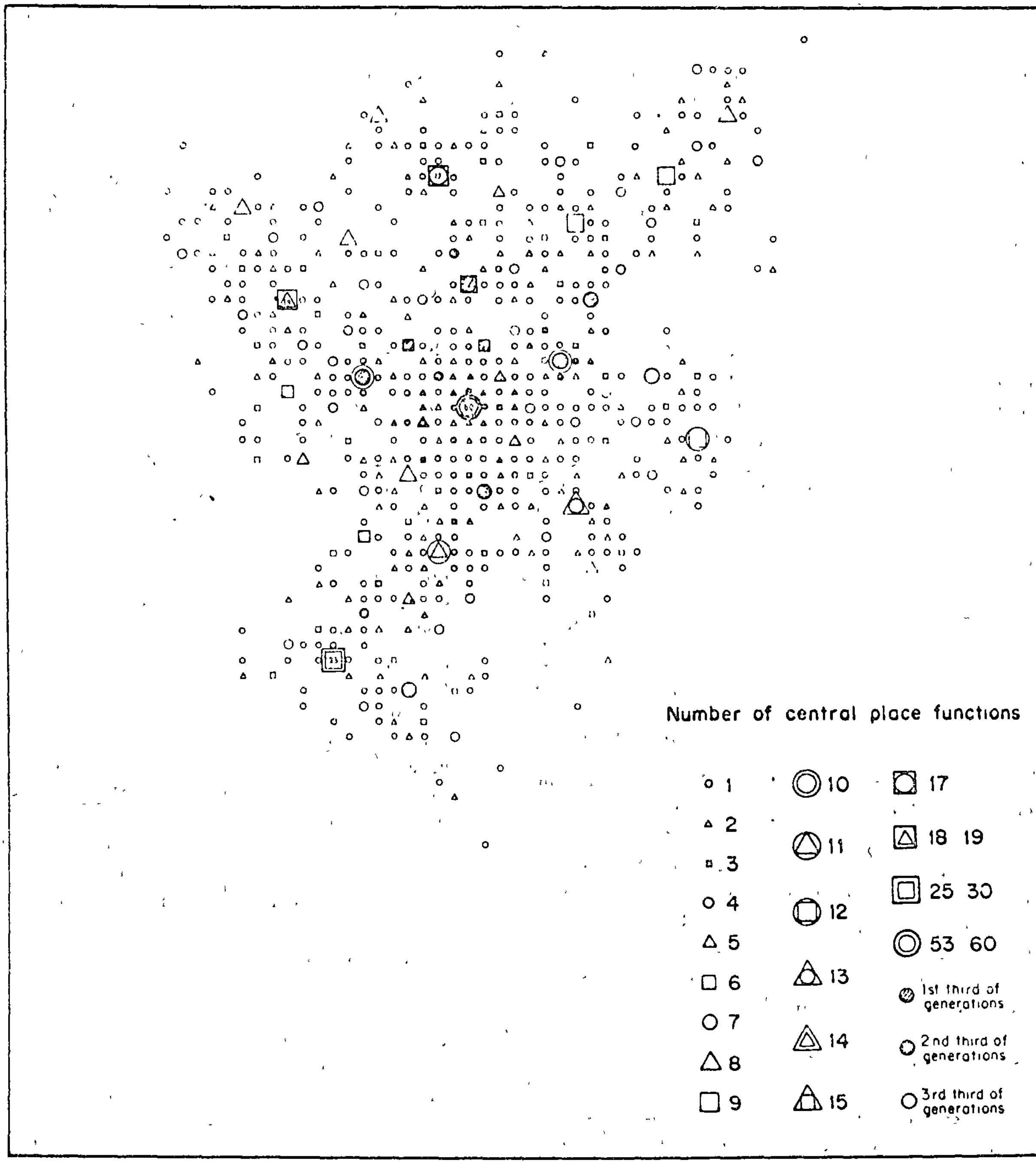


Figure 7 An experimental and hypothetical pattern of central places, as might have developed over perhaps 100 years (from Morrill, 1970)

must not fail to take advantage of their increasing ability to change the orientation from the collector of data to the user

In his interdisciplinary capacity, the cartographer must relate in some degree to the total environment which forms the basis of geographical mapping These data encompass geographical inequity (social/economic/physical), environmental quality, urban problems, population problems, marine resources, and so on, all the relevant issues of today!

In practice cartographers have failed Admittedly, they have tried to encourage an understanding of the relationships between elements within our environment, but the most "important" interactions were probably not always the ones considered. Cartographers have lacked essential initiative: they often took their subject matter passively from parent disciplines, such as geography, which were themselves essentially unconcerned or uncommitted. Thus, in spite of large amounts of money being spent each year in gathering information about the environment, we have few maps relating to relevant topics, and cartographers have contributed little to understanding conditions and relationships that affect people. We ask our administrators, planners and other decision-makers to work in a cartographic vacuum, an environment in which there are few useful maps on the basis of which to make rational decisions and predictions.

The following quote from an account of Mariner 9 prior to its instrumental contact with Mars is presented as one example of the limited vision of cartographers:

> "The Soviets and the U.S. are already fully aware of one dramatic Martian occurrence. Since the end of September, astronomers have observed a dust storm on the planet. Spreading at the rate of 20 or 30 m.p.h., the yellowish cloud now obscures much of the planet's surface and is one of the most severe blowups ever witnessed through terrestial telescopes. Some scientists are delighted with this rare chance of witnessing close up one of Mars' puzzling storms, which seem to occur when the planet moves closest to the sun and the Martian surface heats up. Others are equally concerned that the dust may obscure the view through Mariner's twin cameras. ... says Project Scientist Robert H. Steinbacher: '*The mappers just sick*,' while those people looking at dynamic changes on the planet are hoping the storm lasts long enough for them to study it." [Italics mine]
>
> (*Time*, November 15, 1971, p. 46)

The sad point here is that the cartographers were not prepared for mapping such dynamic and exciting phenomena as the storm pattern itself. Unfortunately, this example is symptomatic of a general deficiency which becomes even more tragic closer to home. The phenomena we map are not always the most crucial ones.

We owe the map user more. At the least we owe him access to relevant information and should modify the cartographic system accordingly. Major structural changes are not necessary, only a change in attitude to benefit the decision maker who is in a position where maps could be useful. In many cases all that is required is a change in emphasis from *physical* to *social*, from *past* to *present*, from *static* to *dynamic*, or from *interesting* to *relevant*.

(3) Scale

Perplexing geographical scale problems have been identified which cartographers must consider explicitly. Two components of geographical information have long been recognized as the site/situation dichotomy. The conflict that results in attempting simultaneous study of both relative position and site intensity is analogous to Heisenberg's *uncertainty principle*. Focus on either site or situation tends to diminish possible resolution of the other. Furthermore, the site/situation dichotomy is scale dependent, that is, it changes with the scale of inquiry.

Geographers involved in modeling spatial systems have found need to formalize particularly the "situation" concept in order to build operational process models. This has been partially accomplished through definition of *mean information fields*, which have also been labelled *geographical neighborhoods* or the *neighborhood of a point* (Figure 8). Understandably, the vast range of geographical phenomena exhibits wide variation in such neighborhood characteristics as form, extent, and intensity. Neighborhoods also overlap, are subject to distortion, and so on. All of this, of course, complicates the fact that any study of the processes that form the geographical environment must relate directly to the appropriate process neighborhoods.

It should now be apparent why data collection schemes have been likened to fishnets. Information is not captured unless its basic neighborhood or period is larger than the mesh size (i.e., sampling interval) of the net. If the data collection interval is small in relation to the relevant geographical neighborhood, an erratic and erroneous pattern is produced. If the basic data interval is large in relation to the relevant spatial neighborhood, the data are smoothed and lost. The important conclusion is that geographical data must be collected at a scale commensurate with the underlying spatial processes if the nature of the pattern is to be preserved in the data. A complication arises, however, if we accept the philosophical assertion that no practical map can possess true absolute scale, this is, represent a one-to-one mapping with the earth's surface. Proper areal aggregation is necessary and is based on several assumptions alluded to above: (i) spatial processes operate at various areal scales to produce geographical pattern complexes that exhibit a range from local variation to regional and global trends, and (ii) geographical sampling presents problems similar to those inherent in function sampling; thus, according to the *sampling theorem*, spatial patterns possessing periods (scales) less than twice the size of a regular data collection grid will be lost (Figure 9). When maps themselves form the base for sampling, it is necessary to coordinate both map scale and scale of individual map measurements with process scales in order to compensate for all of the scale-dependencies of such sampling regimes. Unfortunately, the task of specifying the scale at which a particular geographical pattern should be measured and mapped still poses a difficult cartographic problem, primarily because spatial

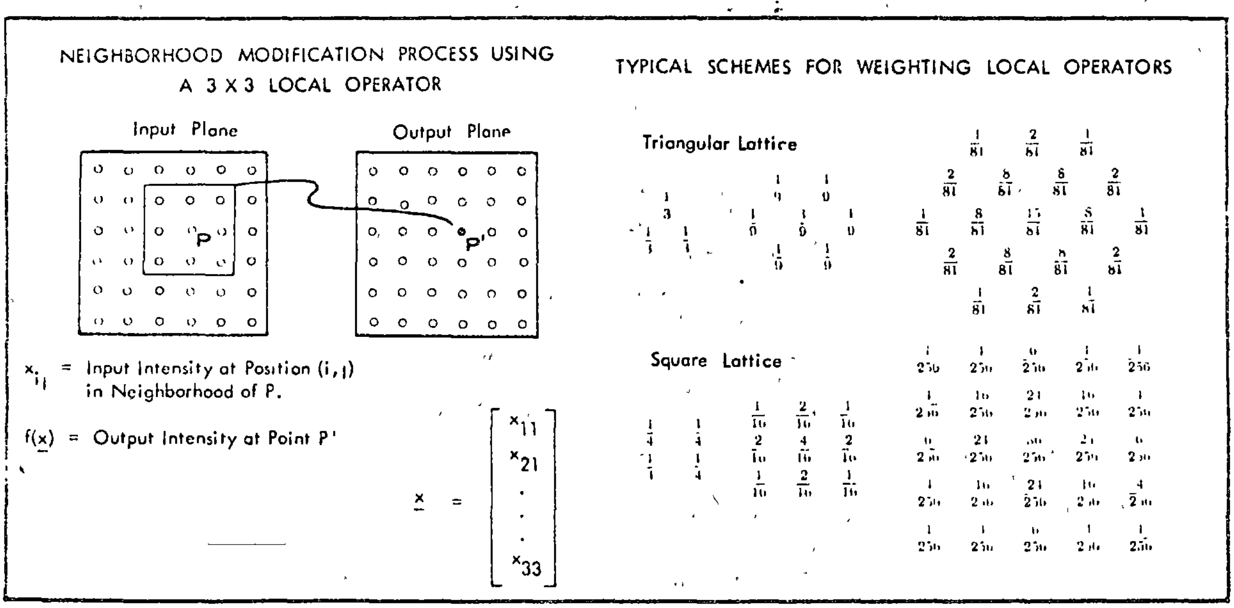


Figure 8 The geographical neighborhood concept with typical weighting operations

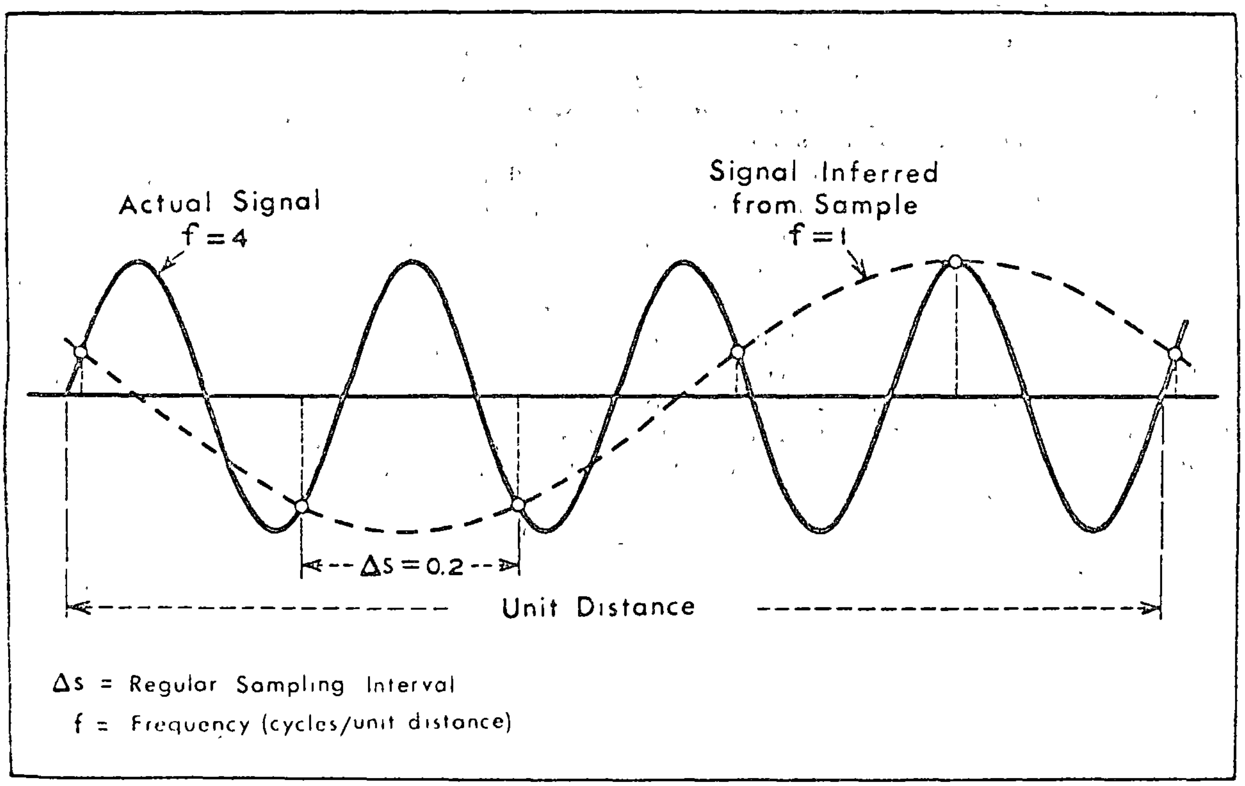


Figure 9 Theory of function sampling (after Blackman and Tukey, 1958)

analysts have yet to formally characterize the pattern scales associated with particular spatial processes.

Communication theorists would say that maps are *band limited*, that is, they focus on a particular spatial scale and blur information associated with other scales. The implication is that map scales must be conscientiously matched to user requirements. Unfortunately, map scales appear to have been too small for many "human" applications at the local level. Planners and various urban decision makers find little use for conventional maps. Scales of spatial processes important to people may well be filtered out of such standard and widely employed mapping series as the U. S. Geological Survey topographic sheets. The conclusion is that many commonly used maps are not really appropriate for the applications attempted. A poor transfer function exists between information presented at conventional map scales and map referent environments.

As an additional consideration, scale dependencies may develop in the data when geographical information is aggregated by area and the observational collection units exhibit different geometries. This is the case with most census data. On the basis of solutions to similar problems in other fields, Tobler (1969b) suggests that it should be possible to design techniques which would permit correction for these observational unit effects and, thereby, recover the true signal, a process labeled *geographical enhancement*.

Obviously the overall scale problem is complex, has important cartographic implications, and is not well understood at this time. We can make several observations. First, even if data were available "at the largest scale ever likely to be desired," the problem of matching level of cartographic abstraction to the relevant geographic neighborhood would still be with us. Establishing map specific data files simply introduces the neighborhood problem earlier in the information processing system. Secondly, given that geographical patterns are scale-dependent, we should suspect that the data input for a good map at one scale is not simply transformed into equally good data for maps at other scales. How to transfer from one scale to another is an intriguing question yet to be answered in a satisfactory manner. The scale question will be taken up again in a later section in relation to specific problems arising in cartographic processing of geographical information.

(4) Sampling and Prediction

Maps are increasingly compiled from spatial samples (population subsets) rather than from population data. Initial preoccupation with aspatial sampling designs borrowed from other disciplines has now largely given way to true spatial (plane) sampling (Figure 10). Points, quadrats, and traverses have been the standard sampling units employed in studies involving spatial distributions. Common procedures for selecting unbiased sample units include *random designs*, where sample elements are drawn independently with equal probabilities using random number tables, *systematic designs*, where sampling follows a consistent and orderly scheme throughout, *stratified designs*, where the sample is proportioned out to a stratification of area based upon existing knowledge of the population, and *hierarchical designs*, where sampling detail changes as a function of the level of population organization (see Berry and Baker, 1968). Sampling poses two basic question for the cartographer. First, what sampling attributes suffice for "adequate" descriptions of spatial patterns? Secondly, what effects do inferential generalizations (which go beyond the sample data) have on subsequent mapping and map interpretation operations, whether they be intuitive or statistical?

The main goal for cartographers in sampling and prediction is to generate information of uniform accuracy and abstraction so that subsequent mappings will be equally generalized throughout. In practice, variable map generalization (or "differential smoothing") becomes a major problem because (i) the spatial patterns sampled are by nature variable, and (ii) even in the densest plane sampling of a continuous distribution, z-values are obtained at only a relatively few points.

Sparse and uncontrolled sample values may suffice when a simple symbol plot (e.g., daily T.V. weather map showing the pattern of low temperatures for U.S. cities) is desired. But, usually, generalization from the sample data is called for and some form of interpolation and extrapolation is necessary to supplement the control with *predicted* z-values at other locations (refer to Figure 3). Since a variety of prediction (i.e., interpolation and extrapolation) methods are available, to be able to select the "best" model it is necessary to know, for each technique, the nature of prediction errors and its influence on derived map parameters (See Morrison, 1971). Evaluation and comparison of the accuracy and range of applicability of various possible methods of prediction thus require estimation of both prediction and propagation errors, each of which is related to the original z value function and the degree to which the sample represents $z = f(x,y)$. The structural nature of spatial distributions is appropriately studied using trend surface analysis and similar techniques. Actual sampling breaks down into problems relating to sample size, units of observation, and positioning design. From a practical standpoint it is generally advisable to collect geographical data at regular spatial intervals in two orthogonal directions in order to make best advantage of modern data processing capabilities.

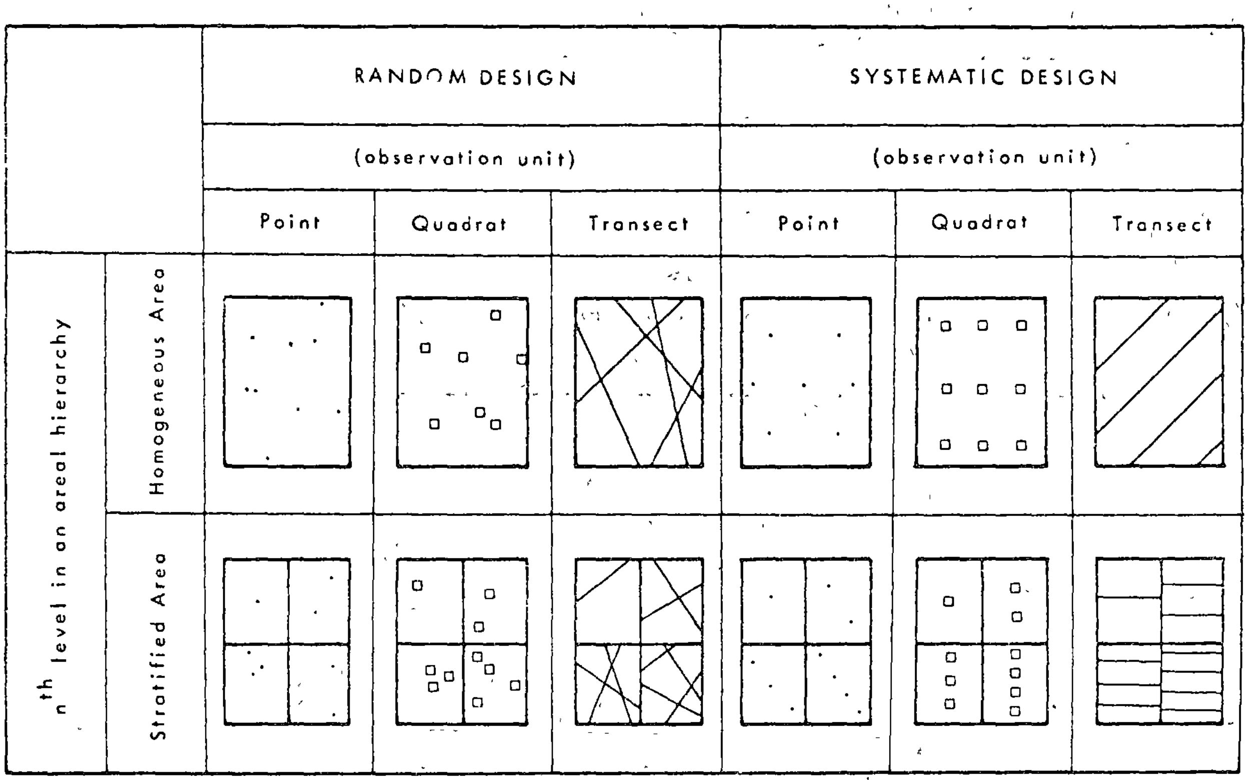


Figure 10 Spatial sampling schemes

Major problems apparently exist in balancing sometimes contradictory sampling criteria such as cost, operational convenience, and accuracy differences. The implications of using various prediction models with sample data drawn from populations exhibiting different structural characteristics clearly provide additional and valuable topics for future cartographic research We are faced with the problem of determining the accuracy of "estimated maps" based on various samples Recent work suggests the possibility of establishing general sampling/prediction model guidelines that will help guarantee the truest representation of the spatial characteristics of the phenomenon. Joel Morrison (1970) formally states the problem in terms of a simply calculated indicator (based on the nearest neighbor statistic) which specifies the degree of confidence one can have in the accuracy of maps produced by given sample procedures executed in various distributional contexts

Sampling also raises interesting questions not encountered when mapping population statistics Possibly the most intriguing question is whether confidence limits regarding the statistical reliability of a sample have a comparable interpretation in the case of subsequent map depictions. For example, what are the implications of convincing demonstrations that visual fidelity is not equivalent to statistical fidelity in mapping? This phenomenon relates to the psychophysics of human vision and warrants a great deal of attention in future research

(5) Measurement

Cartographers are discovering that the mapping operations permitted on given geographical information are dependent on the *level of measurement* achieved Figure 11 is designed to review and summarize what is meant by measurement in cartography and should be studied thoroughly by those unfamiliar with the concept before proceeding with this discussion More fundamental treatment of this important subject can be gotten from Siegel (1956).

Although it is easy to demonstrate that the radically different properties of the *nominal, ordinal, interval*, and *ratio scales* of measurement are of fundamental importance in mapping, their full implications in cartography have yet to be systematically and rigorously defined It is evident, however, that historical improvement in data collection schemes has also progressively raised the levels of measurement attained, and higher levels of measurement permit increased mapping accuracy and complexity The important

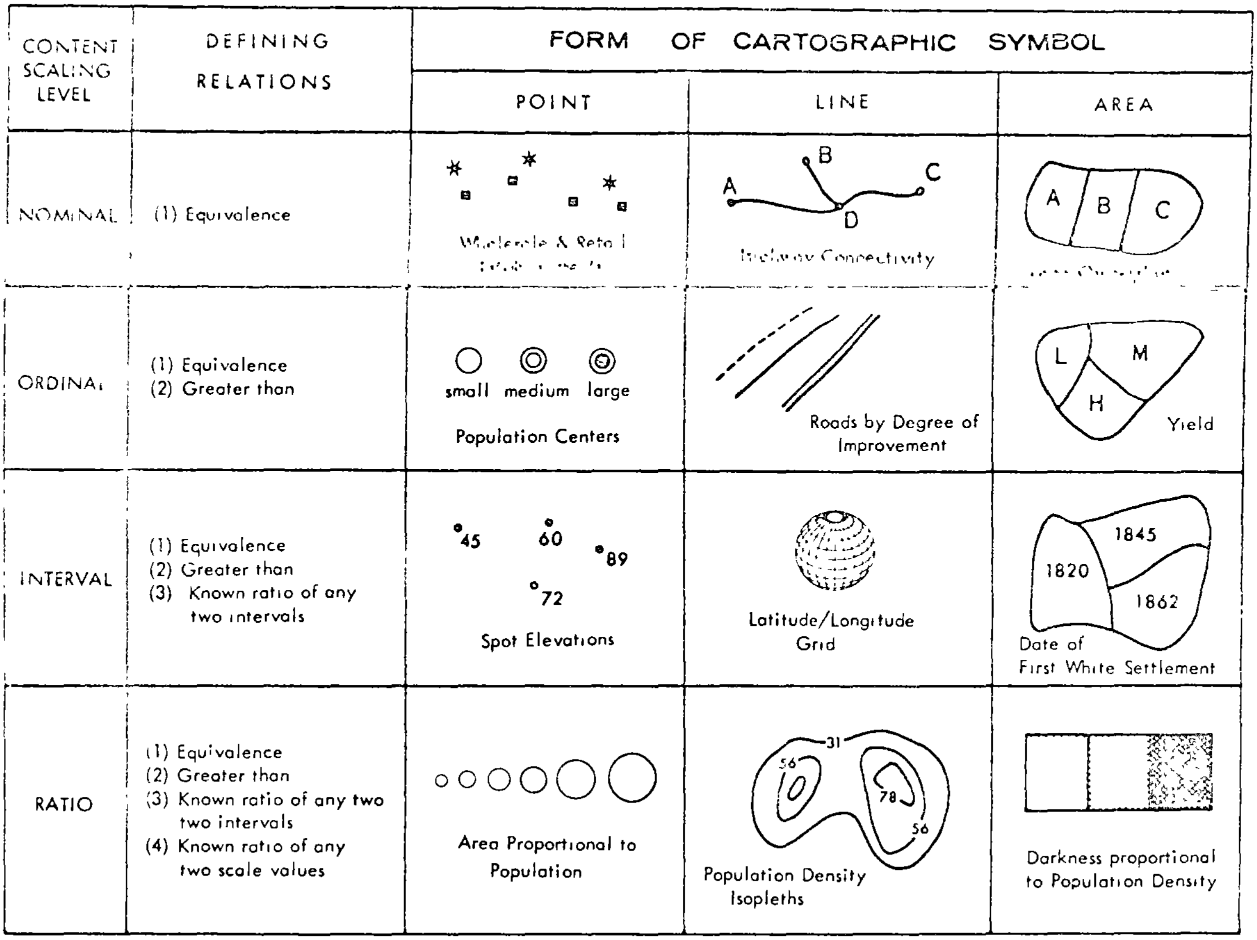


Figure 1.1 Measurement levels underlying map information

point is that all statistical and cartographical operations are permissible at the highest level (i.e., ratio scaling), but at the lowest level of nominal scaling only a very limited range of carto-statistical operations is permissible. Mapping "below" the achieved level of measurement is possible but generally inefficient (except for generalizations) since it involves a *loss* in information, whereas mapping "above" the measurement level is extremely difficult (without making categorizing assumptions which are at best dubious) if not impossible since it implies a *gain* in information. This is equivalent to saying that decreasing levels of measurements entail increasing map generalization and we still are unable to *ungeneralize* a map except in very special circumstances.

It should become quickly apparent that cartographers must exercise special caution when dealing with measurement concepts, since data plotted on maps have an implicit spatial component which permits the map reader to derive information beyond simple content which may be all that was explicitly scaled. For example, an outline map of the United States might at first appear to be strictly classificatory (i.e., nominal), but note that it is at least theoretically possible to compute both the length of state boundaries (i.e., ratio) and state position (i.e., interval). Roads between cities on a map may depict nothing but connectivity (i.e., nominal), yet it would be theoretically possible to derive distance (i.e., ratio) and position (i.e., interval) relationships. Additional possibilities obviously exist. Clearly, scaling on content and geographical ordering are quite separate things, but both represent measurement operations. Maybe the point here is that geographical information is properly ordered when in map form, which means that it possesses the spatial attribute position, from which can be derived direction and distance to related objects.

These few comments on scaling should suffice to convince the reader that cartographers must begin establishing serious formal guidelines to aid in optimizing the playoff between survey restrictions (i.e., time, labor, money), level of measurement, intended short-term map purpose, and potential long-term demands on the data

Thorough understanding of the question of measurement will introduce cartographers to methods which permit extraction of the maximum amount of useful information from each measurement scale. Growing importance of premap data manipulation (next section) suggests that knowledge of mathematical operations permissible on data recorded at each measurement level will also become increasingly more critical to effective mapping.

(6) Preprocessing

Cartographers face severe problems of sifting and filtering huge masses of data. Some form of data reduction is necessary to keep cartographic processing within reasonable limits. Numerous premap data processing operations which permit image improvement (enhancement), noise and redundancy reduction, extraction of salient features, and spatial frequency filtering have been developed in effective response to this data management problem Data transformations in wide variation are now routinely performed either to obtain desired end-products for direct mapping or to facilitate further cartographic processing Preprocessing refers here to any premap transformation of raw data beyond simple mode changes These transformations serve several specific purposes which, though seldom mutually exclusive, are conveniently separated for discussion. One class of transformations is illustrated effectively from the viewpoint of the communication engineer. In the context of map purpose, any set of data may be divided into a useful segment called the *signal* and spurious material called *noise* (Figure 12) Removing various types of noise from the data prior to mapping can often greatly improve the final map result For instance, Blumenstock (1953) suggests that data reliability can be affected by observational, sampling, and bias (or persistent) error resulting from various limitations in data collection. Numerous premap operations are designed to reduce or eliminate noise attributed to this original data error. Use of the principles of stereophotogrammetry to derive reliable planimetry from measurements taken on images distorted in photography is one well known example of premap correction for system error.

Spatial *smoothing* or *filtering* is possibly a more "cartographic" example of premap correction for data error. In this restricted interpretation, the goal of filtering is to isolate and remove that part of the data which can be attributed to the above error types This usually involves filtering high frequency (small scale) components in the data which are assumed to be associated with various random disturbances and are therefore eliminated. Of the numerous techniques for removing spurious short-wavelength components, computing local spatial moving averages and removal of random, high-frequency residuals through some form of *trend analysis* seem to be the most popular (Figure 13).

It is also possible to eliminate information considered superfluous, and therefore distracting, to particular mapping problems. This can be accomplished in several ways. For example, it is possible to enhance (a) the study of major spatial trends by first removing noise due to small scale fluctuations (low pass filtering), (b) the study of small scale features by first removing noise due to major trends (high pass filtering), and (c) the study of intermediate scale features by first removing noise due to higher and lower frequency variation (band pass filtering). The effects of dominant spatial features can be excluded so that the poorly developed features (of any frequency) that they obscure or conceal can be identified and studied more effectively. Reverse smoothers or amplifiers may accomplish the same goal.

Currently the most popular method of selective spatial filtering is geographical trend analysis. This quantitative technique of fitting mathematical surfaces to empirical data permits components of global, regional, and local variation to be identified formally and separated (Figure 14) We can conclude from Tobler's (1969b) review in "Geographical Filters and Their Inverses" that the actual choice between power series polynomial models, double Fourier models, nonlinear "substantive" models, and two-dimensional power spectrum models is critical, extremely difficult, and largely dependent on the nature of the data and the objectives of map analysis. The full effects of spatial smoothing or filtering must be studied until fully understood so that the procedures can be designed (or chosen) to best suit the application. Initially, research on filtering could be directed profitably at identifying meaningful picture "neighborhoods," defining useful local operators, studying the inward propagation of boundary effects, and increasing the interpretability of the results.

A second category of transformations includes the bulk of premap operations intended to accentuate select or derived attributes and relationships in the original data. The range of transgenerations available at many computer centers with the widely used Biomedical Computer Programs and the numerous transformations provided for in the more recently created Statistical Package for the Social Sciences are examples of simple non-spatial transformations. A number of transformations which actually take the spatial variable into account were performed by Tobler (1969a) in an attempt to isolate certain terrain characteristics which could then be correlated with an existing classification of land types. Relevant aspects of Tobler's experiment include local dodging, raw and locally averaged first and second spatial derivatives, local range, and local

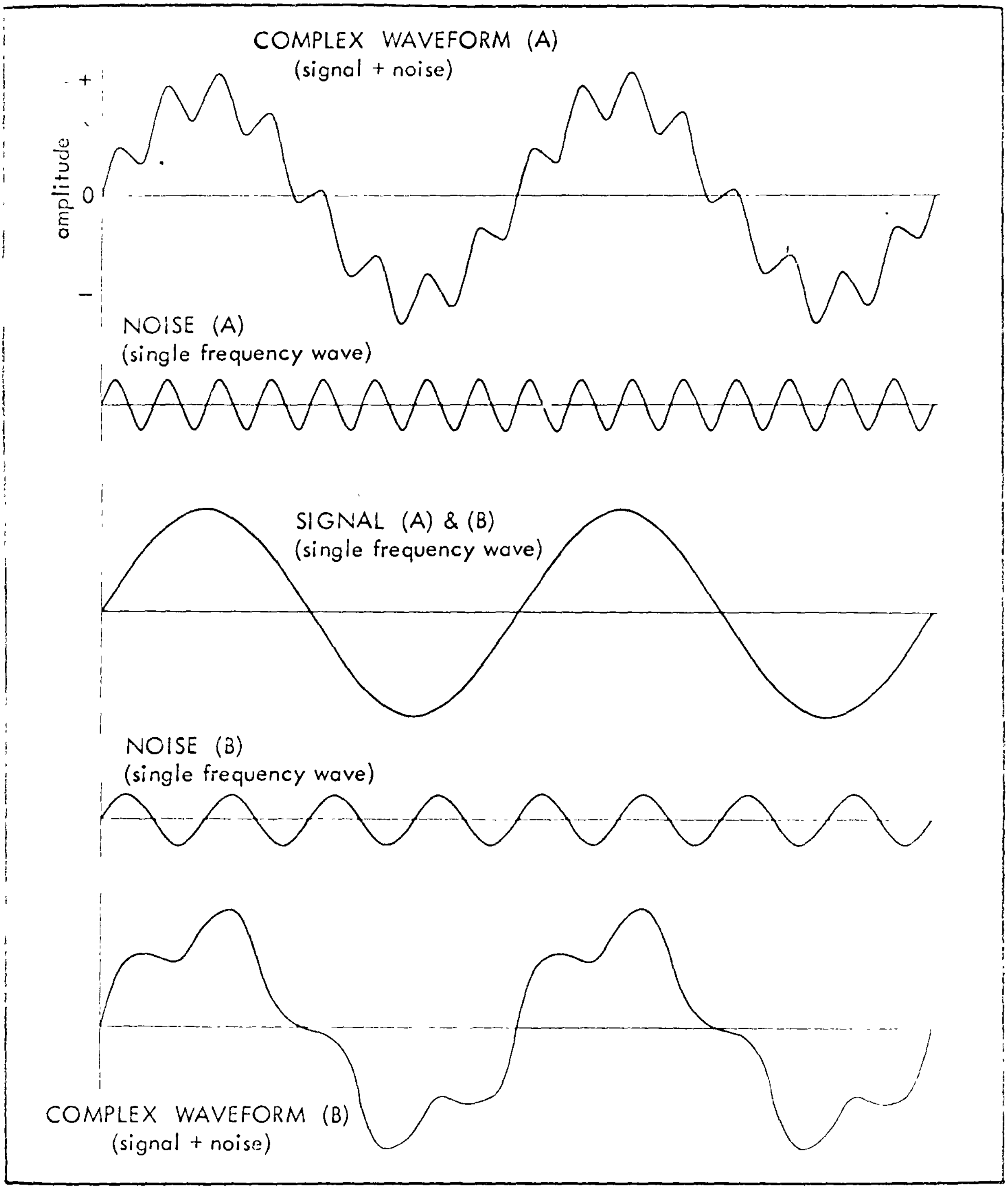


Figure 12 Decomposing data into signal and noise components an example of two functions differing only in fine detail

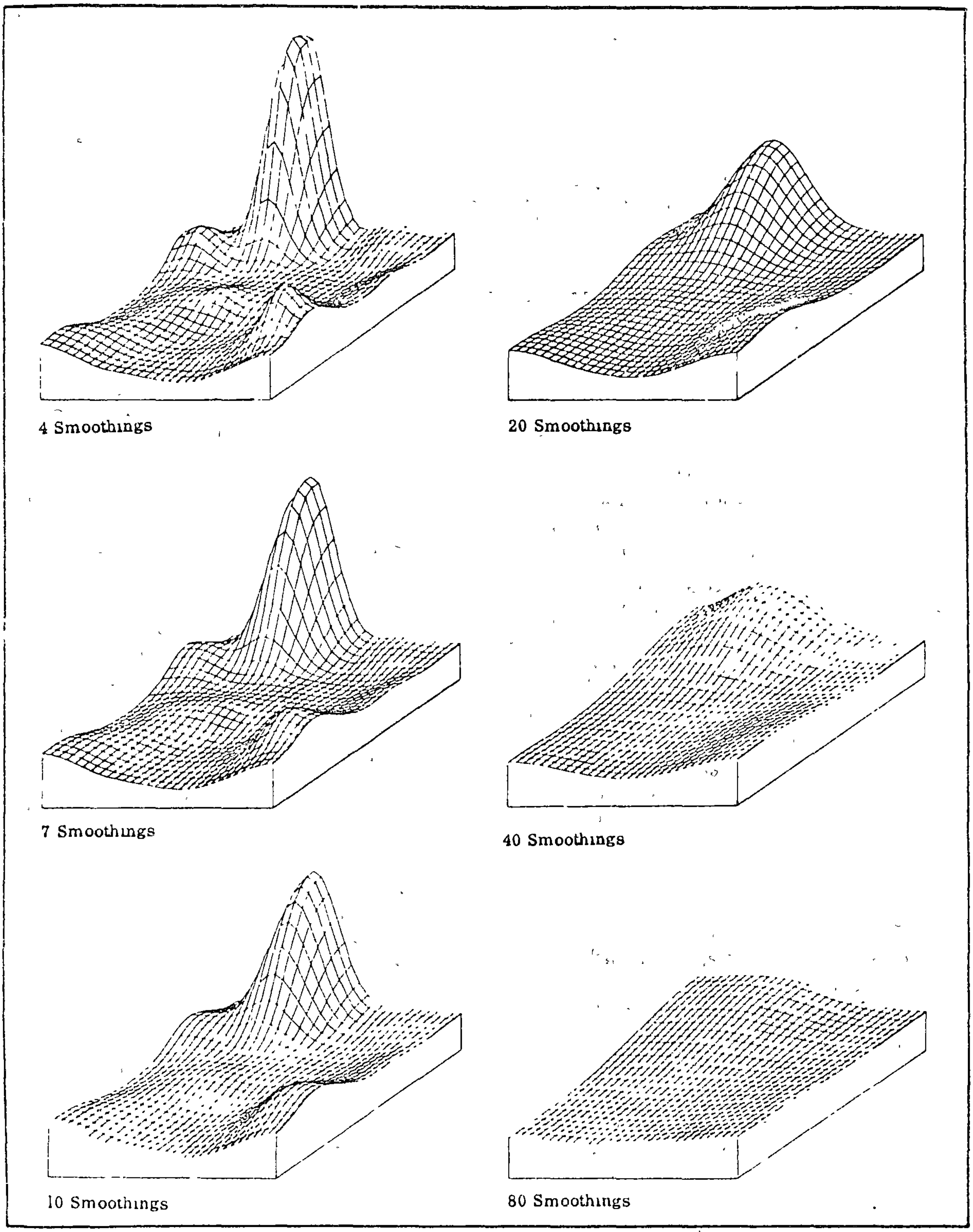


Figure 13 Progressive smoothings of the population density of an eight county region of northern Kansas (from Warntz, 1971)

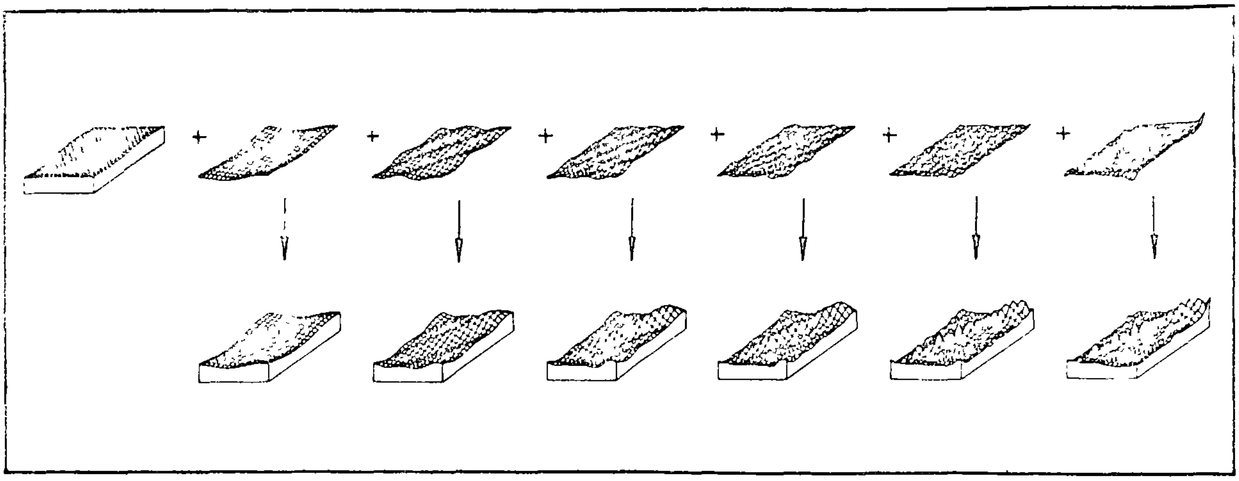


Figure 14 Geographical trend analysis Fourier synthesis of a portion of a mountain range in Ghana (from Tobler, 1969b)

variance (Figure 15) Additional terrain roughness parameters have been calculated as well. Computation of many indices, averages, densities, percentages, and coefficients, along with the generation of autocorrelation surfaces, potential surfaces, and probability surfaces, are further examples of premap operations falling in this class (Figure 16).

A third class of preprocessing techniques is designed to condense multivariate data into either a simple coding or a few "essential" features which account for most of the total variance Forgotson's (1960) review of classifying functions, entropy functions, entropy-ratio functions, facies departures, and variability functions illustrates coding techniques used with multivariate geological data, which permit single variable mappings (Figure 17) Multivariate classifying techniques such as principal components analysis, factor analysis, cluster analysis, and discriminant functions provide optimal variable combinations on the basis of a reduction in total sum or squares criterion (Figure 18)

A fourth class of procedures is intended to produce forms of information that are more suited for subsequent analysis Examples of these premap operations include basic (i.e., statistical) standardization and normalization procedures and transformation of spatial topology (Figure 19)

Obviously the number of possible premap transformations is restricted only by our ability to conceive them, and to date our imagination has not been impressive Some transformations, of course, will naturally be more meaningful than others from one situation to the next Combination of these transformations in "sequence" or in "parallel" also seems desirable (sometimes even necessary) which raises the question of their combinatorial algebra Tobler (1969a) has briefly explored this aspect of premap processing The second derivative, locally averaged first and second derivatives, and low low pass filter (all used in his study) offer fairly complex examples of meaningful concatenations Surface fitting analysis of the results of most of the transformations mentioned above is also permissible and sometimes revealing The extent to which mode changes and true transformations may be combined is well demonstrated by the series (pictorial data–digitizing -mean correction–autocovariance function–Fourier transforms–averaged Fourier transforms–logarithmic transformation–linear transformation) used by Agterberg (1967) in two-dimensional power spectrum analysis In general, the restrictions on permissible combinations of transformations are identical with those on single transformations–the results must be interpretable and meaningful

The logic underlying the astonishing growth in mathematical/statistical preprocessing has much support For one reason, increasing demand for spatial data, coupled with major developments in data acquisition capabilities, have already overwhelmed the cartographic processing system. The obvious efficiency gained through data compaction in premap processing will become even more essential if predicted increases in data inputs to the cartographic system from EROS (Earth Resources Observation System) and similar schemes materialize Premap data manipulation is therefore considered to be of great importance as a partial solution to our information overload problems Secondly, the ability of humans to analyze effectively spatial distributions is alleged to deteriorate progressively as the number of variables increases, interrelationships among variables become more subtle and the magnitude of variations decreases This suggests that cartographic presentation must demand as little mental com-

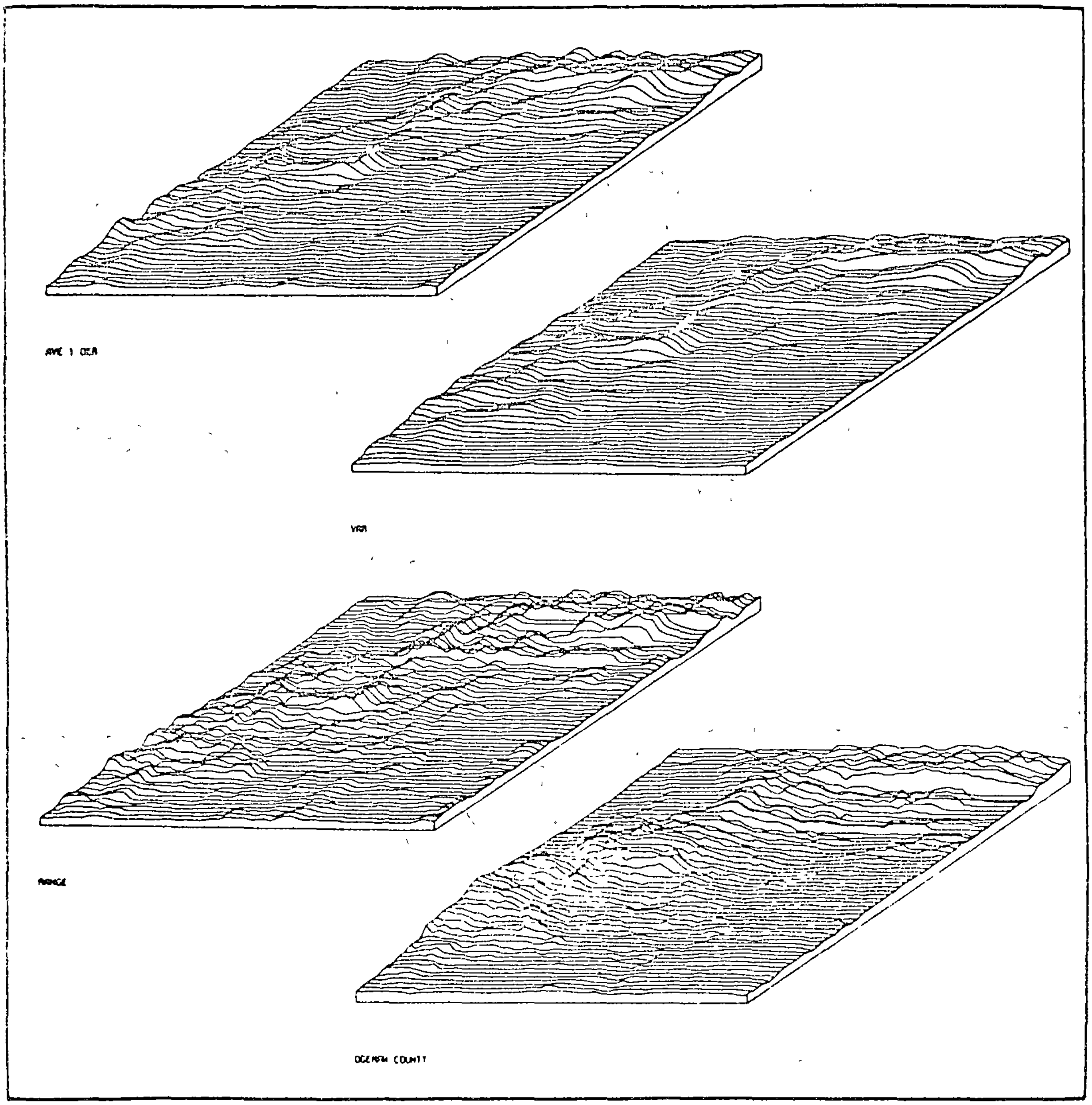


Figure 15 Modifying a terrain surface using local geographical operators (from Tobler, 1969a)

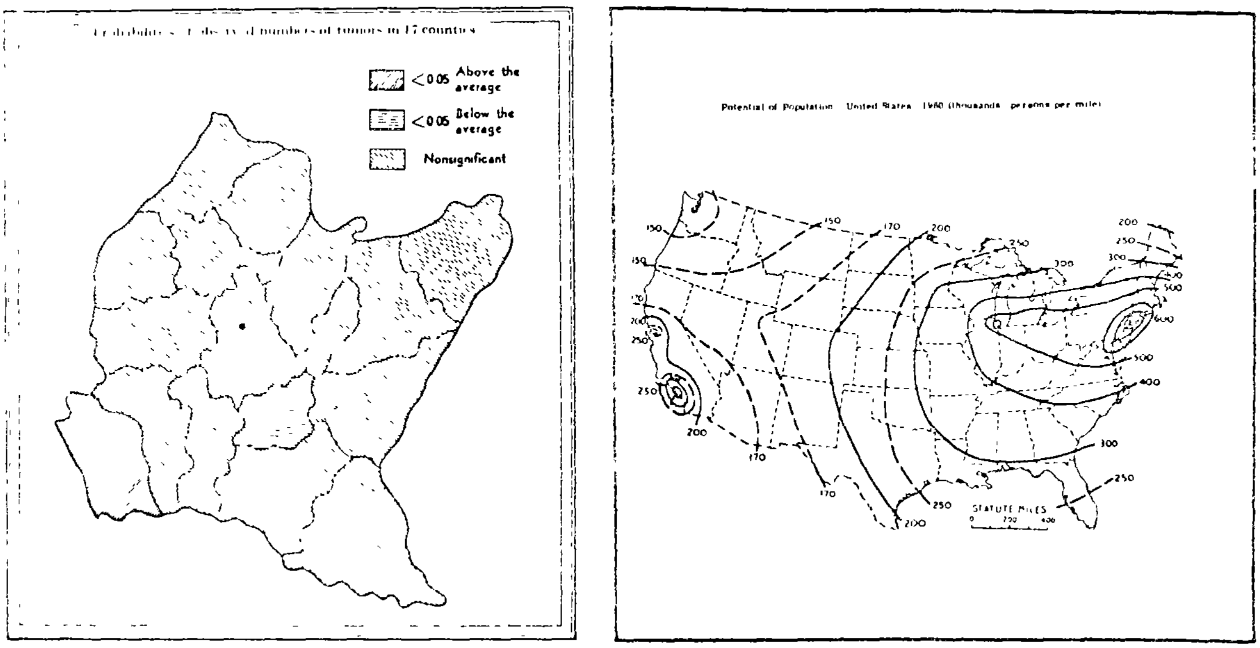


Figure 16. Overall data transformations: probability map (from Choynowski, 1968) potential map (from Neft 1966)

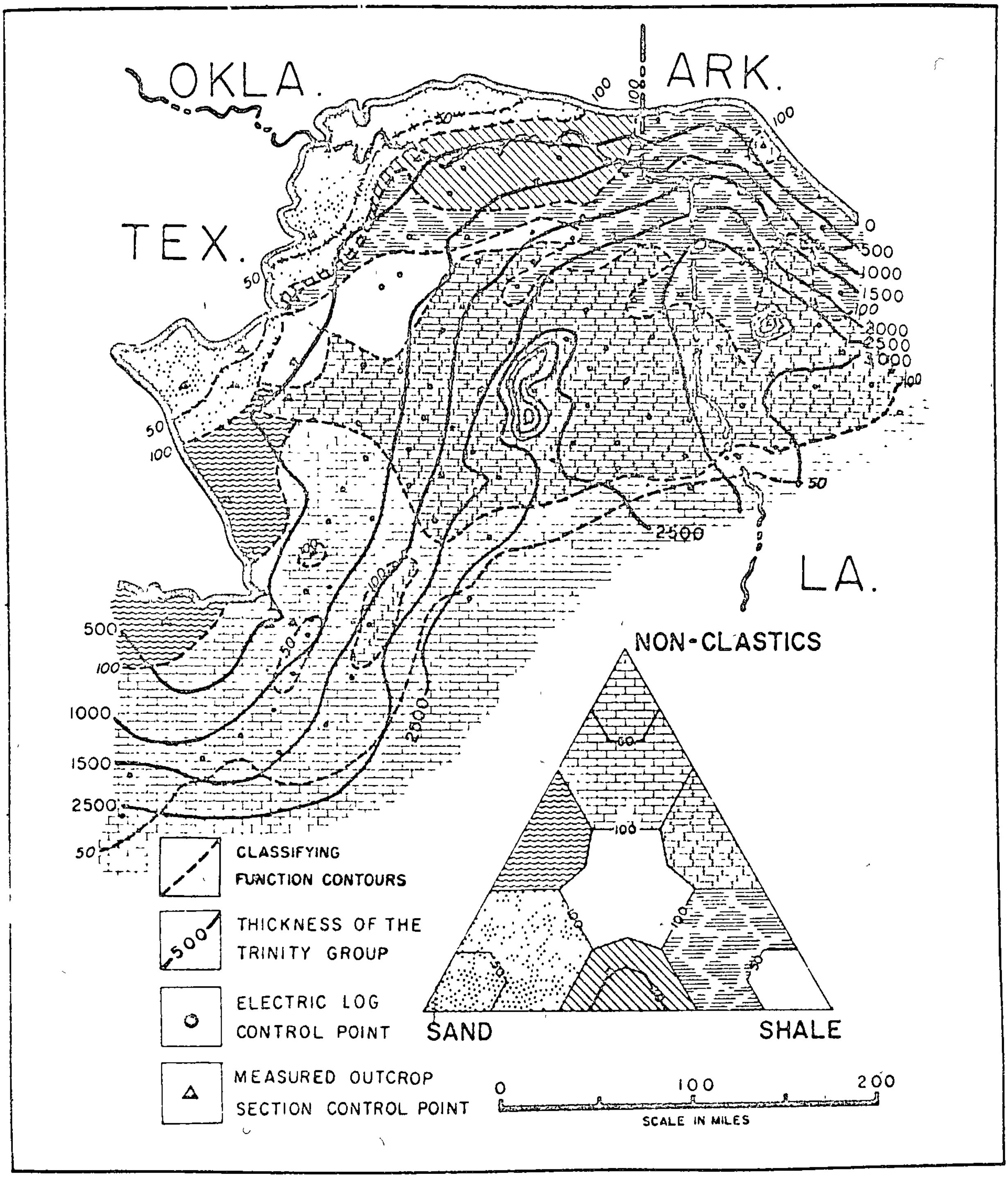
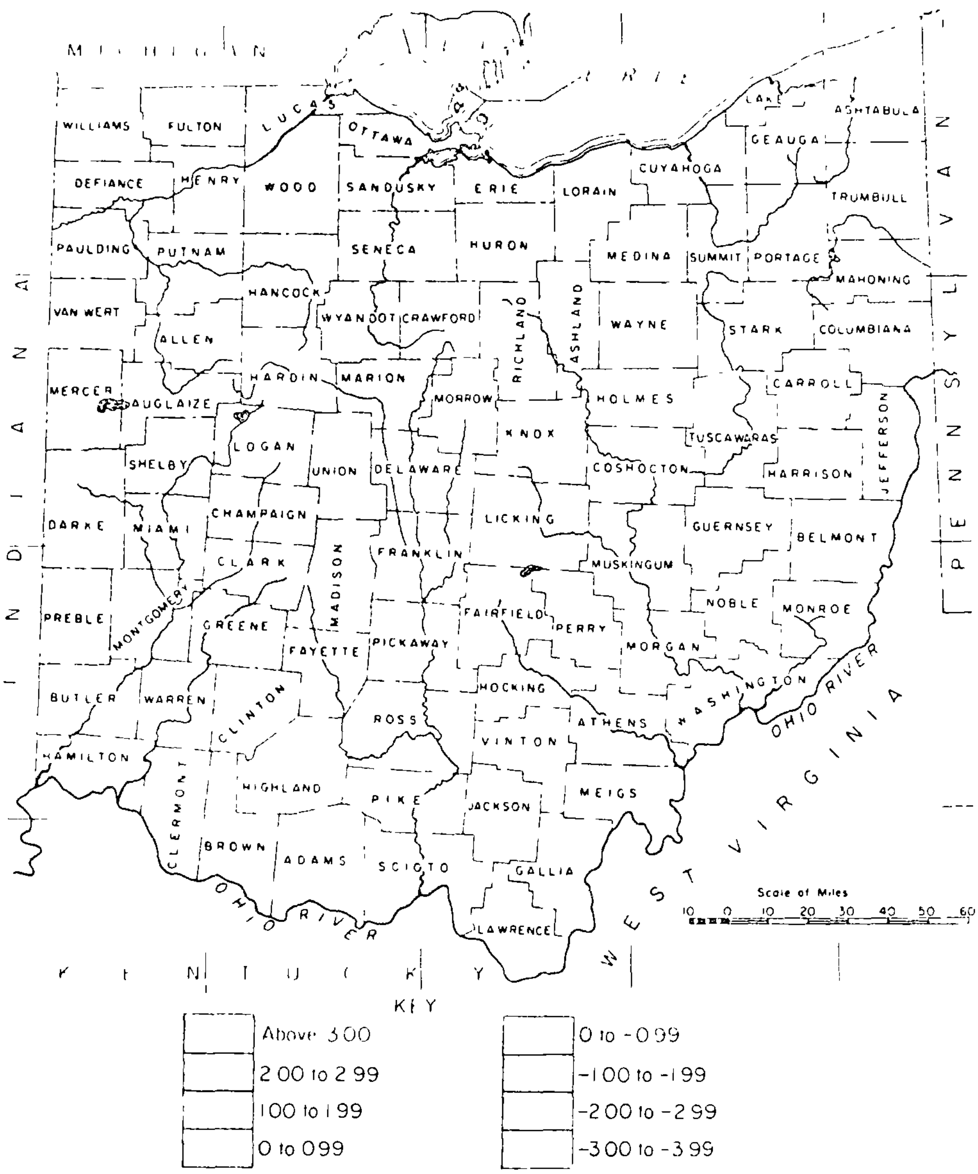


Figure 17. "Isentropic" lithofacies map a geological example of multiple component mapping (from Forgotson, 1960)



Standardized scores on component I

Figure 18 Reducing multivariate data to a few dimensions a factor map of agricultural regions for Ohio depicting the general crop dimension (from King, 1969)

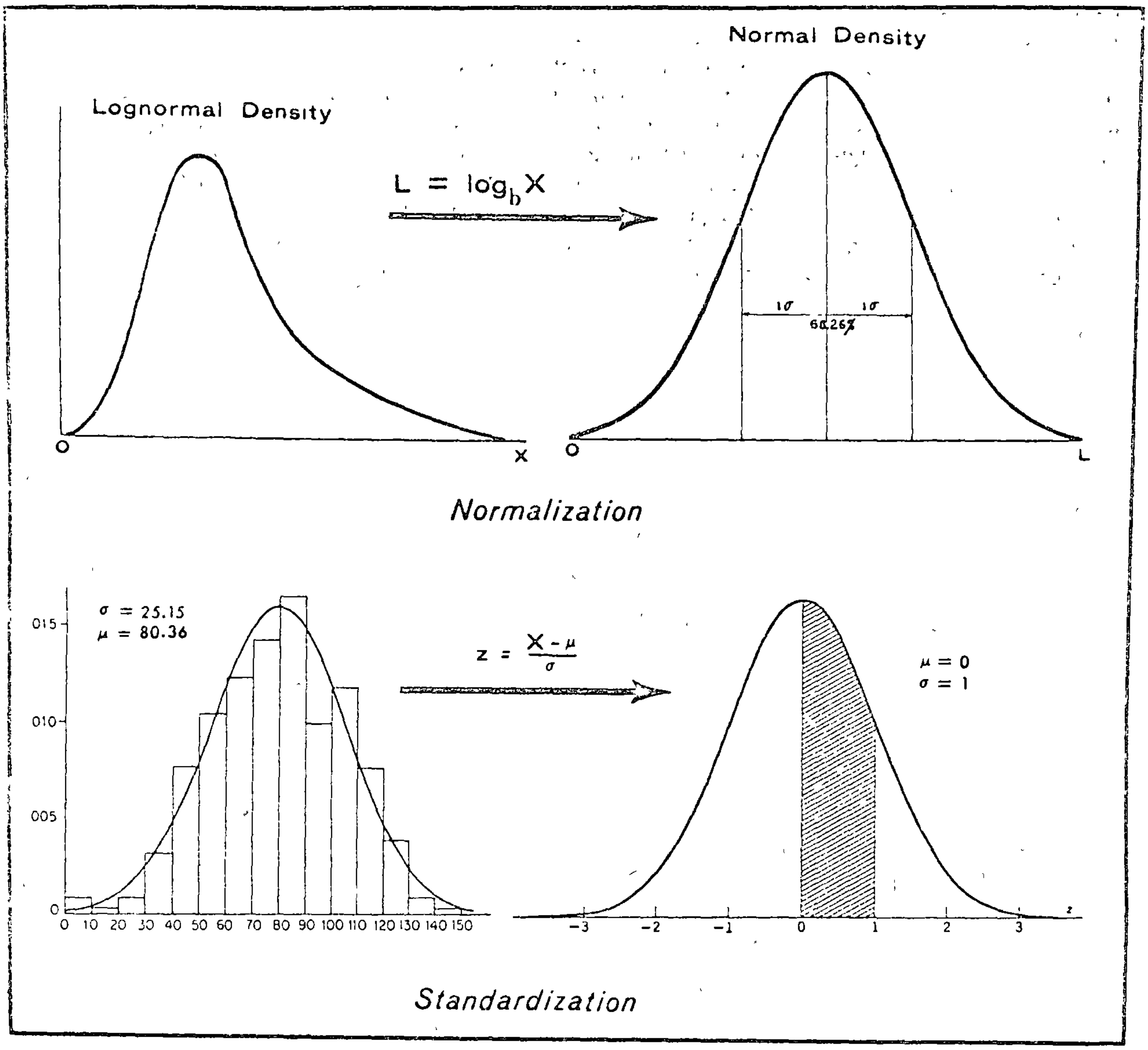


Figure 19 Preparing data for subsequent analysis: normalizing and standardizing transformations

[illegible] and conceptualization as possible if the full potential of creative intuition and decision making is to be realized. If the cartographer can develop more effective data reduction techniques, and the map reader can be taught to understand their underlying concept (i.e., readily decode them), then the amount of information communicated by a single map might be greatly increased.

Finally, maps must be integrated into ever higher levels of analysis if the full potential of visualization provided by maps is to be realized throughout the cyclical (hierarchical) process of scientific inquiry. The consideration of raw data alone, which has been the general rule in the past, will not satisfy the need for maps in future theory formulation and scientific model building endeavors. Cartographers will also have to go beyond descriptive statistics and univariate analysis to statistical inference and multivariate analysis. This, in turn, will require new map inputs, since the conventional method of mapping absolute frequencies or percentages seriously limits the user's ability to make inferences based on mapped variations without the danger of basing conclusions on nonsignificant random fluctuations.

Extensive research is needed throughout the area of premap information processing before the technique can become generally effective. Although the mathematical/statistical procedures are relatively straightforward, few cartographers possess the background necessary for their application. Interpretation of the results—mapped or otherwise—is often extremely difficult as well. The relevant point is that preprocessing provides solution to the problem of data overload only to the extent that effective decoding is possible. Many established techniques are still well beyond the grasp of many cartographers as well as map readers. The research cartographer's task will be to explore the potentialities of each of the numerous techniques and to provide practicing cartographers and map readers with interpretative guidelines which will insure proper and effective use of preprocessed data input.

III. MAPPING

The medium of graphics has long been used to create two-dimensional representation of spatial phenomena for the primary purpose of visualization and, for many, this has also been the essence of "cartography." Recent liberalization of the interpretation of cartography places actual data portrayal, or mapping, in perspective as an integral step in the broader context of cartographic processing. In order to evaluate the actual mapping process it is necessary to assess how well it functions as a communications link in a much broader data handling system. The interdependencies between data collection, mapping, and information retrieval activities weigh heavily in this evaluation.

Conventional mapping is extremely time consuming and costly. In order to keep pace with developments in other fields, map production efficiency and map effectiveness must be continually increased. Automation, along with improvements in materials, equipment, and techniques, should dominate the drive to produce larger numbers of maps in less time and at lower cost per unit. Human intuition and creativity, in conjunction with considerable experimental research, should lead to the production of maps which more effectively serve their intended purpose.

(1) Production Efficiency

Economy in map production is primarily a function of total mapping time, which in turn depends upon labor, material, and equipment requirements. Since *human labor* accounts for the largest proportion of work-time consumption and production costs, yet is generally characterized by low productivity, an obvious way to gain map production efficiency is to reduce the amount of human skill and labor involved in the mapping process. Labor saving has been accomplished by (a) automating traditional labor-intensive mapping procedures, (b) employing manually controlled equipment and utilizing prepared materials which increase human efficiency by reducing task time/effort, or (c) altogether eliminating traditional labor-consuming steps in the mapping process.

(a) Automation

Probably the most exciting and influential recent development in map construction is the trend toward computer-assisted mapping. The bulk of the labor involved in mapping is straightforward, redundant, and well suited for machine operations. High-speed, digital computers and automated drafting machines have already proven especially valuable for routine calculation and mapping chores. Automated devices are, in fact, fast becoming indispensable in meeting demands for more rapid production of wider varieties and greater volumes of maps. Much of the drudgery of routine compilation and the preparation of reproduction copy can now be automated.

The traditional method of treating map inputs as something unique to particular map compilations is inefficient. The conversion of data to the graphic mode always presented serious information storage and retrieval problems. Changing the scale or purpose of a map usually requires going back to the beginning compilation state and starting over. In contrast, computers store geographical information (i.e., position and intensity values) in highly-accessible digital forms. Once a mapping system has been preprogrammed and data have been made machine-usable, little human intervention or effort is required to produce maps automatically in large number and variation from a single set of data. The data are also accessible by multiple users. The benefit of using automated procedures in conjunction with multiple-purpose data files (called *data banks*) can be a tremendous reduction in data handling time and cost, and significant mapping flexibility. Keep in mind that this increased flexibility provides solution to problems of cartographic analysis only to the extent that we exercise *cartographic sense*. There is no guarantee that ten, or even twenty, computer maps will inherently be any improvement over one well thought out manually produced map.

Actually, it is not so much the flexibility as the increased speed of automated procedures that is most exciting developmentally. If maps can be made more quickly, the scope of cartographic analysis can be extended to include shorter-life phenomena. A two-year turn-around time for a standard topographic sheet may suffice in light of the normal rate of change in landform, but air pollution maps of a metropolitan area must be available within a matter of hours if they are to be of more than historical interest. Policy decisions affecting people must often be based on highly specialized, up-to-date maps of ephemeral subject matter and they may be used only a few times at most The slow *response time* and tremendous cost of producing many current maps, such as the USGS topographic sheets, definitely do not satisfy such administration and planning needs

Input to automated systems consists of alphanumeric data or symbols communicated directly through an on-line terminal (which can be an ordinary home telephone) or recorded on Hollerith-type punched cards, punched paper tape, or magnetic tape Computer-generated mapping consists of adding data to a preprinted base map, or reproducing a map on blank recording media such as paper, plastic, film, or cathode ray tube The bulk of automated equipment and associated techniques performs the task of positioning symbols by relative location.

Standard computer printout equipment can function as a primitive but economical class of automated mapping devices The computer may be programmed in two ways. The simplest, but least flexible, procedure is to provide a complete set of instructions which prespecify each printing position For example, map elements to be printed may be punched on Hollerith cards in their proper spatial format and the card deck listed on a printer The disadvantage of this technique is that a new set of instructions is required for each variable, and the projection and scale cannot be changed An alternate approach is to identify information by locational coordinates and then create a map-printing computer program to prepare its own printing position specifications This technique has the advantage that the program can be used to select variables, alter projection and scale, and change plotting symbols A large number of standard mapping programs of this second type are now available and widely used in diverse disciplines

Common line-printer created map formats include outline maps annotated with census data, contour-like symbol strings or bands, isopleth or choropleth appearing symbol fields, and plots of discrete observations (Figure 20) Since common impact line printers are restricted to those symbols found on most standard typewriters, maps produced on them are generally characterized by low *geographical resolution,* but this does not have to be the case. Theoretically it is possible, though not always practical, to make printer maps at large scales and then reduce the output drastically This actually has been done by Yoeli in his analytical hill shading procedure (Figure 21) Although printer maps have proven extremely useful, they are limited to discontinuous symbols and are further restricted by (1) the inability to center these symbols between column positions across the page, or between line positions down the page, and (2) the fact that significant directional distortion can accrue due to the rectangular form of printer positions.

A second class of mapping machines is the coordinate plotter The more flexible systems possess several operating

Figure 20 Isopleth map produced on line-printer

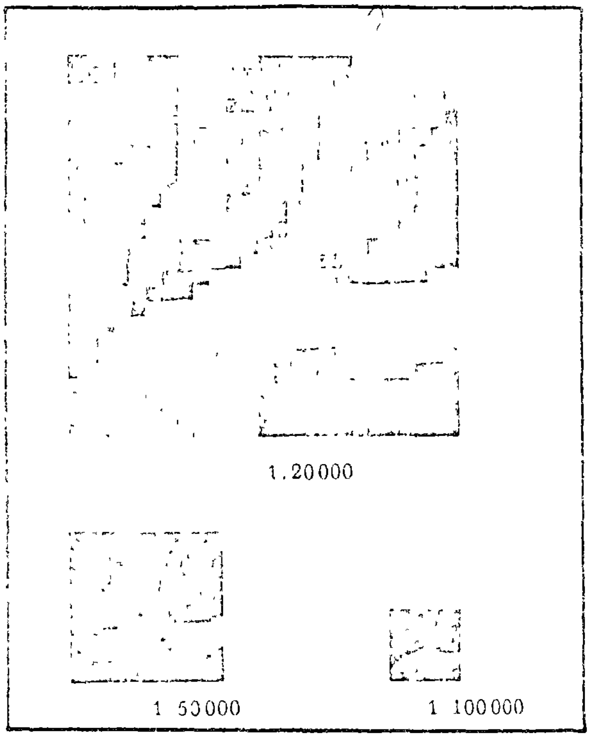


Figure 21 Shaded relief on line-printer (from Yoeli, 1965)

modes For example, by using interchangeable instrument heads, the output format can be an inkdrawn (using ball point, wet ink, or nylon tip pens) or printed original on paper, a scribed negative, or an exposed film positive image. Primitive coordinate plotters--that is, those which only plot symbols can draw, scribe, print, or photographically expose data, symbols, and alphanumeric annotations at specified coordinate positions. These plotters have the same application as computer printers, but possess the advantage of finer resolution, larger choice of symbols and image field, color capability, and greater selection of drafting base materials.

The more sophisticated continuous-curve plotters display even greater flexibility in application. In addition to plotting symbols, these machines can draw, scribe, or photographically expose continuous straight-line segments connecting strings of points which are identified by coordinates (Figure 22). In this fashion, a straight-line approximation of any curve can be produced at tolerance levels unmatched by manual plotting methods. Basic automated scribe systems are already fully capable of producing accurate linework at least equal in quality to good manual scribing. Acceptable gray tone patterns for areal shading have also been produced by these plotters (Figure 23). Another high speed continuous curve plotter uses a laser optical system. A reflected laser beam plots the map image, line by line, on photographic film or paper. The extremely fine resolution permits high density of lines and patterns, and thus makes a wide range of gray shades possible.

Common cartographic applications of continuous-curve plotters include drawing map projections, grid reference systems, world outlines, political and census tract boundaries, perspective diagrams, isolines, transportation networks, and river systems. In most cases the symbol plot capability is used in conjunction with continuous-curve plots.

Computer-controlled electronic mapping systems constitute a third class of automated mapping devices. Map features are traced electronically and displayed by means of a CRT (Cathode Ray Tube), an electrostatic printer/plotter, or similar directly viewed device. Nonimpact electrostatic raster scanners are less than two years old but already appear to be one of the most significant developments in display devices. These machines are capable of proceeding at such phenomenal speeds that they can actually keep pace with the internal speed of modern electronic computers. Equipment varies greatly and is under continuing development but, in general, output can be printed line by line at a current rate of from 300 to 5000 lines per minute (paper speed of 1 to 2 inches/second) with a resolution of 70 to 100 dots or impressions/inch. An 8½" by 11" map can be plotted in less than 7 seconds regardless of complexity (Figure 24). The factors of low cost, hard copy, electronic speed, shading, reliability, and flexibility (one device can often replace the older mechanical line-printers and incremental plotters) associated with electrostatic mapping methods suggest that cartographers are about to realize the full potential of computers in cartographic information processing. The main disadvantage of electrostatic printer/plotters is that each "writing nib" which can produce a dot is associated with a corresponding binary digit within the computer and the computer must first produce the map image dot by dot in binary form.

In contrast to electrostatic printer/plotters, the CRT represents an ephemeral display technique in which the map image is created by deflecting a beam of light in both the x and y directions. CRT's have the symbol plotting capability (black-and-white or color) of continuous, dashed, or dotted lines of variable width, alphanumerics, and programmed graphic symbols of variable size and orientation (Figure 25). The CRT display is characterized by extreme flexibility, high speed, high resolution, relatively high cost (except on a production basis), small image area, and low accuracy. Dynamic CRT displays make it possible to change information content almost instantaneously (including map updating and correction), symbol type, projection, and other map parameters. When employed as

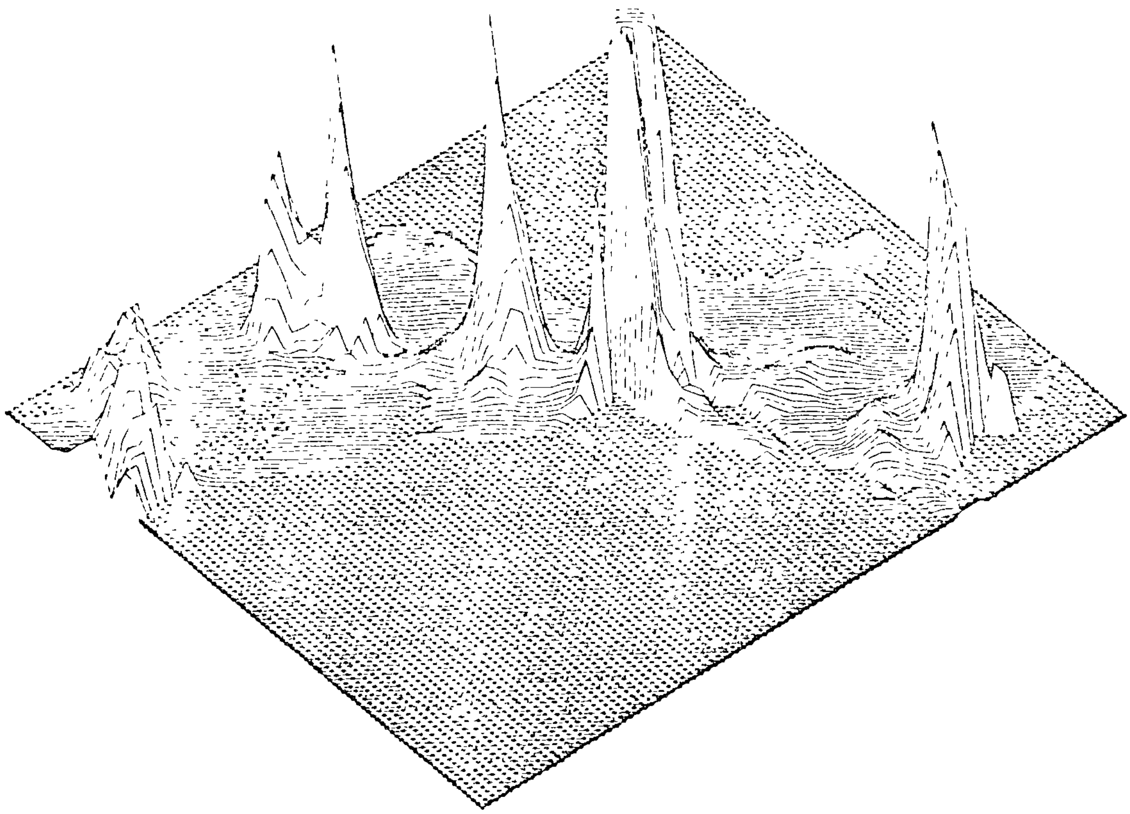


Figure 22 Line plotter map 1960 population density by counties within the Northeast Corridor (from Warntz, 1971)

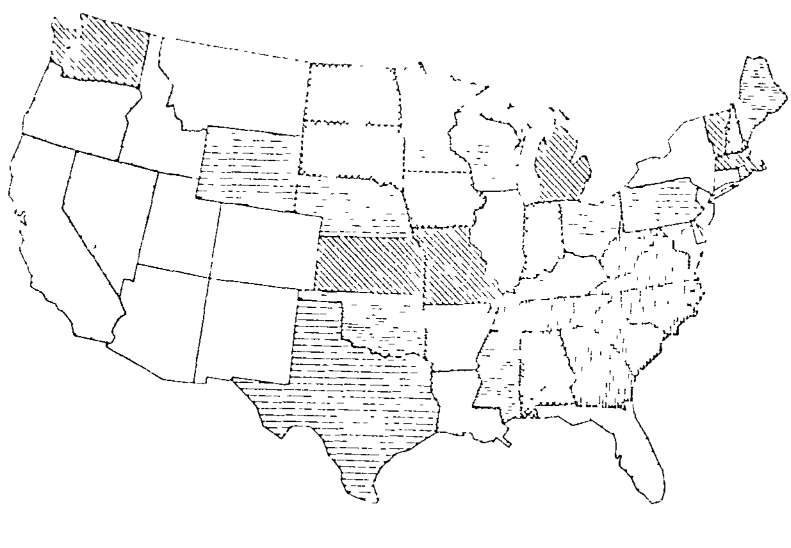


Figure 23 Areal screens on the coordinate plotter an experimental choropleth map (from Cartwright, 1970)

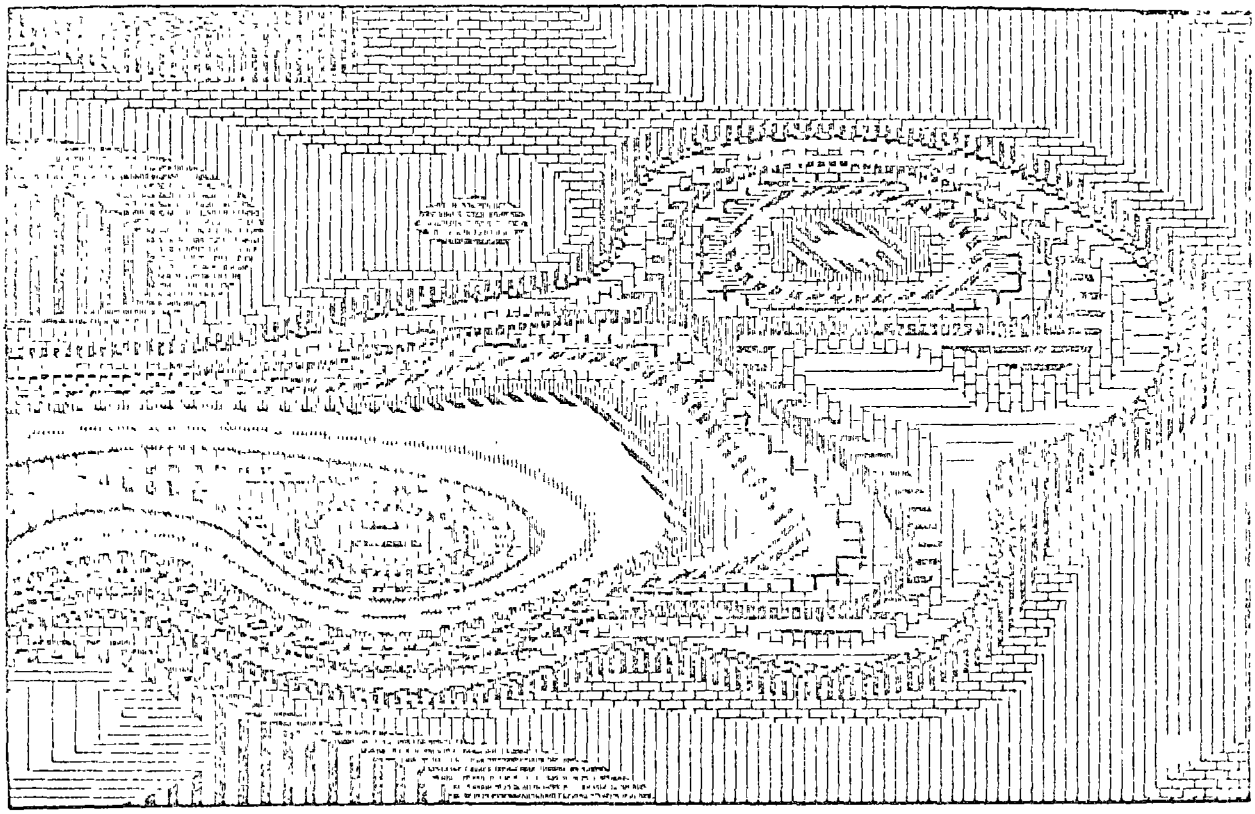


Figure 24 Map-like display produced with an electrostatic raster scanner (from Arnoldy, 1971)

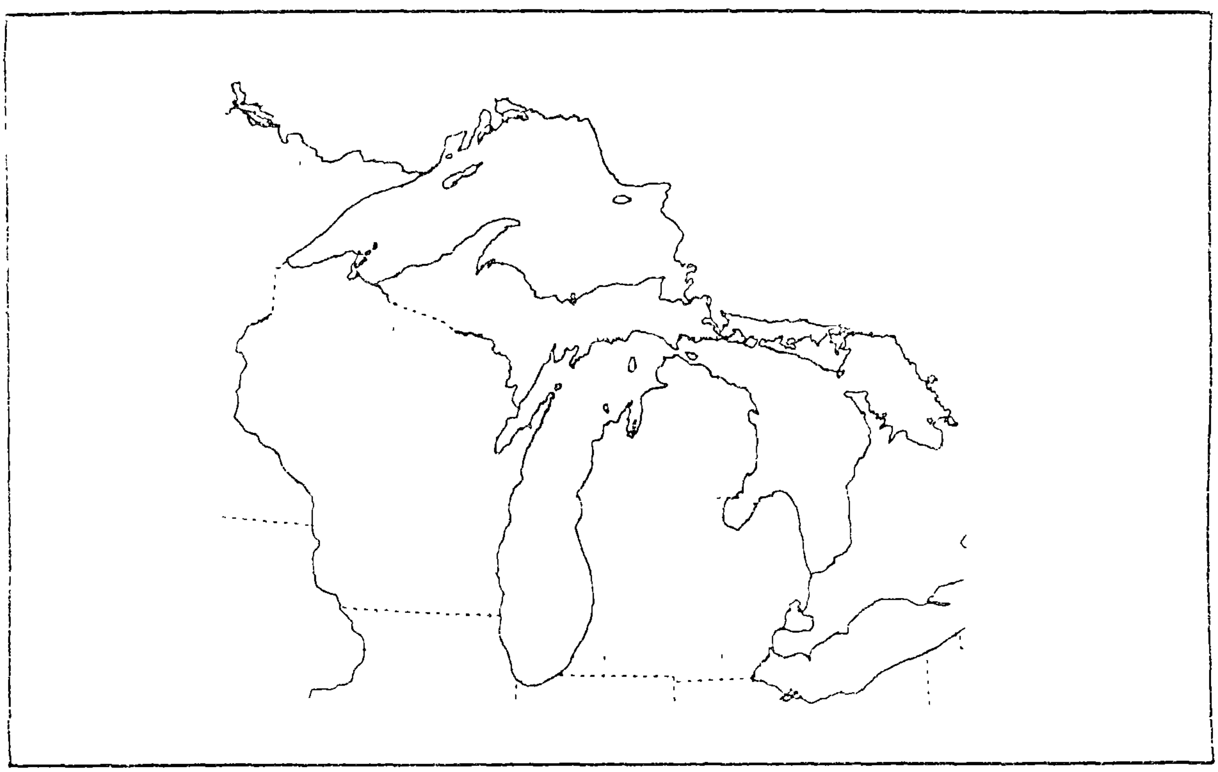


Figure 25 Cathode ray tube outline map of the Great Lakes region (from Hershey, 1963)

interactive computer consoles, CRT devices appear particularly well suited to cartographic experimentation, since they permit the researcher to view in near real time the progressive effects of these parameter alterations In addition to direct viewing, it is possible to record graphically CRT output directly on film or indirectly on, for example, an electrostatic printer/plotter. The production capability of CRT's has proven useful in producing movies or animated maps. Although electronic systems are dynamic and have the advantage of speed over mechanical systems, they are still inferior to mechanical procedures for the production of large maps with extreme accuracy.

A fourth class of "mapping" machines is that used to produce raised relief maps The first step is to carve a three-dimensional representation of a statistical surface from a solid block of material, using a computer-controlled three-axis milling device and digitally stored data. The resulting relief model can be hand "finished" and serve as the end product, or it can be used as a form in the production of relief maps The subject for mapping does not have to be the landform surface.

Basic efforts to automate the cartographic process have involved vast financial and technical resources in both hardware (equipment) and software (program instructions) development and, therefore, have remained the domain of the large mapping organizations which happen to be interested primarily in the preparation of general purpose topographic maps, nautical charts, aeronautical charts, lunar maps, and other nonthematic maps Automation in thematic cartography has essentially involved developing software that could exploit the capabilities of existing combinations of computers and peripheral lineprinters, plotters and cathode ray tube display systems Several large, flexible thematic mapping systems have been or are currently under development, including AUTOMAP at the Central Intelligence Agency, SYMAP at Harvard University, LINMAP and COLMAP at the Ministry of Housing and Local Government in England, the Automatic System of Cartography at the Royal College of Art, and the Canadian Geographical Information System

The ability to develop automated thematic mapping routines on limited budgets utilizing widely available multiple-purpose equipment has encouraged extensive experimentation The proliferation of applications of automation to thematic mapping now encompasses most common display methods The same techniques and practices have been independently developed and applied over and over by scientists with a wide range of interests The general lack of professional communication between these individuals has too often resulted in duplication of effort, a failure to learn from others' mistakes, a lack of standardization (as to program language, computer installation, user access, etc.), and, in general, a tragic waste through redundant use of limited resources This communications breakdown has caused needless delay of full realization of the potential of automated thematic cartography. Although individual mapping routines can sometimes be found in publications by their authors or extracted from collected works of computer programs for environmental scientists, the bulk of existing computer mapping capability is still not available to the individual user and is not fully utilized. The need for a coordinated package of user oriented mapping programs such as might be produced by a summer institute on Thematic Mapping by Computer and which could be simply modified for use on most computer facilities is obvious and urgent. In the meantime the recently established Geographical Program Exchange (GPE) will be a most useful service (see Wittick, 1971).

(b) Equipment and materials

Cartographers are currently faced with an unusually vast array of new drafting and reproduction materials, equipment, and associated techniques which account for substantial savings in labor. The individual and combined effects of these developments, and the resulting implications for geographical mapping, are still not well understood Yet it is possible to identify in superficial fashion those influences which already are fast changing the nature of mapping activities.

Possibly the most obvious labor saving change is that the process of creating a map image with pen, ink, and paper has rapidly disappeared from all except small mapping operations One reason for this trend is that a large variety of high-quality *preprinted materials* suitable for mapping have been made commercially available at relatively low cost At least one of the common brand names, such as Deca-dry (Chart-pak), Formatt, Prestape, Instantype, Letraset, Paratone, Microtype and Normatype, are usually available at local graphic arts supply outlets These adhesive-backed, stick-up materials now include letters, numbers, symbols, lines, and areal tones/patterns, in a wide variety of types and sizes (Figure 26) The cartographer has only to position the chosen materials in the proper map format. Most companies that supply stick-up materials will also accept your own art work from which they will make up sheets of special symbols You can accomplish the same thing locally by running a film positive of special symbols or redundant art work in contact with Dietzen's diazo-sensitized, adhesive-backed film through an Ozalid machine It is common practice for limited production mapping establishments such as university laboratories to send lists of type to commercial typesetters, such as Monsen, who are in the business of producing the desired stick-up materials

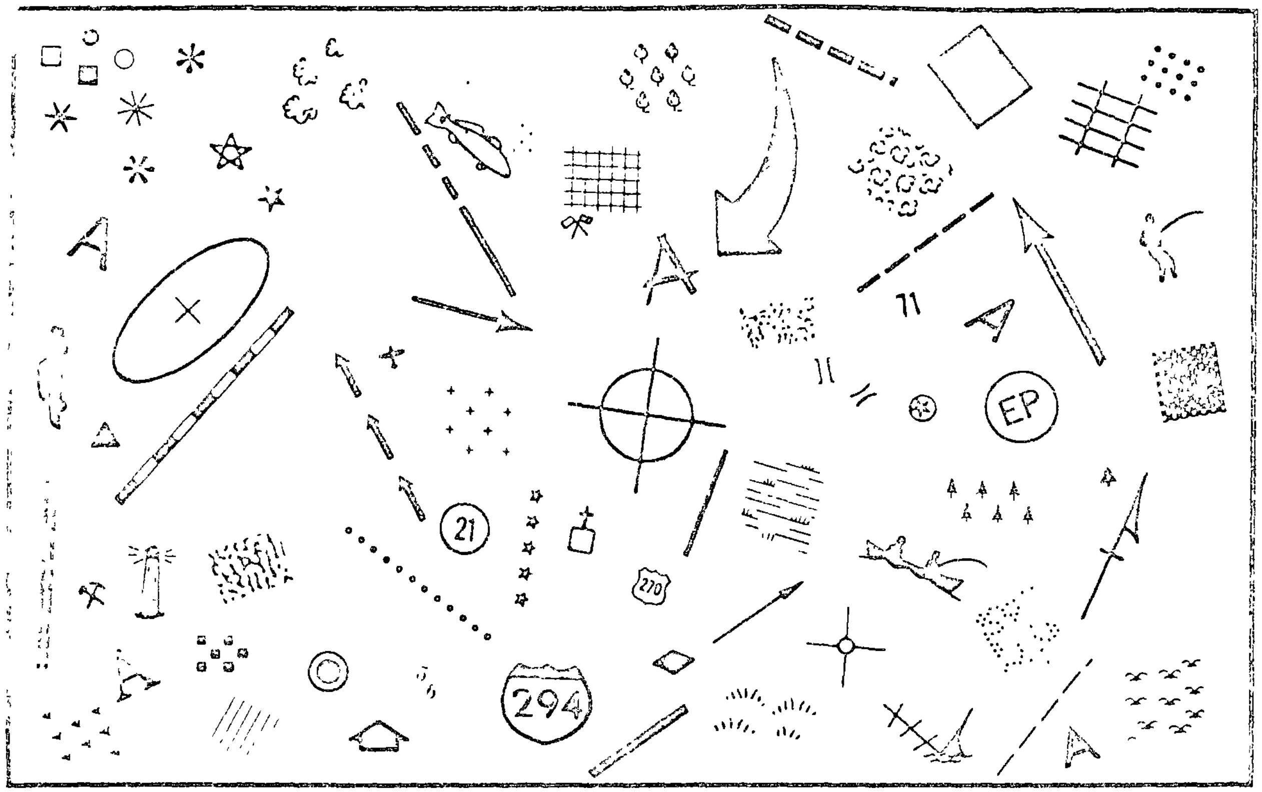


Figure 26 A small selection of commercially available stick-up materials

One university cartography laboratory (University of Wisconsin) is actually large enough to support a phototypesetting system capable of producing adhesive-backed sheets of set type in house. The actual procedure involved in obtaining preprinted materials is not the main point to be made here, however. What is important is that the technique of positioning pressure-sensitive symbols in a map format is extremely fast, produces uniformly high quality products, and requires very little skill on the part of the cartographer.

A second reason for the trend away from ink drafting is the recent development of *scribing*. This process of engraving lines and symbols into an opaque-coated, dimensionally stable plastic sheeting (e.g., *Scribecote*) permits the cartographer to prepare a negative map image which can substitute for a film negative in reproduction processes. The advantage of scribing is that extremely high-quality map negatives can be prepared at reproduction size with significant savings in effort and cost over ink construction techniques. It also takes less time and expense to train people to scribe than it does to train them to draft with ink.

Presensitized scribecoat materials have greatly simplified the preparation of color tint masks. The previously tedious task of preparing open window negatives by hand can now be replaced by a photoetch process (e.g., *Striprite* or *Peelcoat*) which involves three simple steps: (1) transferring the original negative image by photographic contact to the presensitized scribecoat, (2) chemical etching of the image on the scribecoat, and (3) removing or lifting out the scribecoat in the areas where the color tint is to appear.

The substitute of preprinted materials for ink and paper drafting significantly reduces the total cartographic workload. The end result is generally a higher quality product produced in less time and with less skill than maps produced by traditional methods. We should expect many future developments in preprepared cartographic materials, some of which will require radically different procedures from those associated with conventional map compilation of the pre-1970's.

(c) Task reduction

A third mechanism for increasing map production efficiency is to develop new methodologies which actually reduce material or equipment needs through the elimination of traditional steps in the mapping process. Probably the most dramatic recent saving in production time and costs has resulted from *negative scribing*, which eliminates the photographic step in map processing. Research associated with scribing entails the search for better base materials, scribe coats, instruments, and the coupling of scribing techniques to automated cartographic systems.

New *nonprinting processes* which eliminate the necessity of expensive photographic processing can also provide more economical solutions to production problems in many cases. High initial costs have long plagued cartographers, especially when only a relatively few copies of a particular map were required. Nonprinting methods are relatively slow, but have the advantage of low initial cost compared to printing methods and, therefore, are ideal for short runs (i.e., when only one or several copies are needed). A second advantage of nonprinting processes is that extensive photography and darkroom facilities are not required. Since the wide variety of nonprinting procedures are not simply classified, only those techniques which cartographers have found to be commonly available and most useful will be mentioned (Figure 27). One general process produces direct contact positives (trade name Diazo or Ozalid) and negatives (similar to standard blueprints). Sensitized printing papers are used and the copies necessarily are the same size as the original map. A range in paper and print color is possible. A second general process, called photocopying, actually refers to a number of processes based on a modification of the principle of conventional photography. The best known photocopying methods utilize sensitized paper and include the "autopositive," "transfer," and Photostat (trade name) processes. The "autopositive" technique is a contact method which directly produces a positive image. The "transfer" technique is a contact method which requires an intermediate negative step. The Photostat technique involves an intermediate photographic procedure which produces a negative image and permits reduction and enlargement of the original. A third general method, Xerography (trade name Xerox), is an extremely fast electrostatic process which does not require chemicals or sensitized printing paper. A positive copy may be produced in several colors, and reduction and enlargement are possible. A fourth class of so-called proofing materials has been developed which can be used to produce color proofs and contact positives or negatives under standard cartographic laboratory conditions (i.e., without darkroom facilities). Since these are contact processes, reduction and enlargement of the original are not possible. These processes combine the various negative originals by photocomposition to produce a proof copy which facilitates review and editing prior to final map printing.

Continuing developments in presensitized papers and films, new plastics, color proofing materials, and the like, along with the increasing general availability of nonprinting processes, make the production of a few copies of a map much less of a problem than was previously the case. In fact, the most difficult problem facing a cartographer now or in the future is that of selecting the most reasonable

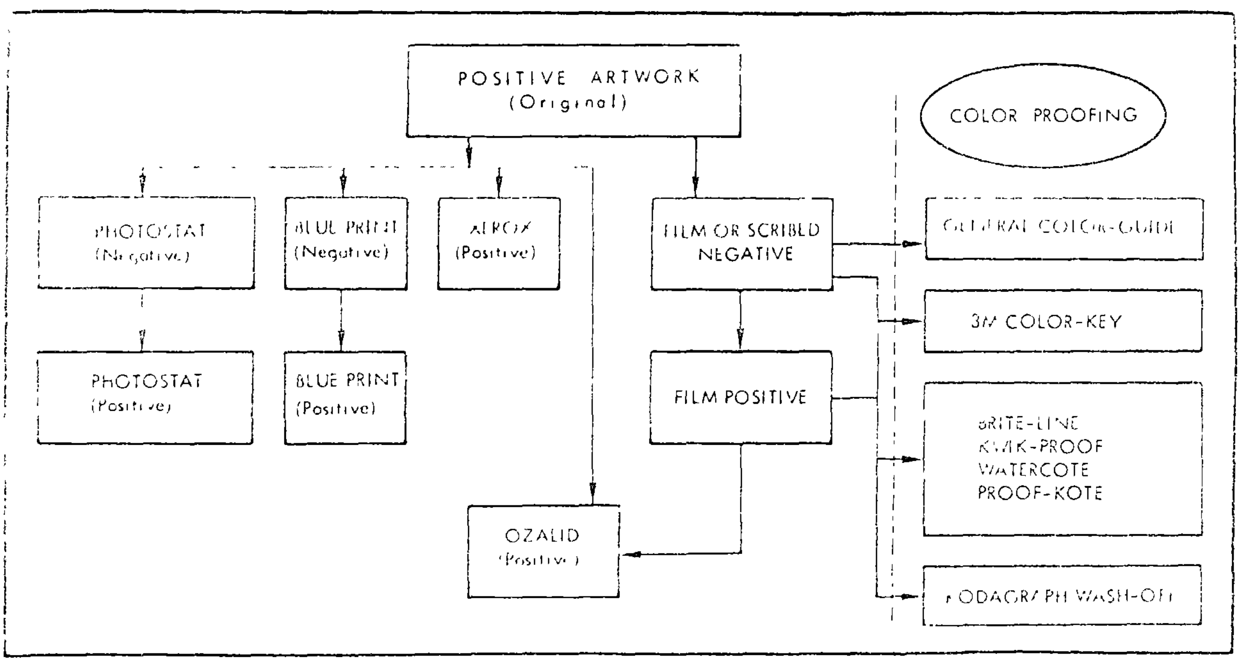


Figure 27. Non printing processes for map reproduction

alternative from the wide choice available. Future cartographers will have to know which mix of processing techniques provides the most economical solution to individual user needs. Often this may mean no map at all, at least in the hard copy sense, but rather a displayed (CRT, television, electroluminescent/thermochromatic, etc.) image that is only temporary in nature and readily disposable once it has served its purpose of augmenting the use of the conventional map.

Photomapping is rapidly gaining acceptance as a technique that (i) can be readily tailored to user planimetric accuracy standards and (ii) permits selective cartographic enhancement of aerial photographs. The difference in production time and cost between photomaps and conventional line maps can often be considerable. The cartographer must, therefore, learn to take full advantage of the complementary nature of most photomapping techniques, since they facilitate selective elimination of needless steps associated with traditional line map production while still satisfying particular user requirements. For example, Lyddan 1971 states that photo-revised USGS topographic maps showing updated changes in a distinctive color have not only cut standard updating costs by 2/3 but also produced a useful and well received product. The task ahead is to find answers to such questions as (a) In what production situation can photomaps be produced more quickly or economically than conventional line maps? and (b) In what situation can photomaps provide information which it would not be possible to show on a conventional line map?

Research associated with the above and related methodologies is of the same essential nature. The general problem is one of exploring the potentialities of each new methodology relative to all others and thereby establishing its practical working limits. The major question involves the balance between expensive equipment and retraining on one hand against gains in production speed and labor efficiency on the other.

Efficiencies related to additional labor-associated procedures can also be mentioned. For example it is convenient to categorize map information by type and then build maps in modular fashion using *overlay methods*. The phenomenal growth in the use of this technique closely parallels the increasing use of color in map reproduction, the design of color projection equipment, and the development of transparent films and plastics that take color symbolism well. Overlaying is particularly valuable in large mapping agencies where it is common to make a number of maps depicting different attributes of the same region. Once overlays of hydrography, transportation, typography, and so on, have been compiled, they often can be used over and over many times and in different combinations and thereby suit various map user needs. Thus, by reducing mapping effort that is limited to a single application, considerable economies in production are gained.

Overlaying can also be a great saving for the map user. The capability of being able to build a series of *tailor-made maps* from a single base map and the appropriate set of overlays provides valuable flexibility and saves the user the expense and burden of needing a whole set of *finished maps* depicting the same information. Superimposing layers of data in this fashion in effect adds another dimension to the limited information-carrying capacity of the two-dimensional map surface. In general, since the notion of information separation and overlaying is well suited to automatic data processing and mapping, we can expect a large amount of attention to be directed toward this concept in the future. The task ahead largely entails the selection and/or design of combinations of symbol schemes which facilitate, rather than distract from, map interpretation.

Another way to achieve mapping efficiency is to utilize map forms which meet specific user requirements rather than satisfy a heterogeneous user audience. It has been common practice for map users who are interested in highly selective, abstracted information to find themselves in a position where they must make do with a map designed with some other purpose in mind. All too often the only source available is a multipurpose map base that is cluttered with extraneous symbols which may actually distract significantly from optimum interpretation and expected impact characteristics of the graphic. In addition to having to deal with a lesser quality product, the map user may pay for compiling mapped information he does not need. Or, as another example, the conventional hard copy, multicolored map is well suited for certain purposes but also very awkward or unnecessary in other situations.

In light of the previous discussion, it is reasonable to conclude that there exists an overall trend toward increased efficiency in map production. Since mapping efficiency is closely associated with map effectiveness, the next section is devoted to consideration of techniques which might facilitate closer tailoring of map form, content, and style to user needs.

(2) Product Effectiveness

The key to producing more effective maps is to utilize better the unlimited possibilities for variation in mapping. The bounds of mapping are set primarily by the limits of human imagination and ingenuity. Close scrutiny of rigid, conventional ways of thinking should reveal fruitful innovative areas. For this purpose, it may be helpful to treat maps as graphic model interpretations of spatial phenomena. This point of view encompasses the powerful concept of *symbolic representation* on one hand and that of *scientific model building* on the other.

(a) Symbolic representation

Man's thinking processes are inseparably linked with visualization processes which are often based on systems of signs or symbols. Over its long history, cartography has empirically developed the use of signs to a high degree of elaboration and complexity. In fact, to the extent that maps communicate geographic information, they comprise a language of graphic portrayal. Additionally, it seems only logical that the concepts and methods developed by linguists and logicians for study of natural (words) and artificial (logic and mathematics) languages would provide appropriate starting points for the description and analysis of the special language of maps. Unfortunately, formal study of the conceptual nature of map symbols in cartographic communication is almost nonexistent and the logical functions and properties of the geographical map as a symbolic system remain essentially unanalyzed. The promise that study of the theory of signs, or semiotics, holds for the deeper understanding and analysis of the symbolic nature of maps is just now attracting attention (see Dacey, 1970). Analysis of the cartographic language from the semiotic viewpoint would involve study of its formal structure (called *syntactics*), its relation to objects designated (called *semantics*) and to the persons who use it (called *pragmatics*). When translated into mapping terms, the previous statement means that we should focus on three major dimensions of cartographic processing: study of the relation of map symbols to their referents, study of the relation of map symbols to one another, and study of the relation of map symbols to their interpreters (Figure 28). Initial work relating to the semantics, syntax, pragmatics, and other attributes of a map language is stimulating but has not yet gotten far. Important differences between the requirements of a map language and conventional linguistic structure have also been noted. In particular, problems of "neighborhood" and "juxtaposition" result from the multi-directional structure of spatial patterns.

Lack of a formal definition of the language of cartography should not prevent us from making a number of observations on the symbolic aspect of cartographic representation. Viewing map expression as but another form of symbolic representation helps illuminate the natural possibilities and limitations in map symbolism. Cartographic representation of geographical phenomena is restricted to the use of point, line, and area emphasizing symbols. But within these constraints there is, theoretically at least, infinite freedom. Probable symbols for depicting referent

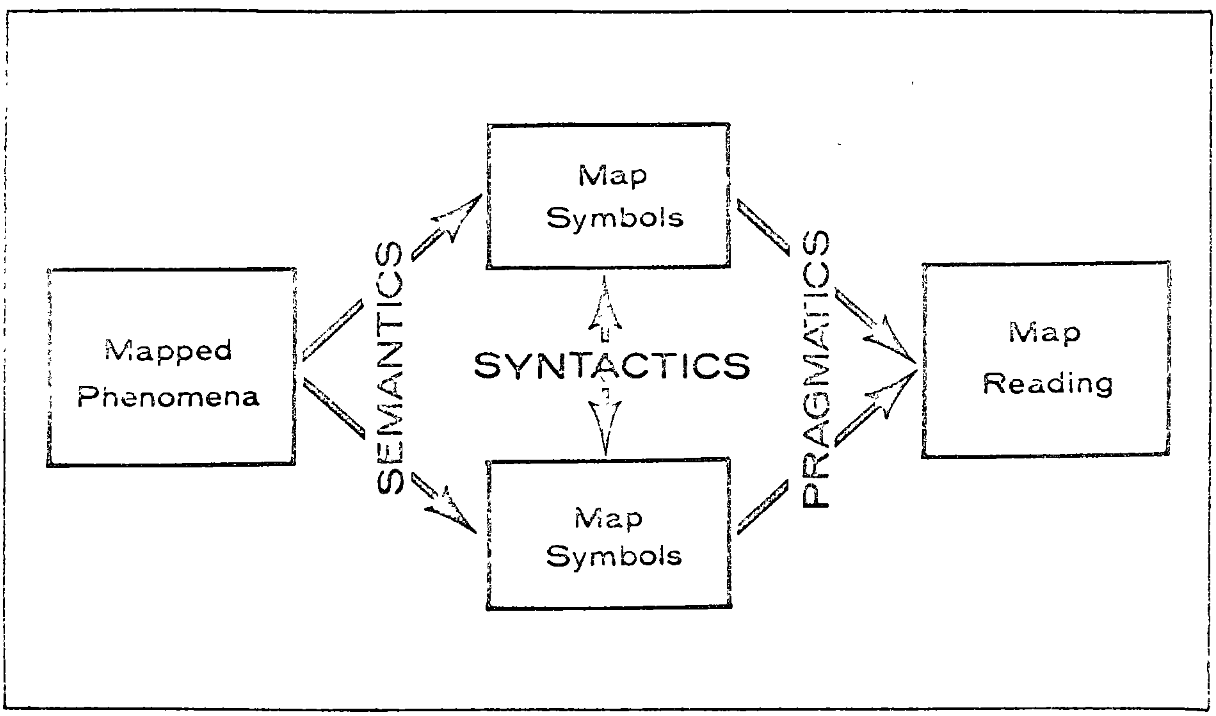


Figure 28. Semiotic: a cartographic interpretation of the theory of signs (after Morris, 1938)

phenomena can be selected and combined in various manners to produce a variety of (map) solutions, each of which can be tested against geographical reality.

Maps have strong intuitive appeal. Their widespread acceptance suggests that spatial phenomena are especially suited for graphical analysis. Map symbols are used to generalize their referents in many ways (Figure 29) At the least abstract level are the "true-to-nature" or *natural* symbols which are realistically designed to physically resemble an image of the phenomena depicted Intermediate level abstractions are accomplished through symbols which generalize essential characteristics of their referents "semi-naturally " Coastlines, rivers, transportation networks, and contour lines are commonly represented semi-naturally At the highest level of abstraction, point, line, and areal symbols are used to generalize their referents "arbitrarily" through cartographic convention Use of simple geometric figures such as circles, squares, and triangles to represent point emphasis data, and the use of various line symbols to depict different political boundaries are examples Unfortunately, the relative information-carrying and transmitting capacity of natural, semi-natural, and arbitrary symbols is unknown Evidence strongly suggests, however, that there exists a need for untrained map readers to have elements more realistic and less symbolic The appeal of natural symbol coordinations with

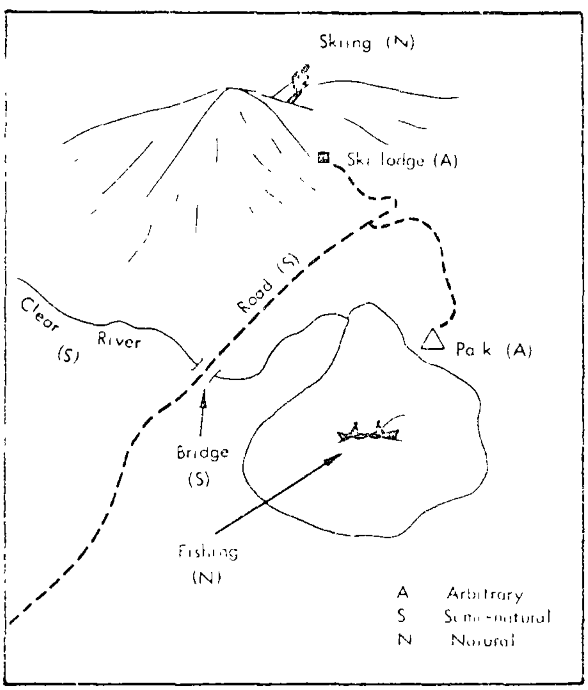


Figure 29 Symbolic coordinations natural, semi-natural, and arbitrary

untrained map interpreters appears greater than that of semi-natural symbol coordinations, which, in turn, are themselves of greater appeal than that of arbitrary symbol coordinations. The reverse appears to be the case with trained map readers. If the responses to map symbols suggested above are actually true, the important conclusion to be made is that maps must be carefully designed with their eventual users directly in mind.

Human, material, and equipment tolerances determine to a large degree the minimum practical size at which symbols can be resolved at particular map scales It would be extremely rare (if not impossible) to produce a map in which all map elements (representing mapped phenomena) were depicted in *true relative scale*, that is, where all map symbols were proportionately correct to the areal extent of their referents A degree of symbolic "distortion" is characteristic of all cartographic representation. For this reason, a *scale-free* map comparable to scale-free statistical measures is seldom, if ever, attained in actual practice. The suggestion is that the scale factor underlying any data collection scheme should be appropriately modified to account for the scale dependent nature of map symbols And this level of rigor cannot be realized until we have extensive research available involving the functional relationship existing between map scale, pattern characteristics (such as complexity, intricacy, etc). and cartographic selection and emphasis

The obvious question that arises is to what degree the language of cartography is understood by its users Centuries of experience with symbolic generalization has led to little beyond intuitive knowledge of the cartographic effectiveness of even such simple and commonly used geometric symbols as dots, circles, triangles, squares, et cetera Information is particularly lacking when the problem is to select effective combinations of these symbols to produce multiple-factor maps More attention has been devoted to pictograph type symbols where the graphic form is chosen to facilitate intuitive symbol/referent identification Yet even pictograph design and deployment have been largely a personal matter on the part of "artistic" cartographers and have seldom been based on the results of experimental evaluation.

Generally speaking, cartographers have poorly exploited the arbitrary nature of the coordination between map symbols and mapped information Regardless of the value of symbol standardization, one must conclude that symbolic conventions have not only discouraged novel map expression but also have inhibited experimentation with alternative symbolisms A less restricted attitude toward symbolism research would surely lead to more effective overall cartographic communication. Fisher's (1970) attempt at cartographic symbol taxonomy may provide a useful, though laborious, approach toward identifying logical gaps in our symbolism and suggest alternative symbol forms More important, however, is to realize that map symbolism is perfectly general and should be designed or selected to suit the purpose of cartographic analysis

Scientists confronted with conceptually difficult processes plot numbers on graphs to "see" what they mean, often under the assumption that even bad graphs may provide more meaning than tidy lists of numbers Normally we need all the insight we can get, and graphics are closely associated with the intuition that lies behind so much creative inquiry The computer business increasingly uses pictorial output Graphics are used in basic research in engineering, mathematics, physics, and other fields as a means of visualizing complex formulas and models The map, as a graphic form of symbolic representation, also serves the primary function of visualization in scientific research (Figure 30)

It appears that maps (or graphics) are not designed, intended, or well suited for precision work One should not expect detailed statistics from mapping The impact of the map is more often of greater importance than the information Maps serve well the need for a general picture of the nature of a distribution or the relationships between several distributions, at least when the patterns are not too large. In fact, quite crude graphics may satisfy (fulfill) the desired objectives of research, as most computer output, including maps well illustrates Even in cases where the approximation achieved by maps is not sufficient, the information obtained from cartographic analysis may provide a useful starting point for numerical methods The point is that the practical importance of maps is based less in their accuracy and more in their clarity, convenient manipulation, and rapid assimilation by the human user When research gets beyond the map related stages of inductive, hypothesis generation and model refinement and into the deductive testing of these notions, the formal logic and objectivity of mathematics and statistics seem more appropriate.

In light of the previous discussion, one cannot help but wonder why such large amounts of cartographic effort have gone into determining the accuracy of measurements on maps, a subject known as *cartometry* A plausible explanation may be that the physical existence of a map gives the illusion of seeming accuracy and makes measurement easy, if not strictly reliable The temptation is too great to ignore Yet research actively associated with cartometric methods seems to be at direct odds with the theoretical basis for graphic symbolism in the first place It is therefore interesting that articles on quantitative map use emphasize the investigation of instrumental and operational errors and, to a lesser degree, errors due to paper deformation and projection distortion, but treat the consequences of variations in symbol representation only in an incidental and non-rigorous way. It is this author's opinion, with certain obvious exceptions, that the sooner map makers and users realize the thematic map (in particular) is not intended to be used as a base upon which accurate measurements should be made and elaborate calculations subsequently based, the quicker we will have effective cartographic analysis. Problems concerning the playoff between positional and magnitudinal accuracy versus map effectiveness are discussed further below.

(b) Scientific model building

The cartographer is faced with an informational environment possessing enormous scope and complexity His task is to select efficiently and effectively from this environment those data which best aid the decision making process of some user To do this it is necessary to resort to *cartographic generalization*—a special case of the more general process of scientific abstraction -which imposes severe spatial, temporal, and content limitations on information to be presented. Since map creation entails the same evaluation, selection, and emphasis that characterizes scientific model building, it is not surprising that, Stephen Toulmin, well known philosopher of science, has noted, maps themselves bear close resemblance to scientific models It is this abstractive, contrast enhancing character of maps which gives them their scientific value The detail and complexity of the real world is partially overcome by creating a highly simplified but apparently rational representation demonstrating only selective features and relationships The map is created to represent a particular fragment of reality determined to be of special interest A single map is only one of many possible model interpretations of reality and must not be considered as unique.

Once a particular spatial phenomenon or combination of phenomena is chosen for cartographic depiction and has been suitably processed (i e., reduced to usable form), it is also subject to an infinite number of graphical (map) treatments or interpretations (Figure 31) This concept is demonstrated simply by the fact that different combinations of samples (type, size, schema), class intervals, unit values, graphic techniques, projections, generalization procedures, and additional *map parameters* can generate quite different overall map images using the same data Cartographers must comprehend this model nature of maps and learn when simple model differences will have a significant effect on map interpretation For obvious reasons, the general failure to grasp this map/model notion has seriously limited the scientific advancement of cartographic analysis

For pedagogic purposes it may be useful to refer to the

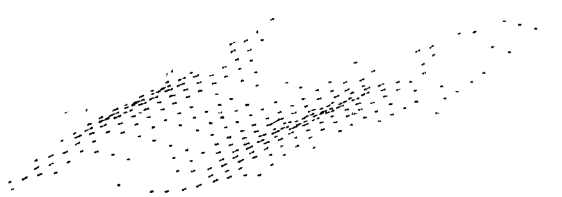
PARTIAL DATA LISTING
(Matrix notation)

I	J	Z
1	1	0.0
1	2	.8
1	3	1.6
1	4	2.1
1	5	2.6
1	6	2.7
1	7	2.7
1	8	2.4
1	9	1.9
1	10	1.3
1	11	.7
1	12	.0
1	13	-.5
1	14	-1.0
1	15	-1.3
1	16	-1.3
1	17	-1.2
1	18	-.9
1	19	-.5
1	20	-.0
2	1	1.6
2	2	2.4
2	3	3.1
2	4	3.7
2	5	4.1
2	6	4.3
2	7	4.2
2	8	3.9
2	9	3.4
2	10	2.8
2	11	2.1
2	12	1.4
2	13	.8
2	14	[illegible]
[illegible]	15	[illegible]
[illegible]	16	[illegible]
[illegible]	17	[illegible]
[illegible]	18	[illegible]
2	19	[illegible]
2	20	1.2
3	1	2.8
3	2	3.6
3	3	4.3
3	4	4.9
3	5	5.3
3	6	5.4
3	7	5.3
3	8	5.0
3	9	4.5
3	10	3.8
3	11	3.1
3	12	2.4
3	13	1.8

DATA MATRIX (Digital map)

0.0	.8	1.6	2.1	2.6	2.7	2.7	2.4	1.9	1.3	.7	.0	-.5	-1.0	-1.3	-1.3	-1.2	-.9	-.5	-.0
1.6	2.4	3.1	3.7	4.1	4.3	4.2	3.9	3.4	2.8	2.1	1.4	.8	.3	.1	-.0	.1	.3	.8	1.2
2.8	3.6	4.3	4.9	5.3	5.4	5.3	5.0	4.5	3.8	3.1	2.4	1.8	1.3	1.0	.9	1.0	1.3	1.7	2.2
3.5	4.3	4.9	5.5	5.8	6.0	5.8	5.5	4.9	4.2	3.5	2.8	2.1	1.6	1.3	1.2	1.3	1.6	2.0	2.5
3.4	4.1	4.8	5.3	5.7	5.7	5.6	5.2	4.6	3.9	3.2	2.4	1.8	1.3	.9	.8	.9	1.2	1.6	2.1
2.6	3.3	4.0	4.5	4.8	4.8	4.7	4.3	3.7	2.9	2.2	1.4	.7	.2	-.1	-.2	-.1	.1	.6	1.1
1.3	2.1	2.7	3.2	3.4	3.5	3.3	2.8	2.2	1.5	.7	-.1	-.8	-1.3	-1.6	-1.7	-1.6	-1.3	-.9	-.3
-.0	.7	1.3	1.7	2.0	2.0	1.7	1.3	.6	-.1	-.9	-1.7	-2.4	-2.9	-3.3	-3.4	-3.2	-2.9	-2.5	-1.9
-1.2	-.5	.0	.5	.7	.6	.4	-.1	-.7	-1.5	-2.3	-3.1	-3.8	-4.3	-4.7	-4.7	-4.6	-4.3	-3.8	-3.2
-1.9	-1.3	-.7	-.3	-.1	-.2	-.4	-.9	-1.6	-2.4	-3.2	-4.0	-4.7	-5.2	-5.5	-5.6	-5.4	-5.1	-4.5	-3.9
-1.9	-1.3	-.8	-.4	-.2	-.3	-.6	-1.1	-1.8	-2.6	-3.4	-4.2	-4.8	-5.3	-5.6	-5.7	-5.5	-5.1	-4.6	-3.9
-1.3	-.7	-.2	.2	.3	.2	-.1	-.6	-1.3	-2.0	-2.9	-3.6	-4.3	-4.8	-5.1	-5.1	-4.9	-4.5	-3.9	-3.2
-.1	.5	1.0	1.3	1.4	1.3	1.0	.5	-.2	-1.0	-1.8	-2.6	-3.2	-3.7	-3.9	-3.9	-3.7	-3.3	-2.7	-2.0
1.3	1.8	2.3	2.6	2.7	2.6	2.3	1.8	1.1	.3	-.5	-1.2	-1.8	-2.3	-2.5	-2.5	-2.2	-1.8	-1.1	-.4
2.5	3.1	3.5	3.8	3.9	3.8	3.5	3.0	2.3	1.5	.8	.0	-.6	-1.0	-1.2	-1.1	-.9	-.4	.3	1.0
[illegible]	3.9	4.3	4.6	4.7	4.6	4.2	3.7	3.1	2.3	1.6	.9	.3	-.1	-.3	-.2	.1	.6	1.3	2.1
[illegible]	[illegible]	4.4	[illegible]	4.8	[illegible]	4.4	[illegible]	[illegible]	[illegible]	1.0	1.1	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1.6	2.1	2.5	2.8	2.9	2.9	2.6	2.1	1.5	.9	.2	-.5	-1.0	-1.3	-1.4	-1.2	-.8	-.3	.5	1.1
.1	.5	1.0	1.3	1.4	1.3	1.1	.6	.1	-.6	-1.3	-1.9	-2.3	-2.6	-2.7	-2.5	-2.1	-1.5	-.8	[illegible]

$$Z_{(X,Y)} = \text{SIN}(X+Y) + 2\text{SIN}(2Y) + 3\text{SIN}(3X)$$



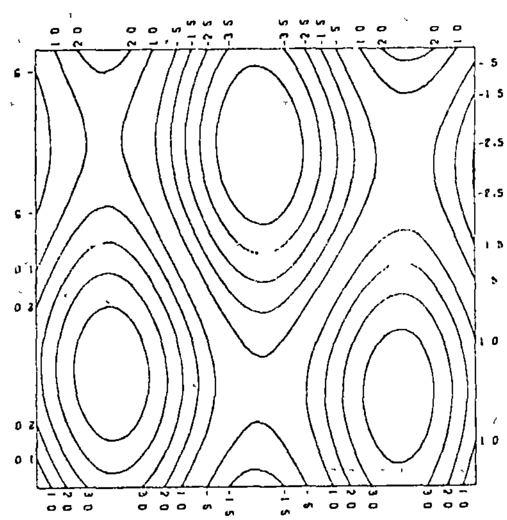
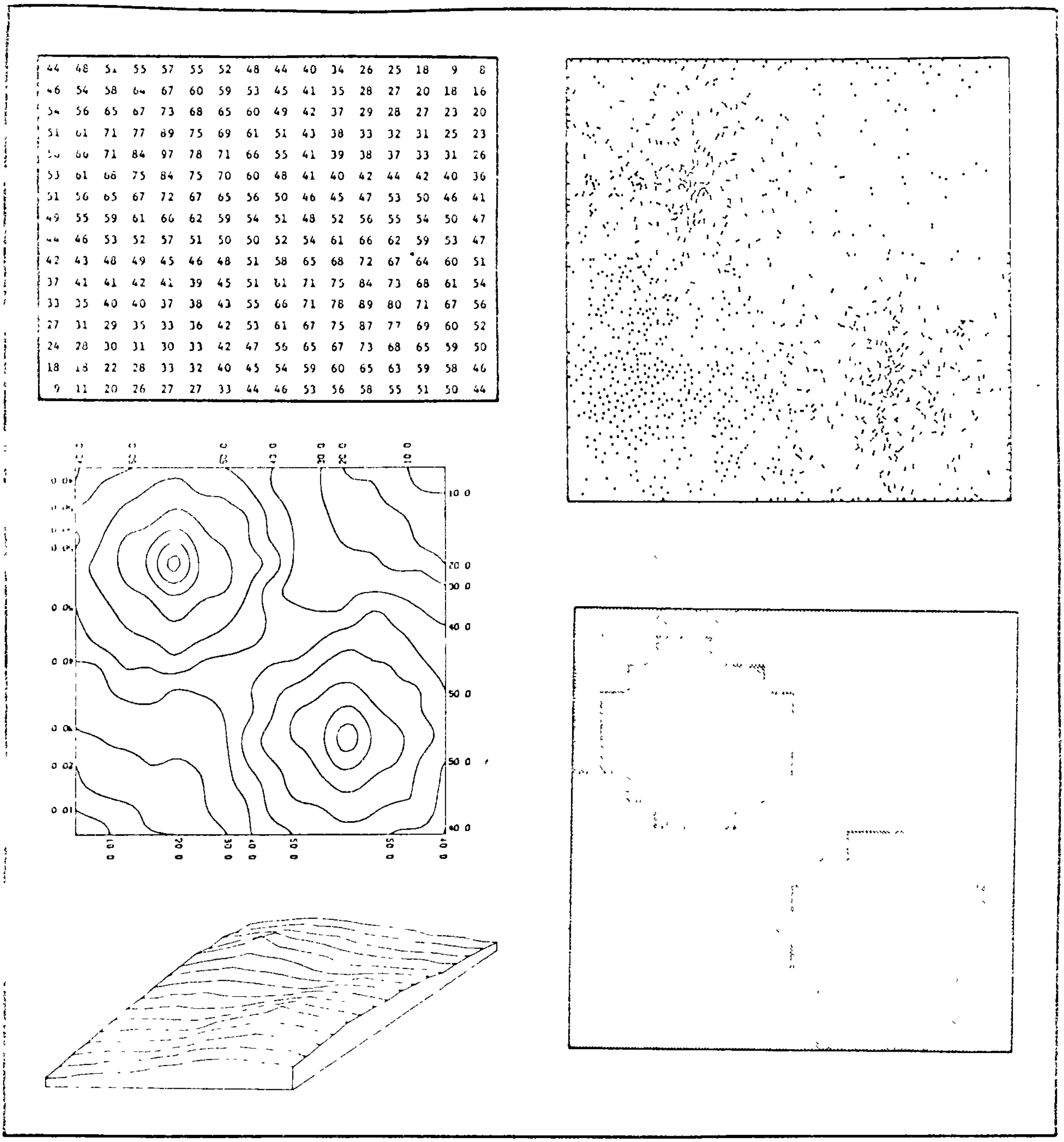


Figure 30. Visualizing structure in data tables, digital maps, functions, graphics



44	48	[illegible]	55	57	55	52	48	44	40	34	26	25	18	9	8
46	54	58	64	67	60	59	53	45	41	35	28	27	20	18	16
54	56	65	67	73	68	65	60	49	42	37	29	28	27	23	20
51	61	71	77	89	75	69	61	51	43	38	33	32	31	25	23
[illegible]	66	71	84	97	78	71	66	55	41	39	38	37	33	31	26
53	61	68	75	84	75	70	60	48	41	40	42	44	42	40	36
51	56	65	67	72	67	65	56	50	46	45	47	53	50	46	41
49	55	59	61	66	62	59	54	51	48	52	56	55	54	50	47
44	46	53	52	57	51	50	50	52	54	61	66	62	59	53	47
42	43	48	49	45	46	48	51	58	65	68	72	67	64	60	51
37	41	41	42	41	39	45	51	61	71	75	84	73	68	61	54
33	35	40	40	37	38	43	55	66	71	78	89	80	71	67	56
27	31	29	35	33	36	42	53	61	67	75	87	77	69	60	52
24	28	30	31	30	33	42	47	56	65	67	73	68	65	59	50
18	18	22	28	33	32	40	45	54	59	60	65	63	59	58	46
9	11	20	26	27	27	33	44	46	53	56	58	55	51	50	44

Figure 31 Possibilities for graphic interpretation the graphic model concept

schematic representation of map abstraction provided in Figure 32. In this illustration I have given the *map universe* or *population* (i.e., the set of all possible maps of all mappable phenomena) the name *Supermap.* The implication is that in a metaphysical sense there exists a single, all-encompassing map representing all possible individual map forms, types or varieties. Any single map constitutes a subset of the map universe characterized by a specific combination of content election, space/time/content abstraction, graphic techniques, cartographic parameters, and display media. This notion is clearly illustrated by considering the following written description of the map presented in Figure 33:

> Figure 33 is a contour map with a constant contour interval of 1,000 feet depicting Mt Rainier, Washington, at a ground scale of 1 250,000. Contour lines were drawn with a #00 Leroy pen on Dupont UC-4 drafting film No reduction.

For this example we have the following specifications:

Content selection	Area of highest elevation in Washington State, U.S.A
Space Abstraction	Area bounded by 46° 46′ and 47° 00′ N lat. 121° 33′ and 121° 55′ W. long.
Time Abstraction	Map data 1963 (data collected 1959)
Content Abstraction	Rough impression of mountain's form and position desired
Graphic Techniques	Contour mapping
Cartographic parameters	1,000′ contour interval (constant) mean sea level base scale 1:250,000
Display media	Cronaflex (UC-4) drafting film Leroy #00 pen Tusche T (ink) No reduction

A schematic diagram of the abstraction process in the above example is provided in Figure 34.

Individual maps are conceptually abstracted from the map universe in much the same way that humans perceive order out of seeming chaos The "whole" is but noisy confusion until focus on an individual element causes that particular item to emerge clear and the other elements to fade into the background. This is precisely what we do in conceiving "unique" maps. The important point, of course, is that we do not forget that the background always exists.

The map/model notion strongly suggests that if the passive approach to map composition which is so prevalent could be transformed into a positive search for understanding, it could potentially lead to the development and fuller utilization of cartographic processing. Encouraging

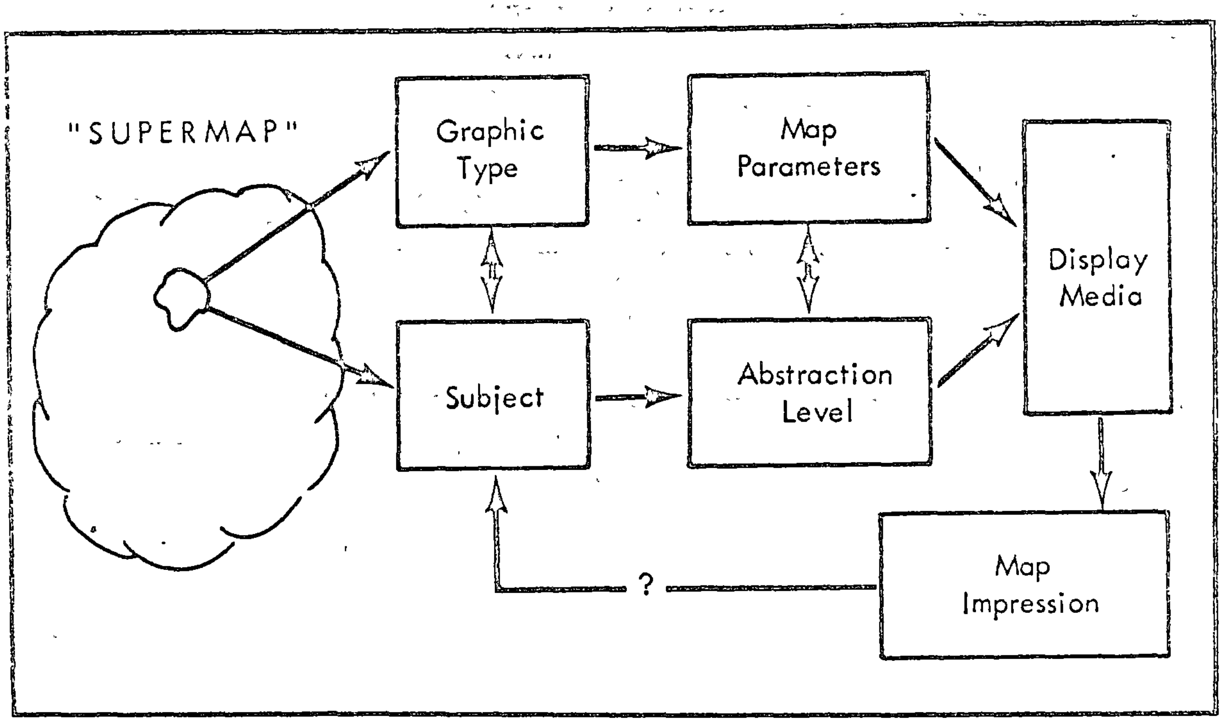


Figure 32 A schematization of the map abstraction process

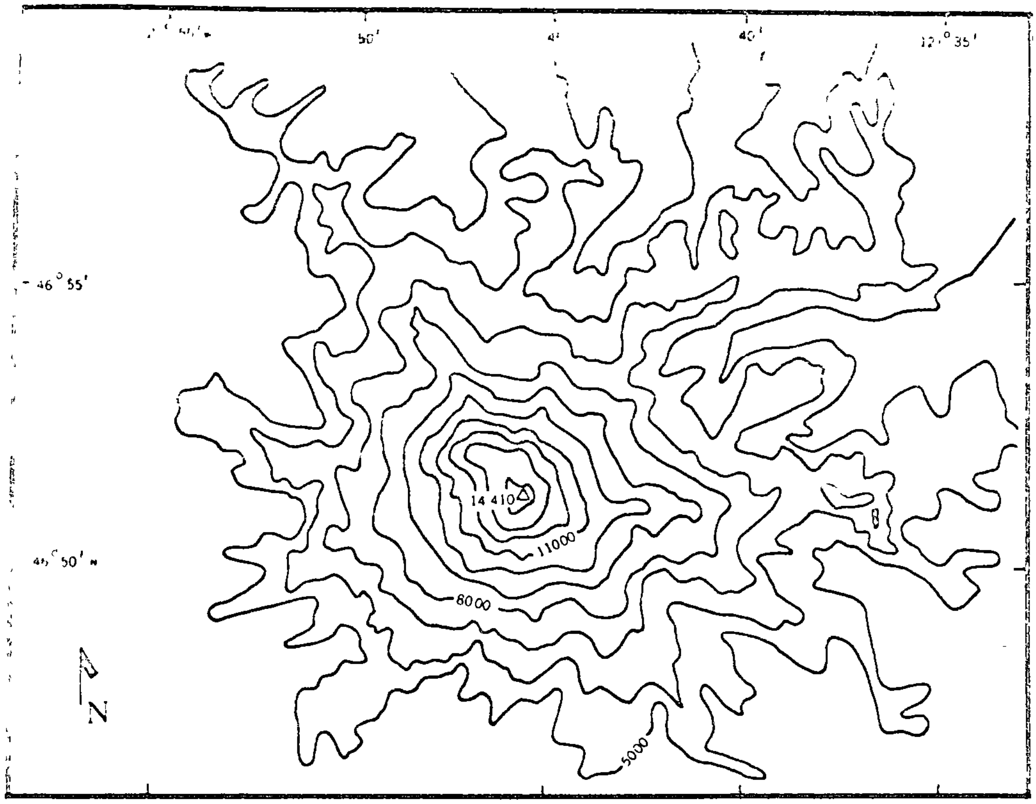


Figure 33 Contour map of Mt Rainier, Washington

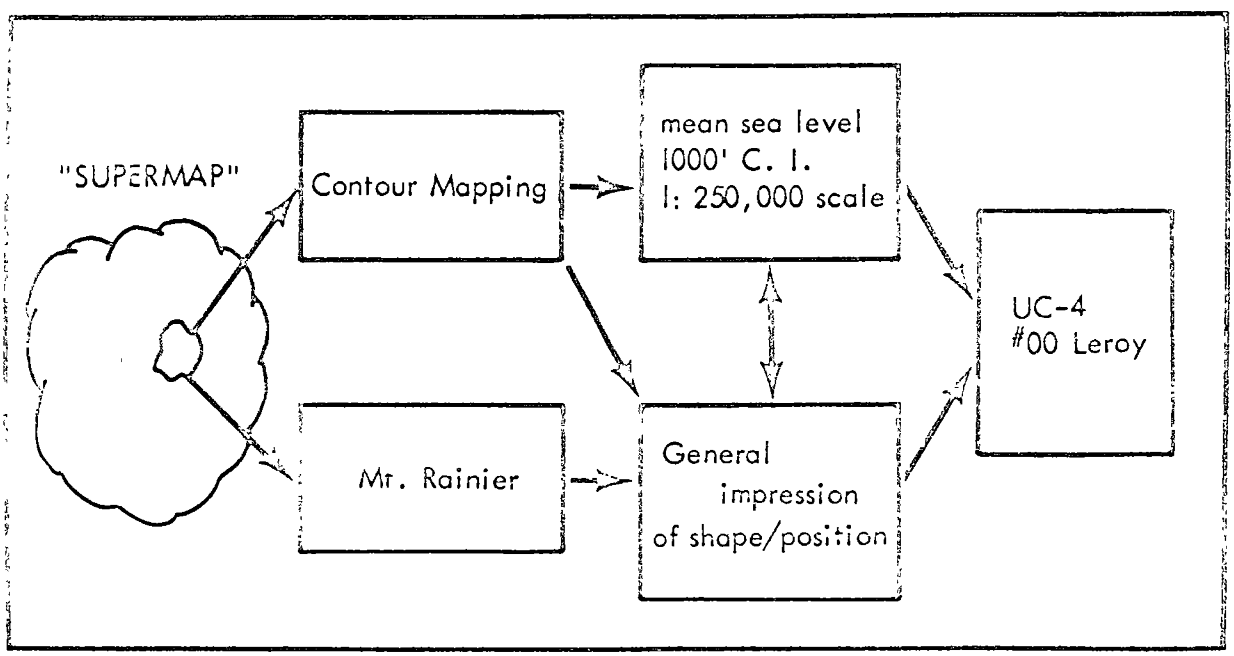


Figure 34 Cartographic abstraction the Mt Rainier example

progress is being made here. Intuitive cartographic decision making procedures are slowly being displaced by more formal (empirical) guidelines established through experimentation with various map parameter combinations. Since better understanding of how the manipulation of these map parameters influences map effectiveness is crucial in our quest for improving cartographic data processing, at least brief consideration of the cartographic research involved seems warranted. The major proportion of this research has involved laborious manual manipulation of experimental designs, although automation now promises to alleviate much of the former drudgery. Only a few of the more significant findings can be reported here.

The geometrical and mathematical principles of *map projections* probably represent the most obvious and best understood model characteristic of maps. Although there is theoretically no limit to the possible number of distinct projections, in practice only a relatively small number of rather different projection "types" are in common use. The conflict between area, shape, distance, and direction preserving properties makes it immediately clear that different but equally "correct" projections can produce images of the same phenomenon that contrast greatly in appearance and usefulness. The conclusion is that the choice of "best" projection is only meaningful in terms of the desired purpose of the cartographic model. Considering the absence of rigorous analysis, current guidelines for selecting proper projections are likely to be based on information which, in some cases, is more illusionary than factual.

For obvious conceptual and computational reasons, maps are traditionally formulated within the constraints of Euclidean geometry. Although the assumption of a Euclidean basis for spatial patterns suffices in the study of many physical phenomena, there is little theoretical basis for this choice of geometry in the study of social/cultural space relations. Doubt surrounding the validity of the Euclidean assumption raises serious questions about the basic usefulness of cartographic analysis in spatial process research. For example, can spatial patterns generated by behavioral processes which are markedly non-Euclidean be analyzed effectively with maps which are by convention based on an irrelevant Euclidean space? The essential problem will become ever more important as cartographers progressively abandon their landform tradition.

Deliberate use of projection distortion attributes to facilitate the study of functional geographical space relationships has led to several novel map transformations. Although it is conventional for maps to be scaled in measures of physical distance (e.g., miles) many geographical problems are scaled in something other than metric space. Two types of map transformations which contain deliberate distortion have been created to simplify the solution to this problem. For certain purposes, *functional distance* measured along established routes, or between known points, may be the primary concern of the map user. For example, a person interested in the time it would take to travel between two points on a standard map must first calculate the mileage and then divide by the estimated average rate of travel. Obviously this is inefficient, since it is conceptually simple to replace physical distance in a network by functional distance scaled in terms of time, effort, cost, or similar units (Figure 35). In fact, an increasing demand for these non-metric map scales in recent years has posed new and difficult problems for cartographers. There are severe constraints on desired distance transformations, particularly when more than just topological relations must be preserved. The most serious limitation undoubtedly is the inability to map non-Euclidean space by conventional methods without losing valuable space relationships. For example, although topological relations are generally preserved, most time-distance maps are only valid when "distances" are read out from one center point. Although such constraints are particularly annoying in light of developing needs they are of course not new map projection problems.

A second form of map transformation which is gaining in popularity involves the deliberate distortion of area. Again the technique is conceptually simple. Conventional geographical (i.e., physical) area is simply replaced by some areal quality (phenomena). The geometry of geographical space may be modified by discrete cells, as in most familiar cartograms (Figure 36). An attempt is generally made to preserve the shape and contiguity relations of areal units while working within the size and shape format of the original map. Even with these restraints there is still an infinite number of solutions to the problem. The aim of cartograms is to produce visualization with impact. The technique is especially useful where great inequalities in the distribution of a phenomenon exist from one region to the next. Cartograms are generally considered to be "expressive," but not particularly academic map renditions, due largely to the uncontrolled nature of their construction.

Geographical theory often includes a "uniform plane" assumption, in spite of the fact that a homogeneous distribution is seldom, if ever, encountered over large enough areas in reality to provide suitable empirical tests of spatial models. This inability of testing theoretical models empirically has led to the systematic creation of uniform density regions through controlled mathematical distortion of distance units in a continuous fashion (Figure 37). Such modification of the geometry of geographical space, in order to smooth the variation in a spatial distribution, is particularly appropriate because once models have been

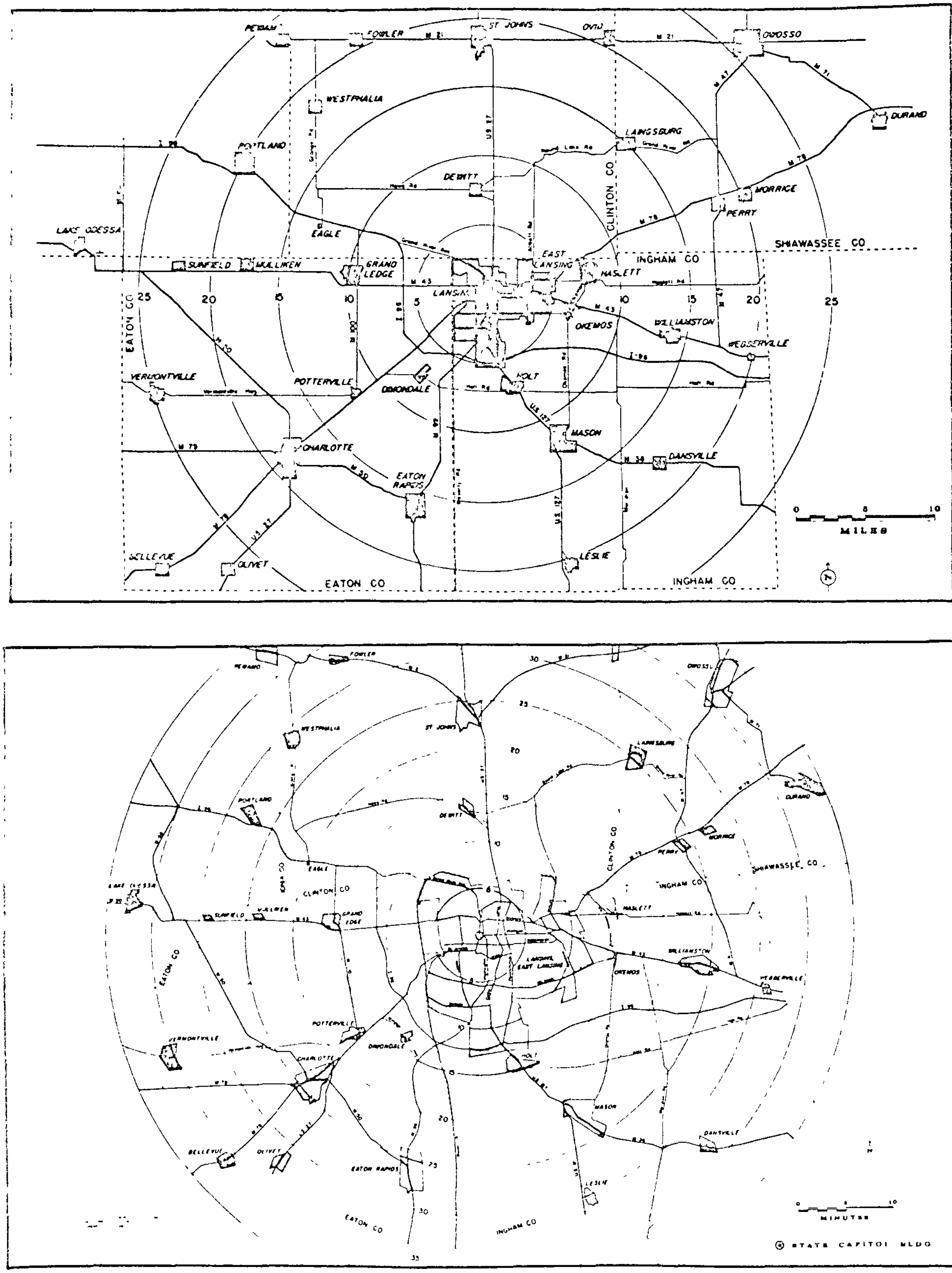


Figure 35 A time-distance transformation of the Lansing, Michigan, area (after Blome, 1963)

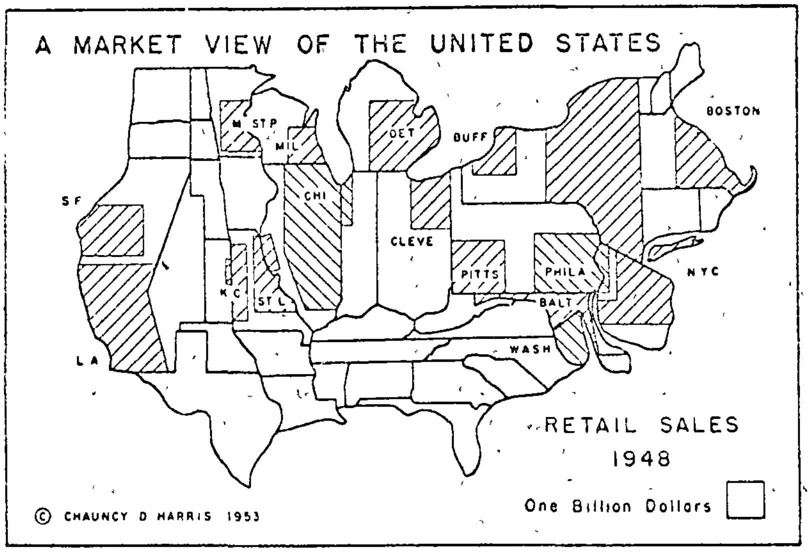


Figure 36 Cartogram areas of cities and states shown in proportion to their retail sales (from Harris, 1954)

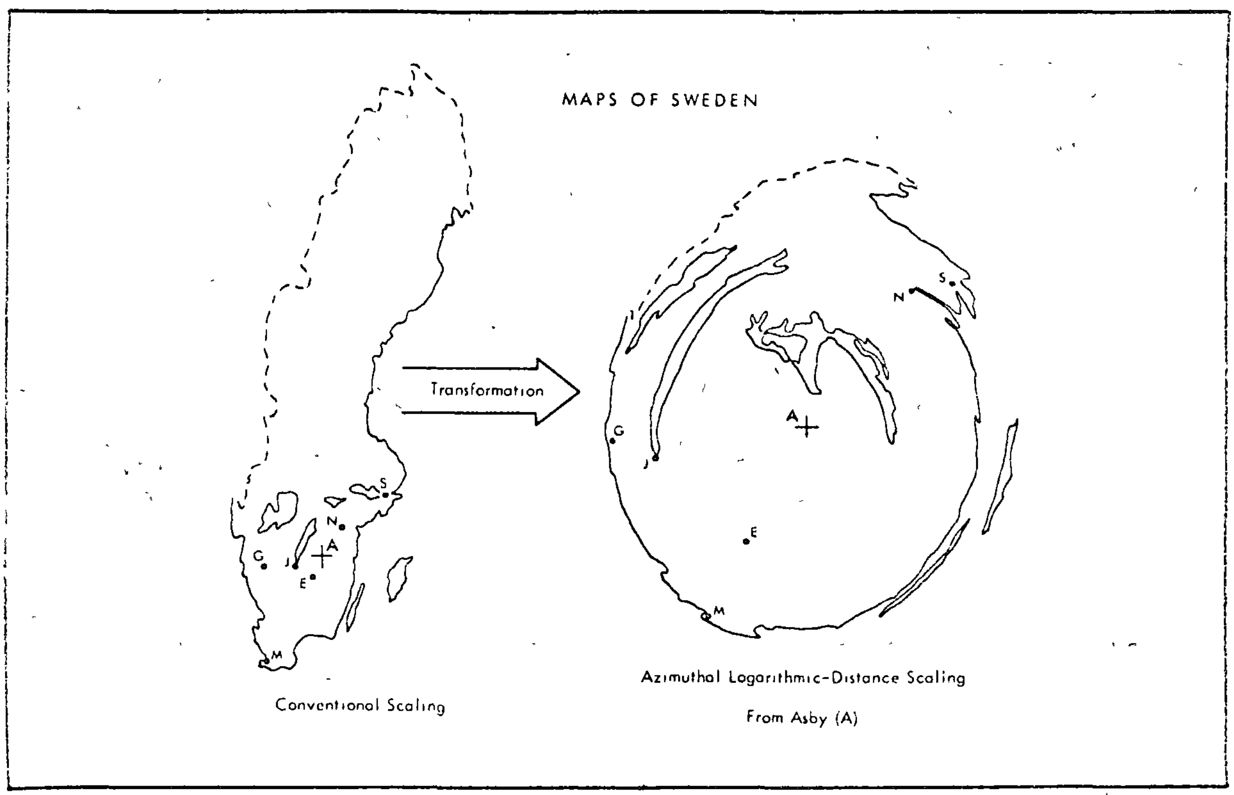


Figure 37. Continuous azimuthal-logarithmic distance transformation of Sweden (after Hagerstrand, 1954)

evaluated, an inverse transformation restores the original geographical background

Unfortunately, cartographers have serious technical problems to resolve before this powerful technique becomes generally operational. And, whatever the form of the transformation, problems arise first in specifying meaningful transformations and secondly in preserving desired properties such as boundary, shape and continuity relations. The search for permissible and meaningful transformations, the degree of control attainable with each, and the reversibility of the transformation processes provide important topics for future research.

Common sense suggests that it is possible to interpret geographical information in a hierarchy of *generalization levels* Stafford Beer, well known in management science, has used the phrase *"cones of resolution"* in reference to this hierarchical notion (Figure 38). The goal in map generalization is to achieve data and message compression while maintaining adequate map quality. In map construction, point, line, or areal symbols are used to reference earth features which are inherently zero, one, two or three dimensional. The actual coordination is arbitrary. What appears on a map as a point could well have areal extent in reality, and what appears as a line on a map at one scale could be depicted as an area on a map at another scale. Symbol/referent coordinations appear functionally related to the purpose of mapping and the scale of consideration. This non-theoretical basis for symbolization surely helps explain the fact that the actual generalization of map symbols continues to present one of the most intriguing and enigmatic cartographic problems The problem goes well beyond simple map enlargement and reduction due to the scale dependencies in map construction Map drafting that "looks good" at one scale seldom suffices at other scales, especially when these other scales differ significantly from the original. This point is easily illustrated if darkroom equipment is available.

The motivation underlying map generalization and the actual techniques employed vary to the extent that quite different end products can result using the same initial data.

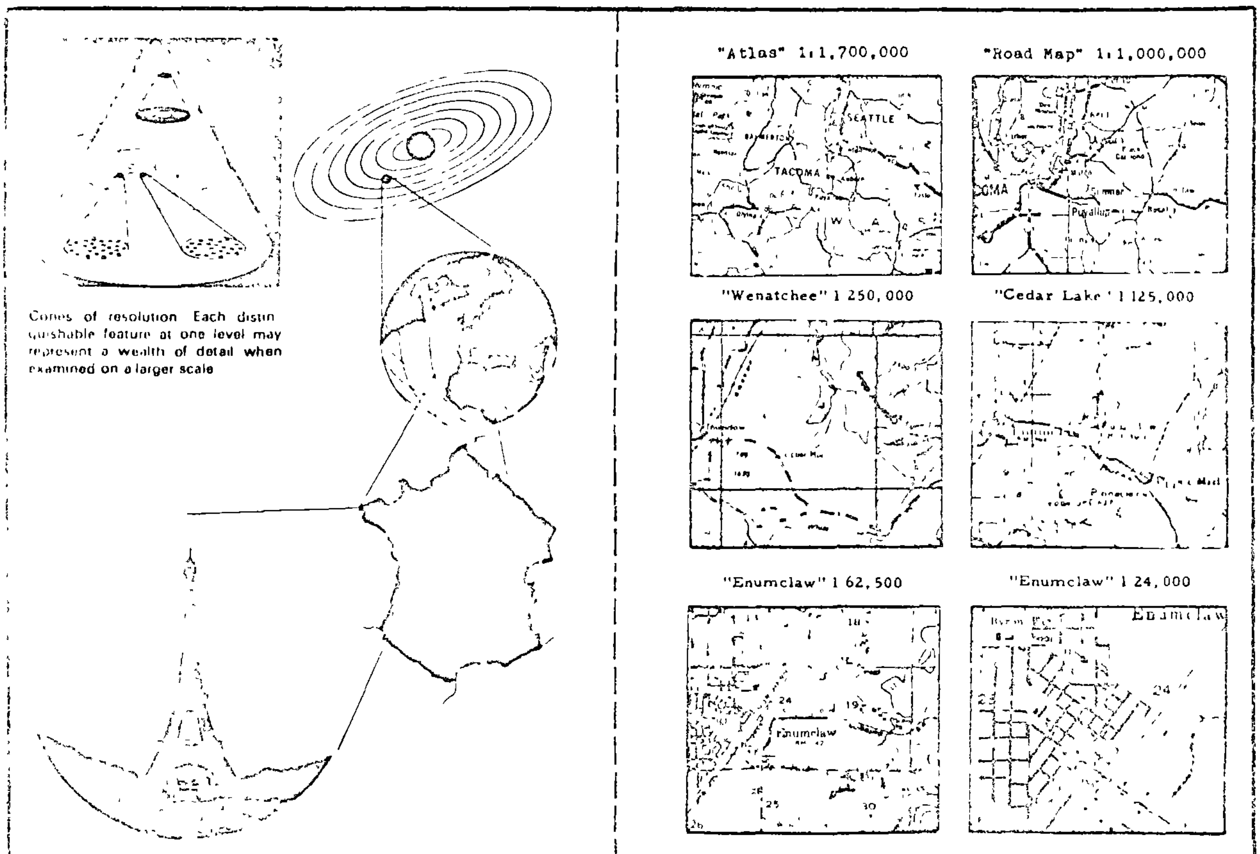


Figure 38 The similarity between levels of map generalization and Stafford Beer's "cones of resolution" (after Beer, 1968)

Previous research has primarily treated the inverse relationship between degree of generalization and map scale, with the goal of establishing guidelines which could be used to establish analytical map generalization functions or to subdue the wide variation inherent in subjective and personal generalization styles. Researchers have tended to focus on one-dimensional (line feature) generalizations as related to coastlines, rivers, transportation networks, contours, et cetera, using good and bad examples from existing maps. A second class of research is analytical map generalization intended to aid map interpretation by reducing complex map patterns to meaningful component configurations (e g , trends or residuals) The bulk of analytical generalization has dealt with the problem of graphical profile and contour smoothing

Recent experiments in map generalization have been facilitated by electronic data processing (EDP) and have often gotten their impetus from the needs of automated mapping. The approach has shifted from direct graphic line generalization to numerical line generalization and three-dimensional "surface" generalization (Figure 39) Augmenting graphical generalization with an analytical approach should not only clarify what is meant by simplifying a map but should also suggest valuable new techniques Those looking for research topics should note that attempts to fully automate cartographic processing have all failed to some degree for lack of objective map generalization guidelines

The way toward major advances in cartographic analysis was opened recently when cartographers realized that map generalization was simply a form of spatial frequency filtering (see Preprocessing above), a technique which has proven valuable in such fields as communication theory and picture processing by computer The process of removing particular structural details from the data through controlled filtering constitutes a formal selective-emphasis mechanism of great promise for improving over-all map interpretation and comprehension Rather than associate map generalization exclusively with scale changes, the principles are employed to filter map information in accordance with map use This notion carries initial selection from reality of phenomena to be mapped (primary generalization) through a secondary generalization process in order to illuminate specific attributes of a phenomenon In this sense it is merely a logical extension of common data classification procedures to a level where spatial frequency is explicitly considered

The nature of *proportional cartographic symbols* has been studied quite thoroughly as to form or geometry positioning, scaling, and graphic enhancement (see Ekman, Lindman, and William-Olsson 1961) Although improvements in each of these aspects is continually sought, the bulk of research entails the search for new symbols which can be more effectively positioned, scaled, enhanced, or coordinated with their referents. Current efforts to develop symbols which more effectively portray environmental pollution fall within this class of research Since actual evaluation of most point emphasizing symbols has been essentially a matter of judged effectiveness, further discussion of this topic is more appropriately treated under Image Processing below.

Researchers have largely neglected premap decisions relating to the selection of *dot map parameters* It is known, however, that the accuracy and effectiveness of dot maps are functionally related to spatial distribution characteristics, dot size, dot number (unit value/dot), visual tone density, geometry (size, shape, arrangement) of data collection units, and the placing of dots (see Dahlberg, 1967). The failure to develop flexible, automated dot mapping routines testifies in part to the lack of rigor in our understanding of the subject of dot mapping In fact, the development of such a computer program would make (i e , require as a basis) a significant contribution to research in this area as well as provide a valuable contribution to automated cartography in general

Premap decisions associated with the construction of *three-dimensional maps* can be isolated and grouped into those involving structured form of presentation, viewing point, vertical exaggeration, and graphic construction techniques The selection of a *perspective* (e g , parallel, angular, oblique) or *non-perspective* (e g , isometric) presentation is still based more on convenience than on rigor We do have some useful guidelines for establishing viewing points in terms of azimuth and angle above the horizon, and for determining approximate degrees of vertical exaggeration needed in different mapping situations in order to produce acceptable (i e , realistic) three-dimensional presentations (see Jenks and Steinke, 1971) But the proper choice of graphic construction or surface enhancement technique has proven difficult to formalize The problem appears to break down into one of weighing such factors as ease of construction (in terms of available facilities), scale and surface characteristics, and the degree of realism (quality of the depth illusion) desired The great diversity of technical options alluded to above not only establishes the potential of three-dimensional mapping but also raises a large number of interpretative problems which remain unsolved

An extensive literature has developed on *quantitative areal maps* which depict quantities by lines of equal value called isarithms (i e , isometric lines or isopleths) or by average value per unit area as illustrated by choropleths Hsu and Robinson (1970), Jenks and Caspall (1971), and Morrison (1971) are particularly good basic references in

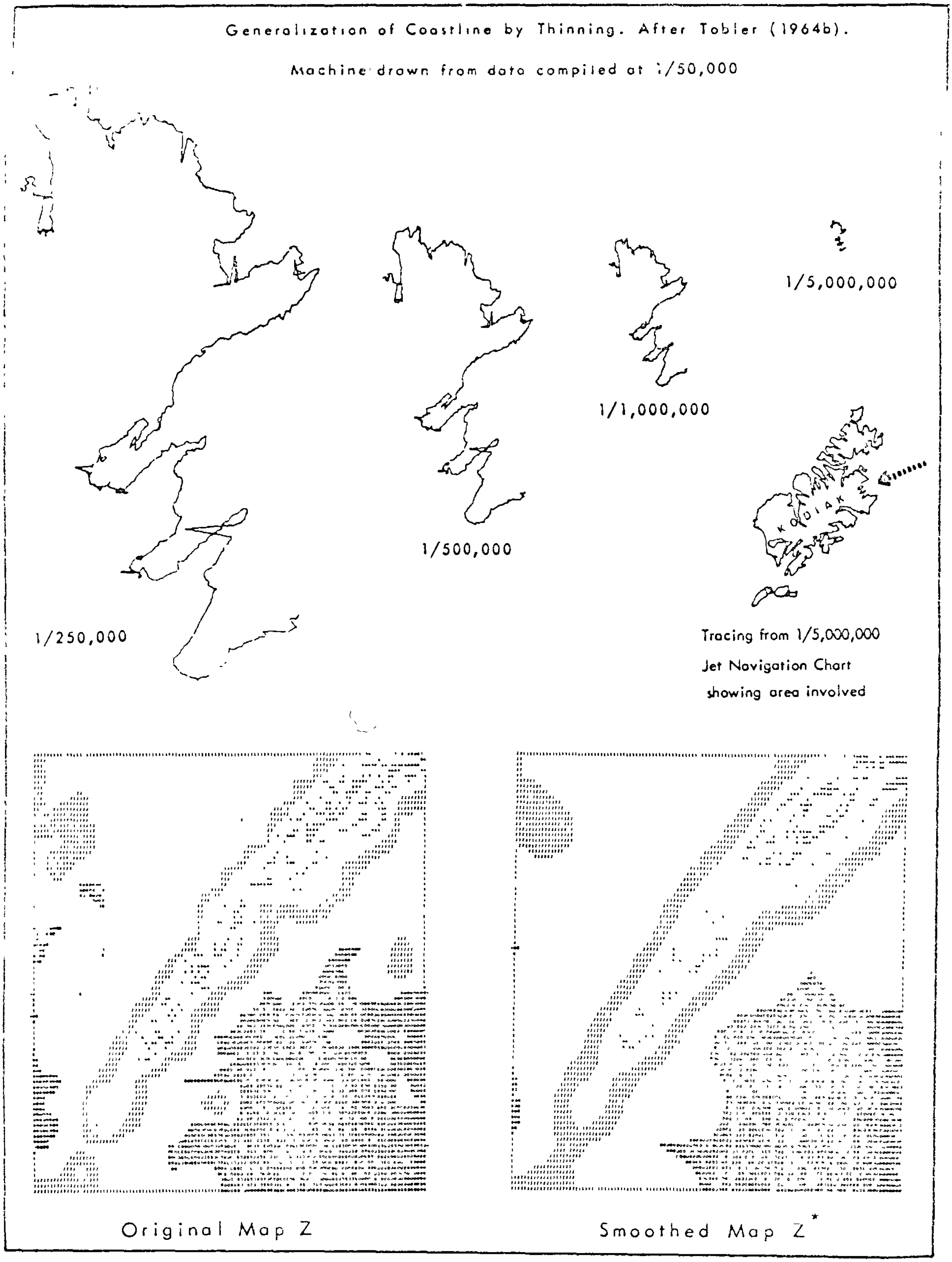


Figure 39 Numerical map generalization of line and surface features (from Tobler 1966)

this area Research has revealed a complex inter-relationship between such factors as (a) the statistical nature of the distribution in relation to the geometry of statistical units (in the case of choropleths) or the number, spacing, and location of control points in conjunction with the interpolation scheme and reference area employed in the case of isarithms, (b) the number of data classes and the criterion for the selection of class limits, (c) map scale, and (d) the output graphic technique, including the selection of screen tones and pattern The full combined influence of these factors on final map appearance is still not understood to the point where premap decision making relating to quantitative areal maps can be formalized, however Practical (i e, operational) guidelines for parameter selection that would guarantee minimum discrepancy between original data and map image are neither exhaustive nor definitive

The use of color in cartography provides an excellent example of the possibilities for variation with a single map element In a purely technical sense, the color on a map reflects a mix of reproduction parameters, including printing inks, papers, and copy screens (dot or line-screens) The result--color--is at least tridimensional. Physicists measure the physical properties of light in terms of dominant wavelength, percent reflectance or brightness, and saturation (purity or chroma) These three attributes of color relate, respectively, to what Professor Robinson (1967) has referred to in the cartographic literature as hue, value and intensity—the so-called psychological dimensions of color. These visual dimensions of color in light of map applications may be scrutinized even further, since the perceptual aspects of each attribute are additionally affected by the physiological characteristics of the perceptual mechanism, the connotative and subjective biases of intellect, and cartographic convention or symbolism Thus we have taken what might appear to be a simple element of map making and shown that the use of color is in fact, subject to a complicated mix of technical and perceptual variations, each of which may drastically alter map interpretation The growing use of color in map making represents a particularly difficult problem, because research on the subject of color applications is scarce in spite of the fact that we seem to have long identified the cartographic parameters of color in at least a theoretical sense.

It should be clear in light of previous discussion that any set of geographical data is subject to many graphical forms of presentation Even if considered in isolation, which is in no way a realistic form of evaluation, none of the above outlined decision making procedures relating to proper choice of map parameters is so static, straightforward, and clear that it does not warrant the scrutiny of future research When it comes to the more important "gestalt" effects of the various decisions all considered together in a final map presentation, our ignorance is staggering Before map data, map elements (symbols), and map users can be functionally and most effectively integrated into the geographic information system, it will be necessary to satisfy the need for research in which the simultaneous interaction of all mapping variables is determined. Our limited understanding of the consequences of parameter changes in association with previously discussed map forms can be useful in providing background information and as a beginning point for future researchers. The notable point is that in the past a tremendous amount of time and labor went into the compilation of a single map interpretation of a distribution. The cartographer's first judgment was assumed to be sound, since it was very unlikely that he would concern himself with compiling additional maps of the same information so that the most suitable graphic could later be selected. Experimentation involving the compilation of large numbers of maps was severely limited for the same reason.

Automation and other new time and labor saving devices are radically changing this situation, since it takes little effort to produce almost any map version of a distribution with equal facility once the data and mapping techniques have been computerized A small number of cartographers already have facilities and mapping routines available which permit them to produce "ideal" maps almost effortlessly through iterative, trial and error mappings with computers Eventually the cartographer will be able to sit before Sutherland's (1970) interactive display device and manipulate map parameters at will in a near real-time environment

Map guidelines can never be more than empirical rules for suggesting what would probably be the best graphical presentation in each different "typical" mapping situation. In an automated mapping environment, however, these guidelines could serve not as final decision criteria but rather as appropriate starting points in an iterative search for the best (or ideal) solution to specific mapping problems. One can actually visualize an environment in which a display screen serves as a dynamic worksheet upon which projection, scale, choice of symbols, degree of generalization, and other map parameters all become dynamic variables to be manipulated and controlled by the cartographic operator When a suitable map image has been derived, it may be projected into a microfilm reader/printer or similar device as separation images for color processing. Computerized cartography also will make it possible to have the statistical effectiveness of various data manipulation procedures simultaneously available at each output step If a statistical best fit solution is the mapping goal, then a simple iterative procedure at an interactive computer console will produce this result.

Clearly, the above discussion does not exhaust all possible instances in which the model nature of maps explicitly enters into cartographic decision making and subsequent map effectiveness. But, without further belaboring the point, it should be apparent that the temporal and spatial scale of geographical phenomena, the complexity of spatial processes and their interrelationships, and the nature of the mapping process all interact to restrict severely the possiblities for constructing "adequate" cartographical models. Maps are only approximations, and, for many research purposes, it is essential that they be constructed and presented in such a manner that their validity might be tested or their accuracy and representativeness known. In contrast, maps are traditionally presented in a deterministic way, as if they fully represent the spatial phenomena depicted without error. Random fluctuations in pattern manifestations of spatial processes and random or systematic distortions of the data arising through cartographic processing are usually not treated explicitly. It has actually been suggested that many "rational" explanations and planning decisions may have been based on map patterns which were, in fact, substantially "noise" effects.

A *stochastic approach* to mapping in which the map user is provided with a statement of the confidence level associated with each map seems more appropriate. In practice, it is probably more realistic to deal with confidence levels associated with individual map elements. The example of contour lines has been used by Professor Tobler (1970) and is illustrative. Considering the numerous

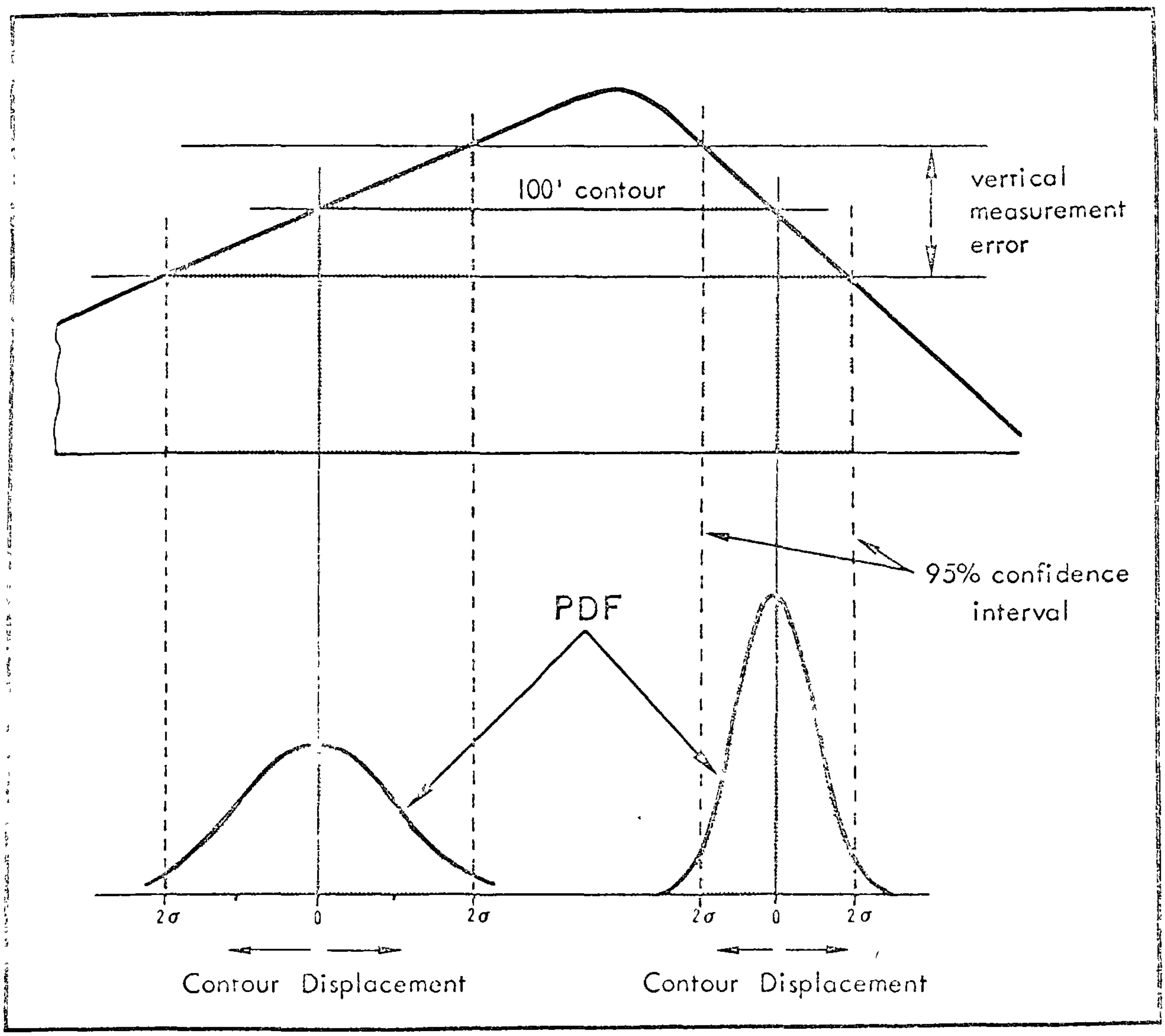


Figure 40 Uncertainty in geographical mapping the probability density function

errors which might accumulate in the process of transforming the elevation of landforms into a family of contour lines on a map, the chances that a single contour line is perfectly positioned (planimetrically) at any point on the map is probably quite small. But small positioning errors should be more common than large errors. Assuming the positioning errors are normally distributed about the true contour position (i.e., there is no uphill or downhill bias), Figure 40 provides a graphic interpretation of the last statement. The vertical component of the probability density function (PDF) provides a measure of the probability of obtaining a positioning error as large as "X" map (distance) units away from the true contour position. Obviously, the thin, solid lines that represent contours on most maps are somewhat misleading! Tobler suggests correcting this by showing contours as fuzzy lines, particularly in those situations where the *standard error* is large in proportion to the scaled width of the line on the map. But what about variable map accuracy? If we also assume that our contour positioning technique produces a constant vertical (elevation) error regardless of landform slope, we can see from Figure 40 that the contour line will be more accurately positioned in the region of steepest slope than in the lowslope areas. To continue with Tobler's solution, we now have a map with thin, solid lines in the areas of steep slope and fuzzy, wide lines in the regions of gentle slope. An alternate approach to the above problem is only to publish maps on which scaled line weights are greater than their probability density functions. This would mean that the less accurate the map, the smaller the maximum allowable scale would be. The main point to be made by this contour example is that all map elements are subject to certain errors which might be studied using a stochastic approach. We should be particularly interested in those cases where it is possible to compute levels of confidence for the map as a unit. The nature of use of graphics in scientific argument suggests that the acceptable level of confidence for maps used inferentially could be somewhat below the standard statistical 1 or 5 percent levels, although maps used for testing or illustrative purposes should probably (ideally) meet these more severe constraints. We must conclude that the use of maps as model representations of reality raises numerous questions of inference and control which warrant discussion and resolution.

So far in this paper the map effectiveness problem has been treated as one of mechanics, and the research topics outlined represent a concern for establishing a closer fit, or correlation, between original data and their cartographic depictions. It seems certain that if this "graphic model" treatment of maps is pursued vigorously by future researchers, it will lead to a wealth of interesting new ideas, useful cartographic methods, and important generalizations. Map effectiveness cannot be properly evaluated, however, without considering the nature of the information retrieval involved in map reading. At this point it is appropriate to consider map reading in a more explicit fashion.

IV. IMAGE PROCESSING

The observed variation pattern in spatial data is often extremely complicated. Geographical maps have traditionally served the need for visualization in attempts at description and explanation of spatial patterns. A logical goal of the *information display* (mapping) process is to produce, as efficiently as possible, the most effective graphic communicator of distributional information. *Image processing* entails the assimilation, manipulation, and analysis of information which is given in spatial pattern form such as a map. I have previously mentioned how information being processed can be distorted by the tools, materials, and techniques of the map compiler as it is converted into graphic form. It should also be apparent that, due to inherent or acquired limitations, the map user further distorts the flow of information in the process of *map reading*.

The lack of correlation between the intended map message and actual retrieved information constitutes an important communications gap. A major portion of recent cartographic research is intended to yield information that may be used to minimize this discrepancy. The research objective can be gained by either training people or machines to be more skillful map readers, or by designing map symbols that can be interpreted more effectively. To be successful both approaches require details of map reading far beyond existing knowledge of the factors controlling the interpretation of map patterns. Since map reading can at least involve visual, statistical, optical or tactual procedures, it is convenient to discuss each of these in turn.

(1) Visual Analysis

Making information visible by displaying it on maps has satisfied the human user in the cartographic processing system. Even though we suspect that the noisiest transform in the cartographic system is the map-to-man communication link, little research has been done to alleviate the problem. Cartographers know relatively little about the traditional procedure of visual map reading beyond the fact that it is subjective and leads to qualitative map description and comparison. Although vague qualitative terms such as dense, complex, hill, plain, correlated, dispersed, et cetera, are useful generalizations for many purposes, they do not characterize their referent phenomena in an objective, scientific manner.

The visual technique has long been criticized for not lending itself to statements of confidence and for its low powers in replicating conclusions. Rigorous evaluation of the results of several individuals is thus tenuous, if not impossible. It is further asserted that visual map reading does not permit the recognition or differentiation of subtleties in configuration and rapidly deteriorates in quality as control gets sparse, pattern grows complex, and the number of variables increases. For these and other reasons we suspect that the results obtained from visual map reading are often imprecise, inconsistent, or more apparent than real. If these deficiencies were, in fact, real, they would distract significantly from the use of maps in scientific enquiry, especially if they went unrecognized.

Human beings are especially well designed to make use of graphic information, regardless of the drawbacks to visual map reading. In terms of the variety of messages on unexpected topics that the human mind easily and rapidly receives through visual techniques in exploring data, there is still no substitute. The visual perceptual mechanism appears particularly well equipped to process spatial patterns such as those found on maps. It is essential that cartographers abandon their characteristically passive role and commence a positive search for understanding which could lead to the development and fuller utilization of the powerful qualities of visual perception in the context of map reading. Improved map communication can come about through map composition which takes into consideration the strengths and weaknesses of the visual perceptual mechanism. The first step toward increasing the effectiveness of visual map reading is to understand and evaluate the interpretative process involved.

The inherent nonlinearity of human vision is possibly the most basic obstacle to quality map reading. Even relatively simple subjective map interpretation judgments may vary significantly from absolute value symbol scaling (Figure 41). For example, where *proportional symbols* are scaled by area or volume, the perception of their magnitude often deviates significantly above or below their absolute value scaling (see Ekman, Lindman and William-Olsson,

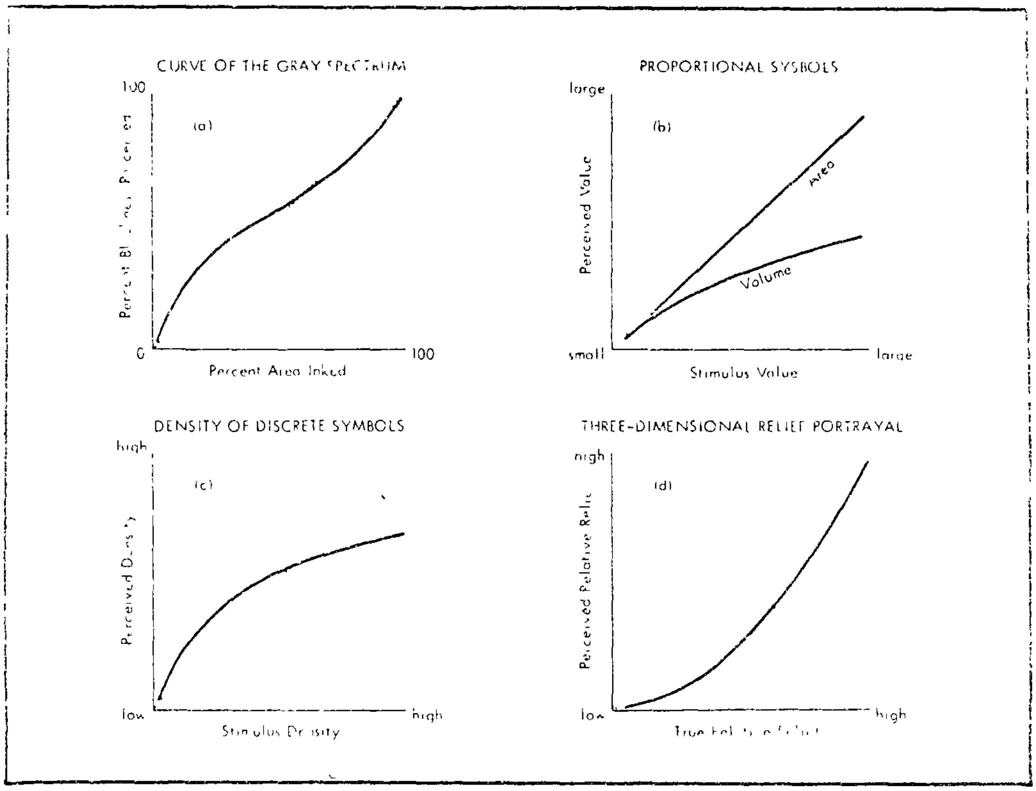


Figure 4.1 Nonlinearities in map reading absolute versus apparent value scaling

1961) Similarly, Jenks and Knos (1961) and others have demonstrated that the perceived area inked on shading screens approximates closely a *gray-scale curve* which is quite different from an actual area inked function Castner and Robinson (1969) have identified a number of dot pattern characteristics which modify the visual impression of area dot symbols Experiments by Robinson (1970) in scaling nonnumerical linear symbols indicate that line weights and forms are not perceived in true relative fashion In summary, some form of data predistortion is obviously required to make visual distortion negligible if a linear map reading relationship is to be closely approximated (see McCleary, 1970) Apparent value scaling has been suggested as an alternative to absolute value scaling, but there is substantial resistance to its general acceptance.

The literature relating to the extent of conceptual and perceptual error in thematic map reading is substantial and growing rapidly Cartographers have given some consideration to the legibility of map symbols in varying map contexts, symbol placement relative to other map elements, the external map reading environment, map user background developmental stages in map reading with age, the relative effectiveness of different map types, and additional and equally intriguing topics In spite of this effort, many of our more obvious techniques still lack research support

Take, for one example, *sectioned-pie symbols* which are so commonly used on maps to show spatial variation in the proportion of particular geographical phenomena (relative to select others) What do we know about pie-graph interpretation? Is shape of symbol significant? If proportional sections of different sized symbols are compared, do they appear proportional, and to what degree can their magnitudes be estimated as well? How does the orientation of a section influence its interpretation and comparison? What role does variable map context (surround) play? For these and many other questions we have few answers, in spite of the ready availability of appropriate psychological testing procedures which could be used to supply the needed information.

Two important notes of caution are stressed at this point First, the information transmitted by each symbol or symbol type is simplest to evaluate in a rigorously controlled, context-free environment, but these strict assumptions must be relaxed to include operational map situations as well In designing and evaluating any system of cartographic symbols, special care must be taken to insure that individual symbols, tested separately for convenience, will suffice functionally when brought together in a given map In the language of the systems designer, the symbols of the map must interface properly, and it is on this aspect of the problem that we are likely to expend much of our future effort. Secondly, only in the uninteresting case of uniformly spaced equal appearing symbols is it reasonable to suppose that all map elements would visually affect one another equivalently. In actual practice, symbol density varies over the map surface, exposing the potential for density dependent information retrieval qualities. The eye is also very selective. It is therefore interesting to speculate whether mutual visual interference between symbols grouped in dense clusters is any greater or less than between dispersed symbols, or whether visual weighting is biased in favor of isolated or grouped symbols.

The actual comparison of several maps is a natural extension of single map reading. Maps to be compared commonly exist as separate items, as superimposed patterns on a single map, or as transparent overlays. The relative effectiveness of these different arrangements has yet to be evaluated. The recent literature of picture processing by computer has created a respectful awareness of the powerful qualities of visual pattern recognition and matching. Cartographers must learn to exploit these human attributes further through more effective map design in order that patterns of greater complexity and exhibiting more subtle differentiations can be processed visually. Formalizing the intuitive and qualitative rotation-translation schema used by skilled map readers would also benefit cartographic training programs, enhance cartographic design, and facilitate the development of automated map reading devices.

Humans appear to exhibit much less facility in estimating the strength and nature of the association between mapped patterns. Although research in the area is still largely fragmented, it is evident that statistical models can be used to study the nature of visual map comparison. McCarty and Salisbury (1961) have used a simple *Pearsonian product moment correlation* model to evaluate visual comparison of isopleth maps, *Kendall's rank correlation* model has been used by Olson (1970) to study the effectiveness of comparing choropleth maps, and a *coefficient of determination* model has been used to evaluate the effectiveness of dot, perspective, and choropleth mapping when the goal is visual map comparison (see Muehrcke, 1969). The conclusion from these studies is that the ability to analyze maps visually varies from graphic to graphic as a function of user background, the data classifying method used, map complexity, map autocorrelation, and the degree of association between maps. Our understanding of map comparison is still far from being either exhaustive or conclusive, but it does appear certain that visual map reading efficiency can be improved vastly by simply specifying certain map reading tasks in conjunction with systematic selection of optimal map parameter combinations. Put another way, properly designed maps can probably increase significantly the complexity and amount of information that the map reader can assimilate visually. But there has not yet been sufficient research to provide rational decision criteria for selecting a particular technique as better than others.

In summary, cartographers know little about how the user goes about reading and responding to mapped information. The current literature of map reading represents only minimal exploration of the vast domain of cartographic psychophysics. Even the research findings that are available require substantial refinement and extension before cartographic methodology can realize the maximum promised benefit. Approaches may vary, but there is general agreement that improved definition of the actual map reading procedure holds considerable promise. The extensive literature on *visual search techniques* (Beller, 1970) and *eye movement parameters* (Yarbus, 1967) available in the psychological journals and the wealth of readily accessible information relating to problems of human factors engineering (see Sinaiko, 1961) need only to be interpreted and applied in the context of cartography. Several uses of the above techniques are immediately apparent. For example, a plot of typical map scanning progress over the map surface as a function of time should indicate the visual weight attached to various symbols and establish the sequence of attention flow between symbols and map zones. The resulting insight into "prime" map space and symbol influences might then be used to design maps that logically direct attention over the map from element to element in such a way that visual weighting corresponds with actual scaling importance. It may also be possible to reveal the extent to which size, shape, location, and density impressions of map symbols are influenced by systematic *field effects* and thus contribute to our understanding of distortions of the visual field (see Ericksson, 1967a and 1967b). Additional applications readily come to mind but cannot be pursued here.

A final note is that visual map analysis cannot be disassociated logically from spatial pattern scale. By alternately blurring or focusing the eye, it can be made to function as a low- or high-pass filter, respectively, but it is doubtful whether this flexibility extends invariably over the range of important spatial scales. It is certain that the human observer is unable to state formally the level of analysis subsumed by this built-in spatial filter and that visual pattern analysis may yield valid interpretations at one scale level but not at another. The implication is clear that even if visual map reading is correct, it is of limited analytical potential unless we can determine the scale of the visual filter(s) operating in each case. Statistical models explicitly incorporating spatial filter mechanisms may provide a way of specifying the scale level in visual map reading and might form the basis of much needed research in this area.

In an effort to minimize the problems associated with visual map interpretation, map readers have adopted quantitative forms of spatial analysis to supplement or substitute for visual analysis.

(2) Quantitative Map Analysis

The basis of objective map reading exists in a variety of descriptive quantitative measures of spatial pattern developed in diverse disciplines. Typical quantitatively oriented map analysis procedures involve a series of digital computer routines which might be conceived to function as map reading machines. Since all quantitative map analysis procedures require numerical data or mathematical formulas as input, they are most efficiently applied directly to unmapped data. Naturally one may question whether map reading is an appropriate term to use when unmapped data are being processed and interpreted. Those who find this terminology confusing might well refer back to earlier discussion of digital and numerical maps.

If numerical matrices and functional expressions are accepted as legitimate map forms, then quantitative processing of these data structures may be viewed as a special case of map reading. When original map data are not available in proper form, they often must be extracted from the map itself prior to quantitative processing. Several methods available for converting geometrical data on conventional maps into numerical form were mentioned previously. The inefficiency inherent in these and other map digitizing methods suggests that the procedure should be avoided whenever it is practical to obtain the original data directly. Earlier discussion of the case against cartometrics further supports this notion. Scaled symbol characteristics and their probability density functions must also be considered (see previous discussion). Suffice it to say here that the consequences of compounding sampling error and bias by converting maps to digits are poorly understood.

The ability to generalize from a numerical map using mathematical methods permits us to supplement traditional map descriptions and comparisons with objective measures of pattern. Quantitative techniques exhibiting a wide range of capabilities have now been borrowed and adapted, or are presently being adopted, specifically for map reading applications. Objective descriptive measures of spatial pattern primitives such as length, shape, intensity, extent, grain, position, gradient, orientation, connectivity, etc., have found extensive map reading use. Obviously, some of these measures are not independent of others. Spatial summary statistics, such as bivariate central tendency and dispersion, are simply computed. The utility of a graph theoretic approach to clarify the structural (topological) and geometrical properties of line networks on maps has been demonstrated. Numerical classification procedures such as grouping or cluster analysis are applicable in regionalization problems.

David Harvey 1968 has conveniently grouped a variety of additional mathematical methods for describing spatial patterns under two broad headings. He refers to those mathematical expressions which objectively describe or generalize about spatial patterns by decomposing them into a limited number of simpler (local, regional, global, random) components as *generalized mathematical measures*. The most important generalized measures are listed in order of increasing complexity as: (a) matrix representation and spatial filtering, (b) power series trend-surface analysis, (c) double Fourier analysis, and (d) two-dimensional spectral analysis. Although the theoretical assumptions underlying the above measures of pattern are not related explicitly to spatial processes, and therefore no specific process interpretation is implied by a good "fit", geophysicists, in particular, have demonstrated that such interpretations can be meaningful when soundly based in spatial process theory.

Harvey's second grouping of quantitative pattern measures, called *specific pattern representations*, includes methods which explicitly involve the comparison of a map pattern with a theoretical pattern derived from specific process assumptions. The most important methods in this grouping are all based on measures of departure from Poisson expectation associated with various random processes and include (a) quadrat sampling, (b) continuity measures, (c) nearest neighbor measures, and (d) linear sequent analysis and occupancy theory. An important question arises in deciding to what extent the hypothesized mathematical processes have geographical process interpretations. Further difficult inferential questions are associated with the scale problem when map patterns are analyzed with models involving hypothesized stochastic processes.

Quantitative map comparison is conveniently treated as two separate problems. Spatial *auto-covariance* or *auto-correlation* measures the degree of internal organization or dependence in a pattern. Constructing *correlograms* by plotting simple correlations between observations at different data intervals (lags) against the intervals themselves can identify scale factors in the pattern (Figure 42). Correlograms may be plotted for spatial transects, which may be roads, or for statistical surfaces. When independence of observation (i.e., auto-correlation not significantly greater than zero) underlies a mathematical expression of pattern, correlograms can quickly indicate whether this model

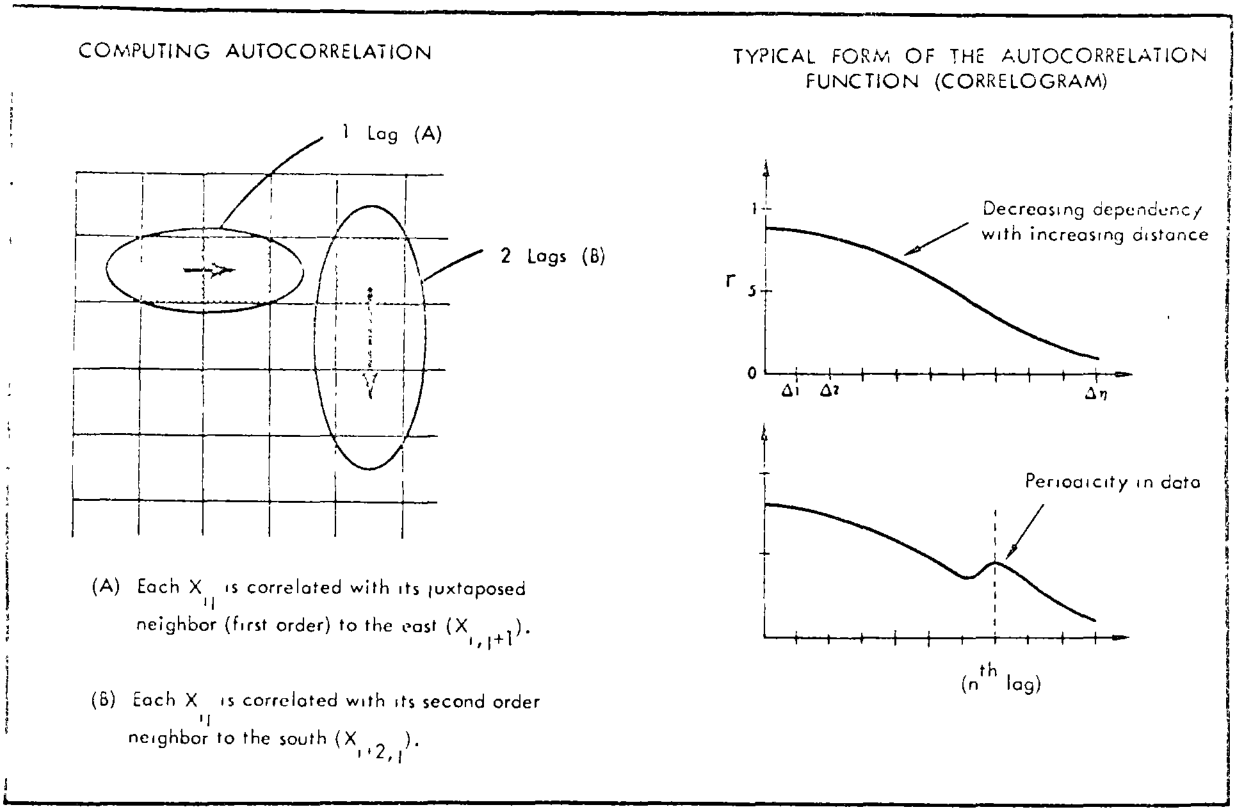


Figure 42 Correlogram analysis near things are more related (dependent) in geography

assumption is met at the desired scale of analysis or, conversely, suggest at what scale(s) of analysis the mathematical statement is meaningful. In addition, the Fourier transform of the auto-covariance function yields the power spectrum which identifies the relative contribution (i.e., importance) of different scale (frequency) bands to the total variance of the pattern.

In contrast to auto-correlation, where a pattern is essentially compared with itself, spatial *cross-correlation* involves simultaneous analysis of two (or more) different patterns. The selection of an appropriate cross-correlation technique from the large number available is primarily a problem of defining just what is to be compared. For example, comparison may involve simple summary statistics such as the mean, variance, etc. It is more common, however, to rely on simple correlation coefficients such as Pearson's *r*, Spearman's *rho*, or Kendall's *tau*, which provide average measures of correspondence between all respective data pairs. Since none of these measures takes the spatial variable into account, their utility in map comparison has often been challenged.

In order to deal explicitly with relative location, increasing emphasis is being placed on pattern decomposition into simpler spatial components and then comparing similar components. Suggested methods of *cross-trend analysis* include (i) computing simple correlations between approximated trend values or residual values, (ii) computing an average measure of the agreement between the direction and value of maximum local gradient at all data-points on paired surfaces, and (iii) computing simple correlations or taxonomic distances between trend-surface equation coefficients. Although the results of cross-trend analysis are promising, some obvious difficulties concerning use of the technique need further study. Of immediate future concern is formalization of gradient analysis and greater understanding of the relative advantages and disadvantages of using orthogonal and non-orthogonal polynomials, and subsequent selection and appropriate weighting of equation coefficients, with different samplings and surfaces.

Spatial pattern correlation is properly considered at a variety of scales or frequencies. Interestingly, a simple extension of spectral analysis, known as *cross-spectral analysis*, permits comparison of pairs of stationary spatial series (Figure 43). The relationship between the two series is measured as a function of (i) the square of the correlation (known as *coherence*) between corresponding frequency components, and (ii) the lag (or *phase-difference*) between identical frequencies. Plotting coherence against frequency

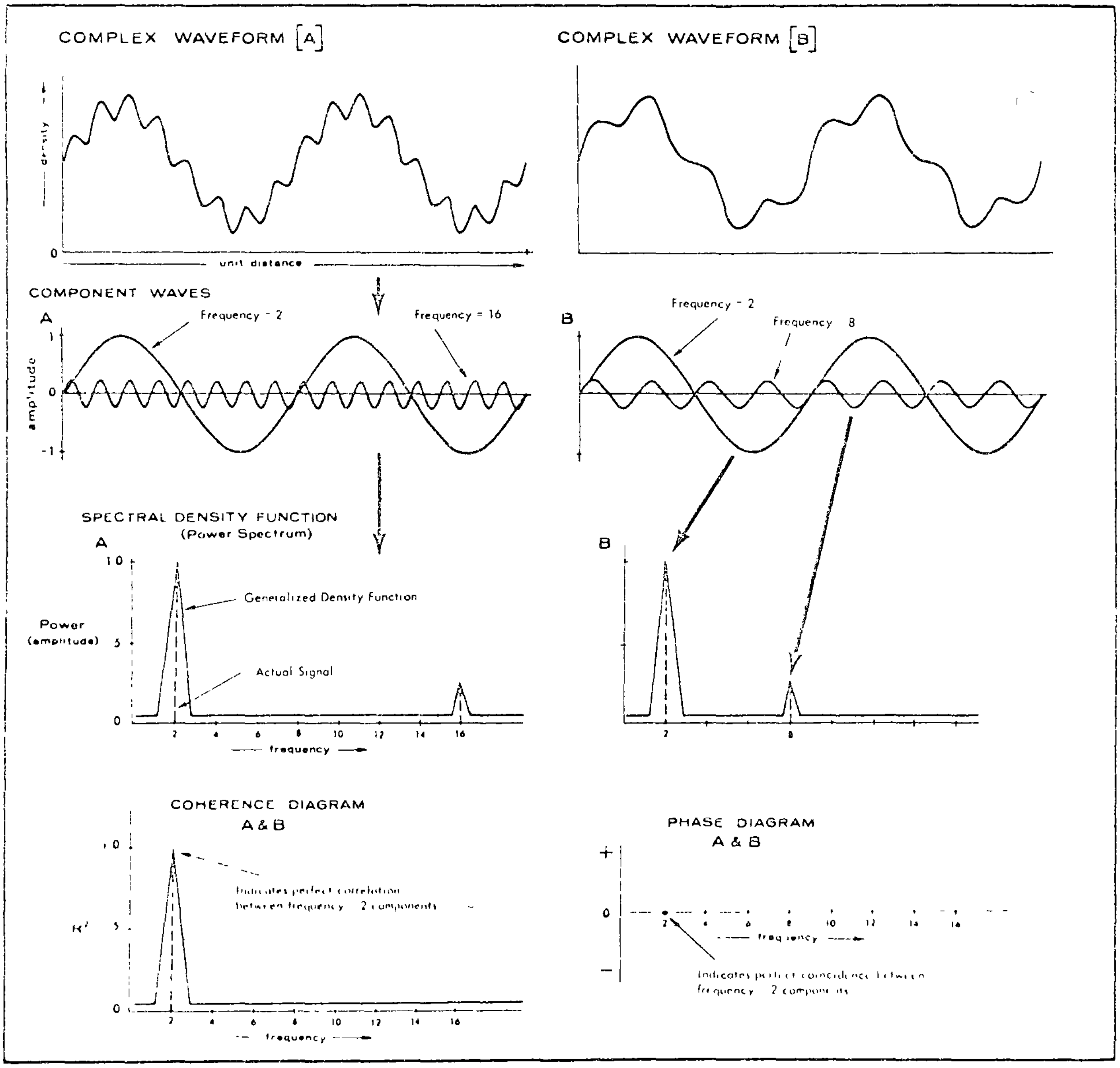


Figure 43 Cross-spectral analysis viewing the relationship between two geographical distributions in terms of the correlation between corresponding scale components

produces a *coherence-diagram* which reveals the pattern of correlations between corresponding components. Plotting phase-difference against frequency produces a *phase-diagram* which reveals lag relationships between identical frequency pairs. A critical research area relating to the use of cross-spectral analysis exists in exploring the reliability of coherence estimates with different spatial samplings (i.e., filterings) and distribution types.

Previous discussion of the *scale problem* established the importance of the scale at which data are collected and symbolized. The scale problem also occurs in the measurement and comparison of spatial patterns. Data extracted from maps as input for quantitative pattern analysis involve a secondary data collection procedure which propagates and compounds the influences associated with primary data collection. The point is that a degree of spatial filtering is inherent in each step of cartographic analysis. Comparing small scale maps is expected to produce somewhat different results from those obtained by comparing large scale maps of the same data. It should now be recognized that a small scale map of a statistical surface is in large degree comparable to a trend map where appropriate "residuals" have been removed. Any scale change is likely to modify the mathematical expression of map pattern. This fact is commonly illustrated by observing the widely different measures of departure from randomness that can be obtained by varying quadrat size. The consequence of this scale dependence is that quantitative generalizations and subsequent interpretations made at one scale do not necessarily hold at other scales, and the scale of analysis must be stated as part of the pattern measure. Researchers have responded to this problem by designing *scale-free* statistical measures of pattern or by using pattern measures such as nearest neighbor methods which minimize the influence of scale.

The scale problem, similar to the projection problem, is not all bad, however. In fact, a major advantage of quantitative map analysis is that it permits detection of scale factors in the pattern which may reflect particular processes. Successive areal aggregation and analysis, iterative quadrat sampling while varying cell size, autocovariance analysis, Fourier analysis and spectral analysis are all particularly sensitive to scale factors in the data. Quantitative analysis also permits matching the level of analysis to particular scales known to characterize given spatial processes. In a sense we have made creative use of the fact that quantitative pattern measures are not independent of the scale of measurement or observation. The similarity between the creative use of scale factors and premap data processing is obvious and is not pursued.

Obviously the comments made so far concerning quantitative map analysis only serve to introduce the concept. For greater depth and understanding, the reader should refer to a good source book such as Leslie King's *Statistical Analysis in Geography*. Unfortunately, the application of most statistical analyses in map interpretation presupposes satisfying one or more distributional assumptions, such as independence of observation, random sampling, normality of data, and linearity in relationships. These restrictive assumptions are seldom achieved in spatial analysis, which raises the question of just how valid the results of quantitative map processing are in specific applications. Although we are not able to specify the consequences of unmet assumptions at this time, it is safe to say that in most cases greater confidence is associated with the use of quantitative procedures in map description than in inferential work. Another important topic for research concerns the yet unknown relative advantages and disadvantages of using one correlation technique over the others in different map reading situations. The essential point regarding quantitative map analysis in general seems to be that the criterion of right or wrong method can only be given in terms of the purpose of the investigation and, even then, given present knowledge, the choice is far from absolute.

Experience has shown that techniques of numerical map analysis seldom reveal information that cannot be ascertained through astute visual map inspection. In fact, the viability of numerical techniques is commonly evaluated, at least initially, against their long established visual counterparts. The implication seems to be that most existing numerical methods of map analysis are essentially routine data-processing techniques and do not function as powerful new research tools. It is possible, however, that, given the inherent sensitivity and repeatability of quantitative analysis it is our inability to ask sophisticated theoretical questions which accounts for the apparently low analytical potential of known procedures. In any case, the importance of these numerical procedures in automated information systems cannot be denied, even at our relatively naive stage of development in this area.

An impressive list of advantages of quantitative map analysis over visual map analysis has also been asserted. In the long run, the most valuable attribute of quantitative techniques appears to be their objective (that is, repeatable, nature). Numerical map processing permits a systematic analysis of spatial pattern data. Mathematical map reading techniques also deal in an explicit and controllable way with problems relating to spatial scale (see above), two-dimensional map pattern classification and form/process relationships, all of which are extremely difficult, if not impossible, to discriminate and control intuitively. Even the most superficial consideration of the data handling capabilities of, for instance, *modern factor analysis*, *multidimensional scaling*, or *canonical trend analysis* suffices to convince most researchers of the superiority of multivariate quantitative analysis over equivalent visual map analysis. Many researchers also assert that quantitative map analysis facilitates the detection of smaller differences and more subtle or complex relationships, and provides a more objective basis for judging the significance of these differences and relationships, than is possible through visual estimates alone.

In spite of the above claims, it is only fair to say that the relative advantages of the visual and quantitative approaches and, therefore, the proper mix of the two techniques in a research design, is still poorly understood. An uncritical attitude toward the role of maps in scientific pursuit has compounded the confusion. Most of the strong and convincing support for the place of maps in geographical science is best viewed as conjecture. It is certain, however, that both quantitative and visual analysis play important roles in the cyclical pursuit of science. Mathematical methods provide a powerful logical framework with which to build theoretical scientific models. Yet in spite of the degree of sophistication attained in the quantitative analysis, there is no way to avoid the subjective element in interpreting the end product. The ready perceptibility of graphic presentation facilitates the human need for visualization and can often clarify or simplify abstract and complex spatial relationships. In fact, subsequent "mappings" of the results of quantitative map analysis are often essential to their interpretation. It is precisely in this interpretative stage that quantitative map analysis is at its weakest level, and only a large amount of research will assure that cartography will benefit fully from the potential degree of rigor inherent in quantitative procedures. In the meantime it seems reasonable to conclude that in well designed scientific research, visual map analysis is particularly suited for the inductive hypothesis generation and model refinement stages, whereas quantitative map analysis is more appropriate for the deductive testing phase.

(3) Optical Analysis

Optics provides a versatile and powerful tool capable of performing a number of linear operations on two-dimensional signals such as maps. Passing light through a reduced transparency of a map produces a diffraction pattern (or 2-D Fourier transform), which displays the distribution of spatial and directional frequencies associated with all elements in the map. Extremely simple lensless optical devices utilizing only a diffuse light source are capable of performing auto- and cross-correlations on map inputs. More complex and versatile configurations employing lenses and coherent optical techniques have been used to perform spectral analysis, various filtering operations, and auto- and cross-correlations of map inputs.

A lack of adequate data has seriously limited mathematical two-dimensional spectrum analysis. To a lesser extent, the computational difficulties associated with such large matrix manipulations have limited mathematical analysis in spite of the efficient Fast Fourier Transform. These problems are overcome by optical configurations which require only precise alignment of a reduced image transparency, lenses, and a coherent light source (e.g., neon laser) along an optical bench (Figure 44). Optical diffraction thus provides us with a direct and practical analytical capability in an area of map study which is largely unexplored.

The size, spacing, and orientation of distributional features over a wide range of scales can be studied by analyzing and comparing their two-dimensional diffraction patterns. However, a number of difficult interpretative problems indicate that new pattern recognition methodologies must be devised before direct spectrum analysis of maps is generally meaningful. Scanning and contouring of diffraction patterns as a preliminary to comparing such patterns has been suggested. It looks as if the old map interpretation problem has returned!

Spatial filtering is an interesting aspect of optical map processing. Using optics it is possible to modulate the input image when it is in the frequency domain and then reconstruct the transformed version of the original image. This technique has been used to improve image quality by sharpening edges, reducing noise, and increasing contrast. The possibility of both directional and spatial frequency filtering to simply remove select elements from the reconstructed image has been demonstrated. Since spatial smoothing and filtering are perfectly general, the reader should refer back to previous discussion of these concepts for mention of known and potential areas of application.

The future role of optical processing in cartography is largely a matter of speculation based on preliminary experiments. In addition to the advantages already mentioned, it is significant that optical techniques permit direct analysis of graphic input with the obvious advantage that spatial parameters are left intact in all ways. The speed of optical processing and its application to graphic materials without the necessity of mode changes (e.g., digitizing) recommend optical techniques as a valuable component in any modern spatial information system. The possible disadvantage of having to interpret a graphic interpretation of a graphic can be distracting and will require the development of new pattern recognition methodologies.

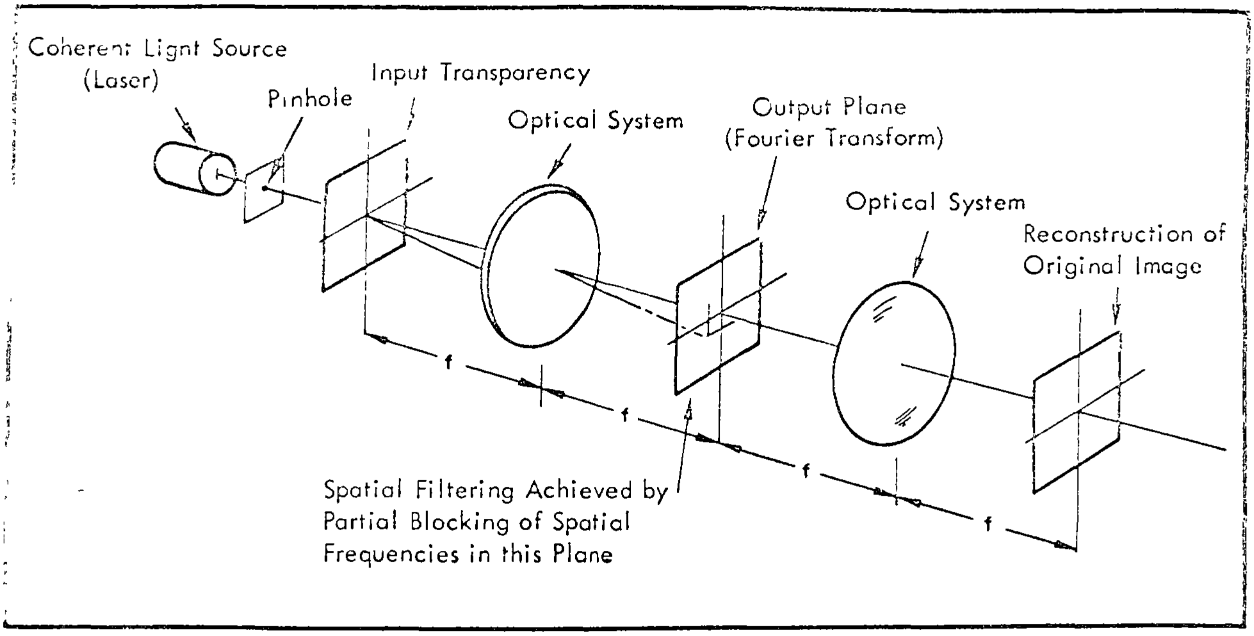


Figure 44 A typical configuration of an optical processing system

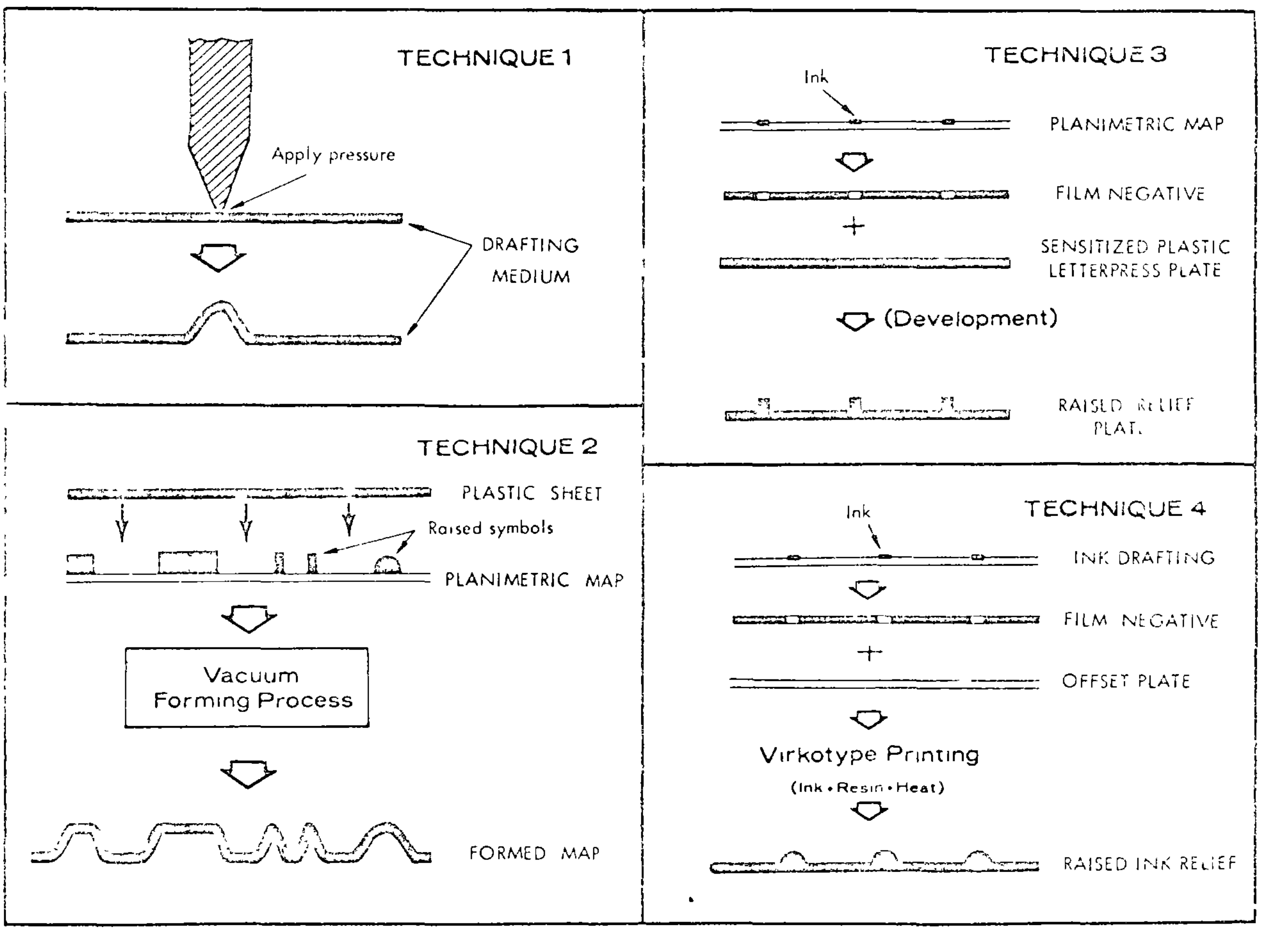


Figure 45 Methods of producing tactual maps for the blind

(4) Tactual Map Reading

The blind have always had special map reading problems. Relevant cartographic activity may be grouped roughly into that portion devoted to determining the needs and capabilities of the blind as map readers, and that portion which emphasizes the actual design and production of tactual maps. Although this distinction helps to identify primary problem areas, in practice the two research goals are inseparable. Blind map readers require map forms which are entirely different from conventional maps. Sadly, it is only in recent years that cartographers have begun to explore systematically the possibility of using the skin as a channel for pictorial material. Already several quite acceptable visual substitute techniques employing embossed graphic symbols are available or under development (Figure 45).

Single one-of-a-kind maps are commonly produced by simply attaching tactually discriminable symbols to a planimetric map base. Multiple copies of this original can be vacuum-formed into plastic sheets using special Thermoform equipment. Maps possessing limited symbol relief can be produced by standard reproduction methods and Virkotype printing which employs special inks that stand up from the map surface in relief. Special plastic media have been developed which respond locally to pressure applied with a pen-like instrument by rising up in relief. These raised drawings can be read easily by the blind and should be particularly valuable as teaching aids. At the University of Washington, we are experimenting with a method for direct conversion of graphic copy (e.g., conventional maps) into tactual display. A film negative of the map is contacted with a special plastic printing plate. Chemical development produces 1/8 inch relief, more than sufficient for direct use as a map by the blind, but also satisfactory as an original for a forming machine or for a printing plate to be used with special raised inks.

Tactile image systems capable of utilizing conventional visual images are also promising. One experimental system uses a subject controlled television camera as a sensory device (eye) and electronically transforms the video image into a tactile display which is transmitted to the back of the subject via a bank of vibrators. Another interesting system is based on a hand-held reading instrument which turns ordinary visual images into vibrations which the blind can translate through touch.

Continuing research on maps for the blind and partially sighted is directed primarily at the identification and production of clearly discriminable symbols. The nature of tactual figure/ground relationships also remains a crucial problem area. The search for new media is unending. Major breakthroughs have often come about through discovery (or development) of a new display medium or technique, which for evaluation required cartographers to go back and research user requirements and responses. Possibly a more theoretical approach would suggest solutions to some of the many challenges which persist.

V. THE OUTLOOK

Thematic cartography has now been pursued sketchily through a formidable array of applications, trends, and problems areas. This hasty treatment of a complex and rapidly developing subject is meant to be suggestive rather than exhaustive. I hope that the notion has been conveyed that there is a conceptual and theoretical scheme of remarkable unifying power underlying relationships between the more or less isolated steps in cartographic information processing. The transition from map making to cartography will gain momentum as cartographers abandon the notion that the map is a unique and final product and explicitly treat cartographic analysis as an integral part of the broader geographical information system. Computer scientists have clearly demonstrated that graphics can serve to monitor complex information processing in addition to satisfying their more traditional illustrative and storage roles.

The complementary nature of the alternative languages of words, logic, mathematics and graphics (including maps) in communicating geographical information must continually be held in perspective. Likewise, different map forms themselves must not be seen as antagonistic devices, but are better regarded as mutually complementary portrayals serving different needs. In this view, map variations are not considered as competitive presentations but as supplementary techniques analogous to multiband sensing. The research potential of a *map interpretation matrix* made up of a number of maps, each of which is most sensitive to a particular aspect of the data, is suggested.

Increased concern with the scientific aspect of maps should result in truly scientific presentations. A statement of accuracy standards and underlying assumptions should become as common a map element as the title, legend, or locational references. But without objective testing of a convincing nature, the prospects for a unified science of cartography seem poor. We still know far too little about the comparative efficacy of various cartographic methods. Our present understanding of *map reading* in particular is dismal. We are not able to say with any confidence what training is required to produce good map interpreters. The information carrying power of even our most common mapping methods is little understood, and almost no research has been devoted to ascertaining the relative effectiveness of various cartographic techniques in different problem situations. Our evaluation of cartographic communication must advance far beyond the few extant studies concerning the psychophysical relation of map symbols. In general, if theoretical principles to guide the choice of cartographic options are to be developed, a more ingenious and perhaps more experimental approach toward formulating and testing meaningful hypotheses must replace our intuitive tradition. In this paper some of the basic issues before cartography have been presented as likely starting points for this research. Many challenges still remain.

Since automation of the cartographic processing system appears to provide at least a partial solution to many crucial mapping problems, we should anticipate an ever greater emphasis in this area. The possibility of automating most aspects of cartography has long been predictable (recall Turing's theorem), and thus the extent of current intrigue with automating various phases of cartographic activity is somewhat surprising from an intellectual viewpoint. In spite of the practical implications of automatically replacing manual methods and multiplying the number of maps produced, it is distressing to see the unlimited promise of computer innovations left unfulfilled. The present fascination for computers and related equipment as an end in themselves is expected to subside eventually and become subordinate to the application of these devices to solving more truly cartographic problems. The capabilities of automated devices and modern communication systems will surely encourage critical evaluation of conventional maps and result in the design and implementation of map forms better suited to contemporary and future needs.

The multidisciplinary nature of cartography in its various aspects dictates that future cartographers possess in-depth knowledge of a wide array of topics and procedures. The need to be able to understand and implement research conclusions as well as make contributions to basic research cannot be overly stressed. Creation of a learning environment that encourages the pursuit of basic research by more cartographers is a top priority. Skillful choice between alternative solutions to given cartographic problems necessitates weighing the relative simplicity, meaningfulness, and interpretability of the various possibilities. These decisions can be made intelligently only after the diverse information from many disciplines which contributes to effective cartographic presentation has been synthesized. Maps are important to many people—geophysicists, meteorologists, geologists, ecologists, oceanographers, geographers, geomorphologists, and others. Cartographic curriculums must become more sensitive to the interests of these disparate fields of interest if cartographic processing is to continue satisfying their diverse needs. The alternative, of course, is to replace proper cartographic analysis of spatial distributions with a better method, and at this time no sufficient substitute is known.

REFERENCES AND SELECTED READINGS

Agterberg, F P, 1967, "Computer Techniques in Geology," *Earth-Science Reviews*, 3, 1, pp 47-77

Arnoldy, Carolyne, 1971, "Plotters One Way to Go–Up!," *Information Display*, 8, No 5 (Sept/Oct), pp. 23-27

Bassett, K , and R Tinline, 1970, "Cross-Spectral Analysis of Time Series and Geographical Research," *Area*, 1, pp 19-24

Beer, Stafford, 1968 *Management Science*, Doubleday Science Series New York, Doubleday and Co

Beller, Henry K , 1970, "Problems in Visual Search," *Paper*, Symposium on the Influence of the Map User on Map Design, Queens University, Ontario

Berry Brian J L, and Alan M Baker, 1968, "Geographic Sampling " in *Spatial Analysis*, edited by J L Berry and D Marble, Englewood Cliffs, Prentice-Hall, pp 91-100

Blackman, R B, and J W Tukey, 1959, *The Measurement of Power Spectra*, New York, Dover Publications, Inc

Blome, Donald A , 1963, "A Map Transformation of the Time-Distance Relationships in the Lansing Tri-County Area," Institute for Community Development and Services, Michigan State University, East Lansing

Blumenstock, David I, 1953, "The Reliability Factor in the Drawing of Isarithms," *Annals of the Association of American Geographers* 43 4, pp 289-304

Cartwright Robert 1970, Paper on experimental coordinate plotter shading presented at the Northwestern Conference on Computer Mapping

Castner, Henry W , and Arthur H Robinson 1969, "Dot Area Symbols in Cartography The Influence of Pattern on Their Perception," *Technical Monograph No CA 4* American Congress on Surveying and Mapping Cartography Division

Choynowski M , 1968, "Maps Based on Probabilities," in *Spatial Analysis*, edited by B J L Berry and D Marble, Englewood Cliffs, Prentice-Hall, pp 180-183

Dacey, Michael, 1970, "Linguistic Aspects of Maps and Geographic Information," *Ontario Geography*, 5, pp 71-80

Dahlberg, Richard E , 1967, "Towards the Improvement of the Dot Maps " *International Yearbook of Cartography*, 7, pp 157-167

Dornbach, John Ellis, 1967 *An Analysis of the Map as an Information System Display*, Unpublished Ph D Dissertation, Clark University

Ekman, Gosta, Ralf Lindman and W William-Olsson, 1961, "A Psychophysical Study of Cartographic Symbols," *Perceptual and Motor Skills* 13 3, pp 355-368

Ericksson, I Sture, 1967a "Field Effects and Two-Dimensional Form Perception," *Scandinavian Journal of Psychology*, 8, 4, pp 216-242

_______, 1967b, "Factors Influencing Two-Dimensional Form Perception," *Scandinavian Journal of Psychology*, 8, 4 pp 209-217

Fisher, Howard T 1970 *Mapping Quantitative Information Issues and Principles* Unpublished paper from the Laboratory for Computer Graphics and Spatial Analysis, Harvard University

Forgotson, James M , 1960, "Review and Classification of Quantitative Mapping Techniques," *Bulletin of the American Association of Petroleum Geologists*, 44, 1, pp 83-100

Gould, Peter, 1966, "On Mental Maps," Michigan Inter-University Community of Mathematical Geographers, *Discussion Paper No 9*

Granger, C W J , 1969, "Spatial Data and Time Series Analysis," in *Studies in Regional Science*, edited by Allen J Scott, London, Pion

Halloway, J Leith, 1958, "Smoothing and Filtering of Time Series and Space Fields," *Advances in Geophysics*, 4, pp 351-388

Harris, C D , 1954, "The Market as a Factor in the Localization of Industry in the United States " *Annals of the Association of American Geographers*, 44, 4, pp 315-348

Harvey, David W , 1966, "Geographical Processes and the Analysis of Point Patterns," *Transactions, Institute of British Geographers*, 40, pp 81-95

_______, 1968, "Pattern, Process, and the Scale Problem in Geographical Research," *Transactions, Institute of British Geographers*, 45, pp 71-78

Hershey, A , 1963, "The Plotting of Maps on a CRT Printer," NWL Report 1844, U S Navy, Dahlgren

Hsu, Mei-ling, and Arthur Robinson, 1970, *The Fidelity of Isopleth Maps An Experimental Study*, Minneapolis, The University of Minnesota Press

Jenks, George F , 1967, "The Data Model Concept in Statistical Mapping," *International Yearbook of Cartography*, 7, pp 186-188

Jenks, George F and Fred C Caspall, 1971, "Error on Choroplethic Maps Definition, Measurement, Reduction," *Annals of the Association of American Geographers*, 61, 2 (June), pp 217-244

Jenks, George F , and Duane S Knos, 1961, "The Use of Shading Patterns in Graded Series " *Annals of the Association of American Geographers*, 51, 3, pp 316-334

Jenks, George F and Theodore R Steinke, May 1971, *Final Report for Three Dimensional Maps*, NR 389-146, NONR 583 (15), Department of Geography, The University of Kansas Lawrence, Kansas, 20 p

Jones, William B , and Roger M Gallet 1962a "Representation of Diurnal and Geographic Variations of Ionospheric Data by Numerical Methods," *Journal of Research*, D, 66D, 4, pp 419-438

_______, and _______, 1962b, "Methods for Applying Numerical Maps of Ionospheric Characteristics," *Journal of Research*, D , 66D, 6, pp 649-662

King, Leslie J 1969, *Statistical Analysis in Geography*, Prentice-Hall, Inc , Englewood Cliffs, N J

Lyddan, Robert H , 1971 "Mapping Goals for the Future," *Technical Papers*, 31st Annual Meeting of the American Congress on Surveying and Mapping, March 7-12, pp 425-434

MacDougal, E. Bruce, 1970, "Spatial Filtering," *Economic Geography*, 46, 2, Supplement, pp 425-434.

McCarty, Harold H , and Neil E Salisbury, 1961, *Visual Comparison of Isopleth Maps as a Means of Determining Correlations Between Spatially Distributed Phenomena*, Iowa City, State University of Iowa

McCleary, George F , 1970, "Beyond Simple Psychophysics: Approaches to the Understanding of Map Perception," *Technical Papers*, 30th Annual Meeting of the American Congress on Surveying and Mapping, March 1-6, pp. 189-209

Morrill, Richard L , 1970, *The Spatial Organization of Society*, Belmont, Calif Wadsworth Publishing Company

Morris, Charles, 1938, *Foundations of the Theory of Signs*, Foundation of the Unity of Science, 1, 2, The University of Chicago Press, Chicago.

Morrison, Joel L , 1970, "A Link Between Cartographic Theory and Mapping Practice The Nearest Neighbor Statistic," *The Geographical Review*, 60, 4, pp 494-510

_______, 1971, *Method-Produced Error in Isarithmic Mapping*, Technical Monograph No CA-5, American Congress on Surveying and Mapping

Muehrcke, Phillip C , 1969, *Visual Pattern Analysis A Look at Maps*, Unpublished Ph D Dissertation, University of Michigan. (available from University Microfilm, Ann Arbor, Michigan.)

_______, 1970, "Trends in Cartography," in *Focus on Geography Key Concepts and Teaching Strategies*, ed by Phillip Bacon, 40th Yearbook of the National Council for the Social Studies, Washington, D C , National Council for the Social Studies, pp 197-225

Neft, David, S , 1966, *Statistical Analysis for Areal Distributions*, Monograph Series No 2, Regional Science Research Institute, Philadelphia

Olson, Judy M , 1970 *The Effects of Class Interval Systems on the Visual Correlation of Choropleth Maps* Unpublished Ph D Dissertation, University of Wisconsin (available from University Microfilms, Ann Arbor Michigan)

Pierce, J R , 1961, *Symbols Signals and Noise The Nature of Communication* Harper Torchbooks, The Science Library, New York, Harper & Row

Reps, Frank J , 1967 Cover illustration of three methods for showing the same geographic data *Photogrammetric Engineering* XXXIII, 3, March

Robinson, Arthur H , 1961, "The Cartographic Representation of the Statistical Surface," *International Yearbook of Cartography* 1, pp 53-63

_______, 1967, "Psychological Aspects of Color in Cartography," *International Yearbook of Cartography*, 7, pp 50-61

_______, 1970, "Scaling Nonnumerical Map Symbols," *Technical Papers*, 30th Annual Meeting of the American Congress on Surveying and Mapping, pp 210-216.

Rosenfeld, Azriel, 1969, *Picture Processing by Computer*, New York, Academic Press.

Siegel, Sidney, 1956, *Nonparametric Statistics for the Behavioral Sciences*, New York, McGraw-Hill Book Company

Sinaiko, H Wallace, ed , 1961, *Selected Papers on Human Factors in the Design and Use of Control Systems*, New York, Dover Publications

Sutherland, Ivan E., 1970, "Computer Displays," *Scientific American*, 222, June, pp. 56-81

Tobler, Waldo R , 1966, "Notes on the Analysis of Geographical Distributions and Numerical Map Generalization," Discussion Paper 8, Michigan Inter-University Community of Mathematical Geographers, Department of Geography, University of Michigan, Ann Arbor

_______, 1969a, "An Analysis of a Digitalized Surface," in *A Study of Land Type*, by Charles M Davis. U S Army Research Office (Durham), Contract No DA-31-124-ARO-D-456, University of Michigan

_______, 1969b, "Geographical Filters and Their Inverses," *Geographical Analysis*, 1, 3, pp 234-253

_______, Oct./Nov. 1970, "Computers as Automatic Draftsmen and Picture Processors," *Research News*, Office of Research Administration, University of Michigan, XXI, 4 and 5, pp 20-22

Tomlinson, R. F. (ed), 1970, *Environment Information Systems*, Proceedings of the UNESCO/IGU First Symposium on Geographical Information Systems, Ottawa

Warntz, William, 1971, "Computer Graphics and Theoretical Geography," in *Research Institute Lectures on Geography* coordinated and directed by Roger A Leestma US Army Engineer Topographic Laboratories, Special Report ETL-SR-71-1, Fort Belvoir, Virginia, pp 79-112

Wiedel, Joseph W , and Paul A Groves, 1970, "Tactual Maps," *International Yearbook of Cartography* 10, pp 116-123

Wittick, Robert, 1971, *Geography Program Exchange*, Computer Institute for Social Science Research, Michigan State University, East Lansing, Michigan (see also the newsletter "Current Research Notes in Quantitative and Theoretical Geography" published irregularly by the same source)

Yarbus, Alfred L , 1967, *Eye Movements and Vision*, New York, Plenum Press

Yoeli P , 1965, "Analytical Hill Shading," *Surveying and Mapping*, XXV, 4, pp. 573-579

ASSOCIATION OF AMERICAN GEOGRAPHERS

COMMISSION ON COLLEGE GEOGRAPHY

Membership – January 1972

Dr. Robert H. T. Smith, Commission Chairman, Queen's University
Dr. John S. Adams, University of Minnesota
Dr. Ian Burton, University of Toronto
Dr. Richard D. Bastyck, Fullerton Junior College
Dr. Paul W. English, University of Texas
Dr. Gordon J. Fielding, University of California, Irvine
Dr. J. Thomas Hastings (Education), University of Illinois
Dr. Leslie J. King, McMaster University
Dr. Richard E. Lonsdale, University of Nebraska
Dr. Ian R. Manners, Columbia University
Dr. Marvin W. Mikesell, University of Chicago
Dr. Harold M. Rose, University of Wisconsin-Milwaukee
Dr. Harold A. Winters, Michigan State University

Ex-Officio Members: President, Edward J. Taaffe; Vice President, Wilbur Zelinsky; Secretary, Wesley C. Calef; Treasurer, Donald A. Wise; Executive Director, J. Warren Nystrom

Dr. John F. Lounsbury, Project Director, Arizona State University, Tempe, Arizona 85281
Dr. Salvatore J. Natoli, Associate Director, Association of American Geographers

Eric Teicholz

GRAPHICS

for Regional Policy Making, a preliminary study

William R. Ewald, Jr., development consultant
1730 K Street, N. W. , Washington, D. C. 20006

August 17, 1973

TABLE OF CONTENTS

PART I

ABSTRACT

This report assumes that this is a time of extraordinary change, that this change is pervasive, and that to manage it will require us to develop new institutions and new processes - as well as to modify ongoing institutions. (The principal investigator's NSF report ACCESS - The Santa Barbara Pilot Process contains a description of such a new institution and process to improve regional policy-making.)

This report also assumes that new analysis/synthesis/communication capacities are needed for these institutions. The complexity and continuous interaction of the many factors involved in regional policy-making seems to call for data, comprehension and communication as never before. Even the definition of policy-makers seems to be changing, to include more than the traditionally recognized officials of government and business. It is now becoming evident that concerned citizens, especially the leadership of citizens' organizations, are also needed as policy formulators, who take part all through the process, from conception, definition of problems, the gathering of data, the analysis of alternatives, decision-making, through to implementation and evaluation.

For a democracy to perform in this technological era of great change it probably must reeducate itself to the task. A result of this line of reasoning is to recognize that education is analysis, synthesis and communication. To put it yet another way, if education is the process of teaching people how to think about things, then the analysis/synthesis/communication process described here is community self-education.

This report identifies and delves into the capabilities of a number of tools for analysis/synthesis and communication, especially computer assisted graphics. It invites consideration of the potential of this technology for the purpose of managing change, but the principal investigator does not expect that simply by employing such new tools that the inexact can be made exact, the unquantifiable quantified or decision-making automated. But I believe that these new electronic tools can contribute to developing a capability that we now lack for managing change.

Part I outlines the features of the analysis/synthesis/communication tools needed to formulate better policy to conserve and develop regions. Part II is a preliminary effort to match that need with technology that is already evolving.

I. GRAPHICS FOR REGIONAL POLICY MAKING

GRAPHICS

This report emphasizes the use of graphics (maps, charts, diagrams, renderings and photographs) for regional policy formulation and decision-making. Does this mean that the alphabet and numerals, graphic symbols in themselves, are to be left out? Actually, words and numbers are not only essential auxiliaries to the successful use of non-verbal symbols, they may commonly provide complete communication within analysis/synthesis/communication systems which offer non-verbal graphic capabilities. The important matter here is whether or not the system provides non-verbal graphic capacity.

Why this emphasis on non-verbal graphic technology? This report assumes that non-verbal graphics can greatly aid analysis, synthesis and communication concerning the many variables involved in formulating regional policy. Although such graphics have been used in planning and policy-making in the past, this report assumes that a whole new language of visual communication is on its way which holds great potential for regional policy-making.

We are moving into an era when color, pattern, motion, sound, and two-way communication will be increasingly available. For policy-making, research, education and management this means a new possibility for comprehension of complexity, without the distortions of over-simplification.

Graphic communication offers a potential for clarifying perception. It offers a way to improve our intuitive intellectual grasp of large complex systems. It provides a new opportunity to perceive the values and attitudes behind conflicting viewpoints. Graphic communication can convey complexity much more quickly, provide spatial orientation for information, and cut through semantic difficulties that confuse different specialities and different political and cultural views. The graphic communications of television are what most people "read," and it offers the potential of being able to communicate the abstractions of future consequences of pending decisions in more understandable form, within limited spans of attention. Further, graphic communication with two-way capability can be queried to suit and enhance the understanding of the user. This is an absolute essential for comprehension of regional problems, unsticking mental blocks that cannot often enough be perceived by other persons.

Regional decision-making, under pressure, that attempts to take into account the complexity of interrelationships among many variables, over different geographic scales and at varying increments of time, while accounting for changing technology, and at the same time acknowledging the changing expectations of people, needs more than reports, rolls of maps or public hearings to illuminate problems, choices and consequences! It seems to call for graphics of all kinds; including computer assisted graphics that are responsive to the viewer.

THE REGIONAL POLICY MAKER Bw1.1

The term, regional policy maker, as used here includes those who contribute to the formulation of policy as well as those who decide policy. Both are considered part of "policy-making." Therefore, interested citizens, as individuals and as organization members, are policy-makers -- as are corporation directors and government officials, elected or not. That is a substantial difference from the more common use of the term and contributes significantly to the definition of the analysis/synthesis/communication process required.

A listing of assumptions about the regional policy maker may be helpful here; he is:

(1) an intelligent, basically honest generalist;
(2) under political and time pressures to make decisions;
(3) inadequately informed and full of his own questions;
(4) impatient with details, but skeptical of summaries;
(5) able to give only severely restricted time for analysis;
(6) suspicious of computers (and their programmers);
(7) uncertain about typing into a computer terminal;
(8) unable to comprehend, unaided, the interrelationships of many variables;
(9) probably limited in his ability to use maps, charts or graphs;
(10) skeptical about the validity of long-range projections (and expert opinions);
(11) limited in his comprehension of development factors, important to the region, which work at different geographic scales;
(12) only partially informed regarding technology and its impact;
(13) not really convinced that major social and economic changes are underway;
(14) a believer in adaptation, not radical change;
(15) seeking respected authority on which to base his decision.

MAKING REGIONAL POLICY

To make regional policy seems to involve different groupings and combinations of people:

(a) an aggregation of public officials without an established regional political constituency;

(b) businessmen with differing short-term concerns, some controlled from outside the region;

(c) citizen-based groups which may or may not be regional in scope, each with different purposes in mind; and

(d) many other competent people, including those in universities and in the professions, who don't yet relate to "region" as important to what they do and how they live.

To communicate concepts, issues, information, analyses, judgments, and feelings about regional issues to these regional policy-makers in such a way that basic choices buried in current decisions are related to alternative futures, seems to call not only for valid analysis and communication, it must also be accepted as credible. Whereas a major aspect of credibility of information is its source, perhaps much more significant than the scant attention it has been given to date, is how regional policy-makers perceive information. This is a matter, not only of how their minds are already "set" when they receive information, but how information is presented to them. The physiology and psychology of how we see and receive information may have an importance to policy-making that is yet untended. Cognitive.

Probably, information, even if it is accurate, valid and reliable, must be assimilated comfortably by a person, to the point it "feels right", before it is credible. When information reaches that state, that of an insight to be trusted, it can affect policy-making. It may be that information must become something akin to "intuition" before it leads to decision. If intuition, in turn, can be defined as unconscious reasoning, then whatever helps that reasoning has a significant effect on credibility.

The important point for the policy-maker besides the accuracy, validity and reliability of the information, analysis and synthesis provided him, is whether it is readily assimilated by him. To assimilate and use information, he probably needs to "sense" as well as to rationally calculate his decisions. In that he can assume that the information and data provided him will almost always be inconclusive, however conscientious the effort, it is especially important to recognize that informasomewhat more comfortable with one view than with others in order to make a decision at all (except the decision not to decide.)

Ultimately, the policy-maker must leap from the best that can be quantified and reasoned, across what cannot be known, to his position or decision. It's a leap of faith, or what policy analyst Sir Geoffrey Vickers terms "the passion of judgment." (Beyond reason alone.)

There are many subjective elements in arriving at a decision or a position. The purpose for bringing more of the capacities of science and technology to regional decision-making is to help synthesize complex problems, to reduce uncertainty, to quantify those elements which are quantifiable, to reduce the number of subjective elements that are so because of ignorance, and to convey analysis and information with credibility, and to provide the forum, the feedback to assure valid means.

Recognizing that all alternatives for each issue cannot be explicitly stated and then precisely quantified, there comes recognition of the need for an accurate, credible process that generates and reinforces sound intuitions in the policy-maker. Much of the process we are seeking to evolve may be in the policy-maker(s) mind(s). Perhaps we should recognize that what we are attempting in computer-assisted interactive graphics is new abilities for feeding information to those minds. An example of that "processing" may be how we sensed Spaceship Earth from the Moon photographs taken by the astronauts. With that accurate graphic communication an entire perspective was perceived. (How puny the words of astronauts compared to their photographs.)

Non-verbal graphic symbols communicate to many levels simultaneously. Such graphic communications may become a "universal" language for policy-makers, managers, workers, learners and consumers about their region. The same information, used differently by researchers, policy-makers and the general public, may be communicated in the same form, through the same technology, using the same systems.

If people do not understand because they fail to perceive the "whole" of complex systems, because they lack a "perspective of reality," perhaps, by using common graphic symbols, people of a region can perceive more perspectives in common, including the understanding that different people, with the same information, may arrive at different perceptions. Maybe the perception of wholes will be among the greatest contributions graphic non-verbal communication can make to constructive regional dialogue.

ANALYSIS/SYNTHESIS/COMMUNICATION TOOLS - DESIRED FEATURES

ACCESS

The companion report ACCESS - The Santa Barbara Regional Pilot Process, proposes a new institution and process to aid and improve regional policy-making concerning the conservation and development of the environment. That report defines the basic characteristics of the analysis/synthesis/communication technology that might be used by such an institution and process. Although use of such technology need not be limited to the ACCESS process, it will be helpful to summarize here the nature of ACCESS (Alternative Comprehensive Community Environmental Study System) which is the context for analysis/synthesis/communication tools proposed here.

ACCESS is:

(a) regional in scope (multi-jurisdictional);

(b) of, by, and for regional policy-makers (and open to use by interested citizens and students);

(c) a permanent, legally organized non-profit institution (with a broad base of financial support and a board that is representative of the region);

(d) a place and a means to store, retrieve, analyze, synthesize, and display regional data and information - economic, environmental, cultural - by means of an interdisciplinary staff, consultants, volunteers and community dialogue (that makes appropriate use of computers and available communication technology);

(e) without operating responsibilities, (the emphasis of ACCESS is on assessment of problems, alternatives and long-range consequences, especially as derived from pending decisions).

The institution and process are cross linked - indivisible. How then could the volume of data and information required be stored, called up, displayed? What sort of analysis/synthesis/communication tools would facilitate comprehension of the link between current regional decisions and future consequences? How could graphic display of information be used to stimulate regional dialogue on Growth and the Quality of Life? Just what sort of technology and what systems would improve regional policy formulation and decision-making of the sort described here?

STORING AND RETRIEVING DATA AND INFORMATION

Traditional library storage techniques, ordered to the special purpose of regional policy formulation, are essential as the beginning. All environmental, economic and cultural data and information - especially in map and plan form - and all operational planning and legal reports for the entire region should be assembled in one place. Probably the library should be both in hard copy and micro-fiche form.

Data related to specific operating functions within the region (such as property descriptions, plats, assessments, zoning, public works, commerce and industry, traffic, utilities, housing, water, sewer) might best be maintained by the organization that has responsibility for and greatest use of that data, provided the data is coded, indexed and put into compatible computer format so that it can be easily accessed from remote terminals, and can be spatially displayed in relation to other data and information.

A study would need to be made of existing sources of data, how it is currently assembled, what scale grid and/or property reference? What systems it could be keyed to? What additional data and information are needed? How it is to be collected and maintained? (Information is defined here as ordered or interpreted data.) What bank of select regional, state, national and world indicators of particular value to regional policy-making would be ordered into a format to be retrieved through computer assisted terminals?

COMPUTER TERMINALS

Present terminals, and the programs they use, have been developed largely of, by, and for middle management for operational purposes. If they are to be used effectively for policy-making, adaptations of certain equipment and programs seems needed. Computer terminals should offer policy-makers.

(a) response to unstructured questions;
(b) visual display that is capable of continuous tone color and movement;
(c) larger-size (than typical now) alphabet letters in both upper and lower case;
(d) graphic as well as alpha-numeric inputs;
(e) hard copy of any display;
(f) audio as well as visual responses;
(g) audio as well as graphic and keyboard inputs;
(h) interchange with other terminals in all modes (audio, graphic, keyboard, etc.);
(i) access to data for personal use at place of work, in designated public places (such as schools, or libraries or community centers) and at home;
(j) access to data for dialogue purposes in specially designed "regional situation rooms".

Some of these capabilities are available now, but not all of them. They would be needed in varying degrees of completeness at various locations. The most completely equipped would be the regional situation room.

REGIONAL SITUATION ROOMS

Each region would need one basic center for the library, core staff and the central regional situation room.

The central regional situation room would provide a specially designed dialogue facility for approximately 15-20 people which would have all needed graphic and analysis support facilities, including television equipment for video-taping and for conveying two-way signals to other situation rooms within (or outside) the region.

As NASA has found, the overhead projected slide delivered ahead of time, serves well as the focus for simultaneous conferences at separate locations interconnected by special high fidelity telephone lines. (See page 16) The program EMISARI, which provides hard copy printout from a keyboard terminal, has increased the efficiency of conferences partially because people using it are conferring from separate locations. (See page 21)

Motion pictures, videocassette and slide projector facilities would be part of the equipment of each situation room where group discussions are held. Except for very large "surround" screen projection or when highest resolution images are required, projected or standard monitor size television offer the potential of coupling to digital computers, and interactive response and control, which supersedes other means of providing graphic display for regional policy-making purposes.

The situation room might very well provide for each participating member the following: a television monitor with keyboard; a Scriptographic or RAND Tablet for graphic inputs. (See page 45) There are other pushbutton controls to register opinion and degree of concern, electronically. The CONSENSOR is a commercial product available for this purpose. Another potential product varies the size of images on multiple projection screens; images are increased or decreased in size according to opinions registered by pressing buttons. (For more on CONSENSOR, see page 65)

Large-scale (10' to 20' in width) projection of television and digital symbols by means of devices such as the Eidophor or GE Light Valve (See page 50) would seem to be desirable for focusing attention on one speaker and one graphic presentation simultaneously, especially if the dialogue in the situation room is itself to be televised.

The computer backup capacity for the situation room could be by telephone line to commercial university time/shared facilities or to the ARPA Network (see page 51). Simple means to make inquiries should be made possible for any user, but for calling up certain information it is probable that at first a number of different programs and computers will be needed. Dealing with different computers and their different programs, at their present stage of development, is like dealing with Swedes, Chinese, Italians, Frenchmen, Englishmen and Argentines all at once. They may, or may not, all "understand" the same basic language; but they are not necessarily individually compatible (interchangeable).

This means having persons present to whom specialized questions could be put. They would operate the computer terminals very much the same as the terminals are manned at the NASA Command-Control Centers, each to its own special areas of operation. In a regional situation room accurate, quality-controlled information might be divided into three or more knowledge areas such as environmental, economic, cultural - each with its "specialist." If questions requiring especially powerful computation capacities were asked, the specialists concerned could turn to a special terminal, if need be, such as PLATO now provides. PLATO permits open-ended querying and dynamic responsive computation with graphic display, but it is not compatible with television. (See page 39)

AUXILIARY FACILITIES

Examples of electronic equipment to be investigated that have contributions to make to the analysis and handling of information include a 1000-line cathode ray tube (CRT) for fast, inexpensive computer-assisted animation, plotters to scan and make up maps, digital storage and manipulation capacity to move instantly from one scale enlargement to another.

III CONCERNS AND CONSTRAINTS

The previous section is only the most tentative sort of definition of the graphic analysis/synthesis/communication tools that seem to offer a useful potential to the proposed regional process worth further investigation. Even so, tentative as this report is it may, somehow, raise concerns or be interpreted to provide an emphasis other than what the principal investigator intends. Therefore, it seems sensible to state here that present expectations for using computers and two-way television should not be confused with their future capabilities. Attempts have been made at this elsewhere in this report by adding "as appropriate" when referring to present applications of computer and two-way television.

What is intended, in part, is to build on the experience of already developed computerized information systems. This is taken to mean, especially for regional policy-makers as defined here, avoiding as "the" answer one big comprehensive regional model, even with a number of submodels in it, (See page 24), except to simulate regional action in order to provide:

(a) perceptions of other peoples' perspectives of the region, through computer assisted role-playing games, and

(b) means to ask better questions and learn more about possible consequences of specific decisions.

The concept that there is one big computer program to be developed that can automatically arrive at the best decision, given any particular conservation or development problem, is not the use being thought of here. Rather, what is proposed is to use a variety of computer programs to store, retrieve, compute, and display information - or to provide other specialized capabilities such as generating computer assisted graphic presentations, including "walks" through proposed projects.

Programming for spatial displays responsive to unstructured query of information bases seems quite important to attempt. For instance, specific information that is instantly converted into graphic form and projected might go a long way toward focusing and deepening dialogue. But all this needs to be accomplished in such a way that the margin of error and degree of uncertainty of the information displayed, commonly a combination of different factors, is also clearly conveyed, preferably simultaneously. It is also true that the elements of computer programs need to be revealed for criticism including values implicitly incorporated into them. Without safeguards like these, the power of graphic presentation may be distorted from its intended use as an analysis/synthesis/communication tool to a sales or propaganda device. It is, therefore, no exaggeration to claim that a whole convention of symbols, color, motion and sound will need to be devised. It may also be clear from this brief description how critical will be the integrity of the institution that manages these facilities, for their proper use and credibility.

The facility described above would have major use by the non-profit regional institution interdisciplinary staff, but there is every reason it should be considered available for use by all policy-makers and their staffs, schools, businesses, civic organizations, etc.

The assumption is that different perceptions derived from display of the same basic regional data, information and analysis - in the situation room designed especially for this purpose (and perhaps televised) -

will help make many more people aware of the "different" realities of the region. In turn, it is hoped this will help move them from advocacy of fixed positions, to dialogue about possibilities. By reducing ignorance, uncertainty and sense of risk, people with naturally conflicting interests may more readily discover what they share, and work towards consensus without domination by any one group or individual.

Use of appropriate technology is proposed primarily to facilitate analysis, synthesis and communication on issues of moment to the region. The institution, process and technology conceived are not believed to be a substitute for political decision-making. Such decisions and business decisions would be made where they are now. The intended effect would be to alter the agenda of policy-makers, to better illuminate their choices and to involve many more people in the total decision-making process. The policy-maker would have a much better informed constituency and vice-versa. Even without the power of decision, a competent, credible, non-profit regional institution with continuity and its own base of communication would provide the region with a significant new means with which to derive decisions.

There is no way as yet to estimate the total cost of the analysis/synthesis/communication system for regional policy-making under discussion here. The next approach is to analyze for a specific regional situation just what technological systems are available and feasible now, what may become available, how extensively they might be used, what the total benefits to policy-making are, what they may help to provide (this might well include assessment of waste, construction and social damage they might prevent), and with whom the use and cost of facilities might be shared (such as television productions, government, research, businesses, schools and universities, civic groups, etc.).

PART II

I EVOLVING CAPABILITIES

This investigator has attempted to analyze the state-of-the-art of computer assisted analysis, synthesis and communication that has future practical application for regional policy-makers, including organized interested citizen-leaders. It appears today that there is indeed helpful electronic capacity now available for this purpose. The direction, at whatever pace it proceeds, is towards computer assisted interactive graphics and two-way television. The former is available now in the form of one color, line, cathode ray tube (CRT) display terminals and the electronic plasma panel of PLATO. Continuous color tone television image terminals are the promise of HumRRO, still under development.

To date, two-way television has been oversold and under-delivered, but that does not negate its potential. This year, for the first time, a number of full-scale community demonstrations will be underway. Three such systems are referred to here. They are concerned with testing consumer interest in using terminals (of various sorts) in the home for purposes such as information, instruction, purchasing and entertainment. For purposes of regional policy-making the technology is available, but it is not believed practical to begin with use of two-way television from regional situation rooms into each home of a community without first testing regional information content and the way it is perceived. An appropriate beginning of the application of this technology to regional policy-making would appear to be by multiple situation rooms of different degrees of complexity within a region, for either separate use or simultaneous interaction.

What follows here is a brief description of each of the technologies or systems investigated in the course of their investigation.

NASA's command-control centers, and its much simpler but vital conference room installations, come the closest to providing a regional situation room prototype. In its need for managing complex "real time" operations, NASA had assembled and developed techniques which suggest how "real time" access to complex data and information bases can be developed. Policy-makers, who never have enough time or total comprehension of their subject matter, need whatever facility that can be provided to reduce the pressure which often forces them to resort to over-simplification in order to reach any decision at all. The cumulative negative effect of decisions so made is evident in the environment all around us. But there is a limit to the time decision-makers can devote to each issue. And there is a limit to the policy formulation and staff analysis that is possible. Still, the necessity for decision persists.

During the year regional policy-makers will be concerned with reports, budgets and projects that may total 350 items, about which all possible information must be at hand including location and timing. Any of these matters may be dropped or picked up several times during the year. There are controlling qualitative aspects of change in the environment to be analyzed and decided upon, such as pollution, housing, parking, street maintenance, waste disposal, congestion, zone changes, accidents, crime, code violations, income levels. Future potentials of the region and its neighborhoods need decisions and correlations with the comprehensive plan, which itself requires continued updating. Formulation of fiscal strategies and decisions concerning bonds, loans and the flow of cash are under continual review. Statistical data concerning schools, hospitals, welfare, transportation, etc., are frequently called for. There are also the operational problems of the various governmental departments.

Perhaps if the analysis/communication institution proposed can develop a process that uses some of the technology discussed here, important help can be provided to meet current regional problems and avoid future crises. Assessing this technology and these systems to learn their contribution to the managing of change in the environment as appropriate is the purpose here. Learning what the appropriate use is will come about only if the situation of the region is well understood and a community dialogue is begun on the Growth and Quality of Life desired there.

REPORT COMPREHENSIVENESS DISCLAIMED

It is important both to the integrity of the 25 graphic-related efforts described here, and to the integrity of this report itself, to disclaim comprehensiveness. This is a preliminary survey, the matter of 25 contracted man/days of work. As such, it does not claim to have identified all capabilities, to have examined those included here in depth, nor even is it claimed they are reported on with great precision, despite the care that has been given.

The 25 different efforts are grouped for convenience, a little arbitrarily. They are described in varying degrees of detail, which in itself is not intended to indicate how they are valued. The objective is to identify certain analysis/synthesis/communication tools and their potential for regional decision-making. Hardware-software systems that for complete understanding require complete reports in themselves are "covered" in a few paragraphs, and then "compared"! (By contrast, NSF is undertaking an extensive 2-3 year review and comparison of the TICCIT and PLATO systems alone.

The utility of this part of the report is that it locates resources, and provides general [illegible] available and evolving. This report was undertaken from the perspective of the regional policy-maker, not the technician. It is written for policy-makers to help orient them to new tools for their work. The risk in such an attempt is the opportunity for misinterpretation. From the detail in which each particular effort reported on here is known at its source, this may have varying degrees of importance. Indulgence in the form of written critique is invited to correct inadvertent inaccuracies or oversights. Consider this copy as one for colleague review, not possible before within the scope of the work as undertaken.

To provide access to all resources referred to, addresses and phone numbers have been given.

II THREE CONFERENCING AND COMMAND-CONTROL SYSTEMS

NASA

National Aeronautics and Space Administration
Washington, D. C. 20546
(Egon E. Kafka, Chief, Skylab Program Planning and Control)
(202-755-3144)
(John Arslanian, TV and Visual Display Manager, Operations Center Branch)
(202-982-6137)

Certain management, monitoring and mission-control operations of NASA provide ready access to volumes of data and information for assembled working groups of people. NASA's means of handling operations for that was examined for what it might have to teach concerning "real time" access to volumes of information for regional policy-making. Two Management Information Centers in NASA's Washington headquarters were visited: SKYLAB and SPACE SHUTTLE, and, in addition, the Goddard Space Flight Center, Combined Mission and Network Operations Center, in Maryland. (DOD "war rooms" provide graphic and computer assisted display, but these are classified security operations. It is assumed that the NASA facilities at Goddard and Houston resemble certain aspects of DOD installations and have available to them much of the same technological knowledge.)

The simpler of the two types of NASA facilities are the Management Information Centers. Perhaps their singular characteristics are reliance on overhead projector graphics, and high fidelity 4-wire dedicated telephone lines interconnecting centers in different locations of the country. Fidelity is said to contribute significantly to voice recognition and ease of exchange between individuals assembled in groups in two or more centers. Before the conference call coded overhead projector slides are delivered to each center which is equipped with two overhead projectors, slide projectors and movie projectors. In addition, back lighted 4' x 6' translucent plexiglass management charts on sliding tracks line one wall.

The Space Shuttle Management Information Center is 20' by 25', carpeted, with a 12' table seating 23 people; the maximum recommended room capacity is 35. Besides the equipment referred to above, there is a standard speaker phone for back-up, and a Magnifax transmitter/recorder is available for facsimile transmission. (The new model can transmit or receive at the rate of one 8 1/2 x 11" sheet of copy in 3-4 minutes.)

NASA has established simple conventions for color, legibility, height-to-width ratio, symbols, and coding of its graphics. It has standard routines for calling conferences and running them, backed by the kind of staff and equipment it takes.

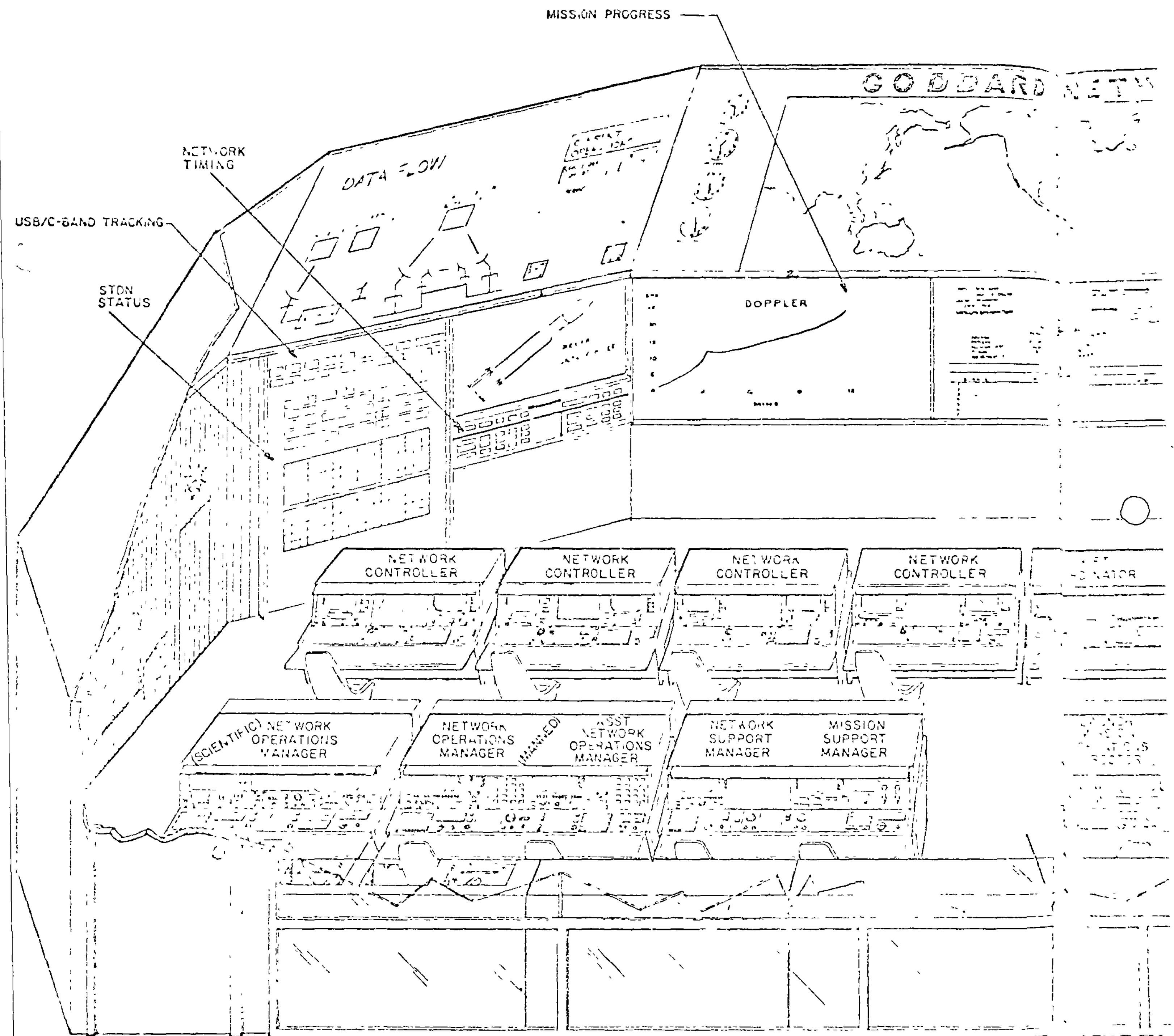
The Operations Control Center at Goddard, on the other hand, is a spectacular. (Even more so is the Manned Space Flight Center at Houston.) The fold-out drawing attached shows the Goddard Center. It is a dimly lit three-storey room with 30 computer assisted monitors and large graphic wall displays, some rear projected displaying digitized information, and with the capability for slides, movies and live television. The need for this display on one wall is to provide a flow of information that is available at one time for all the controllers, managers and directors assembled in the room, monitoring all NASA non-manned space flights simultaneously. The individual display terminals, each separately manned, provide access to specialized areas of information as called for. Looking down into this room from glass-walled roofs from the same perspective as the drawing, are NASA project directors and upper management people.

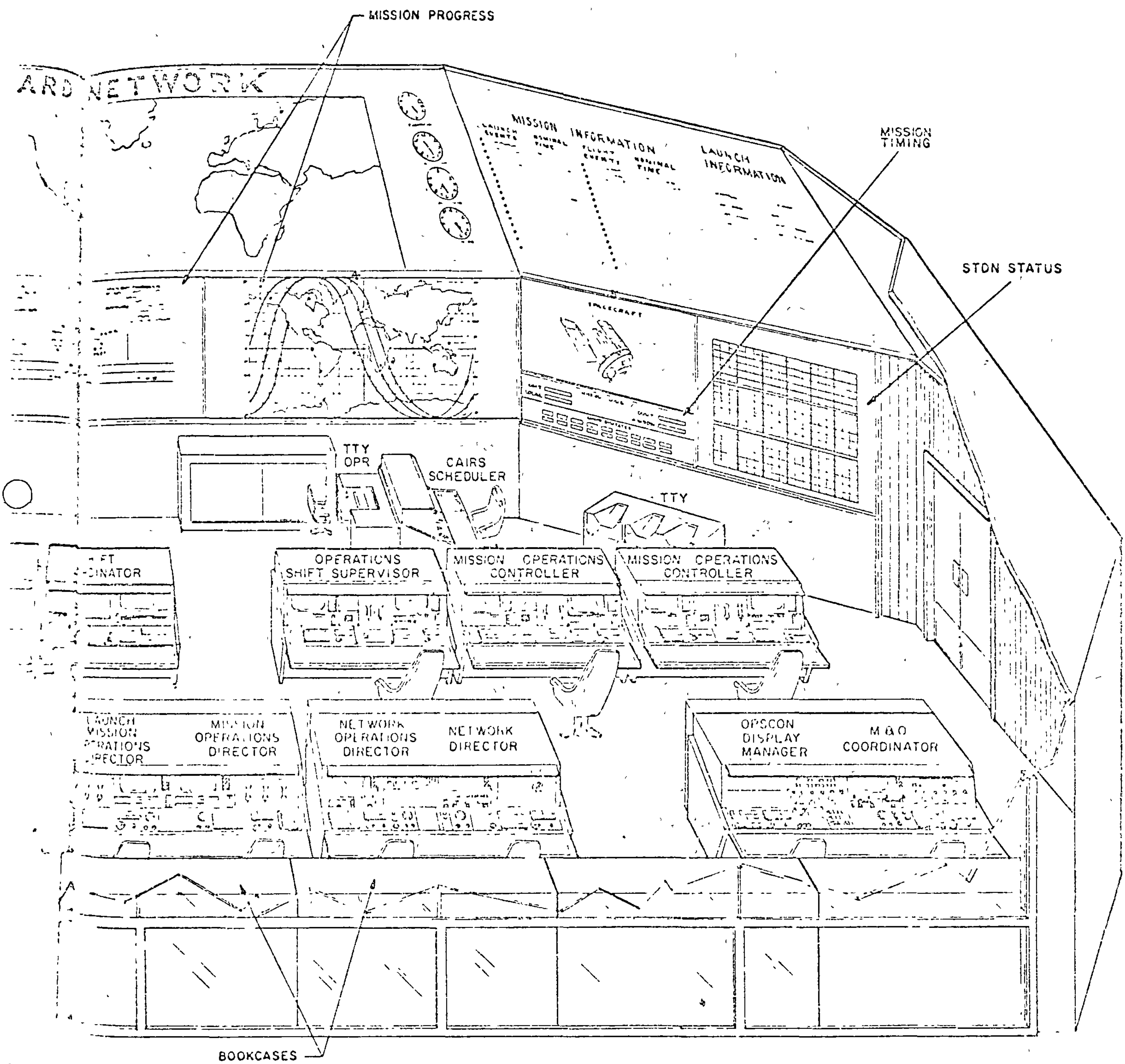
Behind the long display wall is an area for rear projection equipment that seems as large as a theater stage. The basement level is lined wall-to-wall with computer equipment for the display terminals upstairs. The host computer is IBM's largest, the 360/95. The back-up computers are an IBM 360/75 and an IBM 360/65.

Characteristics of Significance - Great technological capacities were developed to support the space program - certainly one of the great success stories in the history of man. But what can this space technology contribute to man's life ON EARTH? Surveillance of the earth by satellite - generating information of ever increasing detail, accuracy and importance - is one example that can be easily understood. NASA has programs to facilitate such technology transfer. Perhaps some aspects of the management of NASA have of themselves a contribution to make. That is what this survey looked into. There do seem to be important similarities between NASA's managing of data and information and the management concerned with regional policy-making.

Dissimilarities too are important to recognize. Great as NASA's space program's successes were (and are), NASA's mission is by order of magnitude(s?) simpler than managing the conservation and development of regions. NASA is a one agency program. NASA programs deal with a fixed number of quantifiable variables. NASA's conflict or trespass over territorial rights is essentially trivial (although tracking stations worldwide are vital). NASA's mission is distinct from, and clearly not competitive with, other on-going operations. NASA's mission was launched with top national priority to "beat the Russians". NASA was well funded under programs that were necessarily committed to a long term before achieving ultimate results (8 to 10 years).

By contrast, regional policy-making is without a constituency. There is a collection of competing public and private entities, not one regional policy-making authority. Regional policy-making must deal with a vast number of variables; some unknown, some unknowable, some inadequately quantified, some not quantified at all. One way or another, almost all regional policy-making has some bearing on the geography of political jurisdictions and


MISSION PROGRESS
NETWORK TIMING
USB/C-BAND TRACKING
STDN STATUS
DATA FLOW
DOPPLER
NETWORK CONTROLLER
NETWORK CONTROLLER
NETWORK CONTROLLER
NETWORK CONTROLLER
(SCIENTIFIC) NETWORK OPERATIONS MANAGER
NETWORK OPERATIONS MANAGER (MANNED)
ASST NETWORK OPERATIONS MANAGER
NETWORK SUPPORT MANAGER
MISSION SUPPORT MANAGER


MISSION PROGRESS
NETWORK
MISSION INFORMATION
LAUNCH INFORMATION
MISSION TIMING
STDN STATUS
TTY OPR
CAIRS SCHEDULER
TTY
OPERATIONS SHIFT SUPERVISOR
MISSION OPERATIONS CONTROLLER
MISSION OPERATIONS CONTROLLER
MISSION OPERATIONS DIRECTOR
NETWORK OPERATIONS DIRECTOR
NETWORK DIRECTOR
OPSCON DISPLAY MANAGER
M & O COORDINATOR
BOOKCASES

property rights. Regional policy-making is related to on-going policy as it is already being made in the region; the business of making policy for an entire region is clearly not distinct from nor free of confusion with existing institutions. Going through the above, it is easy to see why there is no national, state or local consensus on regional priorities for policy-making. Regional policy-making is not well funded. There is no long-term commitment to it.

Making policy for regions is not exactly like making decisions in NASA, but at least there is a similarity to be found in the need to know - a lot - accurately and quickly. NASA must have at its disposal the capacity to store and retrieve, analyze and synthesize, vast quantities of data and information in order to manage its space mission. That has to be "real time" information for monitoring and management purposes in the Operations Control Centers. Availability and accuracy are essential. The information has to be credible and it has to be comprehensible. With it all, NASA has recognized and utilized ways to display information that is vital to the work in both its Management Information Centers and its Operations Control Center.

Regional policy-making, too, requires access to vast quantities of data, information, analysis and synthesis which is credible - and comprehensible. This is not needed by policy-makers in "real time" for the purpose of actually managing regional programs. The "real time" nature of the information needed is due to the constraints on the time of the policy-maker, and those with whom he interacts in policy formulation and making decisions. The complex nature of a region and its many variables and their interrelationships would be better understood, it would seem, if they could be analyzed, synthesized, displayed and understood in unanticipated combinations, quickly. The regional situation room itself being considered here, therefore, is "operational" - "real time".

The technology developed for the Department of Defense and NASA policy formulation and decision-making would seem to offer tools to help manage change in regions that has not yet been thought through and tested - especially with concern for graphics - and with implicit recognition of the actual quality of data and complexity involved.

EMISARI

The Office of Emergency Preparedness
Executive Office Building
Washington, D. C. 20504
(Murray Turoff - 202-395-5143)

The Office of Emergency Preparedness (OEP) has evolved EMISARI, an interactive program for conferring with its U. S. regional offices, and Washington headquarters, via terminal keyboards that print hard copy at the rate of 30 characters per second. The value of this method of calling conferences has proven itself since Phase I of the Administration's price control program. A chart of EMISARI's efficiency claims that 10-20 people at separate terminals - even if in the same building, much less the difficulty and expense of convening them from across the country - are more efficient than all 10-20 assembled for a typical conference in one room.

Computer based group communication permits simultaneous "talking" (typing) and "listening" (reading) as conferees choose. All remarks, additions, corrections, etc. are stored by the computer and made available in hard copy form as needed.

OEP uses a UNIVAC 1108 and has evolved its EMISARI program needs over time with its users in the form of a management system into which entries can be made at any time. Anyone using a connected terminal can update himself on information separately stored in "areas" informally labeled bulletin board, notices, policy, activities, news (clippings), public information, special people file, tables, explanation, estimates, messages, letters. Keyword searches for this information can be made using computer language BASIC.

<u>Characteristics of Significance</u> - The UNIVAC 1108 can handle up to 40 such terminals via normal telephone lines with EMISARI. The user's instruction "manual" consists of two diagrammatic pages of computer commands. It is quickly mastered and policy-level persons find themselves typing at terminals.

A basic point made by Murray Turoff, who explained this management system to me, seems especially pertinent to this investigation: the system evolved under the direction of the users themselves. It wasn't a matter of a year's delay while a consultant went off and developed the special hardware/software combination needed. That was seen as too theoretical, not only because taking a year off before employing this new capacity for Phase I was not possible, but because both the special needs and "comfortableness" of each user needed to be established before there <u>was</u> a system that would work and be used.

NIPS - National Military Command System Information Processing System

IBM
1601 North Kent Street
Arlington, Virginia 22209
(John R. Gerr, Manager, NIPS projects - 703-524-7068)

Federal Systems Division
18100 Frederick Pike
Gaithersburg, Maryland 20760
(Fred H. Badger, Marketing Representative - 301-840-7520)

NIPS is a formatted filing system program developed in 1958 for the Defense Communications Agency and only recently declassified. It is an advanced data management system that works with IBM System /360 computers and is compatible with the IBM 1410 Formatted File System that preceded it. NIPS (or NIPS 360/FFS) provides the ability to structure files, generate and maintain files, retrieve information, and output that information in simple or complex arrays. Our interest here is in the capability for structuring files by geographic coordinates and graphic spatial display of that information after what IBM refers to as "exhaustive diagnostics."

Quoting from the National Military Command System Support Center's Computer System Manual, CSM GD 15A-68, 1 July 1971 (NIPS 360/FFS) Section B, General Description, page 29:

> NIPS exemplifies the heavy-duty file processor which has been the mainstay of the Department of Defense command and control and intelligence data handling. In its current version, this system incorporates a comprehensive on-line capability which further enhances its adaptability to today's processing requirements. The evolutionary approach continues to be the foundation of the system's development. The validity of this approach is increasingly apparent as the number of users and applications increases. For application areas with high-volume and large file-processing requirements, NIPS provides a convenient, efficient, and flexible method of solving the data handling problems in the third generation hardware and software environment.

NIPS is the largest, best established and staffed data management system. It has been refined, evolved, and "de-bugged" for over five years. (The Census Bureau's DIME geo-coding information system, in contrast, is a relatively recent entry and has not had the staffing or de-bugging needed to perfect it.) At the present time, Computer Science Corporation is re-writing NIPS into the computer language COBOL. It will be ready in four years. Undoubtedly, in keeping with DOD policy to date, this will again be an evolutionary advancement and will be compatible with the totally new IBM computer line that, it is said, will be marketed beginning in 1976.

NIPS was designed for the large file capacity needs of DOD, to be able to call up and display information coded on a geographic basis for any given area in the world. Records of 100-150,000 and more are handled. It can more readily manipulate 20,000 and with 5,000 records its performance is outstanding.

Recently, the Department of Transportation's Office of Systems Analysis & Information has examined and tested NIPS for application to its transportation planning needs. Basic publications it has generated in this process include:

> NIPS 360/FFS - AN EVALUATION - Final Report, December 1972;
>
> A SURVEY OF NATIONAL GEO-CODING SYSTEMS, February 1972;
>
> THE NATIONAL GEO-CODING CONFERENCE - Proceedings, May 1972.

Inquiries concerning NIPS civilian applications might best be referred to the Chairman of the Formatted File System (FFS), Commercial Users Group: John Bright, Western Electric, P. O. Box 20046, Greensboro, North Carolina, 27407 - 919-697-3370.

Characteristics of Significance - One doesn't have to read far into the subject of geo-coding before questions are encountered, such as "Should there be one super-system or should interchangeability between a large number of national geo-coding systems be developed?", "What data should be collected and at what geographic scale?", "How is sufficient consistency achieved and for what users?".

NIPS probably offers the most perfected basis to date of geo-coding, storing and manipulating data for the purposes of regional policy-making. Data for specific geographic areas can be searched according to the characteristics of a given problem. Or, according to criteria, geographic areas delineated. The output can be either alpha-numeric or graphic display.

III FT'D REGIONAL APPLICATIONS

VANCOUVER, B.C. - Regional Simulation Study

Resource Science Center
The University of British Columbia
Vancouver 8, Canada
(Robert F. Kelly, Project Coordinator - Gil Evans - 604-228-3131)

The Inter-Institutional Policy Simulation, or IIPS, is a joint project of city, regional, provincial, and federal governments and the University of British Columbia. IIPS is a five-year project begun in June 1970 with the assistance of a $500,000 grant from the Ford Foundation (believed to be less than half the total cost now.)

IIPS' object is to get people working together to build a model of the Greater Vancouver Region using mathematics, logical concepts, the University's IBM 360/67 computer and a project-purchased $47,000 analog computer. (The analog computer allows researchers to simulate a total environment and the digital computer to handle the mass of mathematical data that the new program will generate.)

The first three years work has been an analysis of the region, developing the overall program, and its ten submodels, and the gathering and coding of data. The fourth and fifth year will be spent in testing and refining these models. "By 1975 it is hoped a working model will be available for people to test the possible consequences of alternate policies which would affect the future of the Vancouver area." Though staffed predominately by university researchers and students, and coordinated with governmental agency staff, it has been the hope of IIPS from the outset to provide a "futures" testing vehicle for citizens, politicians and civil servants alike.

IIPS is intended as an early warning system for the region. It is intended as a way to raise alternative choices and test for their consequences. Basically, IIPS is a sophisticated way to learn how to ask better questions about the working of the vast complexity of interrelationships in a metropolitan region of one million people. It is not an automatic decision-maker.

"To operate the IIPS model, one might sit at a keyboard and type in one's questions, ideas and pet policies. The consequences in the simulated world of the model would be shown on a screen in the form of charts, graphs or printed words. Computer terminals are planned at several points in the region ... IIPS strength will be in allowing people to preview the possible outcome if their pet theories were implemented. Its limits are that it will not answer questions in detail ... IIPS will not assume a real planning function; people in the region will do the planning, using IIPS to help understand the complexities of the whole region."

IIPS has developed ten submodels that all react with each other to simulate the functioning of the region. The total model is a complex simplification. It is not reality. It is too simple for that. But it is complex, although intended to be simplified enough to be understandable (by examination) and capable of handling available data.

Submodels of the IIPS are:

- population and demography
- economics
- transportation
- health systems
- pollution
- human ecology
- land classification
- data management
- resources and public services

The submodels are linked by the flow of information from one to another. - "By making certain that each of the submodels provides the information needed by the other submodels, we are ensured that the information flows will be complete and will closely imitate the real flows in the Vancouver Region."

Characteristics of Significance - The following is another quote from the first of the four papers referred to at the end.

"The Vancouver Regional Simulation Study is a bundle of paradoxes. It seems like an exercise in numbers to produce a model - and yet it is not. It seems designed as a service to the bureaucrat and technician - and yet it is not. It seems to assume that the quantifiable variable is by definition the important one - but in fact it does not. Rather, its central purpose is to provide an environment for the institutions and citizens of the Vancouver region to develop a dialogue about the alternate futures open to the region.

Effective dialogue, however, can only be developed if there is a common ground of substance that can trigger and focus the dialogue. We have, therefore, designed a programme in which the first steps were largely technical - basically to develop a simulation model of man/environmental interaction in the urban regional setting of the Lower Mainland of British Columbia. By inter-relating as much as is known of the economic, social, physical and environmental processes in the region within a model, with the unknown and the qualitative outside the model, it was hoped that regional problems could be more explicitly identified and placed in an objective environment for community discussion. In this sense, therefore, that part of the world that can be simulated could become a powerful instrument to explore the consequences of different assumptions and policies, any one of which can generate an alternate future.

We know from experience that even modest simulation modelling efforts to interrelate parts of a system confront the person entering the simulated world with paradoxical and unexpected results. The human reaction to the unexpected and to the paradox is to ask a question, and if there is anything we need now, it is an environment for asking better questions. The key of our approach, therefore, has not been to design a model that will produce unique solutions and in itself specific policies, but rather to produce a process by which the institutions and citizens of the region could pose better questions."

IIPS has discovered that challenging and significant as developing the simulation model is, developing a framework for its responsible use is much more so. "It is this latter challenge that makes the project a fragile and groping experiment."

The University base for the project was vital for this model development (UBC happens to be a university more blessed with a predisposition to interdisciplinary work than most.) But the ability to construct models and deal in conceptual abstraction which is possible in such a place, may permanently outstrip the comprehension of the ultimate user -- the government official and the private citizen. Newspaper accounts, write-ups by University participants and discussions with them indicate the necessity to involve people more effectively from the outset. Citizen understanding and involvement is only now being urgently sought.

"By handling the complex technical detail, the model can free people to concentrate on the non-technical issues - questions of goals, values and the ingredients that define quality of life as each individual perceives it. What is really important, therefore, is not what is put into the model, but what is kept out: not just how to provide a mechanism like a model to handle the quantitative, but also how to design a decision-making framework to handle the qualitative; not just to develop a useable model, but to determine how it is used. The greatest challenge to the project, therefore, is whether, in cooperation with the citizens of this region, technology can be harnessed to man's needs."

"There is also the hope that a way can be devised to allow the transference of the activity to different institutional auspices. With so many institutions involved, so many constraints, and such historically ingrained institutional patterns of behaviour, this is a non-trivial job. It is, in part, what IIPS is all about. How can institutions with such different goals, with individuals of such different motives, come together and act in a cooperative spirit to a common purpose when they have for so long remained separate?"

See these publications which are liberally quoted for this report: Vancouver Regional Simulation Study, 1970 - 1971, M.A. Goldberg, C.S. Holling, and R.F. Kelly; Vancouver Regional Simulation Study - Second Year Report 1971 - 1972 - Section "A" Overview of Project; IIPS Detailed Description of Sub-Models and IIPS, Resource Science Centre, University of British Columbia, Vancouver 8, B. C.

DES MOINES - IMIS - Geoplanning

Integrated Municipal Information System
City Hall
Des Moines, Iowa
(William Batske - Battelle - Northwest, consultant
James Furst - City Planning Director - 515-283-4141)

This fifteen month research and development program, for which the City of Des Moines is the prime contractor, is a project of the Urban Information System Inter-Agency Committee (USAC). USAC is composed of ten federal agencies and departments led by HUD. USAC is assigned the responsibility for refining the objectives of a research program in urban information systems which:

(1) performs research, develops, tests, evaluates and documents systems and sub-system prototypes;

(2) develops prototypes which are transferable to other municipalities;

(3) includes readily comparable elements of data; commonly defined; and

(4) operates at favorable cost/benefit ratios.

The Des Moines project IMIS, for Integrated Municipal Information System, "focuses on the development of geocoding techniques and system methodology in addition to development of a sophisticated Geographical Information, Planning and Analysis System. Using a Master Base Map and Grid System, Map Model System and the Formated File System, the city of Des Moines hopes to demonstrate that municipal data or urban environment data can be easily and effectively integrated with specifically designed geographic structural file."

Des Moines is a metropolitan area of about 200,000 people. For the purposes of this demonstration, a pilot area of about 21 square miles in the southeast quadrant of the city was selected. A Des Moines/Polk County Program/Users Committee of 25 plus was established for the project. The long range objective is to provide data/information geographically coded for the entire region.

Functional areas of city and county government which showed interest in geographic coding and display of information included: - Plan and Zoning Commission, Traffic and Transportation, Community Development, Public Works, Finance, Building Inspection Services, Urban Development, Assessor, County Audits and County Treasurer. To limit the data/information for the pilot area only four agencies were selected: Community Development, Planning and Zoning Commission, Public Works and the Assessors.

Des Moines had already embarked on a major data/information system development program before IMIS. It had acquired, and has since updated, its computer capability, and now has a number of computerized municipal data files. Ultimately the various subsystems could comprise a total integrated system with the ability to cross-reference data between various on-line files. "Many Des Moines agencies need and collect data, but their resources and methods for gathering and maintaining their data vary. The system concept developed by Battelle, GEOPLANS is designed to recognize these limitations.

"~~GEOPLANS is the acronym for GEO graphic PLanning~~ and ANalysis System. ~~It is a system for effectively relating vast stores of~~ data to geographic locations. It provides better and more timely information for more effective decision-making." (CEOPLAN uses the NIPS system referred to in Section II.)

Des Moines is experimenting with "a geographical reference identification method for data storage and retreival and for linking various data source files. GEOPLANS' common identifiers are directly applicable to the needs of spatially oriented planning. This spatial, geographic orientation can be (mapped) State Plane coordinates, legal boundaries, street addresses, or artificially created grids or polygrams."

The Des Moines pilot project is being tested this summer. Publications available for review of this approach include, Geoplanning Research Program, System Conceptualization, Des Moines, April 1973, USAC-DM13-0002, which includes an excellent appendix paper on the state-of-the-art of geocoding technology and urban data systems.

See also GEOPLANS — a Geographic Planning and Analysis System by Battelle Memorial Institute/Pacific Northwest Laboratories, Richland, Washington 99352.

Characteristics and Significance – A primary characteristic of the GEOPLANS design is that while it involves the application of modern automatic data processing, testing it does not require exotic computer equipment or radical changes in current procedures. Present computer equipment, with adaptations, will do. Current agency responsibility for updating and correcting data remains. A common geographic reference index relates data files in different agencies to each other, and is used to call forth data.

The potential appears to be enormous for this system; much more so than previous thoughts of creating one massive data bank for all required regional information. But accurate data base maps are essential for the GEOPLANS system. (New York City is said to have 10,000 different maps and no one base map. The New York City [illegible]ning Commission decided that plotting and digitizing coordinates (NIPS-style) of the various polygons of city blocks, zoning, etc., for so many different base maps – most of which are not compatible with each other – was too slow, too tedious to attempt. Des Moines uses aerial photography for its base map, which was easily controlled and could be digitized readily. Most information reads well at such low density.)

The New York example is brought in here for the purpose of introducing computer graphic scanning as the means of plotting base maps and data. This is now being tested by the New York City Planning Commission (2 Lafayette Street, New York City, Rolf Moulton 212-566-3982). Instead of devising a system of digitized coordinates for the millions of polygons that are involved in New York City, maps are scanned by computer for which programs covering discrepancies have been written. By this much faster means, graphic map information can be input to computers in combination with hand work corrections on the maps. This approach can be used to provide "real-time" maps which integrate basic data.

The important point here is to recognize that the practicality of computer assisted graphics is that (1) one master-base of all data is not needed, but one geographic reference index is and (2) optical scanning by computer may become the most efficient means to input geographically based data into a coded computer format.

SAN DIEGO - IREM

Integrated Regional Environmental Management Project
County of San Diego
Environmental Development Agency
1600 Pacific Highway
San Diego, California 92101
(L. Edwin Coate, Director - Larry Taylor - 714-236-2005)

In 1971, San Diego County consolidated various environmentally related functions, including regional planning and community zoning, into a new Environmental Development Agency. It reports to the county's Chief Administrative Officer who is responsible to the County Board of Supervisors. "Rapid growth and land use planning with insufficient attention to its implications for air quality ..." were cited in a recent report by Director Coate (see first publication reference.) It seems San Diego has 60% more days than Los Angeles when weather conditions could lead to a serious air pollution situation.

The Ford Foundation in April 1971 came in to assist this new effort, to help determine what environmental management is and how effectively it could be accomplished on a regional level by government. A two-year $725,000 grant was made to establish the IREM project staff, located in the Environmental Development Agency. It operates with 30 people, consultants, university contracts, and other governmental agency staff. Both EPA and NSF also came in with additional support funds.

"The IREM project's purposes were to work within county government to:

1. respond to Federal and State legislative requirements;

2. mobilize community resources and support decision-makers;

3. respond to citizen's concerns as articulated by the Board of Supervisors;

4. enhance the regional image, and

5. provide a rational approach to environmental issues."

The program divided itself in two. One section provided regional environmental service such as project environmental impact analysis, community involvement and economic analysis. The other part of IREM was directed to program management which included policy development and technical management of research, development and demonstration projects.

A major effort of IREM was to develop a natural resource inventory for land use and decision-making, and computer graphic techniques were developed by IREM for regional land use projects. Staff and functions defined by the IREM project are to be absorbed by San Diego County.

Characteristics of Significance - To quote further from Director Coates:

"From the IREM case study, it can be concluded that, to be most effective, regional environmental management must relate agencies involved in land use and transportation planning to those which are responsible for regulatory pollution management functions. An effective regional environmental management agency must have linkages to all the key environmentally related institutions in the region. The IREM experience also led to the conclusion that effective regional environmental management had to be carried out with some type of authority or under a definite mandate. Advisory functions and research and development functions were necessary and important, but proved inadequate by themselves. IREM, located as it was in a County government, could not effectively fulfill all of these various criteria.

"The regional concept for environmental management is not only valid; it is essential. We have finally begun to question the efficacy of institutions that deal with environmental quality at all governmental levels. The resulting analysis leads to the conclusion that a new type of regional institution must be created."

See publications: County of San Diego, Regional Issues, Volume 3, Environmental Information System; Basic Concepts, County of

San Diego, Environmental Development Agency, IREM (August 1972); The Coastal Plan of San Diego County, Laboratory for Experimental Design, California Polytechnic, Pomona, for County of San Diego, Environmental Development Agency IREM (August 1972); Geographical Data Handling (Volumes I & II), Edited by R. F. Tomlinson, Symposium Edition, International Geographical Union Commission on Geographical Data Sensing and Processing, Second Symposium, Ottwa, August 1972; A Study Program Organization and Operation, ERMS, San Diego County, IREM Program, December 1971

COLUMBUS, OHIO - Benchmark

The Academy for Contemporary Problems
505 King Avenue
Columbus, Ohio 43201
(Ralph R. Widner, Academy Director - 614-299-3151)

The Academy for Contemporary Problems was established in 1971 as a joint non-profit venture of Battelle Memorial Institute and Ohio State University. Each is committed to $500,000 per year for ten years, and Battelle in addition has financed the construction of a $2,000,000 complex to house the Academy, which will be occupied this Fall.

One of the Academy's first areas of concern is an activity to be supported in several metropolitan areas entitled BENCHMARK. It stems from a Resources for the Future committee position that "too often decision-makers' conception of metropolitan reform and the problems perceived by the metropolitan constituency have been widely disregarded ... BENCHMARK is an attempt to provide continuing data concerning the problems existing and emerging in a metropolitan area, to assist metropolitan decision-makers in making more effective public policy ..." The original test area for this work will be the Columbus metropolitan area which has a population slightly in excess of one million persons.

This public policy research effort will be identified as the Columbus Area Social Profile, or CASP. It "will be a social system -- an organization comprised of research performers, community leaders, neighborhood groups, and individual citizens ... CASP seeks to serve the short-term information needs of public and private organizations and voluntary associations in the Columbus area by providing reliable data on public opinion, attitudes and practices. Data on basic social, economic and ecological conditions will be related to the needs and aspirations of people living in the Columbus area... CASP intends to service a variety of community needs related to self-knowledge and hence to self-government."

CASP is designed to integrate regional information from the wide variety of sources continually providing it. CASP will attempt to provide a comprehensive framework for better understanding issues and resources. To aid in this, CASP is a private social profile of institutions and neighborhoods within the region. It is intended to provide users with a new instrument to audit, over time, trends, aspirations and satisfactions.

The principal goal of CASP is "to provide an informative base that will be maximally useful for diagnosing social strengths in the Columbus area." It has tentatively established goals from which more specific objectives will be derived and pursued with explicit actions. CASP is preparing to expand its preliminary organizational form, inventory institutional and individual resources in the area of value to its mission, gather data, develop an archive, identify issues, etc. Projections of alternative futures for the Columbus area are to be prepared leading to selection of policies. Both integration and evaluation are integral aspects of CASP.

<u>Characteristics of Significance</u> - In the Academy's "mind" is Harold Lasswell's conception of a "social planetarium" for graphically conveying CASP's sort of information for the purpose both of explication and analysis for decision making. Taking a paragraph from a recent Lasswell paper for the Academy:

> If decision-makers are to arrive at a critical assessment of the factors that account for the spreading or the restriction of nations, corporations, and doctrines, they need access to whatever scientific knowledge there is. Now it is no simple matter to apply past laboratory findings to the interpretation of future developments. In the laboratory, conditioning factors are controlled. In weighing the future, the task is to foresee the cluster of factors whose occurrence will influence the outcome. An advantage of the planetarium procedure is that scientific knowledge can be exhibited and employed critically in estimating the probable course of future developments.

There is particular strength in the Ohio State - Battelle-based Academy. It is Battelle's geo-coding capability that is built for the Des Moines region; reported elsewhere in this section. Battelle, one of the world's largest research firms, has a wealth of other hardware/software talent, and in Widner, an experienced regional development executive. At the Ohio State there is an array of talent which includes computer-graphics skills already capable of computer assisted animation.

See publication: BENCHMARK/CASP - A Public Policy Research Effort to Help Improve the Governance of Metropolitan Communities Through Improved Use of Social Intelligence, Mershon Center for Public Policy, Academy for Contemporary Problems, 1973.

NEW YORK - RPA - TV Town Meeting

Regional Plan Association
235 East 45th Street
New York, N.Y. 10017
(William A. Shore, Vice President - 212-682-7750)

With HUD, foundation and other support totaling about $1,600,000, RPA experimented in the Spring of 1973 with an original filmed series of one-hour films for broadcast television on the theme Choices for '76. (Regional Plan Association was started in 1922 by the Russell Sage Foundation. It is probably the most respected volunteer citizens, non-profit planning association in the United States and has pioneered in regional planning.)

The six television programs covered transportation, environment, poverty, housing, cities, and government. Each program was used three times in a series of three-day weekends on each of eighteen television stations in New York, New Jersey and Connecticut, from March 17 to May 14, 1973. A concerted effort was made to organize citizens around television sets in advance to participate in this new form of Town Meeting. They were provided with ballots to "vote" on specific issues as posed during each program.

(A paperback book, How to Save Urban America, (Signet), 1973, was published to coincide with and be used by this series. It was available on newsstands throughout the region.)

From its June, 1973 Newsletter, RPA reports:

"Who took part? The participants had more education and higher income than the Region's average, despite strong efforts to recruit those with an average and below average income and education. Those earning $8,500 - $13,000 were represented in about proper proportion, but those below were very underrepresented, those above overrepresented. A third of the Region has an income above $13,000 while two-thirds of the ballots came from such people. About half the ballots came from persons with college degrees, but only one person in eight over 21 in the Region has a college degree. Black and Puerto Rican votes were 5% - 8% of the total, while they constitute nearly 20% of the Region's population (though special ballots distributed by the Committee on Minority Affairs will add to those received already). As to age, 30 - 44 year olds were overrepresented, over 65 year olds were underrepresented. Inner suburban counties were overrepresented; city and outlying counties generally underrepresented. However, on some two-thirds of the issues, a majority of people (or all but a statistically minor category) voted in favor of policy changes regardless of their age, race, income, educational background or county. Those issues are starred in the vote summaries (not included here).

Participation:

	Watched TV	Returned Ballots
Housing	2.95 million	47,500
Transportation	2.0 "	32,200
Environment	1.75 "	22,500
Poverty	1.5 "	18,500
Cities/Suburbs		12,500

What Do the Votes Mean? We cannot say that the votes tell where the people of the Region now stand, but we can say of the starred issues that people who stop to consider them are ready to support proposed remedies. This suggests that elected officials willing to exercise political leadership probably can achieve policy changes.

What's Next? Regional Plan will report results in public meetings and via mass media. We will get results to officials and candidates for public office and ask their comment. Three of the major candidates for Mayor of New York City have commented already. See New York Times, May 12, 1973. We will try to bring together people who want to take action (on any of the issues) with organizations working on their side. We will try to assist civic action on behalf of a few issues that seem ripe for resolution. At the same time, Regional Plan will prepare a thorough evaluation of the project both to understand its implications and to assist the many groups in other urban regions now considering the same kind of program."

Characteristics of Significance - RPA is now surveying the participants and will be reporting on the effect of the Choices for '76 television series in the Fall. It is significant that 1.5 - 3 million people of the New York region's 18 million took part. Even though not represented in their true proportions, all groups tabulated were large enough in number to provide a scientific sample for cross tabulation by age, income, education, etc.

It appears clear from examining the ballots processed so far, - seven to ten multiple choice questions were asked with each program, - that a majority of viewers seek major development policy changes to achieve better living in the region.

IV NINE SYSTEMS OR COMPONENTS

TICCIT - Computer Controlled Information Television

The Mitre Corporation
1820 Dolly Madison Boulevard
McLean, Virginia 22101
(Richard Morton - Kenneth J. Stettin - 703-893-3500)

TICCIT (for Time Shared, Interactive, Computer Controlled Information Televis on) is a two-way interactive cable television system. It has been funded by NSF over the past 5 years. MITRE's emphasis is on educational and other non-commercial applications of interactive computer assisted communication. It has two developments underway at this time. One is to be a home-use demonstration available to 3500, utilizing Touch-tone telephones and the cable system of Reston, Virginia, beginning in September, 1973. The other is a fully operational demonstration in two community colleges beginning in September, 1974, each with 128 student terminals consisting of TV receivers, keyboards and headsets.

The student terminal uses a standard color TV receiver, head phones and a keyboard for instructing the Data General Nova 800 mini-computer. The home subscriber will use a Touch-tone telephone and a standard color TV receiver to begin with. (Later a twelve or sixteen button keyboard, or a typewriter keyboard will be offered.)

Both configurations offer color TV displays under computer control, letters and numbers and line graphics, in seven colors, and full color movies. Up to 17 lines of 41 characters each may be displayed, 512 distinct characters being definable at any single time. Graphic displays are constructed from straight-line segments drawn on a grid of 200 elements in a vertical direction by 256 elements in a horizontal direction. The color of each character and line segment may be individually specified. Five minutes of full color movies are available per hour. Computer generated graphics are still (not moving) and the audio responses are expected only 10% of the time, varying from one to 40 seconds.

A motivation behind TICCIT is to explore non-commercial applications of computer assisted interactive cable television versus its commercial applications. It relies on already available commercial equipment and is now preparing its development testing phase.

A major impediment for home interactive television has been the need for a "refresh memory" with each terminal. (The refresher memory is the means whereby the 1/60th of a second still images directed by the computer to a terminal are "held" by being repeated ("refreshed") at the particular user terminal until released by another instruction.) This is said to be reduced to one refresh unit per 20 subscribers by assuming that in a normal 10 hour day, interactive television would be used by each subscriber only 30 minutes. (Later, when the price of refreshers drops from $1,000 to $200 or $300, heavy users might acquire their own refresher allowing them to bypass peak hour queues.)

TICCIT is a realtively low-cost system -- a complete system of 128 student terminals with mini-computer will cost $450,000 ($250,000 in moderate quantities) -- at student-per-hour contact cost of less than $1. In the home with a Touch-tone telephone, a subscriber rate of $14 per month is projected.

The TICCIT system uses two Data General NOVA 800 mini-computers. One is a time-sharing mini-computer with storage and other peripheral equipment; the other processes communications with the terminals. TICCIT provides interactive information retrieval displayed on the television screen, which is augmented with certain computation capacities. There is a distance restriction of 1500 feet between classroom and computer. In home cable TV application, that distance is limited to normal "head-end" constraints of 10-15 miles.

The school courseware of TICCIT is developed by a team of specialists, for mass distribution, though instruction is self-paced. The TICCIT approach is that of "instructional technology" which contends that the strategy of teaching is separable from the content. In other words, standardized approaches to interactive learning can be developed to handle almost any subject matter. In this, the dominant role of the individual teacher in the learning process is replaced by "packaged" self-paced instruction.

Characteristics of Significance - TICCIT is ready to be experimented with now, where there is cable television and Touch-tone telephone service. (Its most complete demonstration as a system would be where two-way cable television would be in place in 1974.) At this stage, the $150-200,000 Data General Nova 800 mini-computer, refreshers, etc. are needed for each 200 subscribers.

TICCIT's capacity to compute is not expected to compete with PIATO. It is designed for the growing capacity of mini-computers, retrieving data from their inexpensive storage. TICCIT displays color television images, but there are built-in limitations on motion and color in its present configuration. (What fewer TICCIT terminals - for policy-makers - would make possible has not been estimated.) The "drawing" abilities of TICCIT do not equal the "naturalness" of PIATO, but for some purposes still graphics may be enough.

In TICCIT "normal" motion is restricted to that displayed via videocassettes which are loaded by hand at the "head-end" when called for. However, a stop-motion capability is possible. As to color, although seven can be used, they are assigned within a grid system 200 units high by 256 units wide on the standard TV screen. This means continuous tone color of irregular objects is not possible, nor can colored lines which cross one another maintain their single color.

PLATO - Computer Based Instruction

Computer Based Education Research Laboratory
University of Illinois
Urbana, Illinois
(Donald L. Bitzer - D. Alpert - 217-333-6210)

PLATO is a computer based teaching system which has reached a fourth generation stage in its 13 years of NSF supported development. PLATO IV consists of a Control Data Corporation 6000 computer designed to serve 4000 terminals and a software system which includes its own simplified programming language, TUTOR. Response time of the visual displays it either generates (and they can be animated graphics) on a flat plasma glass panel, or from rear projected slides, is .2 seconds. The latter are projected from one of the 256 colored images on microfiche film. (The 4"x4" microfiches are prepared by University of Illinois, color-corrected for viewing through the plasma panel -- which is slightly green in tint.) Audio messages, over 4000 of them, are accessible in .5 seconds from a record with 21 minutes of sound. (The special computer activated record player is $1500.)

The plasma panel developed in the last year or so is 8 1/2" square and filled with neon. It is transparent, flat glass with 512 transparent electrodes traversing it both horizontally and vertically; 262,144 individual digitally addressable points are under computer control -- which may be activated by the keyset. By touching a special overlay panel, 256 points can be activated. (The extra cost of the touch panel is $600.) All manner of line drawings, charts, graphs, etc., different type sizes, styles can be generated, and in motion. The plasma panel permits connecting the computer to terminals by means of ordinary telephone lines.

The software that is used programs concepts which can be interrogated within the limits of the data base and vocabulary related to specific programs. Programs can be developed at any terminal by anyone (including simple ones by children "instantly"). Familiarity with the system's programming language, TUTOR, is said to be a matter of a few hours.

Servicing 4000 terminals on a time-share basis, with computed responses, including the use of "judgment" algorithms, requires a powerful computer.

Characteristics of Significance - PLATO can "draw" its orange graphics on its terminal panels in such a "natural human" manner, in response to queries put to it, that the effect is disarming. The questioner seems to see into the meaning of his question partly because he is able to type in his own questions and partly the way the graphics are drawn. When a map or an illustration stored on the microfiche is wanted, it "instantly" appears, rear projected into the same plasma panel and, if he chooses, he can then "draw" on it with his finger, (special panel, see above) or by means of keyset inputs to the computer.

No doubt mechanisms to store and "instantly" project 8 mm movies as well as slides could also be developed, but television (and video-cassettes) would require a separate television screen. (This is an important consideration and may call for a hybrid system, once mechanisms for computer accessing video-cassettes and selection from within their programs is achieved.)

One terminal ($5500) and its yearly operating costs of $2200, plus long distance line charges, makes PLATO available to a policy-maker (leased long distance charges to Urbana from the West Coast, unless provided on the Federal Telpak Rate of 30¢/mile/month, would be approximately three times that or $1500 per month). Of course, data, programs, microfiches, etc. would have to be developed for the particular use, but as time goes on, a growing share of that would be available from studies done once, nationwide. And in time a PLATO terminal and the PLATO system could be interfaced into the ARPA network, bringing down costs and tapping expertise and computing capacities all over the country. Also in time, the PLATO computer installation at Urbana would be replicated at various locations throughout the country.

PLATO provides an interaction, particularly with its responsive "drawing" of graphics and slide retrieval, which is intended for one user. Its effectiveness for a large audience, via a television camera focused on the PLATO plasma panel, may be another matter. Also, the bulk of each terminal, approximately a 30" cube, blocks off one user in the same room, from another. PLATO is intended for individual instruction, providing drill, simulation, and means to stimulate critical thinking and it does that very well.

Other than that, PLATO is also a one-to-one learning technology which stresses self-paced learning, utilizing computer assistance and graphic display; TICCIT and PLATO could hardly be more different.

PLATO needs 4000 terminals to begin to be cost effective, unlike the 128 terminal self-contained system of TICCIT. Also PLATO does not use a standard TV monitor. It has a plasma screen onto which still or moving pictures can be displayed (if previously loaded into the individual PLATO terminal). But PLATO does provide moving, line graphics on a time-share basis as a response to the user's own queries. Neither of these is possible with TICCIT. But neither are seven colors possible with the PLATO plasma panel, only the one color, orange.)

Concepts are stored in PLATO and actual computation produces response, not data retrieval. Stored data and pre-programmed questions are stored in TICCIT, supplemented by a limited computational capacity. (It is said storage of data is inexpensive and that a great enough experience will anticipate the bulk of the questions, so that for certain instructional use, TICCIT has adequate capacity. TICCIT is more rote-like than PLATO. Its content matter may be limited to certain areas - which? And will PLATO's sort of movement be missed?)

PLATO is easily "authored" by the user. This both puts the individual professor at the center of his own course development and frees him for work with individual students. He is not displaced by a "presenter", but may lack the imagination and media competence to make full use of his PLATO terminal. (The TICCIT counter to this is that 1 out of 4000 teachers is published and that this is a fair measure of those capable of developing their own courses.)

PLATO is especially surprising in the easy way it "accepts" unprogrammed questions, "deducing" their sense out of misspellings, etc. -- as long as the query is within the program's vocabulary; and it "judges" responses. Of course, it cannot answer every question of every policy-maker, but it can relate the total effect of changing one or a number of variables in a programmed "concept". Perhaps programming is so easily managed that PLATO can be quickly brought to handle such a high proportion of exploratory questions that a policy-maker would be attracted to use it and forgive the times it took a day or a week to respond.

The student-per-hour contact cost at the optimum loading of the 4000 terminals for PLATO IV is down to 35¢ per hour. This is said to be comparable to normal size elementary teacher-led classes. It includes the $2.5 million main frame cost of the computer, two million words of memory and other input-output equipment costing $2 million and another $1.5 million for software (amortized over 5 years), plus operating costs. (Some have questioned if PLATO can handle more than 2000 terminals.)

HumRRO - Color Halftone Area Graphic Environment (CHARGE)

Human Resources Research Organization
300 North Washington Street
Alexandria, Virginia 22314
(Dr. Ron Swallow - 703-549-3611)

HumRRO, established in 1951 by The George Washington University as a non-profit corporation, was under exclusive Army sponsorship for its first 16 years and developed technology of training and education as well as research into motivation, leadership and personnel management. In 1967 this competence was extended to other federal agencies and to state and local governments. The relationship with The George Washington University was terminated in 1969, providing it new flexibility as an independent non-profit corporation.

HumRRO, it can be seen, entered the technology of computer assisted instruction with a particularly rich concern for practical educational applications. The reason for including HumRRO here is that it has recognized the different virtues of TICCIT and PLATO and has developed a system that is said to combine those virtues and enhance them at competitive costs with greater flexibility in the number of users. The HumRRO system uses a standard color television monitor for display, as TICCIT does in a more limited way, and provides for animated response that TICCIT cannot do and exceeds PLATO in realism and speed. (The CHARGE terminal is now being built. The image generator, which requires about $200,000 support and one year's work, has not been funded.) The proposed design is economical for a system with as few as 100 terminals, yet modular for expansion beyond 1000 terminals without duplication of text storage-retrieval subsystems. All text material is centralized with little or no need for films, or visual materials at the terminals. The power of the system's special-purpose hardware produces cost-performance gains which are orders of magnitude above that possible through the use of a general-purpose computer and software.

Essential elements in achieving this advanced design are summarized in HumRRO working papers as follows: (1) terminal architectures which incorporate new solid state devices, i.e., CHARGE terminal (Color Halftone ARea Graphic Environment); (2) special-purpose hardware to take over well defined and stable software functions, i.e., image generator for graphics transformations from 3-D to 2-D; (3) eliminating I/O bottlenecks within the central computer system by using high-speed drum swap and building a few special interfaces where necessary; and (4) the latest computer CPU and RAM components for mini- and midi-computers where more production cost effectiveness can be realized.

Described by means of x-y-z coordinates, 3-D objects can be "stored" in the computer and converted to 2-D, colored shaded perspectives on the CHARGE terminal. The objects can then be either rotated and move themselves, or be moved through, by means of a keyboard or control stick. A new perspective can be calculated, transmitted and displayed in .15 seconds, changing perspectives, moving closer. "All of this can be done in vivid color, with curved surfaces appearing smooth".

The CHARGE terminal "resolution ... is of the order of 2000 horizontally and 1000 vertically, actual picture quality being limited by the color monitor" capable of altering the displayed picture at real time rates, rather than that presented through the monitor ... without flicker. (18 bits of color are used and the gray bits are of the order of a 7-bit log scale.) The terminal is estimated to cost $10,000 or $6,000 in quantity. Up to 200 terminals can be supported by an image generator costing around $100,000, permitting real time image generation in 5% of the terminals.

Evans and Sutherland, and GE have developed graphic terminals that generate color perspectives from 3-D objects. Both are limited to one terminal costing approximately $250,000. (See page 54)

The proposed CHARGE terminal system is said to be 15 times faster than that developed by Evans and Sutherland and its output is a model picture (coded as a set of edges) which reduces the requirement on a terminal refresh buffer from 42.10^6 bit to only $1/410^6$ bits. In contrasting the complexity of the images that can be generated, CHARGE handles 64,000 edges, Evans and Sutherland 3000, GE 250.

Characteristics of Significance - HumRRO's system combines realism in continuous tone color, perspective and movement with responsive capacities which exceed those of PLATO. A virtue for the regional policy-maker using CHARGE is to "experience" proposed developments before they are built and to share this experience with interested citizens. How will new projects appear? What would it be like to move through them? Concepts that require color, perspective and motion to help convey their meaning spatially and over time could be described and graphically displayed.

Further, by being compatible with standard television, one central storage system for a region could convey all manner of information to single terminals, for large screen projection in regional situation rooms or to a whole community television audience. With two-way cable capabilities, many terminals could view the same image at one time and participate in altering it. A further advantage of CHARGE, and TICCIT, is that the "refresh" hardware for a number of terminals can be shared, thereby reducing cost. This is not possible for PLATO in which the refresh unit has to be built into each terminal.

One system could encompass community graphic requirements from education to general community dialogue and policy-making. This one system could handle (1) real time image generation, (2) live broadcasts, and (3) videocassettes.

From HumRRO reports, the following comparison can be made of its CHARGE system to TICCIT and PLATO:

	CHARGE	PLATO	TICCIT
display	tv monitor	plasma panel	tv monitor
minimum number of terminals for complete system	100 - 1000	1000 - 4000	100 -
interconnection	tv cable	telephone cable	tv cable
motion	computer generated in full color, line and area	computer generated line, one color, one gray level	hand access videocassette, full color
3-D to 2-D hardware/software	yes	no	no
terminal image resolution capability (points discernible horiz. & vert.)	2000 h 1200 v	500 h 512 v	320 h 240 v
system core (16 bit words)	128,000	256,000	32,000
swap & job lesson size (16 bit words)	32,000	4,000	8,000
response time	.15 sec.	.15 sec.	1.0 sec.
transmission time	.004 sec.	0.1 - 2.0 sec.	0.1 sec.
executable instructions (user/second)	10,000	1,000	3,000
non-echo jobs (peak-seconds - 1 job/user)	1	2	20

THETA-COM - Subscriber Response System

Theta-Com of California
9320 Lincoln Boulevard
Los Angeles, California 90045
(Marshall Carpenter - 213-641-2100)

The SRS is a two-way interactive computer assisted television communication system, a commercial development financed by Hughes. Communication takes place between a computer complex and subscriber terminals. The subscriber terminal consists of a modem, with no operating controls, and a subscriber console, of which there are two models. One model has a 3-digit board, the other a 0 to 9 numeric keyboard and a paper strip printer. They can be installed on either single or two-way cable systems, and a microwave link has been tested. The SRS is designed to respond in a time-share mode to heavy service traffic (50,000 subscribers) within 2.4 seconds, including the recording of billing subscribers on magnetic tape.

This winter 1000 pre-production models of the SRS will be installed in El Segundo, California homes for a dynamic testing of:

New Services:	premium television, restricted channels, frame grabbing, channel polling, opinion polling, interactive education, and audience participation;
Existing services:	meter reading, shop-at-home, reservation services, emergency services (fire, burglar, police) and various derivations of these services.

The modem and console will be provided to subscribers and actual services delivered on a fee basis. The purpose is to test the reaction of subscribers to the services, determine their profitability and the need for change or modification.

The following is from a Theta-Com technical paper:

The CATV system being installed in El Segundo by Theta-Cable is a two-cable system, shown in Figure 7.[1] The system consists of an "A" cable system and a "B" cable system. The "A" cable trunk line is intended for downstream transmission only in the band from 54 to 300 MHz. The "B" trunk line operates bi-directionally: downstream from 174 to 300 MHz, and upstream from 5 to 108 MHz. The "A" cable distribution system, however, operates bi-directionally with the upstream bandwidth between 5 and 30 MHz. Upstream signals from "A" distribution line are routed through appropriate low pass filters to the "B" trunk where they travel upstream to the head end and are routed to the SRS Local Processing Center. The "A" trunk, and distribution system, is intended primarily for home subscribers while the "B" trunk is intended for municipal, business, and industrial usage, where the greater upstream bandwidth (5-108 MHz) can be utilized for additional data communication and/or upstream video channels.

The "A" cable system will contain 32 trunk amplifiers and 124 line extenders. The "B" cable system will contain 32 trunk amplifiers and only 4 line extenders, in view of the smaller distribution demand anticipated from the specialized users.

The longest cascade in the system consists of 7 trunk amplifiers and 2 line extenders. Total plant mileage is approximately 30 miles. For the trunk, 3/4" foam dielectric cable will be used and 1/2" cable for the distribution system.

For the El Segundo system, Jerrold SP-1/2/5-2W trunk amplifiers and SLE-300-2W line extenders are used.

Characteristics of Significance - Regional policy-makers on FCC standard two-way cable television systems (with TV cameras) could use the SRS for two-way video communication between specific subscribers, such as government agencies and private citizen organizations.

Two-way use of computer assisted instruction would be possible, but limited compared to PLATO. However, special regional videotape movie productions and stored images could be accessed, and for opinion polling each person polled could be identified. The response time of 2.4 seconds for 50,000 subscribers would probably be considered by policy-makers as "too long". It is said that up to .5 seconds it is difficult to distinguish differences in response time. Over .5 seconds, the difference in the lag time of response is noticeable. A dominant proprietary feature claimed of the SRS operation is how it has combined a mini-computer and software to cut response time for large numbers of subscribers.

Initial production terminals will cost $300-400 each and will be available starting in the last quarter of this year.

TOCOM - Total Communication

TOCOM INC.
P. O. Box 47066
Dallas, Texas 75247
(Charles Low, Vice President and Brian Belcher, Manager, Digital Systems)
214-253-3661

TOCOM has developed an interactive home terminal system for two-way cable television, primarily for commercial applications. It is to be first tested at Irving, Texas, near the home plant, with 1500 subscribers.

This is primarily a computer assisted system for communication, not for processing or instruction. (TOCOM's simple 4-key terminal is said to be mass producible at $125 per terminal versus $375 for Theta-Com or, as I have been told, $600 for the RCA terminal). The TOCOM system can respond to 60,000 remote units every 6 seconds. It can be used to provide burglar, fire and "emergency request" services as well as pay-TV, meter reading, information retrieval, medical monitoring, home shopping, keyboard inputs, etc.

TOCOM, designed for the cable system operator, consists of: (1) the home terminals, which operate through two-way coaxial cable with the standard home television monitor; (2) a computer controlled central data terminal.

The home terminal is a combination 26-channel TV converter and a digital transmitter-receiver with its own unique identification at the central computer (for polling, marketing, billing purposes). The initial subscriber capacity of a system is 2000, expandable to 60,000.

<u>Characteristics of Significance</u> - TOCOM claims to be the simplest (4-key), least expensive two-way electronic system at the prototype development stage, ready for manufacture. FCC rulings of March 31, 1972 require all new cable TV systems in the top 100 markets to be built with two-way capabilities; all the existing systems in the 100 top markets are to be converted within 5 years.

Besides the projected home-to-television studio response, systems like TOCOM and Theta-Com could be adapted for use either <u>within</u> regional situation rooms or interconnected situation rooms. Polling of these, limited audiences would seem the most significant first use for regional policy-making purposes.

SCRIPTOGRAPHICS - Data Tablet

Scriptographics Corporation
390 King's Highway
Fairfield, Connecticut 06430
(203-384-1344)

Scriptographics is similar to but different from its competitor the RAND tablet. The Scriptographics tablet senses the position of either a stylus, which can be an inked pen, or a sensor. This signal is conveyed in a stream or point mode, or by remote signals, with a resolution of 100 lines per inch. Data can be presented in Binary or BCD and developed on a cathode ray tube or television monitor. The tablet is 3/8" thick and comes in sizes for 11" x 11" to 36" x 40". Other sizes are available. It can be transparent or not. Prices are in the $2,000 range, plus or minus, depending on quantity, size and options selected.

Characteristics of Significance - This is a simple way to add motion and emphasis to graphic presentations to or by policy-makers. No skill is required. It could become a particularly effective means to communicate via either TV monitor or large-scale projected TV images. On an aerial photograph or chart, for instance, white lines can be "drawn" over the image to add information. With tablets in front of each participant in a regional situation room, any one of them could add to each other's drawing, equation, etc.

EIDOPHOR - Large Screen Television Projector

Datex Division
Conrac Corporation
1600 S. Montana Avenue
Duarte, California 91010
(Kenneth R. Eppele, Marketing Manager - 213-359-5381)

The Eidophor 5070 accepts either the standard red, green or blue television signals or digitized information. Images can be projected up to 30 x 40 feet by this high intensity light system, manufactured by Gretag in Switzerland.

From the Eidophor brochure: "In an EIDOPHOR Large Screen Projector the incoming television signal is modulating an electron beam which in turn deforms a thin oil layer on a concave mirror. This oil layer is the actual picture carrier of the EIDOPHOR system. The brighter a certain spot of the picture should be, the more electrical charges are deposited by the electron beam on that spot and the more the oil layer gets deformed.

The television signal of a complete picture therefore engraves a raster scan relief image into the oil film similar to the flat picture on the screen of a home television receiver. The oil relief on the mirror now deflects the light of a 2.5 kW Xenon lamp and using a dark field optics arrangement containing a mirror bar system it projects the television picture through a lens system onto the large size screen. The outstanding advantage of the EIDOPHOR system is in the use of a separate light source which makes it possible to project large pictures of excellent brightness and resolution. The white Xenon light of the simultaneous unit is split into its red, blue and green components for each EIDOPHOR subunit so that each subunit projects its own color content onto the large screen. Superposition of the three color images produces the true color picture. Automatic electronic registration guarantees a sharp picture at all times. Automatic color correction circuitry further assists in projecting a perfect picture with regard to brightness, color tone and color saturation. The EIDOPHOR simultaneous color system which is compatible with public color television, achieves excellent and flickerfree reproduction of large color areas as well as small details."

The simultaneous color unit, model 5070, has a light output of 3600 lumens, a minimum of 800 lines at the center. (It is said to have a design capability of 2000 lines.) The system consists of two units: the projector which is 41" wide by 44" deep by 74" high and weighs 1060 lbs. and the electronics cabinet 22" wide by 35" deep by 53" high, weighing 465 lbs. The current catalog price is $185,000; Model ED-8, for black and white television, is $65,000.

Characteristics of Significance - NASA uses this system to provide operational information to its large roomful of technicians at the Manned Spacecraft Center at Houston working in normal lighting conditions. It is finding increasing use in large sports arenas for scoreboards and for special effects in television studios. In a regional situation room for television broadcasting, Eidophors can provide a large, brilliant, clear image. It offers the potential for superimposing digital graphics and other data on television images as called for by the viewer. A television camera in a normally lighted room could pick up a large Eidophor image and all the people meeting in that room.

GE LIGHT VALVE - Large Screen Television Projector

Video Display Equipment Operations
Building 6, Room 206
General Electric Company
Electronics Park
Syracuse, New York 13201
(Gerrold P. Gunderson, Sales Manager - 315-456-2562)

GE's large screen television video projector system starts from the same light valve principle as Eidophor, but GE uses a single electron gun for all three colors versus one for each primary color used by Eidophor. Its purpose is the same, but it does not claim the brightness of Eidophor. GE suggests projection of images from 2' to 20', the latter in a darkened room. GE's light source is 750 lumens. The horizontal resolution capability is 600 TV lines minimum. Dimensions for the color model PJ500 are 24" wide by 23" deep by 60" high; weight is 460 lbs. The price is $41,500 and the monochromatic model PJ700 is $30,000.

Characteristics of Significance - GE delivers one quarter the brightness, significantly less color tone quality, and much loss in potential resolution compared with Eidophor - but it does so at one quarter the cost, and requires less complex maintenance. If a large image in a normally lighted room is needed, especially if that image needs to be picked up by a television camera, the GE light valve is inadequate. (NASA's use of the GE Light Valve at the Goddard Space Flight Center is in a large, dimly lit room.)

At a GE demonstration in New York in a darkened room, it was surprising how satisfactory a 10' projected image was when fed from a SUNY videocassette - said to have a resolution equivalent to about 200 lines.

Comparing the GE Light Valve black and white image with a color image at NASA's Goddard Space Flight Center (each 9' in width), it was easy to see how much sharper black and white images are. The major reason is the number of lines of resolution (272) for color versus (800) black and white television images. But there is also bound to be a significant difference in the tonal quality of color images between GE and the Eidophor systems - one electron gun versus three, 700 lumens versus 3200. The two need comparison in the circumstance of a regional situation room to resolve the large price differential. Television coverage, lighting of the room, and detail to be read from the screen are important criteria for such a test.

ARPA NETWORK

ARPA Network Information Center
Stanford Research Institute
333 Ravenswood Avenue
Menlo Park, California 94025
(415-327-0940)

The ARPANET is a government sponsored (Defense Advanced Research Projects Agency) communication system, interconnecting a set of computers across the nation, that provides very fast responses. Its interactive message switching has fostered the development of techniques for computer-to-computer communication.

A goal of the ARPA Network is to provide persons and programs at one location on the Network access to, and interactive use of programs that exist and can be run on other computers wherever they may be in the Network. Over the past three years it has grown to serve over thirty sites, mostly colleges, with over forty independent computer systems connected.

The ARPA Network interconnects by means of wide-band leased lines. Small identical processors, Interface Message Processors or IMPs are placed at each node of the Network connecting each computer center, or Host, to a system of leased 50-kiloband common-carrier circuits. Each IMP can support up to 4 Hosts. A terminal IMP is designated as a TIP. Each TIP can support up to 3 Hosts and 64 computers.

Much time to date has been taken to develop the various protocols to allow communication between the many types of computers, terminals and data formats on the Network.

For government sponsored users, initial costs are $78,000 for a TIP and $54,000 for a 316 IMP. Maintenance costs are $7,000 per year for TIP and $5,000 per year for IMP. Operating costs are $16,500 per year plus 30¢ per kilopacket in excess of 4500 kp in a single month. (A kilopacket is 1000 bits.)

The Network Control Center at Bolt Betanek and Newman, Inc., Cambridge, Massachusetts has overall responsibility for the operation of the communications aspects of the ARPA Network. The Network Center at UCLA regularly monitors traffic and experiments to determine performance characteristics. At Stanford Research Institute, the ARPA Network Information Center helps ARPANET users find resources for their information-handling needs and to help geographically distributed groups collaborate with each other.

Characteristics of Significance - No one region need attempt to amass all the computers, programs and expertise it could use for regional policy-making if it is connected to the ARPA Network. (At present PLATO is not directly connected, but it is technically feasible to do so.) Any user incorporated into the ARPA Network can call on any other user of the Network for a special competence or workload capacity it does not itself possess. Using the variety of programs and computer systems the ARPA Network makes available calls for expertise.

V THREE APPROACHES TO ANIMATION

UCSD - Animated Color Movies ~~Derived from~~ Computer Graphics

Chemistry Department
University of California at San Diego
3262 Urey Hall (Revel Campus)
San Diego, California
(Kent Wilson - 714-453-0200 x1473)

The interest of Kent Wilson and his students has led to a number of short (10 minute) NSF funded color movies to better communicate scientific information. Single frame camera photographs are made, in black and white, of the computer generated line graphics on the face of a cathode ray tube. One by one, the images are photographed and, when run at normal motion picture speed, provide motion. Color is added to the resulting black and white film by means of filters through "aerial photography". Three separate films have been funded by NSF for about $20,000 each and each has taken an elapsed time of 6 months. It is difficult to pre-estimate costs for such work (the subject requirements can vary so), but for students with this knowledge behind them, an elapsed time of 3 to 4 months would be adequate to produce similar productions now.

Of the three films produced so far, the one that shows air pollution over Los Angeles most resembles what a policy-maker and his constituents would use. That film, by means of contour lines that rise and fall over a period of time related to a simplified perspective map of the Los Angeles Basin, shows the status of various air pollutants at specific locations. The effect of the sun daily is clearly shown, as well as where in the Basin the pollution is worse. The movie is derived from data collected at fixed monitoring points in the LA Basin. A program is written which connects by contour lines these points and displays fluctuation of readings in a three-dimensional perspective.

<u>Characteristics of Significance</u> - These films are an excellent demonstration of how understanding can be quickly achieved. Collected data is interpreted into information graphically. A long lecture or a thick report with voluminous tables, with an expert there to read them, is not the way to convey meaning of this sort to a policy-maker--just a short film.

Color is a significant addition to the black and white photography and was added to the film cited above for less than $500. The policy use of this technique would be for specially prepared analytic movie productions in which the factor of time is dynamically shown with motion in relation to a specific geographic area.

UTAH - Halftone Computer Animated Graphics

Division of Computer Science
(Merrill Engineering Building)
University of Utah
Salt Lake City, Utah 84112
(Ivan Sutherland - 801-581-8224)

Graduate students are generating continuous tone animation cells for motion pictures at the rate of 100-200 per hour and believe they could, with other equipment, go to 700 per hour. They use a Digital Equipment Company PDP-10 computer, a 5" cathode ray tube with 1,000 lines and a 35 mm animation camera; a "Watkins-Box" solves the hidden surface problem.

The object being animated is first completely described by polygons located on its surface. (The "points" of the polygon can be numerically defined in relation to each other.) With the assistance of special algorithms (programs or instructions), the area between the points of each polygon is first shaded and then the edge of the plane of abutting polygons is smoothed. An airplane, a face, a building, all sorts of objects have been so described and then caused to "move", one "cell" at a time, by computer instructions which display each cell on the surface of the cathode ray tube which is photographed, one cell at a time, by the animation camera. Additional instructions can work filters into this process which then produces color animation.

This work has developed to this point under the auspices of the Department of Defense through ARPA. Funding has been ended now since DOD claims it has what it needs (simulated aircraft carrier landings from the viewpoint of the practicing pilot in which the ship moves and movement of the plane is responsive to actions taken by the "pilot").

There are 70 students at the Division of Computer Science, about 5 of whom are in halftone computer animated graphics. There is said to be more work to be done to develop more algorithms and especially to devise three-dimensional data measuring and input techniques for this purpose. Continuing funding is now being sought from NSF.

Reports describing this procedure are:

COMPUTER GENERATED ANIMATION OF FACES - Frederic Ira Parke
Computer Science, University of Utah, June 1972, UTEC-CSc-72-120

COMPUTER DISPLAY OF CURVED SURFACES - Henri Gouraud
Computer Science, University of Utah, June 1971, UTEC-CSc-71-113

Characteristics of Significance - This computer generated animation drastically reduces the cost of animation previously done by hand. Both real time and prepared animation can be generated for individual regional study and communication purposes that would have taken too long and been too expensive to even consider before such animation. Aesthetic and purely cognitive concepts too abstract or complex to be visualized or conveyed before will be possible.

HUMRRO - Color Halftone Television Animation

(see description of system in Section IV of PART II)

Characteristics of Significance - At UCSD the result is limited to line, color animation as generated by computer to a standard cathode ray tube (CRT) where it is displayed and photographed one cell at a time. This photography is then run in normal motion picture sequence to simulate action. The UCSD approach is excellent and inexpensive for reproducing line graphics such as diagrams, charts and simplified line drawings on film.

At the University of Utah, Evans and Sutherland, with their $250,000 Watkins Box, etc. can generate, for filming shaded, continuous tone, color animation of irregular objects at costs that are a fraction of hand animation.

HumRRO's system, it is said, will produce color halftone "animation" from computer stored data directly to television monitor display. And it is done with relatively inexpensive terminal and image generator equipment.

The extra capability of the HumRRO system is the "real time" response it can provide once basic images or "sub worlds" are described and stored in the computer system of which it is a part. This permits the viewer to generate graphic response and movement at his command and it too can be used to generate graphic displays of policy analysis. Its natural mode for recording would be on videotape, not film, eliminating processing, delay, expense. It cannot be claimed, however, that television equipment to date can produce the same sharp resolution as film, and in certain cases the preparation of animation ahead of actual use will be desirable or perhaps necessary, making filmed cathode ray tube (CRT) production competitive with 'real time" TV display.

VI TWO RELEVANT CONFERENCES

FRONTIERS IN EDUCATION - Purdue University, April 9-11, 1973

This was the third annual Frontiers in Education conference. But it was the first conference sponsored by both the Educational Research and Methods Division of the American Society of Electrical Engineers and the Education Group of the Institute of Electrical and Electronic Engineers. While there were some industry and consultant representatives present, the 500 attendees were mostly engineering professors. The emphasis of the conference plenary and workshop programs was to identify and discuss innovations of "trail blazer educators".

Stress was placed on breaking down the barriers against assimilating new technology into the classroom. The reluctance, even fear, of professors to change their approach was scored - and the implications of electronic means to facilitate education outside the classroom, as well as in it, were repeatedly brought forth.

The conference focused more effectively on self-paced learning than classroom teaching. Innovative findings and experience in educational technology, sociology, educational psychology, and other allied fields were discussed as they applied to engineering education. But the implications of the new educational technology, by the nature of the interactive technology so often at the center of attention, was much broader than engineering education. It seemed to me that education in general, including education of the regional policy-maker of concern in this investigation, could have been included just as well.

Both the content of engineering courses and the easy familiarity engineers have with technology make it natural for education and technology, and practical communication applications of both to society, to evolve in engineering schools. As Ralph Siu said at the 1967 American Institute of Planners conference, perhaps it is the technologists (the engineers) that will lead the way to a new order of learning, to development of a holistic humanism that integrates art, spirit, science and technology. There are many more such technologists than scientists, artists or spiritual leaders and their ego is perhaps less. If the engineers at this Purdue conference weren't openly enthusiastic about the use of new educational technology, they were at least curious, and at worst, fatalistic about the inevitability of its application - once the cost/benefit ratios were worked out.

The state-of-the-art of educational technology seemed to include the use of slide projectors for some! But use of videocassettes, and the latest in self-paced computer assisted instruction was demonstrated and discussed also. Statements about dull professors being replaced by dynamic "presenters" on videocassettes, produced with the aid of educational psychologists, and media specialists, with the professors serving as course compilers and consultants were flatly stated and stolidly received. It appears the engineers are on the brink of really doing "something about it" (education).

Published: Proceedings of the Third Annual Frontiers in Education Conference, Institute of Electrical and Electronic Engineers, 345 East 47th Street, New York, N. Y. 10017. IEEE Catalog Number 73-CHO 720-3E (350 pp. - $10)

KINOSTATISTICS - Washington, D.C., July 27-29, 1973

Bureau of Social Science Research, Inc.
1990 M Street, N. W.
Washington, D. C. 20036
(Albert D. Biderman - Barry M. Feinberg - 202-223-4300)

BSSR in December 1972 published a 50-page booklet KINOSTATICS - Communicating a Social Report to the Nation by research assistant Barry M. Feinberg. Following up that report, the Symposium-Workshop, sponsored by BSSR July 27-29, 1973 in Washington, brought together a remarkable array of skill in graphic communications to discuss how computers, film or television might be used for quantitative communication.

BSSR's particular concern was focused on communicating "social indicators" - statistical measures that reflect crucial states and trends of the country. The twenty-five participants included two New York film producers and Canadian Broadcasting Company TV producers, interaction computer graphic experts, teachers, a Vice President of Computer Image Corporation whose animation techniques are used in Sesame Street, social scientists, symbol designers, a map information specialist, and consultants. It quickly became clear that the graphics being discussed had more to convey than the statistics of social indicators. Kinographics was suggested as an alternate title for this new field of communication.

Films, computer terminals, maps, slides and videocassettes demonstrated the potential capabilities of "graphic language" for conveying information, especially in an interrelated "systems" sense. A number of basic questions about the use of kinographics were raised. What audience? Researchers or the general public? Would it be used to teach people, or communicate to them? How much simplification before distortion?

Some interactive computer graphics might be just for researchers, but kinographics offer the potential of using symbolic language that is intelligible at a number of intellectual levels to totally different groups. The distinction between teaching people and communicating with them was challenged. It was agreed that simplification is necessary, but must be responsibly done, and hopefully, with interactive computer graphics, at least, subject to testing at various levels of generalization by going to underlying data. The distinction was made between the power to communicate and the power to motivate action.

It was agreed that information could have two distinctly different bases of reference: (1) geography or place, and (2) function, as with generalized national information.

There seemed to be total acceptance of the necessity to develop graphic symbols for social indicators and to test those in State of the Nation-type federal reports. But the need for establishing graphic conventions and facilitating their application seemed to call for consideration of graphic symbols in their total context, which includes environmental and cultural as well as economic aspects. Establishing a clearing house and archive, affiliated with government but organized as an independent non-profit entity and not subject to direct governmental control, was proposed.

Questions about the accuracy of visual communications and their potential impact seemed to call for organizing in a way that precludes the possibility of charges of propaganda.

A demonstration film and a report will be forthcoming from this workshop.

VII VARIOUS OTHER NOTES

(The disclaimer which has already been made about the comprehensivity of this report needs to be "refreshed" for this section especially. Under the man/hour limitations of this contract, it is not possible to report on all the resources for regional graphics. This section briefly comments even more briefly and provides contacts for several more.)

UCD - Computer Driven Regional Model

Department of Zoology
Storer Hall
University of California at Davis
Davis, California 95615
(Kenneth E. F. Watt - 916-752-1558)

Ken Watt has been funded by NSF and the Ford Foundation in his efforts since 1970 to model "society." His Land Use and Energy Components Study to model the State of California has evolved into a series of interacting models that represent global inputs such as weather, resource demands, pollution, population, crop failure, etc. The significance of this work is a demonstrated capacity for simulating the effects of changes in variables upon a modeled system--such as a region.

Queries typed into his San Diego regional model, through a Hazeltine terminal via ordinary telephone lines to a GE time-shared computer (CDC 6600 plus a 7700 and a STAR UNIT located in Indiana), were quickly responded to on the standard green cathode ray tube in capitalized small type face.

The model I saw demonstrated gave a great load of detail in response, and seemed limited in the sort of query it could accept. It seemed programmed to multiple-choice questions, which, step by step, narrowed down to an answer.

This demonstrated ability to model a region, seems to need further development for free inquiry and more dynamic graphic spatial display. The communication link from the model to the policy-maker is quite different than from the model to the researcher or programmer.

The significance of the Davis work seems to be in the extent to which a region can be modeled. Undoubtedly, it provides effective data access to middle management now, even in its type output. How might it be queried more effectively by policy-makers? Is graphic display of its output possible, especially on a spatial basis? Or is this modeling primarily for technical users who need, then, to "repackage" it for policy-makers?

EPA - The Environmental Protection Agency

Office of Research and Monitoring
Environmental Studies Division
401 M Street, S.W.
Washington, D.C. 20460
(Dr. Peter House, Director - 202-557-7479)

The National Environmental Policy Act of 1969 led, in 1970, to the organization of EPA, the Clean Air Act, the Water Quality Improvement Act, and the Resource Recovery Act. EPA's concept of long-range comprehensive environmental planning in such resource terms has since been expanded to include the social, economic and physical factors involved in the term "Quality of Life." The President's Council on Environmental Quality provided this mandate when it said in 1970 "Effective strategy for national environmental quality requires a foundation of information on the current status of the environment, on changes and trends in its condition, on or what these changes mean ..."

Subsequently, EPA organized its complex computer based Strategic Environmental Assessment System (SEAS) "for the assessment of alternative procedures in terms of their long-range impact on the environment" with 10-20 year time horizons.

EPA has sponsored national and regional environmental research conferences, symposiums, and projects. Its ground-breaking importance as the lead federal agency in environmental resources management ties it closely to regional policy-making. Among its publications, see:

> STRATEGIC ENVIRONMENTAL ASSESSMENT SYSTEM (SEAS), May 1973
> Office of Research & Monitoring
> Environmental Studies Division
>
> ENVIRONMENT
> National Conference on Managing the Environment
> May 14-15, 1973 - Ramada Inn
> Washington, D.C.
>
> ANTHOLOGY OF SELECTED READINGS FOR THE SYMPOSIUM ON THE QUALITY OF LIFE CONCEPT - A Potential New Tool for Decision-Makers
> Airlie House - August 1972
>
> QUALITY OF LIFE CONCEPT - A Potential New Tool for Decision-Makers
> c. June 1973

USGS/HUD - San Francisco Bay Region Environment & Resources Planning Study

U.S. Department of the Interior
U.S. Geological Survey
345 Middlefield Road
Menlo Park, California 94025
(Robert Brown - 415-323-2145)

This study, jointly funded by HUD's Office for Research and Technology and the U.S. Geological Survey, began in January, 1970. It is conceived as a broadly based approach to relating physical environment factors, particularly geologic hazards, to regional and urban planning and development.

Original funding was based on estimates of a minimum effort in each of 30 earth-science program elements. Later, other elements of urban land use planning were added. The diversity of environmental considerations treated in the study, it was believed, should develop many principles applicable to other urban regions. It is considered highly experimental concerning the type of physical data collected, the way the data is synthesized, the formation in which the data is displayed, and the lines of communication that will help society utilize it. While there is a national interest here, it deals with the specifics of one region.

See publication: PROGRAM DESIGN, 1971 - San Francisco Bay Region Environment and Resources Planning Study, U.S. Geological Survey, Menlo Park, California 94025.

MRC-TV - Metropolitan Regional Council
Suite 2437
One World Trade Center
New York, N.Y. 10048
(Rodman Davis - 212-466-3850)

MRC-TV, inaugurated in June 1973, is an innovative telecommunications system serving local governments in the New York City metropolitan region. It employs a two-wave microwave television system:

1) from atop the 110-storey World Trade Center building in which MRC is headquartered to studios in each of the nine member communities,

2) between any two community studios, and

3) one-way video from each community studio to its local area, to which an audio return capability can be added.

The system is established by MRC and the nine member communities. NSF is financing a three-phase study of this system to determine its impact. The object is to bring public officials and other groups together via tele-conferencing. One black and white TV camera is installed at each studio location.

See publications: Intergovernmental Communities in the New York-New Jersey-Connecticut Metropolitan Region, 1972, and Communications and Government: A Regional Report, 1972, available from the RAND Corporation, Santa Monica, California 90406. See also MRC's bibliography on Communications, Telecommunications and Innovations.

HARVARD - Computer Graphics Technology

Program on Information Technologies and Public Policy
200 Aiken Computation Laboratory
Cambridge, Massachusetts 02138
(Anthony G. Oettinger, Director - 617-495-3946)

Laboratory for Computer Graphics and Spatial Analysis
Graduate School of Design
520 Gund Hall
48 Quincy Street
Cambridge, Massachusetts 02138
(Eric Teicholz, Associate Director - 617-495-2526)

Information Technology is an interdisciplinary program supported by the John & Mary R. Markle Foundation. It was launched in 1972 with two goals:

1) the development of a coherent understanding of information technologies and their policy implications, and

2) the illumination of public policy alternatives through this understanding.

The program criticizes prevailing views of information functions as fragmented by technology, institutional functions, application, policy issues and/or academic disciplines. It seeks an integrated view of information technology and makes the statement that computer and telecommunication technologies are now distinguishable only by their distinct constituent embodiments. A basic premise is that this inherent logic of information is increasingly unitary. (See its publication on STATUS AND PLANS, February 1971).

The Laboratory of Computer Graphics was established in the Spring of 1965 with a Ford Foundation grant. Under the direction of Howard T. Fisher, the Laboratory developed programs for high-speed, electronic digital computer mapping and new techniques for graphic display. It continues investigation into the uses of graphic analysis, computer graphics in particular. (The Laboratory is best known for its Synagraphic Mapping System, or SYMAP.) The Laboratory is also a service organization to the Harvard Graduate School of Design.

There are numerous laboratory publications available and it reports monthly through its newsletter CONTEXT.

HUD - Policy Planning and Research

Office of Policy Planning
Department of Housing and Urban Development
Washington, D.C. 20410
(Frederick A. McLaughlin, Jr., Director - 202-755-5965)

Community Environment and Utilities Technology Division
(Allan R. Siegel, Director - 202-755-5360)

HUD's role in housing, urban development, and local planning is well known and traces its roots back for decades, although the Department itself was authorized only in 1965. Its role in research and policy planning is not so well known.

HUD has led the ten agencies involved in its Urban Information Systems Inter-Agency Committee USAC since 1968. A major series of experiments in municipal information systems has been instituted through USAC. It has financed research application efforts such as Operation Breakthrough, Des Moine's Geo-Coding project, and the Regional Plan Association's Town Meeting TV series (see Section III), and the Metropolitan Regional Council-TV (this Section) are just three of the research efforts of significance to this regional graphics survey that are HUD-financed.

HUD is sponsoring a demonstration in rural Connecticut led by Dr. Peter Goldmark to show how telecommunications can provide services and amenities to make such areas competitive with cities. It is now also developing a community telecommunications demonstration concept to explore the potential impact of cable television or broadband communications on the urban environment. This concept is especially concerned with testing the delivery of social services via two-way communication and experimental demonstrations.

There is an interest here that has come to see in two-way television communication, means:

1) to relieve social and political alienation of individuals from government and community;

2) to provide immediate response to individual and neighborhood needs; and

3) to reduce costs for quality services to the individual consumer.

A particular concept at work here is a Community Information and Service Center which would provide a central point to which citizens could bring questions and problems relating to a wide variety of municipal services.

NATIONAL ACADEMY OF SCIENCES - Remote Sensing for Policy-Making

Committee on Remote Sensing Program
for Earth Resource Survey (CORSPERS)
National Academy of Sciences
2101 Constitution Avenue, N.W.
Washington, D.C. 20418
(Capt. Winfred Berg, Executive Secretary, CORSPERS - 202-961-1431)

Captain Berg is staffing the NAS Study on Remote Sensing for policy-making for which NASA and ERTA are providing demonstration data for special points in selected regions. Among the issues isolated is what digested data (information) the policy-maker should get in relation to familiar bases. What information is to be supplied? How is the base to be described?

An NAS panel, representative of various applications of remote-sensed data is seeking to determine what information is needed. After that, how to sense from satellites can be resolved. At present ERTS' Satellite Sensor resolution is adequate for regional purposes such as crop control, drainage, blight, etc. An object 200' on a side can be sensed from a satellite 40 miles up. That information is digitized and radioed to earth stations where it is converted into a graphic, continuous tone ("photographic-like" image). DOD resolution photography from satellites, in which film is dropped by parachute, is by one report said to identify objects 50' on a side. Ultimate results from satellites down to one foot will be possible. Present 200' sensed from ERTS is inadequate for most urban planning purposes.

This NAS Study divides world information by remote sensor into three time categories on the basis of the frequency it is collected for use:

FREQUENCY	EXAMPLE OF USE
minutes - hours	hurricanes, weather, pollution, peak-hour traffic
days - months	agricultural crops, commuting traffic
year - years	basic earth science maps, urban development.

The NAS Study will be published by Spring 1974.

CONSENSOR - A Consensus Recording Device

Applied Futures, Inc.
22 Greenwich Place
Greenwich, Connecticut 06830
(W.W. Simmons - 202-661-9710)

The CONSENSOR is a device that assists in the decision-making process by facilitating the determination of consensus. It is designed for use by task groups and committees that meet in conference rooms to make decisions. The device consists of a terminal for each of the participants in the meeting. The terminals are connected by cables to the Master Control unit and Display Panel. Each terminal includes two switches, one to register response, one to register how strongly that response is felt. The Bar Graph Display shows the collective results.

The CONSENSOR is designed to clearly and quickly indicate the overall consensus of the group. It assures the participation of all the members of the group, while allowing the influence of each member's judgment on the overall consensus to have its effect. It also enables the individual participants to express their judgments anonymously, in that the Bar Graph Display only presents the consensus of the group.

A FINAL WORD

This preliminary study cannot be ended without one final apology to those doing work vital to the subject of graphic communications for regional policy-making, for what has been either inadequately described here or omitted altogether. Final

The use of data plotters in transportation planning

As transportation planning and land use studies grew in importance, the amount of data generated and the difficulty of presenting it in usable form increased tremendously

The Federal-Aid Highway Act of 1962 alone, has produced 233 studies to date They range in size of population and area covered from 50,000 people in a few square miles to the New York City metropolitan area of 18 million people and about 8,000 square miles

When electronic data processing techniques first were applied to these studies by computers in the 1950's, the amount of detailed information increased greatly. Enormous amounts of numerical data poured from the machines and a great deal of time was spent preparing, by hand, graphic tabulations and displays required by analysts and clients

Computers soon were put to work to produce graphic data displays, or plots, in less time and with greater accuracy than by hand This report explores the state of the art of graphic data plotting by computer in transportation planning

It describes the computer mapping process, types of hardware, data processing techniques and the development of a geographic base file It also presents the results of an inventory taken to determine the use of data plotters in transportation planning and some illustrations of data plots Finally, the report touches on some new applications that have potential for use in transportation planning.

Computer mapping process

The process of computer mapping is rather straightforward Figure 1 shows the steps leading to an automatically plotted display Hardware used in automatic plotting are the digitizer and the plotting machine. The computer is needed to manipulate data, perform calculations and to develop new data The computer often is used on-line to command the plotter

The digitizer is an electronic drafting machine with a cross-hair sensor connected to an arm for calculating the X and Y coordinates of a point It can be connected to a key punch, tape punch or magnetic tape unit to record the data Digitizers cost from $6,000 to more than $100,000

Digitizing is the assigning of X and Y coordinates to a point These data are required to direct the plotter in locating that point.

Digitizing can be a simple process if used to draw a highway system for a small area. In such cases, the nodes of the links often are digitized by hand and a geographic base file developed that can be applied to the highway network

On the other hand, if a study is sophisticated in its analysis of land use data and the intent is to quantify land use parcels, substantial programming is required to manipulate the data The digitizing process also becomes much more time-consuming, generally requiring the digitizing of points around the perimeter of each parcel Such a task cannot be performed by hand for a large area.

Digitizing costs vary with the method used Digitizing a point by hand and coding an identification for that point costs 20 to 40 cents per point Encoding X and Y coordinates on a digitizer costs one to five cents per point if performed in-house, not including overhead costs

Plotting is the drawing of a map or chart and the display of statistics There are two types, character plotting and line plotting.

Character plotting uses the computer printer or other plotting device to make maps using various characters. The frequency of recurrence of a given phenomenon, say number of trip ends or number of cars parked, is indicated by the intensities of black print.

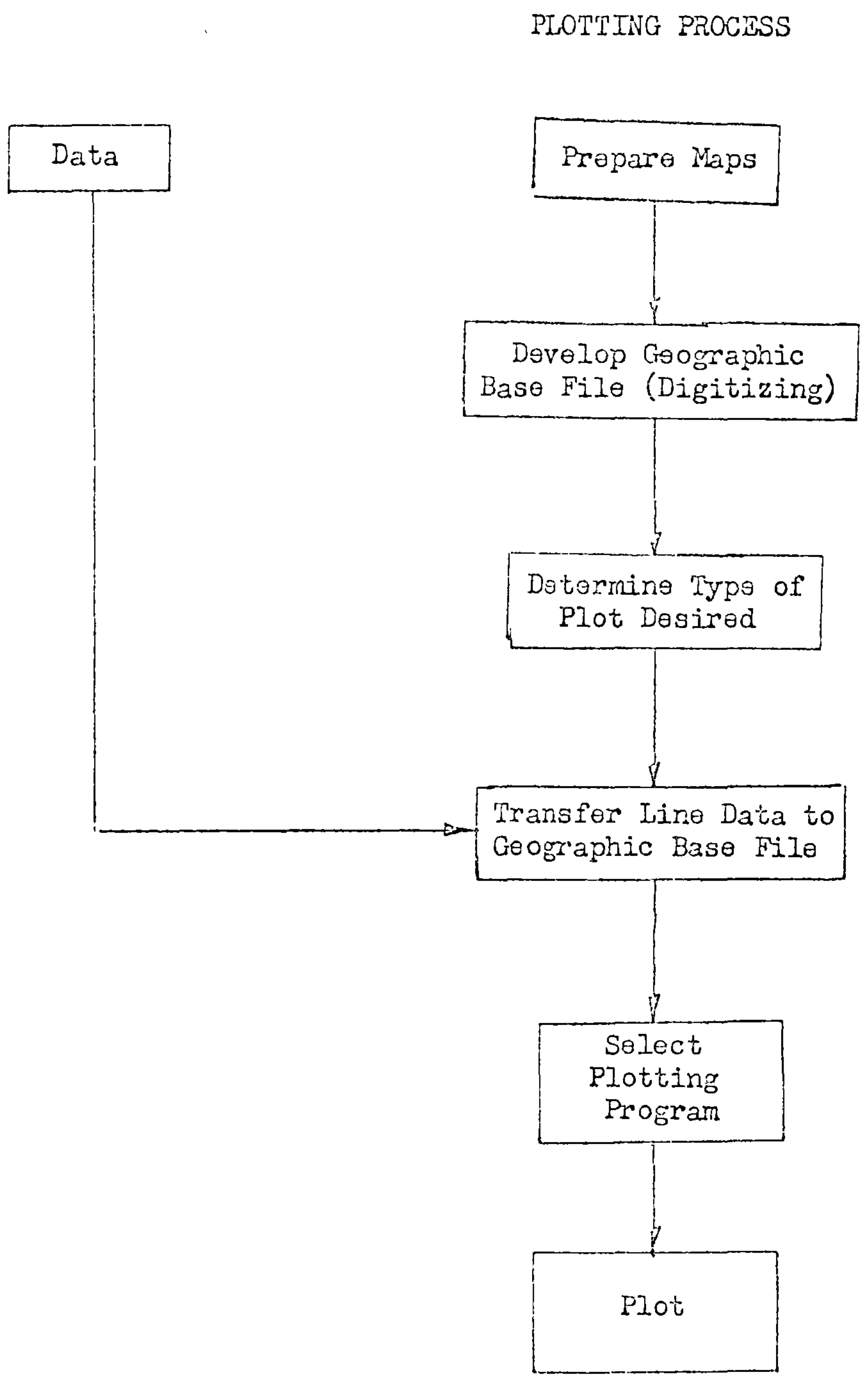


Figure 1

Several programs have been developed to do such plotting. Character or printer plotting is inexpensive and fast. It works well for plotting land use or area statistics and graphs. Figure 6 is an example of printer plotting.

Line plotting is an ink drawing on paper, the pen commanded by the plotter. Some line plotters are very sophisticated and have as many as eight pens so that eight different colors can be used with various widths of lines. This type of plotter comes in two forms, a flat bed that looks like a drafting table and a drum plotter.

Planning data

Data used in transportation planning falls into two general categories: area data such as population of a traffic zone or grid, and linear data such as average daily traffic on a highway system or ridership on a bus system.

Many data sources are used in planning. In the past few years, many sources have become computerized and can be plotted. Caution must be used in the plotting process, however, to select the appropriate data for plotting and to specify the plotting method. In the graphic plotting of Average Daily Traffic (ADT), for example, an appropriate scale must be selected.

Various maps are used in a planning study. Some serve well as base maps for the digitizing process, but others are not accurate enough or are too detailed. Two types are related directly to the two types of data available. They are area maps delineating such factual information as traffic zone boundaries, blocks, tracts and land use parcels and linear maps showing computerized networks, desire lines, spider web networks and other analytic material. In choosing maps to be used for developing a digitizing and plotting base, it is important to digitize all data with the same degree of accuracy.

Planning efforts concerned only with the plotting of highway systems found that the maps were not as accurate as the area data to be digitized.

Geographic base files

Geographic base files are maps in computer form. They consist of records containing location information on all segments within a metropolitan area. A segment consists of a portion of a street between two intersecting streets. Location information contained in segment records for an urban region consists of street name, range of addresses, node numbers for the intersections at each end of the segment, coordinates for the intersections and codes for the areas on either sides of the segment: blocks, census tracts, police districts and planning analysis areas.

Additional codes can be added to segment records for points, the intersections or nodes, lines, the segment or areas lying on either side of the segment.

Since the information on points, lines and areas is incorporated in a geographic base file so that it is interrelated, a geographic base file permits geocoding of point, line and site or area data on a common integrated and comprehensive basis. When a geographic base file is stored in a computer, it can be programmed to use the file in the same way as a map, but at much faster speeds. Input files for automatic plotting can be prepared easily from geographic base files when a plot at the street segment level—a block outline map—is needed because geographic base file records already are at this level.

Preparation of plotter input files at the arterial street network level requires additional work. This consists of identifying the node numbers in the geographic base file corresponding to nodes in the geographic base file. This information can be used to select coordinate values for nodes in the arterial street network. When the arterial street network contains a relatively small number of links, it may be easier to obtain coordinates for nodes manually.

Data to be plotted must be understandable to the plotter or computer printer and the X and Y coordinates must be developed. Various coordinate systems are applicable. Some studies developed a coordinate system that fits available base maps. Others used state plane and other Universal Transverse Mercador (UMT) coordinate systems. One study that used the process to plot airline origin and destination data for the nation has employed a scale of inches to digitate the data with enough leeway in the scale to add countries outside the United States later.

In developing a geographic base file, type of data to be plotted and that might be plotted in the future must be determined beforehand.

After a file has been developed, the data must be edited to make sure that the points have digitized correctly. The more straightforward method is to plot the points at the scale at which they were digitized. Such a plot then can be overlaid on the base map and errors identified readily.

Data to be plotted, whether they result from computer manipulation of source data or are the source data, must be related to the geographic base file. A link on a highway network, for example, usually is coded to identify it from an A node to a B node. To plot this link, the X and Y coordinates of each node must be assigned to that link. This is done through development of a compatibility file.

Automated geocoding

Geographic base files were prepared in this way for more than 200 metropolitan areas by the U.S. Bureau of the Census, in conjunction with the 1970 Census of Population and Housing, local urban transportation study organizations and state transportation or higaway departments.

The geographic base files initially were prepared to facilitate automated geocoding. Since they contain coordinates they also may be

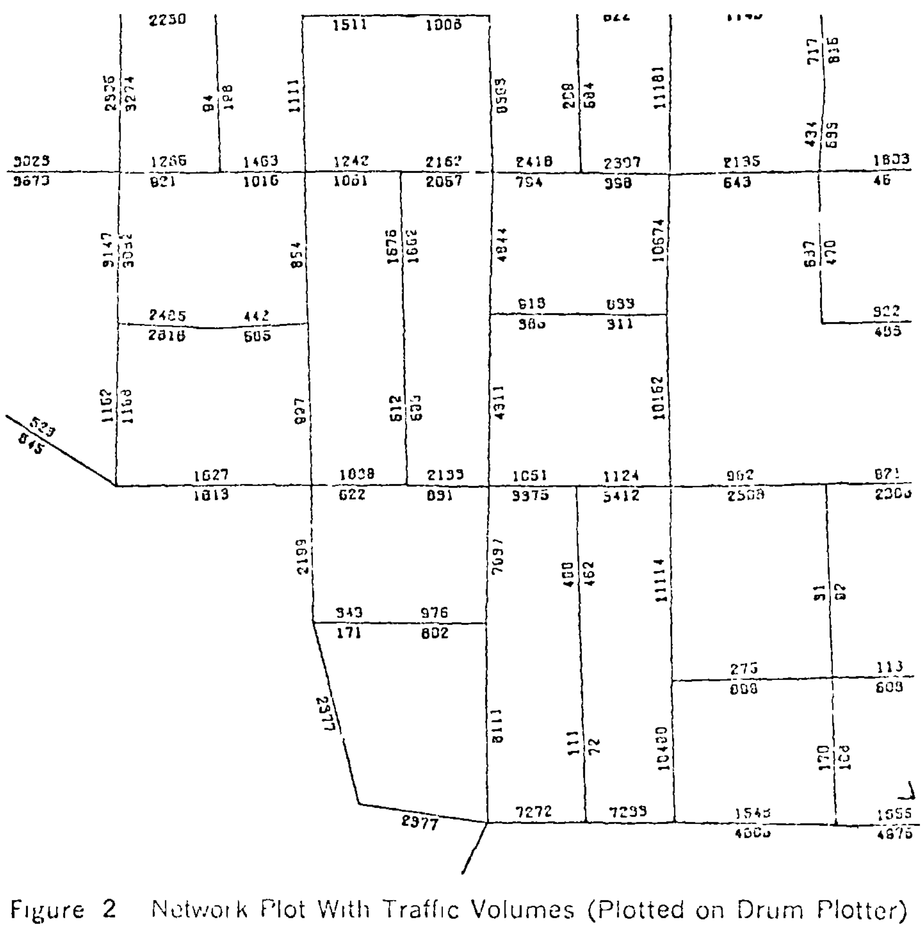


Figure 2 Network Plot With Traffic Volumes (Plotted on Drum Plotter)

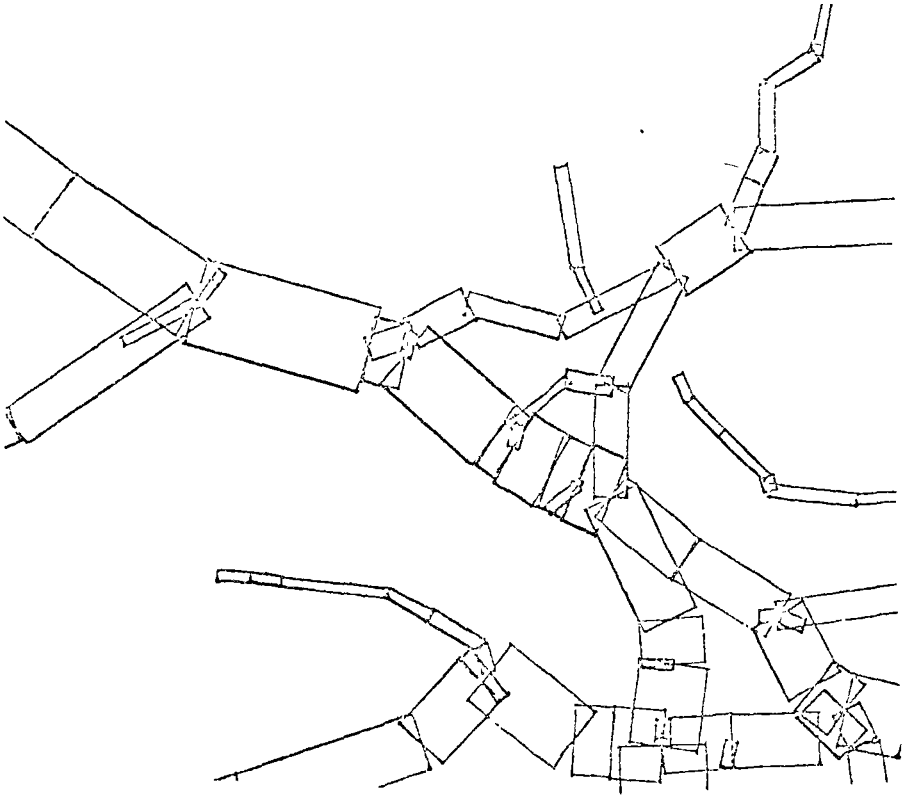


Figure 3 Graphical Network Assignment Plotter on Flat Bed Plotter 1" - 10 000 ADT (Before reduction)

used for preparing input files for automatic plotting of street networks

A questionnaire was developed to determine the use of data plotters in transportation planning The Urbanized Area Transportation Planning Programs Directory lists 233 programs and their key officials

One hundred and fifty-five officials, some responsible for more than one program, received the questionnaire Seventy-eight were returned

Survey results

Number of forms mailed	155
Number returned	78
Number of transportation studies with access to plotter—	
flat bed	7
drum	17
total	24
Number of transportation study organizations plotting	
traffic and networks	21
land use	11
graphs	13
other*	5

*Travel comparisons, census data and regression plots

Plotter used in preparation of—	
analysis	24
displays	17
Transportation studies with access to digitizers	15

The results show that about 30 per cent of the study organizations employ data plotting as an integral part of their planning process Drum plotters are more extensively used than flat bed All of the studies using data plotters develop network plots Land use plots and data graphs also are developed frequently It is interesting to note that the primary use of data plots is for analysis, with display purposes ranked second

Figures explained

Figures 2 through 9 show plots similar to those used in a number of transportation planning studies. Although the figures are printed in black, colors could be used if permitted by the type of plotter available

With the flexibility of plotting in colors, Figure 2—Network plot with traffic volumes—could be drawn by functional classification, street width, or other data in the link file

Node numbers also could be annotated The traffic volumes plotted could have come from traffic distribution models or from actual traffic counts This figure shows link assignments plotted without manual effort. Furthermore, the planner could use this method to post link distances, speeds and other data. A computer program can be used to select from the link file only that information meeting predetermined parameters The process is made more sophisticated by plotted overloaded links according to degree of overloading by using colors on the plotter.

Graphic portrayals of traffic movements, Figure 3—graphic network assignment—are drafted easily with a plotter. This type is an important planning tool.

A number of study teams have taken this type of plot and painted in the areas between lines for presentation at public hearings Projections to various target years can be shown on the same chart in different colors, indicating increases or decreases in traffic volume in a corridor over a period of time.

Many studies have developed and are using spiderweb networks shown in Figure 4 to analyze traffic movements Network plots are extremely valuable in the editing process Uncoded links and links that have coding errors can be found visually In large networks it is nearly impossible to find link errors, especially errors such as link classification

After performing computations on the computer for analysis, it is possible to graphically plot the data as shown in Figure 5—volume-to-capacity ratios. This plot originally was in color, one color representing a volume-to-capacity ratio for a link, e g 0 5 to 0 8, 0 8 to 1 0 The link width denotes the low or high end of the range

Figure 6—SYMAP—uses the computer printer to fill in areas by printing various computer characters The result is a map with greys of various intensities Such a map can be developed rapidly on a printer and much more economically than by plotter

Both table and drum type plotters fulfill the need for drawing lines, but are uneconomical in coving an area to indicate intensity

Figure 7—area by land use class—shows a unique method developed by a study team for quantifying area data. The digitized data is a series of points encompassing a land use parcel. The plotter drew the outlines around each area and consequently drew the same line twice. In this process the computer has a description of each area. Programs were developed to allocate any type of area data to the area digitized. The land use classification shown in the figure can be quantified automatically into traffic zones, if the traffic zones have been digitized. In addition, census data with its tracts or enumeration districts, also could be quantified by traffic zones and cross-referenced with land use data.

A number of studies present data as bar charts. Figure 8 is a plot of town boundaries as a base map on which data are depicted.

Figure 9—daily traffic volumes—shows daily traffic counts from a toll gate plotted for a year. The computer calculates the Y value as a ratio of daily volume to the peak day's volume. This type of plot helps the user understand traffic variations.

New techniques for computer data plotting are developing rapidly. Harvard University's laboratory for computer graphics and spatial analysis lists five computer mapping programs for use in trasportation planning.

SYMAP

Produces maps on a line printer to depict spatially variable data geographically. Contour, conformant and proximal maps can be produced with a wide range of options.

SYMVU

Generates three-dimensional line drawings of continuous surfaces on a pen or CRT plotter. The program is used to prepare perspective views of "surfaces" implicit in various types of maps produced by the *SYMAP* program. Capabilities include hidden line removal and user definition of image size, vertical and horizontal rotation.

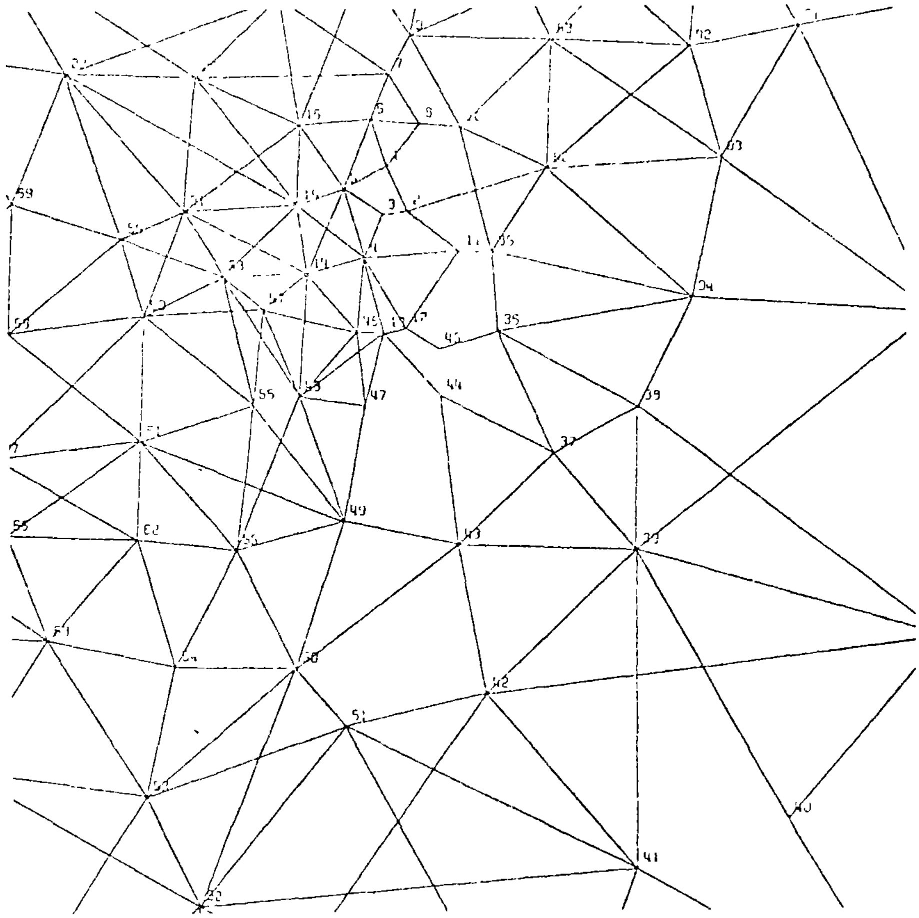


Figure 4 Spiderweb Network (Plotted on Drum Plotter)

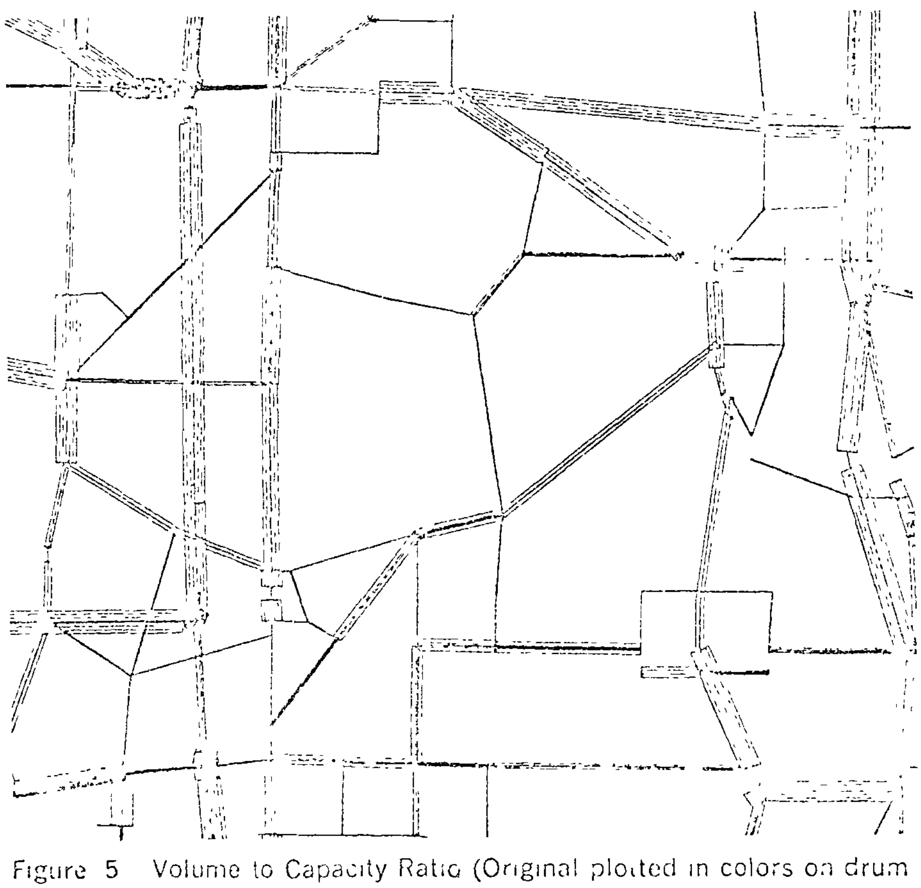


Figure 5 Volume to Capacity Ratio (Original plotted in colors on drum plotter)

Figure 6 SYMAP—Shading Map Showing Five Income Levels (Plotted on printer)

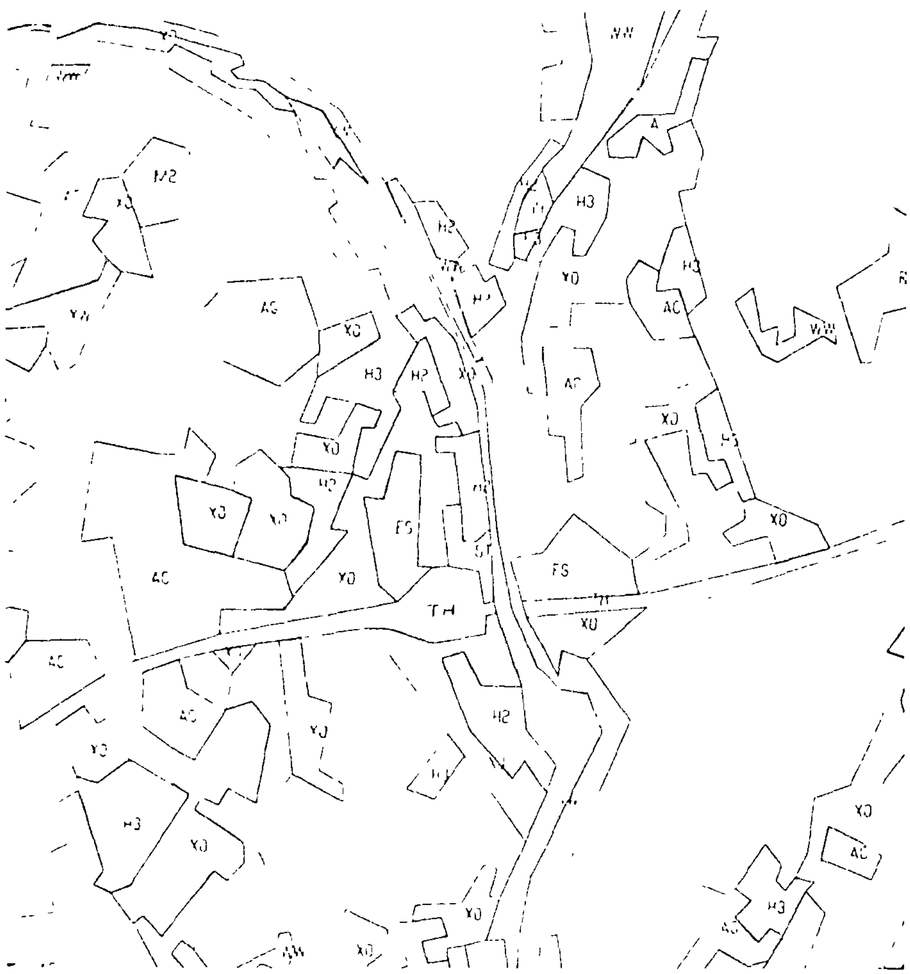


Figure 7 Area by Land Use Class

GRID

Uses a line printer to display data collected on the basis of a rectangular coordinate grid. An array of *GRID* values are categorized, assigned graphic symbols and displayed as a map. The program accepts land use, census or other data aggregated to a grid.

OTOTROL

Produces three-dimensional line drawings of an object, such as an architectural structure, viewed from any point in space. The program does not solve the hidden line problem but it is fast and highly flexible. The user can repeat all or part of the data at any scale and at any vertical or horizontal rotation from the same input data.

CALFORM

Produces shaded conformant maps on a line plotter by using symbolism, usually as shading to represent data values associated with geographic areas, such as census tracts, municipalities or counties. The 1972 National Transportation Study makes extensive use of a three-dimensional perspective histogram plot in presenting its results.

Promising technique

The field of interactive computer graphics opens a new, promising analysis technique for transportation planners involving graphic output and input, as well as direct communication with an analysis system.

Benefits of this direct communication include time savings, development of better alternatives and cost savings. The amount of the benefits will vary with the sophistication of the hardware. Early applications of interactive computer graphics were made with a cathode ray tube, a light pen, a function console keyboard and a computer.

The first efforts occurred in the early 60's. More sophisticated equipment is available today but the principle is the same. Interactive graphics have been applied successfully to design problems including highway design. Several applications also have been developed in urban planning.

Research failed to disclose cur-

... use of interactive graphics in transportation studies. Transportation network manipulation programs have been developed at the University of Washington, however, allowing networks to be built and modified. Research also has developed a network generator and an interactive transit systems analysis package. MIT and the University of Illinois at Chicago Circle also have done intensive work in this field. More research and program development is anticipated.

Conclusions

Computer data plotting is an integral part of many urban transportation studies. Application of computer technology allows transportation planners to analyze and display massive data files quickly and efficiently.

Techniques exist to prepare plots of transportation networks, network loading, land use files, socioeconomic data, and almost any other data that can be coded geographically. New techniques being developed featuring three-dimensional plotting, computer shading and interactive graphics, show the greatest promise.

This report was prepared by Committee 6B-I, Department 6, Transportation Planning, ITE Technical Council. Committee members were Edward G. Bates Jr., chairman, Robert E. Barraclough, William Harting and Ernest Nussbaum.

During development of this report, Department 6 was headed by John D. Edwards (1969-1970), David K. Witheford (1971-1972) and Walter A. Barry Jr. (1972-1973). Standing committee members, in addition to the chairmen, are Charles C. Crevo, George T. Lathrop, James J. McDonnel, Henry L. Peyrebrune, Henry D. Quarby, Walter Rainville, Vukan R. Vuchic and George V. Wickstrom.

Committee 6B-I expresses its appreciation to the following organizations for supplying graphics: Connecticut Department of Transportation, Harvard University, Tri-State Planning Commission, New York City, New Hampshire Department of Public Works and Highways and the Federal Highway Administration.

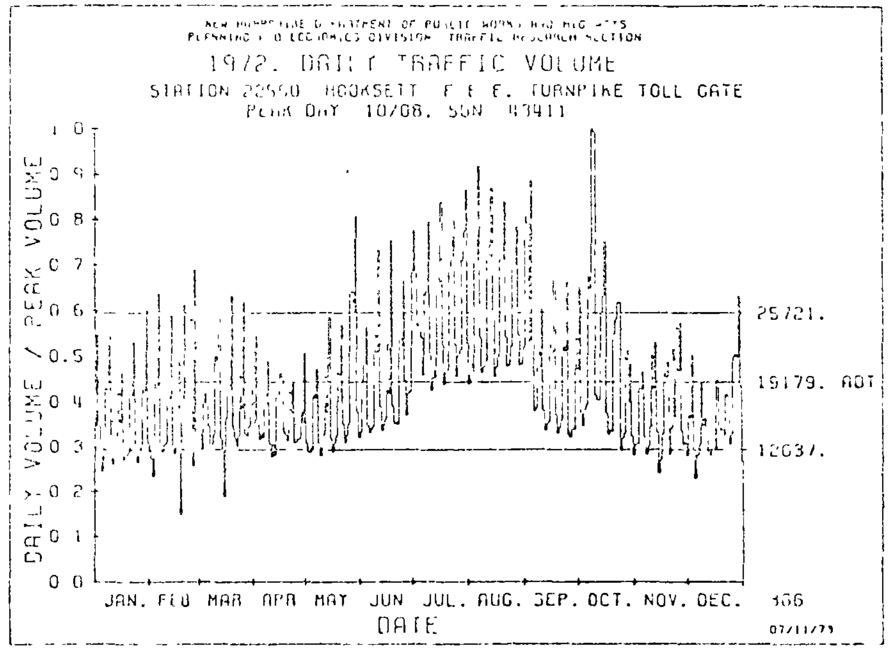


Figure 8 Bar Chart of Population

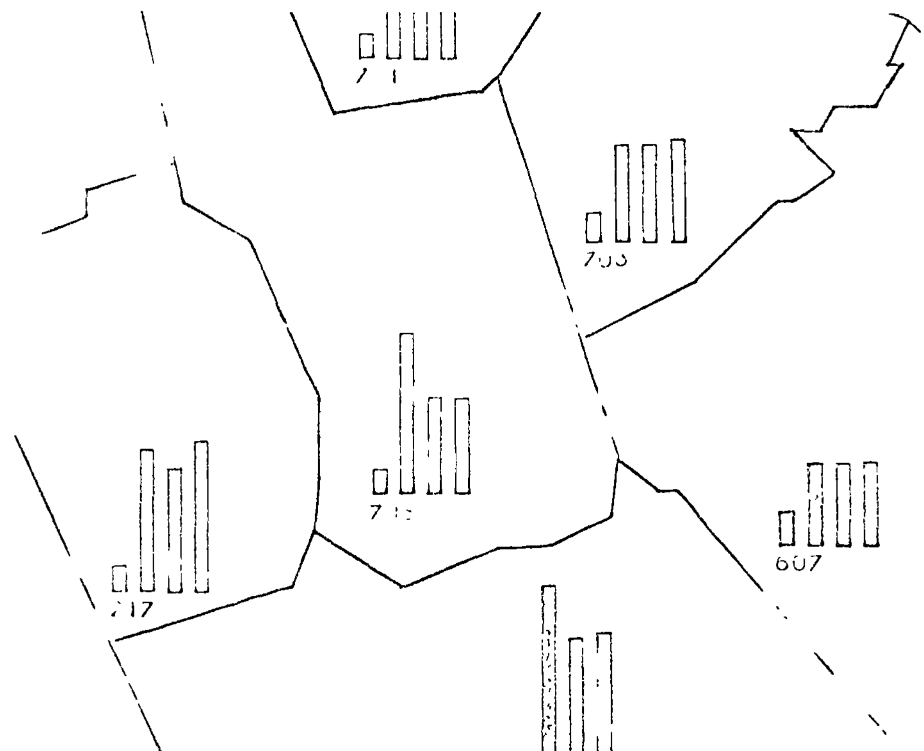


Figure 9 Daily Traffic Volumes (Plotted on Flatbed Plotter)

COMPUTER GRAPHICS IN URBAN AND ENVIRONMENTAL SYSTEMS

by
R. L. PHILLIPS

Reprinted from
Proceedings of the IEEE
Volume 62, No. 4, Apr. 1974

pp. 437-452
Printed in the U.S.A.

Computer Graphics in Urban and Environmental Systems

RICHARD L. PHILLIPS

Abstract—The key role played by computer graphics in urban, regional, and environmental information systems is surveyed by reference to specific examples. While the interactive cathode-ray tube terminal is found to be the most valuable medium, other output devices such as printers and plotters are discussed as well. Elements of computer cartography and data base organization are introduced, followed by discussions of urban information and transportation network design systems, a regional analysis system, and air and water quality information systems. Specific recommendations are given for further research in data base design and display techniques, and for promising new hardware for use in geographically distributed data systems.

I. Introduction

A. Overview

FOR MANY AREAS of investigation that rely upon computer-aided analysis, the use of computer graphics is a luxury—a luxury in the sense that, while invariably valuable, graphics affords only secondary confirmation of the validity or utility of a solution. This situation is often true in the exact sciences where, for example, an absurd solution to a differential equation can be rejected by applying our knowledge of governing physical laws, or where an optimal engineering design can be assured by the application of rigid minimax strategies. Graphics, in such cases, serves primarily as an effective means of displaying the results of a computation, but probably plays little or no intermediate role in realizing the solution. Obviously, the spurious result will be recognized all the more quickly with the aid of graphics, but this is a weak application of a powerful medium.

In this paper we describe an area where computer graphics plays an unusually important and integral role in solving problems. Indeed, some will say graphics is *essential*, not just helpful, for the solution of this type of problem. Broadly, the subject is geographical information systems, but specifically, those systems that pertain to urban and environmental problems. The methods of solution are not usually based upon the exact sciences. In these systems one must deal with such imponderables as social attitudes and political exigencies. There may not be any "right" answers or optimal strategies. Moreover, the user of these systems will generally not be knowledgeable in computer usage; he will be a school superintendent, city planner, or other government official. Therefore, it is imperative that he be presented with alternative solutions to his problem in the "universal language" of graphics.

The aim of this paper is to indicate how computer graphics is being used in urban and environmental systems and related areas. By design, this paper is not a comprehensive survey of such systems. Some applications are not discussed or mentioned, simply because their use of computer graphics does not materially differ from those presented. Before discussing any of the systems, however, we shall define certain device terminology and present some fundamental notions of cartography.

Manuscript received December 13, 1973.

The author is with the Computing Center, University of Michigan, Ann Arbor, Mich. 48105.

B. Graphical Display Devices

Our primary focus in discussing graphics will concern the following four devices.

1) Line printer—every computer installation has one; printer graphics is therefore available to all. In spite of the relatively poor resolution of the printer, plots can be produced quickly and inexpensively. Moreover, it is possible to produce a useful gray scale by selective overprinting, a feature which is valuable for thematic mapping. The printer is a batch-oriented device, and is relatively useless in an interactive system.

2) Digital plotter—marketed by a variety of manufacturers, e.g., Calcomp, Xynetics, Complot. These devices employ a digital stepper motor to move a pen over paper mounted on a flatbed or drum surface. Extremely precise line drawings can be produced by digital plotters at relatively high speed for reasonable cost. Some plotters have multiple-pen arrays for the production of multicolor output. This, too, is a batch-oriented device, and is useful in interactive systems only for producing after-the-fact output.

3) Microfilm recorder—known generically as a COM (computer output to microfilm), this batch-oriented device employs a cathode-ray tube (CRT) display and a high-speed camera to record the output. The camera is usually of the universal variety, i.e., it can be configured to accommodate microfiche or movie film of various sizes. While the final picture quality is not as good as that produced by a digital plotter, the recording speed is much faster, and it provides the capability of producing computer-animated films. The cost of these devices, however, is significant, with some systems costing over 500 000 dollars. Some manufacturers are Stromberg-Datagraphix, Calcomp, and Singer.

4) CRT display—the essential device for interactive work. The CRT terminal allows the user to see an immediate graphic representation of his problem's current status. The refresh-type CRT terminal permits the user to see dynamic or incremental changes on the display, while the storage-tube CRT terminal requires total screen erasure before a fresh display can be generated. The storage-tube terminal, however, can display a more complex scene, and is from three to five times less expensive than the lowest priced refresh display. A discussion of the characteristics of both types of terminal is given by Newman and Sproull [1], while details on CRT's and associated hardware are treated by Davis [2] and Poppelbaum [3].

C. Computer Cartography

At the heart of any display of regional, urban, or environmental information is the map. All the systems we shall describe deal with geographically distributed data, and a map of one kind or another is the most natural vehicle for presenting computer-generated solutions or situations. Certain terminology from cartography will, therefore, be useful in the sequel.

The familiar road map or the political map each belong to a class of maps which we shall call reference or *base maps*

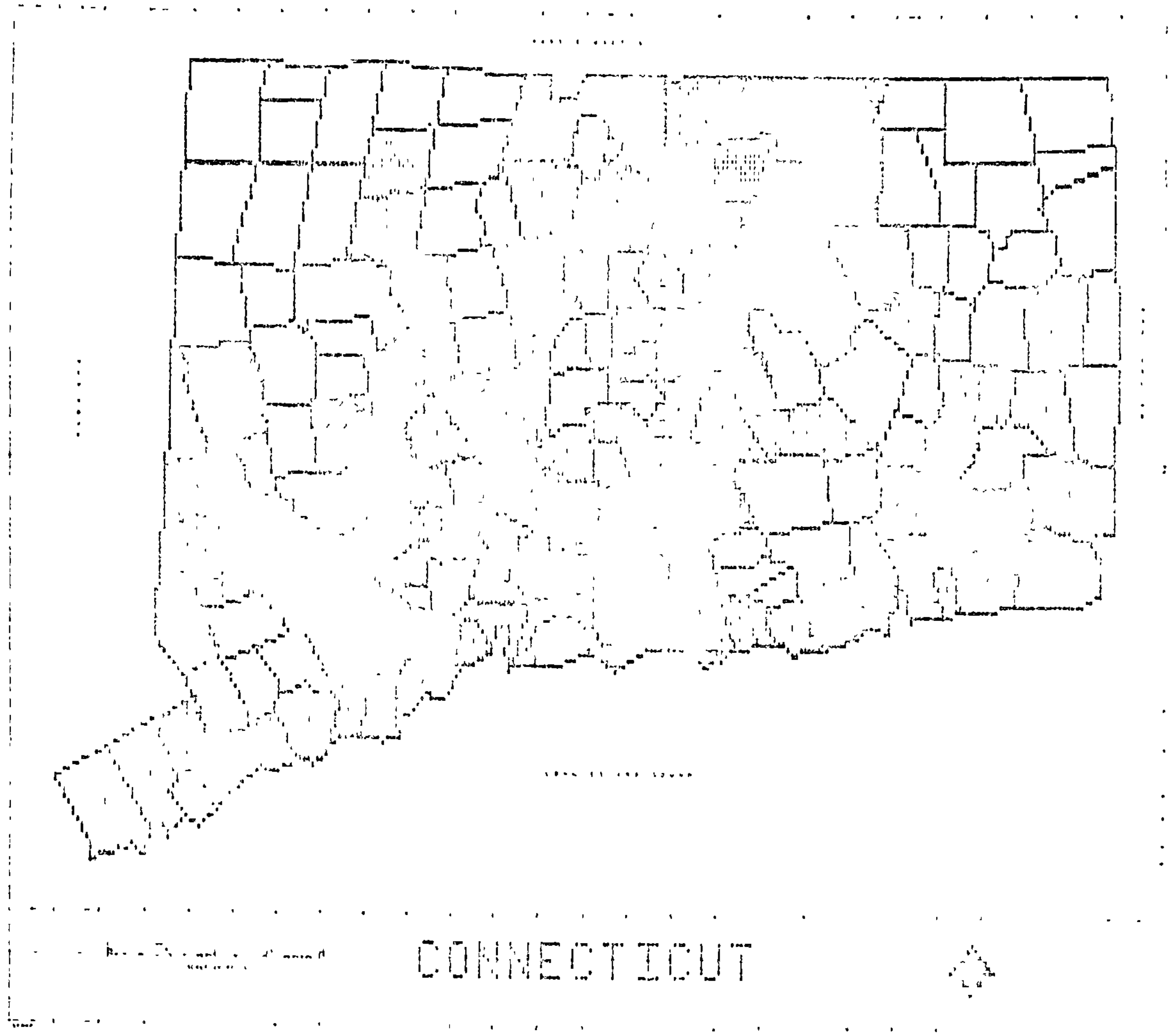


Fig. 1. Choropleth map of Connecticut. (Scale in miles.)

They show features called *geographical entities*, such as county boundaries, road locations, and the courses of rivers. Base maps are useful as locational aids, but for the applications to be discussed in this paper the *thematic map* is the most valuable. These are part of a broad class of maps, whose purpose [4] is to communicate geographical concepts such as the distribution of densities, relative magnitudes, gradients, spatial relationships, movements, etc.

The main task in thematic mapping is to represent on a two-dimensional diagram a third dimension which, in general, represents some statistical quantity. The term *statistical surface* [4] has been applied to this three-dimensional entity, and a variety of techniques have been devised to represent it. A widely used method is called choropleth mapping [1]. An example of a choropleth map is shown in Fig. 1, which indicates by town the average lot size in square feet. In general, darkness correlates with magnitude. The choropleth map's primary objective is to symbolize the magnitude as they occur within the boundaries of a general enumeration district—towns in the case of Fig. 1.

Fig. 1, incidentally, is a good example of a printer plot map. It was produced by a mapping system called SYMAP, distributed by the Laboratory for Computer Graphics and Spatial Analysis at Harvard University, Cambridge, Mass. The Laboratory publishes *The Redbook* [8] which describes several available programs and provides an overview of the various projects that have used the programs.

When representing a statistical surface so as to emphasize its gradients, *isarithmic* mapping is used. A familiar example is the contour map showing the relative elevations and gradients of a land form. Population potential in New York is depicted by isarithms in Fig. 2.

Often the best method of depicting the statistical surface is to use a perspective projection with a suitable viewing point. This can be a relatively expensive computing process, especially when hidden lines are eliminated, but the visual impact and the amount of information transferred by such a display are considerable. A perspective view of the topography of the area of Greater Vancouver, British Columbia, is shown in Fig. 3. When using an interactive display device, the value of such a presentation is even greater because the user can alter his viewing point in order to maximize the information content of the display.

Two other types of thematic mapping [6], although not used in any of the information systems surveyed for this paper, deserve serious consideration for future work. The first involves the deliberate distortion of conventional cartographic projections [7] to study functional rather than metric space relationships. Conventional maps are scaled in terms of physical distance (e.g., miles) while travel distance measured along established routes or between known points, may be the primary concern of the map user. This would be an especially valuable concept for studies of allocation analysis or transportation system design. An example of

Fig. 2 Isarithmic map of New York State (Scale in miles (top), in kilometers (bottom))

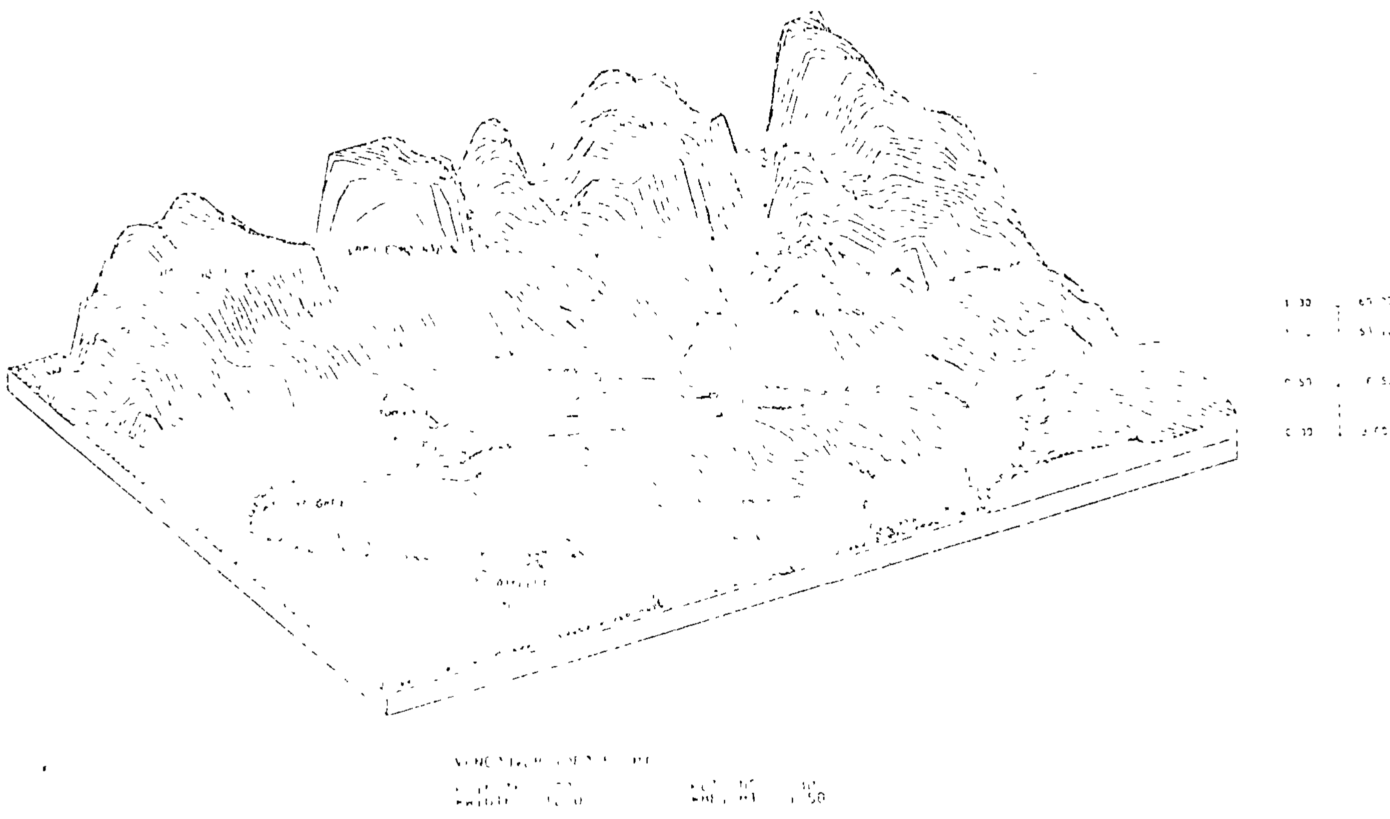


Fig. 3 Perspective view of topography of Greater Vancouver, B. C., Canada

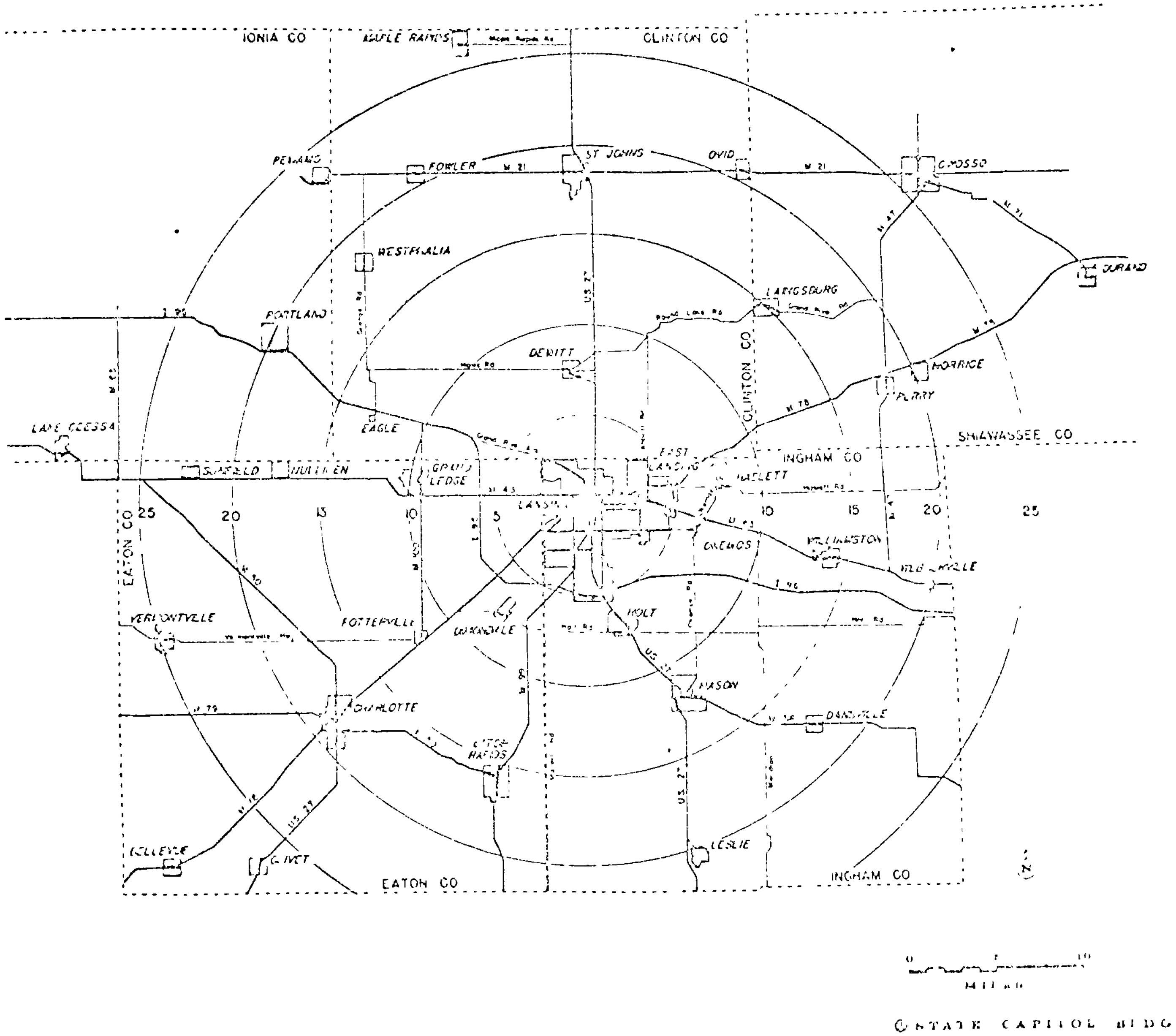


Fig 4 (a)

such a map is shown in Fig. 4, where the projection has been distorted according to a travel-time metric.

A second form of nonstandard map representation involves the deliberate distortion of area. Here one attempts to preserve the shape and contiguity relations of areal units, but conventional physical area is replaced by some areal quantity, such as total population. This type of map is called a *cartogram*, and can be used to display a wide variety of phenomena. A good example is shown in Fig. 5, where areal distortion according to retail sales volume results in an unusual view of the United States. The potential value of these two unusual mapping techniques is pursued in our concluding section (Section V).

For the computer graphics specialist, computer cartography presents some interesting and challenging problems. These include map generalization, distinguished feature identification, name placement, and methods of map digitization, storage, and retrieval. For more information the reader may consult a comprehensive tutorial [8], and a bibliography [9] for additional reading.

D. Data Base Organizations

In what follows we shall refer frequently to different types of cartographic data base organizations that are employed in the design of geographic information systems. A detailed discussion of these organizations and their coupling to computer graphics techniques is important. The selection of an organization has implications not only for the way graphical data will be retrieved and reconstructed for display, but also for the scheme by which the data are originally digitized. The two basic organizations or map encoding schemes are *grid* encoding and *line* encoding.

Grid encoding is most useful for information systems where spatially distributed attributes of certain areas are of interest. This would be the case in land-use studies or regional analyses for data such as soil type, forestation, and terrain slope. Usually, not all of these data types appear on one map, one must collect data from several maps, which are often not constructed with the same scale or projection. In collecting the data, the map is overlaid with a uniform rectangular grid, and the attribute value for each cell within the grid is assigned

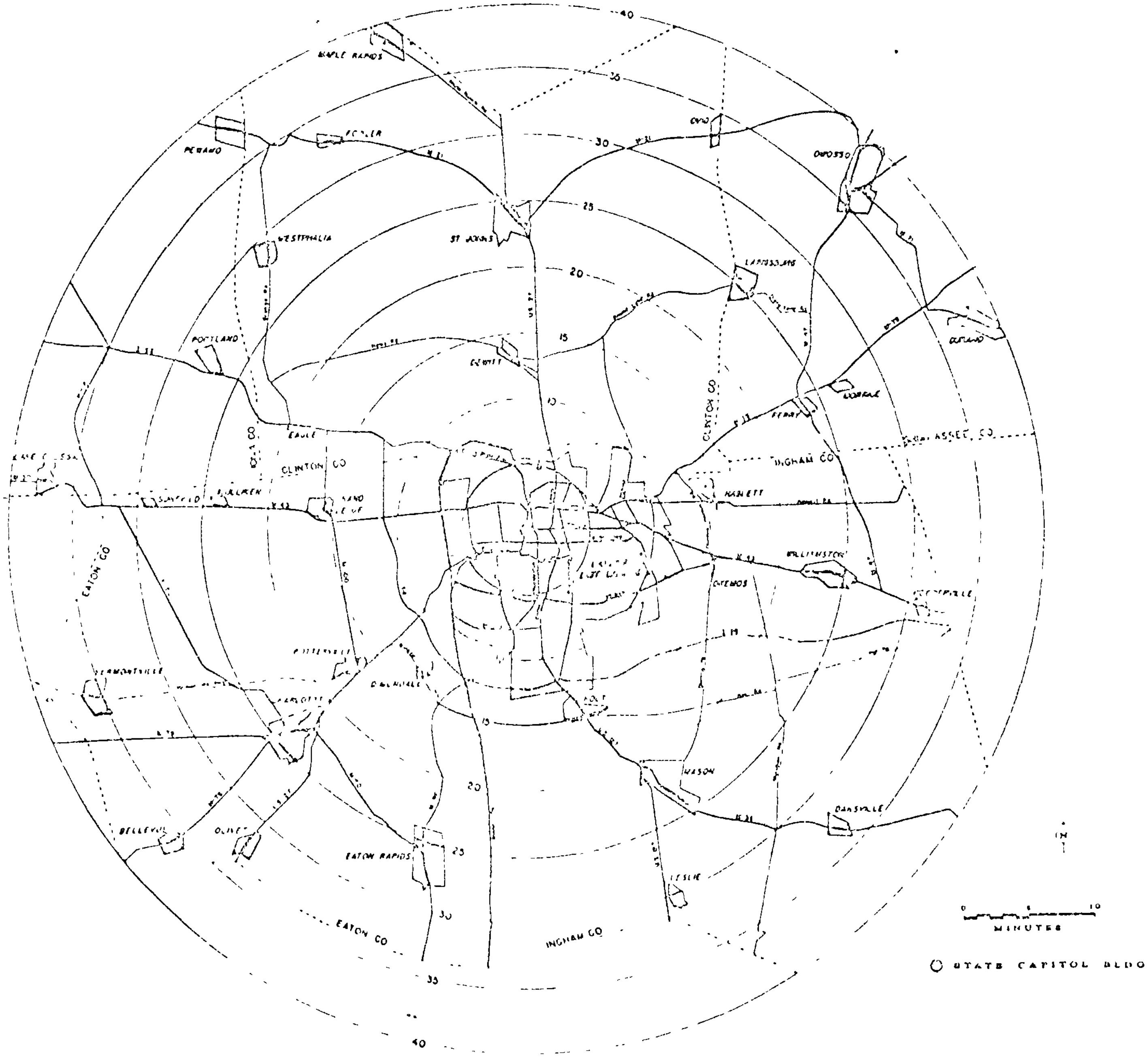


Fig 4 (b)

Fig 4 Projection distortion based upon a time distance transformation (Prepared by Graphics Section, Institute for Community Development, Continuing Education Service, Michigan State University, East Lansing)

Thus a single cell will contain only one type of tree ground cover, or soil type, the attribute is constant throughout the cell To insure accuracy, the cell size is usually chosen to be small enough that it will contain more than 50 percent of one class of a given attribute The cell size specified for a grid system has some obvious consequences The finer the grid, the greater will be the storage requirements for the data Moreover, the designer of the information system must commit himself at the outset to a level of resolution that is the finest he will ever want to consider

The grid system, on the other hand, offers the advantage of easy data accessibility Each cell is assigned an address, and the cell's various attribute values can be retrieved easily by straightforward techniques Also, mapping techniques such as SYMAP [5] are ideally suited to the representation of gridded data, the line printer is, after all a cell-oriented device

The most general type of cartographic data representation is obtained by line encoding In this system one digitizes the boundaries of regions of interest, the paths of rivers, and the routes of roads The data thus obtained form a series of x-y coordinates, stored in terms of the original map projection, or converted to longitude and latitude pairs If digitizing is sufficiently accurate, this system offers the advantage of reproducing the map in its original detail In fact, if some coarseness is tolerable, a map of even smaller scale than the

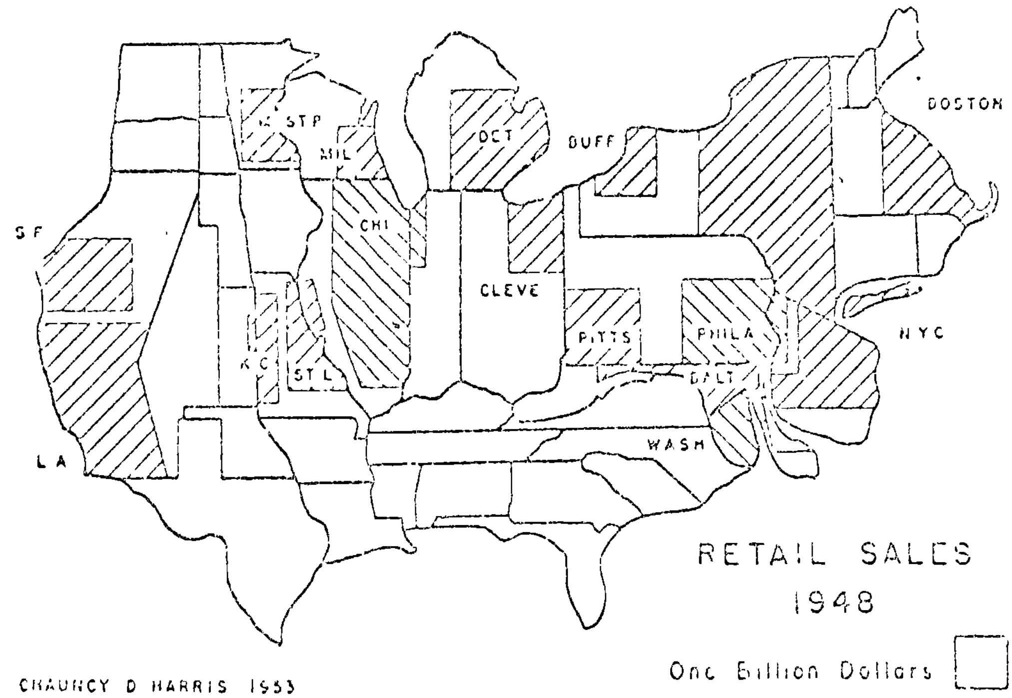


Fig. 5. Cartogram showing a market view of United States.

original can be represented at data display time. The price paid for such accurate map portrayal is having to contend with a relatively difficult data storage and retrieval problem.

Some systems employ polygons as the basic cartographic entities. This poses problems at data collection time, for it is unlikely that a digitizer, manual or automatic, will be able to reproduce a common boundary between polygons with the same digital precision every time the line is traversed. The result is two or more representations of the same line, causing slivers of ambiguity to appear due to the digitizing errors. A more workable approach is to digitize each line of interest only once and store the additional information needed to define the relation between line segments and polygons. For a small computational overhead, the desired polygons can be reconstructed at display time by connecting appropriate vertices.

Both grid and line encoding schemes have been used in the systems to be discussed, but it should be noted that there is need for considerable research into methods of cartographic data representation and retrieval, particularly for line encoding.

E. Applications Overview

Three applications will be discussed. The first concerns *urban analysis* systems, including spatial allocation studies and transportation network design. The second concerns *natural resource* information systems, including land use and regional analysis systems. Finally, the third is about systems that concentrate on *environmental* problems, especially those dealing with air and water quality.

Although we have lumped all of the systems under the general heading of information systems, there is one important distinction between them. Some of them are aimed at just query response, i.e., where various types of data are extracted from files for display. Other systems have a modeling and simulation capability as well. In fact, one system to be discussed is only a simulator, which produces information for subsequent analysis.

II. Urban Analysis Systems

A. General

The general aim of planning-oriented urban programs is to provide appropriate members of local government with information to use in making long- and short-term planning judgments. Questions related to zoning, land use, taxation, and annexation are typical of the many problems confronting urban officials. The 1970 U. S. Census is an excellent source of information for help in understanding and attacking urban problems. A series of reports published by the Bureau of the Census [10] provide details on the available data and their suggested uses. Over 260 standard metropolitan statistical areas participated in the data collection, so information exists for urban centers of all sizes in all locations.

B. Computer Output to Microfilm (COM) and Applications

An example of the active use of census data is afforded by the activities of the Comprehensive Planning Organization (CPO) [11] of San Diego County, Calif. For some time this organization had been using urban development modeling programs but without graphical output. Recently, due to the availability of census data, CPO began investigating the use of computer output to microfilm (COM) for computer mapping. Their goal is to provide the urban planner with an up-to-date planning map which can be produced at his request quickly and inexpensively. An example of such a map is shown in Fig. 6. This is a choropleth map showing changes in occupied dwelling units for a two-year period, plotted by census tract. The map was produced on a microfilm recorder using a mapping system called CALFORM, another package which is available from the Harvard Laboratory for Computer Graphics and Spatial Analysis [5]. The actual map size is 24 in by 36 in, enlarged from a 35-mm frame, for a total production cost of about 20 dollars.

Another area in which CPO finds COM useful is in the plotting of block records contained in the dual independent map encoding (DIME) files [12]. The DIME is a system of map encoding developed in conjunction with the 1970 census

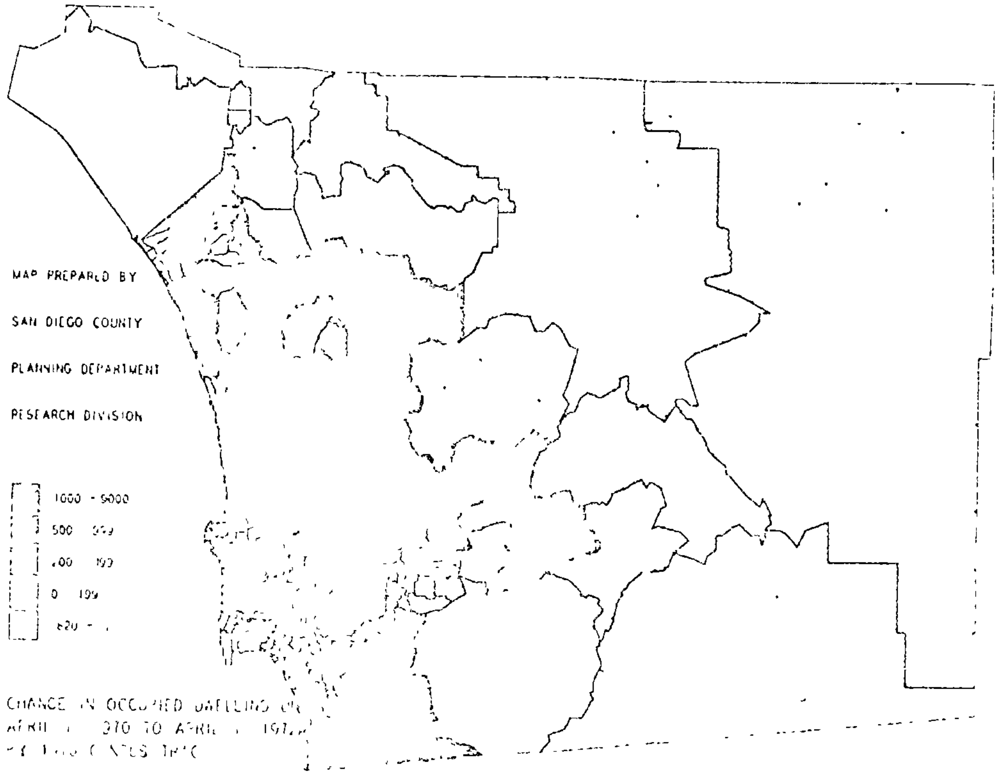


Fig. 6. COM-produced choropleth map of San Diego County, Calif.

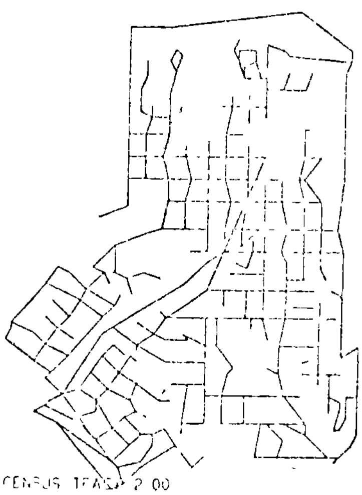


Fig. 7. Street system in San Diego, Calif., from a DIME file.

for use in displaying census data for local areas. Details of the scheme can be found in [12] and [13]. One can retrieve and plot selected lines, or plot areas that are bounded by selected lines, i.e., streets and city blocks. An example of a map derived from DIME data is given in Fig. 7, where the street system within a census tract is displayed.

Another project that involves the production of maps on COM devices is in progress at Lawrence Berkeley Laboratory (LBL) [11] of the University of California. This system is noteworthy because of the use of color in the final map production. Color keying of information has long been recognized by cartographers as an effective way of enhancing human perception of a map's content. The use of color in computer graphics has hardly been considered and the LBL project demonstrates how much a map can be improved by its inclusion.

An example of a map produced by the LBL system is shown in Plate I. Color is not produced directly on the display device but is added photographically by a color separation process. It is claimed that with this system the cost of producing a map frame can be kept as low as 2 dollars and certainly would not exceed 10 dollars. Clearly these maps are of high quality (even place names are included) considering that they are computer produced.

C. *Interactive CRT Terminals*

The use of a CRT-display terminal, whether it be storage or refreshed, enables one to interact directly and immediately with an information system. The feature of interactivity, however, places additional demands upon the designer of the system which do not arise in batch oriented systems. In particular, since rapid response is desirable, one must pay careful attention to the design and configuration of both the cartographic and the statistical data bases [15]. Sequential file organizations, which can be tolerated in batch systems, would severely degrade the performance of an interactive program.

A second consideration in developing an interactive system is the design of a simple easy-to-use query language. Since such programs are to be used by persons with diverse backgrounds, it is important to remember that the user may not be a programmer. Notley [16] argues, moreover, that in addition to being easy to use, the query language and the structure used for representing the geographical data should be integrated. Thus there can be a direct relationship between the current display and the sequence of actions that led to its creation. Some of the systems to be discussed have tried to follow this approach.

An interactive urban mapping system called INTERMAP [17] has recently been developed by a group at the University of British Columbia, Vancouver, Canada. The system uses an ADAGE refresh-type CRT graphics terminal connected

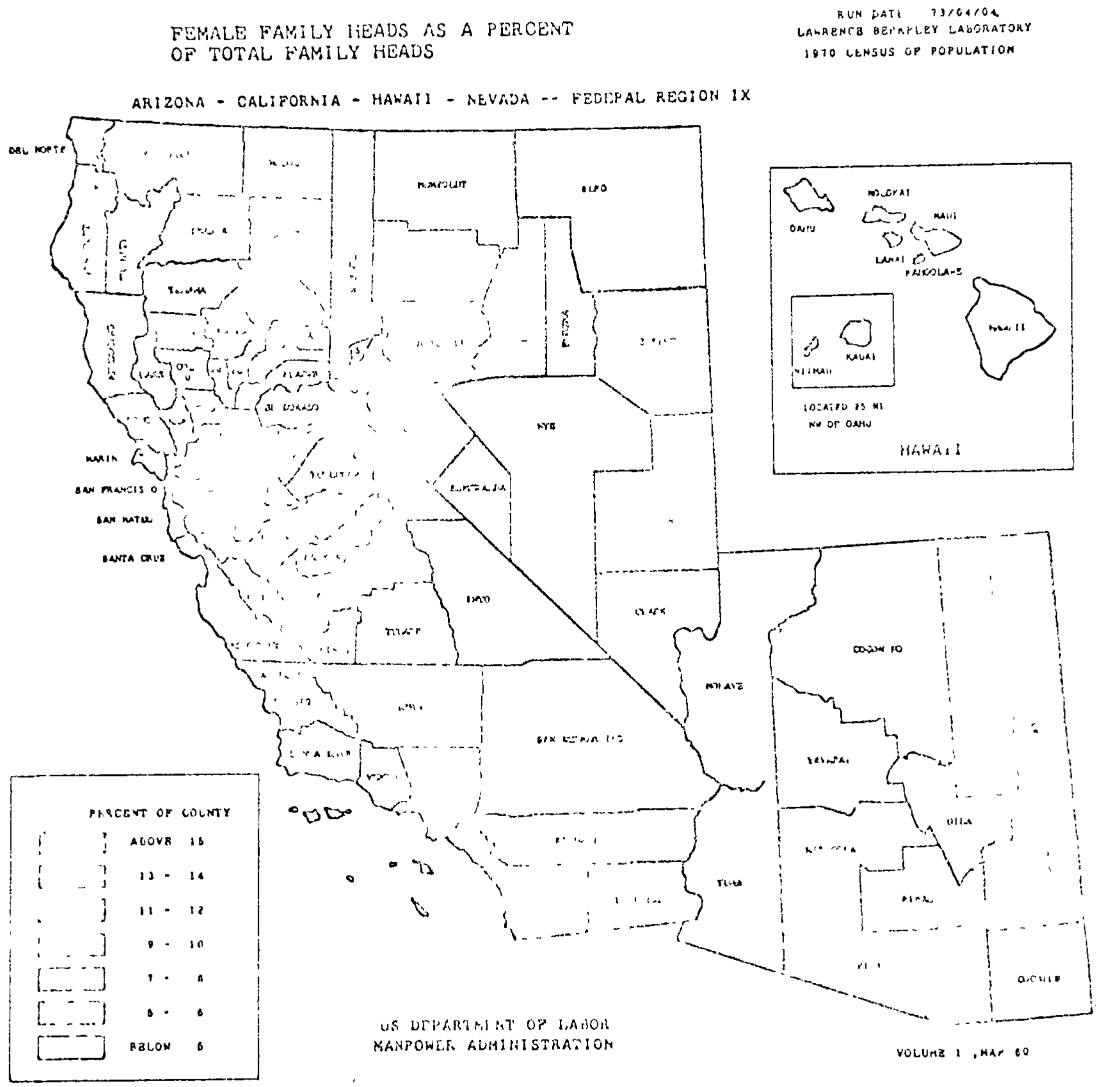


Plate I Color choropleth map of Western United States

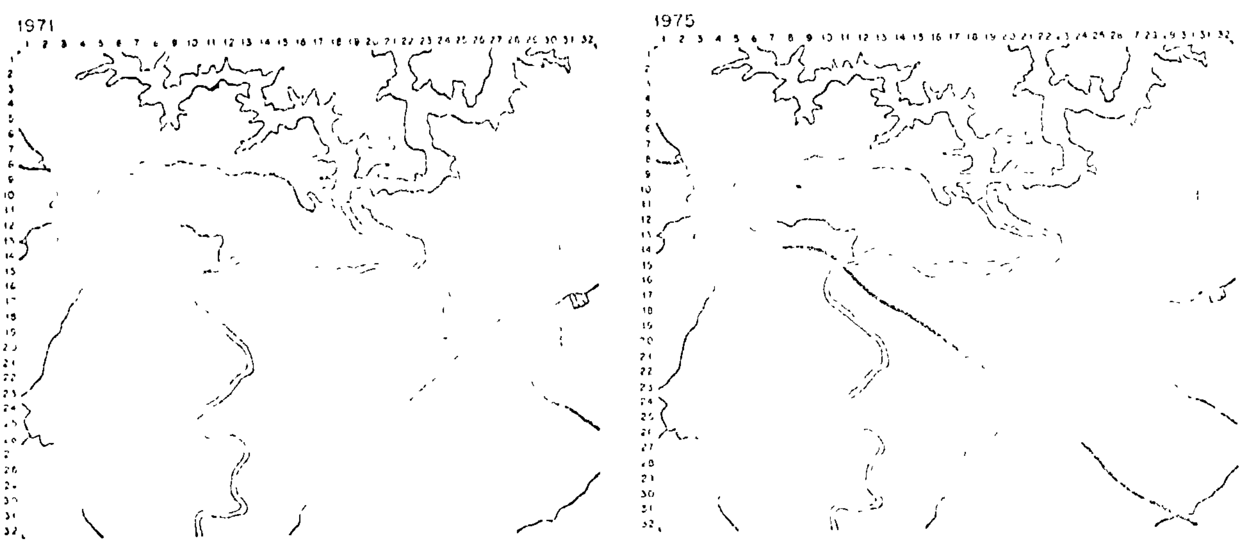


Plate II Color representation of land accessibility index

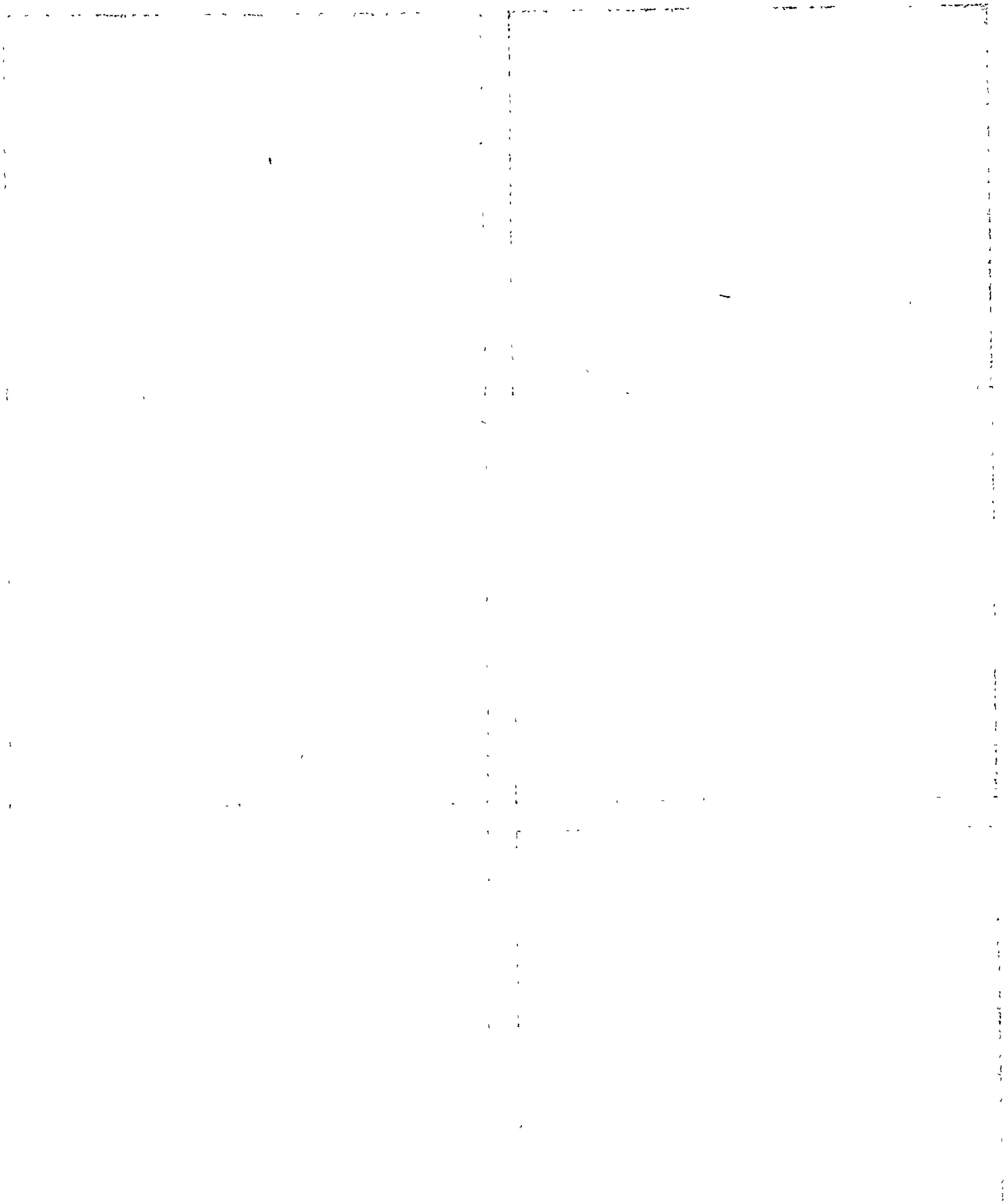


Plate III. Sequence of frames from computer animated film showing variability of SO_2 pollution in St. Louis, Mo.

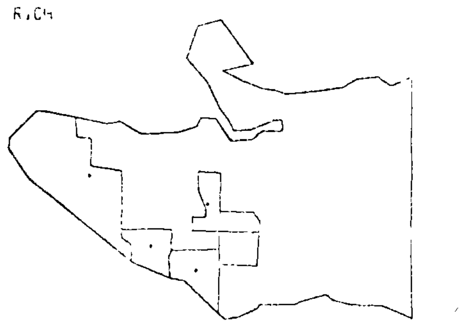


Fig. 8. Census tracts in Vancouver where average annual income is greater than 10 000 dollars.

to a time-sharing computer. The stated aim of the system is "to provide a tool to allow urban researchers to create useful maps in a way that is natural and flexible." Nottey's [16] desiderata for interactive query languages are nicely fulfilled by this system. This was accomplished by a set-theoretic approach to data retrieval where the present status of the display is represented by the system as a subset of the universe of data. Thus when minor changes are to be made to the display, the next set operator need not be applied to the universe, only to the subset.

The query language for INTURMAP is easy to learn and use. One can create an arbitrarily named set, assign to it an attribute value relationship, and display it. For example, suppose the user were interested in seeing the locations of all households where the reported annual income is greater than 10 000 dollars. The sequence

```
CREATE RICH
ASSIGN AV INC-PER-HSLD>10000 TO RICH
DISPLAY RICH
```

would result in the display shown in Fig. 8.

A system similar to the INTURMAP project, but with modeling capability, has been developed by personnel at IBM San Jose Research Laboratories in conjunction with Santa Clara County planners [18], [19]. The system, called the geodata analysis and display system (GADS) can forecast suggested planning strategies for the user's consideration. These forecasts, of course, are only as good as the modeling programs used, but even if there is doubt as to their accuracy, the forecasts can point out alternatives to the planner he might not otherwise have considered.

The GADS program uses an IBM 2250 refresh type CRT terminal. For modeling purposes Santa Clara County has been divided into 334 zones, which are based upon census tracts. A variety of attributes are assigned to each zone, and the user can ask to have all zones displayed that satisfy a user-specified combination of attributes. Zones can be aggregated to higher levels, e.g., planning areas and unnecessary common boundaries between zones can be eliminated. With such a wealth of available display alternatives, the policy innovator can review a projected pattern of land development and quickly assess its desirability. An income distribution map produced by GADS is shown in Fig. 9. The different symbols shown correspond to various criteria imposed by the four queries, which are also shown in the figure.

A different set of problems in urban analysis is being attacked by a group at the Urban Systems Research Center at the University of Washington, Seattle. The problems emphasize spatial allocation analysis, where one seeks to find

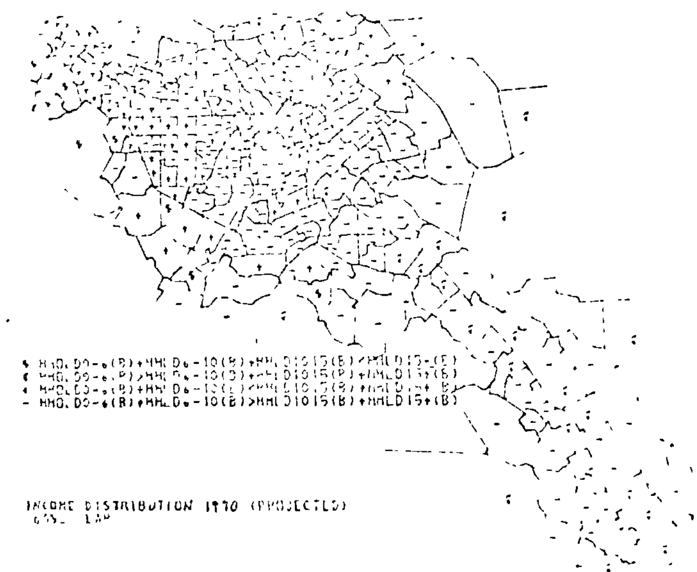


Fig. 9. Census tract display for Santa Clara County, Calif.

an efficient locational pattern for a set of facilities involving the provision of a service to a spatially distributed population. For example, one may wish to find the best locations for ambulance dispatch centers or fire stations, such that some measure of total travel requirements is minimized.

In spatial allocation problems there is a set of defined routes over which the services or commodities can flow. The graphics requirements for such a problem are to display in a meaningful way this network of permissible paths. The data base organization must be capable of defining the connectivity properties of the network, so a line encoding scheme for the map data must be adopted. The DIME [12] system's data, incidentally, would satisfy this requirement.

The group at the University of Washington has worked on a variety of applications [20], [21], including one that involved a particularly adroit usage of graphics [21]. The system is the urban transportation analysis system, called UTRANS, and its purpose is to assist the planner to structure and test a bus transit system design for an urban corridor. It is limited to the case where there are many origins and only one large destination, e.g., a central business district, a large manufacturing plant, large university, or rapid-transit rail station. The program runs on an IBM 1130 connected to an ARDS storage-tube CRT terminal. The user first specifies the location of several park-and-ride lots, specifying lot capacity and parking fees. He then locates bus stops and sets bus fares, frequency of service, number of buses on each route, and other pertinent data. The computer uses these data to produce, in the form of graphical and tabular displays, predicted measures of the bus system's performance. To evaluate means of improving performance, the user can move the location of a parking lot, increase the bus fare, etc., and quickly obtain a new evaluation of system performance. This process continues until a "satisfactory" solution is obtained. Examples of output from several stages of this program are shown in Fig. 10.

D. Digital Plotter Applications

Many of the interactive systems just discussed have the capability of routing the current screen display to a digital plotter for subsequent off-line processing. Thus the digital

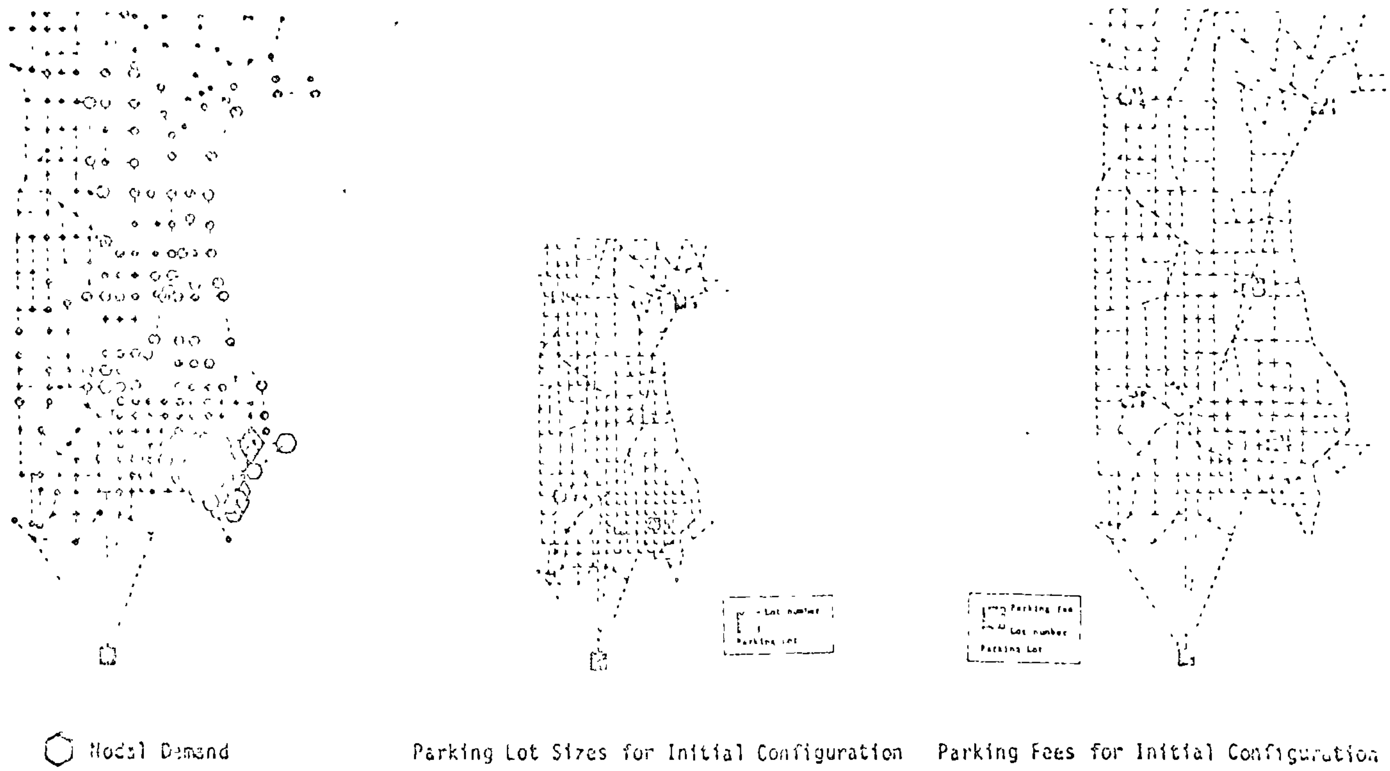


Fig. 10. Selected output from the UTRANS system.

Fig. 11. Street system for school bus routing study.

plotter is not the primary display device but serves only to produce hard copy. In this section we discuss an interactive system that is difficult to distinguish in terms of output devices; it uses them all, but the digital plotter most frequently. The system is called the urban geometric heuristics system, (UGH) [22], [23], and was developed at the Urban Systems Laboratory at M.I.T.

UGH is impressive from many standpoints. There is probably no other urban analysis system that has its flexibility and power. This system is adaptable to any type of urban analysis problem, whether it be policy planning or transportation systems design. This flexibility has its price in terms of both computer costs and ease of use. UGH relies on another large system called ICES [24] and, to a large extent, uses its analytical and file manipulating features. While the command language of UGH is not difficult, it is extensive, and the developers of the system state "in order to use UGH, the user must, of course, understand geometric problem solving, urban information, land information, and his particular problem."

The versatility of UGH is derived from the fact that it is not oriented toward census tracts or street networks: it deals in geometric entities such as points, lines, curves, and polygons. The user of the system gives meaning to those entities, in the context of his application. Thus the system is usable, not only by a planner, but by an engineer or a manager as well.

One exemplary application of UGH [25] involved a school bus routing problem. The aim of the study was to investigate methods of reducing transportation costs by improving the efficiency of school bus usage. Typical output from that study is the display of permissible routes (Fig. 11) and a map of the bus stops finally selected (Fig. 12). The school district for which this study was done actually reevaluated its bus routes resulting in a substantial reduction in transportation costs: the mileage traveled by the buses was reduced by almost [illegible] percent.

III. NATURAL RESOURCE ANALYSIS SYSTEMS

A. *General*

Regional and natural resource analysis systems differ from urban systems mainly in terms of subject matter, strategies, and goals. From a data organization and retrieval standpoint there are few differences; in both cases one is concerned with the analysis of areal units with homogeneously distributed attributes.

The regional, natural resource, or land-use planner is concerned with extrapolating land usage and its associated en-

Fig. 12 Specific bus stop locations for school bus routing study

vironmental impact, both in an assumed steady state and in the presence of actual or anticipated perturbations. Potential applications include evaluating the effect and implication of alternative land use patterns, service, utility, and transportation systems, and measuring the environmental impact on natural systems within the region.

Unlike his urban counterpart, the regional planner may deal with vast areas—an entire county or group of counties, a river basin, or a national park. His time scale for planning forecasts may stretch five years into the future, or it may span decades. Only two major systems will be discussed but the reader is referred to an exhaustive overview [26] of a variety of geographic information systems. Our attention will be focused on systems that exemplify the two major types of data organization - grid and line encoding.

B. A Grid-Oriented System

For the last few years, a large ambitious program on regional environmental systems analysis has been in progress at the Oak Ridge National Laboratory [27], [28], Oak Ridge, Tenn. Analysis is grouped into several areas, including socioeconomic, land-use, ecological, and sociopolitical data. A 6500-mi² region in the East Tennessee Valley surrounding Knoxville is being used to test the system.

For this paper the work of the land-use analysis group is of interest. An extensive system, the Oak Ridge modeling information system, called ORRMIS, has been developed which is based upon the grid or cell method of data encoding (see Section I-D). For the sixteen-county study area of East Tennessee there are 1 843 000 geodetic cells, 3.75″ latitude by 3.75″ longitude, in the data base. Data stored represent such varied information as the type of rock, soil, and vegetation.

ORRMIS offers the user a wide range of data-base query capabilities and modeling possibilities. Further, all types of display devices ranging from a line printer to an interactive CRT terminal are available to the user. An example of ORRMIS output, Plate II, depicts the accessibility index, a weighted sum of factors related to terrain slope, local roads, expressways, etc., for a small portion of the study area. The color-keyed output is clearly cellular in nature; the base map information is provided by an overlay. Other display formats, such as three-dimensional plots and computer-animated films, are also available to the ORRMIS user.

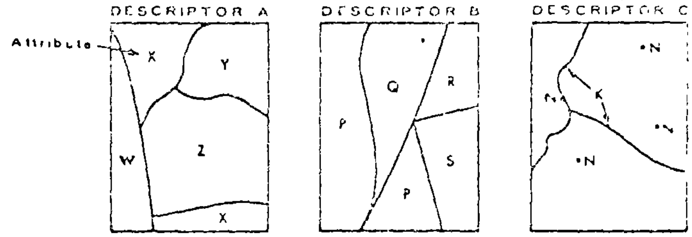


Fig. 13 Example of thematic maps for the same geographic area used for overlay processing

C. A Line/Polygon-Oriented System

The Raytheon Company (Automatic Operation) recently completed a project for the United States Department of Interior for the development of a polygon-oriented information system, the natural resource information system (NRIS) [29]. The purpose of the system is to aid managers in making decisions concerning the disposition of land-related resources such as minerals, water, vegetation, and forests, as well as animal resources such as cattle, protected wildlife, and fish. The data base for this system is derived from a variety of thematic maps, which depict regions of arbitrary shape over which given attributes are homogeneously distributed. The boundaries of these zones are digitized, point by point, so that they can be digitally reconstructed as an n-sided polygon. Recall, however, that the best approach is to digitize each line only once, and reconstruct polygons through a series of identifier codes.

The land manager will want to ask composite questions of his system. These questions amount to a conceptual overlay of any of the maps that constitute the data base. In graphical terms this poses some interesting problems, requiring clever and efficient methods of polygon processing. There has been considerable research [30] in this area aimed at performing extensive polygon processing for modest expenditures of computing time. The problem is exemplified by the three maps shown in Fig. 13.

Imagine that maps A, B, and C represent soil type, ground cover, and forest fire lookout towers. The attributes for map A might be loam, sand, gravel, and clay. The user of the system will make area-related requests such as "produce a map showing all areas where both X (map A) and P (map B) occur." From a set-theoretic data extraction standpoint this is an intersection operation. Another operation, enumerative in nature, would be to find all points N within a certain geographical area.

The NRIS system is capable of producing fairly high quality maps. The system is noninteractive and a digital plotter is used to draw composite maps such as shown in Figs. 14–16. The first two thematic maps should be viewed as problem input; they represent resource value and vulnerability, respectively. Fig. 16 is a derived map which shows ½-h fire control zones, i.e., locations that may be reached by fire fighters within 30 min to prevent the destruction of a valuable resource.

The utility and the versatility of NRIS have been demonstrated on ten USDI data bases ranging from the Great Smoky Mountain National Park planning study to the New York State recreational resource study.

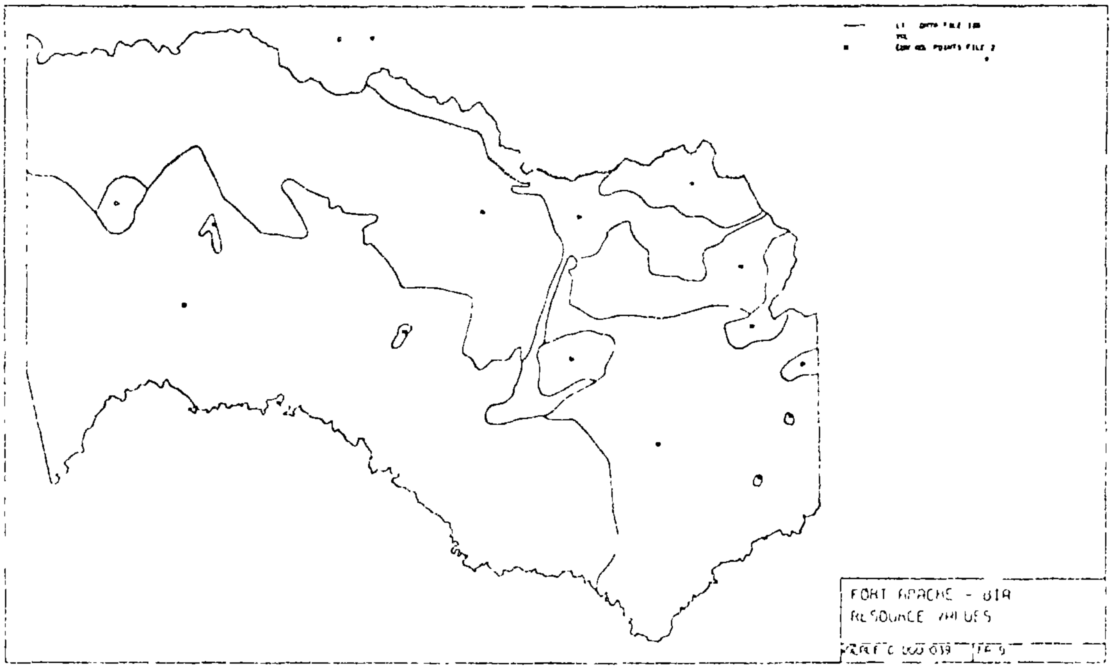


Fig. 14. Map of Fort Apache Reservation showing resource values.

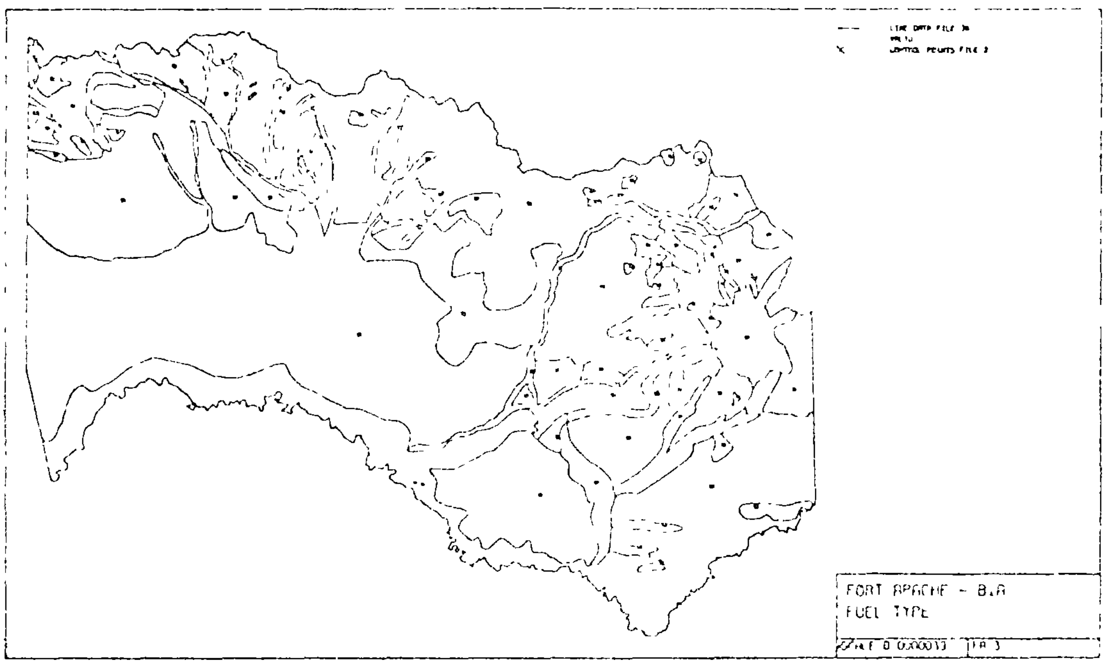


Fig. 15. Fort Apache resource vulnerability map.

IV. Studies of the Environment — Air and Water Quality

A. General

Systems that deal with aspects of air and water quality are unique among general geographic systems because the laws of the exact sciences can be applied to produce quantifiable solutions to environmental problems. In addition, the data upon which these systems rely, unlike demographic and socioeconomic data, are easier to obtain and maintain. However, even though physical laws can be applied to atmospheric and hydrologic processes, the formulation of good models and the application of appropriate boundary conditions pose formidable problems. Moreover, quality-control policies suggested by environmental systems may be difficult to implement because of the need to interact with the political and economic sectors of society. Thus, like all other systems we have discussed, the predictions of air and water quality programs serve only an advisory purpose.

B. Air Pollution Applications

The factors that affect the distribution of pollutants over an urban area are many and diverse. It is not enough to merely identify sources of pollution, but the mode of injection

Fig 16 Composite Fort Apache map showing ½ h fire control zones

of pollutants into the atmosphere must also be considered [31] In addition, the terrain configuration, the diurnal cycle, and the variable meteorological conditions all play an important role in determining pollutant concentrations Notable progress has been made in recent years in accurately modeling these processes by, among others, a group at the IBM Scientific Center [32], Palo Alto, Calif A severe test of their simulation model was recently carried out in New York City [33] This location is particularly challenging because of its unusual meteorological conditions and the complexity of the physical environment In general, good spatial and temporal agreement was obtained between predictions and observed measurements of sulfur dioxide (SO_2)

The results of the New York simulation were displayed in the form of isarithmic maps on an IBM 2250 CRT display Contours of SO_2 concentration were displayed for selected times of day and for selected prevailing wind conditions The display was static, however, because the long computation time precluded the continuous updating of the map For a problem that has significant temporal variations, it is useful to see the dynamic unfolding of events The use of computer-animated films, a display medium not yet discussed, provides that capability Schreiber [34] discusses the value of computer animation as an aid to understanding the results of complicated scientific computations He has produced a film, using the pollution model of [33], that shows the dynamic 24-h variation of SO_2 concentration in St Louis, Mo Moreover, the film has the added dimension of color, making it easy to separate base map data from computed isarithms A series of frames from the film is shown in Plate III

C Water Quality Applications

The United States Geological Survey and the United States Environmental Protection Agency (EPA) have long been collecting water quality data at locations throughout the United States. The current EPA network comprises some 150 000 monitoring stations which report chemical analyses at least monthly—and, in some cases, daily More than 2500 chemical and biological parameters are cataloged for these stations, representing more than $12 \times 10^{?}$ characters of information currently on disk storage This data base, together with its retrieval software, is called STORET [35] Various state and federal agencies have the ability to interrogate this data base and retrieve information useful for long-term water quality monitoring, the assessment of hydrological trends, and the design of expanded monitoring networks

The sheer bulk of STORET, coupled with an inconvenient mode of retrieval, nearly prohibits an interested agency from making timely and meaningful use of the available data Moreover, even if a fast, efficient retrieval program were available, printed output is certainly not the best way to present the information. Strong and significant geographical relationships exist between the data, which cannot easily be detected from a tabular output This suggests the applicability of interactive cartographic display techniques

A system called the INteractive System for Investigation by Graphics of Hydrological Trends (INSIGHT), [36], [37], has been developed to make it possible for a user to display, on a cartographic background, the locations of selected water quality monitoring stations Chemical and biological data acquired at these stations can be displayed graphically or in tabular form Being interactive, the system allows a user to investigate hundreds of water quality situations in the space of an hour or two Then, guided by the displayed information, he can pursue interesting patterns of water quality degradation, rejecting all those cases that seem to be unimportant Thus the user gains an insight which otherwise might be hidden by the volume of the data or the labor involved in interrogation by nongraphic means

Figs 17 and 18 exemplify the type of display that INSIGHT produces The terminal used in the system is a Tektronix 4010 storage-tube CRT, and the figures were taken directly from a hard-copy device attached to the screen A typical base map with hydrological features is shown in Fig 17 The State of Michigan is shown in Fig 18 as it would appear observed from about 400 mi above the state and looking toward a compass heading of 10° The vertical bars rising from the southeastern part of the state represent minimum, maximum, and average values of dissolved oxygen measurements made at the selected stations during their respective periods of record The display method affords a complete geographic-ecological picture of the water quality conditions in selected areas

INSIGHT is significant as 1) it produces good quality displays on an inexpensive terminal, 2) it runs on a multiuser system, not on a dedicated computer, and 3) it is being used routinely and productively by persons who are not computer specialists Collectively, these three points epitomize the attributes of a successful geographic information system.

V Conclusions and Future Research

A Overview

The most formidable obstacle to more effective use of computers in solving urban and environmental problems is the lack of effective "interfaces" between the user and the programs he uses The interface problem is especially acute because *the user is not computer knowledgeable* he is a city planner, a transit system official, or a school superintendent In contrast to the progress in developing the geographically relatable data bases needed by urban and environmental systems, there has been almost no progress either in presenting data in a form useful for problem solving, or in providing the functions to help transform data into useful information Report generators, statistical packages, and simple query systems represent the state of the art These facilities are

Fig. 17. Reference map of Florida showing major hydrological features.

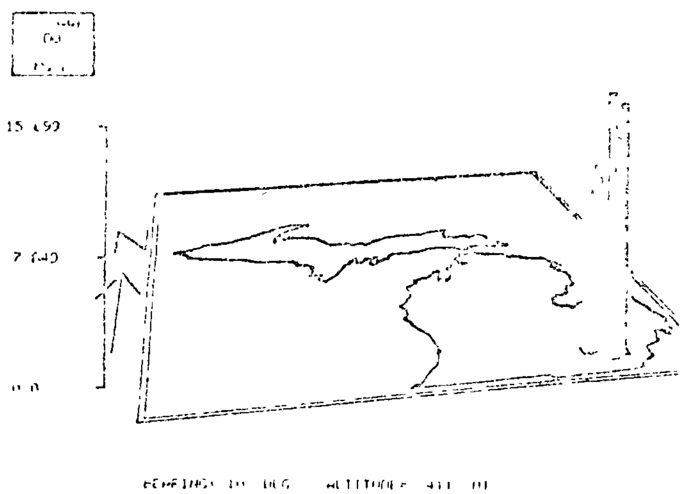


Fig. 18. Perspective map of Michigan showing superimposed water quality data.

usually batch-oriented and inflexible, and require programming skills for effective use. As a result, most societally oriented computer use has been merely for clerical applications, so far too little use has been for problem solving.

One thing that is clear is that computer graphics must play the dominant role in reducing the impedance of the interface between user and data. Aside from the fact that all systems we have considered deal with geographically distributed data, best depicted on a map, it is imperative that many alternative solutions and strategies be presented concisely to the user. This can be achieved through the most perceivable medium—graphics.

In addition to employing perspicuous display techniques, urban and environmental systems must be *truly interactive*. A batch system will be of little value in presenting an array of solution possibilities to the user in such a way that he can react to them immediately. Thus *interactive graphics* is the means for bringing together computer-based data and problem solving users not having a programming background.

Several hardware and software problem areas have been identified, which relate to the effective use of interactive graphics and which will benefit from further research and development.

B. Hardware

Bearing in mind that computer cartography is at the heart of any of the displays considered, we should look toward equipment that will satisfy the axioms of the map maker. One of these concerns detail. Since detail implies the display of a large amount of information, one should consider the new large-screen generation of storage tubes (Tektronix 4014 and 4015) for cartographic applications. It is not reasonable to expect a refresh-type CRT terminal to display a complicated map without exhibiting flicker. Even though dynamic graphics is not possible on a storage tube, it is attractive because of its large capacity for data display. Moreover, the Tektronix 4014 offers higher resolution, 4096 by 4096 points, than is currently available on most other CRT's storage or refresh.

Color has always been important to the map maker; it is no less important for computer-drawn maps. Available color displays are expensive and, particularly for the raster-scan type, the resolution is poor. Moreover, the shadow-mask type of color tube used in present systems always suffers from convergence problems. A new and promising color tube was recently announced by CPS, Inc. It is a single-beam color penetration tube with extremely high resolution. There is as yet no complete display system which incorporates this tube, but with increased use and demand for environmental data mapping it seems likely that soon such a unit will appear on the market, hopefully at a modest price.

C. Software

From a software standpoint many of the systems discussed have not taken sufficient advantage of existing display techniques. For example, not one of the systems surveyed makes use of projection distortion (Fig. 4) to represent pertinent nondistance metrics, the accessibility map from ORRMIS (Plate II) might be more profitably represented in this form. The use of a cartogram, a map where area is is purposely distorted (Fig. 5), would certainly be of great value for urban and regional analysis systems displays. Imagine, for example, the income-distribution map from GADS (Fig. 9) displayed as a cartogram; how much meaningful would be areal distortion rather than the use of symbols. Another useful display technique is the three-dimensional choropleth mapping routine recently devised by Tobler [38] and shown in Fig. 19.

All of these techniques come under the heading of perspicuous data displays. Their realization is not always computationally trivial. For example, for the functional metric map shown in Fig. 4, a road network problem had to be analyzed to produce the anisotropic measures of travel time. To produce a proper cartogram, a set of nonlinear partial differential equations must be solved [39]. Even for a simple map this can be time consuming. Finally, the three-dimensional choropleth map (Fig. 19) requires a large amount of hidden-line processing which, even after so much research in the area, is still not a trivial problem. Since *interactive graphics* is our goal, further research is needed to expedite the computations before such techniques can be effectively exploited.

Additional research in query language design is required. It should be simple and natural for a user to request the display of cartographic information, as well as graphs, tables,

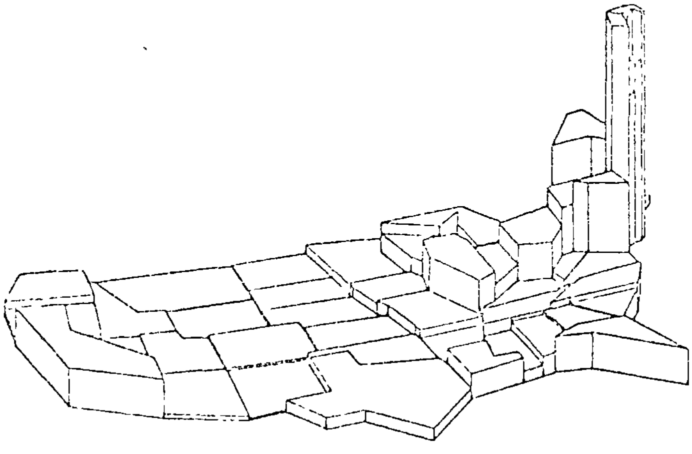


Fig. 19. Three-dimensional choropleth map of U.S. population density

and histograms of operational data. In this regard, the effective use of graphical input devices such as tablets, lightpens, etc., remains to be fully exploited.

Additional work is needed on the encoding, storage, and retrieval of cartographic data. Conventional data base design deals with one-dimensional items, while map data are at least two-dimensional. New approaches for search and retrieval are required [40]. Again, the guiding goal of providing interactive graphics systems implies that speed and efficiency of all software components is of the utmost importance.

References

[1] W. M. Newman and R. F. Sproull, *Principles of Interactive Computer Graphics*. New York: McGraw-Hill, 1973.

[2] S. [illegible], *Computer Data Displays*. Englewood Cliffs, N. J.: Prentice-Hall, 1969.

[3] W. J. Poppelbaum, *Computer Hardware Theory*. New York: Macmillan, 1972.

[4] A. H. Robinson and R. D. Sale, *Elements of Cartography*. New York: Wiley, 1969.

[5] W. Warntz, Ed., *The Redbook* (Selected Projects of the Laboratory for Computer Graphics and Spatial Analysis). Cambridge, Mass.: Harvard Univ. Press, 1971.

[6] P. Muehrcke, *Thematic Cartography* (Commission on College Geography Resource Paper no. 19). Washington, D. C.: Ass. Amer. Geographers, 1972.

[7] W. Tobler, "A classification of map projections," *Ann. Ass. Amer. Geographers*, vol. 52, no. 2, pp. 167–175, 1962.

[8] T. Peucker, *Computer Cartography* (Commission on College Geography Resource Paper no. 17). Washington, D. C.: Ass. Amer. Geographers, 1972.

[9] ——, "Computer cartography: A working bibliography," Dep. Geography, Univ. of Toronto, Toronto, Ont., Canada, Discussion Paper no. 12, 1972.

[10] "Data uses in urban planning," in *Census Use Study Reports, Nos. 1–15*. Bureau of the Census, Washington, D. C., Rep. 9.

[11] L. P. Johnston, "Application of computer micrographics to spatial analysis and planning in San Diego," Comprehensive Planning Organ. San Diego County, Calif.

[12] *The DIME Geocoding System*, Bureau of the Census, Washington, D. C., Census Use Study Rep. 4.

[13] *Proc. 1st Int. DIME Colloq.* (Aug. 27–29, 1972), Bureau of the Census, Washington, D. C.

[14] D. M. Austin, Lawrence Berkeley Lab., Univ. of California, Berkeley, private communication.

[15] K. Dueker, "Spatial data systems, I, II, and III," Dep. Geography, Northwestern Univ., Evanston, Ill., Tech. Rep.; also ASTIA Doc. AD 652005/6/7 (Clearinghouse for Federal Documents, Springfield, Va.).

[16] M. G. Notley, "On the design of a graphic display language for interrogating a large data base," U. K. Scientific Centre/Peterlee (IBM, United Kingdom, Ltd.), Rep. UKSC-0004, 1971.

[17] A. Clement, "INTURMAP, the interactive urban mapping system," Dep. Comput. Sci., Univ. of British Columbia, Vancouver, B. C., Canada, Mar. 1973.

[18] E. J. Cristiani *et al.*, "An interactive system for aiding evaluation of local government policies," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, pp. 141–146, Mar. 1973.

[19] P. E. Mantey *et al.*, "Information for problem solving: The development of an interactive geographic information system," in *Proc. IEEE Int. Communications Conf.* (Seattle, Wash.), June 1973.

[20] J. B. Schneider, "Man-computer synergistics: An aid to the design of urban systems," presented at Confer-in 72, Seattle, Wash., 1972, Confer-in Paper no. 51.

[21] J. B. Schneider *et al.*, "Man-computer synergism: A novel approach to the design of multi-objective, multi-modal urban transportation systems," presented at the Int. Conf. Transportation Research, Bruges, Belgium, June 1973.

[22] B. Schumacker, Ed., "UGH—Urban heuristics users' guide," M.I.T. Urban Systems Lab., Cambridge, Mass., July 1972.

[23] D. W. Richardson, "Application examples of UGH," M.I.T. Urban Systems Lab., Cambridge, Mass., May 1973.

[24] D. Roos, *ICES System Design*. Cambridge, Mass.: M.I.T. Press, 1967.

[25] H. Vaccaro, "The use of urban COGO in school bus routing," M.I.T. Urban Systems Lab., Cambridge, Mass., 1972.

[26] R. F. Tomlinson, Ed., *Geographic Information Systems* (Proc. UNESCO/IGU 2nd Symp. Geographic Information Systems, Ottawa, Ont., Canada, Aug. 1972).

[27] C. W. Craven *et al.*, "Regional environmental systems analysis," Oak Ridge Nat. Lab., Oak Ridge, Tenn., Rep. ORNL-NSF-EP-12, Mar. 1973.

[28] A. H. Voelker and C. R. Meyers, "Computer display in spatial modeling," Oak Ridge Nat. Lab., Oak Ridge, Tenn., Rep. ORNL-NSF-EP-25, Aug. 1972.

[29] D. J. DeCoursey, *The Development of a Natural Resource Information System*, vol. 1. Wayland, Mass.: Raytheon Co., Aug. 1973.

[30] B. K. Aldred, "Points in polygon algorithms," U. K. Scientific Centre/Peterlee (IBM, United Kingdom, Ltd.), Rep. UKSC-0025, 1972.

[31] F. F. Abraham *et al.*, "An introduction to air quality diffusion modeling," IBM Scientific Center, Palo Alto, Calif., Rep. 320-3295, June 1972.

[32] L. J. Shieh *et al.*, "Air quality diffusion model: Application to New York City," *IBM J. Res. Develop.*, vol. 16, pp. 162–170, Mar. 1972.

[33] P. Halpern *et al.*, "Air quality modeling for episodic conditions in New York City," in *Proc. Int. Symp. Air Pollution* (Tokyo, Japan), 1972.

[34] D. L. Schreiber, "Computer animation: An aid to understanding the results of large scientific computations," IBM Res. Lab., San Jose, Calif., Rep. RH-986, Mar. 1972.

[35] R. S. Green, *The Storage and Retrieval of Data for Water Quality Control*. Washington, D. C.: U. S. Gov. Printing Office, May 1960.

[36] R. L. Phillips and D. E. Geister, "Interactive graphics in water quality investigations," *Proc. Ass. Comput. Mach.* (Atlanta, Ga.), Aug. 1973.

[37] R. L. Phillips, "Interactive mapping for environmental systems," in *Proc. Amer. Soc. Photogrammetry/Amer. Congr. Surveying and Mapping Conf.* (Orlando, Fla.), Oct. 1973.

[38] W. Tobler, "Linear operators applied to areal data," in *Proc. NATO Advanced Study Inst. on Display and Analysis of Spatial Data* (Nottingham, England), July 1973.

[39] ——, "A continuous transformation useful for districting," presented at the Conf. N. Y. Acad. Sci., New York, N. Y., Nov. 1972.

[40] J. L. Pfaltz, "Representation of geographic surfaces within a computer," in *Proc. NATO Advanced Study Inst. on Display and Analysis of Spatial Data* (Nottingham, England), July 1973.

UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE INGENIERIA
CENTRO DE PROCESAMIENTO DE DATOS

ALGUNAS APLICACIONES DEL COMPUTADOR DIGITAL A LA CARTOGRAFIA TEMATICA

Ingº Bernardo A. Paris
Laboratorio de Computación Gráfica
Universidad Central de Venezuela

Caracas, Septiembre de 1974

I Jornadas de computación técnica

AVICE — Colegio de Ingenieros de Venezuela

AREA: Aplicación de la Ingeniería, Arquitectura y Ramas Afines.

Trabajo N° 12

ALGUNAS APLICACIONES DEL COMPUTADOR DIGITAL A LA CARTOGRAFIA TEMATICA

Ing. Bernardo A. Paris
Laboratorio de Computación Gráfica
Universidad Central de Venezuela

Hotel Caracas Hilton, 25-28 de Septiembre de 1974

ALGUNAS APLICACIONES DEL COMPUTADOR DIGITAL A LA CARTOGRAFIA TEMATICA

Bernardo A. Paris

Laboratorio de Computación Gráfica
U.C.V.

INDICE

I.- RESUMEN.-

Se da una definición informal de superficie estadística y las distintas formas de representarlas de acuerdo al tipo - de información de entrada. Se extiende el concepto introdu ciendo la variable tiempo. Se hace algunos comentarios sobre el "HARDWARE" necesario para desplegar información con significado espacial. Se describen algunos programas, conceptos y procedimientos relacionados con la producción auto mática de mapas temáticos por medio de computadores. Se - discuten los resultados y las dificultades encontradas en la difusión de éstas técnicas. Se comentan los próximos de sarrollos del laboratorio.

II.- INTRODUCCION.-

El presente trabajo constituye un esfuerzo dirigido a presentar ante los profesionales de la computación todo un con junto de resultados experimentales y no definitivos sobre - algunas aplicaciones de la computación gráfica al campo de las ciencias sociales obtenidos en el Laboratorio de Computación Gráfica de la U.C.V. Concretamente se presentan una serie de programas de computador orientados a producir auto máticamente distintos tipos de mapas temáticos y desplie - gues gráficos de variables de tipo social, físicas, etc.

Los programas creados utilizan como medio de salida un Deli neador Digital (PLOTTER) o la impresora de línea del Computa dor, utilizándose uno u otro dispositivo según convenga a - los objetivos del usuario.

A nivel mundial, éstas técnicas han sido lo suficientemente desarrolladas como para permitir la publicación de Atlas - Geográficos totalmente editados por computadores en Suiza e Inglaterra.

El autor expresa su agradecimiento a los organismos de las II Jornadas de Computación Técnica por el esfuerzo que supo ne un evento de esta naturaleza y por haber brindado la o- portunidad de presentar este trabajo ante la comunidad inte resada en las aplicaciones y usos de los computadores.

III.- DESARROLLO.-

1.- DEFINICION DE ALGUNOS CONCEPTOS

1.1. La Superficie Estadística: Si se tiene un - conjunto de valores (digamos población, indi ce de natalidad etc.) repartidos sobre un á- rea geográfica cualquiera, se podría obtener una visualización del fenómeno bajo estudio - asignado a cada punto (definido por sus coor- denadas cartesianas) una altura relativa al - valor de la variable (Fig. 1a y 1b).

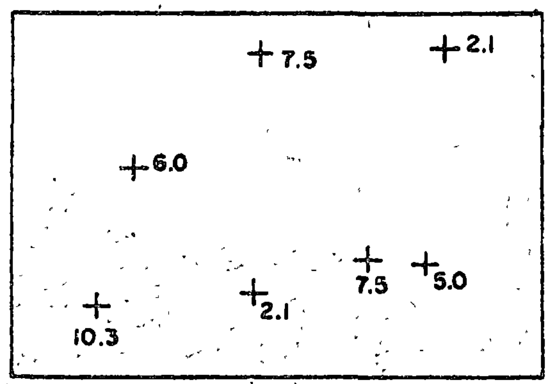


Fig. 1a

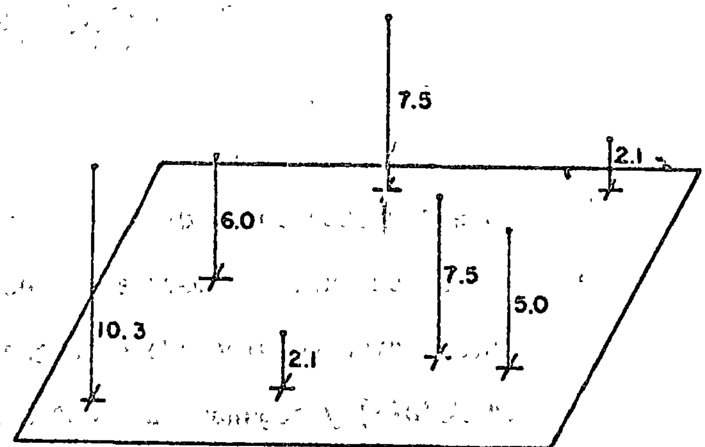


Fig. 1b

Si se hace pasar por cada uno de los puntos una superficie continua se obtiene lo que algunos autores (1) llaman una superficie estadística (Fig. 2). Matemáticamente la superficie puede venir representada por una función - z = f(X, Y) o por arreglo bidimensional de números.

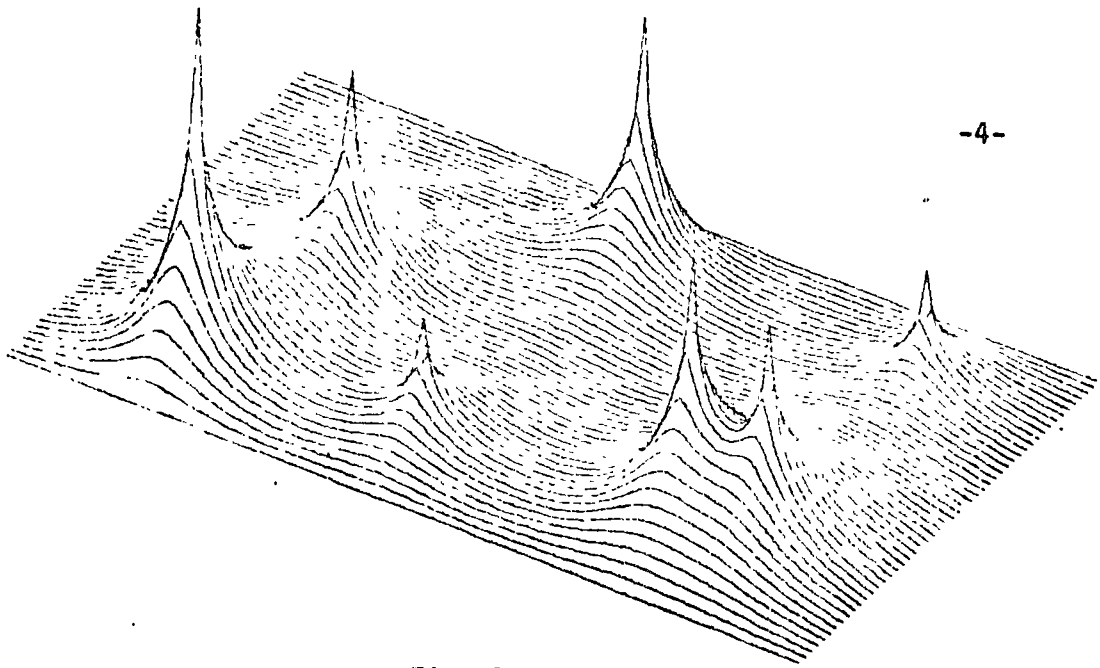


Fig. 2

En el proceso anterior se distinguen dos tipos de puntos. Primero aquellos que tienen información recolectada de alguna forma (DATA POINTS) y segundo los puntos sin información (GRID POINTS). Es precisamente el proceso de interpolar los valores de la segunda clase de puntos a partir de los primeros lo que produce la superficie estadística. La manera en que se realiza la interpolación es referida en el punto 3.3.

La forma especial que se elija para representar la superficie estadística sobre un plano determina el tipo de mapa o vista en perspectiva que se va a producir. Si combinamos lo anterior con el hecho de que la superficie - estadística puede adoptar formas distintas -

en función de que la información recolectada se refiera a puntos geográficos precisos, sub-regiones políticas o arbitrariamente delimitadas (Distritos, Municipios, etc.), se comprende que existan diversos tipos de representación.

1.2. Formas de Representar la Superficie Estadística:

a. Modelos Tridimensionales. A partir de la matriz numérica pueden obtenerse los perfiles presentados por cada columna del arreglo. Esto se consigue imprimiendo un vector alfa-numérico cuya longitud sea proporcional a la magnitud del valor de cada elemento del vector columna de la matriz. Una vez procesadas todas las columnas se obtienen un conjunto de perfiles que ordenados adecuadamente producen el modelo en tres dimensiones. Este método, por lo complicado y lento se utiliza solo en contadísimas ocasiones.

b. Dibujos en Perspectivas. En este caso se trata de obtener una representación bidimensional de un objeto tridimensional. Para ello deben desarrollarse algo rítmos que permitan resolver problemas - planteados por la descripción de los distintos tipos de proyección, rotación de objetos, recta tapada, ventana, etc.

c. Proyecciones Ortogonales. En este caso específico de la producción de mapas temáticos en cartografía.

Si una superficie estadística obtenida - en la forma descrita en 1.1., es cortada por planos equidistantes paralelos a la base se obtienen una serie de curvas de intersección. La proyección sobre el plano base de estas intersecciones son - las ampliamente conocidas "Curvas de nivel" (ver Fig. 4). A nivel mundial se - han desarrollado gran cantidad de algorítmos para generar automáticamente los mapas de contornos correspondientes.

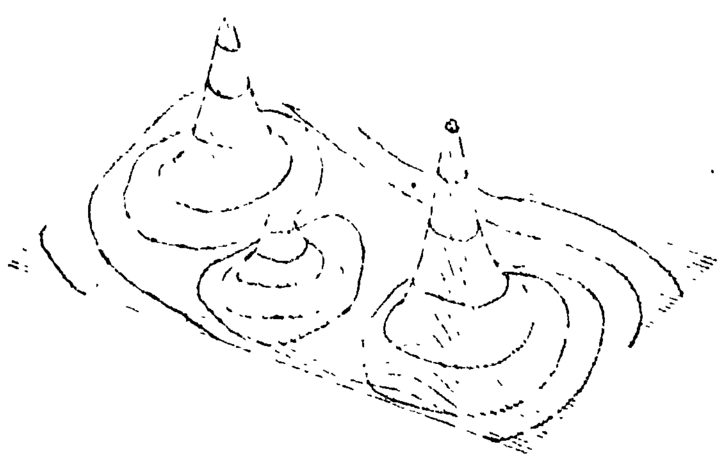


Fig. 4

Si la información recogida sobre el terreno no tiene un carácter puntual sino está repartida sobre un área políticamente delimitada, la superficie estadística adopta físicamente el aspecto de un conjunto de columnas cuyas bases están constituídas por las fronteras de una subregión y la altura es proporcional a la magnitud del fenómeno bajo estudio. Si por ejemplo se tiene un conjunto de Distritos o Municipios, en los cuales se ha contado la densidad de población (hab/Km^2) se podría tener el siguiente diagrama:

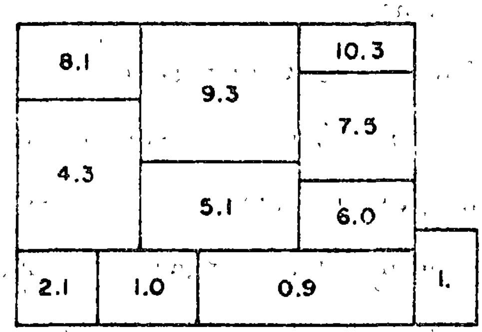


Fig. 5

Realizando el procedimiento anterior de construir columnas proporcionales al fenómeno medido obtendríamos algo parecido a la Fig. 6.

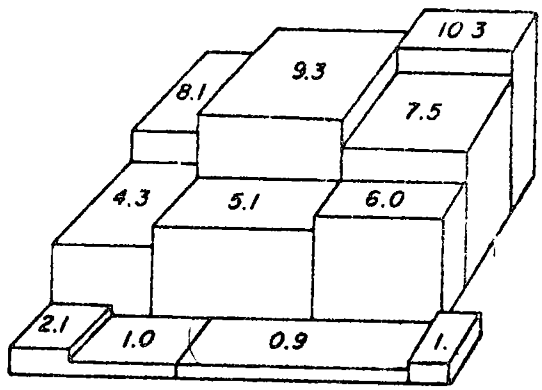


Fig. 6

En este caso la superficie estadística adopta la forma de un histograma tridimensional.

Para visualizar este tipo de representación mediante una proyección ortogonal se adopta el convenio de asignar una gradación de color a cada sub-frontera de la región. El mapa así obtenido se denomina de conformidad o de coropletas.

Una de las formas más inmediatas y elementales que puede adoptar una superficie estadística partiendo de información de tipo puntual consiste en asignar a cada pun to de la región el valor del punto de información más cercano. El algorítmo consiste en analizar cada elemento de la matriz calculando las distancias a cada punto con información. De todos los segmentos obtenidos se elije el menor, y el valor del elemento correspondiente se coloca en el elemento bajo análisis. Este proceso es llevado a cabo en un solo barrido de la matriz, siendo por lo tanto muy rápido. Los mapas así obtenidos se denominan de proximidad (Fig. 7).

Fig. 7

Tal y como se cumple para el caso anterior de mapas de conformidad, para visualizar la superficie estadística en proyección ortogonal se sombrea con una intensidad proporcional a la magnitud del fenó meno estudiado (6).

1.3. <u>Otras Formas de Representación de Información con Contenido Espacial</u>: Otros métodos de establecer correlaciones espaciales de información con la Geografía de una región y que pueden ser automatizados usando un computador son:

a. Mapas de Círculos Tapados. En este caso se dibuja un círculo o esfera cuya área o volumen sea proporcional al valor de la variable; el centro del círculo coincide con la posición del punto de información en el mapa (Fig. 8). Para disminuir la confusión causada por la aglomeración de círculos, se adopta el convenio de que los de menor radio tapan a los de mayor radio.

Fig. 8

b. Mapas de Cubos. El valor de cada punto de información se toma para construir un cubo cuyo volumen es función de la magnitud del fenómeno bajo estudio.

c. Mapas de Puntos Aleatoriamente Distribuidos. Este tipo de técnica nos presenta una aplicación interesante del método de Monte Carlo a la cartografía. Consiste en utilizar un generador de números pseudoaleatorios para obtener las coordenadas X e Y de un conjunto de puntos. El número de puntos a dibujar aleatoriamente sobre un área cuyo centro de gravedad coincide con las coordenadas del punto de información, es proporcional a la variable.

1.4. Imagen Estática y Dinámica de una Superficie Estadística: Hasta este punto hemos visto el aspecto estático de la superficie estadística, es decir que cuando un mapa temático presenta la distribución de una variable a partir de un Censo Nacional, se realiza una "FOTOGRAFIA" de una situación en un punto definido de la escala del tiempo.

Una extensión lógica del hecho de que muchos procesos cartográficos pueden ser automatizados mediante un computador, consiste en introducir el aspecto dinámico del fenómeno bajo consideración. En efecto, es posible generar una secuencia en el tiempo de cualquiera de las representaciones de la superficie estadística tratadas anteriormente. Tal secuencia es susceptible de ser fotografiada cuadro a cuadro por una cámara de cine; la película animada que se obtiene nos presenta la evolución a lo largo del tiempo de la variable (2).

2.- EQUIPOS DE SOPORTE.-

2.1. Computador e Impresora:

El equipo mínimo necesario consiste en un computador y una impresora de 132 caracteres por línea. Mediante un adecuado diseño del SOFTWARE, es posible producir mapas de conformidad, de contorno y de proximidad usando una o más formas o páginas de papel. La técnica consiste en transformar diferentes rangos de valores de la matriz numérica en sobre impresión de caracteres de la impresora. Esto produce una variación tonal proporcional a los valores del arreglo bidimensional. La ventaja de usar una impresora radica principalmente en que el mapa se ejecuta en forma inmediata, sin ningún paso posterior. Las desventajas son:

a. Resolución muy pobre. La relación tamaño del mapa sobre tamaño del carácter - es elevada.

b. El desgaste de la cinta impresora producen mapas de tonalidad variable, disminuyendo así la posibilidad de comparación entre mapas realizados en tiempos diferentes.

c. Desde el punto de vista del programador, el diseño de SOFTWARE es más difícil.

d. No existe posibilidad de utilizar colores.

2.2. El Delineador Digital: La herramienta más útil y poderosa de que se pueda disponer en éste campo lo constituye el Delineador Digital o PLOTTER. Las ventajas son:

a. Gran resolución (de 0.01 a 0.001 pulgadas). Este produce mapas de gran calidad.

b. Posibilidad de utilizar diferentes colores.

c. Diseño de SOFTWARE con grandes posibilidades y programación relativamente fácil.

La única desventaja del PLOTTER es que generalmente esta conectado fuera de línea. Por lo tanto, existe un pequeño retraso producido por el tiempo que se toma un operador en generar el mapa.

Debe mencionarse la existencia del PLOTTER CRT. Presenta la ventaja de producir mapas en forma instantánea, pudiendo programarse para generar secuencias en el tiempo. La desventaja consiste en el tamaño reducido de la pantalla y el tener que fotografiar las imágenes para tener registros permanentes.

2.3. Equipos Auxiliares. En primer lugar citaremos la mesa digitizadora, necesaria para obtener en forma digital las coordenas de los puntos que constituyen los contornos geográficos de las regiones.

Si se van a producir películas animadas es necesario una cámara de cine con capacidad de tomar secuencias cuadro por cuadro.

2.4. Equipo Disponible en el Laboratorio de Computación Gráfica de la U.C.V.:

a. Computador Burroughs 5500. Posee 5 modulos de memoria, discos de cabeza fija.

b. Impresora rápida de 132 caracteres por línea. Capacidad de sobre-impresión.

c. Delineador Digital CALCOMP, modelo 536 de resolución 0.01 pulgadas y 300 incrementos por segundo.

d. Cámara de cine CANON 814.

Debe notarse la falta de una mesa digitizadora. Esto solo ha traído como consecuencia, que las coordenadas del contorno de los mapas y ciudades se recojan usando láminas de papel milimetrado. El proceso es lento y propenso a errores, pero el alto costo del equipo y la poca frecuencia de uso, no justifican en los actuales momentos su adquisición (7).

3.- PROGRAMAS DESARROLLADOS POR EL LABORATORIO DE COMPUTACION GRAFICA DE LA U.C.V.-

3.1. Los programas según el Tipo de Información:

Siguiendo la misma secuencia presentada en los puntos 1.2, 1.3 y 1.4 anteriores, se han desarrollado los grupos diferentes de programas en función de que la información haya sido recogida en forma puntual o se refiera a variables distribuidas en límites políticos. Para el primer caso se ha establecido el uso de una malla de 50 por 120 elementos (Fig. 9). Esto tiene su justificación en el hecho de que se aproxima bastante al tamaño de la forma continua usada por la Computador -- B-5500, a saber, 57 líneas con 132 caracteres. Dado que el ancho de un carácter es menor que la distancia entre líneas, los elementos de la malla utilizada son rectangulares. Esto permite conservar sin deformaciones cualquier contorno geográfico que se intente dibujar mediante la impresora (ver Fig. 19).

Para el segundo tipo de programas, basta dar la magnitud de la variable por Estado, Distrito o Municipio.

3.2. Modelos Tridimensionales de la Superficie Estadística: El programa TRIDIME/ESTADIS produce a partir de una matriz el 50 por 120 elementos, el perfil relativo de cada una de las 120 columnas en la impresora. Mediante un procedimiento manual bastante lento es posible obtener un modelo en tres dimensiones.

Esta técnica tiene aplicaciones en la representación de la topografía de regiones.

3.3. Dibujos en Perspectivas: El programa TRIDIME/PERSPEC, genera mediante el Delineador Digital un dibujo en proyección axonométrica de la superficie estadística. La característica principal radica en la eliminación de todos aquellos segmentos de recta que están ocultos al observador (hidden line problem). Esto proporciona una buena percepción del fenómeno estudiado. En la Fig. 10, puede observarse un intento de representar la población del territorio venezolano mediante una superficie estadística. El observador está colocado en la parte sur-este de Venezuela. Debido a la gran magnitud de la población de Caracas, la población del resto de las ciudades aparece grandemente disminuidas.

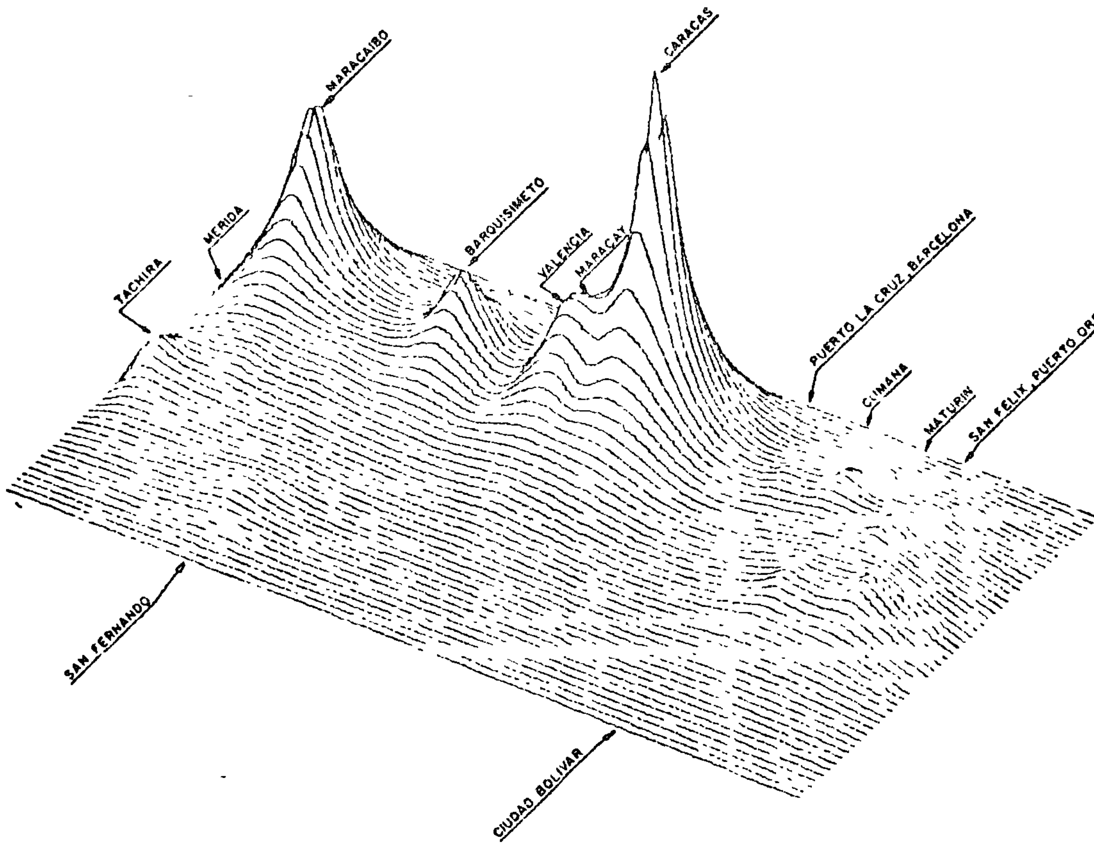


Fig. 10

En el apartado 1.1. tratado al principio se hizo referencia que a partir de los puntos - de información se obtenían por interpolación los valores del resto de los puntos. Existen varios algorítmos para obtener esto. El usado en el laboratorio consiste en asignar a - cada punto de la matriz el promedio de los - cuatro u ocho puntos circundantes.

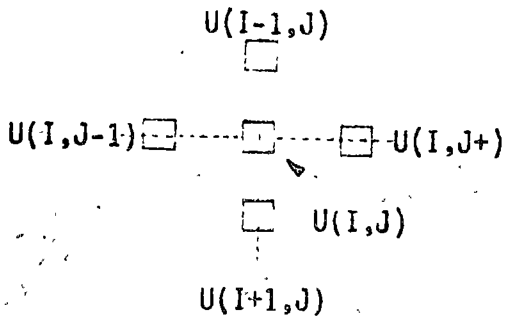


Fig. 11

Dentro de la literatura especializada este - proceso recibe el nombre de "Algorítmo del - Promedio Móvil".

Si trabajamos con una matriz de tamaño M por N los límites de las variables I de fila y J de columna serán 2, M-1 y 2, N-1 respectivamente. Efectuando este "barrido" un número suficiente de veces (excluyendo los puntos - con información) se genera la superficie estadística (o de potencial). A fin de "suavizar" un poco el aspecto de la superficie o - sea, eliminar el "Ruido de Fondo" de los datos de entrada, se vuelve a someter el arreglo al algorítmo del promedio móvil, pero esta vez se incluyen los puntos de información. Si se compara la Fig. 2 de la página 4 - con la Fig. 12 puede notarse la diferencia - obtenida.

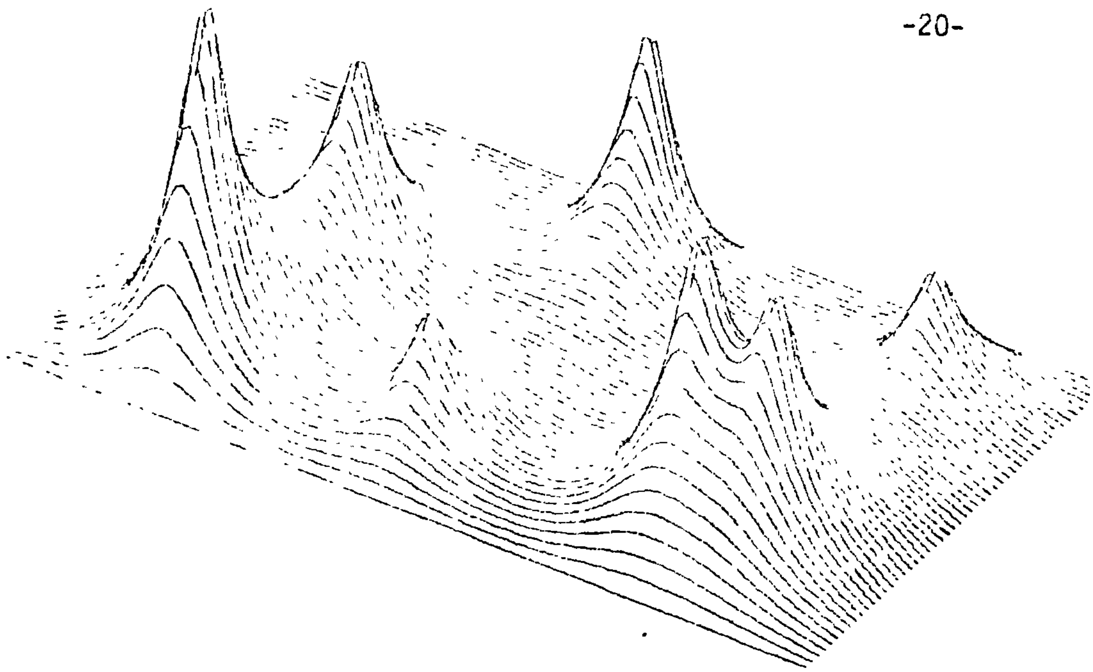


Fig. 12

3.4. Mapas de Contorno: El programa IMPRE/CONTOR genera mediante la impresora de línea mapas - de contorno usando sobre impresión de caracteres (Fig. 13).

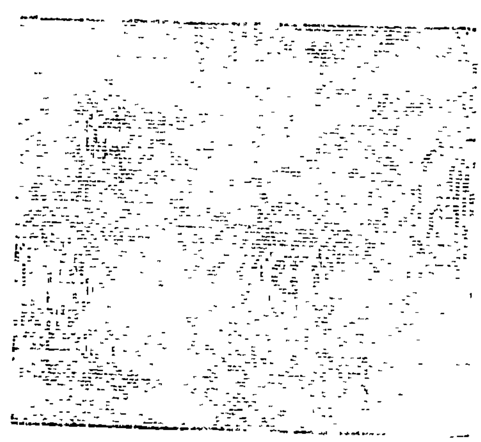


Fig. 13

Cada mapa viene impreso sobre cuatro formas de papel diferente que luego son unidas para formar el mapa definitivo. El "PROMEDIO MOVIL" es el algorítmo usado para interpolar los puntos. El laboratorio no ha desarrollado programas de este tipo con salida por Delineador Digital.

3.5. Mapas de Proximidad: IMPRE/CONTOR genera mapas de proximidad del Territorio Venezolano usando, como en el caso anterior, la impresora de línea (Fig. 14).

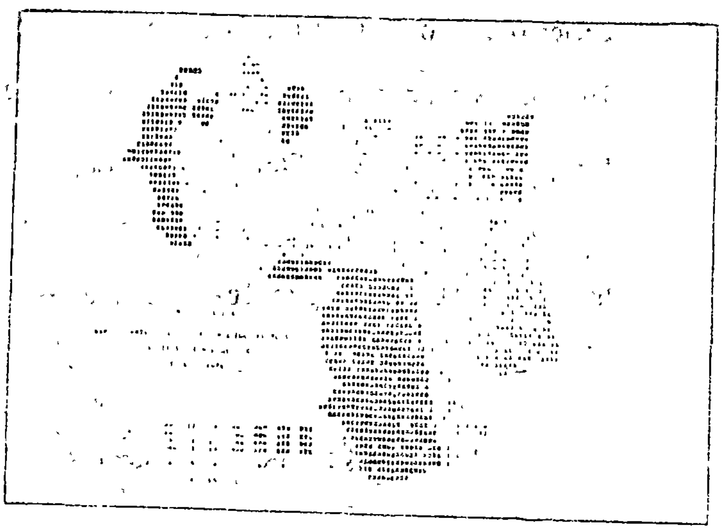


Fig. 14

Uno de los programas a resolver en este caso es como sobre-imponer un contorno geográfico en una matriz de manera tal, que los elementos que no pertenecen al interior queden excluídos de procesamiento e impresión. La recolección de información es efectiva mediante la malla de la pág. 27 (3).

3.6. Mapa de Círculos Tapados: El problema principal consiste en dibujar N círculos a partir de N puntos de información. El radio de cada círculo es proporcional al valor del punto. A fin de evitar la aglomeración de símbolos, los círculos más pequeños tapan a los más grandes en caso de solape (Fig. 15). El programa VENE/CIRCUM genera mapas a tráves del Delineador Digital del Territorio Venezolano para 617 puertos y ciudades.

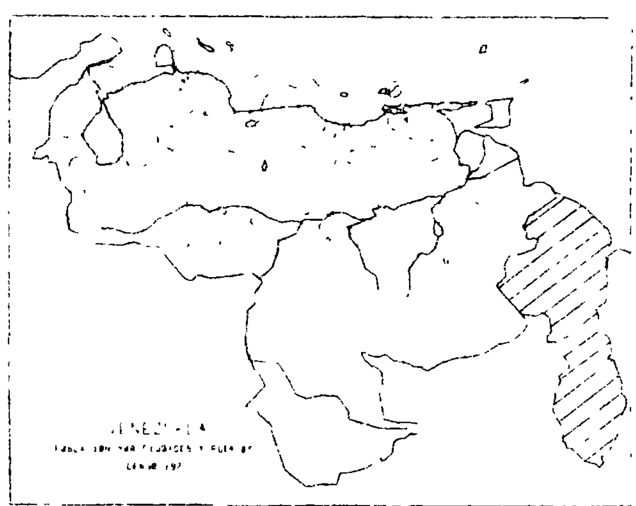


Fig. 15

3.7. Mapas de Cubos: El programa VENE/CUBO genera mapas políticos a nivel de estado. En el centro de cada uno de ellos se dibuja un cubo cuyo volumen es proporcional a la variable. La salida es por "PLOTTER" (Fig. 16).

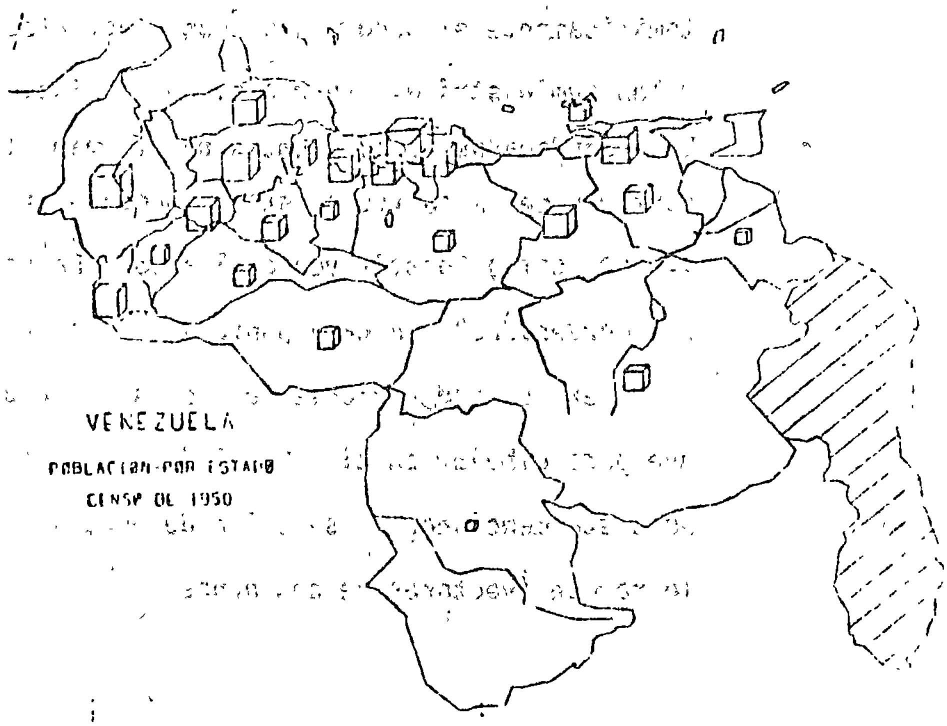


Fig. 16

3.8. Mapas de Conformidad: Los programas VENE/CONFOR y CENTRO/OCCIDEN producen a través del "PLOTTER" mapas de conformidad del territorio venezolano a nivel de Estado y de la zona Centro-Occidental de desarrollo por Distritos.

(Figuras 17 y 18). La información suministrada no tiene carácter puntual, estando ésta relacionada con límites políticos. Diez clases o intervalos categorizan a los datos de entrada. Los límites entre clases los calcula automáticamente el programa si son constantes, o los suministra el usuario si son variables. Las coordenadas de los puntos que forman el contorno del mapa (compuesto de pequeños trazos de recta) constituyen el llamado "Banco de Coordenadas". A cada punto se le asigna un número en orden creciente, de tal manera, que para dibujar en el "PLOTTER" cada contorno basta mencionar la sucesión de puntos que lo forman (vectores de contornos).

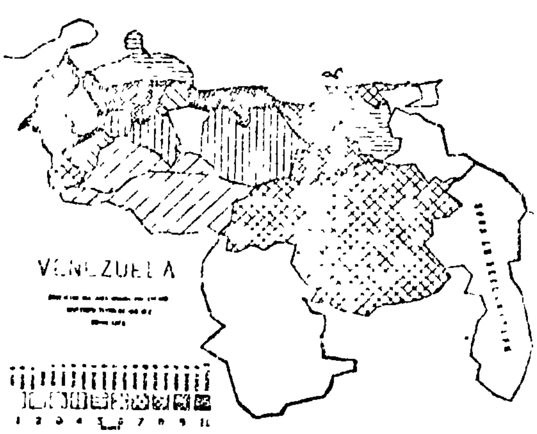


Fig. 17

Fig. 18

3.9. Películas Animadas Generadas por Computador:

El Laboratorio ha realizado algunos experimentos tendientes a producir películas animadas, en las cuales cada cuadro es generado por la impresora de línea o por el "PLOTTER". La película producida muestra varias secuencias de objetos rotando y alejándose, el proceso de someter una matriz al algoritmo del promedio móvil, etc. Todo el proceso de generación de la película es largo y tedioso pero produce resultados muy interesantes al introducir la variable tiempo, lográndose así captar la imagen dinámica de los fenómenos bajo estudio.

4.- COMENTARIOS

El principal obstáculo encontrado por el personal del laboratorio en la difusión de estas técnicas reside en la aparente falta de interés mostrada por parte de los usuarios potenciales. Parte de la actividad del personal se consume en tratar de divulgar los logros obtenidos. Por esta razón el grupo dicta con frecuencia cursos de extensión, charlas, seminarios, etc.

5.- FUTUROS DESARROLLOS.-

El Laboratorio proyecta perfeccionar los programas generadores de los mapas temáticos descritos anteriormente y crear "SOFTWARE" para obtener:

a. Mapas de contorno con salida por "PLOTTER".

b. Mapas de conformidad con salida por "PRINTER".

c. Mapas de puntos usando métodos de Monte Carlo.

d. Programas generales que permitan al usuario definir cualquier contorno geográfico y obtener por "PLOTTER" o "PRINTER", mapas de conformidad, de proximidad o de contornos.

e. Crear los algorítmos necesarios para generar películas animadas por computador que desplieguen el aspecto dinámico de superficies estadísticas a través de los distintos tipos de representación descritos anteriormente.

f. Aplicación de éstas técnicas al despliegue de información proveniente de Banco de Datos Geográficos (5).

LABORATORIO DE COMPUTACION GRAFICA .

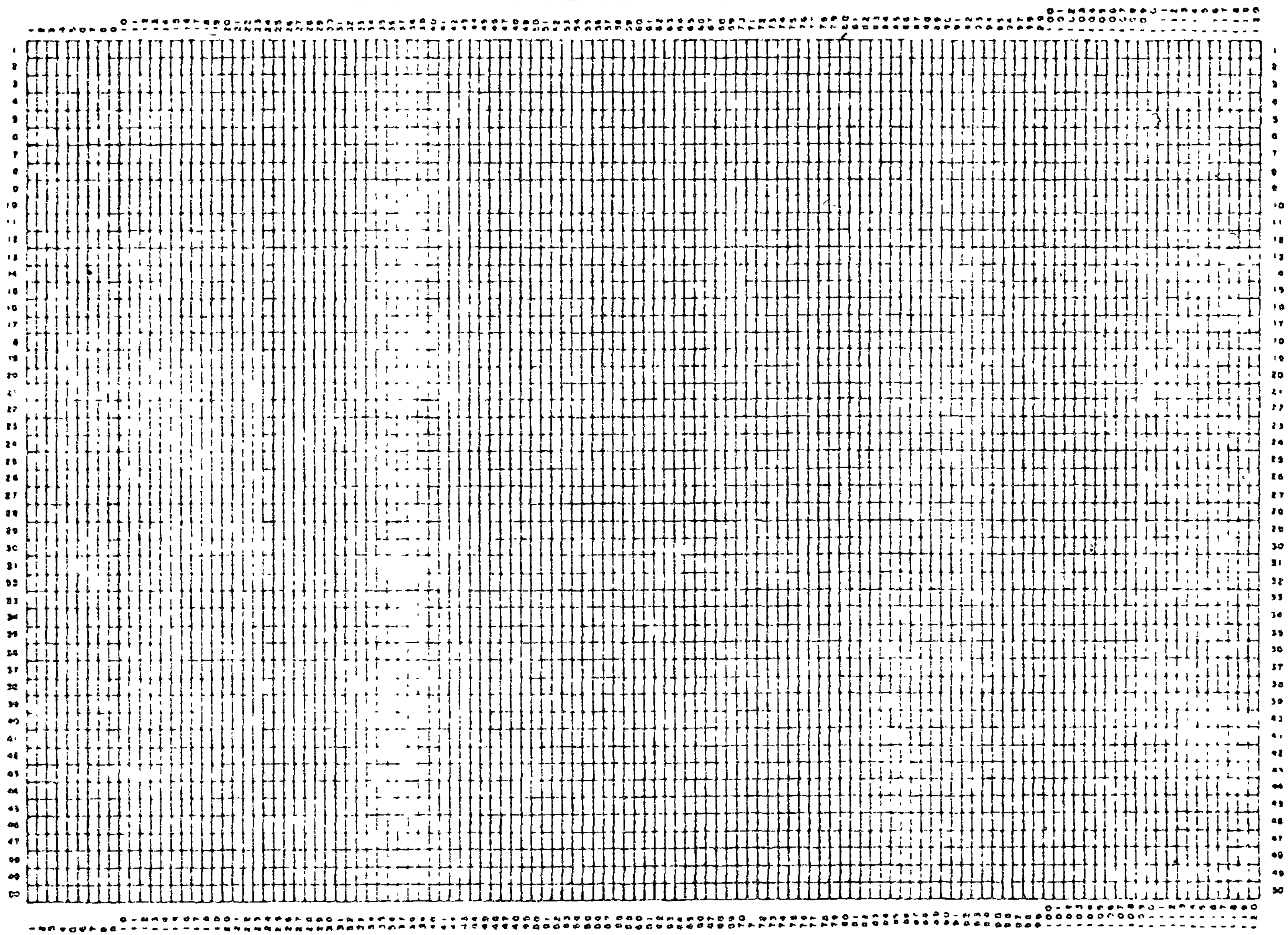


Fig. 19

R E F E R E N C I A S

1.- RCBINSON A. "Elements of Cartography". (1.969). Pág. 150.

2.- WASHINGTON W. M. "The Application of CRT Contour Analysis to General Circulation Experiments". Bulletin of the American Metereological Society. Vol. 49, No. 9 (Sep. 1.968).

3.- PEUKER T. "Computer Cartography". Resource Paper No. 17. Association of the American Geographers. (1.972).

4.- MUEHRCKE P. "Thematic Cartography". Resource Paper No. 19. Association of the American Geographers. (1.972).

5.- MEYER M. "The Statistical Mapping and Geografic Base Files of the V.S. Bureau of the Census". U.S. Dept of Commerce.- Washington D.C. (1.970).

6.- "The Red Book". Laboratory for Computer Graphics and Spa-- tial Analysis. Harvard University, Cambridge. Mass. (1.973).

7.- "Sistema de Delineación Digital". Manual del Usuario. Sección de Sistemas, Centro de Computación. Facultad de Ingeniería. U.C.V. (1.972).

LCG+SA- REPRINTS

CARTOGRAPHIC DATA STRUCTURES

by Thomas K. Peucker and Nicholas Chrisman

Laboratory for Computer Graphics
and Spatial Analysis
520 Gund Hall
Harvard University
Cambridge, Massachusetts 02138

This paper is subject to revisions and is or will be submitted for publication. Any quotation, therefore, needs the approval of the authors.

CARTOGRAPHIC DATA STRUCTURES*

BY Thomas K. Peucker and Nicholas Chrisman

INTRODUCTION

This paper reports on a series of on-going research projects which are concerned with efficient and flexible data structures for geographic and cartographic analysis. The main objective of the research can be summarized in three points:

In most cartographic data banks, the arrangement of the data is guided by the input stage. In other words, little manipulation of the data is usually performed after the data has been input into the system from maps. Secondly, cartographers and computer scientists have made few attempts to combine different types of cartographic information, for example - height with other cartographic features. Therefore, the different types of cartographic entities are stored in different files and it is usually extremely time-consuming to achieve any combination of them. The third point leads into the general problem of cartographic/geographic data banks, which is that the data structure is usually very simple and lacks especially one facet which we consider very essential for most geographic and cartographic analyses. This is an indication of the relative location of a geographic entity, i.e. the position of a geographic entity with respect to its neighboring entities.

This paper therefore will characterize types of existing geographic and cartographic data systems especially with respect to the above three points which we may abbreviate with the terms flexibility, comparability and topology. It then will describe attempts which have been made by the authors to produce data systems which are superior to the existing ones. Throughout the paper the term "neighborhood function" will play a major role and it therefore will be explained in more detail in the following section.

NEIGHBORHOOD FUNCTION

When we are asked for the location of a city, we will give it with respect to a river, a seacoast, a pass, a neighboring larger city, etc.; rarely will we use geographic coordinates as longitude or latitude or map coordinates. In fact, we are taught in elementary geography that the geographic coordinates will tell us little about either the large scale (site) or small scale (situation) characteristics of a place. Similarly, if I describe my position on a piece of terrain, I will not use my map to determine my location within the UTM-grid, rather I will look for nearby relief features (peaks, rivers, slopes, roads, etc.) as orientation characteristics.

The research for this paper has been supported by ONR contract N00014-73-C-0109, and the Laboratory for Computer Graphics and Spatial Analysis, Harvard University.

In contrast, when we create a geographic data bank of any kind, we resort to nothing else than some kind of absolute coordinate system. To carry the idiosyncracy even further, the geographic evaluation and/or mapping system used with the data very often does not allow us to include cartographic features which would give us an indication of relative orientation as streets, rivers, etc. (The programs of the Laboratory for Computer Graphics and Spatial Analysis are some of the laudable exceptions.)

While the human can help himself in his orientation on a map through overlays or map comparison, the computer is in much more trouble to determine relative location. If the relative location of each of, say, 5000 points is to be calculated, the program literally has to compute every point's distance to every other point to find the closest few. To make it worse, there are programs available that do this computation several times within one program run.

As was mentioned before, some indication of the relative location of a geographic feature, also called its neighborhood function, is very important. This neighborhood relationship can be expressed in several ways, as an implicit or explicit function or as a discrete function in the form of a table. The explicit functions can be polynomical or trigonometric equations set for a discrete grid of surface patches which give the form of the surface at each point within the patch. Typical for this approach is the work of Junkins a.o., (1973). Also, two-dimension spline functions fall in this category (Holroyd and Bhattacharyya, 1970). A much more frequent way of defining a neighborhood is by the explicit function in the form of a sort routine which finds the n closest neighbors. This is done for a series of interpolation algorithims to produce a regular grid of points (Shepard, 1968; Heiskanen and Moritz, 1967). The computations increase close to the square of the increase of the number of points, since the search has to be repeated for every point and has to go every time through all points or at least a large number of points.

This search procedure also applies in the case of planar surfaces where neighboring polygons have to be found, a problem which occurs in the case of contiguity constraints in problems of factorial ecology and other regional correlations. Here the problem becomes to sort through all the polygons to find points which a pair of polygons have in common. Again, the problem increases in complexity close to the square of the number of units.

The implicity neighborhood relationships are usually a function which describes the decoding structure of the geographic entities. One very good case is cesribed in Rosenfeld (1969) for different types of neighborhood relationships within a regular grid. The point P_{ij} has the four neighbors (i+1,j), (j,i+1), (i-1,j), (i,j-1) and the eight neighbors (i+1,j), (i+1,j+1), (i,j+1), (i-1,j+1), (i-1,j), (i-1,j-1), (i,j-1), (i-1,j+1).

Neighborhood relationships in the form of tables are very rarely used. This type records the neighborhood function by pointers to neighboring geographic entities. For example, a structure which is built on the basis of Thiessen polygons could have such a structure by simply having the labels of the neighboring points in the record of a point. The most famous of this type of data structure is the DIME file of the U.S. Census which encodes line segments and adds to it the names of the polygons to the left and right of the line segment, and the names of the two nodes at either end. The neighborhood relationship used in the DIME development are derived from a discipline unexplored by most geographers, which is topology.

Neighborhood relationships will be discussed in greater detail during the analysis of different existing data structures. Two types of geographic data bases will be discussed: those defining three-dimensional surfaces and those defining planar surfaces. For both types, a summary of historical development will be presented and it will be shown that although presently at different stages of the development, these two types can be treated as special cases of one topological data structure.

DATA STRUCTURES FOR PLANAR SURFACES

The types of geographic entities on planar surfaces are points, lines and area enclosing lines or closed polygons. The latter might be the most frequently encoded feature in geographic data systems.

The simplest data base system for planar surfaces is to encode entity by entity with little or no regard for entity overlaps or adjacencies. In other words, every polygon in a polygon system is encoded and stored without any regard for contiguous polygons, lines are encoded without regard for the fact that they may intersect or merge with other lines, etc. The results of such an encoding are sliver lines, very confusing and unesthetical graphic results, and especially the impossibility of doing anything with the data base short of an extremely coarse graphic image. (Fig. 1).

The solution to this is to either edit the data or make the encoding or graphic output scale so coarse that the scale steps are within the area of variance of the encoder. The first alternative has been attempted in several cases, the most famous being the MAP-MODEL system (Arms, 1970). The editing is guided by the assumption that every segment has to be represented twice - except for segments on the outer boundary of the set. The editing program therefore sorts through the segments for every single segment and tags those for which is has not found a compliment, so that the user can intervene. Adjacent segments are first defined by the middle point of the segment and then - for a higher accuracy test - by the two end points.

To overcome some of the limitations of independent location structures, systems were developed based on a common location dictionary. This dictionary contains the coordinates of every point on the map. Polygon boundary lists then consist of the labels (location numbers) of the boundary points. Line and point information is handled in the same way. (Fig. 2). A number of programs are based on this structure (CALFORM, Laboratory of Computer Graphics and Spatial Analysis), others have subroutines to convert this type of structure to the simple line list described above (e.g. PDATA for SYMAP, Laboratory for Computer Graphics and Spatial Analysis). Programs have been developed to simplify data-input through automated polygon-identification (Douglas, 1973 in Peucker (ed.) 1973).

The point dictionary data base has the advantage that sliver lines do not occur. However, the problem of neighborhood relationships is not approached any better. Granted, the search for common lines is not anymore by the coordinates of points but by their labels, but this only brings us closer to a solution by a little less computer time. It also creates difficulties. A point directory can and will be accessed in a completely arbitrary order, since there are no restrictions regarding point placement. The standard response to this problem is to make the dictionary core resident which unfortunately will limit the complexity of the map that can be handled in this manner. The shortcoming (of shared data) is augmented by the continued independence of the entities created by the dictionary. Instead of _n_ points with their x and y coordinates, there are still _n_ references to points.

Some of the objections can be eliminated by formulating an intermediate object between the entity and the points used as an addressing scheme (Nake and Peucker, 1972, Peucker (ed.) 1973). A geographic entity can be created from a list of "line segments" which are, in turn, created from references to the point dictionary. This system allows for easy definition of the entities with a minimum of pointers but each entity is still independent in the sense that the entity's neighbors are not known. The direction of access is still from entity to location but not the reverse.

All of the data structures described thus far are of limited flexibility and utility because of the restriction of not knowing neighborhood relationships. In other words, although the absolute location of each geographic entity is known, its relative location, i.e. with respect to neighboring geographic entities, is not known. By adding the topological function of each element to a data structure, large improvements in flexibility and scope of applications can be realized.

If one is concerned about the memory capacity which is needed for the storage of explicit neighborhood relationships, one might consider a system with implicit neighborhood functions by modifying the form of the entities in the encoding stage. Many geographic information systems are in existence which store land-use data in grids of rectangular cells (Hsu, 1973). However, a serious problem is encountered with a regular discrete encoding of planar surfaces; since according to the sampling theorem (Tobler, 1969), the sampling interval has to be half of the size of the smallest features to be encoded. Either the size of the cell-unit has to be very small to be able to encode, for example, urban land-use, or with a grosser cell size, only very slowly changing land-use can be stored. In the first, a very large volume of redundant information would be created in a low-density land-use area and matrix reduction techniques such as run-length encoding (Amidan and Akin, 1970) would only create physical rather than logical compaction. In the second, highly varying land-use features had to be aggregated to a degree which makes the usefulness of the whole system doubtful.

Beyond sampling problems, a grid structure imposes a bias towards specific orientations of features, since diagonals, although physically longer are given the implicit cell relationships of unit distances along the major axes. The grid structure also creates the illusion of working with a discrete point space, rather than a regular areal partitioning.

Recent studies (Switzer, 1974) have developed parameters which determine the degree of inaccuracy of a given sampling mesh. In cases where some types of geographic features vary over small areas and others over very large areas - the typical case for state-wide land-use mapping - the error is averaged over all features and will effect features with a high-frequency variation more than others.

The variation of the mesh size within the information system would be only little help here. The data management and the manipulation and display routines would become more complex, although most likely less than many researches seem to imagine. Exact figures about the difference cannot be given since no literature about such a system is known.

EXPLICIT TOPOLOGICAL RELATIONSHIPS

One of the first known attempts to incorporate explicit topological structure into a geographic data base is the Dual Independent Map Encoding (DIME) system of the U.S. Bureau of Census. The DIME files were originally developed as an automated topological error detection system for the Address Coding effort of the 1970 census. (Fig. 3).

The basic element of the DIME file is a line segment defined by two end points. It is assumed that the segment is straight and not crossed by any line of interest. The metropolitan files usually define this unit as a street block face. Complex lines are represented by a series of segments approximating the line. The segment has two "node" identifiers, along with the coordinates of its two end points and codes for the polygon on each side of the segment.

While DIME topology makes a great deal of information accessible to urban researchers, functional relationships are not made explicit. Segments sharing a node for example must be found by laborious search procedures. The same is true in order to assemble the outline of a polygon. More importantly the DIME structure is cumbersome to use for many cartographic applications involving areas made up of complex lines. As procedures for a "one-shot" effort checking as is the case of Address-Coding and Address-Matching in metropolitan areas it is quite adequate. However, for efficient computer storage and retrieval and for many other applications, improvements must be made.

For example, the reliance on the individual line segment makes the reduction of details for display purposes difficult since line segments cannot be simply deleted without correcting the reference codes for the affected nodes.

At the Laboratory for Computer Graphics and Spatial Analysis, the junior author has developed a data structure designed to contain all the information needed to construct any of the previously enumerated planar structures. The basic object of the structure called POLYVRT is the "chain." Like a DIME segment, a chain has nodes at its two ends, separates two areal objects and is assumed to be uncrossed. It differs in that the POLYVRT chain may be made up of many points where the DIME unit has only two points per line segment. A boundary between two polygons can be referenced by a single chain no matter how complicated, because line detail is topologically unimportant.

The coding of a complicated boundary as a unit is not unique to POLYVRT, as was mentioned before. The project, the "Interactive Map in Urban Research" (Nake and Peucker, 1972; Peucker (ed.) 1973) as well as the World Data Bank I (Schmidt, W., 1969) are composed of "lines", in the latter case some of them contain over 4000 points. However, the chain based system of POLYVRT is a new structure because of the topological role assigned to the chain and the subsequent construction of a list data structure. Based upon this assignment, the topological information about a chain resembles the information on a DIME record however, the distinction between nodes (i.e. points used for more than one chain) and the points internal to a chain allows internal points to be eliminated without influencing the neighborhood relationships. The main innovation in developing a chain representation is that areas of significant line detail may be efficiently handled. Topological checking is reduced from dependency on the number of points to the dependency on the number of boundaries.

In addition to the indication of the relative location of the chain with respect to its neighboring polygons, information is stored in separate lists assembling the bounding chains for every polygon. Thus, searches can take place in two directions, from the chain to the polygon and from the polygon to the chain. This is very important for any type of neighborhood manipulation, since neighboring entities can be found through their "bounding" or "bounded" complements. In other words in order to follow along a group of chains one flips through chain to polygon to the next chain, etc., whereas to traverse a series of polygons one tests for adjacent polygons by going through the chain directory for each polygon.

The POLYVRT program places point information in secondary storage. The three higher level objects (chains, nodes and polygons) are core resident. Only the chain refers directly to the point file in the secondary storage. In addition to a pointer to the location of the points on the point file. The chain record incorporates the name of the chain, the labels of the starting and ending node, the left and right polygons. Conversely, the polygon list consists of the bounding chains in proper sequence. (Figs. 4 and 5).

DATA STRUCTURES FOR 3-DIMENSIONAL SURFACES

Boehm (1967) describes in a very detailed analysis the advantages and disadvantages of different types of encoding of surfaces. He comes to the conclusion that the encoding of surfaces by contours minimizes the storage capacity necessary for a surface whereas a regular grid of surface points minimizes the computing time necessary for several types of manipulations. The types of data banks which Boehm studied were different contour encodings as well as a regular grid structure and a grid structure where the mesh could be increased at places.

Unfortunately, Boehm did not reflect on the reasons why contoured data minimized storage capacity whereas a regular grid minimized computing time. If he had done so it is quite feasible that the development of geographic data bases would have taken different routes. However, it is surprising that this topic has produced little discussion despite the fact that extremely large data banks of terrain (digital terrain models) have been developed.

When encoding surfaces one has to try to adapt the density of points to the variation of the local terrain. The question of how dense the points have to be can be answered by using the philosophy of the sampling theorem, which says that the sampling mesh has to be half of the smallest variation to be detected. For the terrain of a typical map, the smallest variation can vary considerable which results in the need of frequent adjustment of the sampling interval.

In a contour map the density of contour lines changes with the density of relief variation. It therefore fulfills the requirements set by the sampling theorem for a "non-stationary surface", i.e. a surface with changing terrain. For the regular grid on the other hand if the smallest object one wishes to detect anywhere within a study area is of size ("wave lengths") S, then the grid spacing everywhere must be S/2 or less. The regular grid therefore has a high tendency towards redundancy since smoothness of areas of the study area will contain far more points than are needed to accurately portray their form. To improve the "resolution" of a grid by a factor f, the grid spacing must be decreased by this factor and the total number of points is increased by a factor of f^2.

The question has been raised many times in photogrammetry whether horizontal contours are the best representatives of topographic surfaces. It has been stated (e.g. Brandstatter, 1957) that contours do not detect many types of the breaks which are frequent on terrain. Therefore the encoding of surfaces by vertical profiles has been attempted several times in photogrammetry in recent years. Points are encoded only when the slope of the surface changes. In other cases (Sima, 1972) the break lines are encoded in addition to a regular grid. Again in other approaches (Grist, 1972) only the break lines or even only the ridges and channel lines of a surface are encoded. The detail here of course depends very much on the scale used.

When performing numerical computations on the basis of these digital terrain models, the quantity of data involved will be only one determining factor for the amount of programming and computations needed. For most of the numerical computations on surfaces one needs some type of neighborhood function either to compute some surface behavior, like slope, local variation of relief, etc., or simply to find the next unit for the drawing of a contour, vertical profile, etc. For a set of contours it is relatively easy to create a directory which indicates a sequence of contours in a type of tree where the surrounding contour is the base and from there the other contours go off in branches (Morse, 1968). However, if one is to find the neighboring points of a point on adjacent contours, one has to sort through all the points of those two contour lines and compute the distances between all contour points and the point in question to find the shortest distances. This procedure can be quite time-consuming if the contour line is digitized very densely. A regular grid on the other hand has its neighborhood function implicit as has been shown above. Therefore the finding of a neighbor does not involve any search and thus no extra computer time. For a set of irregularly distributed points as they are represented in very simple data structures (SYMAP, Laboratory for Computer Graphics and Spatial Analysis) the creation of a neighborhood function is usually done by finding the closest small number of points where the number varies around 6.

It has been noted that although plane surfaces and 3-dimensional surfaces look rather different it only takes a few assumptions to treat one as the subset of the other. Whereas 3-dimensional surfaces always are, planar surfaces are not always based on interval or ratio data but are often used for ordinal and nominal data. However, as has been shown (Nordbeck and Rystedt, 1972, Rosenfeld, 1969) one can treat ordinal and nominal data as interval data. But even without this conversion we can combine the two types into one simply using different assumptions about neighborhood.

Given a set of n "data points", i.e. points for which x, y, z-coordinates are known, continuously defined surfaces can be created using three different groups of assumptions about the "surface behavior" (Peucker, 1972).

The first group of assumptions is that of a discontinuous surface which says that the surface retains the value of a data point within the neighborhood of that data point, where neighborhood is defined either by a given polygon (the choropleth approach) or by the fact that the area is closer to one data point than to any other one in the neighborhood (the proximal approach). Secondly the assumption can be that each data point represents a sample of a single value of a continuous surface. Neighborhood is then a number of closest neighboring data points and intervening values are interpolated with different types of interpolation procedures. The third assumption is that the data point is a sample of a surface which can contain errors; thus the data point is not necessarily located on the surface, but close to it. This approach needs a further assumption which usually is that the surface is smoother than a surface would be which was constructed through all the data points.

With these assumptions about surface behavior one can combine planar surfaces and 3-dimensional surfaces and treat them as one as has been done in some programs, the most noticeable being SYMAP (Laboratory for Computer Graphics and Spatial Analysis). Whether one agrees with the continuous treatment of discontinuous surfaces or not is irrelevant at this point. It is important, however, that one keep in mind that many cartographic applications have to treat both types together and therefore data structures have to be developed which can handle both types of surfaces at one time. For nonbelievers such a data structure can only serve as a basis for display procedures, for the believers it opens up vast areas of applications for the display as well as the analysis of geographic data.

THE PROPOSED DATA STRUCTURE

To implement the ideas presented here, the senior author is developing a geographic information system for three-dimensional surfaces, and the two authors have developed the concept for a system for planar regions. Both systems are based on data structures with explicit topological neighborhood relationships.

The basic philosophy of both approaches is to separate the data base from the application programs. In the early days of computer cartography, data were intermittently tied to the application program that used it. Now we see that the available data became so voluminous and the application programs so varied that extra efforts in their preparation for more efficient computations seem to be justified. We are well in step with modern computer science to separate the data base from the application programs with the data structure becoming the link between the two.

The creation of a structured data base is nothing new. Any interpolation to a regular grid builds one, any double loop which compares the segments of a polygon with the segments of all other polygons in a set of statistical regions replaces one, often at a very high cost. Our effort is to spend computing time before any application has been performed in the anticipation of heavy uses of the data base. We claim that an efficient structure of the data base can not only speed up computations considerably, but also simplify the production of application programs.

The first data structure, developed under the working title "Geographic Data Structure" (GDS)[1] is based on irregularly distributed points which are assumed to be sample points without sampling errors from a single-valued surface. Two types of structures build the core of the data-bases. The first creates neighborhood relationships by "triangulating" the data set and storing for every point the labels of all points which are linked with the point by a

[1]ONR Contract # N00014-73-C-0109, Thomas K. Peucker, Principle Investigator, Simon Fraser University, Burnaby, Canada

triangle-edge. The second structure is produced by selecting those points of the surfaces which lie along lines of high information content as ridges and channel-lines and defining them by their nodes, which are peaks, passes and pits. This second data structure serves two purposes: first it is a general representation of the surface for rough computations on the surface; second, it is a "directory" into the more detailed first structure.

The creation of the neighborhood relationship is based on the assumption that the majority of data-sets to be handled is of one of two types:

(a) Sets of irregularly distributed points which were digitized with the understanding that every part was significant.

(b) Sets of regularly or irregularly distributed points where it is known that a number of points are redundant and can be eliminated from the set. Typical candidates for this group are the regular grids of points and encoded contours.

The first type of data set is linked by some type of triangulation. At least two approaches exist. The first (Dueppe and Gottschalk, 1970) creates all possible links, chooses the shortest and eliminates all links which intersect with the shortest. This procedure is repeated with the next shortest links until no links intersect. The result is the set of links with the minimum cumulative distance between neighboring points.

The procedure has one disadvantage: since $\binom{n}{2}$ links have to be created, the number of points is therefore limited to only several hundreds. The first of our approaches therefore limits the links to a number of "potential neighbors," among which the shortest link is chosen and intersected with the "second order potential neighbors," i.e. the potential neighbors of the potential neighbors of a point. This procedure on the one hand limits the number of tests for intersections of links to less than $\frac{n \cdot m^2}{4}$ where m is the number of potential neighbors, an arbitrary number, which will be between 8 and 14 depending on the distribution of points. On the other hand, the procedure does not guarantee that only triangles are constructed. Polygons with more than three sides can result, although they are relatively rare. The check for such polygons and their elimination is very easy and fast.

The second possibility is to create a triangulated structure through use of Thiessen polygons. A published solution (Rhynsburger, 1973) intersects for every point the links to every other point midway and chooses the smallest polygon created by the perpendiculars.

Every point which contributes to the Thiessen polygon is a Thiessen neighbor. This procedure can again be simplified by the assumption of a limited set of "potential neighbors." The same checking routines as above have to be applied. They are very simple, however, given the very elegant structure.

Another approach which limits the number of necessary tests but is mathematically correct at the same time, has been developed within the project.[1] The procedure is based on "fields of potential neighbors" which converge very rapidly.

The alternative to triangulation is three-dimensional generalization, i.e. to select from a set of points those which define the structure with the least deviation from the original surface. The basic concept[2] is to approximate the surface of a series of triangles through a selected set of points, where each additional point included into the set is the one which deviates the most from the approximated triangles until the deviations are below a given value (see Peucker, 1974).

In both cases, the triangulation and generalization of the surface, the result is a "linked list" of surface points. The term linked list means that points are linked with one another through pointers. In other words, a point is not only identified by its x, y, z coordinates, but also by a list of the labels of the points which form edges of triangles with the point.

In our case each record consists of the x, y, z coordinates of a point and a reference to the start of the pointers to the neighbors in a pointer-list. The reason for not having the neighborhood pointers with the point record is that the number of points varies considerably (Mark, 1974, in Peucker). Since the record had to be long enough to include all possible numbers of neighbors large parts of the pointer sections would be empty for most of the time. The pointers are sorted, starting with the pointer the least East of North (Fig. 6) of a point.

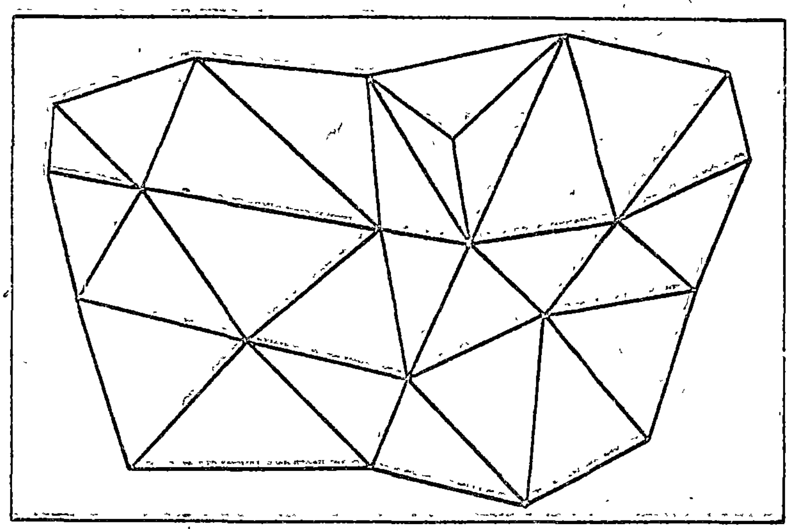
The use of this type of data-structure is very simple and efficient. For every search (profile, contour, etc.) a criterion for edge-intersection is developed. For the contouring it is, for example, the question whether one point of the edge is above the contour level and the other below. A start is found and one end point of the edge considered a reference point and the other a subpoint. The next subpoint is found by looking up the next neighbor in the pointer list. If the test is positive, the intersection is performed and the process repeated. If the test is negative, the reference and subpoints are switched and the process repeated.

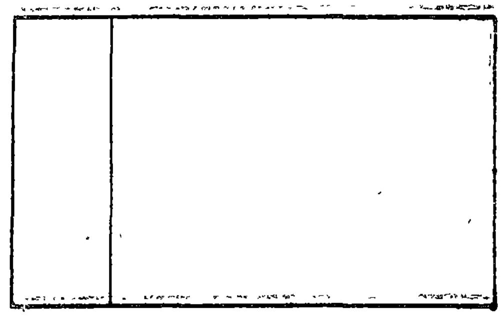
Other procedures are equally simple. To find a triangle for example, one only has to have a reference and a subpoint. The third point is the next label in the pointer list of the reference point after

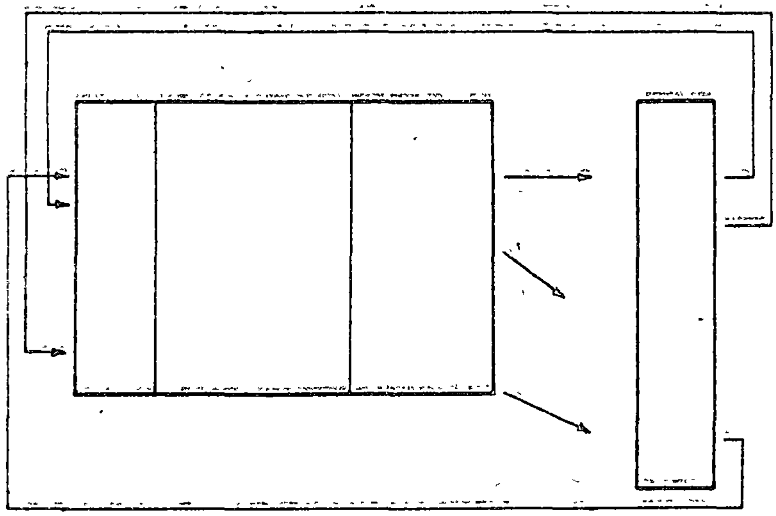
---

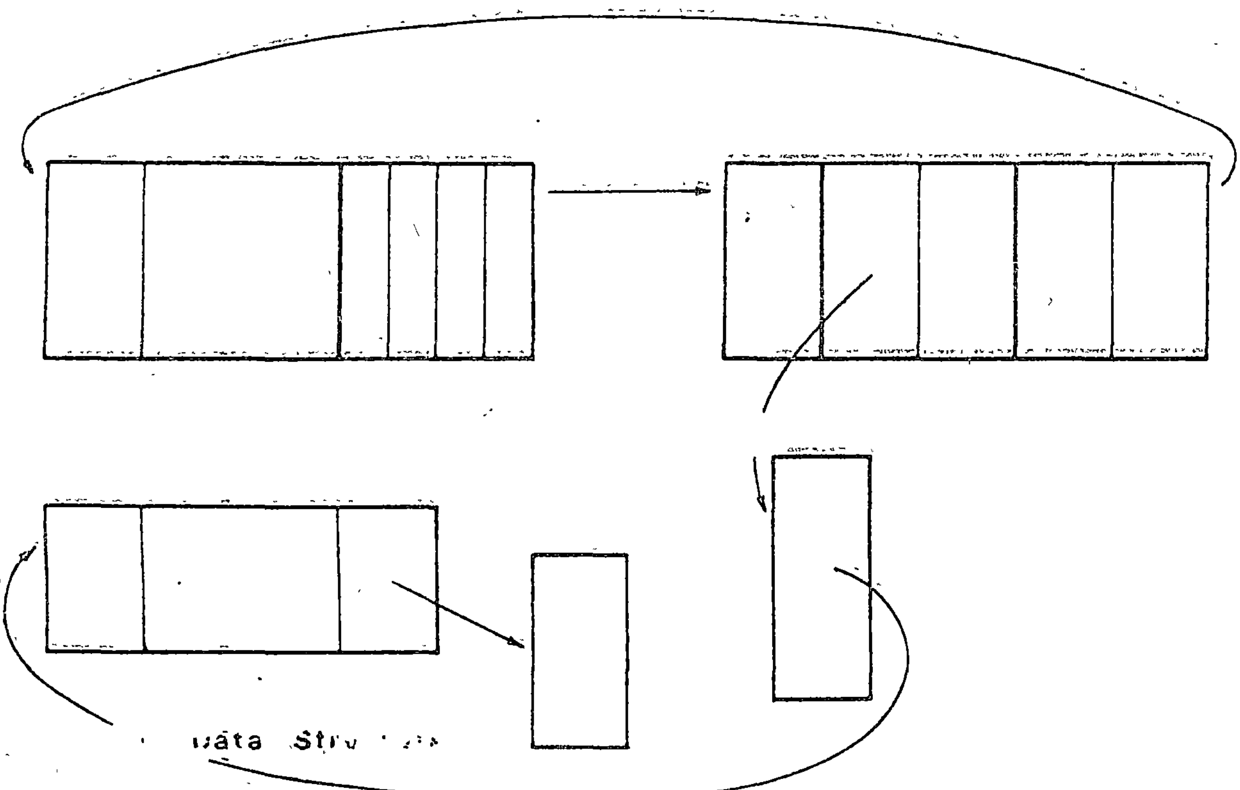
[1] by Kurt Brassel, Harvard University

[2] developed by Randolph Franklin, Harvard University and T.K. Peucker.









ata Stru

A third component of the "Geographic Data Structure" should be mentioned since it shows very well our problem finding geographically logical solutions to computer problems. The problem at hand is the partitioning of the data set. Since with large data sets only portions can be kept in fast memory, the data base is segmented into "pages" which are brought into memory as units. For the "Geographic Data Structure" the paging system can solve several problems inherent in a complex geographic information system.

The boundaries of "patches," as we call the areal extension of a "page," are usually chains already defined for the second structure. Since detail along the chain is of no topological interest, the density of points along the chain can differ for its two sides. In other words, the density of triangles can change from patch to patch. This allows for a very efficient data encoding even in terrain with sudden changes in the "surface behavior" (Peucker, 1972) as at the change from a mountainous area into a plain. (Fig. 7).

Another advantage of the paging-system is the ease of including topographic and planar information. Linking point, line and areal data to the triangulated points would lead to high definitional redundancy. The secondary structure could lead to ambiguities where the terrain is very elongated. Since it is attempted to keep the shape of the patches as compact as possible, the combination of non-terrain data with patch-boundaries seems to be most appropriate.

Since the patch-boundaries are again chains, another virtue comes to light; the patches can be treated as polygons of the POLYVRT and GEOGRAF programs with little difficulties. This link between the two systems lets us hope that eventually we will be able to merge the two systems.

The question is appropriate as to what such a data structure will be able to accomplish. We hope that it will allow us to perform all applications, most of them not only with less programming efforts but also at lower computing costs. (Fig. 8).

We have had practical support for this statement only in the area of surface displays, but since both levels of the data structure are graph-structures, (Pfaltz, 1972) we will be able to rely on all the developments connected with graph theory and network analysis.

A DATA STRUCTURE FOR PLANAR SURFACES

The need to incorporate different types of polygons uncovered one of the limiting assumptions made in POLYVRT. A chain plays a dual role: it is the boundary of two areal entities; and secondly, it is the unbroken unit of point retrieval. This distinction is not significant if one is interested in only one polygon type. The following description of a new proposed system, to bear the name GEOGRAF, makes this distinction.

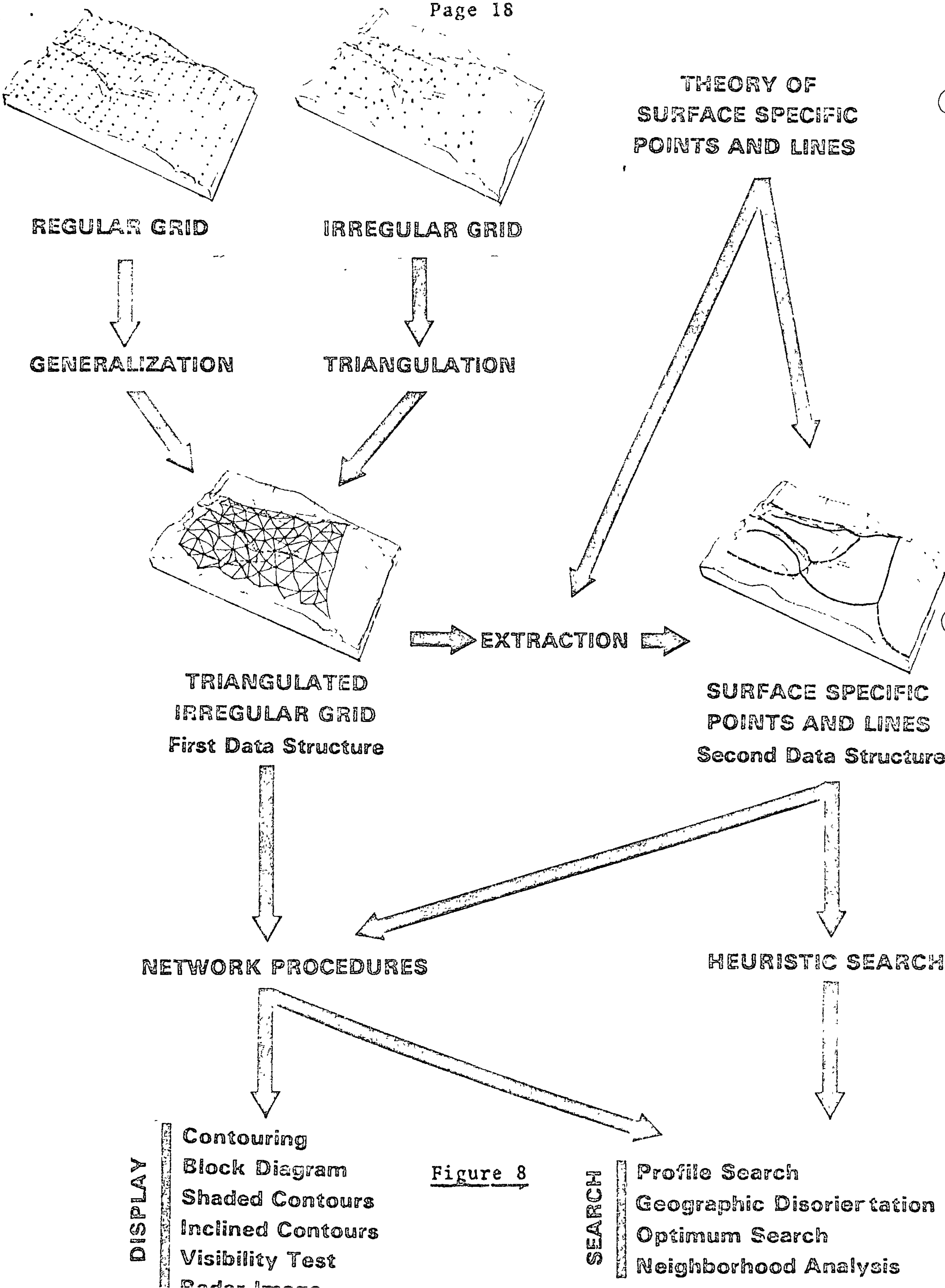

THEORY OF
SURFACE SPECIFIC
POINTS AND LINES
REGULAR GRID
IRREGULAR GRID
GENERALIZATION
TRIANGULATION
EXTRACTION
TRIANGULATED
IRREGULAR GRID
First Data Structure
SURFACE SPECIFIC
POINTS AND LINES
Second Data Structure
NETWORK PROCEDURES
HEURISTIC SEARCH
DISPLAY
Contouring
Block Diagram
Shaded Contours
Inclined Contours
Visibility Test
SEARCH
Profile Search
Geographic Disorientation
Optimum Search
Neighborhood Analysis

Figure 8

Because of the addition of many layers of complexity involving polygon types, the chain cannot remain, to the same degree, the controlling object of the data structure. Just as the notion of an unbroken line is important, so is the notion of an unpartitioned space. In a system which must handle overlapping polygon networks, there is a need to have a root object which described areas uncut by the union of all partitionings. This object is termed the Least Common Geographic Unit (LCGU). The relationship of the LCGU's to all other polygon types will be hierarchical (see Fig. 9).

In turn, the existence of the LCGU will allow for the creation of each polygon system. The LCGU's will be constructed as a POLYVRT polygon directly from chains. In order to allow simple coding of the boundary relationships at each higher level in the structure, the object of the "Chain Group" was invented. This object represents a set of chains which, for a given polygonal level, is a boundary of two areal units. In summary, the higher level polygons are constructed of chain groups which, in turn, are constructed form chains.

Line features can be built up of chains in the same manner as chain groups. Note that each level of polygon, with its chain group listing, and each line feature type, references only the chains themselves. This will allow each system to be considered as a separate directory which is core-resident only when that class of objects is retrieved.

The LCGU has other implications and applications that are useful because of the topological data structure. The LCGU, with its coding for each of the polygon sets, can be combined with contiguity information of linear feature types to produce an Attribute Cross Reference (ACR). The ACR is a table in which all objects (in polygon and linear systems) are cross-referenced to each other to determine nesting. By using attributes of chains (lengths) and LCGU's (areas, population densities, etc.), this cross-referencing capability could assign a string of data, collected by one polygon type, to a string of a second type.

As has been expressed previously, a data structure is of little use without application procedures. Topological manipulation routines are therefore central to the success of this structure. Several examples should be mentioned here:

The intersection of geographic features will rely on topological knowledge to realize economies of scale in processing large files. All operations with lines will actually work with bands, build with endpoints of the line, and the furthest deviants to both sides (if the bands become too wide the lines are split, etc.) With this approach, the windowing process using non-linear windows (which is often the case with geographic coordinates and map projections) becomes quite elegant. Similarly, polygon procedures (point-in-polygon, line across polygon, polygon-over-polygon) allow for gainful application of the topological principle. One constructs chain groups which bisects the universe into parts and sorts the points (or nodes of polygons) into three groups - left, within the band, and right. Only the second group needs more detailed treatment. One then recursively partitions the point set into more halves until one has partitioned the set to the level of the LCGU's.

The second important procedure will be a nested chain-intersection routine. Here, again, the chain-band and its recursive segmentation is used. The number of points which define a chain is constantly increased until the intersection test can be determined without any doubt. The search of a line through a set of polygons will use a graph-search algorithm developed for the GDS (Geographic Data Structure) project. The neighborhood search routines create records of neighbors for every point or line or polygon at any specified depth of neighborhood.

The creation of polygons from points is the major link from the GDS project ot GEOGRAF. The creation of a set of centroids for polygons allows the conversion in the opposite direction. This way, surfaces can be treated as polygonal sets, and can be displayed and manipulated by the routines of GEOGRAF. Conversely, polygonal data can be treated as surfaces for GDS. The neighborhood routines are what make the project an innovation for quantitative geography and planning. Without the topological data structure, neighborhood searches are extremely expensive. However, they are at the same time a most interesting part of urban and environmental analyses once a general overview is obtained.from the data. Without this data structure, research into detail has to retreat to qualitative studies and field analysis, a revealing but cumbersome undertaking which should be started only after full benefit is gained from a detailed quantitative micro-analysis of the data. With the neighborhood analysis, in connection with interactive graphics, a profound evaluation of the data can be undertaken in a very satisfactory user environment.

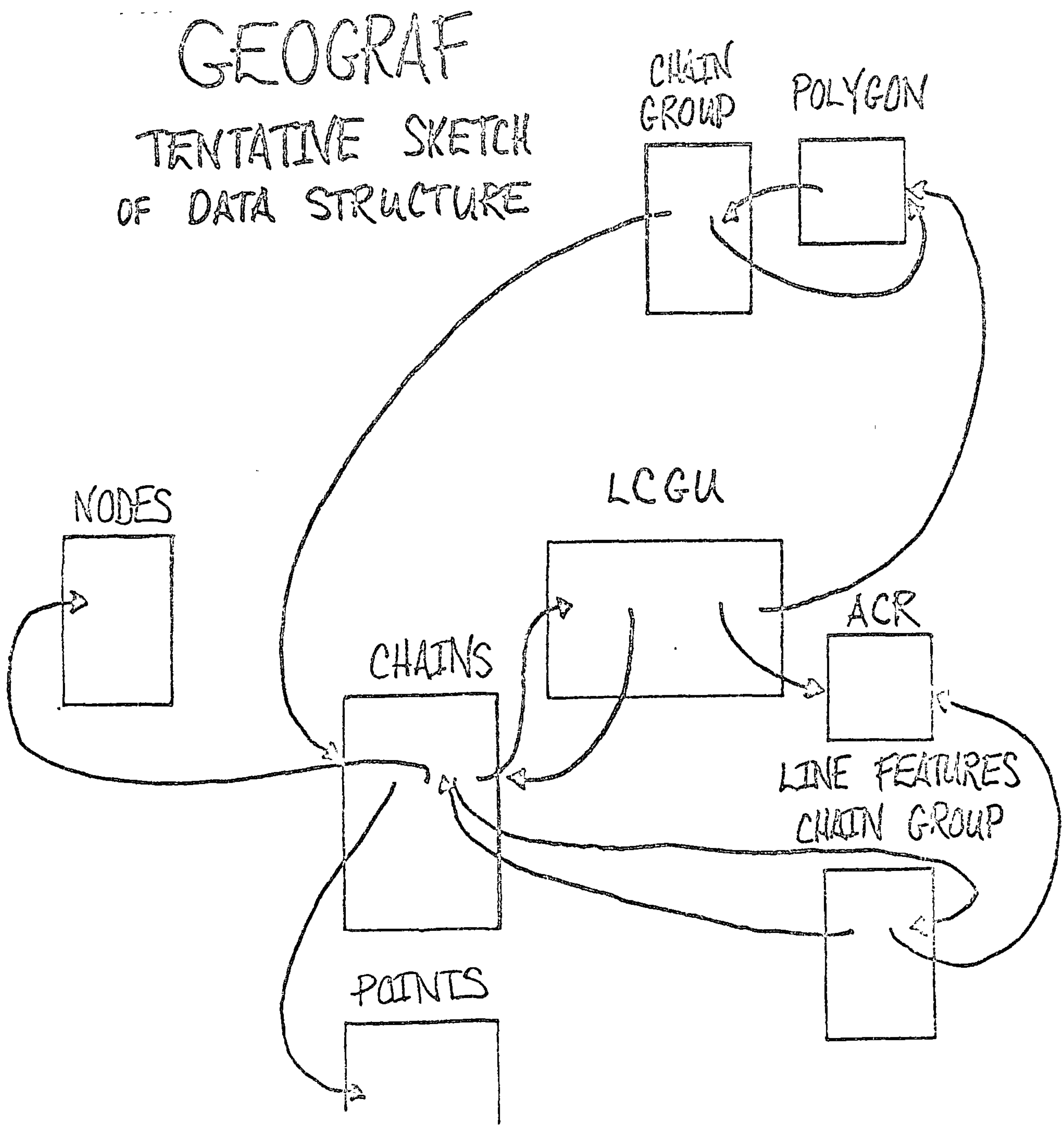
CONCLUSIONS

It might seem as if the proposed systems are presented as the definitive geographical information systems. If the paper makes this impression, we would like to apologize for our excitement which has carried us away to such immoderately sounding statements. But we hope that our research will bring geographic thinking into cartographic data processing and will initiate others to try out the power of the topological approach for the cartographer's and geographer's work.

We firmly believe that basic research and application development are the two sides of a coin and have to go together to get lasting results. In this paper we have concentrated on the theoretical parts of the projects since their development is ahead of the application routines, a fact which should be expected.

The quintessence of our research so far is the hypothesis that topologically structured data bases of three-dimensional and planar surfaces can result in reduced efforts in the development of application routines and a superior numerical performance. We have some indication that the hypothese might be correct; However, the real test will come when the bulk of the application routines is completed.

Figure 9



LITERATURE

Amidon, E.L. and G.S. Aiken, 1970: Algorithmic Selection of the Best Method for Compressing Map Data Strings, Communications, Assoc. for Computing Machinery, vol. 14, 1970, 769-774.

Arms, A., 1970: MAP/MODEL System: System Description and User's Guide. Bureau of Governmental Research and Service, University of Oregon, May 1970, 59 pp.

Boehm, B.M., 1967: Tabular Representation of Multivariate Functions. With Applications to Topographic Modelling. Proceedings, 22nd National Conference, Assoc. for Computing Machinery, 1967, 403-415.

Brandstaetter, L., 1957: Exakte Schichtlinien und Topographische Gelaendedarstellung, Oesterreichische Zeitschrift fuer Vermessungswesen, Sonderheft 18, Wien 1957, 90 pp.

Dueppe, R.D. and H.J. Gottschalk, 1970: Automatische Interpolation von Isolinien bei willkuerlich verteilten Stuetzpunkten, Allgemeine Vermessungs-Nachrichten, vol. 129, No. 10, Oct. 1970, 423-426.

Grist, M.W., 1972: Digital Ground Models: An Account of Recent Research, Photogrammetric Record, vol. 70, No. 4, Oct. 1972, 424-441.

Heiskanen, W.A. and H. Moritz, 1967: Physical Geodesy, San Francisco, Chapter 7.

Holroyd, M.T. and B.K. Bhattacharyya, 1970: Automatic Contouring of Geophysical Data Using Bicubic Spline Interpolation, Geological Survey of Canada, Paper No. 70-55.

Hsu, M-L, et al., 1974: Computer Applications in Land Use Mapping and the Minnesota Land Management Information System, in Davis, J.C. and M. McCullagh (eds.): Display and Analysis of Spatial Data, New York, forthcoming 1974.

Junkins, J.L., G.M. Miller, J.R. Jancaitis, 1973: A Weighting Function Approach to Modelling of Irregular Surfaces, Journal of Geophysical Research, vol. 78, No. 11, Apr. 1973, 1794-1803.

Morse, S.P., 1968: A Mathematical Model for the Analysis of Contour-Line Data, Journal, Assoc. for Computing Machinery, vol. 15.

Nake, F. and T.K. Peucker, 1972: The Interactive Map in Urban Research, Report after Year One, University of British Columbia.

Nordbeck, S. and B. Rystedt, 1972: Isarithmic Maps and the Continuity of Reference Interval Functions, Geografiska Annaler, vol. 52, ser. B, 1970, 92-123.

Peucker, T.K., 1972: Computer Cartography, Association of American Cartographers, College Geography Commission, Resource Paper No. 17, Washington, D.C.

Peucker, T.K. (ed.), 1973: The Interactive Map in Urban Research; Report after Year One, University of British Columbia.

Peucker, T.K., 1974: Geographical Data Structure, Report after Year One, Simon Fraser University.

Rhynsburger, D., 1973: Analytic Delineation of Thiessen Polygons, Geographical Analysis, vol. 5, No. 2, April 1973, 133-144.

Rosenfeld, A., 1969: Picture Processing by Computer, New York.

Schmidt, W., 1969: The Automap System, Surveying and Mapping, March 1969, 101-106.

Shepard, D., 1968: A Two-Dimensional Interpolation Function for Irregularly Spaced Data, Proceedings, 23rd National Conference, Assoc. for Computing Machinery, 517-524.

Sima, J., 1972: Prinzipien des CS digitalen Gelaendemodells, Vermessungstechnik, vol. 20, No. 2, 1972, 48-51.

Switzer, P., 1974: Sampling of Planar Surfaces, in Davis, J.C. and M. McCullagh (eds.): Display and Analysis of Spatial Data, New York, forthcoming 1974.

Tobler, W.R., 1969: Geographical Filters and their Inverses, Geographical Analysis, vol. 1, 234-253.

Warntz, W., 1966: The Topology of Socio-Economic Terrain and Spatial Flows, Papers, Regional Science Assoc., vol. 17, 1966, 47-61.

Laboratory for Computer Graphics and Spatial Analysis
Graduate School of Design Harvard University

Allan H. Schmidt
Acting Director
Eric Teicholz
Assoc. Director

INTERACTIVE MAPPING OF URBAN DATA

Eric Teicholz
Associate Director
Laboratory for Computer
Graphics & Spatial Analysis,
Harvard University
Graduate School of Design

April 1975

This paper originally appeared in the Proceedings of the Second General Assembly of the World Future Society, June 1975.

I. ABSTRACT

The article describes several factors that are contributing to the explosion of statistical and geographical data related to our urban areas and the corresponding interest in automated procedures for the input (capture), analysis and display of spatial data.

Basic automated mapping methods and procedures are described and illustrated using two interactive mapping systems called INPOM and ASPEX, developed at the Laboratory.

Finally some limitations of computer mapping and cost considerations are described.

II. INTRODUCTION

In a recent issue of the New York Times, an article appeared stating that Canada was "going metric" and that this event is expected to increase pressure on the last major non-metric holdout - the United States.

One ramification of this conversion is that large numbers of existing manual maps will become immediately obsolete. Another possible result of this drift towards metric conversion is that the Laboratory for Computer Graphics and Spatial Analysis (the Laboratory) is receiving numerous requests primarily from utility companies and planning agencies, requesting information on how to go about developing automated techniques for the collection (data capture), analysis and display of spatial data.

There are other reasons why there is a great deal of interest in automating procedures relative to the analysis and display of urban information. There is a large increase in both the volume and quality of statistical data. This is due to automated techniques being used by traditional agencies such as the Department of Commerce (who collect and distribute Census data) as well as a result of a host of other governmental and private agencies using computers for collecting statistical data along with geographic identifiers - thereby creating geographic entities which can be used for a variety of analysis and display programs. The Central Intelligence Agency, for example, distributes political boundaries for all countries (World Data Bank I) and will soon release much more detailed breakdowns (World Data Bank II). The United States Geological Survey (USGS) has a multi-million dollar commitment to automate the National Map Series and they claim that all the USGS topological maps will be available in digital form in five to ten years.

Along with the increasing amount of available geographical and statistical data is a demand for new, more detailed and more accurate data on the part of urban researchers. New integrated hardware and software cartographic "turnkey" systems make this data more available and cost beneficial than ever before. Commercial companies such as Computervision (Bedford, Mass.), Applicon (Burlington, Mass.), Calcomp (Anaheim, Cal.), and Calma (Sunnyvale, Cal.) offer such systems. In addition, most major time sharing companies now support remote graphic applications encouraging the development and use of interactive mapping. Finally, dissemination groups such as the Laboratory, the Census Bureau and the Geography Program Exchange (East Lansing, Michigan) are distributing low cost graphic display programs.

III. BACKGROUND

There are several classes (types) of maps that are used for the mapping of spatial data. Base maps display geographic entities such as boundaries (Census tracts, Standard Metropolitan Statistical Areas, blocks), road and river networks or almost any of the twenty-odd cartographic features that are overlayed to produce a USGS topological map. Base maps are normally used to convey locational data but do not convey other types of statistical information.

Thematic maps display geographical concepts such as gradients, density distributions, magnitudes of various attributes or other quantitative or qualitative data. To display geographical concepts, a variety of techniques are employed such as various types of symbolisms, grey tones and color symbols and tones - all of which can be superimposed on a base map.

Statistical surfaces can be represented using chloropleth or isarithmic maps. The former represent statistical variables by conforming to a particular boundary or enumeration district. Input will consist of polygon coordinate data and statistics that relate to the geographical areas. Isarithmic mapping emphasizes gradients such as contours or other isarithms to represent areas and volumes to portray a continuous real (or assumed) statistical surface.

An additional type of map that should be mentioned is called a cartogram which deliberately distorts areas or volumes to represent an aerial quality. The example below, taken from an article by R. L. Phillips in the April 1974 Proceedings of the IEEE (Vol. 62, No. 4, p.442) illustrates a retail market view of the United States employing a program that produces cartograms.

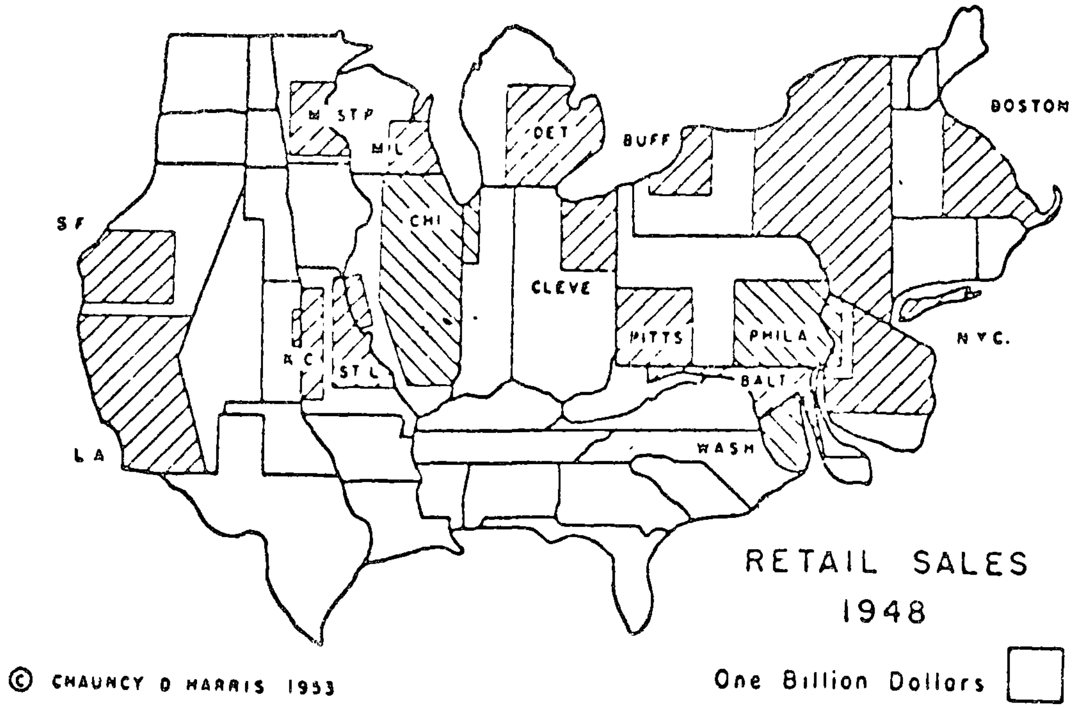


Figure 1: Cartogram Illustrating Retail Sales in the U.S. for 1948

IV. INTERACTIVE MAPPING: The ASPEX and INPOM Programs

Perhaps interactive computer mapping offers one of the most powerful tools for urban researchers to date. One can perform on-line data editing operations using an intermediate display device such as a cathode ray tube (CRT) (either color or black and white), and can selectively retrieve and "massage" data to perform a variety of statistical operations. A user can then alter values, class intervals, symbol and shading types of the output display map. Finally, a user can transform the data, look at each view on a CRT until the desired result is achieved, and then output the final display file to a variety of display hardware devices such as digital plotters, COM (computer on microfilm) plotters, a color matrix plotter (such as the new Color Jet Plotter from Sweden), photoplotters or some other output device. In some cases the resolution of the cathode ray tube itself might be satisfactory. The ASPEX and INPOM illustrations appearing in this article were reproduced from a Tektronix 4610 hard copy output.

At the Laboratory, a variety of ongoing research projects are involved in the interactive capture, processing and display of spatial data.

INPOM

The Interactive Polygon Mapping System (INPOM) is designed to produce maps of countries, states, census tracts and other arbitrarily shaped regions defined within a geographic base file (GBF). INPOM is a two-dimensional mapping program capable of producing conformant base and thematic (shaded) maps. It has the capability of selectively retrieving areas to be mapped, of controlling the degree of detail to be displayed (for outlines) and in the symbolism used to depict data values. The user can zoom in on particular areas of interest, try different types of symbolism, get immediate hard copy from the CRT display, and vary the amount of detail to be displayed.

The flexibility of input is achieved by entering keyword-type commands from the display terminal. The program responds by requesting additional information needed to execute the command. Because of the internal data structure used by the program, it is possible to get listings of points coordinates, chains (the data structure used by the program), single polygons, or user defined regions within the study area.

At present, there are over 30 input commands operating in conversational mode. The commands are entered as 2, 3, or 4-letter nmemonics and the program will respond accordingly by requesting numeric data or alphabetic responses. The numeric responses may be values, coordinates (which may be stored internally if desired), window parameters and the like. All data is free field format so that the user does not have to worry about restrictive fixed field formatting requirements.

Another flexibility of INPOM is that all commands have default conventions or values which the program will preset for the parameters of a command until the user employs that command. Once set, the parameters of a command remain in effect until the command is again specified. Figure 2 below shows the current command file for the INPOM program. Figure 3 (detail level = 1) is a base map of Africa while Figure 4 (detail level = 5) shows a thematic map of Africa showing the gross national product on a per capita basis based on 1970 data. Figure 5 (detail level = 10) shows the same data illustrating the zooming and increased detail level for a section of West Africa.

File Input, Creation and Querying

INC	(INput Chains)
INP	(INput Polygons)
VAL	(input VALues)
FN	(File Name)
RV	(Read Values)
WV	(Write Values)
LIN	(define LINe legends)
CVRT	(input Chains from the POLYVRT program)
REN	(REName chains and or polygons)
INFO	(list INFOrmation on chains, polygons, lines)

Value Level and Symbolism Definition

NL	(Number of value Levels)
LVL	(define value LeVeLs)
SHD	(define SHaDing for levels)
FAS	(FActor Shading density)

Termination

EXIT	(EXIT from INPOM to monitor level)

Windowing, Scaling and Polygon Selection

WORG	(Window ORIGin location)
WSIZ	(Window SIZe specification)
MW	(Move Window across map)
PMM	(Polygon selection by Minimum and Maximum extents)
DMM	(Data coordinate Min-Max selection of polygons)
FDW	(Fill Data Window with partial chains)
FSW	(Fill entire Screen Window with polygons and chains)
FA	(FActor for expanding or shrinking map)
XFM	(Transform point coordinates with respect to a location)

Graphic Manipulation Instructions

OL	(to specify OutLine or shaded maps as output)
DET	(highest DETail level to be drawn)
DL	(Draw Line legends)
DWO	(Draw Window Outline)
PLT	(PLoTing mode for Tektronix 4014 display only)
MAP	(draw a MAP)

FIGURE 2
Command File for INPOM

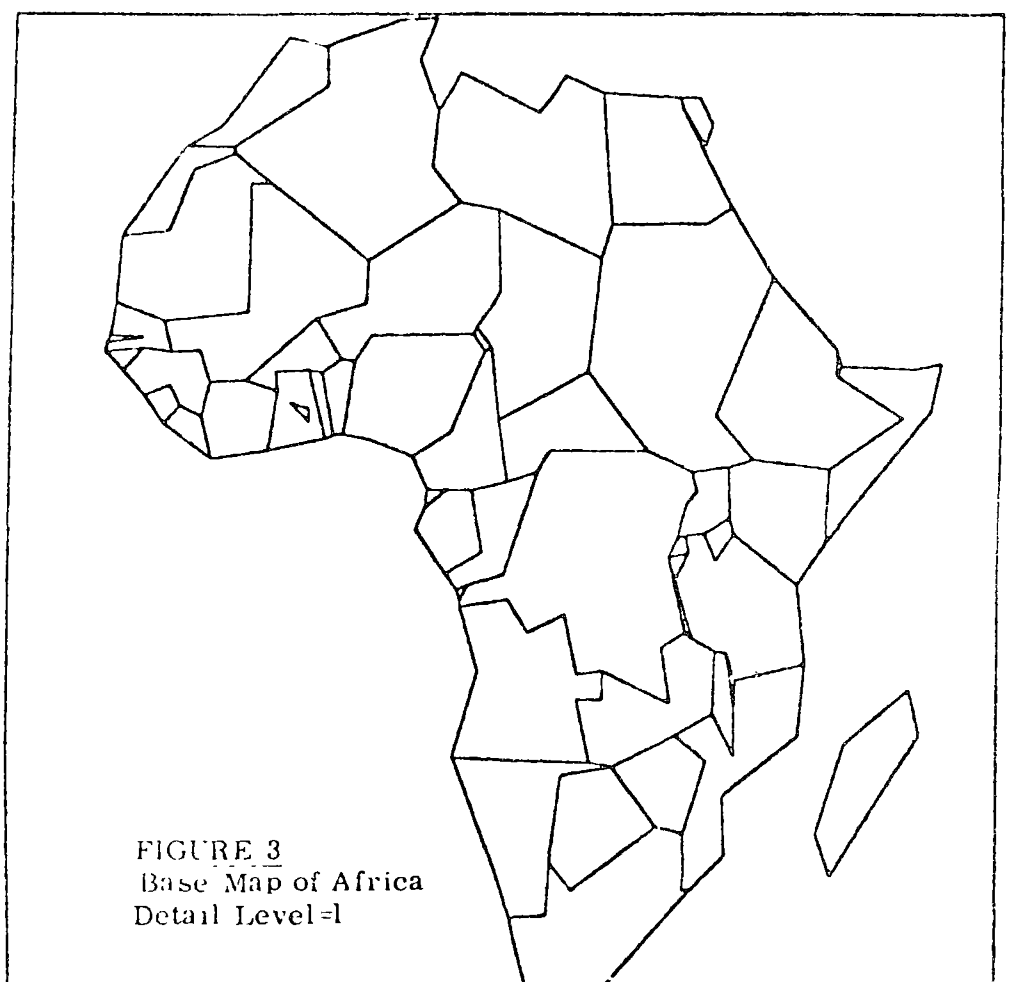
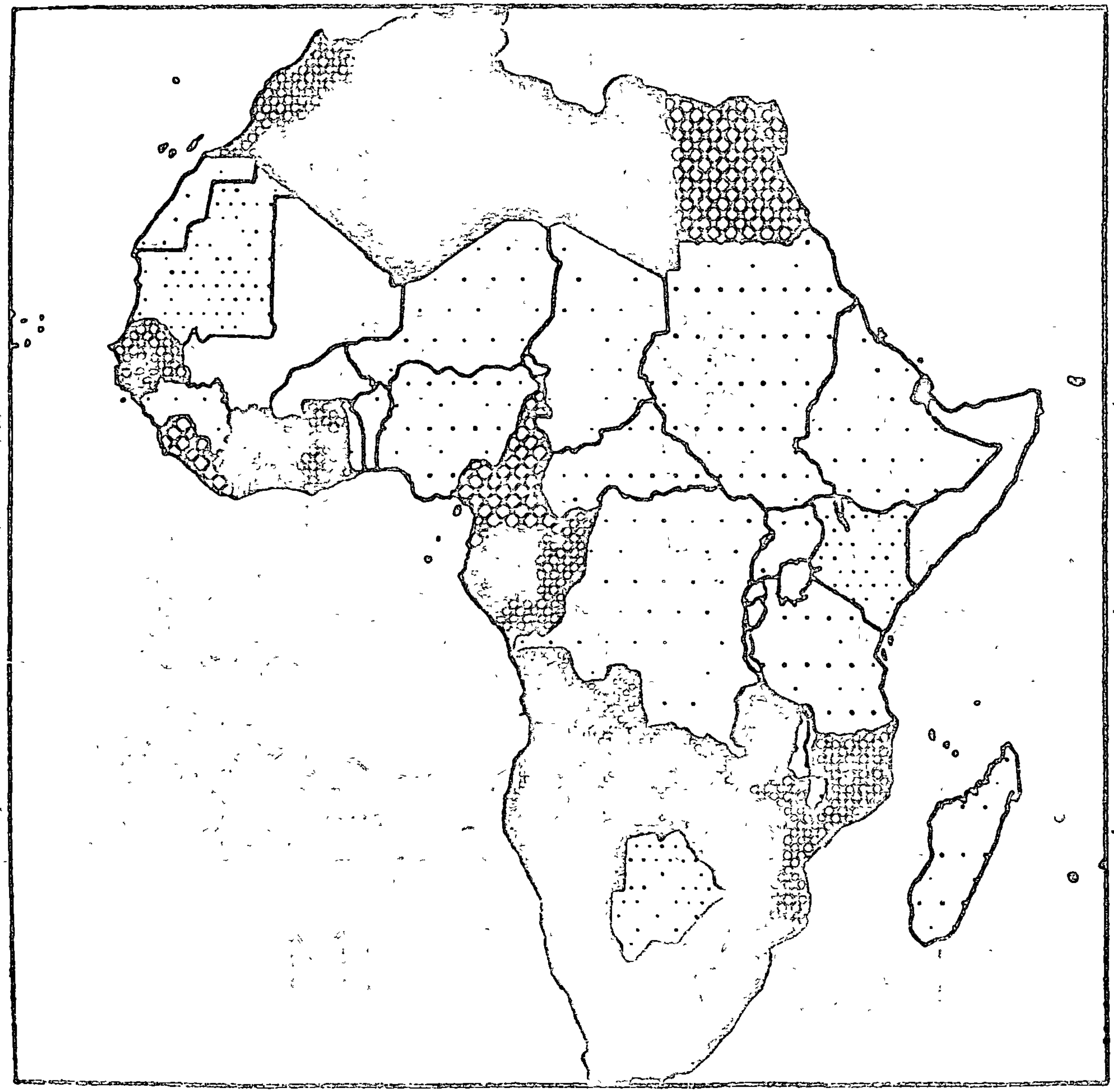


FIGURE 3
Base Map of Africa
Detail Level=1

AFRICA

CHAIN BASE FILE CONSTRUCTED FROM WDB-I



INPOM

LABORATORY FOR
COMPUTER GRAPHICS
AND SPATIAL ANALYSIS

GNP/CAP

GROSS NATIONAL PRODUCT
PER CAPITA (1970)

Figure 4

Thematic Map of Africa
Detail Level = 5

WEST AFRICA

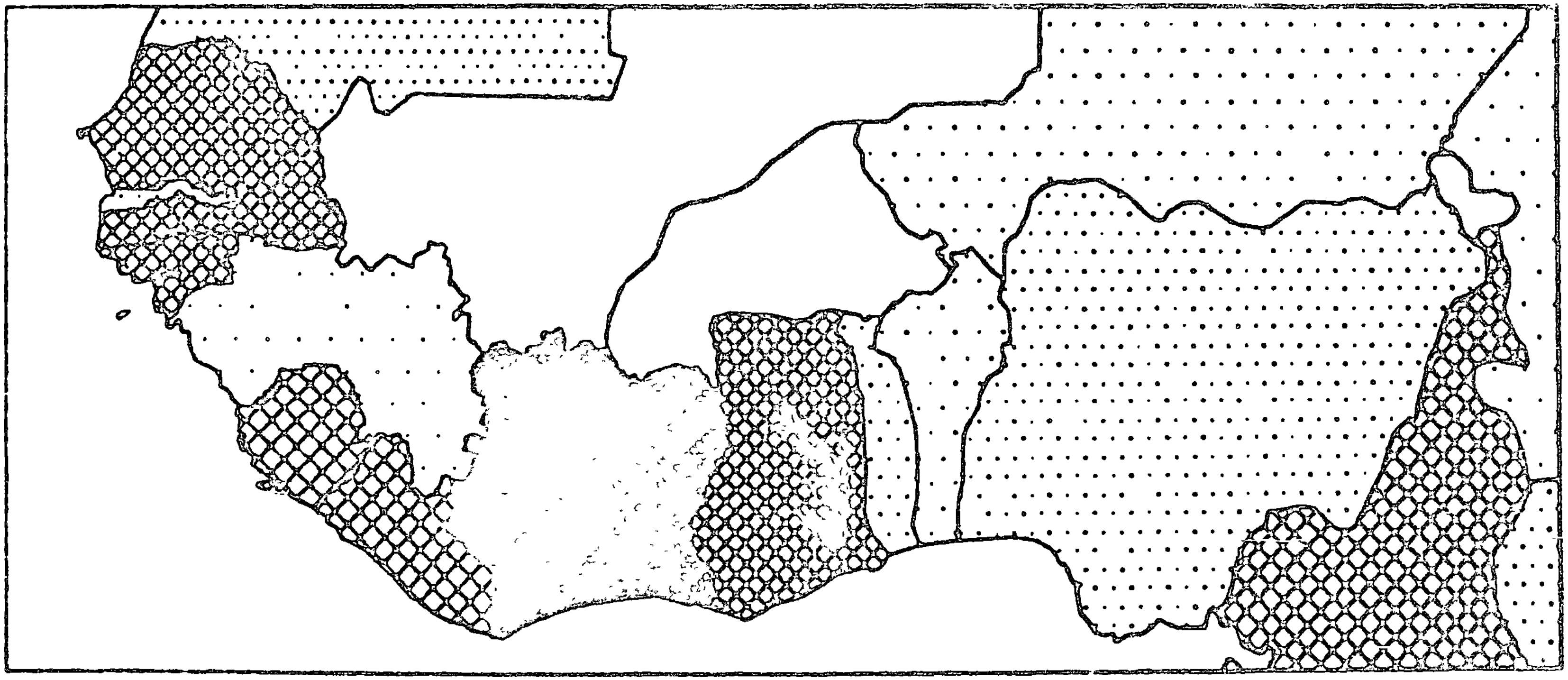


Figure 5

Detail of West Africa
Detail Level = 10

BY INPOM

ASPEX

The Automated Surface Perspective Program (ASPEX) is an interactive (a batch version is also being developed) program that displays three-dimensional representations of statistical surfaces. Such representations of three variables were not very common until the advent of the computer because of their difficulty in construction. Although most people are not particularly accustomed to reading information displayed on surfaces (especially when the information is statistical or mathematical in nature), mathematicians, cartographers and planners are beginning to accept three-dimensional surface representation as a powerful extension of two-dimensional mapping.

The ASPEX program takes a matrix (or array) of data of any size. The program incorporates a free field format command language that operates on nmemonic keywords for the over 70 commands of the program. The commands deal with the following categories:

- -initialization and production (such as display, help, define, expunge, plot, etc.)
- -data input and storage (number of columns, data type, grid input, header information, etc.)
- -data value manipulation (min, max, smoothing, square root, etc.)
- -viewing parameters (including view type such as isometric, planometric, and perspective) and orientation
- -graphic options (including data surface commands such as draw, height, interval, symbol size and cosmetic features such as base information, map scale, title, etc.)

Another important flexibility of ASPEX is the ability to alter the viewpoint so that a user can be located anywhere beyond, above, or upon the surface. Capabilities are also being added to draw features directly on the surface. The scaling of the output plot may be to any predetermined height, width or window size and is accomplished automatically by the program. The three figures below represent different views of the U.S. but generated from the same data base.

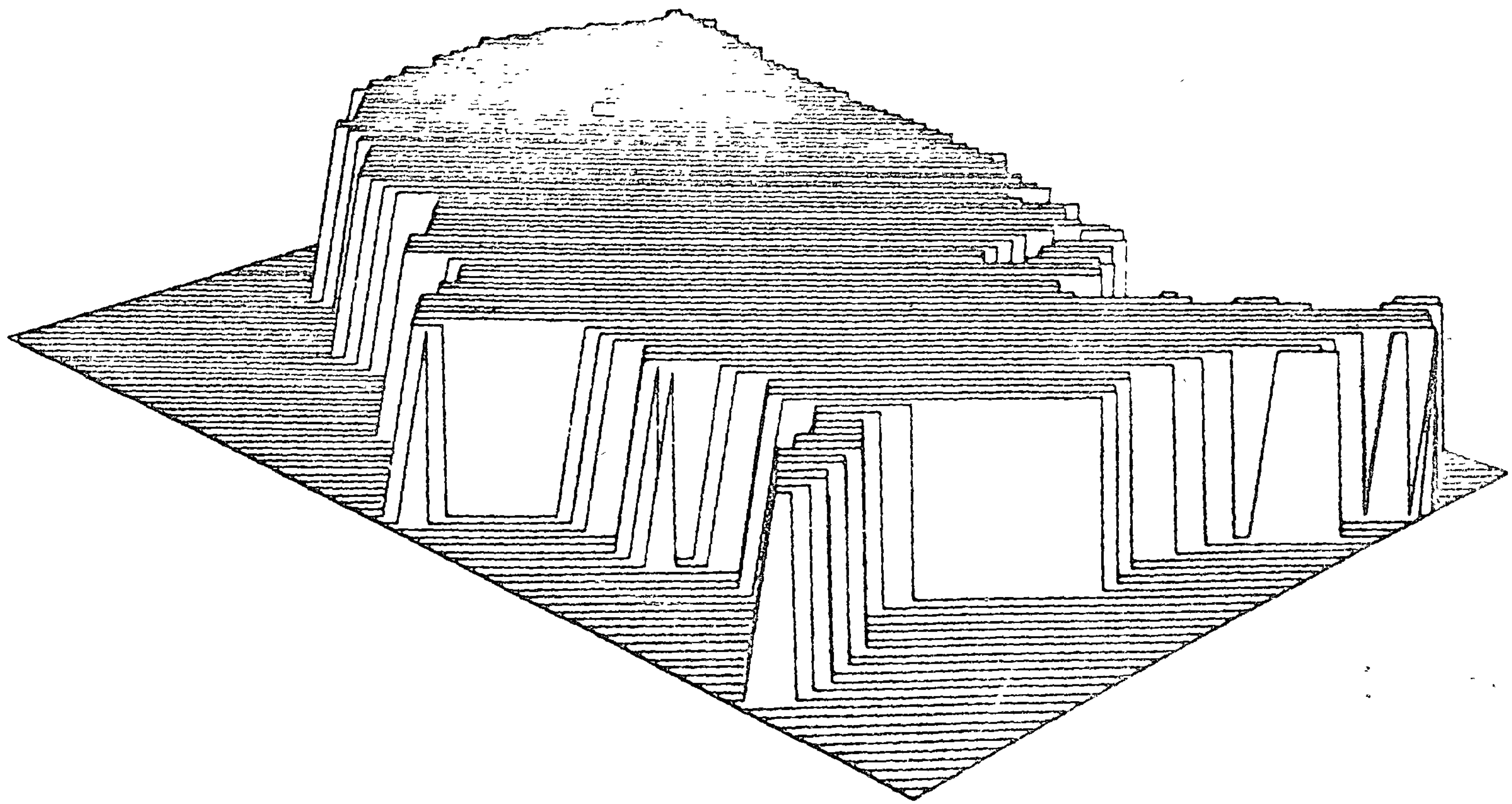


Figure 6

Truncated Height Values

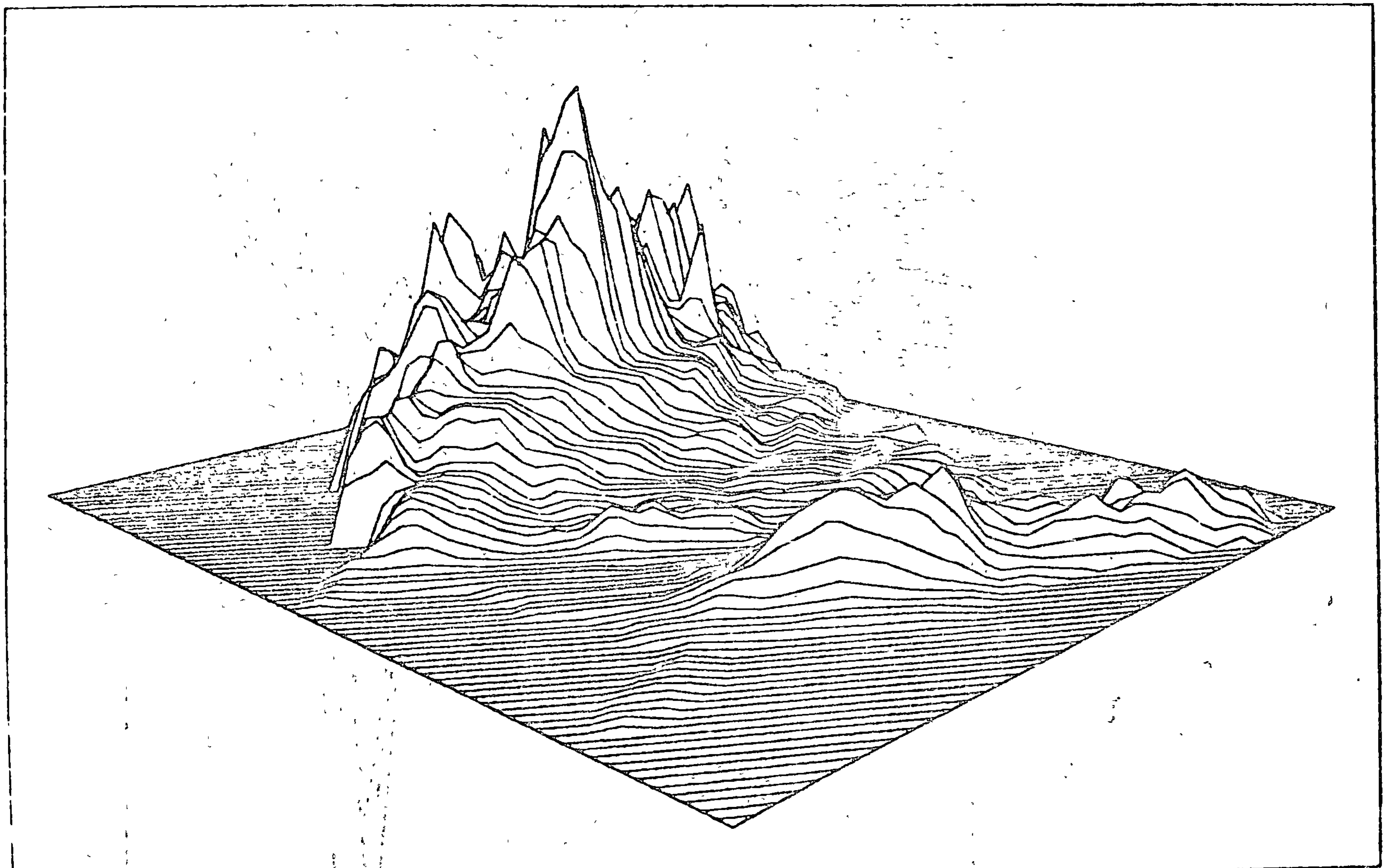


Figure 7

View from the South East

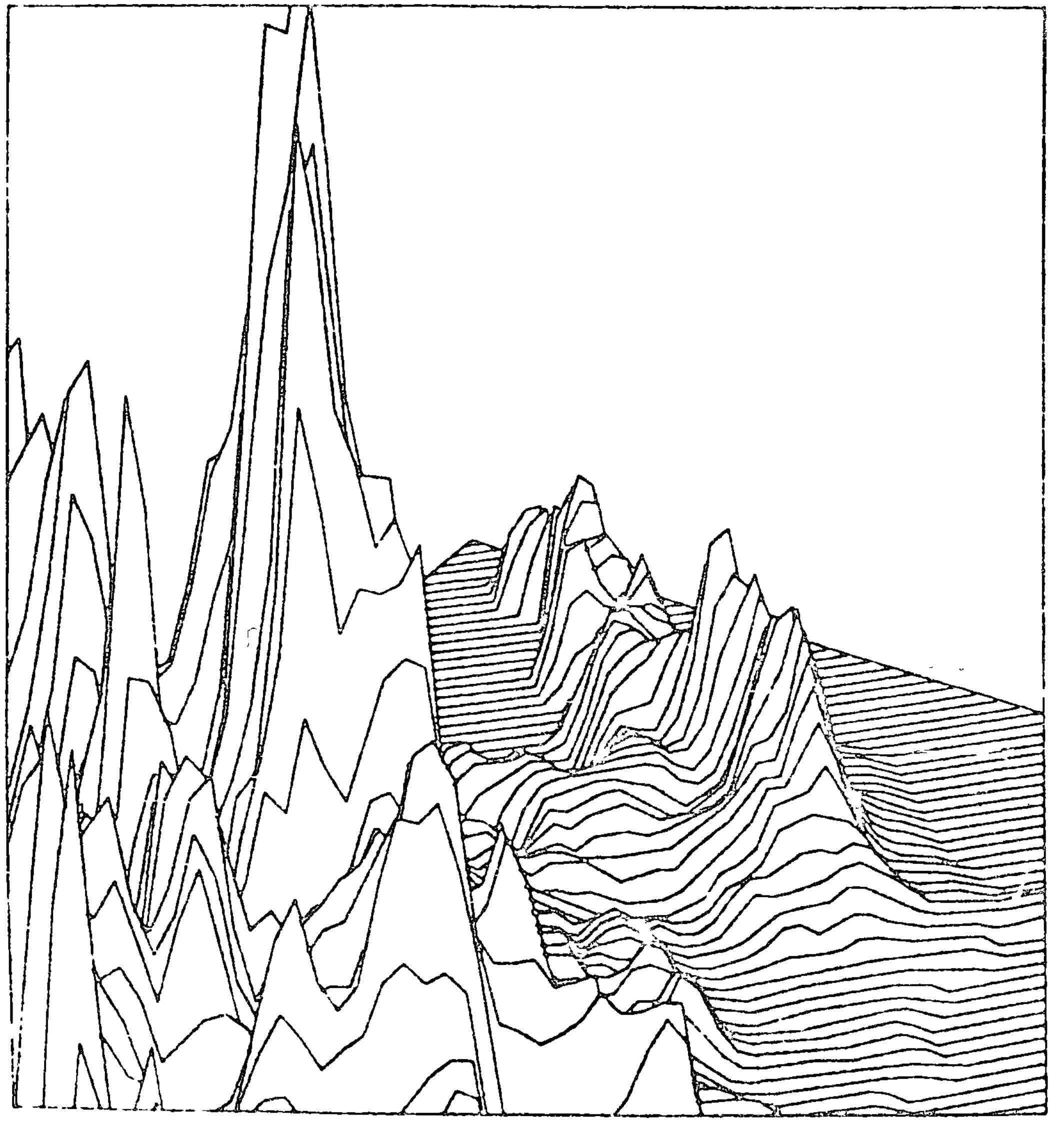


Figure 8

View from the Rockies

V. CONCLUSION

The software to rapidly and economically generate computer maps is clearly available. Unfortunately, this does not necessarily make automated mapping a pragmatic analytic tool. The economics of computer cartography must take into consideration the totality of the process. Questions on input such as where the data is coming from, its reliability, and updating procedures must be considered. One is reminded of the millions of dollars that was spent on data banks in the 1960's that became data dumps in the 1970's.

One must also take into account GBF's from non-contiguous sources that are in different formats. There are also central processor considerations which must reduce the data to its lowest common denominator, restructure it, process it and finally output a display file for a particular output device.

Output devices range from the ubiquitous line printer which produces inexpensive low-resolution maps using over-printing techniques to digital line plotters to COM (computer on microfilm) and expensive photoplotters.

A final word of caution relates to the maps themselves. Being able to produce maps efficiently and inexpensively is no guarantee that the maps will facilitate and improve the decision-making process of urban researchers. One must be able to clearly understand what is being represented by a map and the purpose for which it is intended. Only then will the application of computer technology have any meaning.

ALGORITHMS FOR THE REDUCTION OF THE NUMBER OF POINTS REQUIRED TO REPRESENT A DIGITIZED LINE OR ITS CARICATURE

DAVID H DOUGLAS AND THOMAS K PEUCKER

University of Ottawa/Simon Fraser University, British Columbia

ABSTRACT All digitizing methods, as a general rule, record lines with far more data than is necessary for accurate graphic reproduction or for computer analysis. Two algorithms to reduce the number of points required to represent the line and, if desired, produce caricatures, are presented and compared with the most promising methods so far suggested. Line reduction will form a major part of automated generalization

Lines from maps and photographs are recorded numerically for cartographic manipulation to facilitate their reproduction at different scales and projections and to allow map compilation with other geographic data bases. Usually lines are approximated by straight line segments, and end points of which are recorded by a pair of co-ordinates in either polar or orthogonal measure. The other more important methods by which lines are recorded are chain encoding and skeleton encoding. Chains approximate lines by a sequence of end to end vectors, where the length and direction of the vectors are selected from a fixed, usually four or eight, number of possibilities.[1] Skeleton encoding is directed more at recording closed areas or polygons by filling the area with circles or rhombi of different sizes. The lines forming the boundaries are recorded by implication.[2] The conversion of graphic data to computer readable numerical forms is effected with a co-ordinate digitizer, a bit plane scanner or an automatic line follower. A co-ordinate digitizer converts a pointer's location on a table to x y values which can be written on punched cards or magnetic devices. Polar co-ordinate digitizers, which consist of a slide in a rotating anchor head, record a radius and an angle from a base vector. Another digitizing device consists of a pointer suspended from a pair of retracting wires which activate potentiometers. Conversion of values in one recording co-ordinate system to another can be performed easily with small computer programs

Drum scanners superimpose a vast and very fine grid over the document to be digitized recording a "yes-no" or "on-off" value for each cell location, depending on whether that cell covers a line or not. A trade off is introduced between the fineness of the mesh, implying more computer processing time to reduce the data to forms which are easily handled, and coarseness of the image recorded. On the other hand, the mesh density, being dependent on hardware, is fixed at the time of manufacture and is usually set to be somewhat smaller than the minimum line width. In all cases, the reduction of a bit plane scan, in which lines are represented by clouds of cells containing numerous discontinuities, to chain or vector encoded lines, is a complex process requiring processing time and resources which could only be described as being quite substantial

With a co-ordinate digitizer lines may be recorded in point mode, time or increment automatic modes. Lines recorded in point mode are effectively generalized by the operator who subjectively selects points which best approximate the line to the degree he desires. This presumes, among other things, that he is his own customer. Point digitizing is extremely tedious however, and is unsuitable for anything but the simplest data sets, such as the generalized

David H. Douglas is a lecturer in Geography at the University of Ottawa and is presently on a study leave at Simon Fraser University, Burnaby, British Columbia. Thomas K. Peucker is an Associate Professor at Simon Fraser University, Burnaby, British Columbia

MS submitted June 1973

[1] H. Freeman, "On the Encoding of Arbitrary Geometric Configurations", *Institute of Radio Engineers, Transactions on Electronic Computers*, Vol. Ec-10, 1961, pp. 26-268

[2] J. R. Pfaltz and A. Rosenfeld, "Computer Representation of Planar Regions by Their Skeletons", *Communications of the ACM*, Vol. 10, No. 2, February 1967, pp. 119-122 and 125

outlines of counties or census tracts. Most coordinate x-y digitizers on the market possess, as options, time or increment automatic recording modes. Points are recorded automatically in a given time interval, or after the cursor has moved a preset distance along the x and/or y axis. The prime limiting factor on the speed of recording is the speed of the output device. Magnetic tape transports which record up to 300 characters per second are commonly available, allowing up to 20 or 30 points to be recorded each second. To record coastlines, contour lines, or other lines of high frequency oscillation it is evident that the minimum speed required, given the speed at which an operator can follow a line, is in the order of 5 to 10 points per second, which effectively eliminates paper tape and punched cards as output media. Digitizing onto magnetic tape has more than its share of problems, primarily because there are no foolproof means to ensure the data are correctly recorded at the time of digitizing, and because of the inordinately frequent occurrence of non-confirmable digitizing errors such as line ends which should, but do not meet, lines recorded twice and so forth. The editing procedures necessary are time consuming and clumsy. These problems have been met by elaborate on-line procedures where a mini-computer interfaced to the digitizing table oversees the whole operation, checks and double checks the data recorded, closes loops and signals when it senses a great many errors, such as cursor movement too fast to be accurate.[3]

All digitizing methods, except perhaps for the possible exclusion of point digitizing on a co-ordinate digitizer, record, as a general rule, far more points than necessary to reproduce the line on most graphic devices, even at the scale and resolution of the original line. The elimination of data representing unnecessary points, such as duplicates,

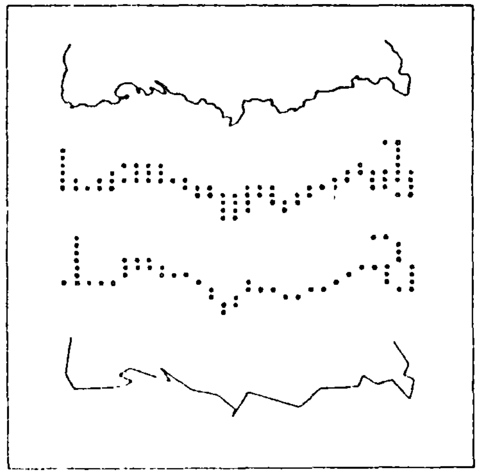


Figure 1 Line represented by 140 points on the plotter and the printer, and the same line represented by 25 points

and points along a straight line, can be of significance, simply because of the diminished storage requirements. As well, the operating speed of many spatial analysis programs and the plotting speed of many graphic devices are related inversely to the number of points to be processed or plotted. Reduction of a line by elimination of unnecessary points representing it assumes a more positive advantage if the line is to be abstracted or caricatured purposely, if the scale of reproduction is to be smaller, or if the output device, such as some Cathode Ray Tube plotters, has a cruder resolution than represented by the original digitized line. Lines which have a higher frequency of oscillation than can be represented within the resolution capability of the graphic device become fuzzy and weak.[4] Figure 1 illustrates line data at the resolution of recording, its reproduction on the computer printer and the reproduction of a greatly generalized version of the line. Given the crudeness of the printer as a graphic device it is evident that the simplified version of the line is preferable to the unsimplified one mainly because of the elimination of most

[3]A. R. Boyle. Computer Aided Compilation, Hydrographic Conference, Ottawa, January, 1970

[4]and are similar in effect to the data clouds recorded by a bit plane scanner

of the double lines and data clouds. Since this line was better represented by 25 points than it was by the original 140 obviously some computer pre-processing was justified

There have been a great many approaches suggested and algorithms programmed to reduce the number of points required to represent numerically recorded lines. Some of these are in regular use within planning agencies and cartographic units. Not all of the methods have been exhaustively tested to measure or judge their cartographic usefulness and there have been few, if any, studies to compare the methods with each other. The methods can be classed broadly into the categories of elimination of points along the line by one or more of a multitude of criteria; approximation of the line with a mathematical function, and deletion of specific cartographic features represented by the line. Of these categories, it would seem that the last one would come closest to duplicating the task as performed by an experienced cartographer as he generalizes

The cartographer attempts to maintain the character and overall impression of an empirically defined, or hand drawn line by selective deletion of some of the details. A fjorded coast is represented by only a few of the actual number of fjords, a delta by only a few of the actual number of channels and so forth. The automation of this approach would rely therefore on the ability to program the computer to recognize specific cartographic features. One attempt is based on an interactive computer program which has the ability to "learn" from the actions of an operator.[5] The operator generalizes a line plotted on a cathode ray screen by signaling the deletion or maintenance of points. As the computer "learns" from what the operator selects it attempts to recognize similar features on its own. This system at its present level of development concentrates on the angular and length relationships of a very small number of segments, but the number of possible ways to represent a single simple class of feature, such as a peninsula, is simply staggering. This interactive system, therefore, represents but a small step towards the solution of a fantastically complex problem

The second group seeks to approximate the points along a line with mathematical functions. This can be done for the whole line at once or it can be done in some piece-wise order taking a small number of connected points at a time. There are several different methods fitting into the latter category. One developed by A. R. Boyle for the Hydrographic Survey of Canada (1972) computes a first order least squares line through a fixed number of points and then steps forward in that direction by a predetermined distance. Two other approaches begin by defining the ends of segments as averages of a fixed number of points along the line. Koeman and Vander Weiden[6] suggest taking the mean while Jancaitis and Junkins[7] take the distance weighted centroid. When these central points are joined the results simulate a piece-wise approximation with functions of the first order. It must be mentioned, however, that the stated purpose of Jancaitis and Junkins was to smooth and not necessarily to reduce the line

The resulting data sets of extracted functions are economical in terms of storage

[5] Andrew H. Clement, "The Application of Interactive Graphics and Pattern Recognition to the Reduction of Map Outlines", Master's Thesis, University of British Columbia, 1973

[6] C. Koeman and F. L. Vander Weiden, "The Application of Computation and Automatic Drawing Instruments to Structural Generalization", *Cartographic Journal*, Vol. 7, No. 1, June 1970 pp. 47–49

[7] James R. Jancaitis and John L. Junkins, *Mathematical Techniques for Cartography*, Final Contract Report for U.S. Army Engineers Topographic Laboratory, Fort Belvoir, Virginia, Contract No. DAAK02-72-C-0256, February 1973, pp. 15–20

space required, but are relatively time consuming in the processing stage. The greater the number of points, the more costly and complex the operation. These functions reproduce lines which are typically much smoother than the lines they represent. In the main they are probably much better suited for smoothing than reduction and have to be considered of limited value for generalizing. Functions extracted in a piece-wise fashion tend to under-represent erratic curves and over-represent smoother curves. Methods which look for central tendencies are inclined to depress the effect of extreme points. Unfortunately, these are often the very points which give character to the line.

Of the group of methods which eliminate points, some concentrate on the points which are to be deleted while others are directed towards selecting those points which are to be maintained. The algorithms directed at deleting points are usually the simplest. In the case of data recorded by time-automatic digitizing a simple test to drop those closer than one resolution unit can eliminate a large percentage of the points recorded. This method can be extended by purposely decreasing the numerical resolution or by establishing a threshold distance. Points closer than this distance to neighbours are dropped.[8] For chain encoding a simple compression on the basis of consecutively equal vectors can also result in significant savings. This can be extended as well for other types of encoding by dropping points whenever the direction of the line is not changed through a threshold angle by the segments subtended on it. The underlying purpose of these methods is to eliminate wasted data space but since the line plotted after this kind of processing would look very much the same as it would before it cannot represent a significant step towards automated generalization.

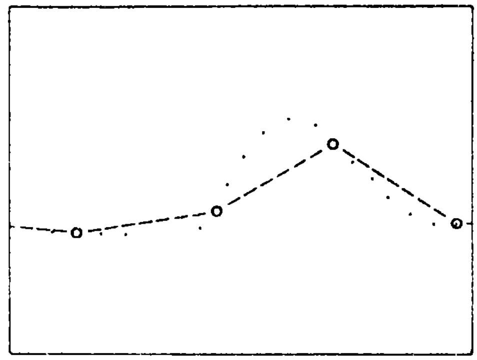


Figure 2. Line reduction by the selection of every sixth point

The simplest and most often used method of line reduction is to delete all but every n^{th} point along the line where n is a fixed integer based upon the desired degree of reduction.[9] The method does not require much in the way of computing resources and it furnishes acceptable results if the digitizing was extremely dense. The primary disadvantage is the frequent elimination or misrepresentation of important features along the line such as promontories, indentations, sharp angles and so forth. A secondary limitation is that straight lines are still over-represented. These shortcomings are made obvious in Figure 2.

The alternative to deleting points is to select them. In the special case of monotonically increasing lines (for instance, just one value of "y" for every "x"), crests and troughs may be selected. The obvious disadvantage here is the omission of points where there is a change of direction but which nonetheless are not crests or troughs. For irregular planar curves, the problem is more difficult. Jarvis con-

[8] W. R. Tobler, "Numerical Map Generalization", *Michigan Inter-University Community of Mathematical Geographers, Discussion* Paper No. 8, Department of Geography, University of Michigan, January 1966.

[9] Experimental Cartographic Unit, Royal College of Art: *Automatic Cartography and Planning*, London, Architectural Press, 1971.

verts the Cartesian to polar co-ordinates and then looks for crests and troughs.[10] This is useful for curves which can be made monotonic by this conversion, but, as for Cartesian measure, the solution cannot be considered general.

One alternative to line generalization which seemed to hold conceptual promise was that method provided by the German firm A E G which supplied the Experimental Cartographic Unit with its GEAGRAPH 1000 plotter and was described by T Lang in 1969.[11] This method was reported as producing acceptable results but was eventually rejected as a general purpose technique by the Experimental Cartographic Unit on the grounds that it required far too much computer time for the on-line processing system being operated at the time The objective of the procedure was to delete points if they were found to lie within a tolerance distance of a straight line segment being tested to represent a portion of the line From one representative point it constructs straight lines to subsequent points until one point between the representative point and the sub-point is further away from the line linking the two than a pre-set tolerance value As soon as this condition is satisfied, the point before the sub-point becomes a new representative point and the procedure is repeated The method gives acceptable results in the case of smooth curves but it does not detect the best representative points on sharp curves and the results are particularly unsatisfying where sharp angles are numerous

The methods proposed in this paper are based on a concept somewhat similar to the pre-set tolerance ideas described by Lang but concentrates rather on the selection of points rather than on their deletion

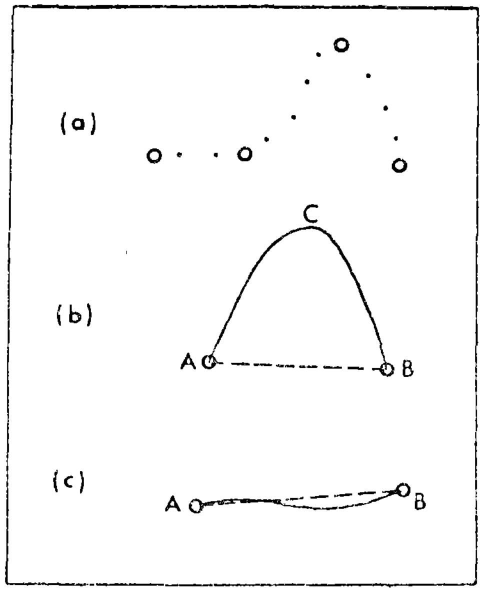


Figure 3 Subjective selection of representative points

Approaches to a computerized solution to many problems begin with an examination of the way one would solve them subjectively Consider the line represented by points illustrated in Figure 3(a) One might choose the encircled points as those which represent the original line to our own requirements of accuracy Perhaps the reason we would select these points and not others might be illuminated by examining the simpler situations in Figure 3(b) and (c) Starting with the obligation to begin with the end points, the question might be "Why would there be a compulsion to insert a point C in (b), where no such compulsion would exist in (c)?" The perpendicular distance of C from the segment A-B may provide a clue This suggests that an arbitrary maximum distance could be established If no point along the line is further than this distance from the straight line segment connecting its end points then the straight line seg-

[10] C J Jarvis, "A Method for Fitting Polygons to Figure Boundary Data", *The Australian Computer Journal*, Vol 3 1971, pp 50 54

[11] T Lang "Rules for Robot Draughtsmen", *Geographical Magazine*, Vol. XLII, No 1, Oct 1969, pp 50-51

ment will suffice to represent the original line. If this condition is not satisfied, then another point along the curved line must be selected and the same test would be carried out with the new segments. The next question is "What point along the curved line should be selected to become the end point of the two new straight segments created?" The obvious answer is the furthest point from the straight segment. Although it is possible that this point may be embedded in a long smooth curve, it is more likely that it is the apex of a relatively sharp angle. As well, this point has already been identified as a result of the distance search, therefore, the benefits associated with its selection far outweigh the possible attraction of selecting some other representative point. In the case of closed loops, where the first and the last point do not define a line then the maximum perpendicular distance from the segment is replaced with the maximum distance from the point. The same process would be repeated with the new segments created until the maximum distance requirement is satisfied for all straight segments.

Two different procedures embodying these principles have been encoded in FORTRAN IV and tested. In addition Method 2 has been encoded as a recursive function in ALGOL W.[12]

Method one begins by defining the first point on the line as an anchor and the last as a floating point. These two points define a straight segment. The intervening points along the curved line are examined to find the one with the greatest perpendicular distance between it and the straight line defined by the anchor and the floater. If this distance is less than the maximum tolerance distance then the straight segment is deemed suitable to represent the whole line. In the case where the condition is not met, the point lying furthest away becomes the new floating point. As the cycle is repeated the floating point advances toward the anchor. When the maximum distance requirement *is* met the anchor is moved to the floater and the last point on the line is reassigned as the new floating point. The repeat of this latter operation comprises the outer cycle of the process. The points which had been assigned as anchor points comprise the generalized line.

Method two is exactly the same as method one except that note is taken of all points which have been assigned as floaters on previous inner cycles. These are stacked in a vector. After the anchor point is moved to the floating point, the new floating point is selected from the top of this stack, thereby avoiding the necessity of re-examining all the points between the floater and the end of the line. This procedure usually results in the selection of a slightly greater number of points than Method 1, but takes approximately 5 per cent of the computing time and is thought to produce better caricatures. This method can also be thought of as taking a logically hierarchical approach to line reduction. On one cycle extreme points are selected and these tested to see if they suffice. If they do not, intermediate points are taken and the same question asked about each of the two new segments produced, and then each of the four new segments are examined, and so on as if in a branching tree. Each branch is terminated when the offset tolerance criterion is satisfied.

To enable valid comparisons four separate subroutines were written on the basis of the procedure described by Lang. One was an exact duplication of that procedure while the other three were combinations of two incorporated modifications.

[12] Andrew H. Clement, "The Application of Interactive Graphics and Pattern Recognition to the Reduction of Map Outlines", Master's Thesis, University of British Columbia, 1973.

The program Lang describes starts by assigning the first point as the anchor and third as a floater The second is tested to see if it lies within tolerance distance of the segment defined by the anchor and the floater If it does, the fourth is assigned as the floater and the second and third are examined and so on The first floating point defining a segment which does not allow all intervening points to satisfy the tolerance criteria causes the anchor to move to the point before the floating point Since selection of the point immediately before the floating point has no cartographic justification, the first modification of the procedure has the anchor point move to the point furthest from the segment The reasoning behind selecting the furthest point is that it is the one most likely to subtend a sharp angle and would therefore have the best chance of properly representing the line The second modification attempts to cut computing time by avoiding unnecessary repeated calculations of distance From Figure 4, it is clear that in most cases the sum of the distances a + b + c is greater than the greatest distance that P_1, P_2, or P_3 lies from the segment P_0P_4 In other words, if a + b + c is less than the tolerance distance then d also would be less than the tolerance Only one distance, rather than all of the intervening ones has to be calculated on each cycle The inner cycle, intended to find the point lying furthest from the segment, is invoked only in the cases where the accumulated total is greater than the tolerance Positive and negative accumulations are kept in separate registers to avoid subtractions from their absolute magnitudes in the case of double curves The maintenance of these running registers is particularly useful when series of points lie along straight lines

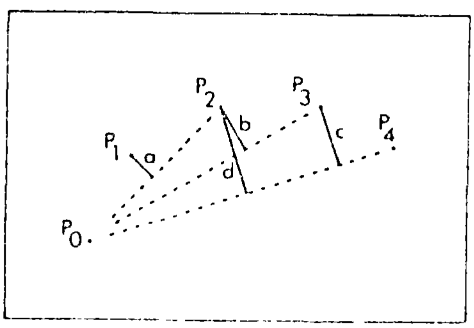


Figure 4 Running registers of accumulated offset distances

The first modification which attempts to select a point which is more rationally defined than simple convenience, has the expected result of approximately doubling the number of points selected and the processing time required to isolate them The second modification definitely reduces the time required to process a given line, especially if a great many points are deleted because they lie along relatively straight segments

All procedures were tested and compared, both for their ability to remove unnecessary points, that is with the offset tolerance set to be less than the resolution of the plotting device (Figure 5), and their ability to produce caricatured representations (Figure 6) All were judged to produce satisfactory results for simple line reduction however the versions of the A F G procedure without the modification to pick the furthest point from the tested segment did not produce satisfactory caricatures because of the tendency to omit and cut corners The methods presented in this paper were tested with substantial data sets and found to be operationally suitable both for simple reduction and in the production of satisfactory abstractions (Figures 5 and 6)

Detailed comparisons in computing time required for each sub-routine were made on the basis of a three inch square and a three inch diameter circle, each made up of 4000 points evenly distributed along its periphery It was felt that the square would give ample opportunity to demonstrate the power of each routine in the

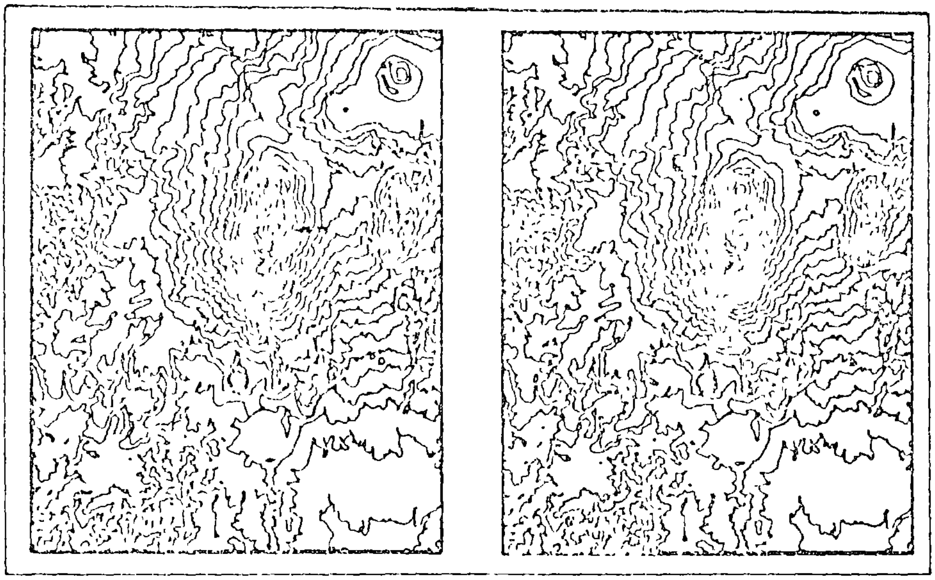


Figure 5. Contours plotted from the original digitized, unduplicated at .001 resolution, 41,311 points (left), and from 7,782 points (right) reduced by Method 2 with a tolerance set to half the resolution of the plotter. The reduction procedure added 16.5 seconds to the 64 seconds required to read and write the data to plot the map on the left. The images may be compared with a simple stereoscope.

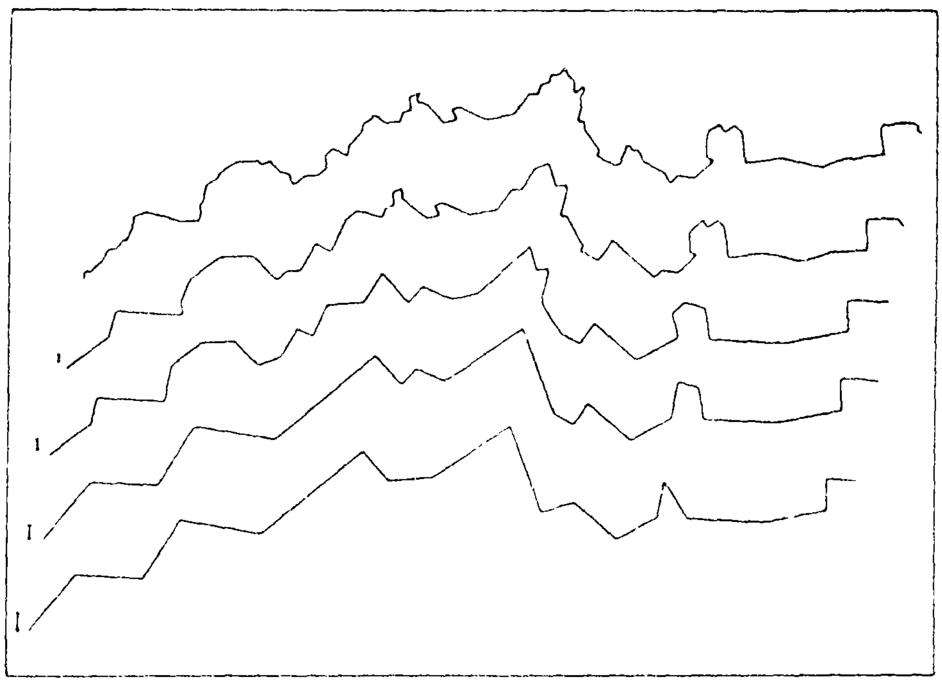


Figure 6. Line reduced and caricatured by Method 2. The tolerance value employed is shown to scale at the left of each caricature which was reduced from the original data represented by the top line.

TABLE 1

PROCESSING TIME REQUIRED TO REDUCE A 3″ CIRCLE AND A 3″ SQUARE MADE UP OF 4000 POINTS EACH EVENLY SPACED ALONG THE PERIMETER TO THE NUMBER OF POINTS INDICATED WITH THE GIVEN OFFSET TOLERANCE

SQUARE	Offset Tolerance (inches)											
	.001		.005		.01		.05		.1		.5	
4000 points	Points	Time	Points	Time	Points	Time	Points	Time	Points	Time	Points	Time
A E G procedure	5	88.4	5	88.6	5	87.3	5	88.3	5	86.4	5	86.9
A E G plus Mod 1	5	87.8	5	88.9	5	88.6	5	88.3	5	89.8	5	113.9
A E G plus Mod 2	5	22.6	5	44.5	5	46.0	5	45.8	5	44.5	5	44.7
A E G plus Mods 1 and 2	5	22.8	5	22.5	5	22.4	5	22.9	5	23.1	5	31.4
Method 1	5	.7	5	.7	5	.8	5	.8	5	.8	5	.7
Method 2	5	.6	5	.6	5	.6	5	.5	5	.5	5	.6
CIRCLE												
4000 points												
A E G procedure	88	5.5	40	11.1	29	14.9	14	32.6	10	42.4	5	97.1
A E G plus Mod 1	171	10.4	77	20.4	55	28.6	25	60.4	18	84.8	5	109.4
A E G plus Mod 2	88	5.4	40	10.8	29	15.4	14	32.7	10	41.6	5	92.2
A E G plus Mods 1 and 2	171	10.6	77	21.5	55	30.0	25	60.9	18	87.8	8	170.2
Method 1	127	25.1	56	10.4	39	7.5	18	3.7	13	2.7	6	1.0
Method 2	129	1.8	65	1.5	33	1.2	17	.9	17	1.0	5	.6

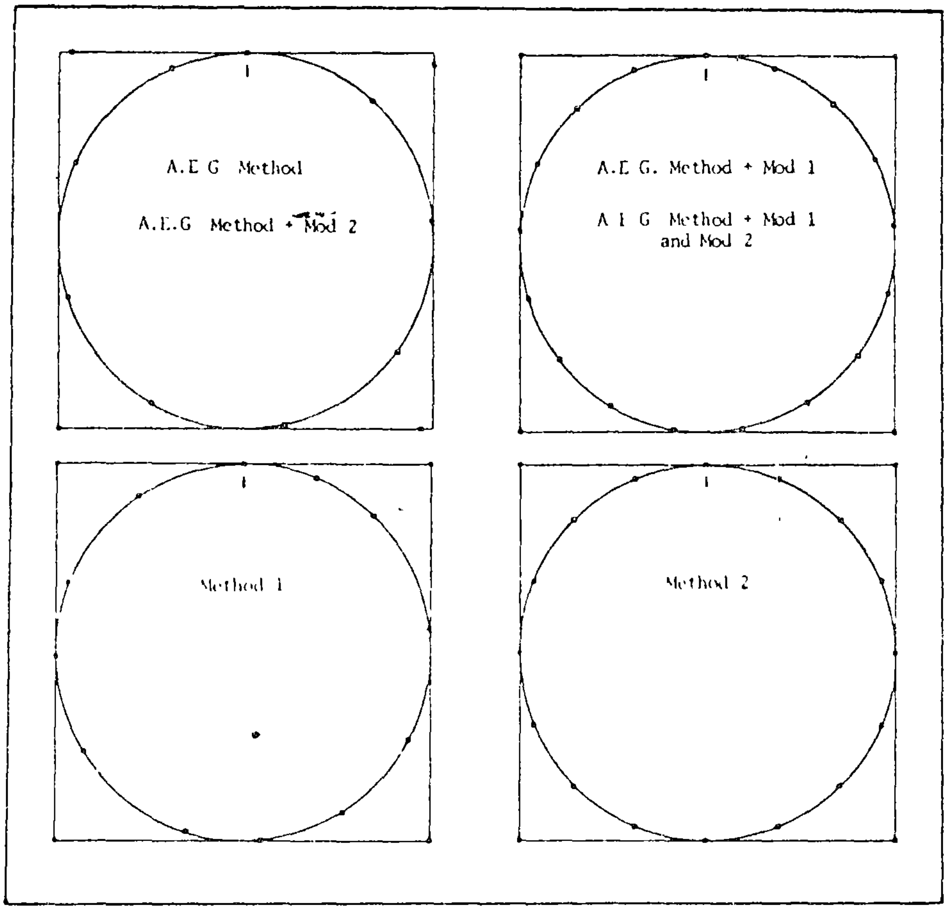


Figure 7 Plotted results for a circle and a square each made up of 4,000 points around its perimeter. The dots on the boundaries indicate the selected points by the indicated procedures. The tolerance value of .1 inch is illustrated to scale at the top centre of each diagram.

case where many points along a line are to be deleted, whereas the circle would be more representative in sinuosity of drawn or empirically recorded lines as far as this timing test was concerned. Table 1 presents results by the established offset tolerance in number of points selected and in seconds central processing unit time required for the reduction procedure (I.B.M. 370/155 under O/S MVT).

Figure 7 illustrates plotted results for a 1/10" offset tolerance with the points selected by each routine marked with a heavy dot. The fact that the routines selected different points and different numbers of points is not unusual or unexpected and similar differences occur in the case of sinuous lines. The shortcoming of the unmodified A.E.G. procedure is evident in the case of points selected to represent the square, which were just less than one tolerance unit from the corners for all but the first and last point.

Each routine selected five points to represent the square and each took approximately the same time regardless of the tolerance, except with the second modification. In this case the first step off the straight line caused the inner cycle to be invoked which found the new anchor point on the first iteration. More iterations were required for the other tolerance limits.

In the case of the circle an increase in tolerance limit caused a decrease in the number of points found to represent it for all methods Those methods which push the examination segment ahead of the anchor points, that is the A F G method with none, one or two modifications, take longer to perform as the offset tolerance is increased This therefore comprises the main reason that they have to be considered unsuitable in an operational context These procedures are fastest if they are unable to delete any points, because in such cases they would have to examine only one point to come to that decision On the other hand if a great number of points are found to be deletable, increasingly large inner cycles are invoked for each advance of the floating point The two methods presented in this paper work in entirely the opposite way and are fastest in the case of lines which are found to be representable with a smaller number of points Presumably this is the object of the effort In all cases Method 2 is seen to take as little as 1 per cent of the time required by the others.

The prime purpose of the routines discussed here is to reduce the number of points required to represent a line and to produce abstractions, or caricatures of the line in cases where these will suffice In many cases these could be considered to be perfectly adequate generalization procedures While the scope of generalization is no doubt much broader, line reduction by means such as those described here, represents an important portion of that topic

RÉSUMÉ Règle générale, les méthodes numeriques enregistrent des lignes avec beaucoup plus de données qu'il n'est necessaire à la reproduction graphique precise ou à la recherche par ordinateur. L auteur presente deux algorithmes pour reduire le nombre de points necessaires pour representer la ligne et produire des caricatures si desire, et les compare aux méthodes les plus prometteuses suggerées jusqu'ici. La reduction de la ligne constituera une partie importante de la généralisation automatique

ZUSAMMENFASSUNG Alle Digitalisierungsmethoden zeichnen in der Regel Linien mit bedeutend mehr Daten auf als fur eine genaue graphische Wiedergabe oder fur eine Computeranalyse notwendig sind Zwei spezielle Rechenverfahren zur Reduzierung der Punktezahl, die zur Darstellung einer Linie benotigt werden und die auch falls erwunscht Verzerrungen produzieren, werden vorgstellt und verglichen mit den bisher am meisten versprechenden Methoden Die Linienreduzierung wird eine grosse Rolle in der automatisierten Generalisierung spielen

RESUMEN Todos los métodos digitales, como regla general, registran líneas que tienen mucho más datos que los necesarios para la reproducción grafica correcta o para el analisis por computadora Se presentan dos algoritmos para reducir el numero de puntos necesarios para representar una línea y si se desea, producir caricaturas, estos se comparan con los métodos mas pometedores sugeridos hasta ahora La reduccion de líneas formara gran parte de la automatizacion en general

Research Notes and Comments

Choropleth Maps Without Class Intervals?

W. R. Tobler

It is now technologically feasible to produce virtually continuous shades of grey by using automatic map drawing equipment. It is therefore no longer necessary for the cartographer to "quantize" data by combining values into class intervals. As a simple illustration an automatic line plotter can be programmed to draw lines virtually any distance apart (Fig. 1). Thus, one can obtain any desired destiny of inked area to white area. For example, if the geographical data, symbolized by z, are normalized to lie in the range from zero to one, then an appropriate spacing of orthogonal lines of width w is given by

$$s = (w/z^x) \cdot [1 + (1 - z^x)^{1/2}].$$

Here an exponent ($x \approx 1.4$) of z has been chosen to approximate the nonlinear response of the human eye [*13*]. The units of the spacing s are those of w. Comparable equations are easily obtained for dashed lines or for dotted masks. Automatic equipment that produces grey areas by modulation of light intensities can produce even more refined displays. There thus results a choropleth map on which the visual intensity is exactly proportional to the data intensity. Since no class intervals have been introduced, there is no quantization error [*1, 2, 11*]. The much studied [*5, 6, 7, 8, 9, 10, 12*] and difficult problem of optimum class intervals is thus circumvented.

Some cartographers will still wish to group their data into classes and will argue that they do this in order to simplify or enhance the map for the user. This, then, is a problem of map generalization and not necessarily one of chosing class intervals. I assume that, by definition, a generalization of a choropleth map is another choropleth map, not a smooth surface as might be built up from modelling clay.

A choropleth map can be generalized in at least four ways. First, by combining

W. R. Tobler is professor of geography at the University of Michigan.

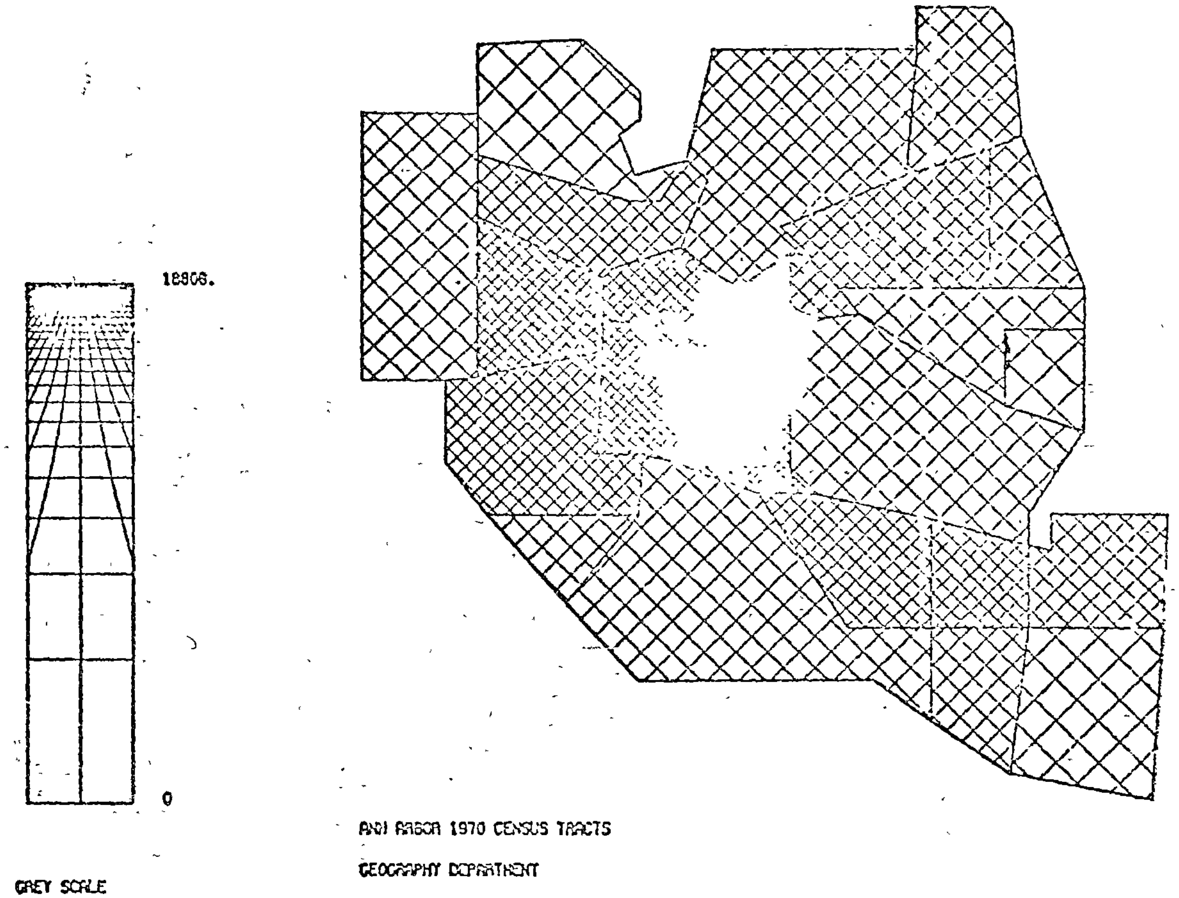


FIG. 1. Choropleth Map Without Class Intervals.

adjacent areal units (units that have similar values are made into new units whose value is some combination of the earlier values, or small units are eliminated, reducing the resolution of the data); secondly, by simplification of the boundaries of the areal units; thirdly, by changing the value of each unit in some manner which depends on the values of the adjacent units [*17*]; fourthly, by quantizing the data more coarsely, i.e., by picking large class intervals, or by using some nonlinear class intervals. As an analogy, one may consider the ways of generalizing a topographic surface: by varying the spacing of the sampling points, by smoothing with a filter, or by choosing a larger or variable contour interval. The latter method is of course comparable to the choosing of class intervals for a choropleth map. Enlarging or modifying the contour interval, without simplifying the contours, does not necessarily improve the map, but may enlarge the quantization error. Taking samples at larger or different spatial intervals is equivalent to filtering using a different two-dimensional Dirac comb [*3*] and thus is a type of smoothing. The more general case is to modify the values of each unit in a controlled manner that depends on the values of adjacent units [*14, 15*]. This is easily achieved by performing the choroplethic equivalent to taking a two-dimensional weighted moving average, as, for example, in binomial filtering [*4, 16, 17*]. Either smoothing or emphasis can be obtained in this manner.

The main argument in favor of using class intervals seems to be that their use enhances readability. This at least is the assertion. It seems equally plausible that this is also true of the three alternate map generalization methods cited above. If the assertion is in fact valid why then is grouping of greys into classes not also (e.g., in addition to spatial filtering) used to enhance aerial photographs, or television? Formulae for the optimal quantization of images are in fact given in the literature on picture processing, where the main difficulty stems from the conversion of continuous images into discrete signals, or relates to transmission band-width and noise reduction studies [*1, 2, 11*]. Typically, a large number (2^6) of levels are recommended, compared with the small (2^2 to 2^3) number used for choropleth maps, though somewhat fewer levels are required for equally satisfactory colored pictures. It is thus not clear why the theory for pictures should differ from the theory for choropleth maps, since both have visual information processing as their ultimate objective. Presumably, both have some domain of validity, but the limits need further exploration.

LITERATURE CITED

1. ANDREWS, H. C. *Computer Techniques in Image Processing*. New York: Academic Press, 1970. Pp. 136-51.
2. DIGIUSEPPE, J. L. *A Survey of Pictorial Data Compression Techniques*. Technical Report No. 16, Contract N. DA-49-083 OSA-3050, University of Michigan, Ann Arbor, University of Michigan, 1969.
3. GOODMAN, J. W. *Introduction to Fourier Optics*. New York: McGraw-Hill, 1968. Pp. 14–25.
4. HOLLOWAY, J. L. "Smoothing and Filtering of Time Series and Space Fields." *Advances in Geophysics*, 4 (1958), pp. 351–89.

5. JENKS, G. F. "Generalization in Statistical Mapping," *Annals*, Association of American Geographers, 53 (1963), pp. 15–26.
6. JENKS, G. F. and M. R. COULSON. "Class Intervals for Statistical Maps," *International Yearbook for Cartography*, 3 (1963), pp. 119–34.
7. JENKS, G. F. and F. C. CASPALL. "Error on Choroplethic Maps," *Annals*, Association of American Geographers, 61 (1971), pp. 217–44.
8. MACKAY, J. R. "An Analysis of Isopleth and Choropleth Class Intervals," *Economic Geography*, 31 (1955), pp. 71–81.
9. MONMONIER, M. S. 'Contiguity-Biased Class-Interval Selection," *Geographical Review*, 62 (1972), pp. 203–28
10. ROBINSON, A. H. and R. D. SALE. *Elements of Cartography*. 3rd ed. New York: Wiley, 1968. Pp. 164–70.
11. ROSENFELD, A. *Picture Processing by Computer*. New York: Academic Press, 1969. Pp. 20–28.
12. SCRIPTER, M. W. "Nested-Means Map Classes for Statistical Maps," *Annals*, Association of American Geographers, 60 (1970), pp. 385–93
13. STOLSSEL, O. C. "Standard Printing Screen System," *Proceedings*, 1972 Fall Convention, American Congress on Surveying and Mapping, Columbus, Ohio, 1972, pp. 111–49.
14. TOBLER, W. R. "Numerical Map Generalization," *Michigan Inter-University Community of Mathematical Geographers Paper No. 11*, Ann Arbor, University Microfilms No. OP-33067, 1966.
15. ———. "Of Maps and Matrices," *Journal of Regional Science*, 7 (1967), pp 275–80.
16. ———. "Geographical Filters and Their Inverses," *Geographical Analysis*, 1 (1969), pp. 234–53.
17. ———. "CHOROS A Computer Program to Apply Linear Neighborhood Operators to Choropleth Maps" Ann Arbor, April, 1972 Available from the Geography Program Exchange, Computer Institute, Michigan State University, East Lansing, Michigan 48823, U.S.A.

EL PROGRAMA SYMVU.

Introducción.

SYMVU es un programa de representación gráfica mediante computadora, escrito con el propósito de generar un dibujo tridimensional de los datos. El programa SYMVU puede ser usado por personas con muy poca experiencia en programación. Solamente son necesarias tres tarjetas de control para la generación de la gráfica. Treinta y dos electivas u opciones están construídas dentro del programa, permitiendo una considerable flexibilidad en la generación de la muestra de los datos. Sin embargo, solamente siete de estas opciones son absolutamente necesarias para la producción de una simple muestra.

El programa SYMVU es usado comunmente para mapas geográficos cuantitativos. A diferencia de los mapas de contorno, SYMVU ilustra los valores absolutos del contínuo espacial de los datos. Los mapas de contorno (ó mapas de isolíneas) usan intervalos redondeados para graficar la muestra de la información cuantitativa. El programa SYMVU tiene también la capacidad de conformar y mapear usando datos generados por el programa SYMAP.

Uno de los hechos distintivos de este programa, es que complementa la tarea de decidir que partes del objeto son vistas y cuales quedan ocultas a la vista.

El programa está escrito en FORTRAN IV y es operado en una máquina IBM-360 usando una memoria de 220K.

Las corridas han sido complementadas exitosamente en tanto en un graficador de 11 pulgadas como en uno de 30 pulgadas, pero debido a restricciones de tamaño en el programa, un graficador de 30 pulgadas presenta una pequeña ventaja.

6.1.1. Principios Básicos.

El programa utiliza datos que están en forma de matriz. Esta matriz información puede ser generada en una o dos formas.

La primera y más común de estas alternativas es utilizar los datos generados por el programa SYMAP como una gráfica bidimensional. El SYMAP interpola datos de localización de área, tal que, produce superficies espaciales contínuas organizadas en forma de matriz. Estos datos pueden entonces ser utilizados por el programa SYMVU para producir gráficas tri dimensionales.

La segunda alternativa es proveer datos directamente en forma matri cial, en tarjetas o en cinta.

6.1.2. Requerimentos de entrada.

Para obtener una muestra gráfica, el usuario debe proveer tres tar jetas primarias de control y tiene la opción de proveer otros tres conjun tos de tarjetas de control y usar una subrutina. Las tarjetas del control requeridas son:

i) Tarjeta de título 1. La tarjeta 1 es la tarjeta de título. Se tiene la flexibilidad de escribir hasta 72 caracteres describiendo el títu lo de la gráfica particular que se está generando. Es importante que este título sea claramente descriptivo para diferenciar el mapa particular que es requerido, de todos los otros mapas de naturaleza similar que han corri do previamente o que pueden ser corridos en el futuro.

ii) Tarjeta electiva 2. Esta tarjeta tiene las variables primarias que especifican el número de renglones y columnas de la matriz y el tipo de vista. También tiene un número de otras variables concernientes a datos de entrada, símbolos, escalas, y el área fuera del área de estudio.

Las electivas 2-1, 2-2, y 2-4 deben de ser especificadas para el pro grama que va a correr.

iii) Tarjeta electiva 3. Esta tarjeta tiene las variables primarias que controlan el tamaño y los ángulos visuales de una gráfica. También tie ne un número de otras variables concernientes a los valores máximo y mínimo contenidos, distancias visuales para vistas no-ortogonales, tamaños de paso, tamaño de símbolo y tamaño de bloc. Las electivas 3-1,3-2, 3-3 y 3-4 deben ser especificadas para el programa que va a correr.

Cada tarjeta tiene un formato diferente de especificación. La primera tarjeta es completamente flexible usando el formato "A". La segunda tarjeta requiere un formato "I" (sin punto decimal). La tercera tarjeta requiere un formato "F" (requiere punto decimal). El contenido y especificaciones para cada electiva están descritos en las siguientes páginas.

Las tarjetas de control opcional son usadas si la matriz de datos está en tarjetas o si los símbolos o letreros se desean en la gráfica.

La Subrutina Data es usada para manipular los valores que son graficados o para leer en una matriz dada en tarjetas.

Si la electiva 21 del programa SYMAP fué usada para generar una cinta conteniendo la matriz de datos, entonces las electivas 2-12 y 2-18 deben ser especificadas para el programa a correr.

6.2 Tarjeta electiva 2

La tarjeta 2 tiene 22 electivas. Todas las variables sobre esta tarjeta son especificadas con un formato "I" (sin punto decimal) y deberán ser perforadas justificadas a la derecha.

Las electivas 2-1, 2-2 y 2-4 deben ser especificadas para el programa a correr.

6.2.1- Electiva 2-1: Rows.

Esta electiva es usada para especificar el número de renglones (excluyendo el borde) en la matriz de datos. El número máximo de renglones es 130. Especificar el número de renglones en las columnas 1 a la 4 ajustado a la derecha. Esta electiva debe ser especificada o el programa no funcionará correctamente.

6.2.2. Electiva 2-2: Columns.

Esta electiva es usada para especificar el número de columnas (excluyendo el borde) en la matriz de datos. El máximo de columnas es 130. Especificar el número de columnas en las columnas 5 a la 8 ajustado a la derecha. Esta electiva debe ser especificada o el programa no funcionará correctamente.

Si un número incorrecto de renglones y/o columnas ha entrado, puede ocurrir lo siguiente:

i) Si el número es menor que el número correcto en fila, solamente la parte especificada del mapa podrá ser graficada.

ii) Si el número es más grande que el correcto, el programa terminará.

6.2.3. Electiva 2-3: View.

Esta electiva es usada para especificar el tipo de proyección que es deseada. El programa SYMVU es capaz de dibujar proyecciones, isométricas, perspectivas de dos puntos, y proyecciones ptanométricas.

La proyección isométrica es generalmente usada para datos geográficos. La perspectiva de dos puntos permite dramatizar la visualización, particularmente para superficies topograficas de paisaje. La vista ptanométrica preserva ángulos derechos y además la forma verdadera del mapa.

Para especificar una proyección isométrica, se pone un 1 en la columna 12. Si se desea una perspectiva de dos puntos, se pone un 2 en la columna 12. Si se desea una proyección planométrica se pone un 3 en la columna 12.

Si la especificación no es hecha, el programa supone que se desea una proyección isométrica.

6.2.4. Electiva 2-4: Line Type.

Esta electiva es usada para especificar la dirección de las líneas para ser dibujadas. El programa tiene la capacidad de graficar las líneas a lo largo de tres diferentes ejes: a lo largo de las columnas, a lo largo de los renglones, o sobre la diagonal de la matriz.

Generalmente, la opción de la diagonal es especificada porque esta permite el máximo de líneas para ser graficadas. Sin embargo, la vista aparece mucho mejor si las líneas son dibujadas a lo largo de los renglones o de las columnas. Normalmente, cualquier ángulo escogido entre 330° y 30° o 150° y 210° deberá de ser dibujado a lo largo de los renglones; cuando el ángulo está entre 60° y 120° o entre 210° y 300° las líneas deberán de ser dibujadas a lo largo de las columnas. Todos los otros ángulos deberán utilizar la opción diagonal.

Si se desea graficar a lo largo de las columnas (el número azimuthal es 90° o 270°) especificar un 1 en la columna 16. Si las líneas son deseadas a lo largo de los renglones (el azimuthal es 0° o 180°) especificar un 2 en la columna 16. Si se desean líneas a lo largo de las diagonales, especificar un 4 ó un 5 en la columna 16. Esta electiva debe ser especificada o el programa no funcionará correctamente.

6.2.5. Electiva 2-5: Smooth.

Esta electiva es usada para especificar el número de suavizaciones binomiales de los datos deseados. Esta opción es raramente usada; es de alguna significación para retratar gráficamente una superficie topográfica basada en observaciones de puntos discretos, que algunas veces exhiben grandes variaciones locales. Especificar el número de suavizaciones deseadas en las columnas 17 a la 20. Cualquier número de suavizaciones puede ser especificado, pero eventualmente la superficie será como un plano.

6.2.6. Electiva 2-6: Pepeat.

Esta electiva es usada para especificar si los datos de la gráfica previa están retenidos, o si un nuevo conjunto de datos esta siendo leido.

El programa puede usar los mismos datos para un número de gráficas con diferentes angulos y elevaciones. Para leer en un nuevo conjunto de datos, especificar un cero en la columna 24: para usar un conjunto de datos previo, especificar un 1 en la columna 24. Si la especificación no es hecha, el para coscurecer el área de no-estudio con simbolismo oscuro, especificar 3 en la columna 32. Si la especificación no es hecha, el programa supone que el área de no-estudio deberá de ser graficada con líneas cero.

6.2.9. Electiva 2-9: Sqrtt.

Esta electiva es usada para especificar cuales de las raices cuadradas de todos los valores de la matriz de datos deberán de ser calculadas y usadas como valores para ser graficados. Esto tiene el efecto de reducir la importancia relativa de valores extremos. Esta electiva es usada para un gran rango de datos, donde la suavización es posible no solamente para los valores más altos sino también para los valores más bajos. (Por ejemplo, si se está estudiando la densidad de población en una región circundante a la ciudad de New York, la cumbre de la población en Manhattan tendrá el efecto de no tomar en cuenta cualquier detalle ocurrido en las áreas circundantes de la ciudad. Por tomar las raices cuadradas de los datos, las variaciones de las superficies circundantes pueden aparecer mas distinguibles).

Para tomar la raíz cuadrada de todos los valores, especificar un 1 en la columna 36.

6.2.10. Electiva 2-10: Samen.

Esta electiva es usada para especificar el mismo factor de escala vertical de la gráfica precedente. Muchas veces es empleada cuando un número de gráficas están siendo hechas por datos en series de tiempo y se desea tener mapas comparativos en escala vertical. Para especificar que el mismo factor de escala "Z" de la gráfica precedente deberá de ser usado, coloque un 1 en la columna 40.

6.2.11. Electiva 2-11: Endlin.

Esta electiva es usada para especificar que los extremos de las líneas que circunscriben la gráfica son suprimidos. Para suprimir los extremos de las líneas, especificar 1 an la columna 44.

para cosourecer el área de no-estudio con simbolismo oscuro, especificar 3 en la columna 32. Si la especificación no es hecha, el programa supone que el área de no-estudio deberá de ser graficada con líneas cero.

6.2.9. Electiva 2-9: Sqrtt.

Esta electiva es usada para especificar cuales de las raices cuadradas de todos los valores de la matriz de datos deberán de ser calculadas y usadas como valores para ser graficados. Esto tiene el efecto de reducir la importancia relativa de valores extremos. Esta electiva es usada para un gran rango de datos, donde la suavización es posible no solamente para los valores más altos sino también para los valores más bajos. (Por ejemplo, si se está estudiando la densidad de población en una región circundante a la ciudad de New York, la cumbre de la población en Manhattan tendrá el efecto de no tomar en cuenta cualquier detalle ocurrido en las áreas circundantes de la ciudad. Por tomar las raices cuadradas de los datos, las variaciones de las superficies circundantes pueden aparecer mas distinguibles).

Para tomar la raíz cuadrada de todos los valores, especificar un 1 en la columna 36.

6.2.10. Electiva 2-10: Samen.

Esta electiva es usada para especificar el mismo factor de escala vertical de la gráfica precedente. Muchas veces es empleada cuando un número de gráficas están siendo hechas por datos en series de tiempo y se desea tener mapas comparativos en escala vertical. Para especificar que el mismo factor de escala "Z" de la gráfica precedente deberá de ser usado, coloque un 1 en la columna 40.

6.2.11. Electiva 2-11: Endlin.

Esta electiva es usada para especificar que los extremos de las líneas que circunscriben la gráfica son suprimidos. Para suprimir los extremos de las líneas, especificar 1 an la columna 44.

6.2.12. Electiva 2-12: Symap.

Esta electiva es usada para especificar el número de líneas verticales por pulgada contenidas en la matriz de entrada.

Si la matriz de datos fué producida por el programa SYMAP, la salida es normalmente hecha en una malla de 8 por 10; hay 8 líneas verticales por pulgada. El módulo del programa SYMAP está basado sobre dos dimensiones de los caracte - res simbólicos comunes en más líneas impresas en el dibujo. SYMVU tiene la capacidad de leer datos que tienen una dimensión vertical de 6 líneas o 10 líneas por pulgada. Esta flexibilidad permite tipos de entrada diferentes al del programa SYMAP. Si el programa GRID es usado, una celda de igual tamaño es el módulo de la matriz. En este caso las escalas para las dimensiones horizontal y vertical de la matriz deberá ser especificada por SYMVU.

Para especificar el número de líneas por pulgada, poner un 6 o un 8 o 10 ajustados a la derecha en las columnas 45 a la 48. Si la especificación no es hecha el programa supone escalas iguales para las dimensiones horizontal y vertical de entrada.

6.2.13. Electiva 2-13: Leg.

Esta electiva es usada para especificar la supresión del control de información normalmente escrito sobre una gráfica SYMVU. Esta información incluye los ángulos de elevación, el ángulo azimuthal, y la información de escala.

Para borrar esta información de la gráfica, especificar un 1 en la columna 50.

6.2.14. Electiva 2-14: Nscal.

Esta electiva es usada para especificar que una escala de alturas deberá aparecer a la derecha de la gráfica. Esta escala tiene medidas en pulgadas en el lado izquierdo y el valor "Z" correspondiente en el lado derecho. En el tope de la escala, la altura más grande y el valor "Z" correspondiente son mostrados. Para que esta escala de alturas aparezca, especifique un 1 en la columna 52.

6.2.15. Electiva 2-15: Flp.

Esta electiva es usada para especificar que se desea un plano flotante como parte de la gráfica. Para incluir este plano, especificar un 1 en la columna 55.

6.2.16. Electiva 2-16: Base.

Esta electiva es usada para especificar si una base, sobre la que una superficie topográfica puede descansar, debe ser construida o suprimida. Cuando los datos para este tipo de superficie están en una matriz rectangular (sin perfiles irregulares), la línea de fondo de la base es determinada por el valor más bajo graficado en la superficie. Cuando un perfil es usado dentro de la matriz (los valores cero son graficados), el programa determina las dimensiones verticales apropiadas para la base. Esta base puede ayudar en la identificación de las relaciones geográficas del área en estudio.

Para suprimir la base, especifique un 1 en la columna 60. Para incluir una base completa debajo de un perfil irregular, especificar un 2 en la columna 60. Para colocar la superficie sobre un bloc, especificar un 3 en la columna 60. Use la electiva 3-10: Size, para especificar la altura de este bloc. Si la especificación no es hecha, el programa supone una base irregular debajo de un perfil irregular; cuando una matriz rectangular de valores es graficada, constituirá una base sobre la cual la superficie total puede descansar. En este caso, la base está determinada por el valor más bajo graficado.

6.2.17. Electiva 2-17: Petdot.

Esta electiva es usada para especificar el número de símbolos que serán localizados en la gráfica. El programa SYMVU tiene la capacidad de graficar ciertos símbolos, geográficamente sobre la superficie; normalmente el programa es limitado en vocabulario a 5 tipos de símbolos. (Se puede alterar levemente para cambiar o manipular el tamaño y el tipo de símbolo, como esta discutido en la electiva 3-9).

Los símbolos son muchas veces usados para localizar ciertos fenómenos, tal como, límites políticos, concentraciónes de población u otros hechos geográficos. Se usan también, para dibujar letreros, tal como, flechas de Norte. Esta electiva especifica el número de símbolos o puntos de la gráfica que el programa está leyendo. Las coordenadas o localizaciones geográficas por renglón y columna, tal como el símbolo actual para cada símbolo, están especificados para entradas de series de datos en tarjetas, siguiendo a las tres tarjetas principales de control. Esto está discutido en la sección sobre las tarjetas opcionales de datos.

Especificar el número de símbolos, para ser usados, en las columnas 61 a la 64, ajustado a la derecha.

Si la especificación no es hecha, el programa supone que no hay símbolos para ser graficados.

6.2.18. Electiva 2-18: Fdata.

Esta electiva es usada para especificar el tipo de datos de entrada y el número de fila. El SYMVU usa datos del SYMAP y otros datos, hasta una subrutina escrita por el usuario.

Normalmente se crea una cinta por la electiva 21 del SYMAP que contiene la matriz de datos para ser graficada. Un número de matrices de datos son muchas veces creadas sobre la misma cinta, con cada matriz especificada sobre una fila separada. La electiva 2-18 del programa SYMVU es usada para especificar la fila sobre la cinta que está siendo graficada.

Si la matriz de datos es reemplazada por una subrutina hecha por el usuario, especificar un cero ajustado a la derecha en las columnas 65 a la 68. Si la matriz de datos está sobre una unidad de salida del SYMAP, especificar el número de fila ajustado a la derecha en las columnas 65 a la 68. Si la especificación no es hecha, el programa supone que la matriz de datos será usada por una suorutina del usuario.

6.2.19. Electiva 2-19: Namin.

Esta electiva es usada para especificar cuales de los valores que caen abajo de un mínimo predeterminado, deberán de formar un conjunto igual al valor mínimo. La electiva 3-5 es usada para colocar el valor mínimo usado por el programa.

Especificar un 1 en la columna 72 para colocar los valores abajo de un mínimo predeterminado en este mínimo. La electiva 3-5, será ignorada si es usada sin esta electiva. Si la especificación no es hecha el programa supone que los valores actuales deberán de ser graficados.

6.2.20. Electiva 2-20: Namax.

Esta electiva es usada para especificar cuales de los valores que caen arriba de un máximo predeterminado, deberán de formar un conjunto igual al valor máximo. La electiva 3-6 es usada para colocar el valor máximo usado por el programa.

Especificar un 1 en la columna 76 para colocar los valores arriba de un valor máximo predeterminado en este valor. Especificar un 2 en la columna 76 si los valores arriba de la especificación máxima no son cambiados, pero el valor máximo especificado en la electiva 3-6 es usado en conjunción con la altura en la electiva 3-4. La electiva 3-6 será ignorada si está usada sin esta electiva. Si la especificación no es hecha el programa supone que los valores actuales deberán de ser graficados.

Cuando los valores extremos son tan grandes como para minimizar la variación en el resto de los, las electivas 2-19 y 2-20 se deberán de usar.

6.2.21. Electiva 2.21: Iform.

Esta electiva es usada para especificar un valor que puede estar en la subrutina Data, para identificar el conjunto de datos apropiado para ser manipulado. Especificar un número entero (de 1 a 99) en las columnas 77 a la 78. Si la especificación no es hecha, el programa supone que un conjunto de datos está siendo manipulado en la Subrutina Data.

6.2.22. Electiva 2.22: Nleg.

Esta electiva es usada para especificar el número de letreros de texto para ser graficados en cada SYMVU. El número máximo es 50. Los letreros de texto son discutidos en la sección sobre tarjetas de datos opcionales.

Especificar el número de letreros para ser graficados, en las columnas 79 a la 80, ajustado a la derecha. Si la especificación no es hecha el programa supone que no hay letreros para ser graficados.

6.3. Tarjeta electiva 3.

La tarjeta 3 tiene 10 electivas. Todas las variables en esta tarjeta están especificadas con un tipo de formato "F" (requieren un punto decimal en todos los campos de especificación). Deberán de ser perforados justificados a la derecha; esto no es obligatorio como lo es con el formato " I" sobre la tar - jeta 2.

Las electivas 3-1, 3-2, 3-3 y 3-4 deberán de ser especificadas para que el programa corra.

6.3.1. Electiva 3-1: Altitude.

Esta electiva es usada para especificar la elevación del punto de vista sobre el plano horizontal. Especificar el punto de vista en grados angulares en las columnas 1 a la 5. Esta electiva deberá de ser especificada o el pro- grama no funcionará correctamente.

Para propósitos generales, una buena altura para empezar la experimenta ción es 45 grados. Para más casos, esto dá una vista optima de la superficie, tal que, se da una buena relación entre las líneas y el espacio en blanco que las separa. De cero a cinco grados la vista de una superficie topográfica del paisaje, normalmente está completamente cerrada al nivel visual del ojo, mientras que una vista cercana a 45 grados aparece al observador como una vista de ojo de pájaro.

Para más propositos de variación visual en la superficie topográfica, es primariamente una función del espacio en blanco entre las líneas trazadas. Porque la sola variación en la dirección de la línea ocurre en la dimensión vertical, la variación máxima en una línea ocurre en la altura de cero grados; generalmente, este ángulo esconde mucho el área de estudio por el factor de línea escondida en el programa.

6.3.2. Electiva 3-2: Azimuth.

Esta electiva es usada para especificar el ángulo horizontal deseado de la dirección de la vista. Especificar el ángulo de vista azimuthal en grados en las columnas 6 a la 10, ajustado a la derecha. Esta electiva debe de ser especifi - cada o el programa no funcionará correctamente.

El programa SYMVU puede girar cualquier matriz de datos 360°, cuando se está usando un sistema de orientación azimuthal. Cuando se especifica un azimuth cero, se está viendo perpendicularmente al fondo del área en estudio. Como se gira la posición de vista hacia la izquierda, se está incrementando hacia 360°. A 90° se mirará perpendicularmente al lado izquierdo la matriz. A 270° se mirará el lado derecho de la matriz.

6.3.3. Electiva 3-3: Width.

Esta electiva es usada para especificar el ancho, en pulgadas, que la matriz de entrada tendrá en la gráfica. Hay dos escalas variables en el programa SYMVU: el ancho y la altura. El ancho se refiere a la distancia a lo largo del fondo de la matriz (las columnas de la matriz de entrada). Bajo situaciones normales, la dimensión del ancho y la dimensión de la altura, cuando se combinan, no deberá exceder a 11 pulgadas.

Especificar el ancho en pulgadas, en las columnas 11 a la 15, ajustado a la derecha. Esta electiva deberá especificarse o el programa no funcionará correctamente.

6.3.4. Electiva 3-4: Height.

Esta electiva es usada para espacificar la altura, en pulgadas, que la matriz de entrada tendrá en la gráfica. Especificar la altura en pulgadas, en las columnas 16 a la 25, ajustada a la derecha. Si se especifica 0.0, la altura será colocada a 3.0 pulgadas. Cualquier altura más grande que 0.0 puede ser especificada, pero la combinación del ancho y la altura no deberá exceder a 11 pulgadas. Esta electiva deberá de ser especificada o el programa no funcionará correctamente.

6.3.5. Electiva 3-5: Amin.

Esta electiva es usada para especificar el valor mínimo que el programa usará para graficar. El usuario normalmente especifica esta electiva, cuando un valor mínimo extremo ocurre dentro de la matriz de datos.

Especificar el valor mínimo deseado en las columnas 26 a la 35, ajustado a la derecha. Si no es hecha la especificación, el programa supone que el valor mínimo actual será usado para graficar. Esta electiva es ignorada a menos que la electiva 2-19 sea también usada.

6.3.6. Electiva 3-6: Amax.

Esta electiva es usada para especificar el valor máximo que el programa usará para ser graficado. Cuando valores de altura extremos ocurren dentro de la matriz de datos, esta electiva es usada para controlar o truncar estos valores. Especificar el valor máximo deseado, en las columnas 36 a la 45, ajustado a la derecha. Si la especificación no es hecha el programa supone que el valor máximo actual se usará para ser graficado. Esta electiva es ignorada, a menos que la electiva 2-20 sea también usada.

En conexión con la electiva 2-20 y la electiva 3-4, esta electiva puede ser usada para colocar la altura igual a algunos valores especificados (el valor máximo), pero permite que el valor máximo actual sea usado para graficar. Esto es particularmente usado para un conjunto de gráficas donde se desea una escala uniforme.

6.3.7. Electiva 3-7: Vdist.

Esta electiva es usada para especificar la distancia de vista para los dibujos en perspectiva. La opción controla la distancia de la que se observa una perspectiva dada. Como se mueve hacia atrás de la ilustración (en otras palabras, incrementa su distancia de vista) los extremos o las esquinas del dibujo son usualmente extendidos hacia afuera, más lejos relativamente de la línea del centro. Con una distancia visual pequeña, el foco está primariamente en el centro de la matriz de datos. Esta variación es vista más rápidamente cuando el ángulo de altitud es muy bajo, tal que, cuando la altitud es cercana el nivel visual del ojo, sobre un dibujo de paisaje natural. Cuando la altitud es muy alta, tal como 45° o 50°, la variación causada por la distancia de vista es relativamente pequeña.

Especificar la distancia de vista, en pulgadas, en las columnas 46 a la 50, ajustada a la derecha. Si la especificación no es hecha el programa supone una distancia de vista de 10.0 pulgadas; esta opción es usada solamente en vistas de perspectiva.

6.3.8. Electiva 3-8: Step.

Esta electiva es usada para especificar el tamaño de paso no-estandar del graficador. Esta electiva es raramente usada, pero puede ser empleada para reducir el tiempo calculado, con una pérdida proporcional en detallar y suavizar los segmentos de línea. Especificar el tamaño del paso, en pulgadas, en las columnas 51 a la 55, ajustado a la derecha. Si la especificación no es hecha, el programa supone un tamaño de paso estandar de 0.05 pulgadas.

6.3.9. Electiva 3-9: Sepsm.

Esta electiva es usada para especificar el tamaño no-estandar de los símbolos (puntos gráfica o letreros) usados en el dibujo. Especificar el tamaño actual, en pulgadas, ajustado a la derecha, en las columnas 56 a la 60. Para propósitos normales, el factor de escala estandar es adecuado. Con excepción del símbolo número 14, que es el más largo, el tamaño estandar de símbolo es 0.25 pulgadas. Ocacionalmente, para propósitos especializados, el usuario puede encontrar una ventaja al variar el tamaño de Símbolo con el que está trabajando. Si la especificación no es hecha, el programa supone un tamaño estandar para los símbolos.

6.3.10. Electiva 3-10: Size.

Esta electiva es usada para especificar la altura de un bloc colocado bajo el área de estudio. Esta electiva es usada junto con la electiva 2-16 (Base) cuando BASE=3. Especificar, en las columnas 76 a la 80, ajustado a la derecha, la altura que se quiera del bloc, en pulgadas, hasta un máximo de 0175 pulgadas. Si la especificación no es hecha, el programa supone una altura estandar de 0.5 pulgadas, cuando BASE=3 en la electiva 2-16.

6.4. Tarjetas opcionales de datos.

Hay un número de tarjetas opcionales de datos, que pueden ser especificadas para el programa SYMVU:

i) Si la matriz de datos esta leyendose en tarjetas en vez de cinta, se necesitan, para cada renglón de la matriz, una o más tarjetas conteniendo los valores para las columnas en el renglón:

ii) Para que cada símbolo sea graficado, se necesita una tarjeta conteniendo la localización y el tipo de símbolo;

iii) Para que cada letrero sea graficado, se necesita una tarjeta conteniendo la localización y el texto del letrero.

Estas tarjetas siguen directamente después de la tarjeta electiva 3 y deben estar en el orden antes dicho.

6.4.1. Matriz de datos en tarjetas.

Si la matriz de datos es suministrada a la computadora, en tarjetas en vez de cinta, la tarjeta disponible para la matriz de enteros sigue inmediatamente después de la tercera tarjeta de control. Normalmente, el programa usa entradas en cinta.

Las tarjetas son leidas bajo el formato especificado en la Subrutina

Data. Habrá un registro - en una o más tarjetas - para cada renglón, secuencialmente, de la cima al fondo de la matriz; cada registro contendrá un valor para cada columna, secuencialmente, de la izquierda a la derecha de la matriz.

6.4.2. Puntos Símbolos.

Para cada símbolo que va a ser graficado, el renglón obligarotio, la columna obligatoria y el símbolo 'type' debe ser especificado en una tarjeta. Las coordinadas obligatorias deben estar en términos de la retícula de entrada de SYMVU. Si el SYMAP es usado para crear la cinta de entrada del SYMVU, estos renglones y columnas obligatorias, deberán de ser copiados de la tabla inmediatamente precedente al mapa producido por SYMAP, para graficar los puntos dato.

Especificar el renglón obligatorio, ajustado a la derecha, en las columnas 1 a la 10, usando un formato "I" (sin punto decimal), y las columnas obligatorias, en las columnas 11 a la 20, también ajustado a la derecha, usando un formato "I". El tipo de símbolo deberá especificarse en las columnas 29 a la 30, ajustado a la derecha, usando un formato I. Si las columnas 29 a la 30 son dejadas en blanco, el símbolo omiso será el 14.

El número total de puntos de la gráfica debe ser especificado en la electiva 2-17.

6.4.3. Letreros.

Para que cada letrero sea graficado, el renglón obligatorio, la columna obligatoria y el texto del letrero actual, deben de ser especificados en una tarjeta. Las coordenadas obligatorias deben estar en términos de la retícula de entrada del SYMVU. Si el SYMAP es usado para crear la cinta de entrada del SYMVU, estos renglones y columnas obligatorias, deben de ser tomados directamente del mapa producido por SYMAP.

Especificar el renglón obligatorio, ajustado a la derecha, en las columnas 1 a la 10, usando un formato "I" (sin punto decimal), y la columna obligatoria en las columnas 11 a la 20, también ajustado a la derecha, usando un formato "I". Las columnas 30 a la 54 son usadas para el texto del letrero; sin embargo, este deberá estar ajustado a la izquierda, tal que, el primer caracter del letrero estará en la columna especificada en las columnas 11 a la 20.

El número total de letreros debe estar especificado en la electiva 2-22.

6.5. Subrutina Data.

Esta Subrutina Fortran puede ser usada para manipular los valores Z y regresar los datos al programa, en la forma en que SYMVU los aceptará. La Subrutina Data es también usada para leer datos de una matriz de datos en tarjetas.

```
      SUBROUTINE  DATA (X, IR, IC, IFORM)
      DIMENSION X (130)
      READ (5,1000) (X (I), I=1, 130)
 1000 FORMAT (40F2. 0/40F2.0/40F2.0/10F2.0)
      DO 101 = 1,130
   10 X (I) = X (I)* 2-1
      RETURN
      END
```

Para el ejemplo mostrado:

— X (130) regresa los valores para un renglón de la matriz, a un tiempo, al programa.

— IR es el número de renglones para ser leidos y regresados.

— IC es el número de columnas (o valores Z) por renglon, leidos por la Electiva 2-2

— IFORM es el valor suplido por el usuario que identifica el conjunto de datos, para ser manipulados por un conjunto particular de proposiciones en la Subrutina DATA. (En este caso, solamente un conjunto de datos es manipulado y IFORM no es usado). IFORM es leida por la electiva 2-21.

6.6. Muestra de Sumisiones.

Hay pocas cosas que guardar en mente cuando se producen más de un SYMVU en una simple corrida.

i) Cualquier vista de la misma superficie, después de la primera vista, puede ser requerida meramente por incluir las tres tarjetas de control inme - diatamente después de las tarjetas de control (y después cualquier tarjeta opcional de datos) para la vista previa.

ii) Cuando se requiere una vista de una superficie, si es la primera, segunda, etc, superficie en sumisión, todas las tarjetas opcionales de control deben ser re especificadas con la excepción de la matriz de datos. En otras palabras, el punto símbolo y los letreros no son llevados de una superficie a la otra, son llevados de una vista de una superficie a la próxima vista de la misma superficie.

iiii) En suma, para las electivas requeridas, otras dos electivas generalmente cambian de vista a vista y de superficie a superficie y deberán de ser revisadas cuidadosamente la Electiva 2-6: Repeat y la Electiva 2-18: Fdata. La electiva 2-12: SYMAP y la Electiva 2-21: IFORM especifican si la entrada está en tarjetas, cinta o disco - usualmente sobre una fila producida por SYMAP. Finalmente la Electiva 2-17: Pltdot y la Electiva 2-22 : Nlog, son usadas cuando los puntos símbolos y los letreros son comprendidos.

iv) Estos principios son resumidos por una sumisión hipotética donde:

— La primera superficie está en disco, dos vistas son deseadas, y no hay símbolos o letreros.

— La segunda superficie está en tarjetas (la subrutina Data deberá usarse para leer los datos), se desean dos vistas, y hay letreros pero no hay símbolos.

— La tercera superficie esta también en disco (deberá estar en una fila que ocurra después de la fila que contiene la primera superficie, remanejado en la mitad de la corrida no es posible aquí), se desean tres vistas, los mismos letreros son usados y los puntos símbolo son añadidos.

6.7. Información técnica.

Esta sección tiene la intención de proveer la información técnica a casos particulares que desean ser implementados por el programa SYMVU - I y su instalación computacional. Tal que, esta sección supone un conocimiento de los sistemas de procesamiento de su centro computacional.

Los ejemplos del lenguaje de trabajo de control fué programado para una IBM 360/65 OS MVT liberando 18 sistemas. SYMVU será computado sucesivamente en FORTRAN IV G y H.

6.7.1. Descripción de la cinta fuente del SYMVU.

La cinta del SYMVU es una cinta de nueve canales escritos en 800 BPI con paridad non.

La primera fila está escrita en el modo EBCDIC con 80 registros de caracteres bloqueado 8000 caracteres por bloc, esto es 100 registros para un bloc, y contienen aproximadamente 16 bloques. Registros del 0000100 hasta 00130000, son comprendidos en el programa fuente del SYMVU FORTRAN IV. Los registros siguientes al programa fuente, son nuestros datos, que pueden ser usados para producir una grafica. Estos registros, 0013990 hasta 00144500, deberán ser metidos en tarjetas y leidos como datos de entrada al SYMVU, a ejecutar en un tiempo. Si no se esta usando los valores sobre la segunda fila que fué generada por el SYMAP, se necesitará la Subrutina Data, para leer en la matriz de valores en tarjetas; estas tarjetas deberán estar colocadas después de la tarjeta Electiva 3.

La segunda fila de la cinta es un conjunto binario de valores, que es la salida de SYMAP. La forma de registro en un bloc variable, con un registro de longitud 524 y un tamaño de bloc de 15724. Esta fila es usada con la muestra de datos de la primera fila.

6.7.2. Configuración mínima de máquina.

La máquina requerida para el SYMVU es:

— una computadora IBM 360.

— 220 K bytes, centro de almacenamiento (superposición SYMVU con un nivel de compilador FORTRAN iv), o 200 K bytes (superposición con un compilador de nivel H)

— un disco de manejo o tres cintas de manejo.

— una cinta de siete canales para un graficador Cal Comp

— un tambor graficador de 11 pulgadas.

6.7.3. Descripción de unidades.

El programa SYMVU tiene afirmaciones con referencia a las unidades lógicas del FORTRAN iv 1,2,3,5,6 y 8. Estas unidades están referenciadas en cuatro clases de afirmaciones FORTRAN.

- - READ/WRITE (*) X,Y,Z, for binary reads and writes
- - READ/WRITE (*), 1040) for EBCDIC read and writes;
- - REWIND *
- - END FILE *

(el asterisco *, refiere el número de unidad).

Las unidades son usadas como sigue:

i) unidades 1,2 y 3 son usadas solamente por intermedio del almacen, y son escritas al final de la fila, requeridas y leidas; la unidad 1 es usada para almacenar los datos de entrada (valores Z), la unidad 2, los puntos símbolos y la unidad 3, los letreros.

ii) la unidad 5 es usada para la tarjeta de entrada.

iii) la unidad 6 es usada para imprimir la salida, describiendo la gráfica.

iv) la unidad 8 es usada por la cinta binaria de entrada generada por el SYMAP; esta es la cinta de los valores Z que son producidos cuando se usa la electiva 21.

Las rutinas del Cal Comp se referenciarán en otra cinta de siete canales, PLOTTAPE. El programa esta colocado para el graficador de 11 pulgadas. Por insertar CALL FACTOR (FACT) como afirmación número 00029250, se puede incrementar el tamaño de la gráfica de salida, con un factor especificado.

6.7.4. Entrada.

La cinta binaria creada por el SYMAP contiene una descripción de la retícula como primer registro. El primer registro consiste en un número entero de renglones en la retícula, el número entero de columnas, el número real de renglones por unidad de coordenadas, el número real de columnas por unidad de coordenadas, el valor mínimo real, el valor máximo real. El resto de la cinta contiene 1 renglón de datos por registro. Por ejemplo, si la retícula es de 130 x 118, habrá 130 registros de 118 cuatro bytes-real por registro.

El usuario deberá usar la Subrutina Data si sus datos no están en esta forma. Una Subrutina Data si sus datos no están en esta forma. Una Subrutina Data mactiva, está incluida en la cinta fuente, pero el usuario pueda proveer su propia subrutina si el formato de sus datos no es comparable a la cinta de salida SYMAP.

6.7.5. Rutinas del graficador Cal Comp usado para SYMVU.

PLOTS (BUFFER; 1024).

- - BUFFER es un trabajo pulidor para las rutinas del graficador
- - 1024 es la longitud de BUFFER en palabras.
- - la pluma original está colocada en el fondo real del papel.
- - combruebe con su propia instalación, para determinar, si es necesario.

LÍMIT (3000,   260)

- - 3000 esta colocado como el número de registros de cinta del graficador para no excederse.

- - 260 es la proposición FORTRAN; número de ramas, por si la rutina de la grafica intenta escribir mas que 3000 registros.

- - el usuario puede cambiar 3000 o borrar la llamada a esta rutina, dependiendo enteramente de sus propias necesidades (esta rutina actua como una protección en caso de que el programa intente graficar punto malo).

PLOT (X, Y, ± IPEN)

- - (X, Y) son las coordenadas, en pulgadas, de la posición en que la pluma es movida de un punto de referencia corriente.

- - IPEN = 2 mantiene la pluma abajo, durante el movimiento.

IPEN = 3 mantiene la pluma arriba, durante el movimiento, para IPEN= -2 o -3 un nuevo origen es definido en la posición terminal, después del movimiento es completado, de otra manera, es como si IPEN fuera positivo.

- - Si IPEN=999 los efectos son los mismos a los de IPEN=3, excepto que un registro busca la dirección 999 es escrito y el dispositivo de salida es cerrado.

SYMBOL (X,Y, HEIGHT, IBCD, ANGLE,± NCHAR)

- - (X,Y) son las coordenadas, en pulgadas, de la esquina izquierda más baja, del primer caracter, para ser producido; HEIGHT, en pulgadas, del caracter para ser graficado; IBCD es el texto; ANGLE es el ángulo en grados, medido del eje de las X; NCHAR es el número de caracteres para ser graficado de IBCD.

NUMBER (X,Y, HEIGHT,FPN,ANGLE, ± NDEC).

- - X,Y, HEIGHT y ANGLE Son lo mismo que en SYMBOL.

- - FPN es un número de punto flotante que es convertido y graficado;

si NDEC=o, la porción entera del número y el punto decimal son graficados; si NDEC=-1, solamente la porción entera es graficada, NDEC -1, NDEC - dígitos son truncados de la posición entera, despues redondeado.

OFFSET (XOFF, XFAC, YOFF, YFAC).

- Los factores enteros son usados por la rutina PLOT cuando IPEN es 2 o 3. Ver la discusión de FACTOR bajo la descripción de unidades.

1.- A- OUTLINE: Este subpaquete es opcional y se usa para especificar el contorno del área en estudio, en mapas de isolíneas o de proximidad, cuando dicha área de estudio no llena por completo el espacio de los bordes del papel donde está dibujado el mapa fuente.

PRIMERA TARJETA: A- OUTLINE en las columnas 1-9, y la letra X en la columna 23, si no se desea la impresión de los datos de entrada. Si se está utilizando el tipo de medicion por renglón y columna perforar en esta misma tarjeta "8" en columnas 31 y 32, y "10". en las columnas 41-43.

ULTIMA TARJETA: 99999 en columnas 1-5.

TARJETAS INTERMEDIAS:

Las tarjetas que se colocan entre la primera y la última registran las coordenadas de los vértices del contorno, utilizando una tarjeta por vértice. El primer vértice que se perfora es el que está por arriba de todos y más a la izquierda. Se procede en seguida en sentido de las manecillas del reloj hasta incluir de nuevo el vértice inicial. La coordenada vertical se perfora como un número decimal en las columnas 11-20. La coordenada Horizontal en las columnas 21-30, también como número decimal. El formato FORTRAN es (10X, 2F10,0). Muchas veces el área en estudio no está definida en un solo contorno, en tal caso deben usarse uno o más contornos, presentados en cualquier secuencia. Ningún contorno poligonal debe tener menos de tres vértices ni más de 100. Sí un contorno poligonal requiriese más de 100 vértices, debe subdividirse en dos o más.

2.- A-CONFORMOLINES (CONFORMANT OUTLINES)

Este subpaquete se usa para escpecificar el contorno de cada una de las zonas que pertenecen al área de estudio y a las cuales se les asocian los valores que vienen registrados en el subpaquete E-VALUES. Sólo se puede asociar un valor a cada zona en cuestión y por lo común solo se requiere de un contorno para delinear la zona.

El contorno de cualquier área se determina del mapa fuente y se especifica en tarjetas, mediante un procedimiento similar al empleado en el subpaquete A-OUTLINE. Generalmente es buena costumbre numerar cada una de las zonas en el mapa fuente empezando con el número 1 y siguiendo bajo riguroso orden progresivo.

Los Vértices de los contornos de zona deben numerarse en el mapa fuente de la misma manera que los vértices del contorno del área en estudio. Se empieza con la zona Nº 1 y así se sigue secuancialmente:

TARJETAS DEL SUBPAQUETE:

PRIMERA TARJETA:	Col. 1-15	A-CONFORMOLINES
	23	X (si no desea la impresión tabular de los datos de entrada)
	26-27	PU(Si se desea que los centros de las zonas de conformación aparezcan perforadas)
	29-30	PR(si se desea que los centros de las zonas señaladas se impriman en el mapa resultado.
	31-32	S.(si se usan mediciones por columna y renglón)

41-43 10. (por la razón anterior)

63 X (si los datos de entrada están registrados en cinta)

ULTIMA TARJETA: 99999 en columnas 1-5

TARJETAS DE DATOS:

Perforar los datos pertinentes a los contornos zonales, sin pasarse a un nuevo contorno antes de completar el iniciado.

EN LA PRIMERA TARJETA DE CADA CONTORNO DE CONFORMACION:

cols. 1-5	Perforar o dejar en blanco el número de referencia asociado al dato, cargado a la derecha
10	"A" o "L" o "P" para indicar si el contorno de la zona va a representarse como una área, una línea o un punto.
11-20 y 21- 30	Coordenadas del primer vértice del contorno de la zona manejada.

EN TARJETAS SUBSIGUIENTES:

Las coordenandas de los vértices que restan (un vértice por tarjeta, perforadas como la primera tarjeta). En un contorno de área debe repetirse el primer vértice para cerrar el contorno.

3.- B- DATA POINTS

Este subpaquete se usa para especificar las coordenandas de los puntos a los cuales se asocian datos: cada punto de éstos debe numerarse en el mapa fuente empezado así: el primero 1 y continuando sin interrupción. Puntos específicos puede ser localizados fuera del área en estudio y así por fuera de la hoja donde está dibujado el mapa fuente.

No se precisa ninguna secuencia especial de posición para los puntos.

PRIMERA TARJETA:

Cols.	1-13	B-DATA POINTS
	23	X (si no se desea la impresión de los datos de entrada)
	31-32	8.(si las coordenandas se presentan en medición de columna renglón)
	41-43	10.(por la razón anterior)

ULTIMA TARJETA: 99999 cols. 1-5

TARJETAS RESTANTES: Se perforan las coordenandas de los puntos, cada uno en una tarjeta separada (no utilizar más de 1000 puntos de referencia en un mapa). Proceder estrictamente en el orden numérico de los números de referencia establecidos en el mapa fuente.

Cols.	11-20	coordenada horizontal
	21-30	coordenada vertical

4.- C-OTOLEGENDS:

Este subpaquete se usa para especificar la posición relativa y el contenido de cualquier frascología especial, numeración u otro simbolismo deseado dentro del mapa o dentro de los bordes de la hoja en que está dibujado el mapa. Puede también proporcionarse información que debe aplicarse por igual a todos los mapas de una serie, tal sería: el título general aplicable al área de estudio, orientaciones, zonas de importancia, ríos, vías de comunicación, etc. Los mapas pueden correrse en diferentes escalas. Es recomendable que la escala sea presentada en una gráfica sin referencia a pulgadas u otras unidades.

El fondo del mapa, área existente entre el borde del mapa y el contorno del área de estudio, puede ser utilizado para leyendas.

En el área de estudio las leyendas deben usarse con moderación. Este subpaquete habilita al usuario para especificar leyendas, tales como, hileras de caracteres, y símbolos que son colocados en el exterior del mapa con relación a los - ejes vertical y horizontal del mapa fuente. Las localizaciones de éstas en el mapa resultado dependerán de las opciones usadas, 1, 13, 14 y 15 del subpaquete F-MAP que establece el tamaño y la escala del mapa. Para determinar las coordenadas pueden usarse cualesquiera de los métodos descritos anteriormente.

PRIMERA TARJETA:

cols. 1-12 C-OTOLEGENDS

23	X	(si no se desea una impresión de datos de entrada)
31-32	8.	(si las coordenadas se presentan en mediación por columna renglón)
41-43	10.	(por la razón anterior)

ULTIMA TARJETA:

cols. 1-5 99999

TARJETAS RESTANTES:

Perforar las especificaciones de la leyenda, como se estable en la tabla. Las tarjetas requeridas para cada leyenda deben ser colocadas en la cuencia correcta.

DEFINICIONES:

Los CARACTERES son designados como sigue:

1,2,3,4,5,6,7,8,9, NUMERICOS

A,B,C,D,E,F,G,H,I,J,K,L,M,N,O,P,Q,R,S,T,U,V,W,X,Y,Z ALFABETICOS

+,.,(,), -, $, *,/, ',' = ESPECIALES

Un SIMBOLO (para los própositos del programa) se compone de cuatro caracteres, impresos uno sobre otro, cualesquiera de ellos o todos pueden ser blanco. Este proceso es denominado "sobreimpresión".

Una hilera "string" es un conjunto ordenado de uno o más caracteres: "CARACAS" o "*/*/*".

Existen tres grandes tipos de leyendas que pueden ser especificados en el subpaquete "C-OTOLEGENDS". Cada tipo está asociado con un tipo distinto de figuera en el mapa fuente, como sigue:

4.1 LEYENDA PUNTUAL

Este tipo de leyenda se asocia a un punto dado en el mapa fuente.

Las leyendas que pueden ser asociadas con este punto son de tres tipos: 1) una hilera horizontal de caracteres. 2) una hilera vertical de caracteres (una hilera de caracteres pue de no incluir sobreimpresiones) o 3) un símbolo simple (compuesto de cuatro caracteres sobreimpresos).

Las leyendas horizontal y vertical requieren dos tarjetas para cada leyenda.

PRIMERA TARJETA:

cols. 1 En blanco si el "String" es horizontal, y un signo "-" si el String" es vertical.

2-3 son blancos

4-5 perforar un número que indica el número de caracteres del "string". Este número no debe exceder a 50.

6-9 son blancos

10 perforar la letra "P", para indicar que la figura asociada es un punto. La localización del punto en el mapa fuente asociado al "string", está especificado por las perforaciones en las columnas 11-20 y 21-30

11-20 perforar la coordenada vertical (medida con referencia al borde izquierdo)

Ambas están expresadas con números decimales.

21-30 perforar el número correspondiente al número de renglones sumados o restados. Ver ejemplo abajo.

Si se dejan las cols. 31-50 en blanco el primer carácter del --

"string" será impreso (en el mapa resultado) en el punto de lo calización asociado al mapa fuente.

SEGUNDA TARJETA:

En la segunda tarjeta el "string" es perforado, iniciando en la columna 1. La perforación no debe extenderse más allá de la co lumna 50, y debe terminar en la columña cuyo número aparece en las cols. 4 y 5 de la primera tarjeta.

EJEMPLO:

Una leyenda de un solo símbolo requiere solamente una tarjeta y es especificada como sigue:

cols. 1-5	blanca
6-9	perforar los caracteres de impresión y/ó sobreim presión para el símbolo deseado
10	perforar la letra "P"
11-20	perforar la coordenada vertical
21-30	perforar la coordenada horizontal

Si el usuario no desea que aparezca el símbolo exactamente al to pe del punto asociado, es decir, se desea desfasar un número de terminado de renglones o clumnas, se debe especificar, como se indica anteriormente, en las columnas 31-40 y 41-50 respectiva- mente.

2.4.2 LEYENDA DE LINEA:

Este tipo de leyenda se asemeja a la leyenda puntual para símbo- los simples, excepto que el símbolo especificado aparece a lo largo de la línea especificada en lugar del punto simple.

Las leyendas de líneas requieren dos o más tarjetas y son especi ficadas como sigue:

cols 1-5 blancos

6-9 perforar los caracteres de impresión y/ó sobre impresión para el símbolo deseado

10 perforar "L" para denotar línea

Perforar las coordenadas del primer punto de línea en las columnas:

11-20 la coordenada vertical

21-30 la coordenada horizontal

No hay capacidad de desplazamiento con las leyendas lineales y de área. Los vértices subsecuentes en la línea (puntos) en los cuales la línea cambia de dirección) son perforados en una trajeta para cada localización, en las columnas 11-20 y 21-30 como para el primer punto.

2.4.3 LEYENDA DE AREA:

Con esta leyenda se determina una área, la cual es llenada con el símbolo especificado. El método para "OUTLINES" tratado anteriormente es usado para determinar el contorno de la leyenda de área. La leyenda de área requiere tre o más tarjetas y son especificadas como sigue:

cols. 1-5 blancos
6-9 perforar los caracteres de impresión y/ó sobre impresión para el símbolo deseado
10 perforar la letra "A" para denotar área.
11-20 perforar la coordenada vertical del primer vértice en el contorno del área.
21-30 perforar la coordenada horizontal del primer vértice

TABLA 2-1

ESPECIFICACIONES PARA OTOLECENDS

1) LEYENDA PUNTUAL: SIMBOLO SENCILLO- Si se desea sobreimpresión una tarjeta

Cols. 6-9	perforar los caracteres de impresión y sobre impresión para el símbolo deseado.
10	perforar la letra "P"
11-20	perforar la coordenada vertical
21-30	perforar el desplazamiento horizontal
31-40	perforar el desplazamiento vertical deseado, expresado por el número de renglones recorridos hacia arriba (precidido por "-") o hacia abajo, refiriéndose, para la medición, al punto asociado en el mapa fuente.
41-50	perforar el desplazamiento horizontal, expresado por el número de columnas recorridas hacia la izquierda (precidido por "-"), o hacia la derecha.

2) LEYENDA PUNTUAL, CARACTER MULTIPLE (Vertical u horizontal) No sobreimpresión- 2 tarjetas.

PRIMERA TARJETA:

cols. 1	para leyenda horizontal dejar blanco y "-" para leyenda vertical
4-5	perforar el número correspondiente a la cantidad de caracteres (que no exceda a "50")
10	perforar la letra "P"
11-20	perforar la coordenada vertical
21-30	perforar la coordenada horizontal
31-40	perforar el desplazamiento vertical como se explicó en el caso anterior.
41-50	perforar el desplazamiento horizontal

SEGUNDA TARJETA:

cols. 1-50	perforar la leyenda deseada iniciando en la col 1 y terminando en la columna cuyo número fue perforado en las columnas 4-5 de la primera tarjeta.

SEGUNDA TARJETA:

cols. 1-50	perforar la leyenda deseada iniciando en la col 1 y terminando en la columna cuyo número fue perforado en las columnas 4-5 de la primera tarjeta.

SEGUNDA TARJETA:

cols. 1-50 perforar la leyenda deseada iniciando en la col 1 y terminando en la columna cuyo número fue perforar en las columnas 4-5 de la primera tarjeta.

3)LEYENDA DE LA LINEA, SIMBOLO SIMPLE-REPETIDO - 2 o más tarjetas:

PRIMERA TARJETA:

Cols. 6-9	perforar los caracteres de impresión y/ó sobre impresión (cualquiera, pueden ser blancos) para el símbolo deseado.
10	perforar la letra "L"
11-20	perforar la coordenada vertical del primer punto de línea
21-30	perforar la coordenada horizontal del primer punto de la línea.

TARJETAS RESTANTES:

perforar las localizaciones de los vértices subsecuentes de la línea, una tarjeta para cada localización, utilizándose, como se explicó, las cols. 11-20 y 21-30. Las cols 1-10 son dejadas en blanco.

4)LEYENDA DE AREA-SIMBOLO SIMPLE-Area del contorno llena 2 o más tarjetas:

PRIMERA TARJETA:

cols. 6-9	perforar los caracteres de impresión y/ó sobre impresión
10	perforar la letra "A"
11-20	perforar la coordenada vertical del primer vértice (correspondiente al punto más alto en el contorno, en el contorno, en caso de no existir más a la izquierda)
21-30	perforar la coordenada horizontal del primer vértice.

TARJETAS RESTANTES:

perforar las localizaciones de los vértices subsecuentes en el contorno usando una tarjeta para cada localización, utilizando las cols. 11-20 y 21-30 como se describió. En la última tarjeta se repite la coordenada de localización del primer vértice - para cerrar el contorno. En estas tarjetas, las cols. 1-10 son dejadas en blanco.

Si se desea una área en blanco en el mapa de contorno, una leyenda de área puede ser utilizada. En este caso, las columnas 6-9 serán dejadas en blanco. El resto del procedimiento será similar al explicado anteriormente. De esta menra obtendremos un mapa resultado con área en blanco, pero se presentará una interpolación en el área. Para prevenir esto, se requerirá hacer uso del subpaquete:

"D-BARRIERS".

5 D - BARRIERS

Este subpaquete se usa para especificar la localización y características de cualquier barrera requerida, para interpolación entre puntos datos. Por ejemplo, en un mapa de densidad de población de la ciudad de Caracas, el valor calculado en un punto cerca del río Guaire, debería, en condiciones normales, depender de los datos dados a ambos lados del río. Pero el río probablemente forma una barrera física, psicológica o política en la distribución de la población, y el usuario podrá expresar este efecto, mediante el uso de una barrera. Existen barreras "impermeables", las cuales no permiten que ocurra la interpolación a través de la barrera; y "permeables" que restringen la interpolación a través de la barrera, pero no la detienen.

5.1. BARRERAS IMPERMEABLES:

Esta barrera divide el área de estudio en regiones completamente separadas. Dos tipos de esta barrera son posibles:

1.- LINEA DIVISORIA: Una barrera de línea divisoria se inicia en el borde del mapa, pudiendo tener cambios en dirección, y, termina en el borde del mapa (puede terminar en el mismo o en diferente borde)

2.- CIRCUITO: Cuando se requiere que una región sea separada de otra que la rodea completamente, entonces debe "colocarese" una barrera de circuito alrededor de la región anterior.

5.2. BARRERAS PERMEABLES: Las barreras permeables se usan para representar la influencia aniquiladora del fenoméno en estudio, de bahías, puertos, lagos, u otros obstáculos.

a) BARRERA POR EFECTOS DE RELIEVE: Una bahía, puerto o boca de río pueden ser pasados como un "relieve" en el área de estudio.

Los puntos sobre el relieve deben llevar una relación de resistencia a la influencia, dependiendo del segmento formado con el punto anterior (esta resistencia no depende de la distancia real).

b) OTRAS CONFIGURACIONES: En general, la resistencia en cualquier vértice terminal con el área de estudio debería ser cero. Un lago de gran extensión necesitará una barrera que se extiende a lo largo de su eje mayor, con cero en ambos extremos y con un incremento en la resistencia hacia el centro, donde un vértice adicional debe colocarse. La resistencia en este vértice deberá ser igual a la longitud del lago.

PRIMERA TARJETA:

cols.	1-10	perforar D-BARRIERS
	23	perforar la letra "X" si no se desea la impresión tabular de los datos de entrada.

Si el método de medición usado es por renglón y columna, perfore el número "8". en las cols. 31-32 y el "10". en las cols. 41-43.

ULTIMA TARJETA:

cols	1-5	perforar 99999

TARJETAS RESTANTES:

Perforar las coordenadas de localización de los vértices de la barrera, seguido por la resistencia asignada.

Si una barrera tiene un extremo en el mapa y el otro en el borde, el primer vértice será el extremo interior.

Para otro tipo de barreras, inicie con el vértice terminal más alto, si ambos se encuentran a igual altura, inicie con el situado más a la izquierda. Perforar los vórtices en orden de aparición, usando una tarjeta para cada vértice; si la barrera es cerrada proceda en el sentido de las manecillas del reloj. Siempre debe repetirse el último vértice de la barrera. Esta repetición indica al computador que la barrera ha sido completada.

cols. 11-20	perforar la coordenada vertical del vértice como un número decimal (medida del borde superior hacia abajo)
21- 30	perforar la coordenada horizontal del vértice - (medida a partir del borde izquierdo)
31- 40	perforar el dato de resistencia como un número decimal)

Si la barrera es impermeable, se debe perforar "-1" en las cols 38-40. Si se desea, pueden emplearse más de dos barreras. Cada - barrera actúa en forma independiente, sin importar si existe coincidencia en alguna porción de su longitud.

6 E - VALUES

Este subpaquete es usado para especificar los valores correspondientes a cada punto de los datos (para un mapa de isolineas o de proximidad) o a cada zona de los datos (para un mapa coropleta).

PRIMERA TARJETA:

cols	1-8	perforar E-VALUES
	23	perforar la letra "X" si no se desea la impresión tabular de los datos de entradas.

ULTIMA TARJETA:

cols. 1-5	perforar 99999

TARJETAS RESTANTES:

Perforar los valores que se desean aplicar a los puntos de los datos dados en el subpaquete "A-CONFORMILINES". Cada valor debe ser perforado en una tarjeta. Si el valor es negativo, perforar un signo menos antes. El orden de las tarjetas perforadas debe sujetarse estrictamente al orden numérico referido al mapa fuente, y estar acorde a los puntos del subpaquete "B-DATA POINTS". Los valores deben perforarse en las cols. 11-20 como un número decimal.

E1 - VALUES INDEX

Este subpaquete se utiliza para cambiar el orden de referencia del valor de los datos en el subpaquete "E-VALUES" descrito en el punto anterior.

PRIMERA TARJETA:

cols. 1-5	perforar E1 -VALUES INDEX

ULTIMA TARJETA:

cols. 1-5 perforar 999999

TARJETAS RESTANTES:

cols.	1-5	perforar el número correspondiente al orden normal del valor del punto.
	6-10	perforar el número correspondiente al nuevo orden deseado.

Para tener una referencia entre los dos valores, deberán perforarse un "1" en la columna 5 y "2" en la columna 10. Si las columnas 6-10 son blancas, el número siguiente será asignado con el orden normal.

7 F - MAP

Este subpaquete da las instrucciones al computador para hacer el mapa, basándose en la información suministrada en los subpaquetes restantes. En este subpaquete se especifica el mapa de acuerdo - con ciertas opciones, que serán explicadas adelante.

PRIMERA TARJETA:

cols. 1-5	perforar F-MAP
23	perforar "X" si no se desea la impresión de los datos de entrada.

ULTIMA TARJETA:

cols.1-5	perforar 99999

SEGUNDA,TERCERA Y CUARTA TARJETAS:

Perforar el título que se desee que aparezca en la parte inferior del mapa. Es importante que esta descripción sea clara y se diferencie de los otros mapas corridos o por correr. Una o más de estas tarjetas pueden ser blancas si se desea, pero las 3 tarjetas deben suministrarse.

TARJETAS RESTANTES:

Deben ser colocadas entre la cuarta y la última y perforar las opciones deseadas. Estas opciones controlan el formato del mapa de salida.

2.8.1.TERMINACION

Una última tarjeta debe ser perforada con"999999" en las columnas 1-6 para terminar el programa. Cuando se corren "mapas múltiples" esta tarjeta se necesita solamente después de terminar el último subpaquete "F-MAP".

El subpaquete F-MAP permite al usuario variar el mapa de salida, de acuerdo con las necesidades o requisitos particulares que se tengan. Esto se logra mediante el uso de varias opciones disponibles. No es necesario invocar una cualquiera o todas opciones para producir un mapa, porque en todos los casos existen opciones estandar que el programa mismo genera. Sin embargo, por regla general, el usuario probablemente se decida por alguna de tales opciones, en particular, por alguna de las diez primeras.

OPCION NUMERO 1: TAMAÑO DEL MAPA PRODUCIDO

Para especificar el tamaño del mapa de salida, perforar el número 1 (identificación de la opción 1) en la columna 5.

Columnas 11-20: dimensión vertical de banda a banda del cuadro del mapa, en pulgadas y como número de punto flotante.

Columnas 21-30: dimensión horizontal del cuadro del mapa, en pulgadas y como número decimal.

No existe un límite establecido para las dimensiones del mapa, pero el programa tiene una medida de seguridad que consiste en que una dimensión mayor de 72" es reducida a 13", y las restantes dimensiones son reducidas proporcionalmente.

Para anular esta medida de seguridad del programa se utiliza la opción 16.

Si no se incluye la opción 1, esto es, no se especifica dimensión, el programa calcula las dimensiones vertical y horizontal en base a las posiciones puntuales extremas (ver opción)

21) y asignará 13" a la mayor dimensión.

OPCION NUMERO 2; PUNTOS EXTREMOS DEL MAPA DE ENTRADA

Esta opción se utiliza para especificar las coordenadas de los puntos extremos que determinan el contenido del mapa que debe producirse. Este contenido puede incluir todo, o cualquier porción, del área que se muestra en el mapa fuente.

De esta manera se puede obtener, ya sean pequeñas porciones del área en estudio o proporcionar un margen alrededor del mapa fuente completo, mediante una juiciosa selección de los puntos extremos.

Si no se especifican estos puntos extremos el programa los genera seleccionandolos de las coordenadas máximas y mínimas que obtiene de un subpaquete previo-A-Conformolines, A-Outline o B-Data Points.

Para establecer los dos puntos extremos que se van a especificar, primero desígnese el área que se va a graficar, por medio de un rectángulo dibujado en el mapa fuente. Las coordenadas de la esquina superior derecha son, respectivamente, los puntos extremos mínimo y máximo que determinan el contenido del mapa que se va a producir.

Las coordenadas de estos puntos deben medirse desde los bordes superior e izquierda establecidos originalmente en el mapa, fuente y especificarlas en las unidades utilizadas en todas las demás coordenadas.

Excepción: Aunque las coordenadas se hayan calculado en el tipo renglón-columna, especifíquense en pulgadas las coordenadas para esta opción.

En una tarjeta, identifíquese el número de opción: "2", perforado en la columna 5.

Columnas 11-20: La coordenada vertical mínima

Columnas 21-30: La coordenada horizontal mínima

Columnas 31-40: La coordenada vertical máxima

Columnas 41-50: La coordenada horizontal máxima

Estándar: El programa selecciona los puntos extremos que determinan el contenido del mapa. El uso de esta opción se recomienda junto con la opción 1, la 13 (escala) y la 14 (avance).

OPCION NUMERO 3: NUMERO DE NIVELES

Para especificar el número de niveles o intervalos de clase, (de 2 hasta 10, inclusive), se divide el valor del rango total de valores con propósito de designación por medio de simbología. Perforese el número de opción: "3" en la columna 5, seguido por el número de niveles deseado, perforados como número de punto flotante en las columnas 11-20.

El número de niveles estándar es 5.

OPCION NUMERO 4; MINIMO DE LOS VALORES DEL RANGO

Para especificar un valor que sea usado como el valor mínimo del rango total de valores, perforar el número de opción: "4" en la columna 5, seguido por el valor mínimo deseado, perforado como número de punto flotante en las columnas 11-20.

El valor mínimo estándar es el valor mínimo de los datos.

OPCION NUMERO 5; MAXIMO DE LOS VALORES DEL RANGO

Para especificar un valor que sea usado como el valor máximo del rango total de valores, perforar el número de opción: "5" en la columna 5, seguido del valor máximo deseado, perforado como número de puntos flotante en las columnas 11-20.

Si el subpaquete E-Values contiene valores más grandes que el especificado, aparececerá en el mapa un área de simbología que consiste en la letra H (high).

OPCION NUMERO 6; RANGO DE LOS VALORES DE LOS INTERVALOS

Esta opción nos permite especificar el rango de los valores para cada nivel o intervalo en los que se va a dividir el rango total de valores, asociados a las zonas o puntos que son dados.

Estándar: A cada nivel o intervalo se le asigna igual rango.

OPCION NUMERO 7: SIMBOLOGIA

Esta opción nos permite especificar simbología diferentes de la estándar.

TARJETA 1: "7" en la columna 5

TARJETA 2: columnas 1-10: Símbolos para cada nivel(caracteres básicos) Columna 1, símbolo para el primer nivel, Columna 2, símbolo para el segundo nivel, etc.

Columnas 11-20: para especificar los símbolos de los puntos dato. (Columna 11; perforar el símbolo que corresponde a los puntos dato del primer nivel: Columna 12: perforar el símbolo que corresponde a los puntos dato del primer nivel. Columna 12: perforar el símbolo que corresponda al segundo nivel, etc.

TARJETAS 3, y 5: Para perforar cualquier carácter de sobreimpresión.

Simbología estándar: Varía de acuerdo con el número de niveles, en la opción número 3.

La tabla muestra el número de niveles que van a ser mapeados, la simbología en general que aparecerá en cada uno de estos niveles y los símbolos correspondientes a los puntos datos.

Una simbología adicional de identificación puede especificarse en las columnas 21-29: La simbología estándar que se presentará incluye:

COLUMNA 21: Simbología general baja. "L" y 3 columnas en blanco

22: Simbología de punto dato, baja "L", 2 en blanco.

23: Simbología general, alta "H" y 3 en blanco.

24: Simbología de punto dato, alta "H", "H", "H" y "/".

25: Simbología de fondo (más allá del contorno del área de estudio). Todas las columnas en blanco.

26: Simbología de isolínea o frontera de conformación. Todas las columnas en blanco.

27: Sin datos (se usa únicamente para problemas surgidos de la especificación de barreras) "N" y 3 blancos.

28: Puntos dati superpuestos. "S" y 3 blancos.

29: Puntos datos con valores inválidos, llamando "datos faltantes"."M" y 3 en blanco.

<u>Simbología no estándar:</u>

El siguiente ejemplo muestra la correcta perforación de las cinco tarjetas requeridas para la opción "7", si se especifican 6 niveles en la opción 3, y se desea el orden inverso de la simbología estándar:

Columna	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
Tarjeta 1						7																				
Tarjeta 2		0	0	0	X	+	.				1	2	3	4	5	6										
Tarjeta 3		X	X	-																						
Tarjeta 4		A																								
Tarjeta 5		V																								

La siguiente ilustración muestra la correcta perforación de las 5 tarjetas para la opción 7 si se desea: suprimir toda la simbología general, hacer negras las isolíneas y tener un fondo (del mapa) con diagonales; al mismo tiempo que se deja estándar todo lo que resta.

Columna	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
Tarjeta 1				7																								
Tarjeta 2									1	2	3	4	5													-	0	
Tarjeta 3																											X	
Tarjeta 4																											A	
Tarjeta 5																											V	

La siguiente ilustración muestra la correcta perforación de las 5 tarjetas requeridas en la opción 7, si se desea correr un mapa de proximidad, en el cual se usen los símbolos A,B,C,D,y E para las clases de datos.

Columna	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
Tarjeta 1			7																							
Tarjeta 2	A	B	C	D	E					A	B	C	D	E												
Tarjeta 3										/	/	/	/	/												
Tarjeta 4																										
Tarjeta 5																										

OPCION NUMERO 8: PARA SUPRIMIR LAS ISOLINEAS

1 tarjeta perforando en la columna 5: "8"

OPCION NUMERO 9: PARA SUPRIMIR LAS BARRAS DEL HISTOGRAMA

1 tarjeta perforando en la columna 5: "9"

OPCION NUMERO 10: TEXTO (de 3 a 32 tarjetas)

Esta opción permite dar, bajo el mapa, información suplementaria, además de la contenida en las leyendas y el título del mapa.

Tarjeta 1: Columna 4-5 "10"

Tarjeta última: Columnas 1-5: "99999"

En no más de 30 tarjetas: Columnas 1-72: Información suplementaria que pueda ser de tulidad para referencias faltantes, tales como fuente de datos, fecha, nombres de las personas que intervienen, etc.

Estándard: No tener ningún letrero.

OPCION NUMERO 11: VALOR REAL IMPRESO EN CADA PUNTO DE REFERENCIA

Una tarjeta para hacer que los valores de los datos aparezcan en la porción del punto de referencia correspondiente.

Columna 4-5: "11"

Estandard: El valor del punto de referencia no aparece.

OPCION NUMERO 12: REPETICION DE VARIAS OPCIONES

La opción 12 puede utilizarse paraindicar que en el siguiente o en los siguientes subpaquetes, de una serie de mapas sometido a procesamiento, las opciones no estandard que se usaron en el mapa precedente, se deben respetar.

Se pueden agragar también algunas nuevas opciones, o modificar las anteriores. Sin embargo, esta opción no debe utilizarse si cualquier opción no estandard se habrá de suprimir.

Para respetar todas las opciones no estaddard que se usaron en un mapa precedente, perfórese la misma identificación de la opción (el 12) en las columnas 4-5.

Estandard: Proporcional las tarjetas requeridas para cada opción no estandard que se use en cada mapa.

OPCION NUMERO 13: ESCALA DEL MAPA

Para especificar la escala del mapa de salida, perforar en las columnas 4-5 el número 13.

Columnas 11-20: Escala, número de pulgadas que se utilizan como unidad de medida en el mapa fuente.

Si la escala del mapa se especifica como "1".0" cada pulgada en el mapa de salida, representará una unidad de medida en el mapa fuente. Si no se especifica la escala, ésta se calculará del tamaño del mapa y de los puntos extremos.

Estandard: La escala se establece automáticamente de otras opciones.

OPCION NUMERO 14: MARGEN

Esta opción se usa para obtener un márgen (no necesariamente en blanco) entre los puntos extremos del área en estudio y los bordes correspondientes de la hoja donde está dibujado el mapa.

Puede también usarse para estrechae el contenido del mapa resultado.

El programa SYMAP, normalmente encuadra el contenido del mapa (determinado por los puntos extremos) del tal manera que ajuste exactamente dentro de los bordes del papel. Si se "recorre" (avanza) el borde del papel en una distancia positiva, el programa supone que el borde está más cercano al centro del papel en la cantidad especificada. En consecuencia aparece el márgen debido a la diferencia entre el borde supuesto y el real.

Por otro lado, una especificación negativa, hace suponer al programa que el borde está más allá del centro que el borde real, por tanto trunca el contenido del mapa.

Para obtener un márgen de media pulgada entre el área de estudió (determinada por los puntos extremos) y todos los bordes del mapa, especificar "0.5" en cada campo de columnas.

Formato:

en una sola tarjeta:

Columna 4-5 : "14"

" 11-20: Lo que se va a recorrer el borde superior

" 21-30: Idem. el borde izquierdo

" 31-40: Idem. el borde inferior

" 31-50: Idem. el borde derecho.

Estandard: Los puntos extremos están en la posición de los bordes del papel.

OPCION NUMERO 15: NUMERO DE CARACTERES DE SALIDA (Por pulgada)

Columnas 4-5: "15"

Columnas 11-20: Número de líneas por pulgada (verticalmente). El número debe de estar cargado a la derecha del campo.

Columnas 21-30: Número de líneas por pulgada (horizontalmente). El número debe estar cargado a la derecha del campo

Estandar: 8 líneas/pulgada, y 10 columnas/pulgada.

OPCION NUMERO 16: MAPAS DE GRAN TAMAÑO

Si se desea que haya dimensiones mayores de 72", perforar en las columnas 4-5: "16".

Esta opción se usa para suprimir una medida de seguriada del programa, la cual impide la impresión de un mapa mayor de 72" en cualquier dirección.

Estandard: La impresión de un mapa con ninguna dimensión mayor de "72 pulgadas".

OPCION NUMERO 17: PARA SUPRIMIR LAS TABLAS DE DATOS DE SALIDA PARA " CONFORMOLINES" Y PUNTOS BAJO ANALISIS.

Perforar en las columnas 4-5, "17"

Estandard: Una impresión tabular de los datos de dalida, inmediatamente antes del mapa.

OPCION NUMERO 18: PARA HACER INVALIDOS LOS VALORES DE MENOS CERO Y VALORES QUE NO SE ESPECIFICAN.

Perforar en las columnas 4-5: "18"

En cualquier punto bajo análisis, cuyo valor asignado se considere inválido, parecerá el símbolo M" en lugar del símbolo numérico que normalmente indicará el nivel asignado al punto.

En la opción 7 se puede especificar simbolísmos diferentes de los estandard, pra datos inválidos.

Estandar: Todos los valores se consideran válidos y se utiliza para los cálculos, incluyendo valores que no se especifican (a los que se les asigna el valor -0.00)

OPCION NUMERO 19: VALOR MINIMO DE DATOS INVALIDOS

Los datos numéricos que aparecen en el subpaquete E-VALUES y que sean inválidos y no se usan para los cálculos.

Perforar en las columnas 4-5: "19" y en las columnas 11-20; el valor mínimo deseado.

Estandard: Todos los datos numéricos se consideran válidos.

OPCION NUMERO 20: VALOR MAXIMO DE DATOS INVALIDOS

Los datos numéricos iguales o menores que el valor máximo considerado se consideran inválidos y no se utilizan en los cálculos. Perforar en las columnas 4-5: "20" y en las columnas 11-20: el valor máximo deseado.

Si se especifica el mismo valor en las opciones 19 y 20, solamente ese valor no se usará en los cómputos.

Si el valor inválido mínimo especificado es menor que el valor máximo especificado, sólo los valores comprendidos en el intervalo MIN Y MAX no se utilizan en los cálculos

Si el valor inválido mínimo entre los dos valores se usaran en los cálculos.

OPCION NUMERO 21: GRABAR EL MAPA EN CINTA

Con propósito de uso con otros programas, se pueden grabar los valores de posición del mapa.

Columnas 4-5: "21"

Columnas 19-20:"1" para obtener la impresión de los valores.

Cuando el mapa se está generando, el programa calcula un valor para cada posición, utilizada para definir el mapa. Esta opción graba en cinta estos valores, para uso posterior como datos de entrada en otros programas.

Estandard: No imprimir o grabar estos valores en cinta.

OPCION NUMERO 22: ISOLINEAS CONTINUAS

Para garantizar la aparición de isolíneas, que de otra manera se suprimirían, permitiendo la aparición de simbología descriptiva, cuando se presenta la situación de que el espacimiento entre pun tos y zonas bajo análisis, sea inadecuado para que aparezcan tanto la las isolíneas como la simbología.

Perforar en las columnas 4-5: "22"

Estandard: Suprimir las isolíneas, cuando por su inclusión, resulta una pérdida en la simbología descriptiva.

OPCION NUMERO 23: SUPRESION DEL SIMBOLO DE PUNTOS CON VALORES INVALIDOS

Perforar en las columnas 4-5 "23"

Estandard: El símbolo "M" aparece en cualquier punto que tiene un valor inválido asociado.

OPCION NUMERO 24: SUPRESION DE LA EQUIVALENCIA NUMERICA

Para supromir la interpretación numérica que aparece en la línea que reza: " ABSOLUTE VALUE RANGE TO ECH LEVEL" PERCENTAGE OF TOTAL ABSOLUTE VALUE RANCE APPYING TO ECH LEVEL"

Perforar en las columnas 4-5 "24"

OPCION NUMERO 25: SUPRESION DE SIMBOLOS DE PUNTOS DE ANALISIS

Perforar en las columnas 4-5: "25"

Estandard: Los puntos objetivos aparecerán en el mapa de conformación.

OPCION NUMERO 26: ALINEAR LA SOBREIMPRESION

Esta opción se utiliza únicamente para ciertas computadoras (IBM 7094) cuya línea de sobreimpresión aparece desfasada de la línea de impresión origianal.

Perforar en las columnas 4-5: "26"

Estandard: La impresión y la sobreimpresión son automáticas en las computadoras IBM 360 o IBM 370

OPCION NUMERO 27: DISTINGUIENDO EL TIPO DE MAPA

Para pedir que el programa identifique el mapa deseado como un mapa de isolíneas, cuando se dan datos tantos para mapas de isolíneas como de conformación, Perforar en las columnas 4-5: "27".

Estandard: Se produce un mapa de conformación si se ha incluido un sobpaquete A-CONFORMOLINES en el paquete **SYMAP**.

OPCIONES PARA LA EXTRAPOLACION DEL RANGO

El programa SYMAP versión V, para obtener mapas de isolíneas, tiene la capacidad de extrapolar más allá del valor más alto dado para una colina (cerro) o más bajo del mínimo valor dado para un valle (cauce). Las siguientes tres opciones controlan esta cuestión.

OPCION NUMERO 31: EXTRAPOLACION FRACIONAL RELATIVA

Perforar en las columnas 4-5: "31", y en las columnas 4-5: "11-20: La fracción del rango de valores que se desea extrapolar. Perforar como número decimal menor que 1.

Estandard: Extrapolación fraccional de 0.1.

OPCION NUMERO 32: MINIMO ABSOLUTO DE EXTRAPOLACION

Perforar en columnas 4-5: "32", y en columnas 11-20: Valor mínimo, debajo del cual no se desea extrapolar.

Estandar: Valor mínimo de extrapolación igual al mínimo de los datos numéricos.

OPCION NUMERO 33: VALOR MAXIMO DE EXTRAPOLACION

Perforar en las columnas 4-5: "33", y en las columnas 11-20: El valor máximo, más allá del cual no se desea extrapolar.

OPCIONES DE BUSQUEDA DE RADIO

Las siguientes dos opciones especiales permiten al usuario tener un control preciso sobre la selección de puntos para interpolar. Un "radio de búsqueda" (vecindad) es la distancia máxima donde el programa "busca" puntos entre los cuales realiza interpolación.

OPCION NUMERO 34: RADIO INICIAL DE BUSQUEDA

Perforar en las columnas 4-5: "34", y en las columnas 11-20: El radio inicial de búsqueda, siempre positivo.

Esta opción normalmente se usará para precisar el radio de búsqueda inicial de un mapa previo, después de modificar el número de puntos o la posición de los puntos cerca de la frontera del área de estudio.

Estandard: La "búsqueda inicial" se basa en el número y dispersón de los puntos, de tal manera que, como promedio, se localizan 7 puntos en el radio seleccionado.

OPCION NUMERO 35: RADIO MAXIMO DE BUSQUEDA

Perforar en las columnas 4-5: "35", y en las columnas 11-20: El radio máximo de búsqueda, perforado como número decimal.

El radio máximo de búsqueda no debe ser menor que el radio inicial de búsqueda.

Estandard: No se tiene un límite para el radio de búsqueda, pero el programa no buscará más allá de ciertos rangos.

OPCION NUMERO 36: NUMERO DE PUNTOS PARA INTERPOLAR

Si se desea especificar el número máximo y mpinimo de puntos en los cuales debe buscarse la interpolación.

Perforar en las colomnas 4-5: "36", y en las columnas 11-20: El mínimo deseado (0-10). y en las columnas 21-30: el máximo deseado (no mayor de 10) cargado a la derecha del campo.

Estandard: Mínimo:4, máximo 10 y promedio: 7.

OPCION NUMERO 37: INDEPENDENCIA

Perforar en las columnas 4-5: "37"

Esta opción sirve para especificar en forma independiente el valor asociado a cada posición. Impide la suavización a través de líneas creadas por barreras o por usos especiales de las opciones 35 y 36. Generalmente esta opción es innecesaria.

Estandard: Calcula los valores en forma independiente, únicamente para 1/6 de los caracteres.

OPCIONES PARA MAPAS DE PROXIMIDAD.

La combinación de las opciones 31, 36 y 37 se usa para especificar al mapa de proximidad.

El mapa de proximidad se basa en el concepto de "el vecino más cercano", por medio del cual la simbología para cada carácter de posición se determina por la simbología aplicable al punto de análisis (punto dato) más cercano.

La información "puntual" se usa tanto para mapas de ISOLINEAS como para mapas de PROXIMIDAD, sin embargo, los niveles que distingue en un mapa de ISOLINEAS no aparecen en un mapa.

El mapa de PROXIMIDAD es especialmente útil, para representar datos cualitativos, particularmente cuando la definición precisa de la frontera zonal no es de primordial importancia.

Formato:

Las opciones 31, 36 y 37, perforadas en las columnas 4-5 de 3 tarjetas sucesivas.

CORRIENDO VARIOS MAPAS.

Se requiere un subpaquete F-MP para cada mapa deseado. Los subpaquetes de datos, una vez registrados, se usarán para mapas subsecuentes, hasta que se modifiquen mediante la inclusión de subpaquetes del mismo tipo (con datos diferentes), o hasta que sean suprimidos mediante la inclusión de subpaquetes (sin datos) que contengan únicamente las tarjetas primera y última, requeridas.

El siguiente es un ejemplo de un procedimiento para correr varios mapas en un solo proceso:

MAPA 1: Incluye los subpaquetes A-OUTLINE (1), B-DATA POINTS (1), C-OTOLEGENDS (1), E-VALUES (1) y F-MAP (1).

Para un MAPA 2: Un mapa similar con diferentes valores, puede producirse agregando solamente los subpaquetes E-VALUES (2) y F-MAP (2). El subpaquete F-MAP (2) contendrá ajustes en los títulos, y puede variar en algunas opciones.

Para el MAPA 3 de la serie el usario podrá desear introducir una barrera a la interpolación, para representar el posible efecto de una carrera que pasará por el área. Solamente sería necesario agregar los subpaquetes D-BARRIERS (1) y F-MAP (3).

En el MAPA 4 se podría desear usar un tercer conjunto de datos y suprimir la barrera que se introduciera en el mapa 3. Para lograr la supresión, se compondría un subpaquete D-BARRIERS con sólo dos tarjetas. (D-BARRIERS y 99999).

Se emplea un procedimiento similar para correr una serie de mapas coropletas. Así mismo el procedimiento es válido para casos que incluye tanto mapas de ISOLINEAS como coropletas.

Existen ciertas circunstancias especiales bajo las cuales el programa supone que se desea un mapa coropleta e ignorará los subpaquetes A-OUTLINE y B-DATA POINTS.

Por otro lado, a menos que aparezca un subpaquete A-COFORMOLINES, el programa supone que se desea un mapa de ISOLINEAS, y en tal caso hará uso de todos los subpaquetes.

Para procesar mapas no relacionados en una misma corrida, debe agregarse una tarjeta antes de los subpaquetes de datos de cual quier mapa que no esté relacionado con el precedente.

REPETICION DE OPCIONES

Al preparar el subpaquete F-MAP debe recordarse que las opciones no estandard se se repiten automáticamente del mapa precedente. Para lograr tal resultado debe incluirse en el nuevo subpaquete F-MAP una tarjeta por cada opción que se desea repetir.

Formato:

Columnas 4-5: número que identifica la opción

Columnas 6-9: "SAME"

Si se desea repetir TODAS las opciones no estandard usadas en el mapa precedente, perfórese el número identificador de la opción "12" en las columnas 4-5. En el nuevo subpaquete F-MAP se puede a gregar y modificar opciones no estandard.

EL PROGRAMA GRID.

INTRODUCCION.

GRID es un programa para computadora que ha sido creado con el objeto de proporcionar un medio de gran eficiencia para la representación gráfica de información recolectada usando como base una malla de coordenadas rectangulares. El programa GRID está diseñado para que pueda utilizarse por personas con muy poca experiencia en programación. Sin embargo, con frecuencia es necesario que el usuario especifique sus propios formatos mediante la subrutina FLEXIN y esto requiere conocimientos elementales de FORTRAN IV.

1.1. PRINCIPIOS BASICOS.

Cada dato está asociado a una célula de la malla. Es muy importante que los valores de los datos sean procesados en el orden correcto, en virtud de que el programa acepta los datos en el orden en que se imprime el mapa.

Utilizando el proceso estandar de impresión, el programa empieza por la parte superior del mapa y procesa los datos horizontalmente renglón por renglón y de izquierda a derecha en cada renglón.

El tamaño y forma de la malla debe ser especificada por el usuario usando la opción 1.

En el proceso de creación de mapas, los valores reales de los datos se generalizan en grupos. Cada grupo tiene asociado un símbolo único. Los grupos en los que deben localizarse los datos y los símbolos asociados pueden también ser especificados por el usuario utilizando las opciones de la 3 a la 6.

Se dispone de dos tipos de simbología:

Una escala de gris que va del oscuro al claro ó un mapa punto en células de media pulgada de lado. Se pueden imprimir también las coordenadas de cada célula en la malla. La mayoría de las opciones de SYMAP para generalización de escala estan también disponibles en GRID.

1.2. ALIMENTACION AL PROGRAMA.

Para obtener una representación gráfica (ó un mapa), el usuario debe proporcionar tres conjuntos de instrucciones y tiene la opción de proporcionar un cuarto conjunto.

Las instrucciones se preparan en los siguientes paquetes:

Paquete de DATOS (usualmente una cinta por separado), paquete del MAPA, paquete de CONTORNO IRREGULAR (opcional) y la subrutina FLEXIN.

i) El paquete de DATOS contiene los datos ó información numerica que genera la representación gráfica.

El programa está diseñado para un máximo de 10 000 células de malla, pero la opción "conjunto de datos múltiples" permite manejar un número ilimitado de células de datos.

ii) El paquete del MAPA permite especificar la forma precisa del mapa de salida en términos de las diversas opciones.

iii) El paquete de CONTORNO IRREGULAR permite especificar las fronteras del area de estudio en el caso de que se esté manipulando un área que no contiene contornos irregulares.

iv) La subrutina FLEXIN es una subrutina en FORTRAN que sirve para especificar el formato de los datos.

En las secciones siguientes se describe primeramente el contenido y después el formato de cada conjunto de requisitos de entrada del programa.

2 2. PAQUETE DE DATOS.

El programa GRID proporciona dos procedimientos diferentes para la entrada de datos:

2.1. OPCION A PARA DATOS.

La opción A usa al GRID como un programa independiente, en el cual, la subrutina FLEXIN se utiliza para:

i) Leer los datos de un archivo.

ii) Ejecutar cálculos estadísticos sencillos en los datos para generar el valor que se vá a mapear.

En este caso, los datos se procesan célula por célula. Esta opción permite un amplio rango de flexibilidad en la organización de los datos.

Se recomienda que las personas que posean poca experiencia con computadoras utilicen esta opción.

Un archivo de datos, generalmente contiene muchas variables diferentes referidas a cada célula de la malla. Cuando no existe suficiente memoria disponible para almacenar cada variable, será necesario leer el archivo de datos básicos cada vez que se haga un mapa. Cuando el número de datos es grande, es impráctico el utilizar un archivo de tarjetas para cada mapa por separado, por lo que se recomienda que tales archivos se graben en disco ó cinta. El programa GRID releerá automáticamente un archivo que se lee en la unidad 12 de entrada/salida en FORTRAN lógico entre mapa y mapa.

Esta opción se activa al especificar un número mayor que cero en el campo 1 de la opción 2.

2.2. OPCION B para datos.

La opción B utiliza al GRID como la "etapa de trabajo final" en una serie de "etapas" que están compuestas de rutinas para manipulaciones estadísticas complejas.

En este caso, los datos utilizados para crear la representación gráfica se transfieren al programa GRID en la forma de arreglos binarios, con un arreglo (ó registro lógico) por cada renglón del mapa. El programa [illegible] un valor real en el arreglo para cada célula en un renglón.

Esta opción se utiliza en forma automática si no se especifica la opción 2. También es activada si aparece un cero en el campo 1 de la opción 2.

2.3. ASIGNACION DE DATOS A LOS NIVELES DEL MAPA.

Internamente, el programa GRID asigna el valor asociado a cada célula de la malla a un cierto nivel ó grupo. El número máximo de niveles es 10, numerados del 0 al 9.

Cuando se utiliza la opción de MAPA PUNTUAL el número máximo de niveles es 20, numerados del 0 al 19.

2.4. CONJUNTOS DE DATOS MULTIPLES.

En su forma estandar, este programa está limitado a procesar 10 000 células dato. Un programador experimentado puede ajustar este límite combinado el tamaño de "COMMON P (10 000)". Un circuito interno construido dentro del programa permite graficar al usuario tantas células dato como sean necesarias para un mapa. Para lograr tal objetivo, se pueden dividir los archivos que se salgan del limite en conjuntos que no tengan más de 10 000 células cada uno.

Cuando es preciso el utilizar conjuntos de datos múltiples, es necesario que se especifiquen los rangos máximo y mínimo de los datos ya que cada conjunto se procesa por separado. El número de conjuntos se especifica en el campo 3 de la opción 2.

Se sugiere que los conjuntos de datos referidos a un contorno irregular se organicen de tal modo que el primer conjunto este referido a la sección superior del mapa y el último a la sección inferior, asegurandose de que cada conjunto abarque totalmente el ancho del mapa .

3. PAQUETE DEL MAPA.

Este paquete instruye a la computadora para que dibuje un mapa basado en los datos proporcionados. Especifica la forma precisa del mapa en términos de una serie de opciones disponibles.

Es importante recordar que:

i) Una vez especificada cualquier opción, ésta seguirá siendo válida en los mapas subsecuentes a menos que sea cambiada.

ii) Las opciones 1 y 7 deben incluirse en el primer mapa de una serie, ya que el programa no crea ninguna condición estandar en relación a estas opciones.

En la primera tarjeta, perforar 'MAP' en las columnas 1-3

En la última tarjeta, '99999' en las columnas 1-5.

En la segunda, tercera y cuarta tarjetas, perforar el título que desea aparezca al pie del mapa.

Una ó más de estas tarjetas pueden dejarse en blanco si se desea, pero es imprescindible que las tres tarjetas aparezcan en el paquete.

En las tarjetas restantes perforar las opciones deseadas.

El programa generará opciones estandar para las opciones no especificadas.

3.1. FORMATO ESTANDAR.

Para todas las opciones, excepto las 7, 10 y 13 se usa un formato estandar. Tal formato es:

i) El número de la opción se perfora como un entero en las columnas 4 y 5 (cargado a la derecha).

ii) Columnas 6-10 en blanco.

iii) Se definen seis campos como sigue:

campo	cols.
1	11-20
2	21-30
3	31-40
4	41-50
5	51-60
6	61-70

3.2. OPCION 1.

Malla (1 tarjeta)

Mediante esta opción se especifican los parámetros de la malla rectangular que entrará en el mapa.

<u>campo 1</u> : número de renglones de células de malla.

<u>campo 2</u> : número de columnas de células de malla.

<u>campos 3 y 4</u> : el tamaño de cada célula de malla en términos del número de caracteres en la dirección vertical (campo 3) y horizontal (campo 4). (Recordando siempre que un caracter mide 1/8 "verticalmente 1/10" horizontalmente.

Los números que van en los cuatro campos descritos deben llevar punto decimal.

3.3. OPCION 2.

Datos (1 tarjeta).

Esta opción controla las formas alternativas para la alimentación de los datos.

Para activar la alternativa A de datos, se debe perforar un número mayor que cero en el campo 1. Si se quiere cambiar a la alternativa B, entonces perforese un cero en el campo número 1. El campo 2 no se usa.

Si se utiliza la opción de conjunto de datos múltiple, especifique en el campo 3 el número de conjuntos de datos a ser mapeados.

<u>El estandar es un conjunto de datos y alimentación de datos con la opción B.</u>

El número especificado en el campo 1 (para la opción A de datos) es transferido a la subrutina FLEXIN como el valor de IFORM. El uso de IFORM. El uso de IFORM es discutido en la sección sobre la subrutina FLEXIN.

3.4. OPCION 3.

Número de niveles (1 tarjeta).

Para especificar el número de niveles ó intervalos de clase en los que se divide el valor del rango total (desde 2 hasta 10), perfore el número deseado de niveles mediante un número decimal en el campo 1. El estandar es 10 niveles.

3.5. OPCION 4.

Valor mínimo del rango (1 tarjeta).

Para especificar un número para usarse como el valor mínimo del valor del rango total, perforelo como un número decimal en el campo 1. El estandar es usar el valor mínimo de los datos. Para regresar al estandar perfore 1.0 en el campo 2.

3.6. OPCION 5.

Valor máximo del rango (1 tarjeta).

Para especificar un número para ser usado como el valor máximo del valor del rango total, perforelo como un número decimal en el campo 1. El estandar es usar el valor máximo de los datos. Para regresar a este estandar especifique 1.0 en el campo 2.

3.7. OPCION 6.

Valores de los rangos de los intervalos. (1 a 2 tarjetas).

Esta opción controla el rango de valores para cada nivel ó intervalo.

El valor del rango total de los datos (modificado por el mínimo y maximo de los opciones 4 y 5) será dividido entre el número de niveles especificado en la opción 3. El estandar es tener rangos iguales para cada nivel ó intervalo. (Ver ejemplo 1).

Para especificar el rango deseado para cada nivel, se usan valores proporcionales al tamaño de los rangos deseados. Estos, deben perforarse como números decimales en el campo 1 para el nivel 1, etc. (Ver ejemplo 2).

Solamente si existen más de seis niveles, continue con una segunda tarjeta, perforando el número para el séptimo nivel en el campo 1, el del octavo en el campo 2, etc. (Ver ejemplo 3).

Existe un máximo de 10 niveles para el simbolismo de la escala en gris y 20 niveles para el simbolismo de mapas puntuales.

Para regresar al estandar, especifique 0.0 en el campo 1.

3.8. OPCION 7.

Simbolismo (5 tarjetas)

Esta opción especifica el simbolismo de la escala en gris que será impreso en el mapa. Ya que no está almacenado ningún simbolismo estandar en el programa, esta opción debe ser incluida en el primer mapa de la corrida. Todas las 5 tarjetas deben ser incluidas cada vez que sea usada.

En la primera tarjeta perfore el número 7 en la columna 5 para identificar la opción.

En la segunda tarjeta perfore en las columnas dadas a continuación los caracteres básicos deseados. Puede usarse cualquier caracter de impresión.

En la tercera, cuarta y quinta tarjetas perfore en las columnas dadas a continuación los caracteres deseados para la sobreposición. Si no se desea ninguna sobreposición, estas tres tarjetas podrán dejarse en blanco.

Columnas 1-10 son usadas para especificar el simbolismo general para cada nivel (columna 1 para el simbolo del primer nivel, etc.

Columnas 11-20 son usadas para especificar el simbolismo especial para los puntos señal respectivos. (columna 11 para el simbolo de los puntos señal en el primer nivel, etc.) El punto señal es el caracter central de una célula de la malla.

Columna 21 es usada para especificar el simbolismo para un valor menor que el mínimo especificado en la opción 4.

Columna 22 es usada para especificar el simbolismo para el punto señal de un valor bajo.

Columna 23 es usada para especificar el simbolismo para un valor mayor que el máximo especificado en la opción 5.

Columna 24 es usada para especificar el simbolismo para el punto señal de un valor alto.

Columna 25 es usada para especificar el simbolismo de base; el simbolismo que aparece fuera del contorno del area de estudio.
(Ver ejemplo 4 para simbolismo de escala en gris para 10 niveles).

3.9. OPCION 8.

Punto señal (1 tarjeta).

El punto señal es el caracter central de una célula de la malla. El simbolismo especial especificado en la opción 7 es impreso en este punto señal. Para suprimir la impresión de simbolismo especial en el punto señal especifique 1.0 en el campo 1. Si se desea restablecer el punto señal en mapas subsecuentes, especifique 0.0 en el campo 1.
Cuando se hace un mapa con una malla de 1 caracter, el punto señal es suprimido automáticamente y debe ser restablecido para mapas subsecuentes.
El estandar es el simbolismo especial en el punto señal.

3.10. OPCION 9.

Histograma (1 tarjeta)

Esta opción controla la impresión al pie del mapa.

Especifique 1.0 en el campo 1 para generar un histograma al pie del mapa, el cual muestra la frecuencia de células de la malla en cada nivel.

Especifique 1.0 en el campo 2 para suprimir la información numérica que es impresa con los niveles. El estandar es no histograma ó diagrama de barras y la inclusión de información numerica. Para regresar al estandar, especifique 0.0 en el campo relevante.

3.11. OPCION 10.

Texto (3-32 tarjetas).

Si se desea información explicativa adicional a aquella contenida en el título del mapa, esta opción pueda usarse hasta 30 líneas de texto debajo del mapa.

En la primera tarjeta perfore el número de identificación de la opción 10 en las columnas 4 y 5.

En las tarjetas siguientes (no más de treinta), las cuales se insertan entre la primera y la última, perfore en las columnas 1-72 cualquier información suplementaria útil para referencias futuras.

En la última tarjeta, perfore ENDTEXT en las columnas 1-7.

El estandar es no tener texto.

3.12. OPCION 11.

Registro de datos (1 tarjeta).

Si se desea un listado de los valores dato, antes de escalarlos, perfore 1.0 en el campo 1. Si se desea un paquete de tarjetas perforadas de los valores dato, perfore 1.0 en el campo 2. Si se desea un paquete de tarjetas perforadas de los números de los niveles a los cuales han sido así asignados los datos, perfore 1.0 en el campo 3.

El estandar es no impresión ni perforación. Para regresar al estandar, especifique 0.0 en el campo relevante.

3.13. OPCION 12.

Mapa Puntual (1 tarjeta).

Como una alternativa al simbolismo normal, se puede producir un mapa puntual usando células de malla de 4 X 5 y el símbolo ø. El rango de los datos puede dividirse en 20 niveles (1[8]. si se especificó un valor máximo en la opción 3). El número de caracteres impresos en la célula es igual al número del nivel: si el valor cae en el nivel 1, solo se imprime 1 de los 20 caracteres, pero si cae en el vigésimo nivel, se imprimen todos los 20 caracteres.

Esta opción suprime las especificaciones del tamaño de las células de la malla en la opción 1 y el número de niveles en la opción 3.

Para especificar simbolismo PUNTUAL, perfore 1.0 en el campo 1. Para restablecer el simbolismo de la escala en gris (especificado en la opción 7) perfore 0.0 en el campo 1.

El simbolismo de la escala en gris es estandar.

3.14. OPCION 13.

Numeración de la malla (1 a 2 tarjetas).

Esta opción genera números de renglones y columnas sobre los cuatro lados de la malla para ayudar al usuario a localizar células individuales sobre el mapa.

La célula superior izquierda de la malla es llamada Célula de referencia de la malla (CRM), la cual provee las coordenadas apartir de las cuales son numerados todos los renglones y columnas.

Si las coordenadas de la CRM no son espacificadas, el programa supone que sean:

columna = 1

renglón = N

donde N es el número de renglones especificado en la opción 1.

Especifique 1.0 en el campo 1 para la numeración de la malla. En el campo 2 especifique el numero de columna de la CRM y en el campo 3 su número de renglón.

El estandar es no numerar la malla. Para regresar al estandar en mapas subsiguientes, especifique 0.0 en el campo 1.

Para algunos usos especializados, la malla básica puede ser subdividida en partes, tales como medios ó tercios y usarse un sistema de numeración no-continuo. El número de subdivisiones es especificado en el campo 1. (Ver ejemplo 5.)

3.15. OPCION 14.

Datos pre-escalados (1 tarjeta).

Esta opción no toma en cuenta la rutina que asigna los valores dato a los niveles. Para activar esta opción, especifique 1.0 en el campo 1. Esta opción suprimirá automáticamente la información numérica debajo del mapa.

El estandar es que el programa escale los datos.

Para restablecer el procesamiento normal en mapas subsecuentes especifique 0.0 en el campo 1; la información numerica será restablecida usando la opción 9.

3.16. OPCION 15.

Marcador de tiempo (1 tarjeta).

Esta opción origina la ejecución de tiempos para las diferentes etapas del programa que serán impresos junto con la información del paquete del mapa. Para activar esta opción especifique 1.0 en el campo 1. El estandar es no tener los tiempos impresos.

La última tarjeta del Paquete del Mapa debe tener 99999 perforado en las columnas 1-5.

4. CONTORNOS IRREGULARES.

Aunque los datos hayan sido recolectados sobre la base de una malla rectangular, el contorno del área de estudio puede no ser rectangular.

Existen en GRID dos métodos para manejar este problema:

4.1. Llenando el rectángulo.

El programa espera leer un valor dato por cada célula. Cuando existe un contorno irregular, el usuario puede completar el rectángulo con registros de datos (generalmente un registro de dato para una célula de malla por tarjeta) que indica que la célula debe ser impresa con el simbolismo de base. Al leerse los valores en la subrutina FLEXIN, la ocurrencia de células blancas ó de base debe ser inspeccionada. Cuando ocurre una célula de base se le debe asignar un valor dato de - 999999.0. Este valor activa la rutina de simbolismo de base y hace que la célula sea impresa de esta manera.

El indicador de base más fácil de codificar es un cero ó blanco a menos que el cero sea un valor válido. Por simplicidad, puede codificarse directamente el valor de - 999999.0.

4.2. Paquete de Contorno irregular.

Para simplificar el manejo de contornos irregulares se ha construido una pequeña rutina dentro del programa, de tal forma que el usuario puede especificar la forma del contorno sin tener que llenar el rectangulo con registros de datos. El contorno irregular es especificado en términos del número de células a partir de los ejes verticales de la malla—límites izquierdo y derecho— que van a ser dejados en blanco en cada renglón. El simbolismo de base será asignado automáticamente a esas células.

Esta información es dada a la computadora en un paquete separado llamado Paquete de Contorno Irregular, y se especifica como sigue:

En la primera tarjeta se perfora IRREGULAR OUTLINE en las columnas 1–17.

En la última tarjeta se perfora 99999 en las columnas 1–5

Entre la primera y la última tarjeta se perfora una serie de tarjetas con el siguiente formato:

En las columnas 1–5 el numero de renglones sucesivos para los cuales se repite el formato particular.

En las columnas 6–10 el número de células blancas al comienzo del renglón; y,

En las columnas 11–15 el número de células blancas al final del renglón.

Estos son números enteros; deben estar justificados a la derecha y no tener puntos decimales. Como el programa procesa las tarjetas en orden, la primera tarjeta se refiere al renglón (ó renglones especificados en las columnas 1–5) superior, la segunda tarjeta al segundo renglón (ó primer cambio de formato).

(Ver ejemplo 6).

Este paquete debe preceder al primer paquete de mapa al cual se refiere. Una vez que ha entrado, será usado en los mapas sucesivos hasta que sea remplazado por un nuevo paquete ó suprimido por un paquete en blanco, que contenga solamente la primera y la última tarjetas, el cual restablecerá la malla rectangular como el contorno.

Esta rutina está limitada a manejar solamente irregularidades contiguas a un eje vertical de la malla.

5. SUBRUTINA FLEXIN.

FLEXIN es una subrutina en FORTRAN IV que es usada para especificar instrucciones acerca de los valores dato a ser mapeados para cada célula de la malla. Estas instrucciones pueden especificar:

i). Si el valor a ser mapeado está localizado en una tarjeta de datos ó en un archivo de datos en cinta ó disco; ó,

ii). Los análisis estadísticos que van a realizarse sobre una variable, ó variables, para derivar el valor a ser mapeado.

Esta subrutina es llamada por el programa principal una vez para cada célula dato que vá a ser mapeada. Cada vez que es llamada, lee la tarjeta de datos ó archivo que se refiere a la célula dato.

(Ver ejemplos 7 y 8).

Estos ejemplos intentan demostrar la utilización de los argumentos de la subrutina (IFORM, T, FIRST). El usuario familiarizado con FORTRAN IV puede desarrollar análisis más sofisticados y rutinas estadísticas para aplicar a sus datos.

6. CORRIDAS EN COMPUTADORA.

Después que han sido preparados los paquetes, deben ser puestos en el orden correcto junto con las tarjetas de control necesarias.

El orden normal de los paquetes es:

Tarjetas de control

Programa en FORTRAN (incluyendo la Subrutina FLEXIN)

Mas tarjetas de control

Datos sobre los cuales vá a operar el programa.

Los datos sobre los que opera el programa consisten en:

Paquete de Contorno irregular, paquete de mapa y los datos de entrada.

Estos paquetes deben estar en el orden correcto:

i) El paquete de Contorno irregular debe preceder al paquete de Mapa al que se refiere. Una vez que un paquete de Contorno irregular ha sido especificado, será usado para todo paquete de mapa hasta que sea suprimido.

ii) Cada vez que el programa lee un paquete de Mapa intenta hacer un mapa. No existe límite para los paquetes de Mapa a utilizarse en una sola corrida.

iii) Si los Datos de entrada están en tarjetas, estos deben seguir inmediatamente al paquete de Mapa a que se refieren.

El final de los datos de entrada se señalan con una tarjeta con END perforado en las columnas 1-3 siguiendo inmediatamente al último paquete de mapa ó la última tarjeta de datos, si los datos están en tarjetas.
(Ver ejemplo 9).